Cell, Vol. 55, 577–587, November 18, 1988,

A Protein Component of Drosophila Polar Granules Is Encoded by *vasa* and Has Extensive Sequence Similarity to ATP-Dependent Helicases

Bruce Hay, Lily Yeh Jan, and Yuh Nung Jan
Howard Hughes Medical Institute
Department of Physiology
and Department of Biochemistry
University of California
San Francisco, California 94143

Summary

Determinants of pole cells, which are precursors of the germ line, are provided maternally and are localized to the posterior pole of the Drosophila egg, as are polar granules. It has been hypothesized that certain RNA molecules associated with polar granules may be necessary for pole cell determination. Using a monoclonal antibody (Mab46F11) against polar granules, we have cloned the gene for one of their components. This gene turns out to be *vasa*, which is required maternally for the formation of polar granules and germ cells. This polar granule component shows significant sequence similarity to eIF-4A, a translation initiation factor that binds to mRNA, and to other helicases.

Introduction

The precursors of the Drosophila germ line (pole cells) form at the posterior pole of the embryo because of the action of maternally supplied and posteriorly localized cytoplasmic determinants (Jazdowska-Zagrodzinska, 1966; Okada et al., 1974; Illmensee and Mahowald, 1974; Warn, 1975; Illmensee et al., 1976; Niki, 1986; Frohnhöfer et al., 1986). Several observations associate these determinants with polar granules, which are densely staining cytoplasmic organelles without limiting membranes (Mahowald, 1962, 1968). First, polar granules are localized in the posterior pole plasm (germ plasm) in late-stage oocytes and early embryos (Counce, 1963; Mahowald, 1968). Second, both polar granules and the pole cell determinants appear to be associated with maternal RNA. UV irradiation or inhibition of protein synthesis early during embryogenesis blocks pole cell formation. Moreover, poly(A)$^+$ RNA isolated from late-stage oocytes or early embryos can restore the ability of UV-irradiated embryos to form pole cells, although these pole cells do not develop far enough to produce germ cells (Okada and Kobayashi, 1987). Interestingly, polar granules appear to be associated with RNA (Mahowald, 1971) only at those stages of oogenesis and embryogenesis when the polar plasm is capable of inducing germ line development, but not at late stages, e.g., after they are segregated into pole cells. Finally, the loss of maternally supplied activity of genes in the *grandchildless-knirps* class causes the absence of polar granules as well as the inability to form pole cells (Boswell and Mahowald, 1985; Schüpbach and Wieschaus, 1986; Lehmann and Nüsslein-Volhard, 1986; Nüsslein-Volhard et al., 1987). These observations are consistent with the hypothesis that polar granules are associated with a selective group of maternal mRNA molecules that are translated in early embryos to give rise to germ line determinants (Mahowald, 1968).

To elucidate the role of polar granules in germ cell determination, it is necessary to analyze the components of polar granules biochemically. We report here the cloning of a gene coding for one such component, a 72–74 kd protein recognized by the monoclonal antibody Mab46F11 (Hay et al., 1988). The deduced amino acid sequence suggests that this protein may be capable of binding and/or unwinding RNA.

Results

Isolation of cDNA for the Polar Granule Antigen and Localization of the Transcripts

Monoclonal antibody Mab46F11 binds to polar granules at the posterior pole of late-stage oocytes and early embryos, as well as to other germ line–specific structures throughout the fly life cycle (Hay et al., 1988). To clone the gene for this antigen, we screened an ovarian cDNA (Lambda ZAP) expression library with Mab46F11. Two classes of positive cDNA clones were isolated; clones within each class hybridize to one another and are probably derived from a single gene. Because cDNAs in the first class hybridize to transcripts expressed exclusively in germ cells (see below), whereas cDNAs in the second class hybridize to RNA throughout the embryo (not shown), we have only investigated clones in the first class and have used these cDNAs as hybridization probes to isolate full-length cDNA clones from an ovarian λgt11 library kindly provided by Dr. Laura Kalfayan. We have previously found a nuclear antigen that is present in somatic cells of late-stage embryos and is stained weakly by Mab46F11 (Hay et al., 1988), but we have not determined whether this nuclear antigen is related to the cDNAs of the second class.

cDNA from the first class (hereafter referred to simply as the cDNA) hybridizes to transcripts that show essentially the same distribution, notwithstanding one exception, as the strong Mab46F11 staining in embryos, ovaries, and testes (Figures 1 and 2). In the embryo, strong Mab46F11 staining is first localized to the posterior pole and is subsequently associated with pole cells (Hay et al., 1988). As shown in Figure 1, high levels of hybridizing transcripts are found uniformly distributed throughout the preblastoderm embryo. By the time of blastoderm formation, hybridizing transcript is essentially absent. Hybridization signal does not reappear until roughly stage 12 or early stage 13, the end of germ band shortening (staging according to Campos-Ortega and Hartenstein, 1985). When it does reappear, label is found over pole cells, which are arranged in bilateral clusters in the posterior third of the embryo (Figure 1C). At this and later stages of embryogenesis, pole cells are the only cells that contain

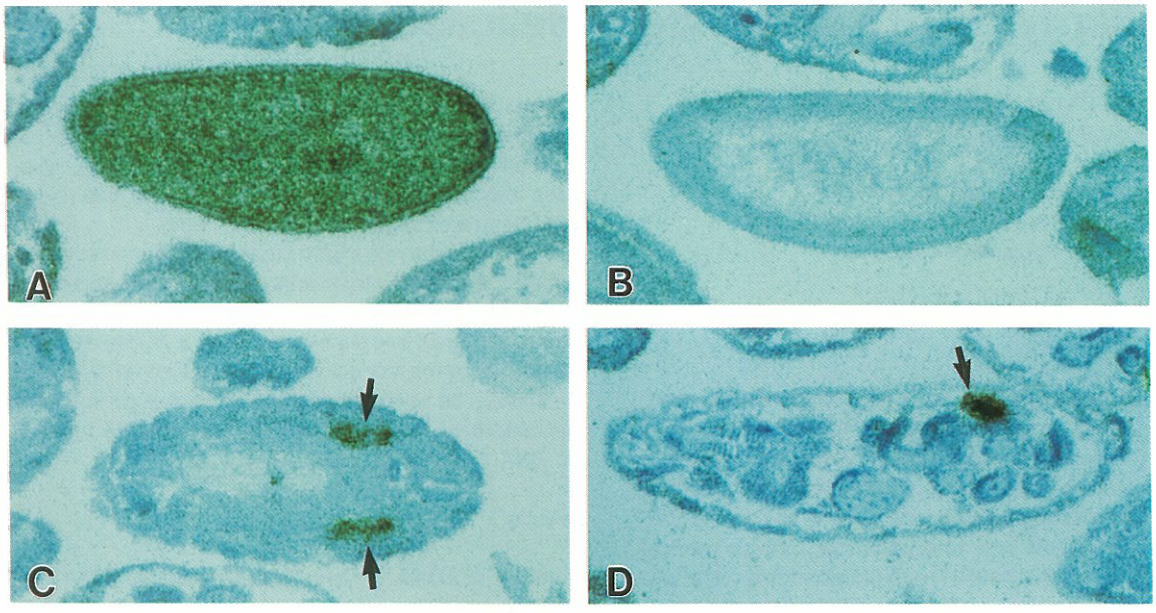


Figure 1. Tissue In Situ Hybridization of Embryos with cDNA Isolated with Monoclonal Antibody Mab46F11

In each panel, anterior is to the left. (A) Parasagittal section through a preblastoderm embryo showing high levels of hybridization uniformly distributed throughout the embryo. (B) Parasagittal section through an embryo at the beginning of blastoderm formation. Hybridization intensity has decreased uniformly throughout the embryo to background levels. (C) Late stage 12, early stage 13 embryo. The only labeled cells are the pole cells (arrows). This is about the earliest stage that we can detect the hybridization signal in pole cells during embryogenesis. (D) Semi-parasagittal section through a first-instar larva. Hybridization is restricted to the pole cells (arrow).

hybridizing transcripts and show strong staining with the Mab46F11 antibody. This temporal distribution of the transcript suggests that zygotic transcription of this gene occurs primarily, if not exclusively, in pole cells (Figure 1).

