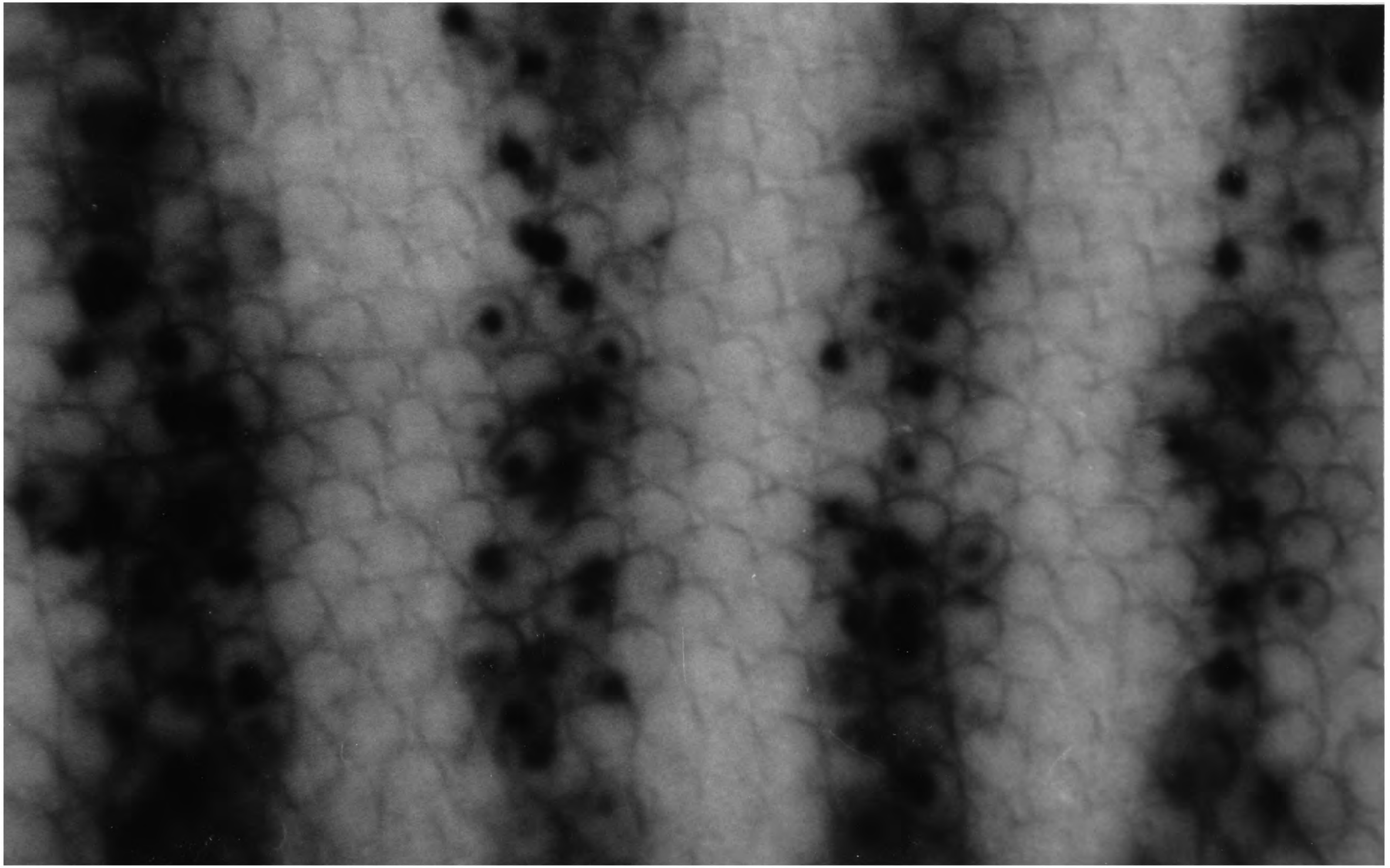


Intracellular Message Localisation in *Drosophila melanogaster*

Ilan Davis
New College
Oxford

Submitted to the University of Oxford
for the degree of Doctor of Philosophy
Trinity 1990

Imperial Cancer Research Fund, Developmental Biology Unit,
Department of Zoology, University of Oxford

To my Parents

As flies to wanton boys are we to the gods,
They kill us for their sport.

(Gloucester in King Lear)

TABLE OF CONTENTS

PREFACE

I performed the experiments described in this thesis at the Developmental Biology Unit of the Imperial Cancer Research Fund in the Zoology Department, University of Oxford, from October 1986 to July 1990. I would like to thank the Imperial Cancer Research Fund for their generous financial support.

I am greatly indebted to my supervisor, David Ish-Horowicz for providing an extremely stimulating intellectual environment and for much support and advice.

I would also like to thank all my colleagues at the DBU for their considerable support.

In particular, I would like to thank Phil Ingham, Susan Parkhurst, Guy Riddihough, Alicia Hidalgo, Sheena Pinchin, Mark Wainwright, Julian Lewis, Paul Martin, Mike Peacy, Keith Morris, Kate Davis, Pam Hickey, Ann Yates, Ian McKay and Jonathan Slack.

Finally, I am grateful to my parents and Alex for their unwavering support and encouragement.

ABSTRACT

The blastoderm embryo of *Drosophila melanogaster* consists of a unicellular syncytium with a large number of peripheral nuclei. The cytoplasm surrounding each peripheral nucleus is compartmentalised into apical periplasm above each nucleus and basal periplasm below it. The expression of different genes in the syncytial blastoderm is crucial for the genetic control of development. The pair-rule genes are involved in controlling the pattern of metamerisation of the embryo. Pair-rule mRNAs are expressed in alternate metameres, in a pattern of stripes. Within each stripe, mRNA is found in the apical periplasm of the syncytial blastoderm.

By analysing the distribution of mRNA of a number of hybrid constructs, I show that the 3' untranslated part of three pair-rule genes are required for the apical localisation of their transcripts. A 1.2kb region in the 3' end of *fushi tarazu* (*ftz*), a 700bp region in the 3' end of *hairy (h)* and a 160bp fragment of the 3' untranslated part of the *even-skipped* (*eve*) pair-rule gene are shown to contain apical localisation signals.

I show that the mechanism of apical localisation is unlikely to involve a cytoplasmic process and that the 3' untranslated part of the *bicoid* (*bcd*) gene contains sequences necessary for apical localisation. I propose that apical localisation involves a nuclear mechanism which exports mRNA from the apical side of the nuclear membrane. I demonstrate that apical localisation is achieved by an RNA-mediated process and not by a DNA-mediated mechanism.

Finally, I demonstrate that the intracellular localisation of transcripts encoding cytoplasmic proteins influences the distribution of the protein in the periplasm. I propose that the function of apical localisation is to limit the diffusion of pair-rule proteins so that the pattern of protein expression resembles precisely the transcriptional domain.

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Asymmetric intracellular transcript localisation may be a general mechanism for targeting proteins to specific subcompartments in the eukaryotic cell. In several cases, mRNAs are localised at subcellular sites where the cytoplasmic proteins encoded by the transcripts function.

The *Drosophila* blastoderm embryo develops as a unicellular syncytium containing a diverse range of well characterised transcripts, many of which adopt distinct asymmetric patterns of intracellular localisation. *Drosophila* provides many powerful genetic and molecular techniques which could be used to determine the mechanism and function of the intracellular distribution of these transcripts.

During the blastoderm stage, an extensive hierarchy of genes controls the process of segmentation. Mutual interactions between these genes establish progressively finer domains of expression in different parts of the embryo. The expression domains define a molecular prepattern which determines the formation of segments and the assignment of their unique identities.

The pair-rule genes are members of the hierarchy which are expressed in seven stripes in the syncytial blastoderm. Within each stripe, pair-rule transcripts are asymmetrically distributed in regions of cytoplasm closely associated with the blastoderm nuclei.

This thesis is concerned with three aspects of the intracellular localisation of pair-rule transcripts: the sequences responsible for intracellular localisation, the cellular mechanism for apical localisation and the function of localisation. The thesis draws together two related disciplines: developmental genetics and cell biology. A full evaluation of the work described below, necessitates a review of early *Drosophila* development as

well as a discussion of subcellular localisation of mRNAs and proteins in the eukaryotic cell.

In this chapter, an outline of early *Drosophila* development is presented, followed by a description of the genetic control of pattern formation. The pattern of intracellular transcript localisation of several genes expressed in the *Drosophila* blastoderm is described before considering the intracellular localisation of mRNAs and proteins in a number of vertebrate cell types. At the end of the introduction, a number of models for the mechanism of apical localisation are proposed. These models are tested in subsequent chapters.

Pattern formation in early development

One of the central questions in biology is how a single cell, the fertilised egg, develops to form a multicellular organism. This process involves the formation of complex pattern from simple initial conditions. The study of the molecular genetics of *Drosophila* development has shed substantial light on the problem. In particular, the identification of approximately 50 mutations (Nüsslein-Volhard and Wieschaus, 1980; Nüsslein-Volhard *et al.*, 1984) affecting the larval cuticular pattern has led to the cloning and molecular analysis of genes which set up segments and determine their identity in the *Drosophila* larva (reviewed in Ingham, 1988). The strength of this work has stemmed from the combination of genetic analysis of mutants with a molecular characterisation of the pattern of expression of the genes in wild-type and mutant backgrounds.

The unfertilised egg is inhomogeneous; during oogenesis an asymmetric egg shell and an unequal distribution of molecular components are laid down. These asymmetries are vital as they trigger the cascade of gene expression which forms the presumptive body plan of the embryo. Before

examining the genetic control of development we must consider how the asymmetric distributions of molecules are established during oogenesis.

The stages of oogenesis

During oogenesis the female produces a large number of eggs which are fertilised with stored sperm just before they are laid. There are two ovaries, consisting of a cluster of 15-17 ovarioles, each of which contains a string of oocytes leading to the oviduct where mature eggs are released.

The oocytes originate at the tip of the ovariole by oogonial stem cell divisions giving rise to cystoblasts, each of which undergoes 4 sequential mitotic divisions (Mahowald and Kambysellis, 1980). The resulting 16 cells are connected by a characteristic series of cytoplasmic canals which are the result of incomplete cytokinesis; each canal lies at the site of an earlier cleavage furrow. Two of the cells develop as pre-oocytes, each of which has a synaptonemal complex required for chromosome pairing during meiosis. One of these becomes an oocyte, the other reverts to the nurse cell pathway adopted by the other 14 cells, and loses its synaptonemal complex (Mahowald and Kambysellis, 1980).

The development of the oocyte is divided into 14 stages (King, 1970). In stages 1-6 the oocyte and nurse cells increase in volume by approximately 40 times but remain equal in size to each other. In stages 8-12 the oocyte doubles in volume every two hours. The oocyte increases in volume 1500 times during stages 1-12. During stages 13 and 14 the vitelline membrane and chorion are secreted on the outside of the oocyte to form a mature egg. The egg is asymmetric in its appearance having two distinct perpendicular axes, the antero-posterior (A-P) and the dorso-ventral (D-V).

Most of the increase in the volume of the oocytes is due to transfer of cytoplasm from the nurse cells. The cytoplasmic material passes through

channels large enough to transport mitochondria. rRNA is transferred in assembled ribosomes; small amounts of mRNA are transferred as well as certain proteins. The flow of cytoplasmic material into the oocyte is crucial for establishing the asymmetries which trigger early embryonic development.

Early embryonic development

The mature egg is fertilised before being laid and begins a remarkably rapid development which leads to a crawling larva in twenty four hours. Early development occurs as a syncytium; the embryonic nuclei undergo nuclear cleavage with no cytokinesis during the first 3 hours of development (at 25°C). 13 nuclear divisions occur before the onset of gastrulation which coincides with the formation of cell membranes.

During the first eight divisions the nuclei are situated in the centre of the embryo surrounded by yolk. Most of the nuclei then migrate to the periphery while some remain in the centre to form polyploid yolk nuclei. On reaching the periphery, those nuclei that migrate towards the posterior pole become the pole nuclei and are the first nuclei to cellularise. The pole cells are the future germ cells of the fly which will emerge after the larva pupates. The remaining peripheral nuclei undergo five more synchronous nuclear divisions to form approximately 5000 blastoderm nuclei (Campos-Ortega and Hartenstein, 1985).

The first 13 nuclear divisions are globally synchronous and occur in waves emanating from the centre of the embryo and ending at the two poles. After the 14th interphase the divisions are no longer globally synchronous; the embryo can be divided into at least 25 mitotic domains which are locally synchronous and occur in reproducible position in the A-P and D-V axes. Most domains occupy two bilateral clusters of cells in mirror-image

symmetry, some of the domains occupy a single region that spans the dorsal or ventral midline and a few groups occur in metamerically repeated clusters. Within each domain the division occurs in a wave like pattern emanating from the centre of the cluster and ending at its boundary (Foe, 1989).

The late syncytial blastoderm consists of a single layer of approximately 5000 nuclei. Gastrulation begins soon after the end of cellularisation and involves the involution of regions of the blastoderm at specific furrows. Invaginations at the ventral furrow form mesoderm, ectoderm and neural ectoderm. The endoderm (gut) is generated later in development by further involution. The germ layers thus formed are the presumptive segmented thoracic and abdominal region of the larva and fold back on themselves to form the extended germ-band (Campos-Ortega and Hartenstein, 1985, and see Fig. 1.1). At this stage overt morphological signs of segmentation first appear as small furrows in the extended germ-band. The formation of the head is more complex and is not addressed in this thesis. Cuticle is secreted by the epidermis at the end of embryonic development (stage 16).

Experiments, which use x-ray induced mitotic crossing-over to generate clones of genetically marked cells in the wing, show that the adult segments are divided into anterior and posterior compartments derived from clonally distinct groups of cells (Garcia-Bellido *et al.*, 1973). Cells which express the *engrailed (en)* gene form the posterior compartment of a segment (see later for a description of *en*). Mitotic recombination is difficult to observe in the larva because only two to three cell divisions occur after the end of the blastoderm stage. However, it is suggested that anterior and posterior compartments also exist in early development because *en* is expressed in a 'compartment-like' manner at all stages.

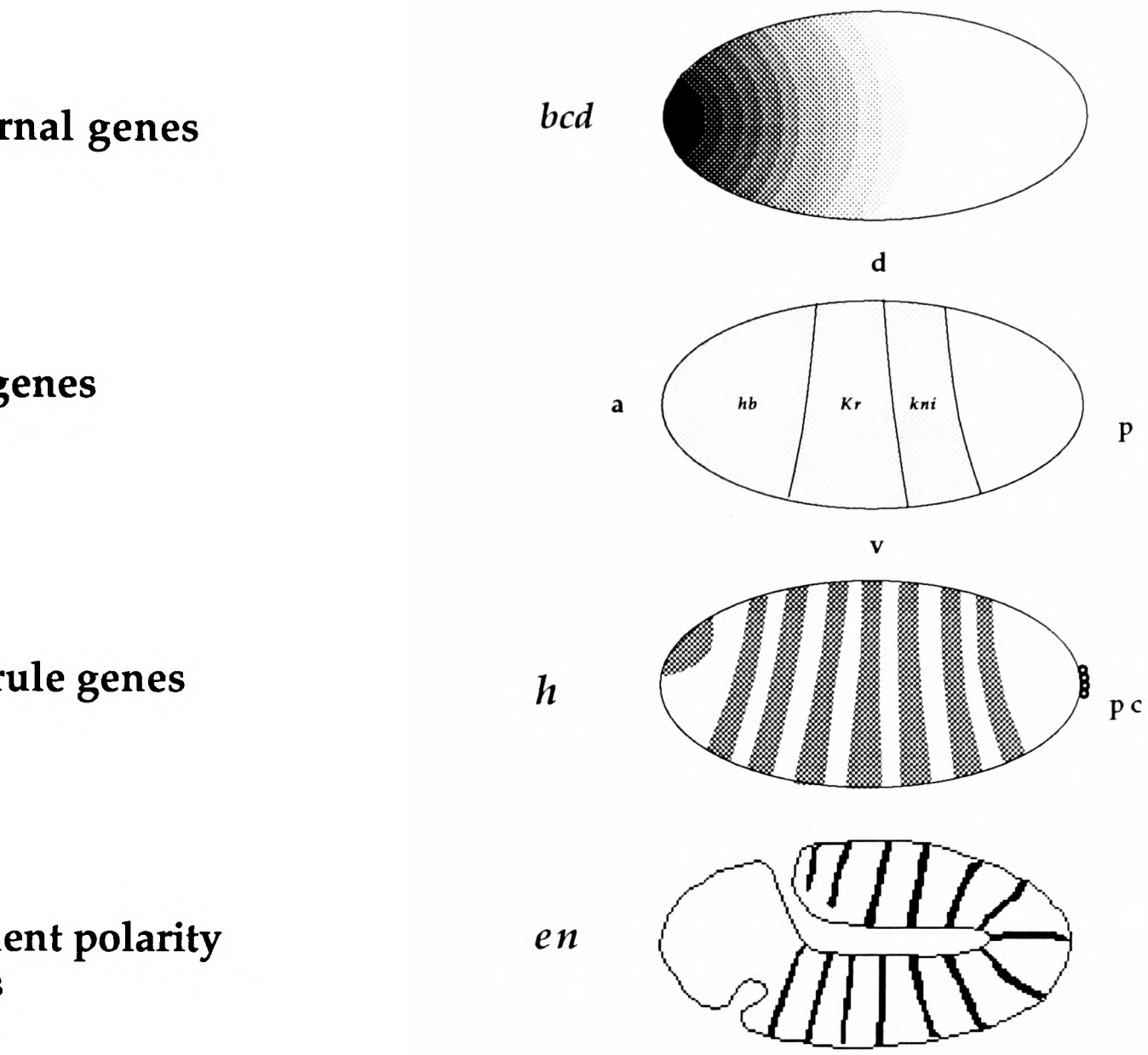


Fig 1.1

The hierarchy of genes which controls the formation of segments and the asignment of their unique identities. A simplified diagrammatic description of the pattern of protein expression of representative genes of each group is shown. The top three embryos are of blastoderm stage, the bottom two embryos at germ band extension. The maternal *bcd* gradient is interpreted by the gap genes to form broad overlapping domains of expression. The pair-rule genes interpret the pattern of gap gene expression and mutually interact to form a seven striped pattern (and a head patch in the case of *h*). The segment polarity genes interpret the overlapping domains of pair-rule genes to form fourteen stripes at germ band extension. *en* is expressed in the posterior compartment of each segment which is the anterior of each parasegment. The homeotic genes respond to information at all levels of the hierarchy and determine the identity of segments in the extended germ band. The pattern of *Ubx* expression is shown only in the ectoderm where it is strongest in parasegment 6. pc: pole cells, 6: position of parasegment 6 a: anterior, p: posterior, d: dorsal, v: ventral (all subsequent embryos are shown in the same orientation).

The larva and adult fly are divided into segments. Parasegments are metameric units which are one compartment out of register with segments, so that the anterior compartment of each parasegment is the posterior compartment of each segment. In the early embryo, the parasegment is the predominant metameric unit as judged by the pattern of expression of developmental control genes in the embryo (Martinez-Arías and Lawrence, 1985).

The pattern of gene expression evolves extremely rapidly in the first 3 hours of development. During this period, approximately 50 genes involved in the control of development interact in a hierarchy to define progressively smaller regions of the embryo.

Mutations altering larval segmentation.

Before considering how genes control early development it is worth reviewing the different classes of mutations which have been classified on the basis of alterations to the larval cuticle phenotype.

In 1980, the first extensive mutagenic screen for genes affecting metamerisation was undertaken. The screen revealed approximately 15 zygotic genes which, when mutated, altered the number and polarity of the larval segments (Nüsslein-Volhard and Wieschaus, 1980). The segmentation genes, as they were termed, could be classified into three distinct groups on the basis of changes in the pattern of segmentation. One group, the gap genes, cause large deletions 2-6 segments wide. Mutations in the gap gene *hunchback (hb)* for example, lead to the deletion of all the thoracic segments, some gnathal and one abdominal segment (Lehmann and Nüsslein-Volhard, 1987).

Mutations in the second group, the pair-rule genes, cause reiterated homologous deletions in alternate segments. The domains deleted in each

of the pair-rule genes are different but in some cases overlapping. The *hairy (h)* pair-rule mutant phenotype, for example, has deletions of alternate segment wide domains consisting of the anterior part of every other even numbered segment and the posterior part of every other odd numbered segment (see chapter 7).

The third group, the segment polarity genes, induce reiterated small deletions within each segment and the mirror-image duplication of the remaining part. *patched (ptc)* for example, leads to the deletion of the posterior part of each denticle belt and its replacement by a mirror-image duplication of the remaining anterior part of of each denticle belt (Nakano *et al.*, 1989). Subsequent screens revealed further members of these three groups of genes and have probably uncovered the majority of such genes (Jürgens *et al.*, 1984; Nüsslein-Volhard *et al.*, 1984; Wieschaus *et al.*, 1984).

In addition to the zygotic segmentation genes, a number of maternal effect mutations have been described which alter the pattern of segmentation in the larva. These are divisible into three distinct classes: the anterior group, the posterior group and the terminal group of genes. As their names imply the classes of mutations lead to deletions in the anterior, posterior and terminal regions of the larva respectively (reviewed in Nüsslein-Volhard *et al.*, 1987).

Several genes have been isolated which, when mutated, alter the identity of segments but do not affect the number of segments. These genes were termed homeotic genes as they lead to transformation of one parasegment to another parasegment in a different part of the body. The product of *Ubx*, the best characterised homeotic gene, is required for the assignment of the identity of parasegment 6. In a *Ubx* null mutant parasegment 6 is transformed to parasegment 5 (reviewed in Akam, 1987).

The formation of segments and the assignment of their identity is a process confined to the A-P axis of the embryo. Patterning in the D-V axis is regulated independently of the A-P axis, by a large set of maternal and zygotic genes. Mutations in these genes lead to transformations from dorsal to ventral characters or *vice versa*. The control of the D-V axis is not considered in this thesis.

The genetic control of segmentation

Many of the segmentation and homeotic genes have now been cloned and the patterns of their mRNA and protein distribution have been analysed in great detail. An extensive network of interactions has been drawn-up from studies of the expression of genes in a variety of mutant backgrounds (reviewed in Ingham, 1988). The hierarchy of genes is clearly vital to the control of early development since the expression data is backed by strong genetic evidence for the role of each gene, as well as evidence from the functional effect of ectopically expressing some of these genes (Ish-Horowicz and Gyurkovics, 1988; Ish-Horowicz and Pinchin, 1987).

The expression and genetic data reveal the *de novo* formation of a molecular prepattern which delineates a presumptive body plan in the blastoderm embryo (see Fig.1.1). The whole process is triggered by maternal effect genes that broadly specify the A-P poles of the embryo. The gap genes respond to the maternal information and are expressed in broad domains several segments in width. The pair-rule genes are expressed later than the gap genes in a reiterated pattern and specify fine spatial information due to subtle differences in their relative phasing. The segment polarity genes respond to this combinatorial information and are expressed in most segments and together specify the identity of individual cells within each segment (Ingham *et al.*, 1988). The homeotic genes respond to information

from many levels in the hierarchy of segmentation genes so that each gene is expressed at its highest levels in a particular segment, thus assigning it a unique identity.

Maternal morphogens.

bicoid (bcd), is the best characterised of the anterior group of maternal effect genes. In embryos from mutant mothers homozygous for *bcd*, the head and thorax are lacking and are replaced by a telson, a posterior terminal structure. There is overwhelming evidence to suggest that *bcd* is the morphogen which specifies the A-P axis (Driever and Nüsslein-Volhard, 1989; Frohnhöfer and Nüsslein-Volhard, 1986). Injection of cytoplasm from the anterior of a wild-type embryo can rescue the mutant phenotype of a *bcd* embryo if injected into the anterior of the embryo. Anterior cytoplasm induces ectopic anterior structures at the site of injection in the posterior of wild type embryos. The extent of the transformation to anterior structures and rescue of mutant phenotypes is dependent on the amount of cytoplasm injected and on the number of maternal copies of the wild-type *bcd* in the donor embryo (Frohnhöfer and Nüsslein-Volhard, 1986).

bcd mRNA is localised in a tight crescent at the anterior of the blastoderm embryo where it is translated. As a result, the highest levels of *bcd* protein are found in the apical pole. The protein diffuses from the site of translation and is degraded in all parts of the embryo, thus forming a steady-state exponential gradient.

The amount of *bcd* mRNA in the anterior pole and thus the peak and steepness of the morphogenetic gradient can be manipulated by genetic means. Such experiments show that a shift in positional values occurs in the embryo and is directly correlated with the concentration of the *bcd* protein in the gradient (Driever and Nüsslein-Volhard, 1988a; Driever and

Nüsslein-Volhard, 1988b). It follows that the levels of *bcd* protein determines positional values along the A-P axis.

