

A small predicted stem-loop structure mediates oocyte localization of *Drosophila K10* mRNA

Thomas L. Serano^{1,2,†} and Robert S. Cohen^{1,*}

¹Department of Biochemistry, University of Kansas, Lawrence, KS 66045, USA

²Department of Biochemistry and Molecular Biophysics, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032 USA

*Author for correspondence

[†]Present address: Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720, USA

SUMMARY

The establishment of dorsoventral polarity in the *Drosophila* oocyte and future embryo is dependent on the efficient transport of *K10* mRNA from nurse cells into the oocyte. To investigate the *cis*-requirements of *K10* mRNA transport, we used a transgenic fly assay to analyze the expression patterns of a series of *K10* deletion variants. Such studies identify a 44 nucleotide sequence within the *K10* 3' untranslated region that is required and sufficient for *K10* mRNA transport and subsequent localization to the oocyte's anterior cortex. An inspection of the 44 nucleotide transport/localization sequence (TLS) reveals a strong potential for the formation of a stem-loop secondary structure. Nucleotide substitutions that interfere with the predicted base-pairing of the TLS block mRNA transport and anterior localization. Conversely, mutations that alter the base composition of the TLS while maintaining predicted base-pairing do not block mRNA transport or anterior localization. We conclude that *K10* mRNA transport and anterior localization is mediated by a 44

nucleotide stem-loop structure. A similar putative stem-loop structure is found in the 3' untranslated region of the *Drosophila orb* mRNA, suggesting that the same factors mediate the transport and anterior localization of both *K10* and *orb* mRNAs. Apart from *orb*, the *K10* TLS is not found in any other localized mRNA, raising the possibility that the transport and localization of other mRNAs, e.g., *bicoid*, *oskar* and *gurken*, are mediated by novel sets of *cis*- and *trans*-acting factors. Moreover, we find that the *K10* TLS overrides the activity of *oskar cis*-regulatory elements that mediate the late stage movement of the mRNA to the posterior pole. We propose the existence of a family of *cis*-regulatory elements that mediate mRNA transport into the oocyte, only some of which are compatible with the elements that mediate late stage movements.

Key words: *Drosophila K10* gene, mRNA localization, mRNA transport, *Drosophila* oogenesis, RNA secondary structure

INTRODUCTION

Examples of mRNA localization as a means for protein targeting have increased greatly in recent years (reviewed in Macdonald, 1992; Ding and Lipshitz, 1993; St Johnston, 1995). It is a particularly common mechanism in developmental systems, where it is often desirable to store positional information in an inactive form. It is also widely used in large cells, e.g., eggs and neurons, possibly because it is more energy efficient to localize a relatively small number of mRNAs and translate them many times than it is to independently localize a large number of protein molecules. Despite its widespread occurrence and biological importance, the mechanism of mRNA localization is poorly understood compared to our understanding of other protein targeting mechanisms (e.g., see Pelham and Munro, 1993).

The richest identified source of localized mRNA is the *Drosophila* oocyte, in which over a dozen localized mRNAs have been found (reviewed in Macdonald, 1992; Ding and

Lipshitz, 1993). The localization of these mRNAs is a dynamic process that begins with mRNA synthesis in nurse cells and transport to the oocyte's posterior pole. mRNA synthesis and transport are tightly coupled events inasmuch as mRNAs do not accumulate in nurse cells to detectable amounts. In most cases, mRNA synthesis and transport extend from just prior to egg chamber formation though ~stage 7 of oogenesis (see King, 1970, for a complete description of egg chamber formation and the ensuing 14 stages of egg chamber maturation). At stage 8, transported mRNAs are moved from the oocyte's posterior pole to the oocyte's anterior cortex, where mRNA-specific sorting events occur. For example, *oskar* (*osk*) and *gurken* (*grk*) mRNAs are moved to the oocyte's posterior pole and anterodorsal corner, respectively, during late stage 8/early stage 9 (Kim-Ha et al., 1991; Ephrussi et al., 1991; Neuman-Silberberg and Schüpbach, 1993). Other mRNAs, e.g., *fs(1)K10* (*K10*), *bicoid* (*bcd*) and *orb*, remain at the oocyte's anterior cortex until stage 10B or later (Cheung et al., 1992; Berleth et al., 1988; Lantz et al., 1992).

Pharmacological studies indicate that mRNA transport and subsequent localization events within the oocyte are mediated by microtubules (Pokrywka and Stephenson, 1991, 1995; Theurkauf et al., 1993). Each nurse cell-oocyte complex is connected by a single microtubule array (Theurkauf et al., 1992). Through stage 6, the array's organizing center (minus-end) is positioned at the oocyte's posterior pole, and its plus-ends extend into nurse cells. During stages 6-8, this array is dismantled and a new one, nucleated along the oocyte's anterior cortex and terminated at the oocyte's posterior pole, is formed (Theurkauf et al., 1992). The dynamics and polarity of microtubules suggest that mRNA transport and anterior localization are both powered by a minus end-directed motor, e.g., a member of the dynein superfamily, while late stage (i.e., stage 9) movement of certain mRNAs to the oocyte's posterior pole is powered by a plus-end, kinesin-like, motor. Consistent with this idea, kinesin's motor subunit, when fused to β -galactosidase (β -gal), targets β -gal enzyme activity to the posterior pole of stage 9 and older oocytes (Clark et al., 1994).

The observed overlap between the localization patterns of different *Drosophila* mRNAs is suggestive of a modular design of *cis*-regulatory elements. Specifically, one might expect all localized mRNAs to contain an element that mediates association with a minus-end motor. Some mRNAs, like *osk*, that move to the oocyte's posterior pole at stage 9, might be expected to also contain an element that mediates association with a plus-end motor. Additional elements that mediate the association of mRNAs to non-microtubule components of the cytoskeleton might be required to maintain localization of some mRNAs, particularly those that remain localized during cytoplasmic streaming stages (i.e., stages 11-12), when microtubules are reorganized yet again (Theurkauf et al., 1992). In direct support of a modular design of mRNA localization elements, Kim-Ha et al. (1993) have identified distinct elements within the *osk* 3' untranslated region (3'UTR) for mRNA transport and for movement to the oocyte's posterior pole. Distinct regulatory elements have also been identified in the *bcd* 3'UTR; the 53 nucleotide BLE1 mediates *bcd* mRNA transport and localization to the oocyte's anterior cortex, while three non-contiguous sequences mediate late-stage *staufer*-dependent localization to the oocyte's anterior cytoplasm (Macdonald et al., 1993; Ferrandon et al., 1994).

In the most streamlined version of a modular regulatory logic for mRNA localization, all similar localization steps would be mediated by the same set of *cis*- and *trans*-regulatory factors. However, this is apparently not the case. For example, *bcd* mRNA transport, but not the transport of other examined mRNAs, is mediated by *exuperantia* gene activity (Berleth et al., 1988; St Johnston et al., 1989). Furthermore, there are no obvious sequence homologies between the localization control regions of similarly localized mRNAs. The lack of such sequence homology could, however, simply reflect the physical nature of mRNA localization elements. For example, they may specify RNA secondary structures, which could be difficult to detect by conventional searching algorithms – both because many different sequence combinations can specify a particular structure and because a single structure could be composed of widely separated sequences. Thus, the degree to which the similar movements of different mRNAs are controlled by the same elements remains unknown. In cases where different mRNAs do use different elements for similar

movements, it is of interest to know if the different elements have evolved for trivial or functionally significant reasons.