In the adult female the Mab46F11 antigen is found in germ line stem cells, in nurse cells (sister cells of the oocyte), and at the oocyte posterior pole. In the adult male the antigen is present in germ cells in the testes at early stages of spermatogenesis (Hay et al., 1988). With the exception of late-stage oocytes, which show a uniform distribution of the hybridizing transcript (as in early embryos), transcript hybridizing to the cDNA is found exclusively in cells of the germ line lineage, paralleling the localization of the Mab46F11 antigen (Figure 2).

Although the Mab46F11 antigen is found sharply localized to the posterior pole of late-stage oocytes and early embryos, the transcript that encodes this antigen appears to be uniformly distributed in the cytoplasm of these cells (Figure 1). This difference may be accounted for by the following scenario. The Mab46F11 antigen is likely to be synthesized in the nurse cells and then transported into the oocyte, where it is targeted to the oocyte posterior pole (Mahowald, 1962; Hay et al., 1988). The transcript for the Mab46F11 antigen, however, probably flows into the oocyte with the bulk of nurse cell cytoplasm during the second half of oogenesis (Mahowald and Kambysellis, 1978) and does not appear to be targeted to any specific location.

Consistent with the observed distribution of transcripts for the Mab46F11 antigen, Northern analysis shows the presence of a single major transcript of about 2.2 kb, which is abundant in ovaries (not shown), female flies, and early embryos (Figure 3). Much lower levels of expression can be detected in late embryos (Figure 3) and in larvae (not shown).

vasa Is the Gene for the Mab46F11 Antigen

Since the Mab46F11 antigen is a component of the polar granules, a prime candidate for its structural gene might be one of the *grandchildless-knirps* class of maternal-effect gap genes, because mutations in these genes cause females to give rise to embryos without polar granules or pole cells (Boswell and Mahowald, 1985; Schüpbach and Wieschaus, 1986; Lehmann and Nüsslein-Volhard, 1986; Nüsslein-Volhard et al., 1987). Of the mutations in this class that we have tested (*tudor, valois, oskar, staufen, vasa*), only the *vasa* mutation ($vasa^{PD23}$) removed the Mab46F11 antigen. Early embryos from $vasa^{PD23}$ mutant mothers lack the 72–74 kd protein (Hay et al., 1988) as well as Mab46F11 staining (Figure 4). In $vasa^{PD23}$ adult females, immunoreactivity is limited to the germarium of the ovary. Substantial levels of the antigen, however, are still present in the adult male germ line of this *vasa* allele (data not shown). Since $vasa^{PD23}$ is a hypomorphic allele (R. Lehmann, personal communication), these observations are compatible with the possibility that *vasa* codes for the Mab46F11 antigen.

The cDNAs isolated with Mab46F11 hybridize to the 35BC region on the left arm of the second chromosome (Figure 5), the same region that contains *vasa*. Further-

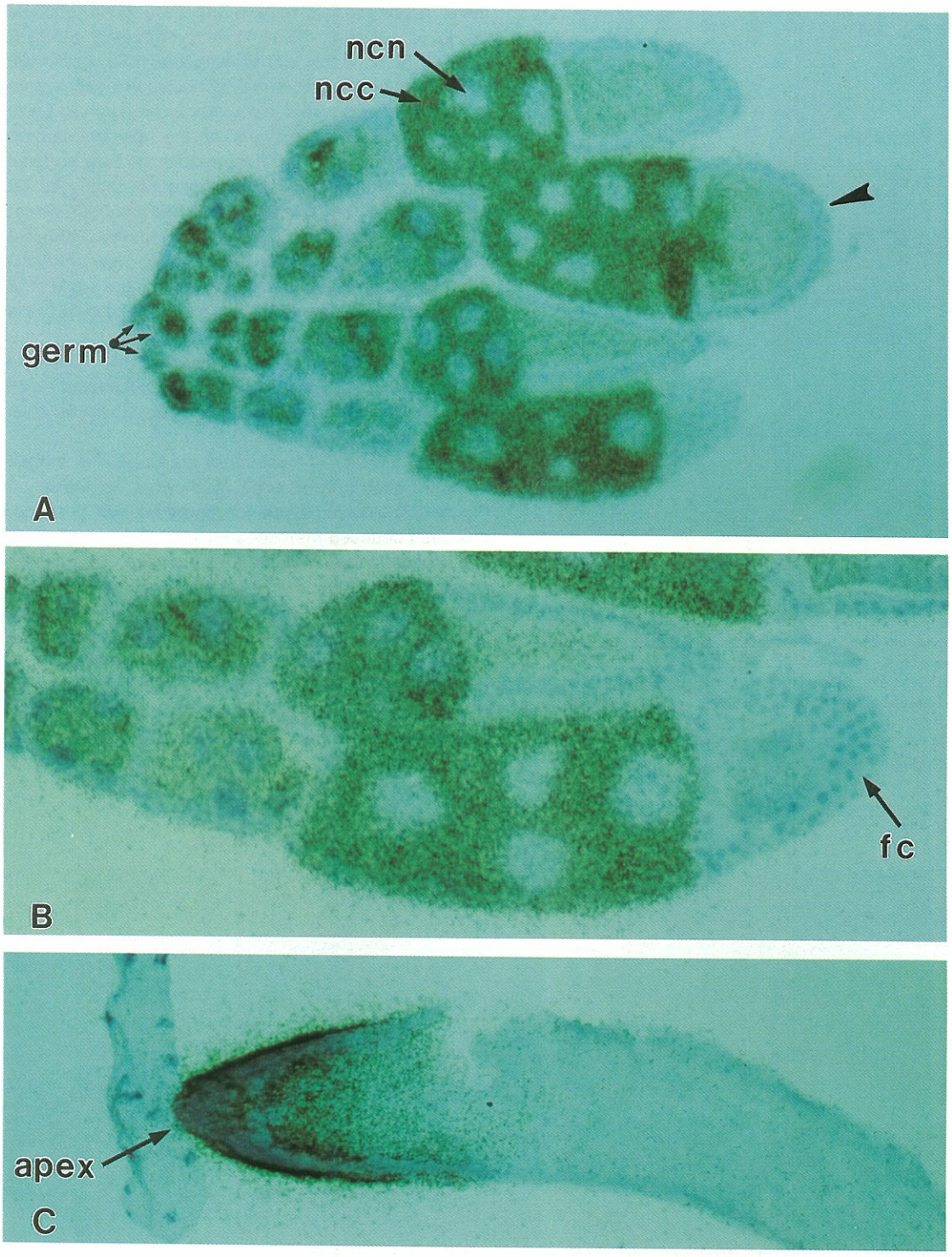


Figure 2. Tissue In Situ Hybridization of Adult Ovaries and Testis with cDNA Isolated with Monoclonal Antibody Mab46F11

(A) Low-magnification view of a section through an ovary. The germarium (germ), containing germ line stem cells and proliferating cysts (Mahowald and Kambysellis, 1978), is located to the left. Later stages of oogenesis are located to the right. High levels of hybridization are found in the germarium and the nurse cell cytoplasm (ncc). Signal is absent from the nurse cell nucleus (ncn), however. At these stages (pre-stage 9,10), the oocyte is relatively devoid of hybridization. The arrowhead points to an oocyte in which transport of transcript from nurse cells to the oocyte is beginning. Transcript accumulates throughout the oocyte beginning around stage 9 such that by stage 14, the mature oocyte, large amounts of transcript are found throughout the oocyte, similar to what is seen in the cleavage stage egg shown in Figure 1A.