The *bcd* protein contains a homeobox domain and acts as a transcription factor by binding upstream of *hb* and possibly other genes (Driever and Nüsslein-Volhard, 1989; Struhl *et al.*, 1989). *bcd* is thought to act via a number of thresholds for activating the transcription of different gap genes. The final pattern of gap gene transcription is produced by the mutual interaction of gap genes (see Fig. 1.1).

The posterior group genes all share the same mutant phenotype in which the posterior part of the larval cuticle is deleted (reviewed in Nüsslein-Volhard *et al.*, 1987). Injection experiments which use wild type cytoplasm to rescue the mutant phenotype indicate that posterior determinants behave differently from *bcd*. Although the rescuing activity is localised in the posterior of the donor embryo, the cytoplasm must be injected into the abdominal region of the embryo and the site of injection has little effect on the position of the posterior pole (Nüsslein-Volhard *et al.*, 1987). Rescue experiments performed with cytoplasm from donor nurse cells reveal that the rescuing activity is present in all the mutants except for *nanos (nos)*. It has thus been suggested that *nos* is the posterior signal through which all the other posterior genes exert their effect. However, experiments which give rise to embryos which lack maternal *hb* and *nos* lead to wild type cuticle pattern (Hülskamp *et al.*, 1989). It is suggested that *nos* is not a posterior morphogen and that its function is to remove maternal *hb* mRNA from the posterior of the embryo. The other posterior group genes are thought to be involved in the localisation of the *nos* activity to the abdomen.

The hierarchy of zygotic genes

The gap genes interpret the maternal cues and are expressed in overlapping wide regions in the embryo (see Fig. 1.1). These domains of expression are dependent on activation and repression of gap genes by *bcd* as well as mutual inhibition by adjacent gap genes (Gaul and Jäckle, 1989; Pankratz *et al.*, 1989). There are five well characterised gap genes: *hb, Krüppel (Kr), knirps (kni), giant (gt),* and *tailless (tl). hb, Kr* and *kni* are homologous to 'zinc-finger' transcription factors and are thought to bind upstream of pair-rule genes (reviewed in Ingham, 1988).

The domains of RNA expression of the gap genes *Kr* and *hb* are much narrower than the domains deleted when they are mutated. The protein expression domains are much wider than the region where RNA is transcribed and are expected to be as wide as the deletion domains (Gaul and Jäckle, 1989; Pankratz *et al.*, 1989). The difference is probably due to diffusion of protein out of the region where the genes are translated. Indeed it has been argued that the subtle gradients formed at the edges of the expression domains extend beyond the threshold of detection and are crucial for gap gene function. It is suggested that gap gene transcription factors bind upstream of different genes at different protein concentrations in the gradient.

The pair-rule genes play a central role in the refinement of spatial information from the broad bands of expression of the gap genes to the single cell wide domains of segment polarity gene expression (reviewed in Ingham and Gergen, 1988). Mutations in the eight known pair-rule genes lead to deletions in alternate segments. The size and position of the regions deleted is different in each of the cases described. There is a close correlation

between the regions deleted when the pair-rule genes are mutated and their pattern of expression of proteins and mRNAs in the blastoderm.

fushi tarazu (ftz), evenskipped (eve), hairy (h), runt (run) and paired (prd) are the best characterised pair-rule genes. Their transcripts are first expressed at cell cycle 13 of the blastoderm embryo and resolves into seven stripes, 3-4 nuclei in width at cycle 14. *h* and *run* have a head patch in addition to the seven stripes (reviewed in Ingham, 1988). The patterns of expression of the pair-rule genes are thought to be transcriptionally regulated by gap genes activating or repressing particular stripes as well as by mutual interactions between the pair-rule genes. It is suggested that combinations of transcription factors bind to stripe elements near the promoters of the genes so as to activate or repress transcription in those stripes (Pankratz *et al.*, 1990).

Pair-rule proteins are expressed in seven stripes which correspond accurately with those of the transcripts (reviewed in Ingham, 1988). *h* protein is also expressed in a headpatch at the blastoderm stage (Hooper *et al.*, 1989). The protein domains of the different genes cover the entire central segmented region of the embryo in partially overlapping regions. The relative phasing of pair-rule gene domains of expression are believed to be crucial for precise activation of segment polarity gene expression.

ftz, eve and *h* encode nuclear proteins with a short half-life of approximately 5min (Edgar *et al.*, 1986). *ftz* and *eve* have homeobox DNA binding motifs. Transient expression assays in *Drosophila* tissue culture cells using a fragment of the *en* promoter region suggest that *ftz* and *eve* are transcription factors which bind near the *en* promoter (Kyuhyung *et al.*, 1989).

h has substantial homology to helix loop helix proteins such as the *N-myc* proto-oncogene and *daughterless,* a *Drosophila* sex determination gene

(Rushlow *et al.*, 1989). This suggests that *h* is also a transcription factor although this has not yet been directly demonstrated.

The analysis of the pattern of expression of different pair-rule genes in mutant embryos of other pair-rule genes reveals that the eight known pair-rule genes are not all equivalent; they form a hierarchy. *eve, h* and *run* are primary pair-rule genes at the top of the hierarchy and are unaffected by mutations in the other pair-rule genes. *ftz, prd, oddpaired, sloppypaired, oddskipped* are secondary pair-rule genes at the bottom of the hierarchy and their expression is altered by mutations in the primary pair-rule genes (reviewed in Ingham, 1988).

The combination of expression studies in mutant embryos with transcription factor analysis in tissue culture, as well as other lines of evidence, has lead to the following hypothetical interactions. *h* represses *ftz* in the interstripes and *ftz* appears to activate *en* (Howard and Ingham, 1986). *eve* also appears to activate *en* but other pair-rule genes may repress *en*. It is suggested that the presence of different combinations of pair-rule genes in different regions of the embryo causes *en* to be expressed in fourteen single cell wide stripes (Ingham *et al.*, 1988). *ftz* and *eve* are both thought to repress *wingless (wg)*, another segment polarity gene, so that it is only expressed in single cell wide stripes where neither *eve* nor *ftz* is expressed. In this way overlapping domains of pair-rule gene expression are combinatorially interpreted to define the expression pattern of some of the segment polarity genes within individual segments.

Segment polarity genes are involved in the signal transduction and cell to cell communication which determines the future fate of individual cells. Mutations in these genes lead to deletions in parts of segments and the replacement of the deleted parts by a mirror image duplication of other parts of the segment.

The best characterised segment polarity genes are *en, wg, gooseberry (gsb)* and *ptc* which are first transcribed at the syncytial blastoderm in regions corresponding to different reiterated parts of parasegments. *en, wg* and *gsb* are first detected just before the beginning of gastrulation in fourteen stripes (Baker, 1987; Baumgartner *et al.*, 1987; Weir and Kornberg, 1985). In contrast, *ptc* is first transcribed all over the late stage 14 blastoderm and resolves into fourteen stripes at germ-band extension (Nakano *et al.*, 1989).

Each *en* stripe (except stripe 2) is one cell wide and corresponds to the anterior compartment of each parasegment, while each *ptc* stripe is three cells wide and is found in non *en*-expressing cells corresponding to the posterior compartment of each parasegment. The *wg* stripes are one cell wide and are in the adjacent cell, anterior to the *engrailed*-expressing cell. The *gsb* stripe is two cells wide and includes both the engrailed-expressing and the wingless-expressing cells (reviewed in Ingham, 1988).

en encodes a homeodomain-containing protein and has been shown to act as a transcription factor in *Drosophila* tissue culture cells (Desplan *et al.*, 1985; Jaynes and O'Farrell, 1988). *ptc* is predicted to be a transmembrane protein and may be involved in a signal transduction pathway (Nakano *et al.*, 1989) while *wg* is a secreted protein found outside the cell (van den Heuvel *et al.*, 1989).

The twelve known segment polarity genes are likely to be involved in a complex signal transduction and cell communication process which enables interpretation of the combinatorial information provided by the pair-rule genes and gives positional values to individual cells within each segment. Some of the cells could be specified by direct interpretation of pair-rule expression at the blastoderm stage (Ingham *et al.*, 1988) while others may require cell-cell communication at a later stage (Martinez-Arías *et al.*, 1988).

The homeotic genes are involved in the determination of the identity of different segments in the larva and fly. They mainly lie in two chromosomal locations: the antennapedia complex (AP-C) and the bithorax complex (BX-C). Each gene is expressed at a different position along the body and determines the identity of segments in that region. The position of the genes within the complexes matches in most cases their order of expression in the body (Akam, 1987).

In summary, the segmentation genes form a complex hierarchy of genes which respond to maternally encoded asymmetries in the egg and generate a detailed prepattern of individual cell identities in each segment of the embryo. The *bcd* protein is the primary morphogen whose gradient is interpreted by gap genes to define large overlapping regions in the embryo. These regions are refined by the pair-rule genes which specify overlapping domains of gene expression approximately a segment in width. Different cells in the embryo express different combinations of pair-rule genes which refine the pattern of segment polarity gene expression thus specifying cell identities in the embryo. The homeotic genes respond to information at several levels in the hierarchy and specify the identity of the segments formed by the segmentation genes.

The climax of the process of refining spatial information in the embryo coincides with the cellularisation of the blastoderm when membranes are formed around each peripheral nucleus. The presence of cell membranes radically alters the nature of intracellular movements of molecules controlling development. Before cellularisation the syncytial blastoderm is a single giant cell in which proteins can diffuse; after cellularisation development involves cell to cell communication as well as cell migration and molecular traffic is largely confined within each small blastoderm cell.

To understand the role of intracellular molecular trafficking in development, we must examine the cellular structure of the blastoderm.

The cellular structure of the blastoderm embryo:

The *Drosophila* blastoderm consists of 5000 peripheral nuclei in a syncytium. Cellularisation occurs over a period of 15-30 min during the 14th interphase (which I refer to as blastoderm stage 14). Annular constrictions (membrane furrows) appear in the outer membrane and deepen to separate the elongating nuclei laterally (Foe and Alberts, 1983). Cellularisation is completed just before gastrulation begins, when a membrane is formed basally, although cytoplasmic connections remain between the cells and the yolk while gastrulation is taking place (Rickoll, 1976, and see Fig. 1.2).

Different cytoskeletal proteins show different distribution within the syncytial blastoderm. Actin lies underneath the outer membrane and surrounds the membrane furrow; it is implicated in membrane furrow formation (Miller *et al.*, 1985; Warn, 1986; Warn and Magrath, 1983; Warn and Warn, 1986). Spectrin is co-localised with actin at this stage, although its function is unclear (Pesacreta *et al.*, 1989). A large number of actin binding proteins have been isolated which in general co-localise with actin, although their function has yet to be defined (Miller *et al.*, 1989). Two perpendicular centrioles are located above the nuclei during interphase and migrate to two opposite poles on either side of the nucleus during cell division (Karr and Alberts, 1986; Kellogg and Alberts, 1988). Microtubules polymerise from these centrioles so as to extend and engulf the nucleus. During nuclear elongation (stage 14), microtubules move down to reach the basal periplasm (Karr and Alberts, 1986). A large number of microtubule associated proteins (MAPs) have been isolated and antibody labelling indicates that they are localised in many distinct patterns in the nucleus and

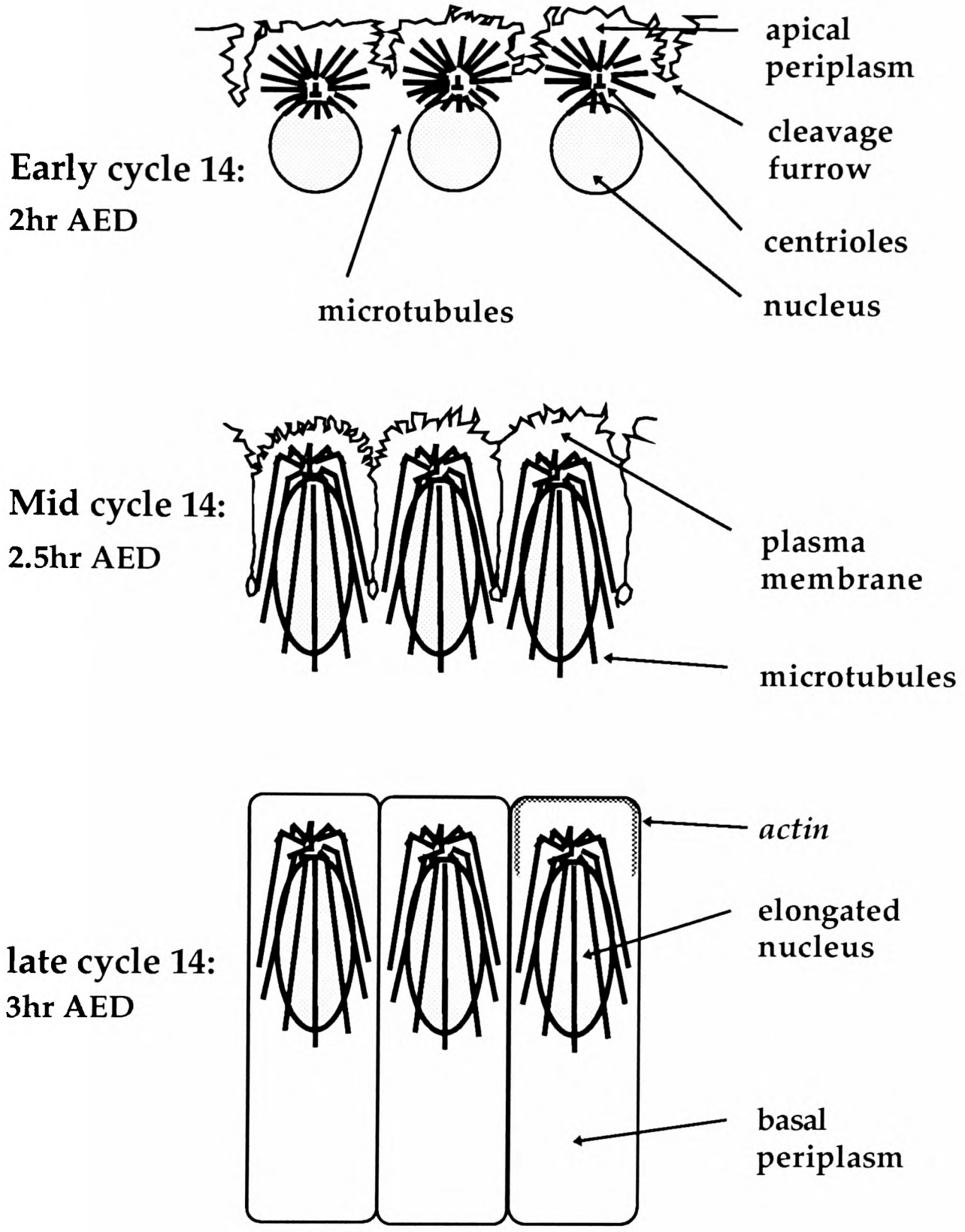


Fig 1.2
The cellular structure and cytoskeleton of the *Drosophila* blastoderm in the forteenth interphase (stage 14), 2.5 to 3hr after egg deposition (AED). Centrioels are situated on the apical side of the nuclei at the site from which microtubules emanate around the nucleus. *actin* is found at the apical membrane and is responsible for constrictions in the plasma membrane which move down to form cleavage furrows compartmentalising the nuclei. Throughout these stages the cytoplasm is divided into two distinct regions; the apical periplasm above each nucleus and the basal periplasm below it.

cytoplasm, although their function has yet to be defined (Kellogg *et al.*, 1989).

The cytoplasmic domain around each blastoderm nucleus is divided into two distinct regions by microtubules and actin as well as by the peripheral nuclei (Warn and Warn, 1986). The cytoplasm on the periphery of the embryo, exterior to the nuclei, is termed the apical periplasm, while the cytoplasm underneath the nuclei is termed the basal periplasm (Fullilove and Jacobson, 1971, and see Fig. 1.2). The blastoderm cells have an apical-basal polarity which may be important for development. Different gene products are required in different parts of the cell suggesting that the intracellular localisation of different macromolecules is of great interest.

Intracellular mRNA localisation in the *Drosophila* oocyte and blastoderm

There are several proteins which are required in different parts of the unicellular syncytial blastoderm embryo. The mRNAs encoding these proteins are present in restricted areas in the oocyte and embryo and involve two very different processes of intracellular mRNA localisation. The first is that of maternal mRNA prelocalised in the oocyte, a giant cell with a single nucleus. The second process involves the localisation of zygotically transcribed mRNA within the multinucleate syncytial blastoderm.

***bcd* mRNA** is localised in the anterior tip of the oocyte in four distinct stages (St Johnston *et al.*, 1989). In the first stage *bcd* mRNA is found localised in a ring in the anterior of the stage 6 oocyte. In the second stage, *bcd* mRNA is also localised in the apical cytoplasm of each nurse cell. The mRNA becomes localised to the anterior cortex of the oocyte as the nurse cells contract during the third stage of *bcd* localisation. This is thought to involve *bcd* mRNA transport from the nurse cells to the oocyte. During the

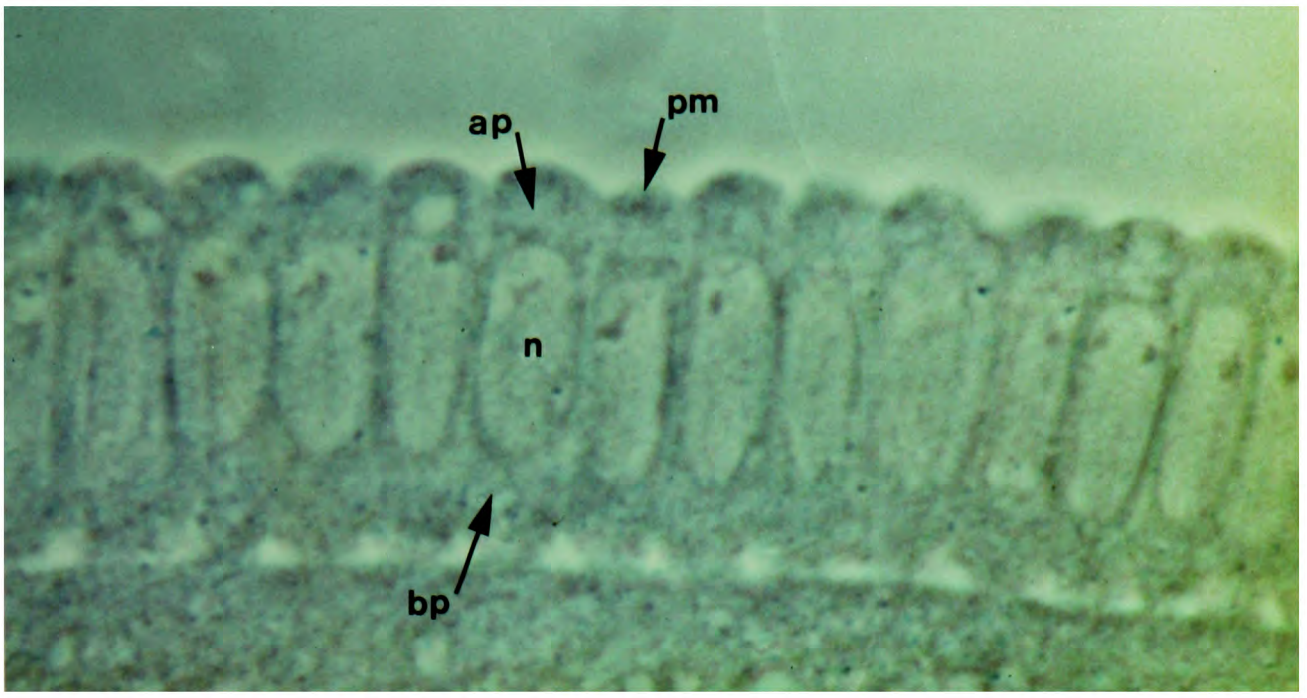
fourth phase of localisation the mRNA becomes localised to a broad region at the dorsal side of the anterior pole. *bcd* localisation is disrupted by mutations in 3 maternal genes *exuperantia (exu), swallow (sww)* and *staufen (stau)* (Berleth *et al.*, 1988; Frohnhöfer and Nüsslein-Volhard, 1987) at stages one, three and four respectively (St Johnston *et al.*, 1989).

Toll (Tl) (a gene involved in the determination of the D-V axis), actin *5C* and *α4-tubulin* are maternally expressed genes whose transcripts are distributed in all parts of the oocyte. While *Tl* mRNA is found in the basal periplasm of all parts of the blastoderm embryo (Gerttula *et al.*, 1988), actin *5C* (Burn *et al.*, 1989) and *α4-tubulin* (Mathews *et al.*, 1989) mRNAs are present in both the apical and basal periplasms. Asymmetric intracellular transcript localisation is not confined to maternally expressed genes; zygotically transcribed mRNAs also display a variety of different patterns of intracellular distribution.

Zygotic transcripts of different classes of genes are found to have distinct patterns of intracellular localisation (see Fig. 1.3). The mRNAs of the gap genes *Kr, hb* and *kni* are found in both the apical and basal periplasms. In contrast, the transcripts of *ftz, eve, h, run* and *prd* are asymmetrically distributed in the apical periplasm above the blastoderm nuclei (see Fig. 1.4 and: Ingham and Gergen, 1988; Ingham *et al.*, 1985a; Kilcherr *et al.*, 1986; Macdonald *et al.*, 1986; Weir and Kornberg, 1985). This intracellular distribution is present throughout the blastoderm stage, from the time the pair-rule genes are first expressed (Ingham and Gergen, 1988; Ingham *et al.*, 1985a; Weir and Kornberg, 1985).

The segment polarity genes are less consistent in their intracellular localisation. When *gsb* (Baumgartner *et al.*, 1987), *ptc* (Nakano *et al.*, 1989) and *en* (Weir and Kornberg, 1985) are first detected, their mRNAs are found in both the apical and basal periplasms. In contrast, when *wg* is first

a)



b)

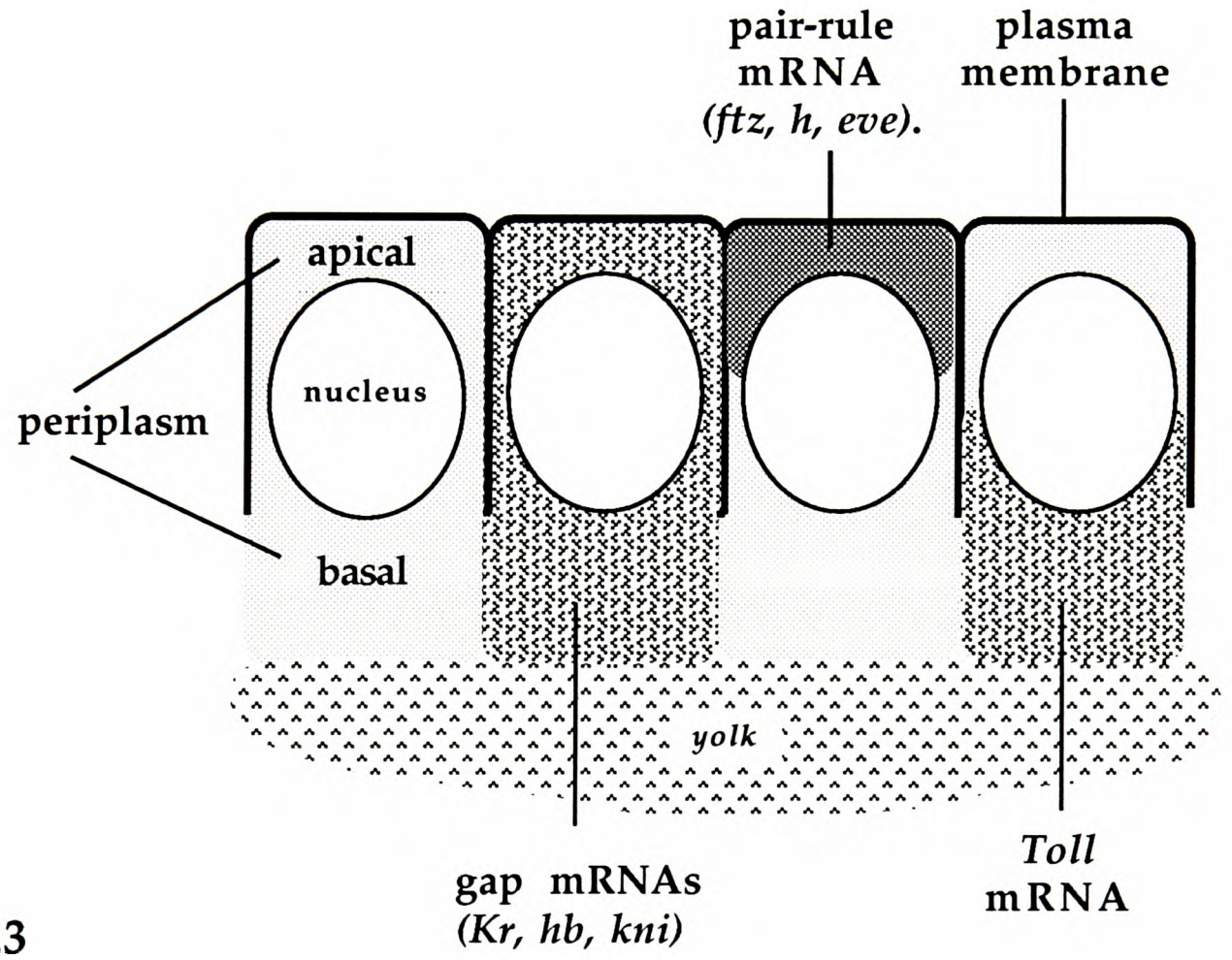


Fig. 1.3

a) The morphology of the late syncytial blastoderm embryo. Arrows show: (n) nucleus, (ap) apical periplasm, (bp) basal periplasm and (pm) plasma membrane.

b) Intracellular message localisation in the syncytial blastoderm. mRNA shows three patterns of distribution: apically localised, basally localised and unlocalised. Periplasm is shown in light shading, RNA in darker shading and membranes in bold lines.