In this paper, we identify and characterize the *cis*-regulatory sequences of *K10* mRNA localization. *K10* mRNA is synthesized in nurse cells, transported to the oocyte's posterior pole and moved to the oocyte's anterior cortex at stage 8, where it persists through stage 10B (Cheung et al., 1992). Consistent with *K10*'s relatively simple mRNA localization pattern, we find that a single 44 nucleotide sequence located within the *K10* 3'UTR is necessary and sufficient for all *K10* mRNA localization steps. The activity of this sequence, called the transport/localization sequence (TLS), is apparently dependent on its folding into a stem-loop secondary structure. All mutations that disrupt TLS's predicted base-pairing pattern also disrupt its mRNA transport and anterior localization activity. Conversely, all mutations that alter TLS's primary sequence, while maintaining its predicted base-pairing pattern, have little or no adverse affect on mRNA transport and anterior localization. The small size of the TLS, coupled with the observation that all mutations that decrease the TLS's transport activity similarly decrease its anterior localization efficiency, suggests that the TLS is a single regulatory element. That is, the TLS appears to interact with a single regulatory protein or protein complex, presumably one that facilitates association of the mRNA with a minus end-directed microtubule-based motor protein (see above). A sequence structurally similar to TLS is also found in *orb* mRNA, which exhibits an mRNA distribution pattern similar to that of *K10* (Lantz et al., 1992), suggesting that TLS is a conserved element mediating mRNA association with a minus-end motor. Curiously, the TLS apparently overrides the posterior localization control sequences of *osk* mRNA: the substitution of the TLS for *osk*'s putative transport control sequences renders *osk*'s posterior localization control sequences inactive. We suggest that mRNAs that have different final destinations within the oocyte require functionally different transport control elements.

MATERIALS AND METHODS

Fly stocks

The wild-type stock is Oregon R. The *K10* allele is *K10^{LM00}* which was provided by T. Schüpbach and is described in Cheung et al. (1992). All markers and balancers are described in Lindsley and Zimm (1992).

Transformation constructs

All of the constructs described in the text are derived from the same ~4.4 kilobase *K10* minigene fragment. The fragment starts at nucleotide (nt) 1 and extends to nt 5363, except the intron has been omitted. Nucleotide positions are from Prost et al. (1988), with sequence corrections described in Cohen and Serano (1995). For further reference, the *K10* 3'UTR begins at nt 3048 and extends to ~nt 4499, where the poly(A)-addition recognition sequence (AATAAA) starts at nt 4474.

Internal deletions of the 3'UTR

The following constructs contain deletions (deleted segment indicated by numbers) of the 3'UTR: *KASSt*, 3112-3661; *KASH*, 3112-3970; *KASD*, 3112-4366; *KΔHD*, 3970-4366; *KAStH*, 3661-3970; *KΔA*, 3661-3749; *KΔB*, 3750-3824; *KΔC*, 3825-3895; *KΔD*, 3896-3970; *KΔAB*, 3661-3824; *KΔCD*, 3825-3970; *KΔBCD*, 3750-3970. The construction of *KASSt*, *KASH*, *KASD*, *KΔHD* and *KAStH* was facilitated

by the presence of naturally occurring *SalI* (S), *StuI* (St), *HpaI* (H) and *DraI* (D) restriction sites at nts., 3112, 3661, 3970 and 4366, respectively. The *KΔA*, *KΔB*, *KΔC*, *KΔD*, *KΔAB*, *KΔCD* and *KΔBCD* constructs were made by joining appropriate PCR-amplified fragments (details available upon request).

3' truncations

The *K26*, *KSSr26*, *KSH26*, *KStH26* and *KSD26* constructs were made by truncating the *K10* minigene fragment (described above) at *SalI*, *StuI*, *HpaI*, *HpaI* and *DraI*, respectively (coordinates listed above), and thus, lack *K10* poly(A)-addition and transcription termination control sequences. Poly(A)-addition and transcription termination control sequences from the *hsp26* gene were then spliced onto the truncated minigene fragments. The *hsp26* sequences are defined by a 623 nt *SacI*-*Clal* fragment and extend from 317 nt upstream of the *hsp26* translational stop codon to 193 nt downstream of the *hsp26* poly(A)-addition recognition sequence (Southgate et al., 1983). Endogenous *hsp26* mRNA is produced in nurse cells, but not transported or localized (Zimmerman et al., 1983). In addition to the 3' truncation, *KStH26* is internally deleted for the *K10* 3'UTR *SalI*-*StuI* region (nts 3112-3661).

TLS substitution and deletion mutants

Kstem5', *Kstem3'*, *Kstem5'3'*, *Kloop* and *Khub* differ from the 4.4 kb *K10* minigene fragment (described above) only in the TLS, as shown in Fig. 7. These mutations were introduced by PCR-mediated, site-directed mutagenesis (details available upon request) and verified by dideoxy-sequencing.

K10-oskar fusion constructs

The *K-osk*, *KoΔ* and *KoΔ+T* constructs were made by truncating the *K10* minigene fragment at the *SalI* site, nt 3112 (see above). Following truncation, *osk* 3'UTR sequences, with or without the *K10* TLS, were added to the 3' end of the minigene. For reference, the *osk* 3'UTR corresponds to the region 2551 to ~3593 [nt numbering from Kim-Ha et al., (1991)]. *K-osk* contains the *osk* *DraI*-*XbaI* fragment (*osk* nts. 2796-3658). The *osk* 3'UTR sequences of *KoΔ* and *KoΔ+T* were generated by deleting the *SphI*-*SacII* region (nts 3083-3231) from the ~4 kb *StyI*-*AspI* fragment, which extends from nucleotide 2471 to ~3 kb downstream of the *osk* poly(A)-addition site. In *KoΔ+T*, the *osk* *SphI*-*SacII* fragment is replaced by the *K10* TLS (sequence shown in shaded region of Fig. 5A).

Other constructs

KΔSD+A and *KΔSD+T* were made by substituting a single PCR-amplified copy of subregion A and a synthetic TLS (sequences given in Fig. 5A), respectively, for the *SalI*-*DraI* region (nt 3112-4366) of the *K10* minigene. *KΔStH+T* was constructed by substituting a single synthetic copy of the TLS for the *StuI*-*HpaI* region (nt 3661-3970) of the *K10* minigene. *KA26* is identical to *K26* (described above), except that the former contains a PCR-amplified copy of subregion A (described above) immediately upstream of the *hsp26* sequence.

70zStH26 consists, from 5' to 3', of a *hsp26*-*sgs3* nurse cell-specific promoter fragment (Frank et al., 1992), all but the first six amino acids of the protein coding region of the *E. coli* *lacZ* gene, the *K10* *StuI*-*HpaI* fragment (nucleotides 3661-3970) and the *hsp26* *SacI*-*Clal* fragment (described above).

Transformations

All constructs described above were cloned into the pCaSpeR4 transformation vector (Pirrota, 1988), with the exception of *70zStH26*, which was cloned into the Germ70 transformation vector (Serano et al., 1994). Constructs were introduced into *w¹¹¹⁸* flies by P element-mediated transformation (Rubin and Spradling, 1982; Spradling and Rubin, 1982). Transposase activity was provided by the p13pwc helper plasmid (Cohen and Meselson, 1985). Several independently transformed lines were generated and analyzed for each construct,

with the exception of *KoΔ*, in which only one line was analyzed. While there were line to line variations in the amount of mRNA produced by a given construct, no differences were seen in the mRNA distribution patterns. The transgenes were introduced into a *K10^{LM00}* mutant background by standard genetic crosses.