(B) Higher-magnification view of a glancing section through a single egg chamber showing the lack of label associated with the somatic follicle cells (fc).

(C) Whole-mount squash of an adult testis showing hybridization limited to the apical tip (apex). This is the region in which germ line stem cells and early-stage spermatocytes reside.

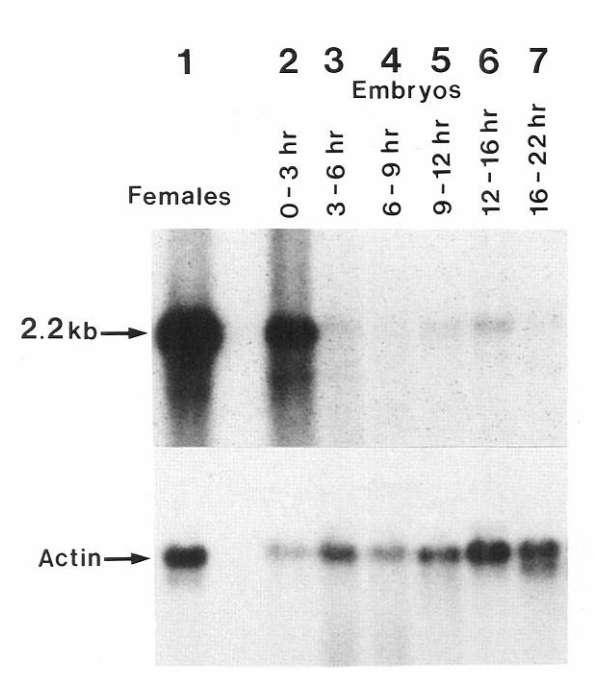


Figure 3. Northern Blot of Maternal Poly(A)$^+$ RNA from Female Flies and Poly(A)$^+$ RNA from Different Embryonic Stages at 25°C

An EcoRI fragment that contains the amino half of the protein coding sequence in the full-length cDNA was used as a hybridization probe. Lane 1 shows an abundant transcript of roughly 2.2 kb in female flies. This transcript is also abundant in early (0–3 hr) embryos (lane 2) but occurs at much lower levels in 3–6 hr and 6–9 hr embryos (lanes 3 and 4). Poly(A)$^+$ RNA from later stages (9–12 hr, lane 5; 12–16 hr, lane 6; 16–22 hr, lane 7) is overloaded (note actin controls below) to reveal the hybridization signals. Embryogenesis takes 22 hr.

more, deletions that remove the *vasa* gene, such as *Df(2L)75c* (35A1–2; 35D4–7) and *Df(2L)A72* (35B2–3; 35B7–8) (Schüpbach and Wieschaus, 1986; Lasko and Ashburner, 1988), remove the hybridization signal, whereas nearby deletions that leave *vasa* intact, such as *Df(2L)64j*, do not remove the signal on the polytene chromosome (data not shown). These results strongly suggest, but do not prove, that *vasa* is the gene coding for the Mab46F11 antigen. The *vasa* gene has just been cloned in a chromosome walk (Lasko and Ashburner, 1988). Sequence comparisons reveal that the *vasa* gene product is essentially identical and that its gene has the same organization as the gene we have isolated (see below). Therefore, *vasa*, which is required for the formation of polar granules (Schüpbach and Wieschaus, 1986; Nüsslein-Volhard et al., 1987), is most likely the gene coding for the 72–74 kd Mab46F11 antigen component of the polar granules (Hay et al., 1988).

The organization of the gene for the Mab46F11 antigen, hereafter referred to as *vasa*, is shown schematically in Figure 6. The coding sequence for the Mab46F11 antigen has been deduced from sequencing a cDNA that apparently contains the entire coding sequence, several overlapping incomplete cDNAs from a second library, and corresponding genomic DNA. The first methionine codon in the open reading frame is flanked by nucleotide sequences that agree with the consensus sequence for the ribosome initiation site (Cavener, 1987), and the stop codon is followed by a consensus polyadenylation signal and a poly(A) tail (Figure 6). Conceptual translation yields a protein of 648 amino acids (Figure 6), so that the primary

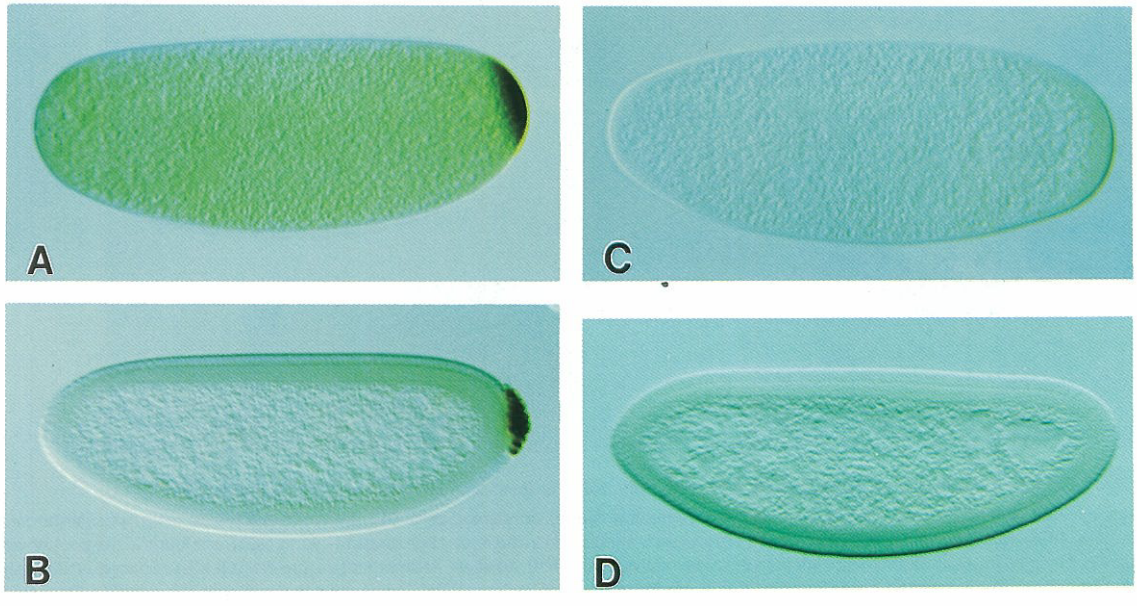


Figure 4. Monoclonal Antibody Mab46F11 Staining of Whole-Mount Wild-Type Embryos and Embryos Derived from Homozygous *vasa*^{PD23} Mothers

(A) Wild-type cleavage stage embryo showing Mab46F11 antigen localized to the polar plasm. (B) Wild-type embryo following pole cell formation. The Mab46F11 antigen has been incorporated into pole cells. (C) Cleavage stage *vasa*^{PD23} embryo showing absence of localized Mab46F11 immunoreactivity. Also note that the background level of antigen throughout the embryo is decreased relative to the wild type. (D) *vasa*^{PD23} embryo stage similar to the embryo in (B). Pole cells are absent and Mab46F11 immunoreactivity is undetectable.