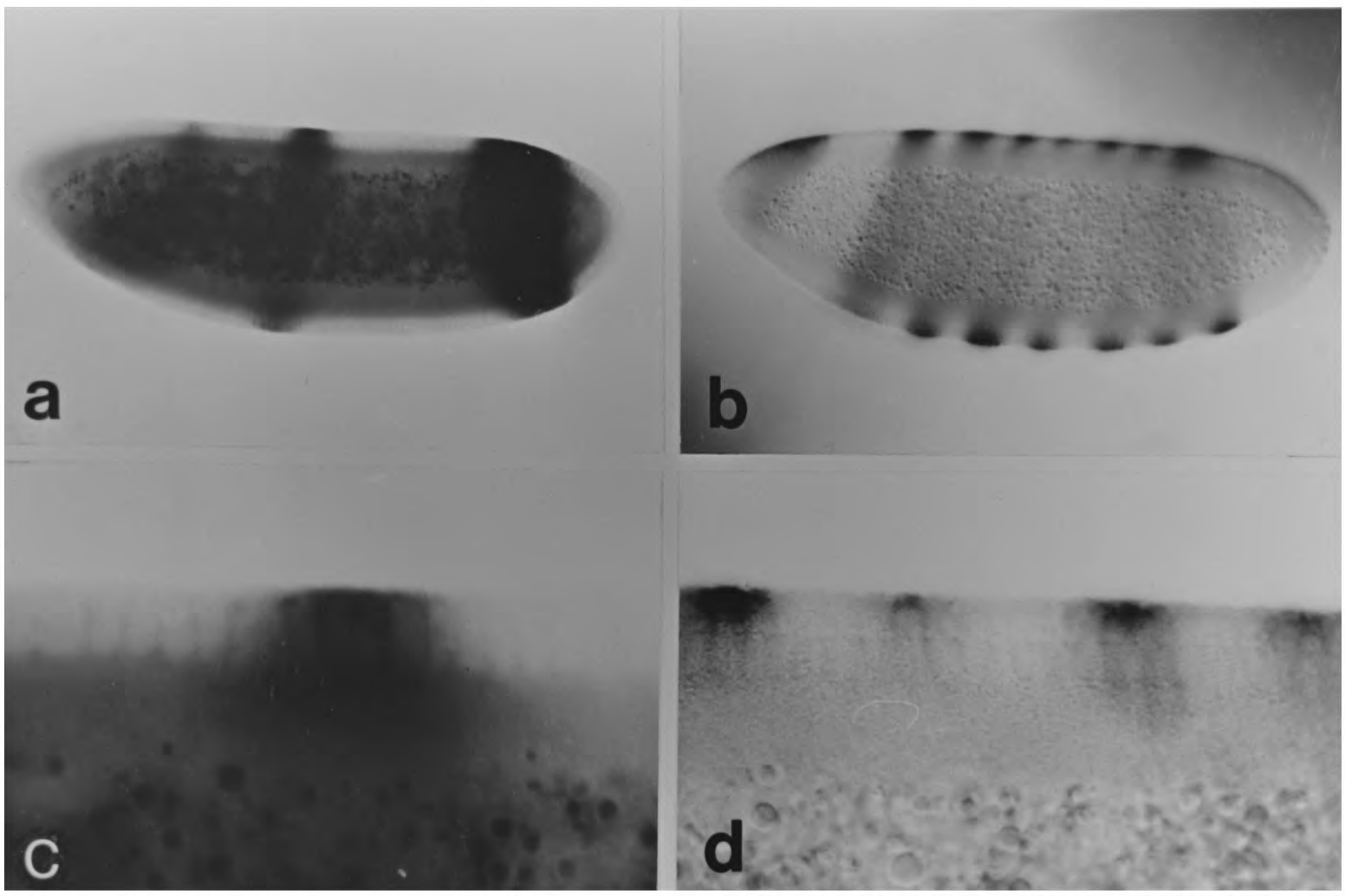


Fig. 1.4
In situ **hybridisation with *h* and *hb* DIG-labelled probes on wild-type 2-4h embryos, illustrating that *h* mRNA is apically localised and *hb* message is unlocalised. a) and c) are probed with *hb* probe. b) and d) are probed with *h* probe. Embryos are visualised under Nomarski optics. c) and d) are high power views of parts of the same embryos seen in a) and b) respectively.**

detected, it is localised in the apical periplasm (Baker, 1987, and A. Hidalgo personal communication).

When the homeotic genes *Ubx* (Akam, 1985) and *zen* (Rushlow *et al.*, 1987) are first expressed in the blastoderm embryo, their transcripts are found in the apical and basal periplasms .

It follows that intracellular mRNA localisation is different in different classes of segmentation genes. There is no published work on *Drosophila* which addresses the mechanism or function of apical localisation in any detail. Before suggesting some testable models for the mechanism of localisation, it is worth considering what happens to mRNA after being transcribed.

RNA processing

From the moment it is transcribed, and while it is being processed, RNA is bound to proteins (see chapter 7). The processing of RNA involves cleavage and polyadenylation in the 3' end (see chapter 6) as well as splicing out of introns by spliceosomes formed on nascent RNA. The addition of a poly(A) tail could be related to the splicing reactions taking place in the spliceosomes and may be equally complex. It is currently unclear whether or not the export of mRNA is intimately connected with RNA processing.

Nucleocytoplasmic communication in other systems

The export of mRNA from the nucleus to the cytoplasm is part of the bidirectional transport of molecules across the nuclear membrane: mRNA, tRNA, rRNA and snRNA are exported while snRNP and specific nuclear proteins are imported. Nuclear pore complexes (NPCs) are the key components which determine transport between the nucleus and cytoplasm. While NCPs allow free diffusion of small molecules, they must distinguish between the different types of macromolecules which are

imported and exported from the nucleus (Ringertz, 1986). It is worth digressing to consider the nuclear import of proteins, as more is known about this process than about mRNA trafficking.

Nuclear proteins have a short amino acid sequence, the nuclear tag, which is required to direct them to the nucleus. Uptake of such proteins involves two steps. The first is the recognition of the nuclear tag and binding to the outside of the nuclear membrane, an energy-independent process. The second step is the uptake of the protein through nuclear pores, an energy-dependent process which is blocked at low temperatures (Richardson *et al.*, 1988).

NPCs are the key components which regulate trafficking of molecules in and out of the nucleus. Wheat germ agglutinin inhibits uptake of nuclear proteins but does not affect the recognition step or binding at the nuclear membrane and does not interfere with mRNA export (Finlay *et al.*, 1987). An N-acetylglucosamine-bearing NPC protein has been shown to confer the ability to take up nuclear proteins in a reconstituted nuclear membrane *in vitro* and is believed to be involved in the recognition of the nuclear tag. However, lack of the protein does not affect the structure of the nuclear pore complex as seen under the electron microscope (Finlay and Forbes, 1990). This protein could thus be the one to which wheat germ agglutinin binds when blocking the import of nuclear proteins. It is not known whether or not different types of pores exist to import and export different classes of nuclear proteins and mRNA species.

Nucleocytoplasmic transport of mRNA involves three stages: transport in the nucleoplasm, transport across the nuclear membrane and transport in the cytoplasm (reviewed in Agutter, 1985; Clawson *et al.*, 1985). At all stages RNA is thought to be bound to proteins and is perhaps tethered to the cytoskeleton so that RNA is not freely diffusible. Injection of tRNA, 5S

RNA or poly(A) coated gold particles into the nuclei of *Xenopus* oocytes has shown that, in general, RNA is exported through the same pores as those through which nuclear proteins enter the nucleus. The gold particles used were much larger than the effective pore size available for free diffusion thus implicating a transport mechanism rather than free diffusion (Dworetzky and Feldherr, 1988; Newmeyer and Forbes, 1988). Experiments *in vitro* have shown that the transport of RNA across the nuclear membrane is ATPase dependent (Clawson *et al.*, 1978).

Different species of RNA molecules show a different pattern of nuclear-cytoplasmic transport. tRNA is transported into the cytoplasm after transcription and processing in the nucleus (Zasloff, 1983). snRNAs are transcribed in the nucleus and exported to the cytoplasm where they bind to snRNP proteins and return to the nucleus (De Robertis, 1983; Guddat *et al.*, 1990; Konings and Mattaj, 1987; Mattaj and De Robertis, 1985).

mRNA has been found to be associated with cytoskeletal elements in the cytoplasm (Fey *et al.*, 1986; Lawrence and Singer, 1986) and with the nuclear matrix in the nucleoplasm (van Eekelen and van Venrooij, 1981). Although the techniques used in such biochemical studies are often subject to artefacts, some of these cellular and nuclear components, in particular the cytoskeleton and nuclear scaffold, may be involved with mRNA export from the nucleus and transport in the cytoplasm and nucleoplasm. Electron micrographs of mRNA molecules leaving the nucleus of a *Chironomus* larval salivary gland indicate that ribonuclear proteins (RNP) bind to mRNA in the nucleus and detach as the RNA is transported across nuclear pore complexes. As the RNA emerges into the cytoplasm it appears to bind to cytoplasmic RNA binding proteins and ribosomes (Stevens and Swift, 1966).

A thoroughly researched example of the nuclear export of mRNA is provided by the *rev* gene of the Human Immunodeficiency Virus (HIV), a member of the lenti-virus family of viruses and the causative agent for acquired immune deficiency syndrome (AIDS) (reviewed in Pavlakis and Felber, 1990). *Rev* protein is required for viral replication and facilitates the export of transcripts containing a Rev responsive element (RRE). (Olsen *et al.*, 1990). Mutations which alter the secondary structure of RRE, disrupt the interaction with *rev* protein and secondary mutations which restore the secondary structure also restore *rev* interaction with the RRE (Olsen *et al.*, 1990). *Rev* protein binds to a 223 bp RRE *in vitro* and protects a 33bp fragment from ribonuclease digestion (Heaphy *et al.*, 1990).

Intracellular transcript localisation in vertebrates

We should also consider what is known about intracellular transcript localisation in other systems. There are several known cases of mRNAs which are localised to a particular subcompartment of the cytoplasm of large cells. It is suggested that the proteins encoded by these mRNAs are restricted to the same subcompartment and function in the same region of the cytoplasm.

Skeletal muscle cells (myoblasts) are a giant syncytium containing hundreds of nuclei distributed evenly in the cytoplasm and they therefore resemble, in this respect, the syncytial blastoderm. The mRNA transcribed from each nucleus remains in the cytoplasm that is close to that nucleus (Hall and Ralston, 1989). However, the distribution of the proteins translated from the mRNAs, depends on the type of protein. In general, cytoplasmic proteins diffuse far beyond the region surrounding the nucleus but nuclear proteins are found predominantly in the nucleus which transcribed their mRNAs, and in lower levels in a few adjacent nuclei.

In myoblasts, mRNAs and the proteins they encode are in many cases restricted to the region of cytoplasm where they are required to function (Ralston and Hall, 1989). For example, the myoblast acetylcholine receptor and the mRNA encoding it are concentrated at its site of function near the postsynaptic membrane of the neuromuscular junction (Merlie and Sanes, 1985).

Nerve cells consist of a small cell body with two types of long cytoplasmic extensions: numerous dendrites and a single, longer, often branching axon. Ribosomes and mRNA are found only in the cell body and dendrites; they are conspicuously absent from the axon, so proteins needed in the axon are translated in the cell body and transported to their site of action. Some classes of mRNA are transported as polyribosomes down the dendrites to the synapses, where ribosomes are clustered so that the mRNAs can be translated to give synaptic proteins (Davis *et al.*, 1987; Garner *et al.*, 1988).

Both axons and dendrites contain microtubules but their polarity is different in the two types of cytoplasmic extensions. Axonal microtubules are uniform in their polarity; (+) ends point towards the growth cone. Dendritic microtubules are non-uniform in their polarity; equal numbers of (+) microtubule ends point to the cell body and the growth cone (Baas, 1988). It is speculated that the difference in polarity of microtubules accounts for the contrasting distributions of ribosomes in the axon and dendrites; organelles are actively transported along microtubules towards their (-) pole (Baas, 1988).

Asymmetric intracellular transcript localisation is also found in fibroblasts which are very small cells in comparison to neurons and myoblasts. actin mRNA is localised near the lamellipodia of fibroblasts, where actin is translated and polymerised to facilitate changes in cell shape (Lawrence and Singer, 1986).

The examples described above include zygotic transcripts in a variety of cell types. Some maternal mRNAs are distributed asymmetrically in the oocyte of several species, the best understood case being the *Xenopus* oocyte.

The *Xenopus* oocyte is not only morphologically inhomogeneous, but also contains an asymmetric distribution of macromolecules, which defines the animal-vegetal (A-V) axis, one of two orthogonal primary axes of embryonic development. The animal hemisphere gives rise to the ectoderm and the vegetal hemisphere gives rise to endoderm. The mesoderm is formed by an inductive influence of vegetal tissue on equatorial animal tissue mediated by at least two mesoderm inducing factors (MIFs). It has been speculated that at least two MIFs are found; an FGF-like protein (Slack *et al.*, 1989) and a TGFβ-like protein homologous to mammalian Activin-A (Smith *et al.*, 1990). Transcript localisation in the *Xenopus* oocyte is worth considering in detail, in order to determine whether a mechanistic similarity exists with mRNA localisation in the *Drosophila* blastoderm.

mRNA localisation changes with different stages of *Xenopus* development before and after fertilisation (Capco and Jeffrey, 1982). Differential animal/vegetal screens have revealed four localised transcripts. Three maternal transcripts are prelocalised to the animal hemisphere (Rebagliati *et al.*, 1985; Weeks *et al.*, 1985) one of which is the a subunit of mitochondrial ATPase (Weeks and Melton, 1987). The fourth transcript, *Vg1*, is prelocalised to the vegetal hemisphere. *Vg1* mRNA becomes localised to a narrow crescent in the vegetal hemisphere from an initial uniform distribution in the oocyte (Melton, 1987).

Vg1 protein is a member of the TGFβ–like family of molecules and may be implicated in the process of mesoderm induction. Like heterologous TGF-β1, *Vg1* enhances mesoderm induction by FGF but it has not yet been shown to functional a primary MIF (Yisraeli *et al.*, 1990). *Vg1* protein is confined to

the vegetal hemisphere throughout development and is associated with membrane fractions (Dale *et al.*, 1989). *Vg1* mRNA is thus localised in the cytoplasmic site at which the protein may function in mesoderm induction.

Vg1 mRNA injected into *in vitro* cultured oocytes, becomes localised in an identical way to the *in vivo* localisation of the transcript (Yisraeli and Melton, 1988). Two distinct phases of *Vg1* localisation can be distinguished using cytoskeletal inhibitors *in vitro*. The first step involves movement of *Vg1* mRNA from the animal to the vegetal hemisphere in early oocytes and is actin dependent; it is inhibited by cytochalasinB which depolymerises actin. This movement is unaffected by nocodazole and colchicine which cause microtubule depolymerisation. The second stage involves mRNA binding to vegetal cortex in mid to late oocytes and is microtubule-dependent; it is inhibited by nocodazole and colchicine and is unaffected by cytochalasinB. It has therefore been concluded that actin transport to the vegetal hemisphere followed by binding to microtubules leads to the formation of the observed tight crescent of *Vg1* mRNA (Yisraeli *et al.*, 1989).

In the asymmetric mRNA localisation described above, the mRNA is located close to the site of accumulation and function of the protein it encodes. My working hypothesis is that a similar situation may occur in the *Drosophila* blastoderm. The localisation of the mRNA to the apical periplasm leads to the translation of pair-rule proteins in the apical periplasm so that the proteins predominantly enter the nucleus which transcribed them. This would explain why the diffusion of pair-rule protein is limited in comparison with that of gap-gene proteins. My thesis does not address the question of the function of apical localisation directly, but it does consider whether the apical distribution of mRNA influences the apical localisation of protein (see chapter 8).

We are now in a position to formulate some models for the mechanism of localisation of pair-rule transcripts in the *Drosophila* blastoderm (see Fig. 1.5; see chapters 5, 6 and 7).

Models for specific pair-rule mRNA localisation in *Drosophila*

Selective mRNA degradation

One plausible model for the localisation of pair-rule transcripts is that mRNA is unlocalised in the cytoplasm but is selectively stabilised in the apical periplasm and is quickly degraded in the basal periplasm (see Fig. 1.5).

Vectorial transport in the cytoplasm

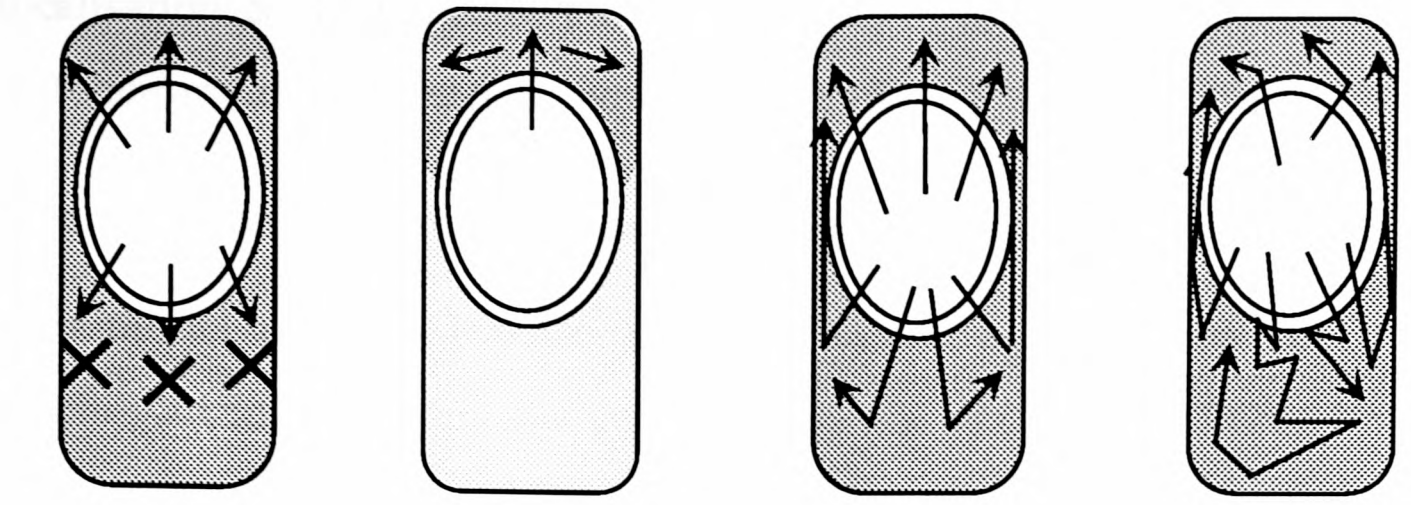
A different cytoplasmic mechanism for localisation is the transport of mRNA from the basal to the apical periplasm. Cytoplasmic mRNA transport may be involved in the localisation of maternal transcripts; *Vg1* mRNA localisation in the *Xenopus* oocyte for example. The model proposes active cytoplasmic transport along cytoskeletal fibres from the basal to the apical periplasm (see Fig. 1.5).

Free diffusion in the cytoplasm and mRNA binding in the apical periplasm

Another cytoplasmic model for apical localisation involves free mRNA diffusion and its attachment to a specific localisation factor in the apical periplasm (see Fig. 1.5).

Vectorial mRNA export out of the nucleus

An alternative class of models involves a nuclear mechanism. An attractive nuclear mechanism is the export of mRNA from the apical side of the nucleus directly into the apical periplasm. The mechanism of export of mRNA out of the nucleus and the exact site of transcription of particular genes in the nucleus are both poorly understood. Consequently, apical

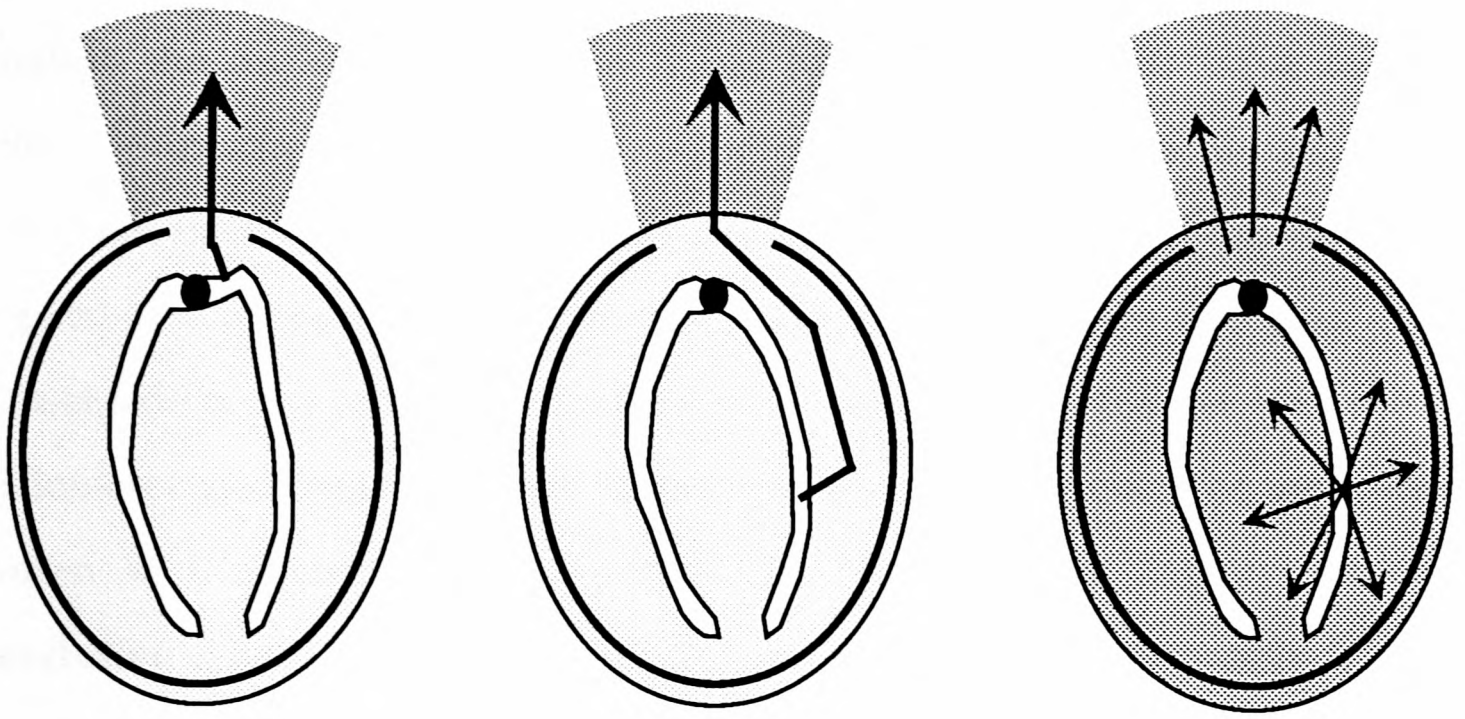


1) Selective Degradation.

2) Directed Export from the Nucleus.

3) Transport in the Cytoplasm.

4) Random diffusion and apical anchoring.



2a) Gene gating

2b) Transcript gating

2c) Diffusion and selective export.

Fig 1.5

A schematic description of possible models for the apical localisation of pair-rule genes. 1) The transcript is exported from all parts of the nucleus and is specifically degraded in the basal periplasm. 2) The transcript is only exported through nuclear pores located on the apical side of the nuclear membrane: 2a) DNA sequences interact with nuclear pores on the apical side so that RNA is immediately exported after transcription. 2b) The transcript is transported along tracks in the nucleoplasm to the apical side and exported through particular pores. 2c) The transcript diffuses throughout the nucleoplasm but can only be exported through pores in the apical side. 3) The transcript is exported from all parts of the nucleus and is transported from the basal to the apical periplasm. 4) The transcript is exported from all parts of the nucleus and diffuses randomly until it sticks to anchoring elements in the apical periplasm.

localisation could be achieved in several different ways corresponding to different models for nuclear export of mRNA.

Free diffusion in the nucleoplasm and selective apical export

This model involves random diffusion in the nucleus until the mRNA reaches nuclear pores on the apical side of the nuclear membrane (see Fig. 1.5).