In situ hybridization

In situ hybridization to whole-mount ovaries was carried out according to Tautz and Pfeifle (1989) with modifications described in Cheung et al. (1992). Digoxigenin-labeled DNA probes were produced by random-priming according to Feinberg and Vogelstein (1983). The *K10* probe corresponds to nucleotides 757-1763 (Prost et al., 1988). The *lacZ* probe was derived from sequences in the pMC1871 plasmid (Pharmacia; see also Serano et al., 1994). In situ hybridization to mRNA produced by *K10* transgenes was carried out in a *K10^{LM00}* background, which produces no detectable endogenous *K10* mRNA (Cheung et al., 1992). Photography was carried out as previously described (Serano et al., 1995).

RESULTS

In previous studies, we mapped *K10* mRNA transport/localization control sequences to the ~1400 nucleotide (nt) 3'UTR of the mRNA (Cheung et al., 1992). To delimit further the *K10* mRNA transport/localization control sequences, we constructed a series of *K10* deletion mutants lacking various portions of the *K10* 3'UTR. Unless otherwise noted, each deletion was made within the context of the same 4.4 kilobase *K10* minigene that we and others have previously shown possesses wild-type *K10* gene activity in transgenic flies (Haenlin et al., 1987; Serano and Cohen, 1995). The transgenes were introduced into a *K10^{LM00}* mutant background which produces no endogenous *K10* mRNA (Cheung et al., 1992), and tested for their ability to produce transported/localized mRNA by in situ hybridization.

The *K10* poly(A)-addition and transcriptional termination control regions are not required for *K10* mRNA transport/localization

We first tested whether the *K10* poly(A)-addition and putative transcription termination control regions are required for mRNA transport/localization. We made a *K10* construct, called *KSD26*, in which the 3' end of the *K10* gene, starting at a site 108 nt upstream of the poly(A)-addition control site (AATAAA), was replaced with the poly(A)-addition and transcription termination control sequences from the 3' end of the *hsp26* gene. As seen in Fig. 2B, the *KSD26* mRNA distribution pattern is indistinguishable from that of wild-type *K10* mRNA. The mRNA is transported to the posterior pole of the oocyte during stages 1-7 and subsequently localized to the oocyte's anterior cortex during stages 8-10B. In contrast, a control transgene, called *K26*, which contains the same *hsp26* 3' end fragment, but lacks all but the first 65 nt of the *K10* 3'UTR (Fig. 1), produces mRNA that is retained in nurse cells until nurse cell regression, i.e., until the indiscriminate transfer of nurse cell cytoplasm to the oocyte at stage 11 (Fig. 2C). We conclude that the *K10* poly(A)-addition site and putative transcription termination control region are not required for *K10* mRNA transport/localization.

While *K10* transgenes containing *hsp26* poly(A)-addition and putative transcriptional termination control sequences

produce efficiently transported and anteriorly localized mRNA, they exhibit reduced gene activity compared to their wild-type counterparts; *K10* mutants carrying one copy of the *KSD26* transgene produce eggs that hatch with a frequency of ~20-80%, while *K10* mutants that carry one copy of a *K10* transgene containing an intact 3' end typically produce eggs that hatch at a frequency of $\geq 95\%$ (data not shown). Based on immunocytochemical analyses (data not shown), we attribute the reduced activity of the *KSD26* transgene to inefficient

translation of its mRNA. Importantly, the distribution pattern of protein produced from *KSD26* mRNA is indistinguishable from that of endogenous *K10* mRNA (data not shown).

The 309 nt *StuI-HpaI* region is necessary and sufficient for *K10* mRNA transport/localization

We next analyzed a series of transgenes lacking different internal portions of the *K10* 3'UTR. As summarized in Fig. 1, these studies identify a 309 nt *StuI-HpaI* fragment that is both necessary and sufficient for *K10* mRNA transport/localization. This is best illustrated by two constructs: *K Δ StH* and *KStH26*. *K Δ StH* specifically lacks the *StuI-HpaI* region and produces transcripts that are retained in nurse cells until nurse cell regression (Fig. 2D). In contrast, *KStH26*, which contains the

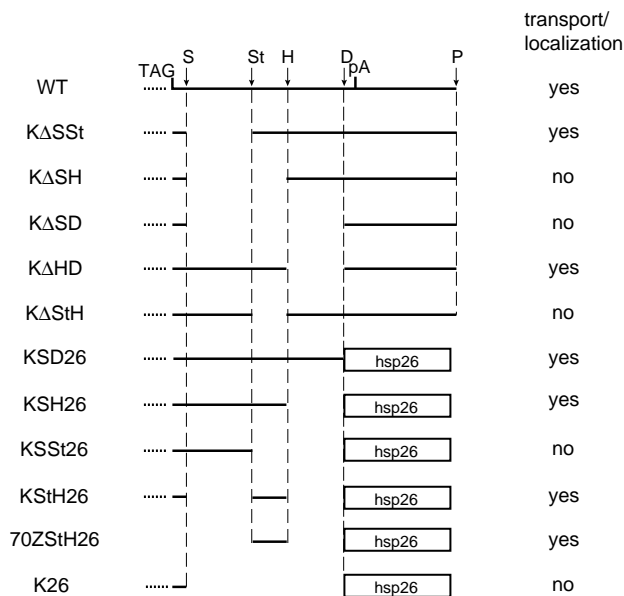


Fig. 1. Low-resolution mapping of *K10* mRNA transport/localization control sequences. All deletions (with the exception of 70zStH26) were created in the context of a *K10* minigene fragment possessing wild-type *K10* gene activity (see Materials and Methods). The constructs differ only in the region shown. The black lines represent *K10*-derived sequences, where deletions are represented by gaps in the line. The approximate positions of the *K10* translation stop codon (TAG) and poly(A)-addition site (pA) are indicated. The bottom six constructs contain *hsp26*-derived poly(A)-addition and putative transcription termination control sequences (labeled rectangle). The ability of the transgenes to produce transported/localized mRNA, as determined by in situ hybridization, is summarized to the right of the constructs (see also Fig. 2). The transcriptional control and protein coding regions of the 70zStH26 construct consists of the *hsp26* nurse cell enhancer with an adjoining *Sgs3* promoter followed by the *E. coli lacZ* protein coding region. S, *Sall*; St, *StuI*; H, *HpaI*; D, *DraI*; P, *PstI*; WT, wild type.