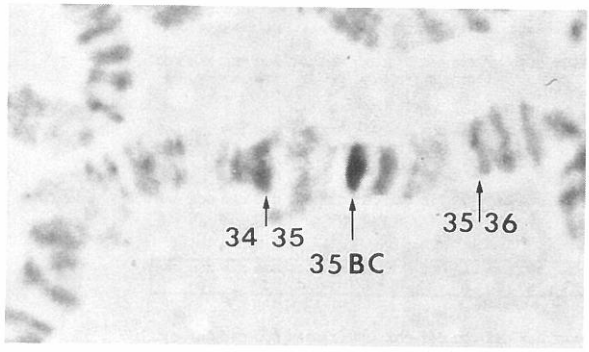


Figure 5. Chromosome In Situ Hybridization with Mab46F11-Positive cDNA

Hybridization signal localizes to a single site at 35BC on the left arm of the second chromosome.

translation product has a predicted molecular mass of 71 kd. This predicted size is very similar to the size (72–74 kd) of the Mab46F11 antigen as measured on SDS–polyacrylamide gels (Hay et al., 1988), in agreement with the assignment of the *vasa* gene product as the Mab46F11 antigen.

The Predicted *vasa* Protein Sequence

We have noticed three regions of interest in the predicted *vasa* protein. A large internal domain of *vasa* has sequence similarity with eukaryotic translation initiation factor 4A (eIF-4A; Nielsen et al., 1985; Nielsen and Trachsel, 1988) and a nuclear antigen, p68, whose appearance is associated with cell proliferation (Ford et al., 1988; Figure 7). This sequence similarity has been noticed independently by Lasko and Ashburner (1988). Second, the carboxyl terminus of *vasa* contains multiple negatively charged residues (Figure 6), as in several single-stranded nucleic acid binding proteins. Finally, near its amino terminus the *vasa* protein contains a fivefold tandem repeat of the heptad sequence F(or S)RGGE(or Q)GG (Figure 8).

Sequence Similarity with eIF-4A and Other Helicases

The *vasa* product shows 27% amino acid identity with eIF-4A and 31% identity with human nuclear antigen p68 (Figure 7). The function of the latter protein is unknown. The activity of eIF-4A, on the other hand, has been well characterized. eIF-4A is a component of a high molecular weight protein complex (called eIF-4F) involved in 5′ cap recognition and the binding of mRNA to ribosomes. It has been suggested that eIF-4A functions to bind and unwind mRNA in an ATP-dependent manner; i.e., that eIF-4A is an ATP-dependent RNA helicase (Ray et al., 1985; Lawson et al., 1986; Abramson et al., 1987; Rhoads, 1988).

Sequence motifs thought to be required for ATP binding (Walker et al., 1982) are present in eIF-4A as well as p68 (Ford et al., 1988) and *vasa* (motifs I and II in Figure 7). In addition to these two motifs, five others have been found in a large number of ATP-dependent helicases (motifs Ia, III, IV, V, and VI; Hodgman, 1988a, 1988b; Lane, 1988). Six of these seven motifs can be readily identified in *vasa* (Figure 7). The extensive sequence similarity between *vasa* and eIF-4A and the presence of these motifs strongly suggest that the *vasa* product is also an ATP-dependent nucleic acid binding protein.

The Acidic Carboxyl Terminus of *vasa*

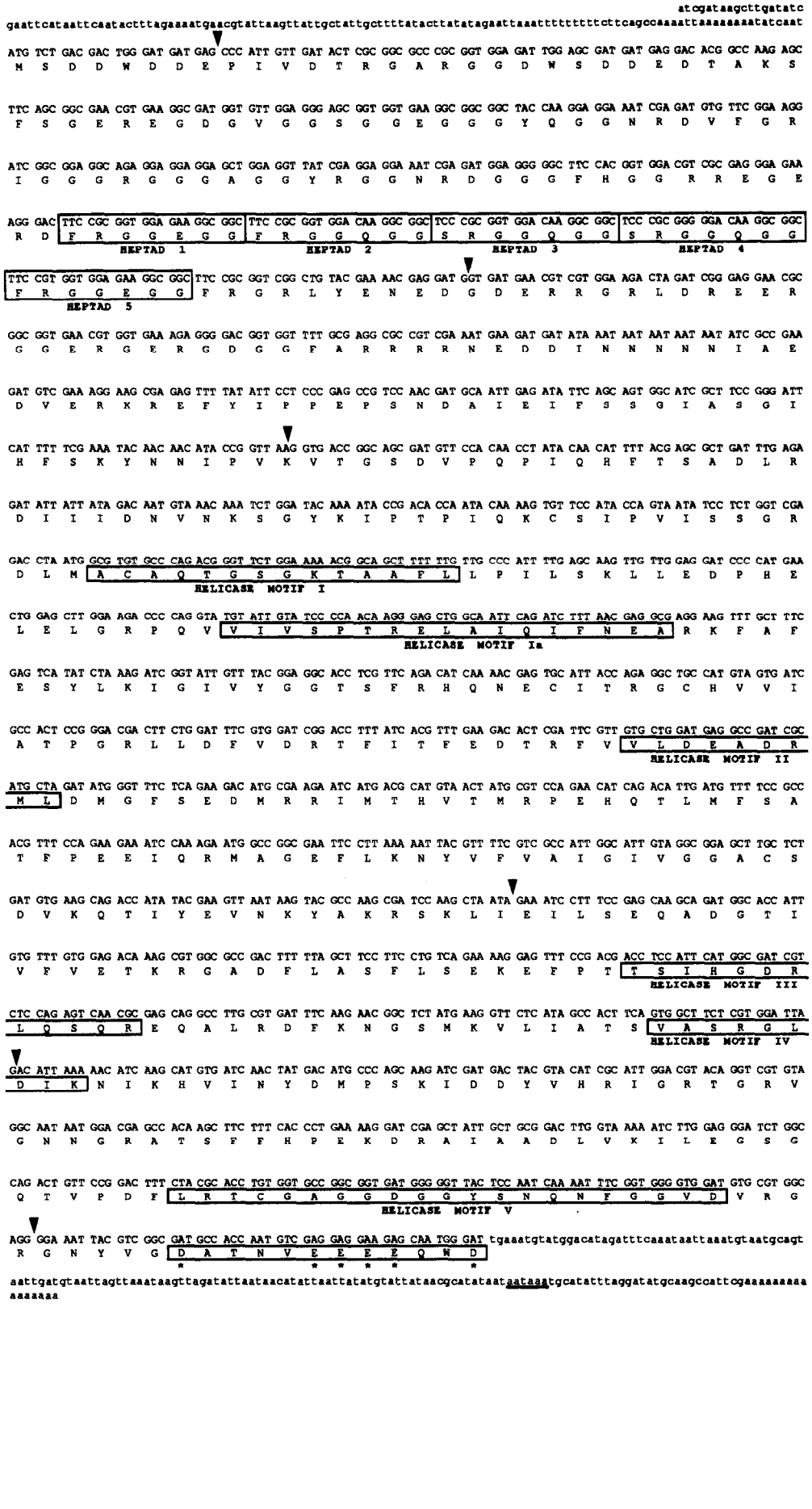
The carboxyl terminus of the *vasa* product is negatively charged; six of the last 12 amino acids are glutamate or aspartate residues (Figure 6). This feature is interesting in light of the sequence similarities between *vasa* and known helicases, since similar acidic carboxyl termini have been found in other RNA binding or single-stranded DNA binding proteins (Williams et al., 1983; Chase and Williams, 1986; Swanson et al., 1987).