Gene-gating

Experiments which analyse the three-dimensional position of chromatin within the nucleus of different tissue types in *Drosophila* have revealed that chromatin is only partly topologically ordered in the nucleus. The chromocentre, which consists of a cluster of centromeres, is located on the apical side of the nuclear membrane and the telomeres are clustered on the basal side of the nuclear membrane. The chromatin which lies between the telomeres and centromeres is found in many different parts of the nucleus often at the periphery and individual chromosomes occupy mutually exclusive domains. The three-dimensional structure of the nucleus varies between individual cells of the same cell type as well as between cells of different tissues (Hochstrasser and Sedat, 1987a and see chapter 7).

On purely theoretical grounds, the 'gene-gating hypothesis' suggests that particular DNA sequences, near or in genes, may interact specifically with a subset of nuclear pores so that the transcripts of these genes are exported through those pores (Blobel, 1985; Mirkovitch *et al.*, 1984). Gene-gating is an attractive idea in the light of evidence suggesting that chromosomes adopt a non-random conformation in the nucleus. Particular genes may be preferentially associated with particular parts of the nuclear membrane (Hochstrasser and Sedat, 1987a). Gene-gating could then explain the apical localisation of pair-rule mRNA if all the pair-rule genes were situated at or

near particular DNA sequences which direct them to a subset of pores on the apical side of the nuclear membrane (see chapter 7).

Transcript-gating

If the pair-rule genes lie within regions of chromatin with a variable topological location then they may be transcribed at a highly variable position in the nucleus. The 'transcript-gating hypothesis' (which I propose in chapter 7) states that RNA sequences mediate the transport of mRNA from its site of transcription along tracks in the nucleoplasm to pores on the apical side of the nuclear membrane (see Fig. 1.5). Apparently similar tracks have been observed in the nuclei of chicken fibroblast cells infected with viral mRNA which fails to exit from the nucleus. Viral transcripts accumulate in curvilinear tracks which may well represent the path taken by mRNA from the site of transcription to a particular region of the nuclear membrane (Lawrence *et al.*, 1989, and see chapter 7).

In the thesis, I deal with three aspects of the intracellular distribution of pair-rule transcripts by making particular constructs of hybrid genes and analysing the intracellular localisation of their transcripts and the proteins they encode. The first is the determination of the localisation signal in pair-rule genes, and the second is the elucidation of the cellular mechanism for localisation using these sequences. The third, which I address indirectly, is the function of localisation.

Before undertaking experiments to analyse the mechanisms and function of apical localisation of pair-rule transcripts, we have to determine which sequences are responsible for the localisation. The more narrowly we can define the regions required for localisation, the more likely we are to arrive at a mechanism and function for localisation. In the first experimental chapter, I describe experiments which show that the sequences necessary

and sufficient for the localisation of pair-rule mRNA lie in the 3' untranslated part of the genes. In the subsequent chapters I narrow down this sequence and describe experiments which test the involvement of the different mechanisms for localisation discussed above.

CHAPTER 2: MATERIALS AND METHODS

BUFFERS AND REAGENTS

List of suppliers of chemicals:

Acetic Acid	Fisons
Agarose	Pharmacia
Alkaline Phosphatase Staining Kit II	Vectastain
Avidin/Biotin peroxidase kit	Vectastain
Biotinylated secondary antibody	Vectastain
Caesium chloride	Fisons
EDTA	Fisons
Ethanol	James Burrough
Foetal calf serum	Gibco
Glycine	Fisons
JB4 methacrylate mount	Polysciences

All other chemicals or reagents were supplied by BDH.

Stocks and media

L-broth, bacterial growth medium: 10g tryptone, 5g yeast extract, 5g NaCl in one litre of distilled water. 100µg/ml Ampicillin was used.

Fly food: Standard soya flour/molasses medium was used (Nüsslein-Volhard *et al.*, 1984).

Apple-juice plates were used for egg collections (Wieschaus and Nüsslein-Volhard, 1986)

Bacterial strains: all plasmids were grown in TG2 strains of *E.coli* on L-broth or L-plates with 100µg/ml of ampicillin or kanamycin.

Solutions and Buffers

TE: 10mM Tris pH8, 1mM EDTA pH8

10XSSC: 3M NaCl, 0.3M NaCitrate pH7

Injection Buffer: 5mM KCl, 0.1mM NaPhosphate pH6.8.

PBS: 130mM NaCl, 7mM Na_2HPO_4, 3mM NaH_2PO_4.

PTW: 0.1% Tween20 in 1XPBS

Fix: 3.7% formaldehyde, 0.1M PIPES pH7, 2mM $MgS0_4$, 1mM EGTA pH8

NBT buffer: 0.1M NaCl, 50mM $MgCl_2$, 0.1M Tris pH 9.5 and 1% Tween20.

Hybridisation solution: 50% formamide, 5XSSC, 100mg/ml sonicated boiled salmon sperm DNA, 100µg/ml tRNA, 50µg/ml heparin and 0.1%Tween 20.

Cloning Methods

CsCl purification of plasmid DNA, minipreparation of plasmid DNA, restriction digests, agarose gel preparation, gel purification of DNA fragments, ligations and transfromation of competent cells with plasmids were performed using standard methods (Sambrook *et al.*, 1989).

Constructing *LT*-based plasmids

The *pLT (LT)* vector was constructed in three cloning steps from *pUChsneo*, a transformation vector containing P-element ends for transposition into the fly genome and a neomycin resistance gene for selection of transformed flies (Steller and Pirrotta, 1985). In the first step, the BamHI to EcoRI fragment of the 3' end of the human *β-globin* gene was inserted into the BamHI and EcoRI sites of the polylinker of *pUChsneo*. Two partly complementary synthetic oligonucleotides (oligos) were then introduced into the SalI and BamI sites in the polylinker of *pUChsneo*. The overhanging ends of the oligos had SalI and BamHI sticky ends but have an

incomplete recognition site for the enzymes; on ligation to the SalI BamHI sites in the polylinker the two restriction enzyme recognition sites were destroyed (see Fig. 2.1). The oligos contain one KpnI site, one SalI site and one NotI site in that order, starting from the 5' end. The oligos also contain a protein stop codon directly after the SalI site (see Fig.2.1). The KpnI to SalI fragment of *FG2* containing the *ftz* upstream region fused to the *lacZ* gene was inserted into the KpnI and SalI sites of the oligo so that the stop codon in the oligo is in frame with the *lacZ* open reading frame (see Fig. 2.2).

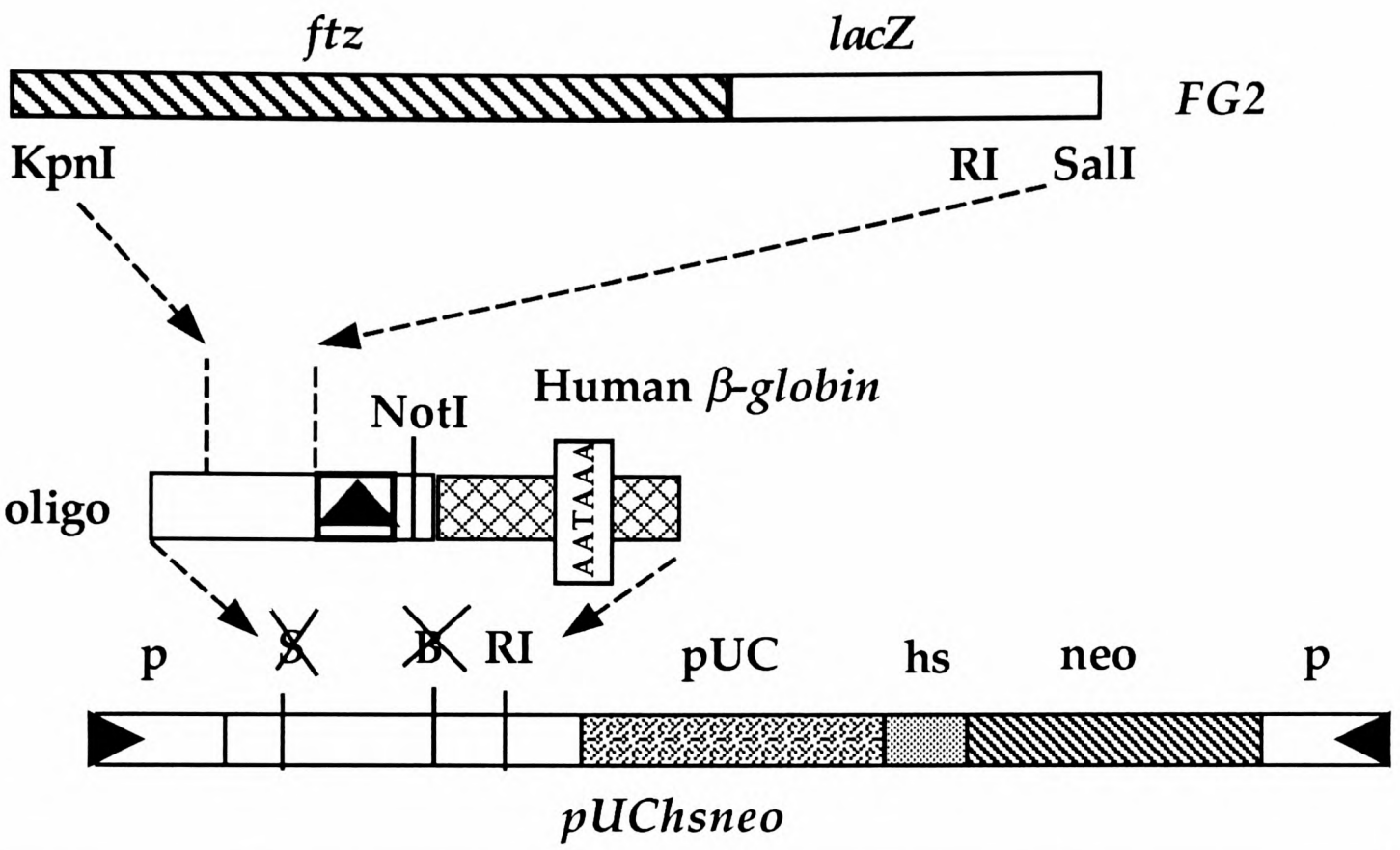
The fragments of *ftz* inserted into *LT* were determined by convenient restriction enzyme sites and were first subcloned into *HSS7-Sal*, a kanamycin-resistant vector, in order to create NotI sticky ends (see Fig. 4.3). The fragments of the *eve* 3' untranslated inserted into *LT* were produced using appropriate oligonucleotide primers using the polymerase chain reaction (PCR) since no useful unique restriction enzyme sites were present in the 3' end.

PCR fragments into *LT*

The 5' PCR primers started with a non-complimentary sequence containing 4bp followed by a SalI site. After the Sal I site a stop codon was introduced in frame in order to terminate translation of the *lacZ* gene in *LT*. There followed a stretch of 15 bp complimentary to either *eve* or *ftz* sequences. The 3' primers were similar to the 5' primer except they contained a NotI site instead of the SalI site and did not contain the stop codon. The PCR product was cleaned with a G-50 column, digested with NotI restriction enzyme and gel purified before being ligated to *LT* digested with SalI and NotI and similarly treated. The position of the primers used for cloning pieces of *ftz* and *eve* by PCR are shown in Fig. 4.7 and 4.9.

a) The sequence of the oligonucleotide inserted into the *pUChsneo* polylinker:

b) Constructing *LT:*



pUChsneo

p	P-element ends	~~S~~	site destroyed in cloning
hs	*hsp-70* promotor	**S**	SalI
neo	neomycine resistance	**B**	BamHI
RI	EcoRI		translation Stop
pUC	pUC 8 sequences		

Fig. 2.1: Constructing *LT*, the localisation tester.

a) The partly complementary oligonucleotides inserted into the SalI and BamHI sites of the *pUChsneo* polylinker destroying the sites on insertion.

b) Constructing the *LT* vector by inserting the BamHI-HidIII fragment of the Human *β-globin* 3' end followed by the oligo described in a) and the SalI-HindIII fragment of *FG2* containing a *ftz/lacZ* fusion.

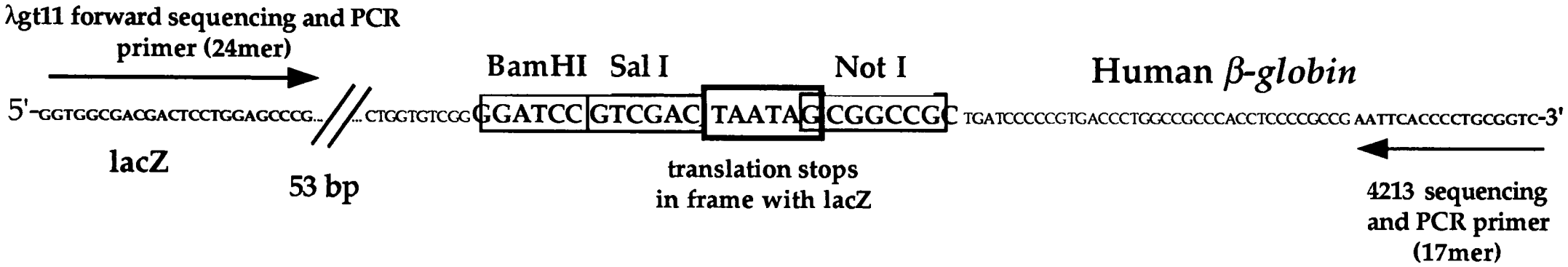


Fig. 2.2: The *LT* Vector:

The sequence, important restriction enzyme sites and oligos for sequencing and PCR.

Polymerase Chain Reaction:

The cetus GeneAmp DNA Amplification Reagent Kit was used essentially as described by the manufacturers with taq polymerase, a thermo-stable DNA polymersae isolated from *Thermus aquaticus*.

For mutagenesis by PCR the reaction was performed in two stages (see Fig. 2.3). The first stage involved two reactions from internal mutagenic primers to two external primers (*e3-eΔ3'* and *e2-eΔ5'*). In the second reaction the two products, which overlap by the length of the mutagenic primers, were mixed and used as a template for a PCR reaction performed with the two external primers *e3-e2* (see Fig. 2.3). The final PCR product contained a 2bp mutation, due to a missmatch in *eΔ3'* and *eΔ5'* (see Figs. 2.3 and 2.4).

For mapping polyadenylation sites cDNA (kindly donated by Gary Paterno) of wild type poly(A)$^+$ RNA (made by Susan Parkhurst) and a PCR reaction performed on the first strand with an internal primer and a poly(T) primer. Because of the high A-T content of the primer 2-3mM Mg^{2+} was used in the reaction buffer instead of the usual 1.5mM Mg^{2+}.

DNA Sequencing:

I used the Sanger dideoxy method (Sanger *et al.*, 1977), with the help of the "United States Biochemical Corporation Sequenase Kit Version 2.0". *LT*-based or Bluescript double stranded plasmid templates were prepared as a miniprep and denatured in 0.2M NaOH, 0.2mM EDTA pH8 for 5min at room temperature (RT). The sequencing reactions were performed essentially as described by the manufacturers.

Several constructs were sequenced in order to confirm their structure. These included the *LT* vector, *LTeΔ*, *LTe32* & *FG2*. Bluescript-based

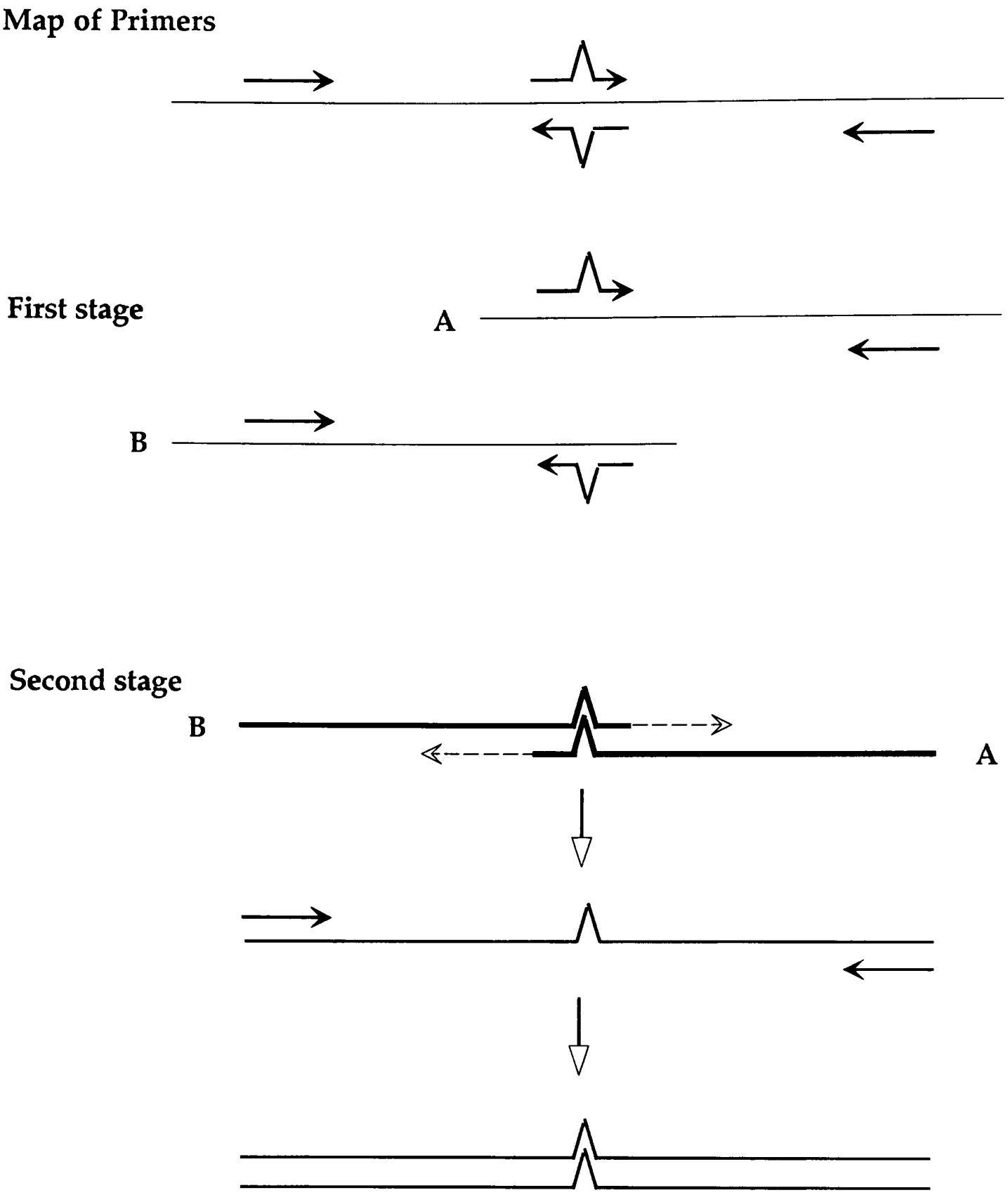
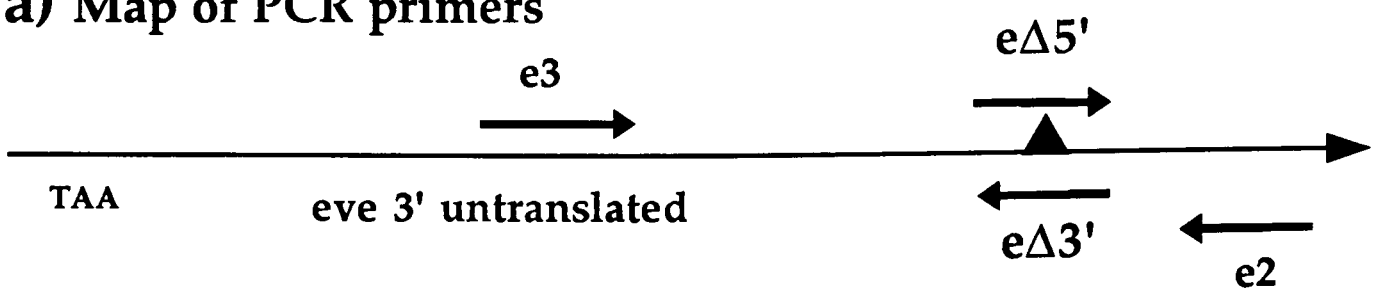


Fig 2.3: Strategy for mutagenesis using PCR .

The scheme used to generate a two base pair mutation in the polyadenylation signal: from AATAAA to AATGCA. The mutation is introduced by means of a missmatch in the PCR primers used in the first stage. Large arrows signify oligonucleotide primers. Dotted arrows signify filling in by Taq polymerase. Kinks in lines indicate missmatches in the primers which become incorporated into the PCR product (Higuchi et al ,1988).



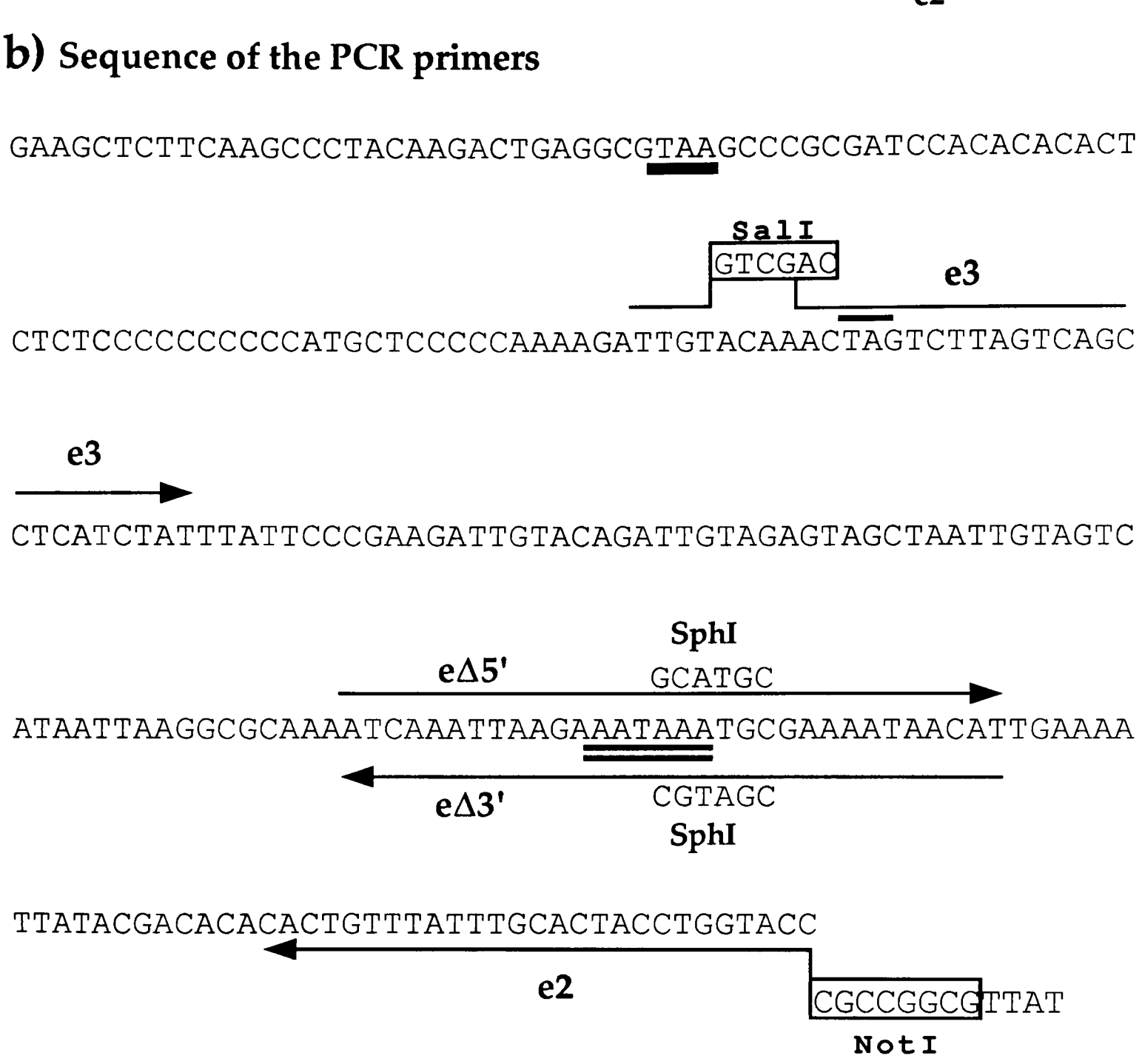


Fig. 2.4: Primers used for PCR mutagenesis.

The PCR primers used to mutate the polyadenylation signal of *eve:* creating an SphI site (AATAAA to AATGCA). The PCR reaction was performed in two steps: 1) PCR reactions using primers e3-eD3' and e2-eD5' 2) 1/50 of each of the products of the first reaction were used as template with fresh primers e2-e3. The principle behind the scheme is illustrated in Fig. 2.1. Arrows pointing to the right contain sense sequence, while those pointing to the left contain anti-sense sequence. The black triangle indicates the position of the AATAAA.

plasmids were sequenced to determine the polyadenylation site of *eve* and *ftz*.