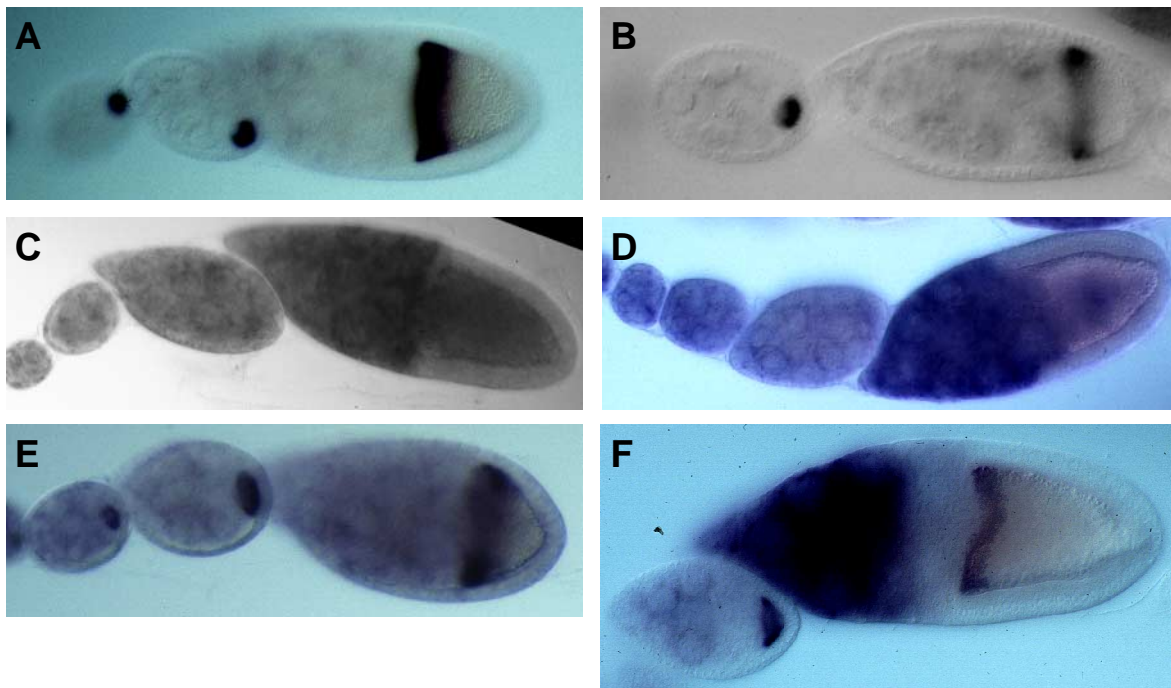


Fig. 2. Whole-mount in situ hybridization of *K10*-deletion and *K10-hsp26* fusion genes. Egg chambers are arranged from left to right in order of increasing stage of development following whole-mount in situ hybridization with a digoxigenin-labeled *K10* (A-E) or *lacZ* (F) probe. Egg chambers carry the following transgenes: (A) none (wild-type control); (B) *KSD26*; (C) *K26*; (D) *K Δ StH*; (E) *KStH26*; (F) *70zStH26*. Transgenes listed in B-E are in a *K10^{LM00}* background, which does not produce any detectable endogenous *K10* mRNA (Cheung et al., 1992). In this and all other figures, anterior is to the left.

StuI-HpaI region but lacks all other portions of the 3'UTR (except the first 65 nt), produces transcripts that are transported and anteriorly localized within the oocyte (Fig. 2E). We conclude that most, if not all, of *K10*'s transport/localization control sequences are contained within the 309 nt *StuI-HpaI* region.

To test whether the *StuI-HpaI* region is sufficient to direct the transport/localization of a heterologous mRNA, we constructed *70zStH26*. The only *K10* sequence contained in this *lacZ*-tagged transgene is the 309 nt *StuI-HpaI* region (Fig. 1). As seen in Fig. 2F, *70zStH26* mRNA is efficiently transported and localized. We conclude that the 309 nt *StuI-HpaI* region is sufficient to direct the transport and anterior localization of a heterologous mRNA.

A 92 nt subregion of *StuI-HpaI* is necessary and sufficient for *K10* mRNA transport/localization

To map more precisely *K10* mRNA transport/localization control sequences, we conceptually subdivided the 309 nt *StuI-HpaI* region into four non-overlapping subregions, denoted A-D. We deleted each subregion individually, or in combination, within the context of an otherwise intact *K10* 3'UTR. As summarized in Fig. 3A, our analysis of these constructs indicates that *K10* mRNA transport/localization control sequences are

fully contained within subregion A. When we delete subregion A (see e.g., *KΔA*), only non-transported/non-localized mRNA is produced (Fig. 4A). Conversely, *K10* transgenes lacking subregions B, C and/or D (see e.g., *KΔBCD*) produce efficiently transported/localized mRNA (Fig. 4B). These results show that subregion A is required for *K10* mRNA transport/localization and that subregions B-D are dispensable.

To test whether subregion A is sufficient for mRNA transport/localization, we constructed *KΔSD+A* and *KA26* (Fig. 3B). The 3'UTRs of these transgenes contain the first 65 nt of the *K10* 3'UTR followed by a single copy of subregion A. The two constructs differ in that the poly(A)-addition site and putative transcription termination control regions of *KΔSD+A* and *KA26* are derived from the *K10* and *hsp26* genes, respectively. Constructs identical to *KΔSD+A* and *KA26*, but lacking subregion A, produce transcripts that are not transported or localized (see *KΔSD* and *K26* in Fig. 1). As seen in Fig. 4C (and data not shown), *KΔSD+A* and *KA26* produce mRNA that is transported and anteriorly localized. We conclude that most, if not all, of *K10*'s mRNA transport/localization control sequences are contained within subregion A. We cannot, however, rule out the possibility that there are redundant copies of a second essential sequence, one or more in the region extending from the first nt of *K10* mRNA to 65 nt downstream of TAG, and one or more within subregions B-D.

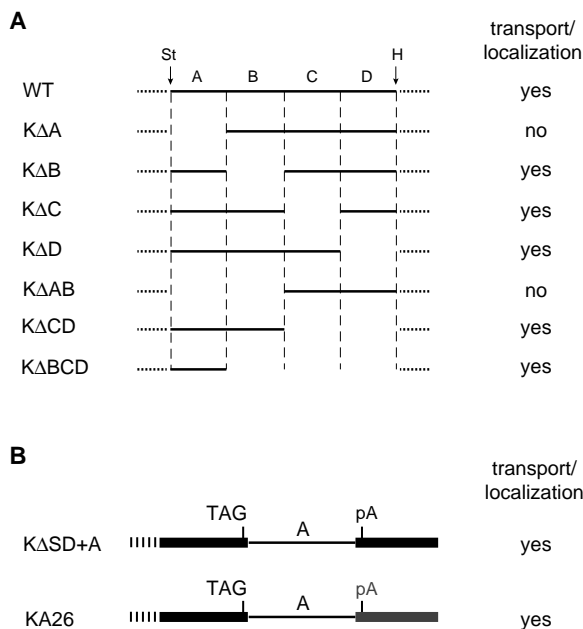


Fig. 3. High-resolution mapping of *K10* mRNA transport/localization control sequences. All deletion constructs were created in the context of the same *K10* minigene described in Fig. 1 and differ only in the region shown. (A) The black lines represent *K10*-derived sequences, where deletions are represented by gaps in the line. The 309 nt *StuI-HpaI* region (see Fig. 1) was subdivided into four non-overlapping subregions, named A, B, C and D, which were deleted individually or in combination as shown. (B) Subregion A (thin line) was inserted between the *K10* transcriptional control/protein coding region (thick black line), and poly(A)-addition site-containing region from the 3' end of the *K10* (thick black line) or *hsp26* (thick gray line) genes. The ability of the transgenes to produce transported/localized mRNA, as determined by in situ hybridization, is summarized to the right (see also Fig. 4). Symbols are as in Fig. 1.