The F(or S)RGGE(or Q)GG Heptad Repeats

The amino-terminal region of *vasa* is glycine-rich (61 glycines in 200 residues, or 30%) and very hydrophilic (76, or 38%, charged residues). In particular, a fivefold tandem repeat of the heptad F(or S)RGGE(or Q)GG is found (Figure 8). Glycine-rich sequences are found in a number of likely RNA binding proteins, including proteins in the nucleolus (Lischwe et al., 1985; Jong et al., 1987; Lapeyre et al., 1987) and heterogeneous nuclear ribonucleoproteins (hnRNPs; Cobianchi et al., 1986; Kumar et al., 1986; Cruz-Alvarez and Pellicer, 1987). Direct association of these proteins with RNA has been demonstrated (Thomas et al., 1981; Bugler et al., 1987; Schenkel et al., 1988) or implicated (Jong et al., 1987; Parker and Steitz, 1987). The glycine-rich regions in these proteins often contain multiple aromatic residues and arginines (Lischwe et al., 1985; Cobianchi et al., 1986; Kumar et al., 1986; Cruz-Alvarez and Pellicer, 1987; Jong et al., 1987; Lapeyre et al., 1987). The clustering of aromatic residues and basic residues in these proteins, as well as in proteins with the RNP consensus sequence (Swanson et al., 1987), has been suggested to be indicative of single-stranded nucleic acid binding domains, as demonstrated for Escherichia coli single-stranded binding proteins, bacteriophage T4 gene *32* protein, and bacteriophage fd (M13) gene *5* protein (O'Connor and Coleman, 1983; Chase and Williams, 1986).

Discussion

We have cloned the gene for a polar granule component recognized by Mab46F11 (Hay et al., 1988). This gene was subsequently identified as *vasa* because its sequence is essentially identical to that of *vasa* (Lasko and Ashburner, 1988). Since the *vasa*^{PD23} mutation removes polar granules (Schüpbach and Wieschaus, 1986) as well as the *vasa* protein in oocytes and early embryos, most likely the *vasa* protein is either a major component of polar granules or a component important for their integrity. The failure of these mutant embryos to form pole cells, then, is consistent with the hypothesis that polar granules are involved in the determination of pole cells (Mahowald, 1968).


HEPTAD 1
HEPTAD 2
HEPTAD 3
HEPTAD 4
HEPTAD 5
HELICASE MOTIF I
HELICASE MOTIF Ia
HELICASE MOTIF II
HELICASE MOTIF III
HELICASE MOTIF IV
HELICASE MOTIF V

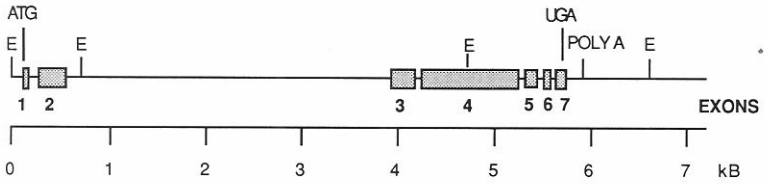


Figure 6. Genomic Organization and Coding Sequence of the *vasa* Gene

The schematic diagram above shows the genomic organization of the *vasa* gene. The coding region is divided into seven exons, separated by five small introns, of 53, 58, 57, 63, and 70 bp, and a large intron of about 3.5 kb. The positions of EcoRI sites used in subcloning genomic fragments are indicated (E), as are the initiation codon (ATG), termination codon (UGA), and the poly(A) splice site (POLY A).

The sequence of the *vasa* cDNA as well as the locations of the six introns (arrowheads) is presented at left. cDNAs were sequenced on both strands. The one cDNA isolated from the Kalfayan library contains the entire sequence shown. Several incomplete cDNAs, corresponding to the amino-terminal half of the Kalfayan clone down to the end of exon 4, were isolated from the Lambda ZAP library. The largest Kalfayan clone (2229 bp) is similar in size to the single transcript identified by Northern analysis, indicating that this clone encodes a nearly full-length or full-length cDNA. The genomic clone spans the three EcoRI fragments shown in the schematic drawing. All of these were sequenced with two exceptions. Only the ends of the 3.5 kb intron were sequenced; genomic sequence corresponding to the 30 bp 5′ of the polyadenylation site of the cDNA was not. We also note that 12 out of 55 bp from bases 2583–2638 differ in the genomic and cDNA clones. The cDNA sequence is shown here. All but the last 30 nucleotides of the cDNA sequence have been confirmed by sequencing genomic DNA (on one strand). Boxed regions refer to features discussed in the text and Figures 7 and 8. These are the five heptad repeats, the six helicase motifs (I, Ia, II, III, IV, V), and the acidic tail (with asterisks beneath charged residues). The consensus polyadenylation sequence in the 3′ end of the sequence is underlined.

We have noted a number of minor differences between our sequence and that of Lasko and Ashburner (1988), including sequences of the introns. Amino acid changes are also noted: #252, Y for F in Lasko and Ashburner (1988); #569, R for C; #581, H for D. A tandem repeat of 39 nucleotides (#539–#577) is present in Lasko and Ashburner (1988) as well as in one, but not in the other, of our cDNA clones. Because only one copy is found in our genomic DNA sequence, only one copy is shown here.

Sequence of the *vasa* Protein Suggests That It May Be an RNA Binding Helicase

The *vasa* sequence contains three features that are interesting in light of the association between polar granules and RNA (Mahowald, 1971) and in light of the implication of maternal RNA in pole cell determination (Okada and Kobayashi, 1987).

First, *vasa* has extensive sequence similarity with the translation initiation factor eIF-4A and contains six of the seven motifs identified in a number of helicases. This suggests that the *vasa* protein may be able to associate with, and to unwind, RNA. Second, the highly acidic carboxyl terminus is reminiscent of the carboxyl termini of several single-stranded nucleic acid binding proteins. Some of these carboxyl termini appear to be involved in the interaction between the nucleic acid binding protein and other proteins (e.g., between the T4 gene *32* product and other proteins involved in DNA replication, or between the high-mobility group nonhistone proteins and histones); such interactions may alter the nucleic acid binding affinity (Burke et al., 1980; Reeck et al., 1982).

Finally, the amino-terminal region of *vasa* contains a fivefold tandem repeat of F(or S)RGGE(or Q)GG. It has been pointed out that from purely physical considerations, the extent of a heptad repeat in any non-α-helical structure is unlikely to exceed two or three heptads (Cohen and Parry, 1986). The presence of five tandem heptads in *vasa* would therefore suggest an α-helical structure, although the presence of multiple glycine residues may weaken this prediction. These heptads and the flanking sequences are arranged in such a way that in an α-helix of 12 turns there would be six arginines along one face next to four phenylalanines and two serines (Figure 8B). Bearing in mind the involvement of basic residues and aromatic residues in the interaction with single-stranded nucleic acids (O'Connor and Coleman, 1983; Chase and Williams, 1986), we are tempted to consider the possibility that a single-stranded nucleic acid (e.g., RNA) interacts with the structure of heptad repeats so that the phosphate backbone has electrostatic interactions with arginines and the nucleotide bases interact with the phenylalanines by stacking, or with the serines by forming hydrogen bonds.

This model further raises the intriguing question of whether the postulated association with RNA has any specificity for primary sequences or secondary structures. Sequence-specific binding domains have been found at the amino-terminal region, separate from the putative helicase domain, of the SV40 large tumor antigen and of the transcription termination factor rho (Brennan et al., 1987; Gish and Botchan, 1987; Dombroski and Platt, 1988; Goetz et al., 1988). Such a model for the heptad repeats should be regarded with great caution, however, since no examples of a single α-helical structure interacting with nucleic acids have been shown. Besides, other possible functions for the heptad repeats exist, such as formation of a coiled-coil structure, although the heptad sequences in *vasa* do not agree very well with the preferred arrangement of heptads found in coiled coils (Cohen and Parry, 1986).