Transgenic flies

All the plasmids injected into flies were based on the *pUChsneo* vector (Steller and Pirrotta, 1985). CsCl-purified DNA (0.5mg/ml) was injected into 0-1h embryos in the presence of 0.1mg/ml *pπ25.7wc* helper plasmid DNA in injection buffer.

I injected plasmids into a *bw;st* stock using the following selection scheme:

	bw/bw ; st/st		(white eyes) DNA injected into 0-1h embryos
	⇓		and allowed to develop on normal food
G0	*bw/bw ; st/st*	x	+/+ ; +/+
		⇓	selected on G418 food
mapping cross: G1		*bw*/+ ; st*/+* ♂	x *bw/bw ; st/st* ♀
G418 food		⇓	
		G2	

If G2 were *bw* and *bw;st*, then the insert is on the second chromosomal; if *st* and *bw;st*, then on the third chromosome. If there were only female progeny (when putative transformant was male), then the insert was on the X chromosome. If there were *bw*, *st* and *bw;st* progeny, then a double insert or X chromosome insert occurred. The transformants were mated to an appropriate balancer stock on G418, and homozygous stocks were established if they were viable. In several cases the identity of the plasmid inserted was confirmed by PCR using appropriate primers (see Fig. 2.2) on DNA extracted from the balanced lines.

Collecting embryos

Fly stocks were grown according to the standard methods described in (Wieschaus and Nüsslein-Volhard, 1986). Embryos were collected on yeasted apple juice plates at 25°C, 60% humidity. They were dechorionated in 14% sodium hypochlorite for 2min at RT and thoroughly washed with water before fixation.

Whole-mount *in situ* hybridization with non-radioactive probes

The procedures used were based on a recently described protocol which uses DIG-labelled probes (Tautz and Pfeifle, 1989). Unless otherwise stated, all rinses were for 2min and washes for 20min.

Fixation: Dechorionated embryos were transferred into 5ml heptane and 5ml of fix were added. The embryos were gently agitated for 20mins and then transferred into a 1:1 mixture of heptane and 90%MeOH/50mM EGTA pre-cooled on dry ice. After 10mins of frequent vigorous shaking, the embryos were quickly returned to RT. Devitellinised embryos were removed from the bottom of the tube and washed with MeOH. The embryos were rinsed in 9:1 MeOH/fix followed by 5:5 MeOH/fix and post-fixed for 20min in fix at room temperature. The embryos were dehydrated in 30%EtOH, 50%EtOH, 70%EtOH for 10min in each and were stored for long periods at -20°C.

Proteinase digestion: The embryos were rinsed in 50%PTW/50%EtOH and three times in PTW. They were proteinase digested for 5min in 50µg/ml non-predigested Proteinase K (stored as frozen aliquots of 25mg/ml in water) in PTW. The embryos were rinsed twice in 0.2% glycine (10%stock) and twice in PTW. They were fixed again for 20min in 3.7%formaldehyde in PTW and then rinsed five times in PTW.

Hybridisation: The embryos were rinsed in 50%hybridisation solution /50%PTW and in hybridisation solution and then prehybridised in fresh hybridisation solution for at least 2h at 50°C. The probe was denatured at 94°C for 2min and 5µl were added to 50µl hybridisation solution, mixed thoroughly several times and hybridised overnight at 50°C. The probe was prepared using the "Boehringer DIG DNA Labelling Kit No. 1093657". 0.1µg of gel purified DNA fragment, 300bp to 3kb in size, was denatured at 94°C for 5min, quick-frozen on dry ice and melted on to ice. The DNA template was made up to 15µl with distilled water, and 2µl primers (vial 5), 2µl dNTP's (vial 6) and 1µl Klenow DNA polymerase (vial 7) were then added before incubating at 14°C overnight or for 48h. The probe was made up to 50µl with TE and used without further precipitation. The probe was stored at –20°C and could be reused undiluted after the hybridisation.

Antibody staining: Following hybridisation the embryos were washed in hybridisation solution and in 50%hybridisation solution/50%PTW at 50°C. They were washed five times for 30min each in PTW at 50°C and rinsed in PTW at RT. The embryos were incubated in PTW with a 1:2000 dilution of alkaline phosphatase coupled anti DIG antibody for 4h at room temp. A 1:200 dilution of antibody was preabsorbed overnight against 100µl of fixed 0-12h wild type embryos in 1ml PTW at 4°C.

Staining: The embryos were rinsed twice in PTW washed five times with PTW for a total of 3h, and rinsed in NBT buffer if Boehringer blue stain was used, or in 0.1M Tris pH9.5 if Vectorstain Kit II (brown) was used. The embryos were incubated in the appropriate alkaline phosphatase staining solution for 30min to overnight. The staining was stopped by rinsing twice in PTW containing 10mM EDTA pH8. The Boehringer blue stain was made by adding 4.5µl NBT (75mg/ml) and 3.5µl X-phosphate (50mg/ml) to 1ml of NBT buffer. The staining solution of kit II was made by adding two

drops of each of solutions 1, 2 and 3 (supplied with the kit) to 5ml of 0.1M Tris pH9.5. The Boehringer blue stain was more stable than the blue stain (kit I), during mounting in organic solvents. However, the blue stain was found to be more intense in prolonged incubations when assaying non-abundant signal.

Mounting: Embryos were mounted in JB4 methacrylate resin (Polysciences) for whole-mount viewing. Embryos were rinsed in 100% EtOH for 1min and in a small volume of JB4-A+catalyst for 2-5mins. The embryos were then mounted in JB4 (A+B+Catalyst).

In situ hybridization with radioactive probes on wax sections

Was performed using standard methods (Ingham *et al.*, 1985a).

lacZ and *h* antibody labelling:

A mouse polyclonal antibody (m3) raised by Sheena Pinchin against *h* and a rabbit polyclonal antibody against *lacZ* protein (beta 2, kindly provided by John Seeley, NIMR, Mill Hill) was used with a biotinylated second antibody and Diamino benzidine substrate (Macdonald and Struhl, 1986). The staining solution was buffered in CAT buffer pH5.5 instead of pH7. Embryos were mounted in JB4 methacrylate as described above in the section: "Whole-mount *in situ* hybridization with non-radioactive probes".

Photography

The black and white photographs were taken with Kodak Technical-Pan with Zeiss Normaski or dark-field optics. The films were developed with HC110 developer at dilution-B for Normaski and dilution-F for dark field. Colour photographs were taken under similar optical conditions with Fujichrome 64T slide film and printed on Cibachrome paper.

CHAPTER 3: THE 3' END IS RESPONSIBLE FOR APICAL LOCALISATION

INTRODUCTION

The pair-rule genes are part of a hierarchy of segmentation genes which are vital for setting up the pattern of segments in the embryo. They are involved in the process of refining spatial information from the coarse expression patterns of the gap genes to the fine pattern of the segment polarity genes.

The mRNA of *ftz, eve, h, run* and *prd* are asymmetrically distributed in the apical periplasm above the blastoderm nuclei. This intracellular localisation is present throughout the blastoderm stage, starting at the time the pair-rule genes are first expressed. Gap gene mRNAs, on the other hand, are symmetrically distributed (unlocalised) so they are found in the apical and basal periplasm of the blastoderm. Other genes expressed at the syncytial blastoderm show one of three intracellular distributions of mRNA: apical, basal or unlocalised (see chapters 1 and Fig. 1.3).

Asymmetric localisation of mRNA is a potential point of post-transcriptional regulation of the expression of a gene and has been shown to occur in a number of other systems (see chapter 1). However, the mechanism of subcellular transcript localisation remains obscure in most cases. It is possible that the powerful genetic and molecular techniques available in *Drosophila* may be used to shed light on the mechanism and function of localisation and may help to isolate universal components of transcript localisation.

To study the mechanism and function of intracellular transcript localisation in the blastoderm, the region of the pair-rule genes containing sequences responsible for apical localisation must first be determined. There is no *a priori* reason to favour any part of the pair-rule genes as candidate

localisation sequences, so different regions of the pair-rule were tested for their localisation potential.

STRATEGY

A variety of fusion constructs, including fragments of pair-rule genes fused to a number of other genes, were therefore analysed to establish the region necessary and sufficient for localisation. The constructs contained fragments of *ftz, eve, h* fused with *α1-tubulin, hsp70, hsp27* and *β-globin* in a variety of permutations. The patterns of intracellular localisation of the constructs were analysed by whole-mount *in situ* hybridisation using DIG-labelled probes on whole-mount embryos and ^{3}H-labelled probes on wax tissue sections.

RESULTS

ftz localisation sequences

The upstream control region of *ftz* contains all the sequences necessary for its correct spatial and temporal expression and for its tissue-specific expression in the nervous system (Hiromi and Gehring, 1987; Hiromi *et al.*, 1985). The intracellular localisations of the transcripts of several hybrid constructs containing the upstream control region of *ftz* were analysed to determine which region includes sequences necessary and sufficient for localisation.

When the 6kb upstream control sequences of *ftz* are fused onto the bacterial *lacZ* reporter gene encoding β-galactosidase, the reporter gene is expressed in a seven striped pattern similar to that of endogenous *ftz* (Hiromi *et al.*, 1985). However, the intracellular localisation of the mRNA of such fusion genes depends on the other sequences present in the constructs.

ftz/lacC contains the *ftz* upstream control region including the 5' untranslated part of the gene fused onto the *lacZ* coding region. The construct contains an *hsp70* 3' untranslated region and polyadenylation signal (see Fig. 3.3 and Hiromi *et al.*, 1985). The intracellular distribution of *ftz/lacC* was determined using both DIG-labelled probes on whole-mount embryos and tritiated probes on wax tissue sections. Both techniques showed that *ftz/lacC* mRNA is unlocalised, being present in both the apical and basal periplasm of the syncytial blastoderm (see Figs. 3.1, 3.2, 3.3 and 3.4). It follows that sequences necessary for localisation must lie in the coding region or in the 3' untranslated part of *ftz*.

The fact that both *in situ* techniques gave the same results for apically localised *ftz* transcripts and unlocalised *ftz/lacC* mRNA transcripts (see Fig. 3.2, 3.3 & 4.2) indicates that whole-mount *in situ* hybridisation can be used to determine transcript localisation; tissue sections are not necessary. This is an important point since the whole-mount *in situ* technique is substantially faster than *in situ* hybridisation on tissue sections; more constructs could be analysed in whole-mount than by tissue section.

FG2, resembles *ftz/lacC* except that the *hsp70* 3' end is replaced by the 3' end of the *ftz* gene including some of the coding region fused in frame with the *lacZ* coding region (Fig. 3.3). *In situ* hybridisation to whole-mount embryos reveals that the mRNA of *FG2* is correctly localised to the apical periplasm (see Fig. 3.4). The results indicate that the 3' end of the *ftz* gene contains sequences necessary for apical localisation.

The results are confirmed by *LTfSH2*, a construct (described in chapter 4) containing the same *ftz* upstream sequences and *lacZ* coding region as *FG2*. *LTfSH2* is apically localised, indicating that the 3' end of *ftz* (SalI-HindIII) contains sequences necessary for localisation.

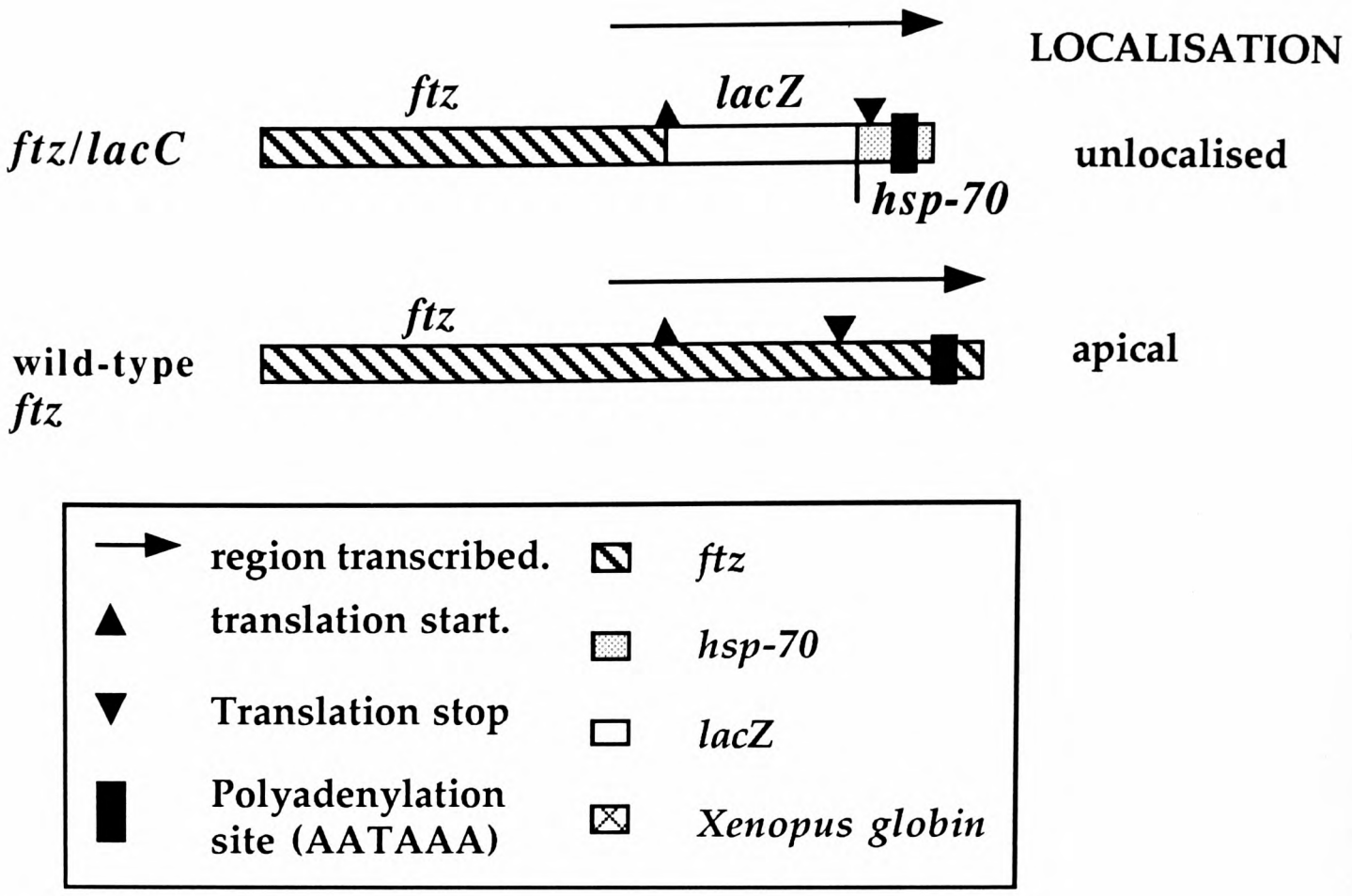
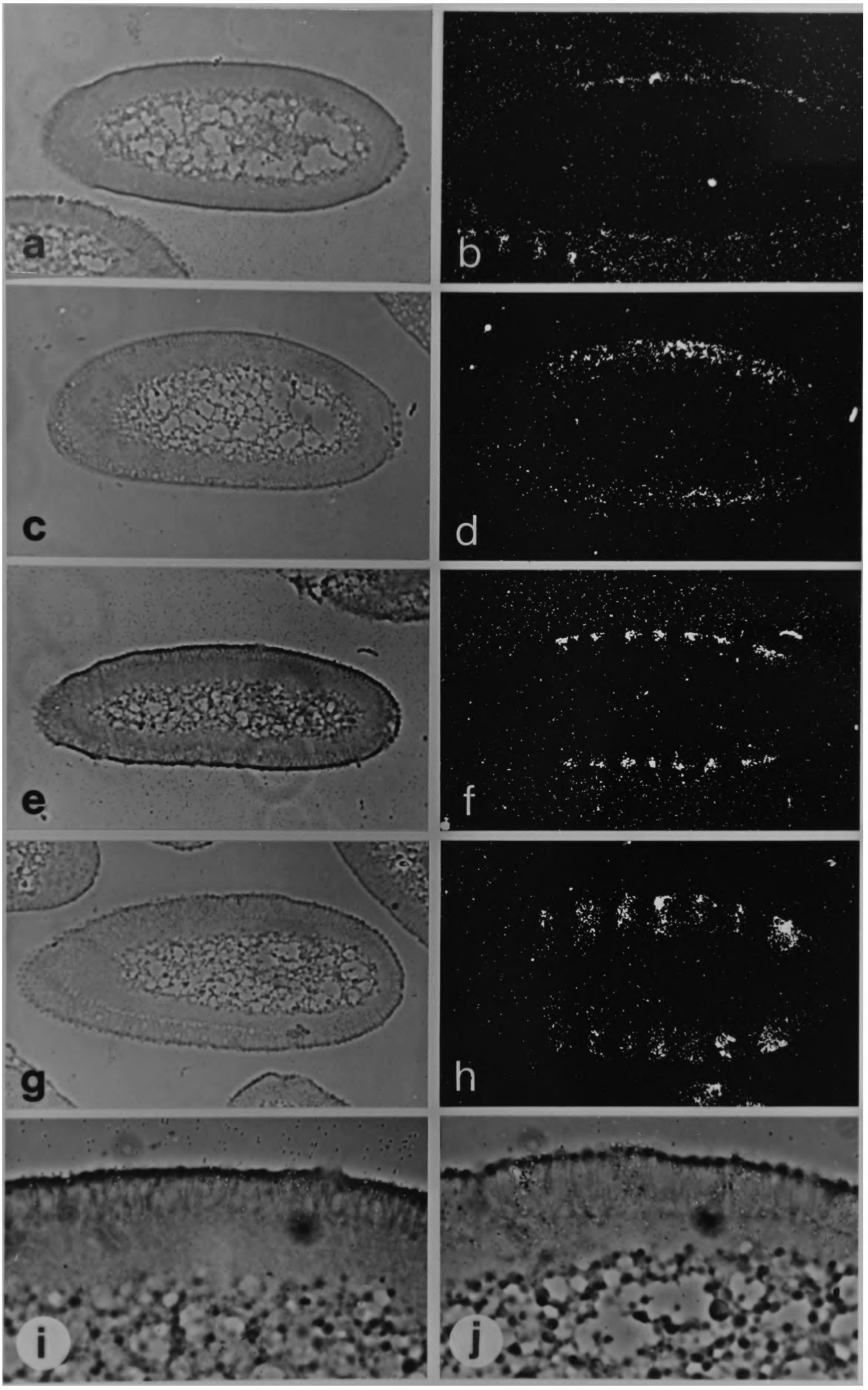


Fig 3.1

Fusion construct, *ftz/lacC,* analysed by *in situ* hybridisation in Fig. 3.2. The construct consists of the upstream promoter elements of *ftz* fused onto the coding region of the *lacZ* reporter gene with an *hsp-70* polyadenylation signal. *lacZ* message and protein are expressed in a striped pattern similar to *ftz*.

Fig. 3.2 (opposite)

***In situ* hybridisation using ^{35}S-labelled probes on wax embedded *ftz/lacC* embryo tissue sections, showing that sequences in the 3' end of *ftz* are necessary for localisation. a), b), e), f) & i) are probed with *ftz* probe and show that ftz mRNA is found in the apical periplasm. c), d), g), h) & j) are probed with a *lacZ* probe and show that ftz/lacC mRNA is unlocalised. a), c), e), g), i) & j) are photographed under phase contrast conditions. b), d), f) & h) are photographed under dark-field conditions.**


a
b
c
d
e
f
g
h
i
j

In chapter 8, *FG2* is shown to contain no translation stop codon after the *lacZ* coding region, so that the *lacZ* protein is fused in frame with 3' end of the *ftz* protein. It is formally possible that the 3' end of the *ftz* gene must be translated for it to function in apical transcript localisation. However, this proposition is refuted by *LTfSH2*, which contains the same localisation sequences as *FG2* but these are not translated. It follows that translation of the region of *ftz* containing localisation sequences is not necessary for apical localisation.

HSF is another construct containing *ftz* sequences. It has an *hsp70* promoter driving the *ftz* coding region and the 3' untranslated region of *ftz*, including a *ftz* polyadenylation signal (Fig. 3.3). The construct also includes a *Xenopus β-globin* polyadenylation signal beyond the *ftz* 3' untranslated region and polyadenylation signal (Struhl, 1985). The intracellular localisation of *HSF* mRNA was analysed by whole-mount *in situ* hybridisation of embryos heat-shocked for 20min at 37°C. *HSF* is transcribed by all the peripheral nuclei, but mRNA is localised to the apical periplasm associated with each peripheral nucleus (Fig. 3.4).

The 5' untranslated part of *ftz* is missing, indicating that it is not required for intracellular localisation. There are no *ftz* sequences downstream of the SphI site present in *HSF*, so sequences necessary and sufficient for localisation must lie upstream of the SphI site (see Fig. 3.3).

ftz^{W20} is a mutation with a break point in the *ftz* intron which eliminates the 3' part of the *ftz* coding region as well as all the 3' untranslated part of the gene (Weiner *et al.*, 1984). The mutant gene contains all the *ftz* 5' untranslated and most of the coding region. The mRNA of the mutant gene is distributed in both the apical and basal periplasm (Ingham and Gergen, 1988). The results confirm those described above indicating that the 5' untranslated part of *ftz* gene is not sufficient for apical localisation of *ftz*

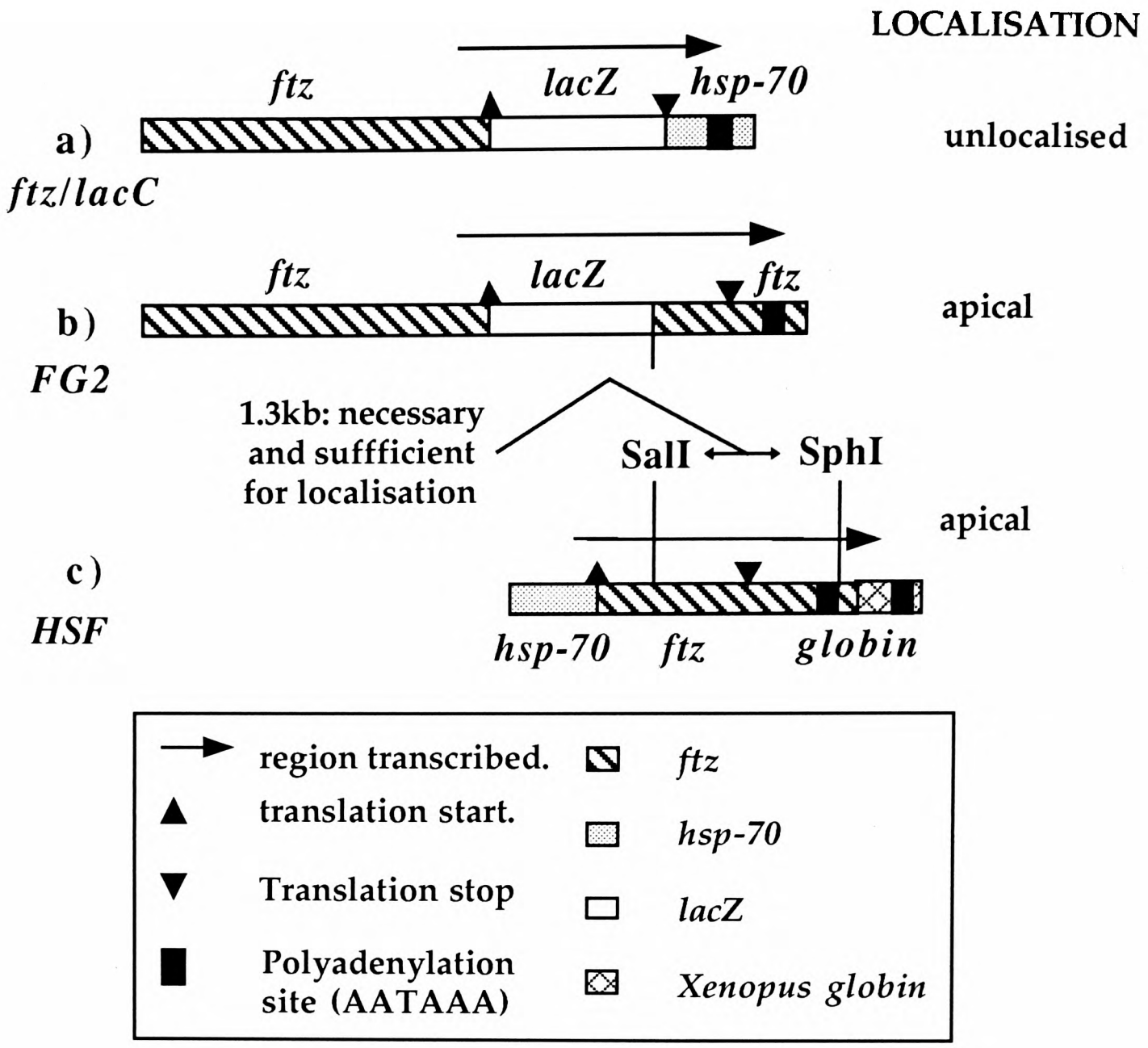
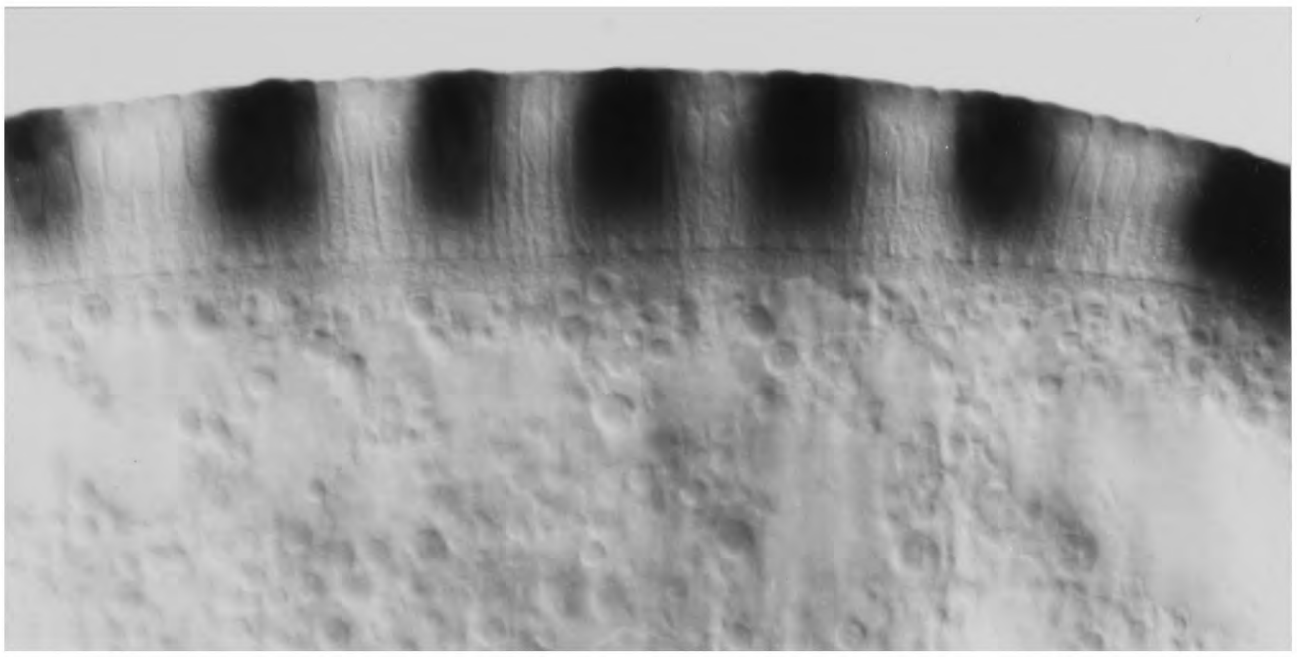