The identification of the 44 nt *K10* mRNA transport and localization sequence (TLS)

Upon visual inspection of the 92 nt subregion A, we identified a 44 nt sequence that has a potential to form a stem-loop

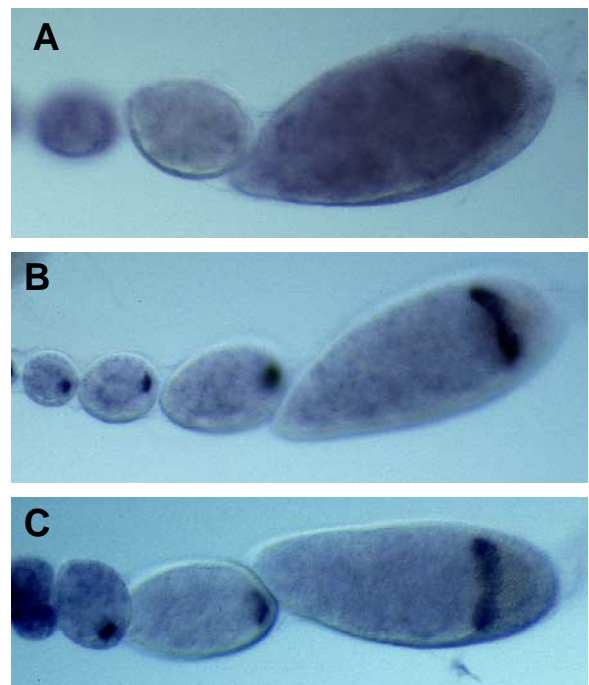


Fig. 4. Whole-mount in situ hybridization of *K10*-deletion transgenes. Egg chambers are arranged from left to right in order of increasing stage of development following whole-mount in situ hybridization with a digoxigenin-labeled *K10* probe. Egg chambers carry the following transgenes in a *K10^{LM00}* background: (A) *KΔA*; (B) *KΔBCD*; (C) *KΔSD+A*.

secondary structure (Fig. 5A,B). We decided to test whether this sequence, called the *K10* transport and localization sequence (TLS), is sufficient to direct mRNA transport and localization. We constructed two transgenes, *K Δ StH+T* and *K Δ SD+T*, in which the TLS replaces the 309 nt *StuI-HpaI* region and the ~1250 nt *Sall-DraI* region of the *K10* 3'UTR, respectively (Fig. 5C). As seen in Fig. 5C (and data not shown), both *K Δ StH+T* and *K Δ SD+T* produce transported/localized mRNA. In addition, *K10* mutants carrying one copy of either *K Δ StH+T* or *K Δ SD+T* produce eggs that typically hatch at a frequency of 80-95%. These results, together with those of the experiments described below, indicate that the 44 nt TLS contains most, if not all, of the mRNA transport/localization control information present in the entire *K10* mRNA.

The TLS is sufficient for the transport and anterior localization of recombinant *osk* transcripts, but overrides *osk* posterior localization sequences

We next wanted to see if the TLS could substitute for the transport/anterior localization control region of a mRNA whose final destination within the oocyte is different than *K10*'s. We chose *osk*, since its mRNA transport and localization control sequences had been previously defined. Specifically, Kim-Ha et al. (1993) identified a ~150 nt *SphI-SacII* region that mediates *osk* mRNA transport and a separate region that mediates late stage posterior localization. We first replaced *K10*'s 3'UTR with *osk* 3'UTR sequences to determine whether *osk* transport and localization control sequences retain activity when fused to *K10* protein coding sequences and produced under the transcriptional control of the *K10* promoter/enhancer. This control construct, called *K-osk*, produces mRNA that, like wild-type *osk*, is transported to the oocyte's posterior pole during stages 1-7, transiently localizes to the oocyte's anterior cortex during early stage 8, and moves back to the oocyte's posterior pole during late stage 8/early stage 9, where it persists through the remainder of oogenesis (Fig. 6A,B and data not shown; Kim-Ha et al., 1991; Ephrussi et al., 1991).

We then constructed two *K10-oskar* hybrid transgenes, called *Ko Δ* and *Ko Δ +T*, that lack *osk* mRNA transport sequences, but retain *osk* posterior localization sequences (Fig. 6A). In *Ko Δ +T*, the deleted *osk* transport control sequences are replaced with the *K10* TLS, while in *Ko Δ* there is no such replacement. As seen in Fig. 6B, *Ko Δ* mRNA is distributed in a manner similar to that reported for other *osk* mRNAs that lack *osk* transport control sequence, but retain posterior localization control sequences (see *olc15* in

Kim-Ha et al., 1993). *Ko Δ* transcripts are detected in nurse cells through early stage 8. Beginning at the end of stage 8, and extending through the end of stage 10B, low amounts of *Ko Δ* transcripts are also seen at the oocyte's posterior pole. Presumably, low amounts of *Ko Δ* transcripts diffuse from nurse cells into the oocyte during early stages of oogenesis and become detectable in the oocyte after stage 8 as they concentrate at the posterior pole via the intact posterior localization control sequences. In contrast to *Ko Δ* mRNA, *Ko Δ +T* mRNA is efficiently transported into the oocyte during early stages of oogenesis (Fig. 6B). However, while *K-osk* and wild-type *osk* mRNAs accumulate only transiently at the oocyte's anterior cortex during stage 8, *Ko Δ +T* mRNA remains there through at least stage 10B (Fig. 6B; detection becomes problematic after stage 10B due to the secretion of the egg chorion). We conclude that the *K10* TLS can direct the transport of a heterologous mRNA. In addition, the *K10* TLS apparently overrides the activity of *osk* mRNA posterior localization control sequences.

The stem-loop secondary structure of the *K10* TLS is apparently essential for TLS function

As mentioned above, the *K10* TLS was originally identified by its striking potential to form a stem-loop secondary structure. To determine whether TLS activity is dependent on such a