Possible Relation between Polar Granules and Pole Cell Determinants

Maternal RNA has been implicated as being important in germ cell determination because it can partially rescue the ability of UV-irradiated embryos to form pole cells (Okada and Kobayashi, 1987). Since only the posterior cytoplasm (the pole plasm) from late-stage oocytes or early embryos is capable of inducing the formation of pole cells (Illmensee and Mahowald, 1974; Illmensee et al.,

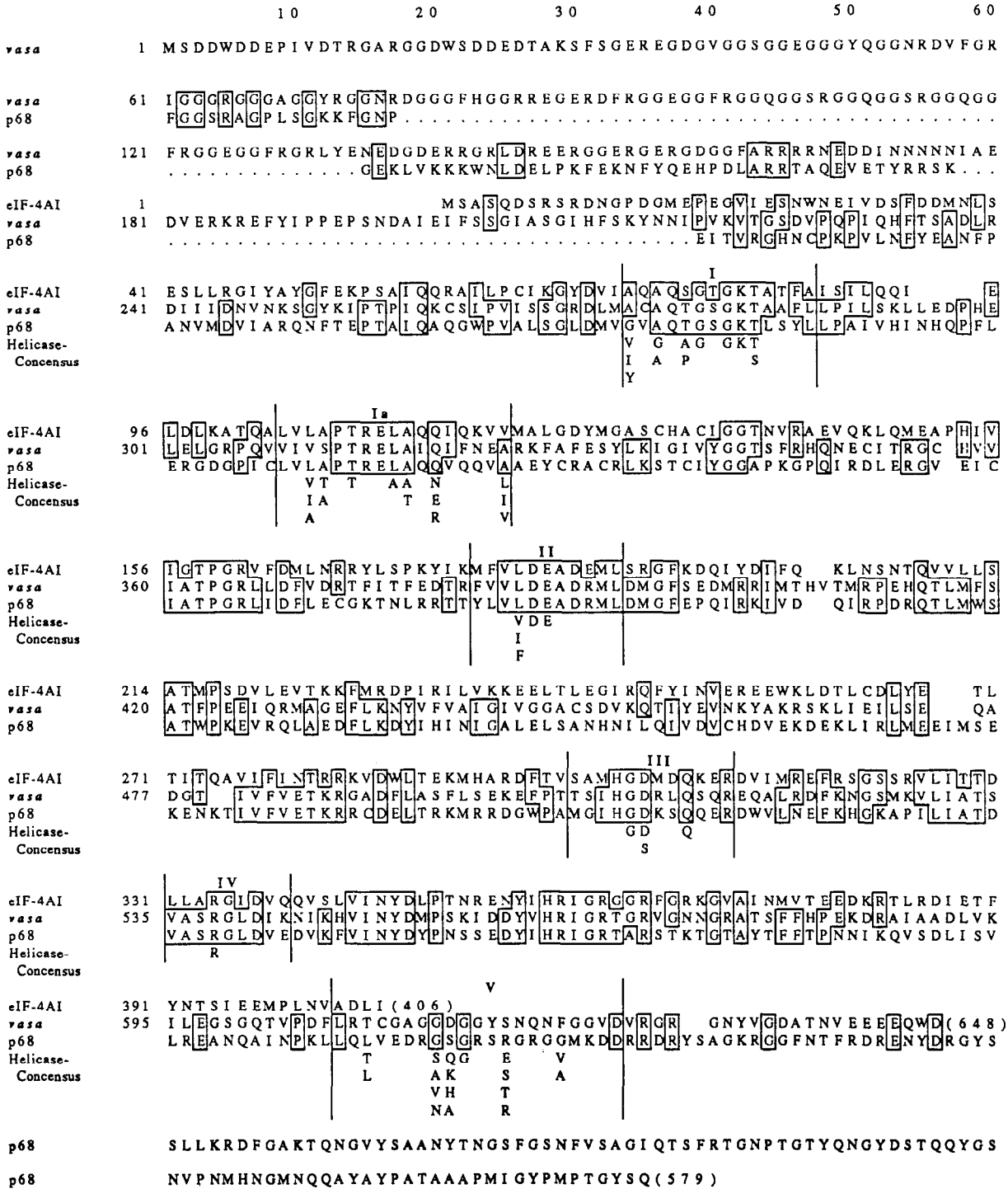


Figure 7. Homology of *vasa* Coding Sequence with Murine eIF-4AI and Human p68, and Identification of Domains Shared by a Large Number of Helicases

Alignments have been performed visually between eIF-4AI, p68, and *vasa*. Boxed regions indicate identity with the *vasa* sequence. Vertical lines and Roman numerals show our alignment of six domains within the *vasa* sequence identified as common to a large number of helicases (Hodgman, 1988a, 1988b; Lane, 1988). Letters below the protein sequence alignments indicate the Hodgman (1988a, 1988b) consensus sequences (in single-letter codes for amino acids) for these domains of helicases.

1976), either the relevant maternal RNA is restricted to the posterior pole or it has wider distribution but is only functional in the posterior pole. In the former case, the posterior localization of these maternal RNA species may be attributed to polar granules, e.g., binding of specific mRNA by the *vasa* protein. (A precedent has been found in Xenopus oocytes, where a protein binds to specific mRNA [Crawford and Richter, 1987].) Translation of such

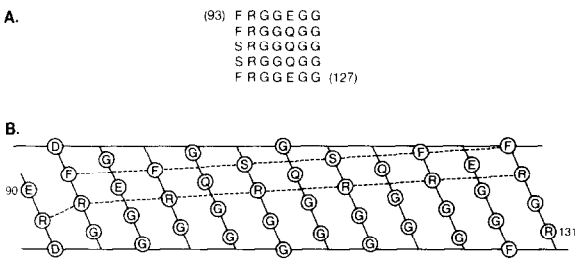


Figure 8. Fivefold Tandem Heptad Repeats in the *vasa* Sequence

(A) The fivefold heptad repeat F(or S)RGGE(or Q)GG in *vasa*. Numbers in parentheses indicate the positions of the first and the last amino acids of the repeats. (B) Helical-net analysis of residues 90 through 131, including the heptad repeats. Dotted lines mark the row of arginines and the adjacent row of phenylalanines and serines.

mRNA in the embryo may follow either its release from polar granules or interactions between polar granule components and other cellular constituents. If the relevant maternal RNA is not restricted to the posterior pole, one could imagine that *vasa*, or other polar granule components, has to bind to these species of maternal RNA to allow for their translation (Kozak, 1983, 1988). It remains to be determined whether one or a combination of mechanisms such as those described above, or others, are important in pole cell determination.

Posterior Localization of the *vasa* Protein

Although the *vasa* protein is sharply localized to the posterior pole as soon as it becomes detectable in the oocyte (Hay et al., 1988), the *vasa* transcript is uniformly distributed in late-stage oocytes and early embryos (Figure 1). Therefore, the targeting machinery must operate at the level of the protein but not of the *vasa* mRNA. The *vasa* protein may be either directly interacting with the targeting machinery or associated with other proteins that are targeted to the posterior pole—possibly other components of polar granules. Indeed, our preliminary studies of maternal-effect mutations of the *grandchildless-knirps* class (*tudor, valois, oskar, staufen*; unpublished results) indicate that these genes may be involved in the establishment of posterior localization or its maintenance.

Experimental Procedures

Stocks

vasa^{PD23} and deficiency stock *Df(2L)75c* were provided by Drs. Trudi Schüpbach and Eric Wieschaus, Princeton University. Deficiency stocks *Df(2L)A72* and *Df(2L)64j* were obtained from Dr. Michael Ashburner and the Bowling Green stock center.