Fig 3.3

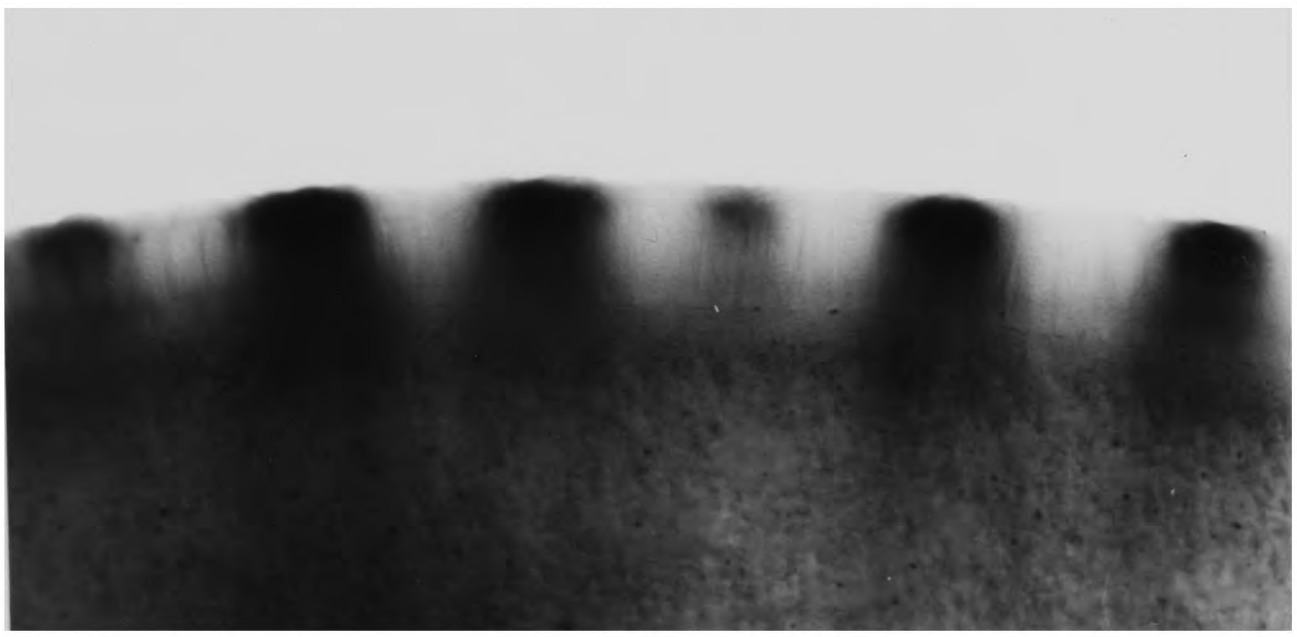
Fusion constructs analysed by *in situ* hybridisation showing that the 3' end of the *ftz* gene is necessary and sufficient for apical localisation. a) *ftz/lacC* as described in Fig. 3.1. b) *FG2* contains the same *ftz/lacZ* fusion as *ftz/lacC*. However, *FG2* contains part of the *ftz* coding region fused in frame with the *lacZ* gene together with the 3' untranslated and polyadenylation signal of *ftz*. c) *HSF* contains the *hsp-70* promoter fused to the *ftz* coding region and 3' untranslated polyadenylation signal of *ftz*. The construct also contains the polyadenylation signal of *Xenopus globin*. The constructs define a 1.3kb region in the 3' untranslated part of *ftz* which is necessary and sufficient for localisation.

Note that in Fig. 3.4a and in a number of subsequent photographs of strongly stained whole mount embryos, localised constructs may give the misleading impression of being unlocalised because of the difficulty of obtaining a true optical section under Nomarski optics. Also note that localised transcripts have distinctly different appearances in early and late stage 14 embryos because of the difference in the thickness of the apical periplasm at these two stages.

a)



b)



c)

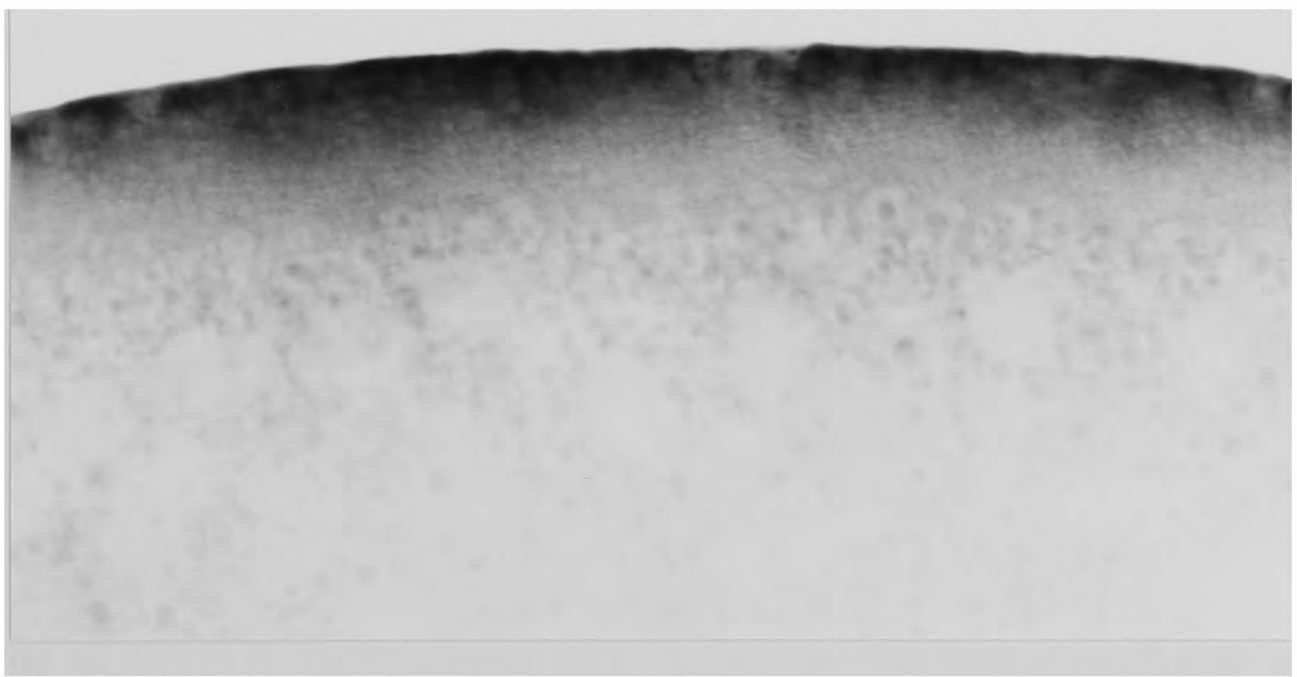


Fig 3.4:

Whole mount *in situ* hybridisation using DIG-labelled probes revealing the intracellular localisation of the constructs in Fig. 3.3 and showing that the 3' end of *ftz* is necessary and sufficient for localisation. a) *ftz /lacC* embryos with *lacZ* probe, b) *FG2* embryos with *lacZ* probe and c) heat-shocked *HSF* embryos (20min at 37°C) with *ftz* probe.

mRNA; sequences necessary for localisation are present in the 3' untranslated part of *ftz*.

The constructs analysed contain various parts of the pair-rule gene *ftz* fused onto genes that do not show apical localisation. The results show that a 1.2kb region between the SalI and SphI sites in the 3' end of *ftz* is both necessary and sufficient for the correct localisation of *ftz* mRNA in the blastoderm. However, it is possible that each pair-rule gene has localisation sequences present in a different part of the gene. I therefore analysed other pair-rule genes to determine if localisation sequences are always present in the same part of the pair-rule genes.

_h_ localisation sequences

Like the other pair-rule genes, *h* transcript is localised to the apical periplasm of in the syncytial blastoderm (see Fig. 1.4). The intracellular distribution of the transcripts of three hybrid genes containing some parts of the *h* gene were analysed by whole-mount *in situ* hybridisation, to determine which region was necessary for *h* localisation. In addition, the transcript distribution of a mutation which deletes part of the *h* coding region and 120bp of the 3' untranslated part of *h* was analysed.

HSH21 and *HSH33* are two heat-shock constructs containing the upstream region of *hsp70* fused onto the *h* coding region. The fusions do not include any 5' untranslated sequences of the *h* gene. *HSH21* is fused onto the 3' untranslated part of the *hsp70* including the polyadenylation signal. *HSH33* contains all the *h* 3' untranslated end, including the polyadenylation signal (see Fig. 3.5).

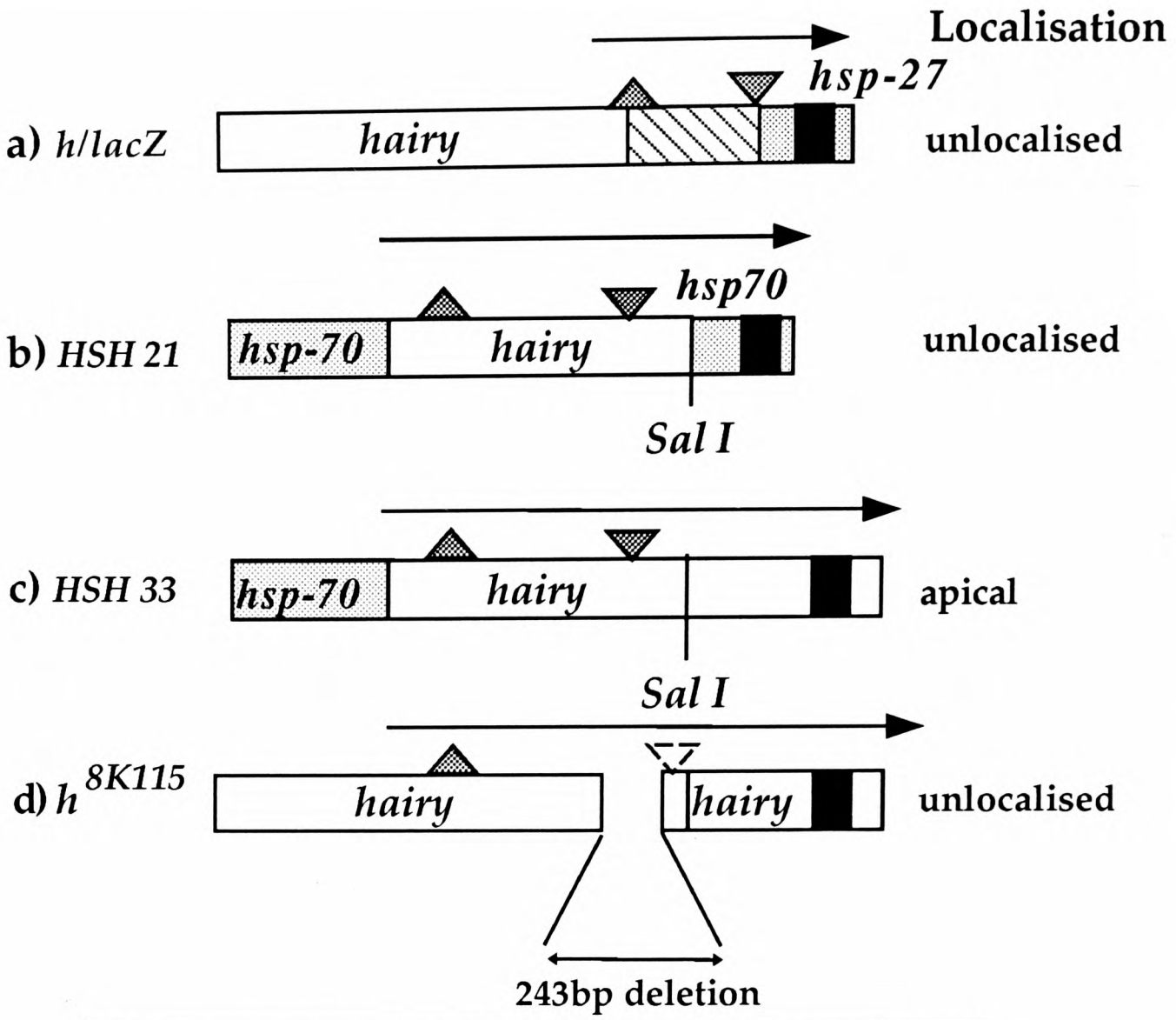
After 20min heat-shock at 37°C, both constructs are transcribed by all the peripheral nuclei, but the two construct differ in their intracellular distribution. *HSH33* mRNA is localised in the apical periplasm above the

peripheral nuclei while *HSH21* mRNA is found in both the apical and basal periplasm associated with each peripheral nucleus (see Fig. 3.6). The results suggest that 600bp of the 3' end of the *h* gene (SalI-XhoI), including the polyadenylation signal, are necessary for localisation. The upstream control region and 5' untranslated part of *h* contains no sequences necessary for localisation (Rushlow *et al.*, 1989).

h/lacZ contains the upstream control region of *h* sufficient to specify expression of the gene in seven stripes but insufficient to specify expression in the head-patch. The construct includes part of the 5' untranslated region of the *h* gene, but no coding or 3' sequences. The 5' end of *h* is fused to the *lacZ* coding region, which in turn is fused in frame with a nuclear localisation tag. The 3' end of the fusion construct consists of the 3' untranslated part of the *hsp27 Drosophila* heat-shock gene (Riddihough *et al* in preparation-see Fig. 3.5).

Whole-mount *in situ* hybridisation reveals that *h/lacZ* mRNA is expressed in seven stripes and lacks a head-patch. *h/lacZ* mRNA is found in both the apical and basal periplasms, confirming the results found with *HSH* constructs (see Fig. 3.6).

h^{8K115} is a deletion of 114bp in the 3' part of the coding region and 120bp of the 3' untranslated region of the *h* gene (Wainwright, personal communication-see Fig. 3.5). *In situ* hybridisation to tissue sections with tritiated *h* probe reveals that the transcript transcribed by the homozygous mutant is unlocalised, being present in both the apical and basal periplasm of peripheral nuclei in each stripe and the head-patch (see Fig. 3.7). Alternate sections were probed with a *ftz* probe so as to identify the homozygous mutant embryos by the lack of *ftz* striping in late cycle 14 blastoderm embryos (Howard and Ingham, 1986 and see Fig. 3.7). The results indicate that sequences necessary for localisation are found in the



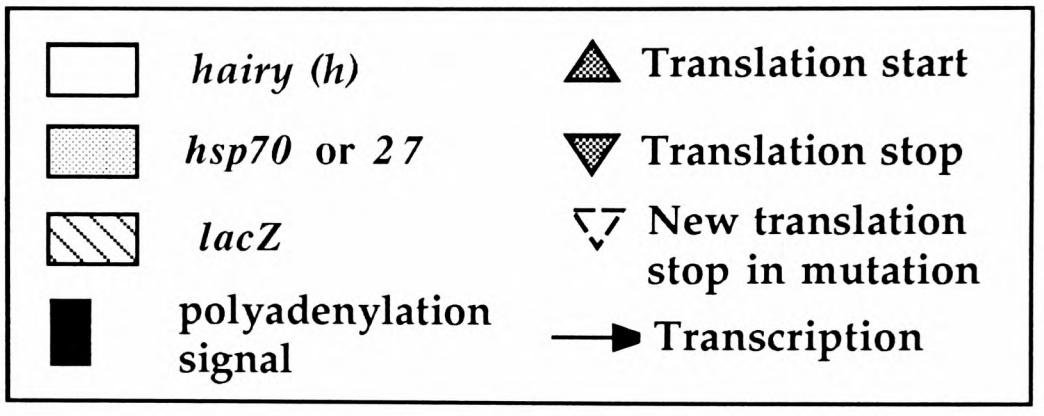


Fig. 3.5

The structure of constructs analysed in Fig. 3.6, revealing that a fragment of the 3' untranslated part of *h* is necessary for apical localisation.

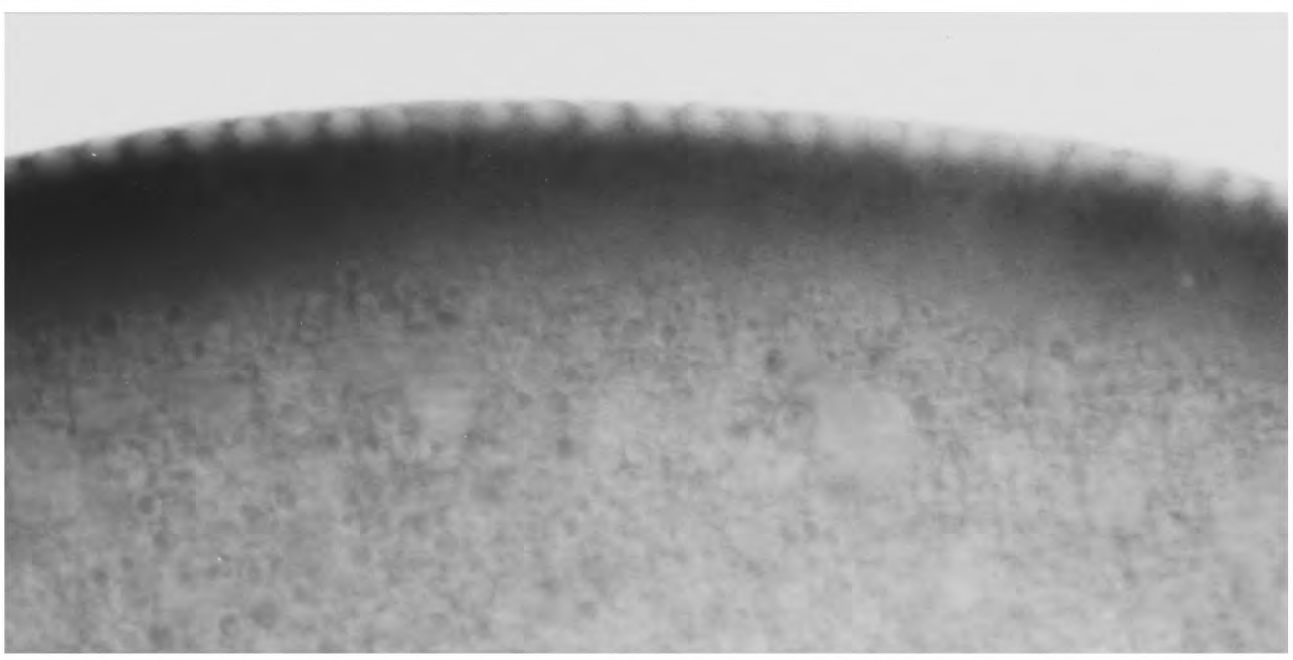
a) *h/lacZ* is a fusion gene made by G. Riddihough, containing the *h* upstream promoter elements controlling the expression of the *lacZ* reporter gene. It contains the *hsp27* polyadenylation signal.

b) & c) are heat shock *hairy* (*HSH*) constructs containing the *hsp70* promoter fused to the coding region of *h*. *HSH21* an *hsp70* polyadenylation signal, *HSH33* contains the entire 3' untranslated part of *h* including the polyadenylation signal. The embryos were heat shocked for 20min at 37^{O}C.

d) h^{8K115} is a *h* mutation in which 243bp are deleted in the coding region and in the 3' untranslated part of the gene. The mutation introduces a new stop codon 6bp after the deletion (M. Wainwright, personal communication).

a)

b)



c)

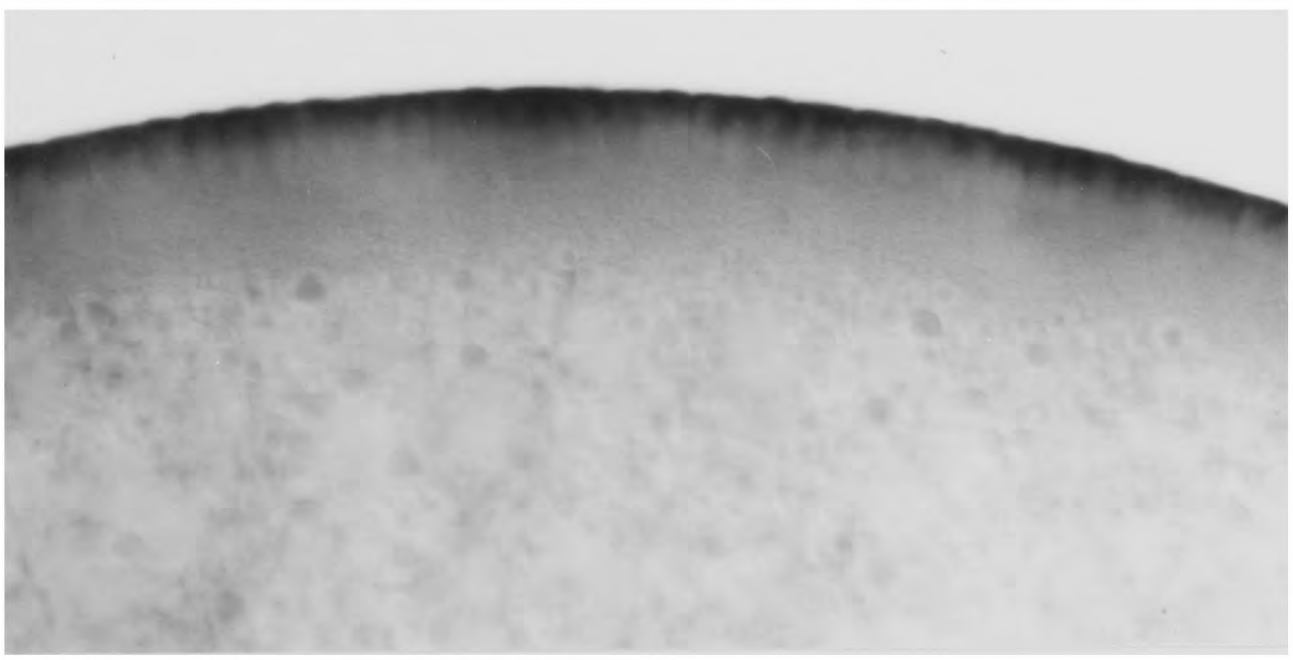


Fig 3.6:

Whole-mount *in situ* hybridisation using DIG-labelled probes revealing the intracellular localisation of the constructs in Fig. 3.5. and showing that the 3' untranslated part of *h* contains sequences necessary for localisation. a) *h/lacZ* embryos with *lacZ* probe, b) *HSH21* and c) *HSH33* heat-shocked embryos (20min at 37°C) with *h* probe.