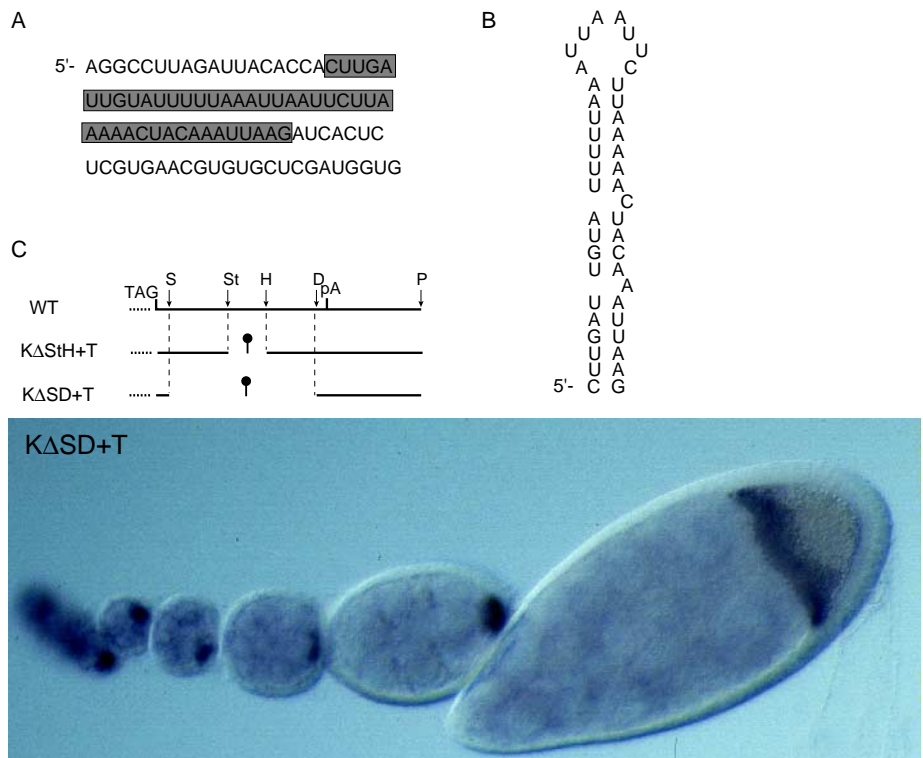


Fig. 5. The sequence, putative secondary-structure and transport/localization activity of the *K10* TLS. (A) The sequence of subregion A corresponds to nucleotides 3658-3749 of Prost et al. (1988) with corrections described in Cohen and Serano (1995). The shaded region corresponds to the 44 nt TLS. (B) The predicted stem-loop secondary structure of the TLS. (C) Schematic diagram and representative mRNA distribution pattern of constructs containing the TLS. The TLS (represented by the small stem-loop) was added to *K Δ StH* and *K Δ SD* (see Fig. 1) to give *K Δ StH+T* and *K Δ SD+T*, respectively (symbols are as in Fig. 1). Both transgenes produce transported/localized mRNA, as shown here for *K Δ SD+T* (see Fig. 2 legend for in situ hybridization details).

structure, we introduced small mutations into the TLS, designed to disrupt its predicted secondary structure and primary sequence, or only its primary sequence.

The first two constructs, *Kstem5'* and *Kstem3'*, contain mutations that disrupt the TLS primary sequence as well as its putative stem-loop structure. In *Kstem5'*, we replaced five internal uracil residues on the 5' side of the stem with five adenine residues (Fig. 7). Reciprocally, in *Kstem3'*, we replaced five internal adenine residues on the 3' side of the stem with five uracil residues (Fig. 7). As seen in Fig. 7, *Kstem5'* and *Kstem3'* mRNAs are neither transported nor localized. In addition, *K10* mutants carrying one copy of either *Kstem5'* or *Kstem3'* produce eggs that typically hatch with a frequency of $\leq 10\%$ (data not shown). Thus, two separate five nucleotide substitutions in the TLS block its ability to direct mRNA transport/localization, even in the context of an otherwise wild-type *K10* transgene. These results show that the TLS is essential for *K10* function and mRNA transport/localization.

The uracil-to-adenine and adenine-to-uracil substitutions described above could block transport/localization because they alter the TLS primary sequence or because they alter the predicted stem-loop structure. To distinguish between these possibilities, we combined the uracil-to-adenine substitutions of *Kstem5'* with the compensatory adenine-to-uracil substitutions of *Kstem3'* to make *Kstem5'3'* (Fig. 7). *Kstem5'3'* thus carries the primary sequence lesions of both *Kstem5'* and *Kstem3'*, but has the potential to form the same stem-loop structure predicted for wild-type TLS. As seen in Fig. 7, *Kstem5'3'* mRNA is transported/localized almost as efficiently as is wild-type *K10* mRNA. Moreover, *Kstem5'3'* transgenes restore almost complete fertility to *K10* mutant stocks. *K10* mutants that carry one copy of the *Kstem5'3'* transgene produce eggs that typically hatch at a frequency of $\geq 80\%$ (data not shown). These data indicate that extensive alterations of the TLS primary sequence have little or no adverse affect on TLS transport/localization activity, provided that they do not disrupt the predicted stem-loop secondary structure. We conclude that TLS activity is likely to be dependent on the formation of the predicted stem-loop secondary structure.

While the above experiments provide strong evidence that TLS activity is dependent on an encoded stem-loop secondary structure, they do not provide much insight into the nature of the interactions between the TLS and the transport/localization machinery. Given the selectivity of mRNA transport and localization, it is likely that such interactions include base-specific recognition events (St Johnston, 1995). Because the major groove of double-stranded RNA is deep and narrow

(Alden and Kim, 1979; Seeman et al., 1976), we thought it likely that base-specific interactions would be restricted to the single-stranded regions of the predicted stem-loop secondary structure. To look for such interactions, we made *Kloop* and *Kbub*. In *Kloop*, we introduced transition mutations into all eight positions of the predicted TLS loop (Fig. 7). In *Kbub*, we removed the two nucleotides that bubble out from the 3' side of the predicted stem (Fig. 7). As seen in Fig. 7, *Kloop* and *Kbub* both produce transported/localized mRNA, although like *Kstem5'3'*, the transport/localization and rescue activities of *Kloop* and *Kbub* mRNA are slightly less efficient than wild type (Fig. 7 and data not shown). These results indicate that the predicted single-stranded regions of the TLS do not make essential base-specific contacts with the transport/localization machinery. The possibility that other regions of the predicted stem-loop structure, viz., the stem termini, are engaged in such base-specific recognition events is discussed below.

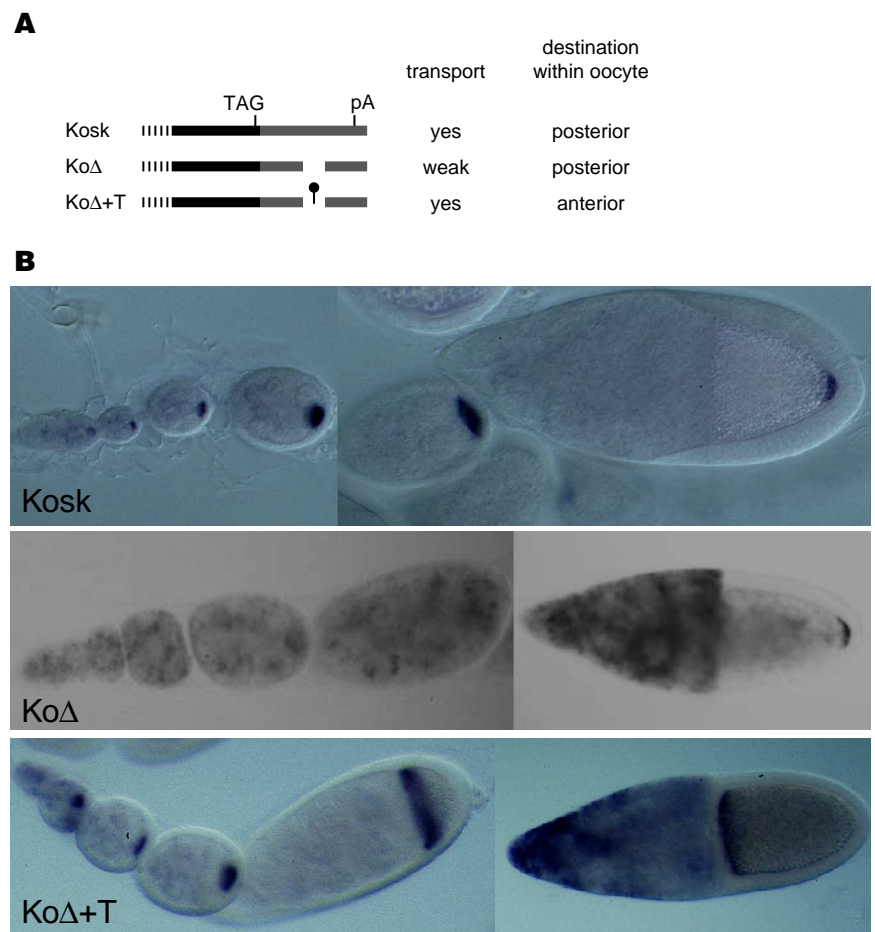


Fig. 6. Schematic diagrams and mRNA distribution patterns of *K10-oskar* hybrid transgenes. (A) The thick black lines represent *K10*-derived sequences. The thick gray lines represent *osk*-derived sequences. The interruption in the line represents the deletion of the ~ 150 nt *SphI-SacII* region, which contains *osk* mRNA transport control sequences (Kim-Ha et al., 1993). The small stem-loop represents the 44 nt *K10* TLS (sequence shown in Fig. 5). The ability of the transgenes to produce transported mRNA (transport) and the final destination of the mRNA within the oocyte (destination within oocyte) are indicated to the right of the diagrams. (B) In situ hybridization montages showing the distribution patterns of *K-osk*, *KoΔ* and *KoΔ+T* mRNA (see Fig. 2 legend for in situ hybridization details).