Immunocytochemistry

Mab46F11 is a monoclonal antibody that recognizes polar granules and other germ line–specific structures. Immunocytochemical techniques were as described in Hay et al. (1988).

cDNA Isolation and Sequencing

RNA was isolated from hand-dissected ovaries using the hot-phenol method (Schwarz et al., 1988). Poly(A)$^+$ RNA was isolated by two cycles of adsorption to and elution from oligo(dT)–cellulose (Aviv and Leder, 1972). This RNA was used to construct a Lambda ZAP expression library (Stratagene). cDNAs for the Mab46F11 antigen were isolated from the expression library using standard techniques for plating, isopropyl thiogalactoside induction, and transfer to nitrocellulose filters (Huynh et al., 1985; Stratagene protocol). Monoclonal antibody Mab46F11 was diluted 1:10 in 10 mM Tris–HCl (pH 8.0), 150 mM NaCl, 0.05% Tween 20 (TBST), and incubated for 1 hr with the filters. Three 15 min washes at room temperature in TBST were followed by a 1 hr incubation with rabbit anti-mouse IgM (Zymed) diluted 1:1000 with TBST. Washes as above were followed by incubation with goat anti-rabbit antibody conjugated to alkaline phosphatase (Promega Biotec) diluted 1:7500 in TBST. Development of the alkaline phosphatase reaction product followed the Promega Biotec protocol.

cDNAs were subcloned from Lambda ZAP using a plasmid rescue procedure (Bluescript manual, Stratagene). cDNAs from the Lambda ZAP ovarian library were used as hybridization probes to screen a λgtII library made from ovarian poly(A)$^+$-enriched RNA (Dr. Laura Kalfayan, University of North Carolina, Chapel Hill) and a genomic Lambda Dash library (Stratagene) using standard techniques (Maniatis et al., 1982). cDNA and genomic DNA were subcloned in Bluescript (Stratagene).

Sequencing of single-stranded DNA was carried out using the Sequenase kit (United States Biochemical) and oligonucleotides corresponding to previously determined sequences. Sequences were determined for both strands of cDNAs and one strand of the genomic DNA. *vasa* sequence data have been submitted to the GenBank/EMBL data library (accession no. J03147).

In Situ Hybridization to Chromosomes

In situ hybridizations to polytene chromosomes were performed as described in Papazian et al. (1987).

Northern Blot Analysis

RNA extraction was either by the hot-phenol method (Schwarz et al., 1988) or by the guanidinium method (Chomczynski and Sacchi, 1987). RNA was electrophoresed on formaldehyde–agarose gels, blotted onto Nytran membranes (Schleicher & Schuell), and hybridized as described in Vässin et al. (1987).

In Situ Hybridization to Tissue

Ovaries and embryos were fixed, embedded in paraffin, and sectioned according to Vässin et al. (1987). Testes were dissected and processed through the methanol fixation step described in Hay et al. (1988). They were then postfixed in 4% paraformaldehyde in PBS (130 mM NaCl, 10 mM sodium phosphate [pH 6.8]) for 15 min and processed for hybridization as with the other tissues. Probes for hybridization were ^{35}S-labeled antisense RNAs derived from cDNA sequences using either the T3 or T7 promoter of Bluescript. Probes were reduced to a mean length of 100–200 bp using base hydrolysis (Cox et al., 1984). Hybridizations, washes, and signal detection were as described in Vässin et al. (1987).

Acknowledgments

We thank Drs. P. Lasko and M. Ashburner for communicating their studies of *vasa* prior to publication, Dr. L. Kalfayan for sending us the ovarian cDNA library, Drs. T. Schüpbach, E. Wieschaus, and M. Ashburner for mutant stocks, and Drs. B. Alberts, E. Giniger, C. Guthrie, R. Lehmann, and P. Walter for helpful comments on the manuscript. We are most grateful for the excellent technical help of R. Tuma and L. Ackerman, for the help of Drs. H. Vässin and M. Caudy with Northern analysis, and for the unfailing assistance in manuscript preparation by L. Schulte and M. Los Huertos. Y. N. J. and L. Y. J. are Howard Hughes Medical Institute investigators.

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

Received September 19, 1988.

References

Abramson, R. D., Dever, T. E., Lawson, T. G., Ray, B. K., Thach, R. E., and Merrick, W. C. (1987). The ATP-dependent interaction of eukaryotic initiation factors with mRNA. J. Biol. Chem. *262*, 3826–3832.

Aviv, H., and Leder, P. (1972). Purification of biologically active globin

messenger RNA by chromatography on oligothymidylic acid–cellulose. Proc. Natl. Acad. Sci. USA *69*, 1408–1412.

Boswell, R. E., and Mahowald, A. P. (1985). *tudor*, a gene required for assembly of the germ plasm in Drosophila melanogaster. Cell *43*, 97–104.

Brennan, C. A., Dombroski, A. J., and Platt, T. (1987). Transcription termination factor rho is an RNA–DNA helicase. Cell *48*, 945–952.

Bugler, B., Bourbon, H., Lapeyre, B., Wallace, M. O., Chang, J. H., Amalric, F., and Olson, M. O. J. (1987). RNA binding fragments from nucleolin contain the ribonucleoprotein consensus sequence. J. Biol. Chem. *262*, 10922–10925.

Burke, R., Alberts, B., and Hosoda, J. (1980). Proteolytic removal of the COOH-terminus of the T4 gene *32* helix-destabilizing protein alters the T4 *in vitro* replication complex. J. Biol. Chem. *255*, 11484–11493.

Campos-Ortega, J. A., and Hartenstein, V.C. (1985). The Embryonic Development of *Drosophila melanogaster* (New York: Springer-Verlag).

Cavener, D. R. (1987). Comparison of the consensus sequence flanking translation start sites in *Drosophila* and vertebrates. Nucl. Acids Res. *15*, 1353–1361.

Chase, J. W., and Williams, K. R. (1986). Single-stranded DNA binding proteins required for DNA replication. Annu. Rev. Biochem. *55*, 103–136.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol chloroform extraction. Anal. Biochem. *162*, 156–159.

Cobianchi, F., Sengupta, D. N., Amudzka, B. Z., and Wilson, S. H. (1986). Structure of rodent helix-destabilizing protein revealed by cDNA cloning. J. Biol. Chem. *261*, 3536–3543.

Cohen, C., and Parry, D. A. (1986). Alpha-helical coiled coils—a widespread motif in proteins. Trends Biochem. Sci. *11*, 245–248.

Counce, S. J. (1963). Developmental morphology of polar granules in *Drosophila*, including observations on pole cell behavior and distribution during embryogenesis. J. Morphol. *112*, 129–145.

Cox, K. H., DeLeon, D. V., Angerer, L. M., and Angerer, R. C. (1984). Detection of mRNAs in sea urchin embryos by *in situ* hybridization using asymmetric RNA probes. Dev. Biol. *101*, 485–502.

Crawford, D. R., and Richter, J. D. (1987). An RNA-binding protein from *Xenopus* oocytes is associated with specific message sequences. Development *101*, 741–749.

Cruz-Alvarez, M., and Pellicer, A. (1987). Cloning of a full-length complementary DNA for an *Artemia salina* glycine-rich protein. J. Biol. Chem. *262*, 13377–13380.

Dombroski, A. J., and Platt, T. (1988). Structure of rho factor: an RNA-binding domain and a separate region with strong similarity to proven ATP-binding domains. Proc. Natl. Acad. Sci. USA *85*, 2538–2542.

Ford, M. J., Anton, I. A., and Lane, D. P. (1988). Nuclear protein with sequence homology to translation initiation factor eIF-4A. Nature *332*, 736–738.

Frohnhöfer, H. G., Lehmann, R., and Nüsslein-Volhard, C. (1986). Manipulating the anterior-posterior pattern of the *Drosophila* embryo. J. Embryol. Exp. Morphol. *97* (Suppl.), 169–179.

Gish, W. R., and Botchan, M. R. (1987). Simian virus 40-transformed human cells that express large T antigens defective for viral DNA replication. J. Virol. *61*, 2864–2876.

Goetz, G. S., Dean, F. B., Hurwitz, J., and Matson, S. W. (1988). The unwinding of duplex regions DNA by the simian virus 40 large tumor antigen-associated DNA helicase activity. J. Biol. Chem. *263*, 383–392.

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

Hodgman, T. C. (1988a). A new superfamily of replicative proteins. Nature *333*, 22–23.