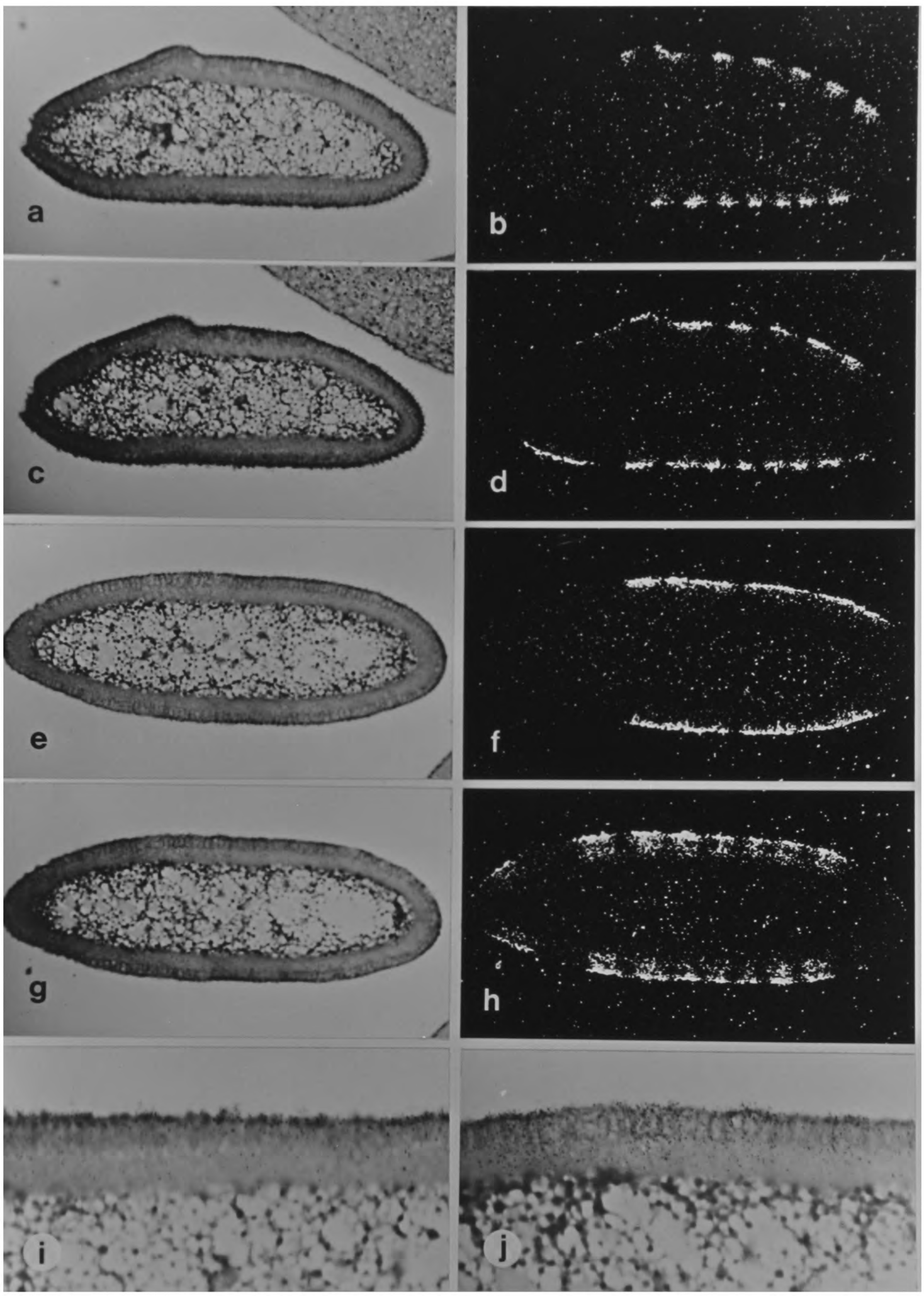
Fig. 3.7

In situ hybridisation with *h* and *ftz* ^{3}H labelled radioactive single stranded RNA probes showing that *h* transcript is unlocalised in h^{8k115} embryos. The 243bp region deleted in the mutation thus includes sequences necessary for apical localisation.

a), c), e), g) are photographed under phase contrast, showing the morphology of the embryo. i), j) are high powered bright field views of the embryos shown in d) and h) respectively.

b), d), f), h) are taken under dark-field conditions, revealing only the grains in photographic emulsion coating the section. The grains indicate where radioactive probe hybridises to RNA.

b) and d) are alternate sections of a wild type embryos probed with *ftz* and *h* probes respectively. f) and h) are alternate sections of a h^{8k115} homozygous embryos probed with *ftz* and *h* probes respectively. The results show that in h^{8k115} embryos the *h* transcript is unlocalised while *ftz* stripes are broader than usual.


a
b
c
d
e
f
g
h
i
j

the mutant, as well as the region defined by in situ hybridisation on the the *HSH* constructs.

The experiments described above show that an extensive region of the 3' untranslated part of *h* is necessary for the correct localisation of *h* transcript in the blastoderm.

α1–tubulin Localisation sequences

eve/lacZ is a fusion gene containing some of the upstream control region, all of the 5' untranslated part of the *eve* gene and 22 *eve* amino acids fused in frame with the *lacZ* coding region. The 3' end of the construct contains 600bp of *α1–tubulin* 3' untranslated sequences including a polyadenylation site (Lawrence *et al.*, 1987). The analysis of the intracellular distribution of *eve/lacZ* mRNA reveals that it is apically localised (see Figs. 3.8 & 3.9). There are two possible explanations for this result. Either the *eve* sequences included in the construct or the *α1–tubulin* 3' end contains sequences necessary and sufficient for apical localisation. I therefore tested two other hybrid constructs containing the *α1–tubulin* 3' end to determine whether *α1–tubulin* contained sequences sufficient for apical localisation.

hb/lacZ is a fusion construct containing the upstream region of the *hb* fused onto the *lacZ* coding region and containing an *α1–tubulin* 3' end including a polyadenylation signal. Two alternative constructs were tested by *in situ* hybridisation to whole-mount embryos. The two constructs differed in the length of the *hb* sequences present in the construct; *HB4.2* and *HB0.8* contain 4.2kb and 0.8kb of the upstream region of *hb* respectively. Neither *hb* nor *lacZ* contains apical localisation sequences, and yet the results show that the transcripts of both constructs are localised to the apical periplasm (see Fig.3.9). It can be concluded that the 3' untranslated part of the *α1–tubulin* gene contains sequences necessary and sufficient for localisation.

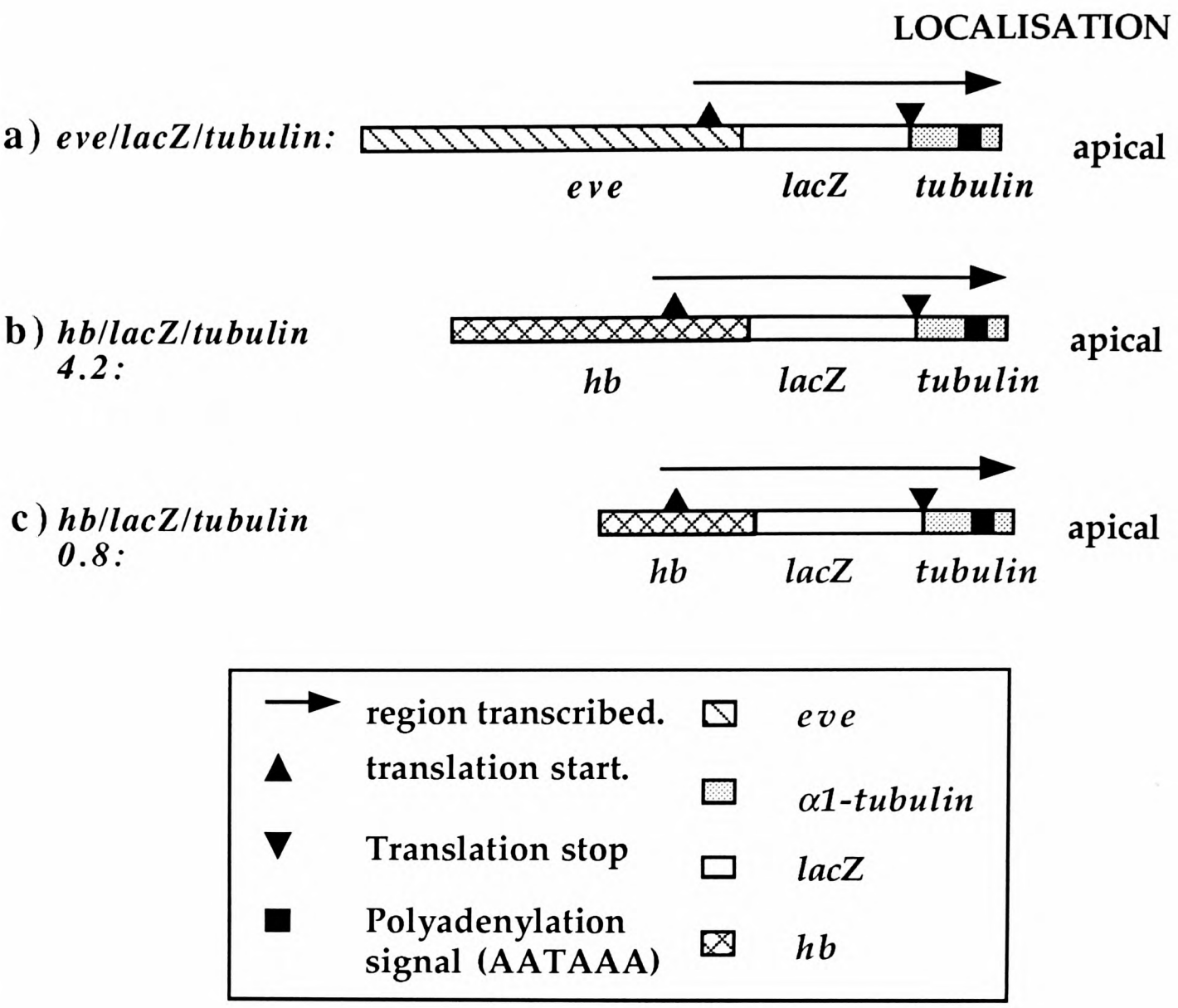
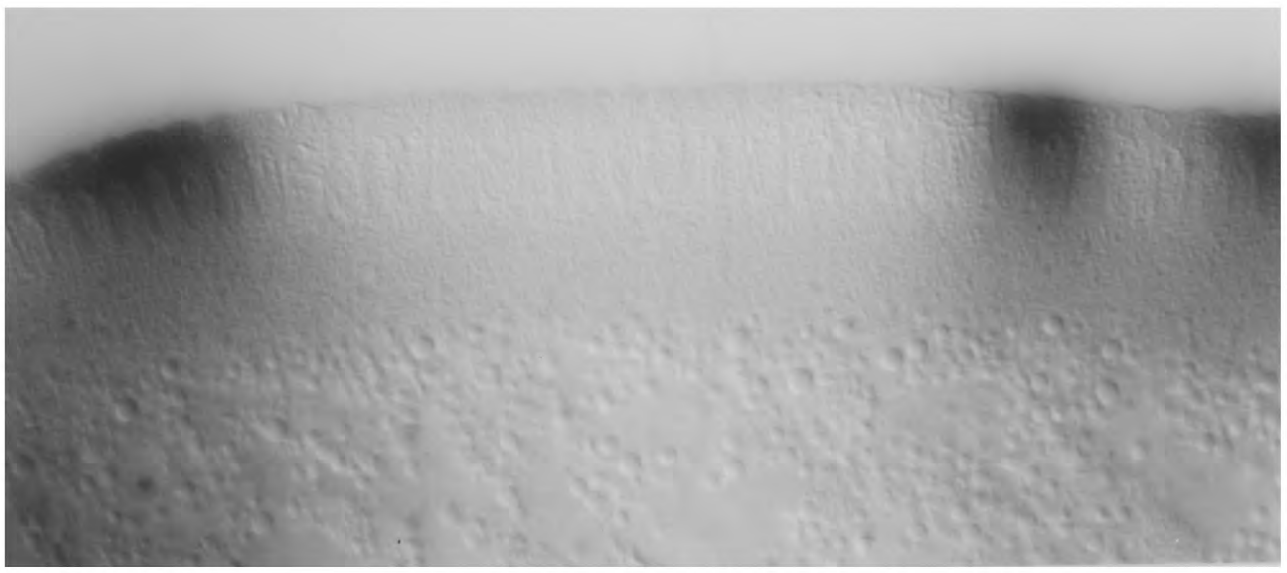


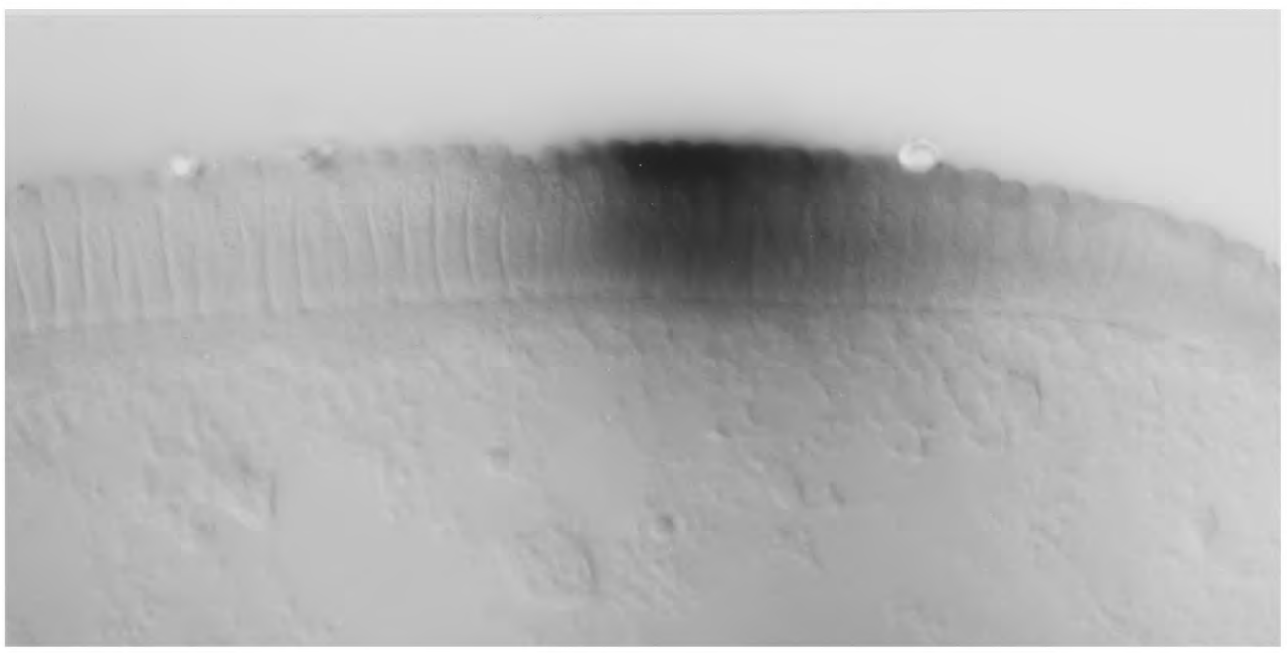
Fig 3.8

Constructs containing the 3' untranslated part of *tubulin* and either the *eve* or the *hb* upstream control sequences regulating the expression of the *lacZ* reporter gene. The results show that the 3' untranslated end of tubulin contains sequences sufficient for apical localisation
The *eve /lacZ* construct contains the upstream control region of the *eve* gene sufficient for the expression of some but not all of the seven stripes. The *eve* sequence used in the fusion contains 22 *eve* amino acids fused in frame at an AvaII site with the *lacZ* gene (Lawrence et al, 1987). It contained 800bp of the 3' untranslated part of the *tubulin* gene including the polyadenylation signal. The *hb /lacZ* constructs (Struhl et al, 1989) contain the upstream region 0.8kb or 4.2kb of the upstream control region of *hb* together with the first 75 amino acids of the coding region of *hb* fused in frame with *lacZ* coding region. The 3' end of both constructs consists of the 3' untranslated part of the *tubulin* gene including the polyadenylation signal.

a)



b)



c)

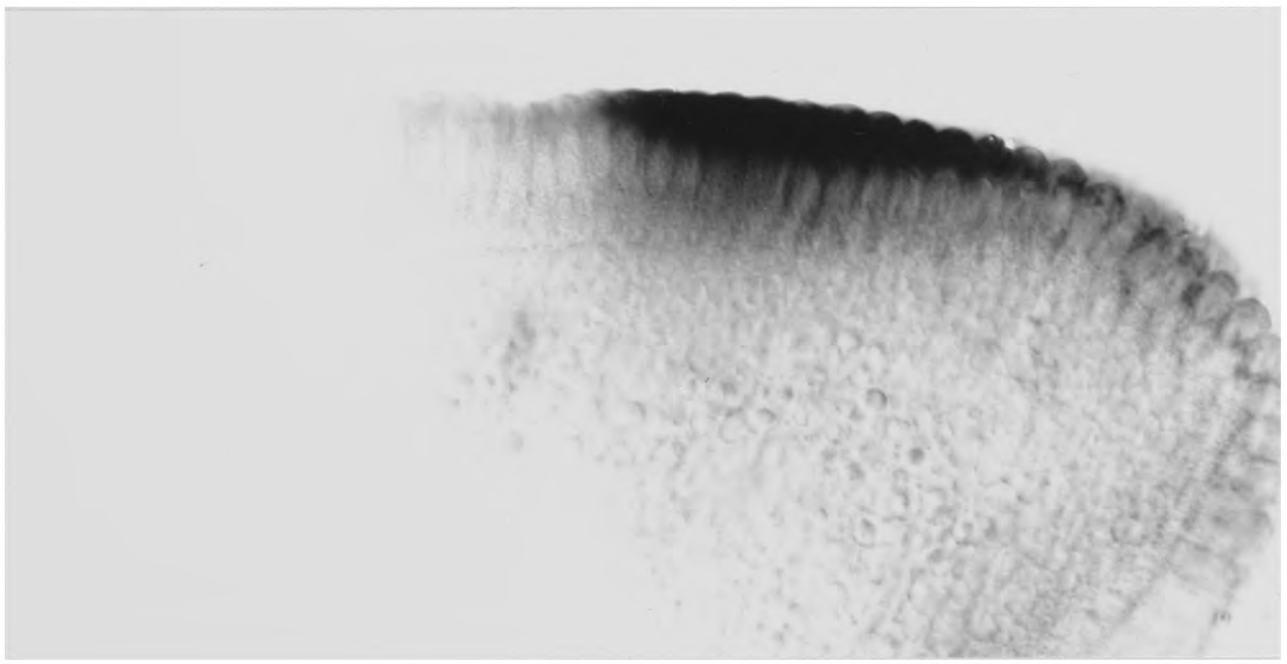


Fig 3.9

Whole-mount *in situ* hybridisation using DIG-labelled probes revealing the intracellular localisation of the constructs in Fig. 3.8. The results show that the 3' untranslated part of the *α1-tubulin* gene contains sequences sufficient for apical localisation. a) *eve/lacZ*, b) *HB4.2* and c) *HB0.8* embryos probed with *lacZ* probe.

DISCUSSION

The analysis of a variety of fusion constructs has revealed that the intracellular localisation of zygotic mRNA transcribed in syncytial blastoderm embryos is dependent on sequences in the 3' untranslated part of construct. In the case of the pair-rule genes, the region containing sequences necessary for apical localisation includes part of the coding region. However, it is unlikely that the coding region is in any way necessary for localisation; in the next chapter it is demonstrated that part of the 3' untranslated region of *eve* is necessary and sufficient for apical localisation. It is likely that the three genes *ftz*, *eve* and *h* share a common mechanism for localisation and that their localisation sequences are interchangeable.

The pair-rule genes are not the only known case of asymmetric transcript localisation which depends on sequences in the 3' untranslated part of the gene. The *Drosophila bcd* gene is localised to a tight cap at the anterior of the oocyte in the mother. *bcd* mRNA is localised in the anterior pole of the *Drosophila* embryo (see chapter 1). The 3' untranslated part of the *bcd* gene is necessary and sufficient for the localisation of its mRNA (Macdonald and Struhl, 1988). *Vg1* mRNA becomes localised in a tight crescent shaped band in the vegetal hemisphere of the *Xenopus* oocyte (see chapter 1). The 3' end of the *Vg1* gene is responsible for the localisation of its mRNA and translation of the *Vg1* protein does not need to be translated for localisation to occur (Yisraeli and Melton, 1988).

Both *bcd* and *Vg1* are maternal transcripts found in the oocyte. The mechanism of localisation in these mRNAs may or may not be related to the zygotic localisation of pair-rule genes. However, it is striking that the regions responsible for localisation appear to be in a similar part of the

genes. In chapter 5, a comparison is attempted between the localisation of maternal *bcd* mRNA and zygotic pair-rule gene mRNA.

In the case of *bcd*, a large region of secondary structure has been proposed to explain why the region cannot be reduced beyond 630bp (Macdonald and Struhl, 1988). It is possible that the apical localisation sequences found in the 3' end of the pair-rule genes must also fold into a complex RNA secondary structure in order to function. It must be noted in this context that secondary structure prediction programmes are notoriously unreliable; a large region of RNA will always form some secondary structure, but its biological significance remains doubtful without further experimentation (see appendix B).

3' untranslated end of *α1-tubulin* contains apical localisation sequences. Maternal *α1-tubulin* transcripts may be localised to the apical periplasm but these experiments have not yet been performed. Microtubules are found more extensively in the apical periplasm than in the basal periplasm; the apical localisation of *tubulin* mRNA may deliver tubulin protein to the site of its polymerisation. Alternatively, it is possible that *α1-tubulin* fortuitously contains correct localisation sequences, or that the localisation sequences are involved in a different cellular mechanism which uses similar cellular components.

The next question to be addressed is the minimal size of the region required for apical localisation. It is possible that the region necessary for localisation must form a large secondary structure, as appears to be the case for *bcd*. The second possibility is that the sequences necessary for apical localisation are smaller than, for example, 100bp. In such a case the localisation signal may operate on the level of primary RNA or DNA sequences interacting with a restricted number of cellular components in a sequence specific manner.

The experiments described in the next chapters address some of these questions.

In summary, the 3' untranslated parts of *h*, *ftz* and *α1–tubulin* contain sequences necessary and sufficient for apical localisation. It seems likely that the 3' untranslated part of all the pair-rule genes contain sequences required for the apical localisation of mRNA. In the next chapter a vector, which itself is not localised and can be used to test the presence of localisation sequences in fragments inserted into the 3' end of the construct, is used to define the smallest region sufficient for apical localisation.

CHAPTER 4: DEFINING THE *eve* AND *ftz* LOCALISATION SIGNAL

INTRODUCTION

In the previous chapter it was shown that sequences in the 3' untranslated end of *ftz* and *h* are necessary and sufficient for apical localisation of pair-rule mRNA. The 3' untranslated regions included in the transcripts of *ftz* and *h* are 480bp and 830bp respectively. I decided to try to define the smallest region of the 3' untranslated end necessary for apical localisation of at least one pair-rule gene.

Of all the pair-rule genes, *eve* and *prd* have the shortest 3' untranslated region: approximately 250bp in length (Frigerio *et al.*, 1986; Macdonald *et al.*, 1986). If the *eve* or *prd* localisation signals were also present in the 3' untranslated part of the gene then narrowing them down would be easier than in *ftz* and *h*. I decided to concentrate on narrowing down the *eve* apical localisation sequence, since *eve* is better characterised than *prd*. At the same time, I attempted to define a small region of *ftz* sufficient for localisation which could be used in sequence comparisons with the *eve* localisation sequences.

STRATEGY

The *pLT* (or *LT*) plasmid vector was constructed with a *ftz* upstream control region driving the expression of the *lacZ* reporter gene at the blastoderm stage. *β-globin* 3' sequences were included to ensure efficient polyadenylation. The *LT* transcript is distributed in the basal periplasm so DNA fragments can be inserted in the 3' end of the vector to test whether they can direct apical localisation of the hybrid transcript.

RESULTS

Constructing *LT*, a basally localised plasmid vector for testing apical localisation sequences

LT was constructed in order to test for the presence of localisation sequences in different parts of the pair-rule and other genes. The plasmid is based on *pUChsneo*, a plasmid vector containing the neomycin resistance gene under the control of a heat-shock promoter and including P-element inverted repeats for transposition (Steller and Pirrotta, 1985). The vector enables the introduction of foreign genes into a polylinker and their stable integration into the *Drosophila* genome (see chapter 2 for details).