orb sequence resides in a 280 nt region of *orb* mRNA that is sufficient for *orb* mRNA transport and anterior localization (Lantz and Schedl, 1994). These findings, together with the similarities of *K10* and *orb* mRNA distribution patterns through stage 11 (Cheung et al., 1992; Lantz et al., 1992), suggest that *K10* and *orb* mRNAs are recognized by the same component(s) of the transport/localization machinery. We have been unable to identify candidate TLSs in any other mRNAs known to be localized within the *Drosophila* oocyte (see Fig. 8 for details). This raises the possibility that the transport of at least some mRNAs is mediated by factors different than those that mediate *K10*, and possibly *orb*. We cannot rule out, however, that other localized mRNAs share a common secondary structure motif with the TLS that is not easily identifiable by a comparison of primary sequences.

While most localized mRNAs appear to be transported into the oocyte in a similar fashion, our results suggest that the mRNA transport control sequences of different mRNAs are not functionally equivalent. For example, the *K10* TLS can substitute for *osk* mRNA transport control sequences, but it is apparently incompatible with the activity of *osk* mRNA posterior localization control elements. The observed incompatibility may occur for superficial reasons, e.g., the inclusion of the TLS in the *osk* 3'UTR may alter the secondary structure of *osk* posterior localization control sequences. Alternatively, *osk* mRNA transport control sequences, unlike the *K10* TLS, may facilitate the release of the mRNA from a minus end-directed microtubule motor at stage 8, thus allowing *osk* posterior localization control sequences to mediate the association of *osk* mRNA with a plus end-directed microtubule motor.

mRNA transport and localization control elements and RNA secondary structure

The TLS has a predicted secondary structure consisting of a 8 base loop and an imperfect, 17 base pair stem – one strand of the stem contains two one-base bubbles (Fig. 5B). Consistent with the idea that TLS activity is dependent on the formation of such a stem-loop structure, we find that TLS transport/localization activity is abolished when base changes are introduced into one of the two stem strands alone, but not when simultaneously introduced into both strands in a compensatory manner that preserves the predicted secondary structure. Other mRNA transport/localization control sequences have also been proposed to form secondary structures. Most notably, the 625 nt *bcd* mRNA localization signal is predicted to form an elaborate secondary structure, consisting of five large stems (Macdonald and Struhl, 1988; Macdonald, 1990; Ferrandon et al., 1994). This is supported by the finding that, while the primary sequence of the *bcd* localization signal varies significantly between a number of *Drosophila* species, the putative secondary structure remains conserved (Macdonald, 1990). In addition, *stau* protein, which mediates *bcd* and *osk* mRNA localization, exhibits double-stranded RNA-binding activity in vitro (St. Johnston et al., 1992). Together, these findings suggest that secondary structure formation is a common feature of mRNA transport/localization control sequences.

Given the selective nature of mRNA transport/localization, it is likely that recognition of the TLS by the transport/localization machinery requires base-specific contacts. A complete base substitution of the predicted loop, or the removal of the two single nucleotide bubbles from the 3' side of the stem, have

only a slight negative affect on TLS activity, indicating that these regions do not provide essential base-specific contacts. It could be that residues of the stem itself make base-specific contacts with the transport/localization machinery. The minor groove of double-stranded RNA is wide and accessible, but it contains few hydrogen bond acceptor and donor groups and the potential for base-specific hydrogen bonding is very low (Seeman et al., 1976). One possible base-specific hydrogen bond acceptor/donor group in the minor groove is the carbonyl group of any uracil immediately 3' to adenine (Rosenberg et al., 1973). The TLS stem contains two such groups. The predicted *orb* stem (Fig. 8) also contains several such groups, but at non-conserved positions. Assuming that the overall conservation between the predicted stem loop structures of *K10* and *orb* is not fortuitous, it would thus appear that the minor groove does not play a major role in the base-specific recognition of the stem by the transport/localization machinery. In contrast to the minor groove, the major groove of double-stranded RNA has a rich, base-specific collection of hydrogen bond acceptor and donor groups (Seeman et al., 1976). While the major groove of double-stranded RNA is narrow and generally not accessible to small, amino acid-size molecules (Alden and Kim, 1979), chemical acylation studies show that bases positioned near a loop or bubble of two or more nucleotides are accessible, presumably due to a local distortion of the helix (Weeks and Crothers, 1993). Accessibility extends into the duplex about 2 nucleotides on one strand and about 4 nucleotides on the other. The bases adjacent to the TLS loop are conserved in *orb* (Fig. 8) and thus are good candidates for base-specific interactions with the transport/localization machinery. Residues at the base of the stem are not conserved and thus apparently less likely to make specific contacts with the transport/localization machinery.

The protein components of the transport/localization machinery have remained elusive. Genetic analyses have identified only two genes, *Bic-D* and *egalitarian* (*egl*), that are required for *K10* mRNA transport into the oocyte (Suter and Steward, 1991; Cheung et al., 1992). However, it is unclear whether *Bic-D* and *egl* encode proteins that bind directly to the TLS. These genes are also required for oocyte determination and microtubule organization, developmental events that are likely to be a prerequisite for mRNA transport (Schüpbach and Wieschaus, 1991; Theurkauf et al., 1993). Thus, biochemical approaches may be most useful for identifying the proteins that directly contact the *K10* TLS.

We thank Paul Macdonald for providing the *osk* clone and Vicki Corbin, Xiangyi Lu, Doug Ruden, Kathy Suprenant, Rob Weaver and members of the Cohen laboratory for helpful discussions and comments on the manuscript. T. L. S. was supported in part by NIH training grant 5T32EY07105. R. S. C. was supported by a grant from the NSF (IBN-08821) and by the General Research Fund of the University of Kansas.

REFERENCES

- Alden, C. J. and Kim, S.-H. (1979). Solvent-accessible surfaces of nucleic acids. *J. Mol. Biol.* **132**, 411-434.
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nüsslein-Volhard, C. (1988). The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO J.* **7**, 1749-1756.