Hodgman, T. C. (1988b). A new superfamily of replicative proteins. Erratum. Nature *333*, 578.

Huynh, T. V., Young, R. A., and Avis, R. W. (1985). Constructing and screening cDNA libraries in λgt10 and λgt11. In DNA Cloning: A Practical Approach, Vol. 1, D. M. Glover, ed. (Oxford: IRL Press), pp. 49–78.

Illmensee, K., and Mahowald, A. P. (1974). Transplantation of posterior pole plasm in *Drosophila*. Induction of germ cells at the anterior pole of the egg. Proc. Natl. Acad. Sci. USA *72*, 1016–1020.

Illmensee, K., Mahowald, A. P., and Loomis, M. R. (1976). The ontogeny of germ plasm during oogenesis in *Drosophila*. Dev. Biol. *49*, 40–65.

Jazdowska-Zagrodzinska, B. (1966). Expermimental studies on the role of polar granules in the segregation of pole cells in *Drosophila*. J. Embryol. Exp. Morphol. *16*, 391–399.

Jong, A. Y. S., Clark, M. W., Gilbert, M., Oehm, E., and Campbell, J. L. (1987). *Saccharomyces cerevisiae* SSB1 protein and its relationship to nucleolar RNA-binding proteins. Mol. Cell. Biol. *7*, 2947–2955.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbiol. Rev. *47*, 1–45.

Kozak, M. (1988). A profusion of controls. J. Cell Biol. *107*, 127.

Kumar, A., William, K. R., and Szer, W. (1986). Purification and domain structure of core hnRNP proteins A1 and A2 and their relationship to single-stranded DNA-binding proteins. J. Biol. Chem. *261*, 11266–11273.

Lane, D. (1988). Enlarged family of putative helicases. Nature *334*, 478.

Lapeyre, B., Bourbon, H., and Amalric, F. (1987). Nucleolin, the major nucleolar protein of growing eukaryotic cells: an unusual protein structure revealed by the nucleotide sequence. Proc. Natl. Acad. Sci. USA *84*, 1472–1476.

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

Lawson, T. G., Ray, B. K., Dodds, J. A., Grifo, J. A., Abramson, R. D., Merrick, W. C., Betsch, D. F., Weith, H. L., and Thach, R. E. (1986). Influence of 5′ proximal secondary structure on the translational efficiency of eukaryotic mRNAs and on their interaction with initiation factors. J. Biol. Chem. *261*, 13979–13989.

Lehmann, R., and Nüsslein-Volhard, C. (1986). Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in Drosophila. Cell *47*, 141–152.

Lischwe, M. A., Ochs, R. L., Reddy, R., Cook, R. G., Yeoman, L. C., Tan, E. M., Reichlin, M., and Busch, H. (1985). Purification and partial characterization of a nucleolar scleroderma antigen (M_r = 34,000, pI, 8.5) rich in N^G,N^G-dimethylarginine. J. Biol. Chem. *260*, 14304–14310.

Mahowald, A. P. (1962). Fine structure of pole cells and polar granules in *Drosophila melanogaster*. J. Exp. Zool. *151*, 201–216.

Mahowald, A. P. (1968). Polar granules of *Drosophila*. II. Ultrastructural changes during early embryogenesis. J. Exp. Zool. *167*, 237–262.

Mahowald, A. P. (1971). Polar granules of *Drosophila*. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis. J. Exp. Zool. *176*, 345–352.

Mahowald, A. P., and Kambysellis, M. P. (1978). Oogenesis. In Genetics and Biology of *Drosophila*, Vol. 2, M. Ashburner and T. R. F. Wright, eds. (New York: Academic Press), pp. 141–224.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Nielsen, P. J., and Trachsel, H. (1988). The mouse protein synthesis initiation factor 4A gene family includes two related functional genes which are differentially expressed. EMBO J. *7*, 2097–2105.

Nielsen, P. J., McMaster, G. K., and Trachsel, H. (1985). Cloning of eukaryotic protein synthesis initiation factor genes: isolation and characterization of cDNA clones encoding factor eIF-4A. Nucl. Acids Res. *13*, 6867–6880.

Niki, Y. (1986). Germline autonomous sterility of the P-M dysgenesis hybrids and their application to germline transfers in *Drosophila melanogaster*. Dev. Biol. *113*, 255–258.

Nüsslein-Volhard, C., Frohnhöfer, H. G., and Lehmann, R. (1987). Determination of anterior-posterior polarity in *Drosophila*. Science *238*, 1675–1681.

O'Connor, T. P., and Coleman, J. E. (1983). Proton nuclear magnetic resonance (500 MHz) of mono-, di-, tri-, and tetradeoxynucleotide complexes of gene *5* protein. Biochemistry *22*, 3375–3381.

Okada, M., and Kobayashi, S. (1987). Maternal messenger RNA as a determinant of pole cell formation in *Drosophila* embryos. Dev. Growth Differentiation *29*, 185–192.

Okada, M., Kleinman, I. S., and Schneiderman, H. A. (1974). Restoration of fertility in sterilized *Drosophila* eggs by transplantation of polar cytoplasm. Dev. Biol. *37*, 43–54.

Papazian, D. M., Schwarz, T. L., Tempel, B. L., Jan, Y. N., and Jan, L. Y. (1987). Cloning of genomic and complementary DNA from *Shaker*, a putative potassium channel gene from *Drosophila*. Science *237*, 749–753.

Parker, K. A., and Steitz, J. A. (1987). Structural analyses of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. Mol. Cell. Biol. *7*, 2899–2913.

Ray, B. K., Lawson, T. G., Kramer, J. C., Cladaras, M. H., Grifo, J. A., Abramson, R. D., Merrick, W. C., and Thach, R. E. (1985). ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. J. Biol. Chem. *260*, 7651–7658.

Reeck, G. R., Isackson, P. J., and Teller, D. C. (1982). Domain structure in high molecular weight high mobility group nonhistone chromatin proteins. Nature *300*, 76–78.

Rhoads, R. E. (1988). Cap recognition and the entry of mRNA into the protein synthesis initiation cycle. Trends Biochem. Sci. *13*, 52–56.

Schenkel, J., Sekeris, C. E., Alonso, A., and Bantz, E. K. F. (1988). RNA-binding properties of hnRNP proteins. Eur. J. Biochem. *171*, 565–569.

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

Schwarz, T. L., Tempel, B. L., Papazian, D. M., Jan, Y. N., and Jan, L. Y. (1988). Multiple potassium-channel components are produced by alternative splicing at the *Shaker* locus in *Drosophila*. Nature *331*, 137–142.

Swanson, M. S., Nakagawa, T. Y., LeVan, K., and Dreyfuss, G. (1987). Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. Mol. Cell. Biol. *7*, 1731–1739.

Thomas, J. E., Raziuddin, Sobota, A., Boublik, M., and Szer, W. (1981). A RNA helix-destabilizing protein is a major component of *Artemia salina* nuclear ribonucleoproteins. Proc. Natl. Acad. Sci. USA *78*, 2888–2892.

Vässin, H., Bremer, K. A., Knust, E., and Campos-Ortega, J. A. (1987). The neurogenic gene *Delta* of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats. EMBO J. *6*, 3431–3440.

Walker, J. E., Saraste, M., Runswick, J. J., and Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes. EMBO J. *1*, 945–951.

Warn, R. (1975). Restoration of the capacity to form pole cells in UV irradiated *Drosophila* embryos. J. Embryol. Exp. Morphol. *33*, 1003–1011.

Williams, K. R., Spicer, E. K., LoPresti, M. B., Guggenheimer, R. A., and Chase, J. W. (1983). Limited proteolysis studies on the *Escherichia coli* single-stranded DNA binding protein. J. Biol. Chem. *258*, 3346–3355.