- Cheung, H.-K., Serano, T. L. and Cohen, R. S.** (1992). Evidence for a highly selective RNA transport system and its role in establishing the dorsoventral axis of the *Drosophila* egg. *Development* **114**, 653-661.
- Clark, I., Giniger, E., Ruohola-Baker, H., Jan, L. Y. and Jan, Y. N.** (1994). Transient posterior localization of a kinesin fusion protein reflects anteroposterior polarity of the *Drosophila* oocyte. *Curr. Biol.* **4**, 289-300.
- Cohen, R. S. and Meselson, M.** (1985). Separate regulatory elements for the heat-inducible and ovarian expression of the *Drosophila hsp26* gene. *Cell* **47**, 737-743.
- Cohen, R. S. and Serano, T. L.** (1995). mRNA localization and function of the *Drosophila fs(1)K10* gene. In *RNA Localization* (ed. H. D. Lipshitz). In press. Austin: R. G. Landes.
- Cooley, L. and Theurkauf, W. E.** (1994). Cytoskeletal functions during *Drosophila* oogenesis. *Science* **266**, 590-596.
- Ding, D. and Lipshitz, H. D.** (1993). Localized RNAs and their functions. *BioEssays* **15**, 651-658.
- Ephrussi, A., Dickinson, L. K. and Lehmann, R.** (1991). *oskar* organizes the germ plasm and directs localization of the posterior determinant *nanos*. *Cell* **66**, 37-50.
- Feinberg, A. P. and Vogelstein, B.** (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**, 6-13.
- Ferrandon, D., Elphick, L., Nüsslein-Volhard, C. and St. Johnston, D.** (1994). Staufen protein associates with the 3'UTR of *bicoid* mRNA to form particles that move in a microtubule-dependent manner. *Cell* **79**, 1221-1232.
- Frank, L. H., Cheung, H.-K. and Cohen, R. S.** (1992). Identification and characterization of *Drosophila* female germ line transcriptional control elements. *Development* **114**, 481-491.
- Haenlin, M., Roos, C., Cassab, A. and Mohier, E.** (1987). Oocyte-specific transcription of *fs(1)K10*: a *Drosophila* gene affecting dorsal-ventral developmental polarity. *EMBO J.* **6**, 801-807.
- King, R. C.** (1970). *Ovarian Development in Drosophila melanogaster*. New York: Academic Press.
- 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.
- Kim-Ha, J., Webster, P. J., Smith, J. L. and Macdonald, P. M.** (1993). Multiple RNA regulatory elements mediate distinct steps in localization of *oskar* mRNA. *Development* **119**, 169-178.
- Lantz, V., Ambrosio, L. and Schedl, P.** (1992). The *Drosophila orb* gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. *Development* **115**, 75-88.
- Lantz, V. and Schedl, P.** (1994). Multiple cis-acting target sequences are required for *orb* mRNA localization during *Drosophila* oogenesis. *Mol. Cell. Biol.* **14**, 2235-2242.
- Lindsley, D. L. and Zimm, G. G.** (1992). *The Genome of Drosophila melanogaster*. New York: Academic Press.
- Macdonald, P. M.** (1990). *bicoid* mRNA localization signal: phylogenetic conservation of function and RNA secondary structure. *Development* **110**, 161-171.
- Macdonald, P. M.** (1992). The means to the ends: localization of maternal messenger RNAs. *Sem. Dev. Biol.* **3**, 413-424.
- Macdonald, P. M. and Struhl, G.** (1988). Cis-acting sequences responsible for anterior localization of *bicoid* mRNA in *Drosophila* embryos. *Nature* **336**, 595-598.
- Macdonald, P. M., Kerr, K., Smith, J. L. and Leask, A.** (1993). RNA regulatory element BLE1 directs the early steps of *bicoid* mRNA localization. *Development* **118**, 1233-1243.
- Neuman-Silberberg, F. S. and Schüpbach, T.** (1993). The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF α -like protein. *Cell* **75**, 165-174.
- Pelham, H. R. B. and Munro, S.** (1993). Sorting of membrane proteins in the secretory pathway. *Cell* **75**, 603-605.
- Pirrotta, V.** (1988). Vectors for P-mediated transformation in *Drosophila*. In *Vectors: A Survey of Molecular Cloning Vectors and Their Uses* (eds. R. L. Rodriguez and D. T. Denhardt), pp. 437-456. Boston: Butterworths.
- Pokrywka, N. J. and Stephenson, E. C.** (1991). Microtubules mediate the localization of *bicoid* mRNA during *Drosophila* oogenesis. *Development* **113**, 55-66.
- Pokrywka, N. J. and Stephenson, E. C.** (1995). Microtubules are a general component of mRNA localization systems in *Drosophila* oocytes. *Dev. Biol.* **167**, 363-370.
- Prost, E., Deryckere, F., Roos, C., Haenlin, M., Pantescio, V. and Mohier, E.** (1988). Role of the oocyte nucleus in determination of the dorsoventral polarity of *Drosophila* as revealed by molecular analysis of the *K10* gene. *Genes Dev.* **2**, 891-900.
- Rosenberg, J. M., Seeman, N. C., Kim, J. J. P., Suddath, F. L., H. B. and Rich, A.** (1973). Double helix at atomic resolution. *Nature* **243**, 150-154.
- Rubin, G. M. and Spradling, A. C.** (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348-353.
- Schüpbach, T. and Wieschaus, E.** (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**, 1119-1136.
- Seeman, N. C., Rosenberg, J. M. and Rich, A.** (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA* **73**, 804.
- Serano, T. L., Cheung, H.-K., Frank, L. H. and Cohen, R. S.** (1994). P element transformation vectors for studying *Drosophila melanogaster* oogenesis and early embryogenesis. *Gene* **138**, 181-186.
- Serano, T. L., Karlin-McGinness, M. and Cohen, R. S.** (1995). The role of *fs(1)K10* in the localization of the mRNA of the TGF α homolog *gurken* within the *Drosophila* oocyte. *Mech. Dev.* **51**, 183-192.
- Southgate, R., Ayme, A. and Voellmy, R.** (1983). Nucleotide sequence analysis of the *Drosophila* small heat shock gene cluster at locus 67B. *J. Mol. Biol.* **165**, 35-57.
- Spadling, A. C. and Rubin, G. M.** (1982). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**, 341-347.
- St Johnston, D.** (1995). The intracellular localization of messenger RNAs. *Cell* **81**, 161-170.
- St Johnston, D., Driever, W., Berleth, T., Richstein, S. and Nüsslein-Volhard, C.** (1989). Multiple steps in the localization of *bicoid* RNA to the anterior pole of the *Drosophila* oocyte. *Development* **107 Supplement**, 13-19.
- St Johnston, D., Brown, N. H., Gall, J. G. and Jantsch, M.** (1992). A conserved double-stranded RNA-binding domain. *Proc. Natl. Acad. Sci. USA* **89**, 10979-10983.
- Suter, B. and Steward, R.** (1991). Requirement for phosphorylation and localization of the Bicaudal-D protein in *Drosophila* oocyte differentiation. *Cell* **67**, 917-926.
- Tautz, D. and Pfeifle, C.** (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81-85.
- Theurkauf, W. E., Smiley, S., Wong, M. L. and Alberts, B. M.** (1992). Reorganization of the cytoskeleton during *Drosophila* oogenesis: implications for axis specification and intercellular transport. *Development* **115**, 923-936.
- Theurkauf, W. E., Alberts, B. M., Jan, Y. N. and Jongens, T. A.** (1993). A central role for microtubules in the differentiation of *Drosophila* oocytes. *Development* **118**, 1169-1180.
- Weeks, K. M. and Crothers, D. M.** (1993). Major groove accessibility of RNA. *Science* **261**, 1574-1577.
- Zimmerman, J. L., Petri, W. and Meselson, M.** (1983). Accumulation of a specific subset of *D. melanogaster* heat shock mRNAs in normal development without heat shock. *Cell* **32**, 1161-1170.