The 3' end of the human *β-globin* gene was inserted into the polylinker of *pUChsneo*. Two partly complementary synthetic oligonucleotides (oligos) were then introduced upstream of *β-globin* in the *pUChsneo* polylinker. The oligos contained KpnI, SalI and NotI sites in that order, starting from the 5' end, as well as a protein stop codon directly after the SalI site. The KpnI to SalI fragment of *FG2* containing the *ftz* upstream region fused to the *lacZ* gene, was inserted into the KpnI and SalI sites of the oligos, so that the stop codon in the oligo is in frame with the *lacZ* open reading frame (see Fig. 4.1 and chapter 2).

In situ hybridisation using DIG-labelled probes revealed that *LT* mRNA is found exclusively in the basal periplasm (see Fig.4.2). It follows that there are three distinct patterns of transcript distribution in the blastoderm cytoplasm: apically localised, basally localised and unlocalised. The *LT* vector is thus suitable for testing the ability of a sequence inserted into its 3' end to confer apical localisation.

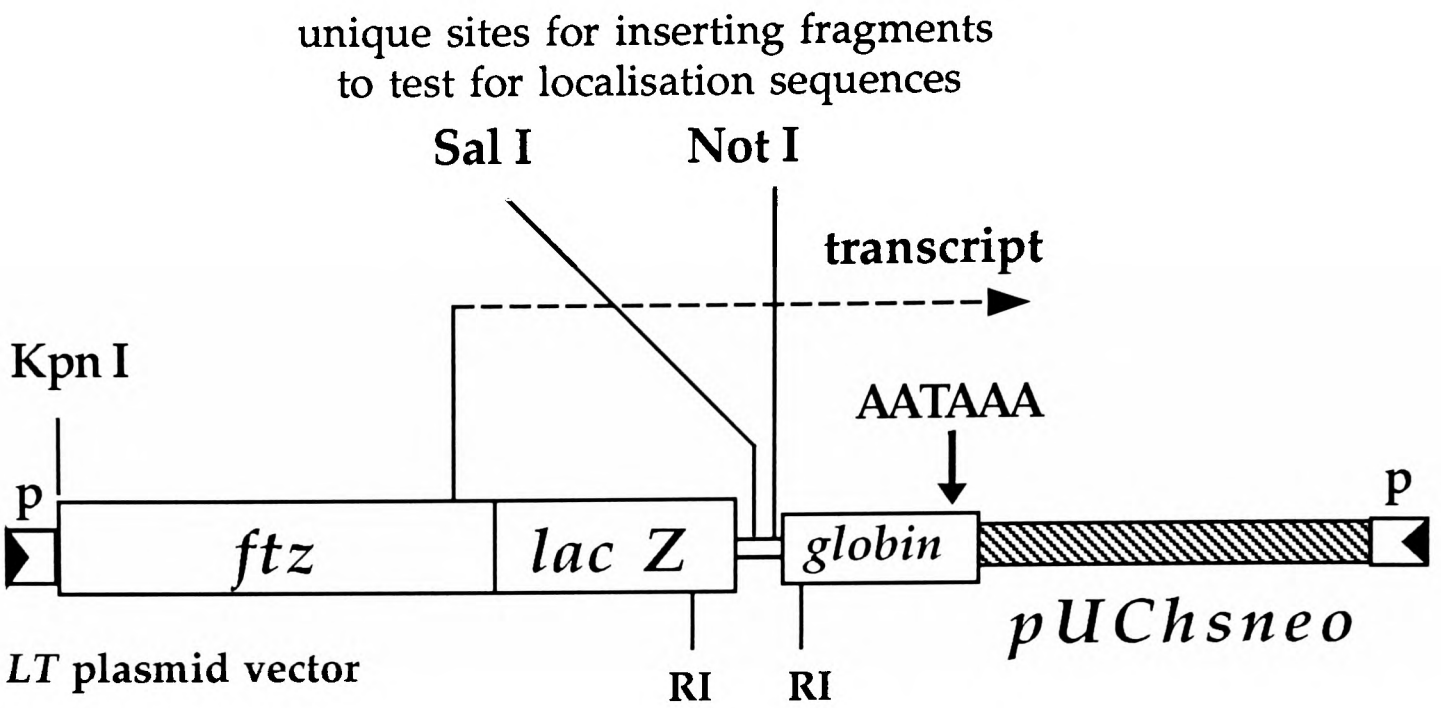
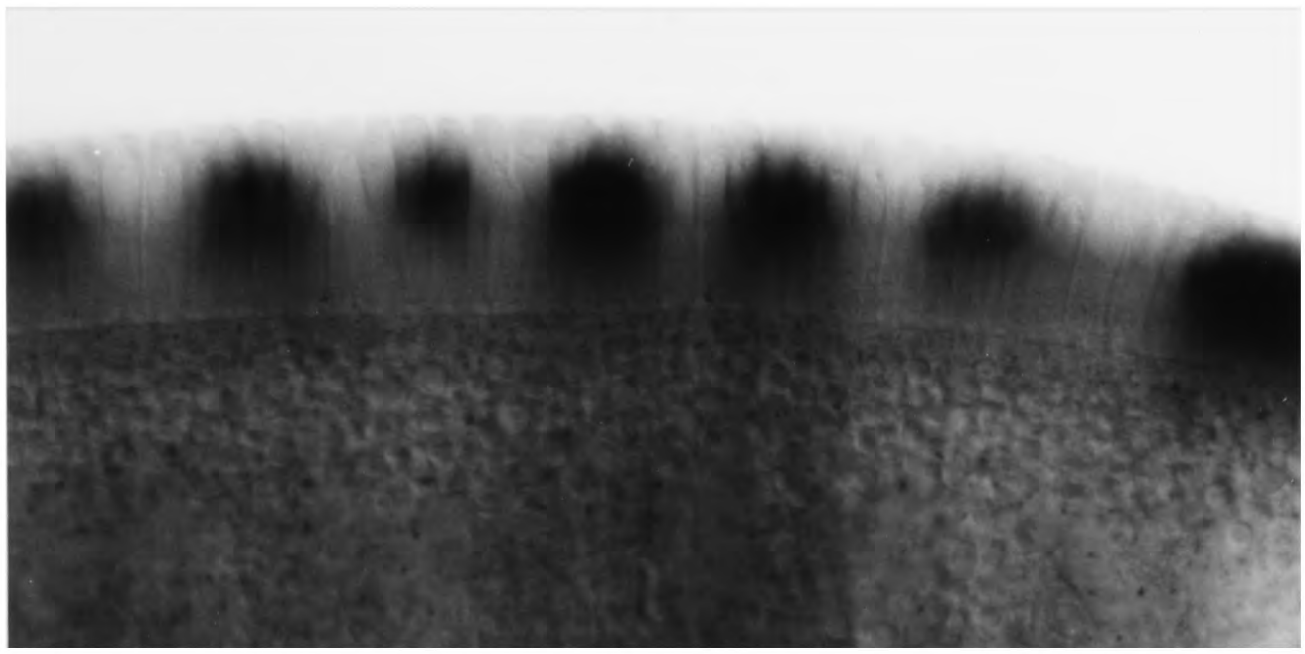


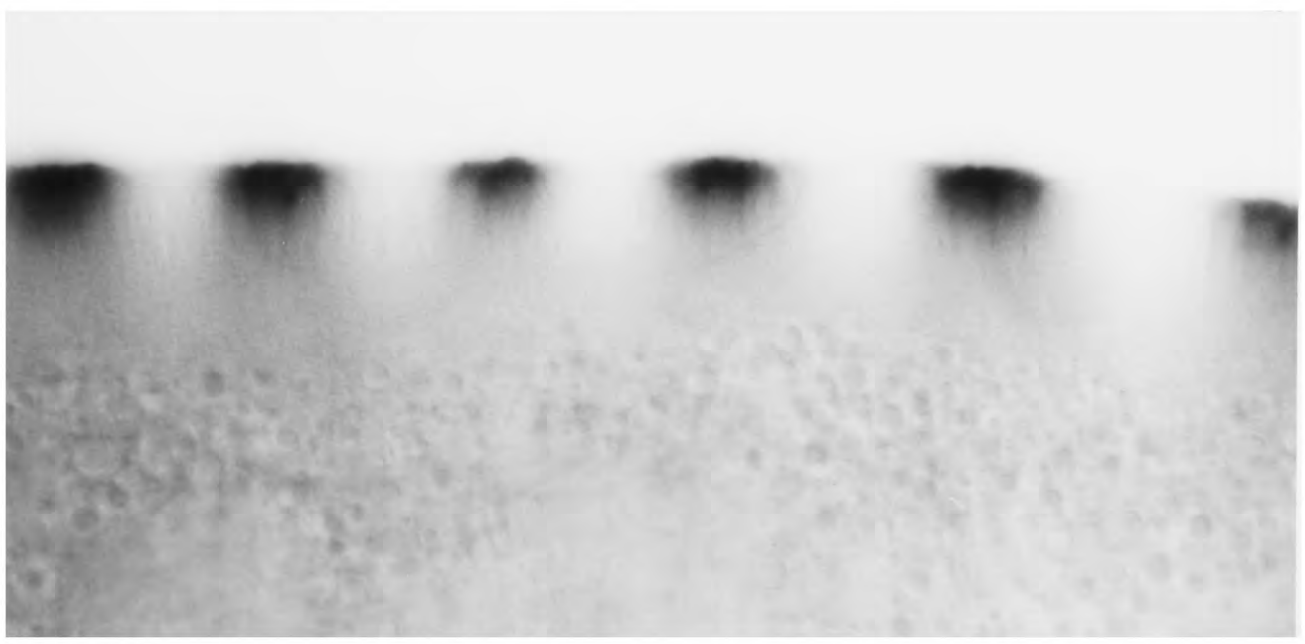
Fig 4.1

The *LT* plasmid vector is based on the *pUChsneo* plasmid vector which enables integration into the *Drosophila* genome by P-element transposition and selection of transformed lines by neomycine resistance. *LT* has the *ftz* upstream control region driving the expression of the *lacZ* reporter gene in a *ftz*-like pattern of seven stripes at the blastoderm stage. A translation stop codon is present in frame with the *lacZ* reporter gene to ensure that a normal β-galactosidase protein is translated for X-gal staining. The *β-globin* 3' untranslated sequences are included after the *lacZ* coding region to ensure correct termination and polyadenylation. The transcript of the vector is distributed in the basal periplasm of the syncytial blastoderm so that DNA fragments can be inserted into two unique sites in the 3' end of the vector to test for the presence of apical localisation signals in fragments. (see chapter 2 for details of making *LT*). P represents P-element inverted repeats and RI are EcoRI sites.

a)



b)



c)

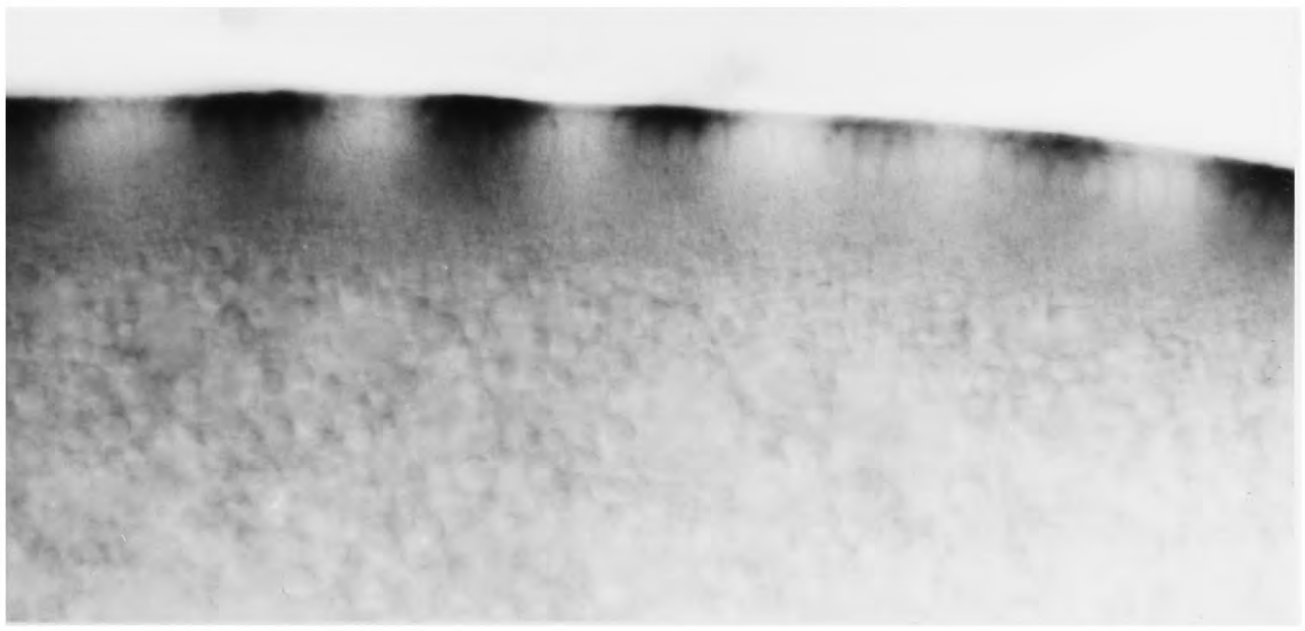


Fig. 4.2

The intracellular localisation of *LT* mRNA compared with that of *ftz* and *eve* transcripts. *LT* transcript is localised to the basal periplasm and pair-rule message is localised to the apical periplasm. Whole mount *in situ* hybridisation was performed using the following DIG-labelled probes: a) *lacZ*, b) *ftz* and c) *eve*. a) is an embryo transformed with the *LT* plasmid, b) and c) are wild type embryos.

Strategies for inserting fragments of *ftz*, *eve* and *hsp70* into *LT*.

The SalI and NotI sites in the 3' end of *LT* are unique so that sequences may be inserted into the sites in order to test their ability to confer apical localisation. The fragments inserted into *LT* were made using two very different strategies. The first strategy used conventional cloning techniques to subclone the 3' end of *ftz* (HindIII to SalI) into the polylinker of a kanamycin-resistant vector (see Fig. 4.3). The 3' untranslated fragment of *ftz* could then be excised with NotI and inserted into the NotI site of *LT*. Both orientations were recovered to give *LTfSH* (see Fig. 4.6) and *LTfHSINV* (see chapter 7). The subclone was digested with EcoRI (RI) and religated, thus deleting all the *ftz* sequences 3' of the RI site. The SalI to RI fragment of *ftz* was then excised with NotI and inserted into the NotI site of *LT* (see Fig. 4.6).

The second strategy used appropriate PCR primers in *ftz*, *eve* and *hsp70* to generate fragments of the 3' untranslated regions of the genes. The PCR primers were designed with SalI and NotI restriction enzyme sites so that the PCR product could be inserted into the SalI and NotI sites of *LT*. A protein stop codon was introduced by the 5' primers to terminate translation of *lacZ* (see Figs. 4.4 & 4.5).

The intracellular transcript localisations of the *LT*-based constructs generated by the two strategies were examined using whole-mount *in situ* hybridisation with DIG-labelled *lacZ* probes (see chapter 2).

***ftz* localisation sequences**

It was important to establish that the *LT* vector is indeed suitable for testing the ability of DNA fragments inserted into the vector to confer apical localisation. The most satisfactory test was to insert a fragment already

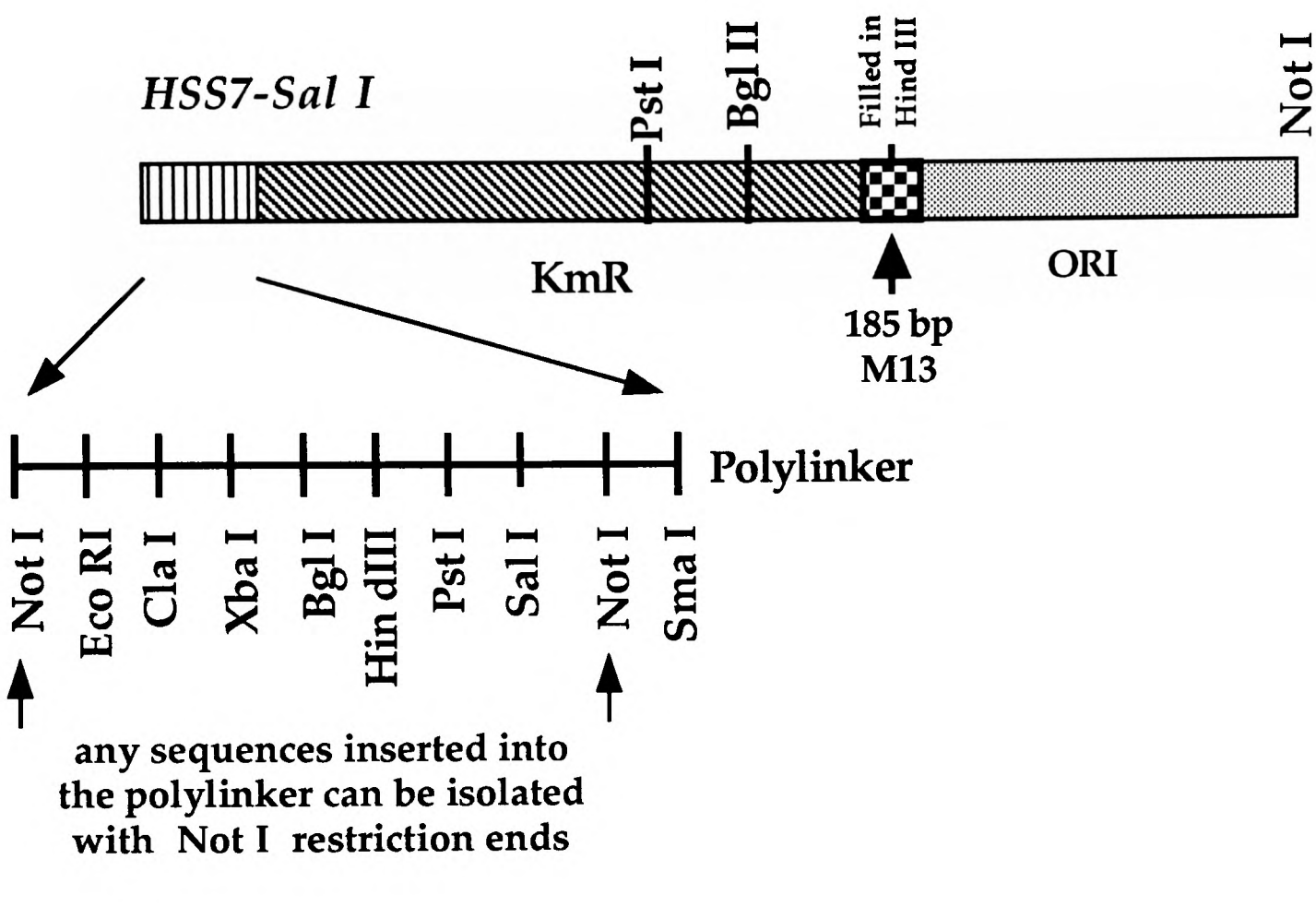


Fig 4.3

HSS7-Sal, a kanamycin-resistant plasmid used to insert the 3' end of *ftz* and create a variety of deletions in the 3' end of the gene. Deleted fragments are excised and gel purified with NotI sticky ends. KmR is the kanamycine resistance gene. ORI is the origin of replication of the M13 plasmid.

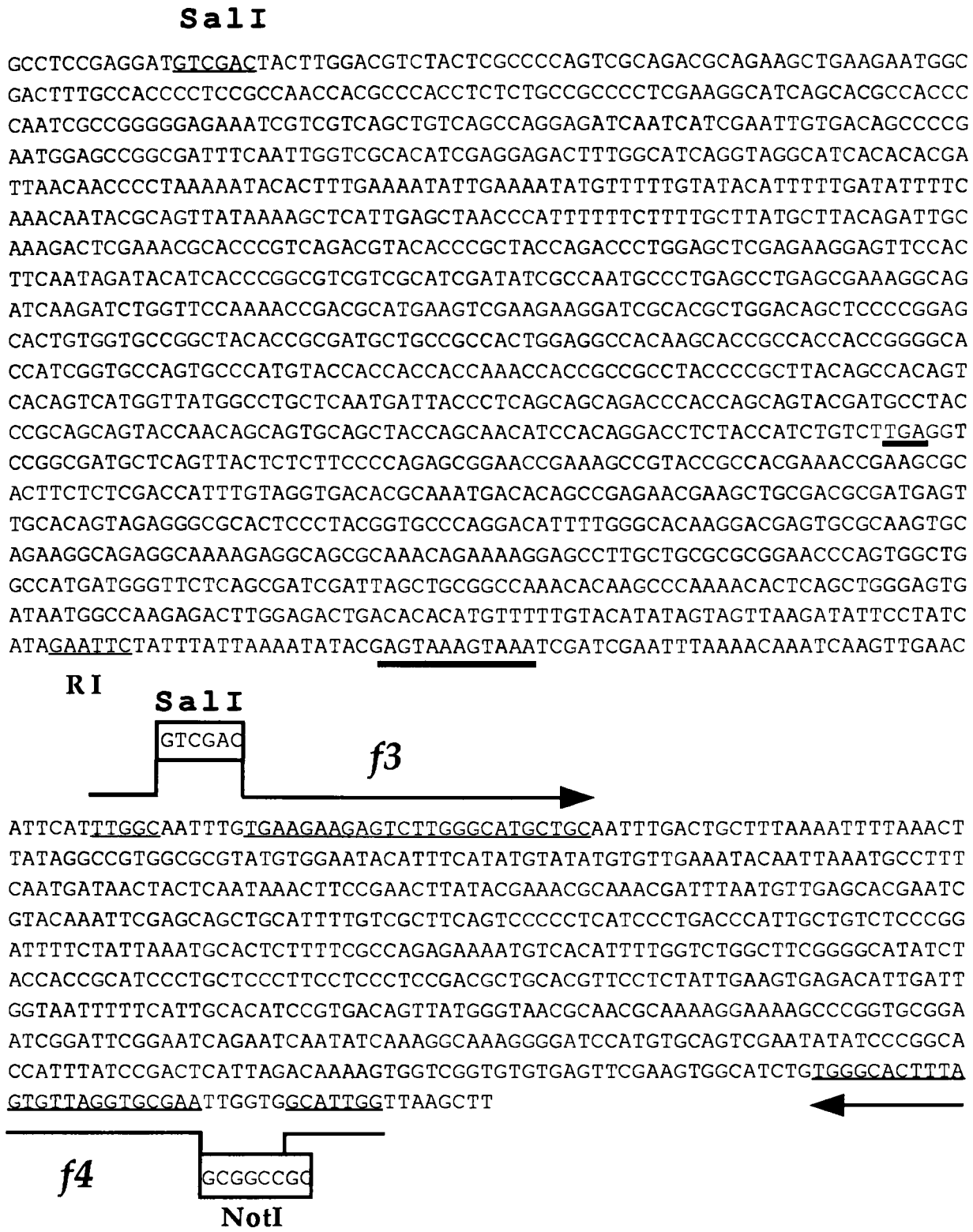
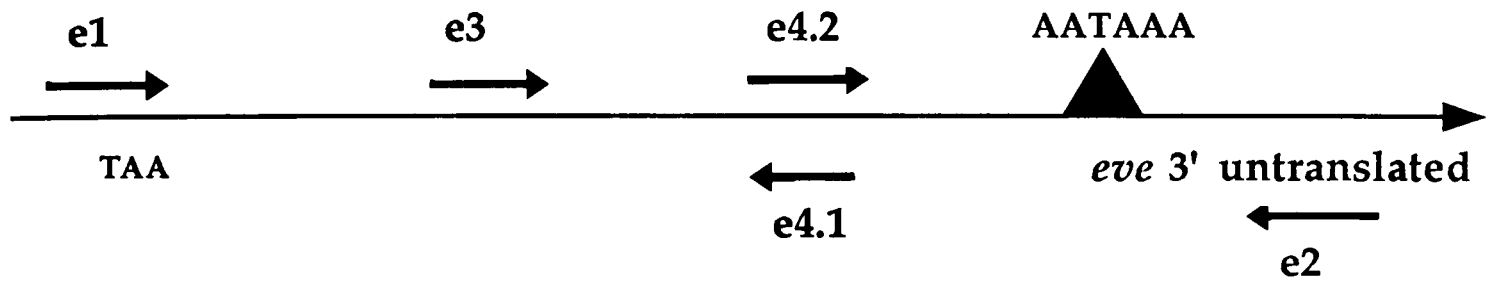


Fig. 4.4

A map of the PCR primers used to generate a fragment of the 3' untranslated part of *ftz*. Arrowed lines represent parts of the primers which are homologous to *eve* sequences. Sequences in boxes are either SalI or NotI sites, bold lines indicate the polyadenylation signal (see chapter 6) and termination codon. Arrows indicate 5'→3'. If the arrows point to the right, the primers contain the sense strand, if they point to the left the primers contain the antisense strand. Restriction enzyme sites are underlined. (see Laughon and Scott, 1984, for the *ftz* sequence).

a) Map of PCR primers



b) Sequence of the PCR primers

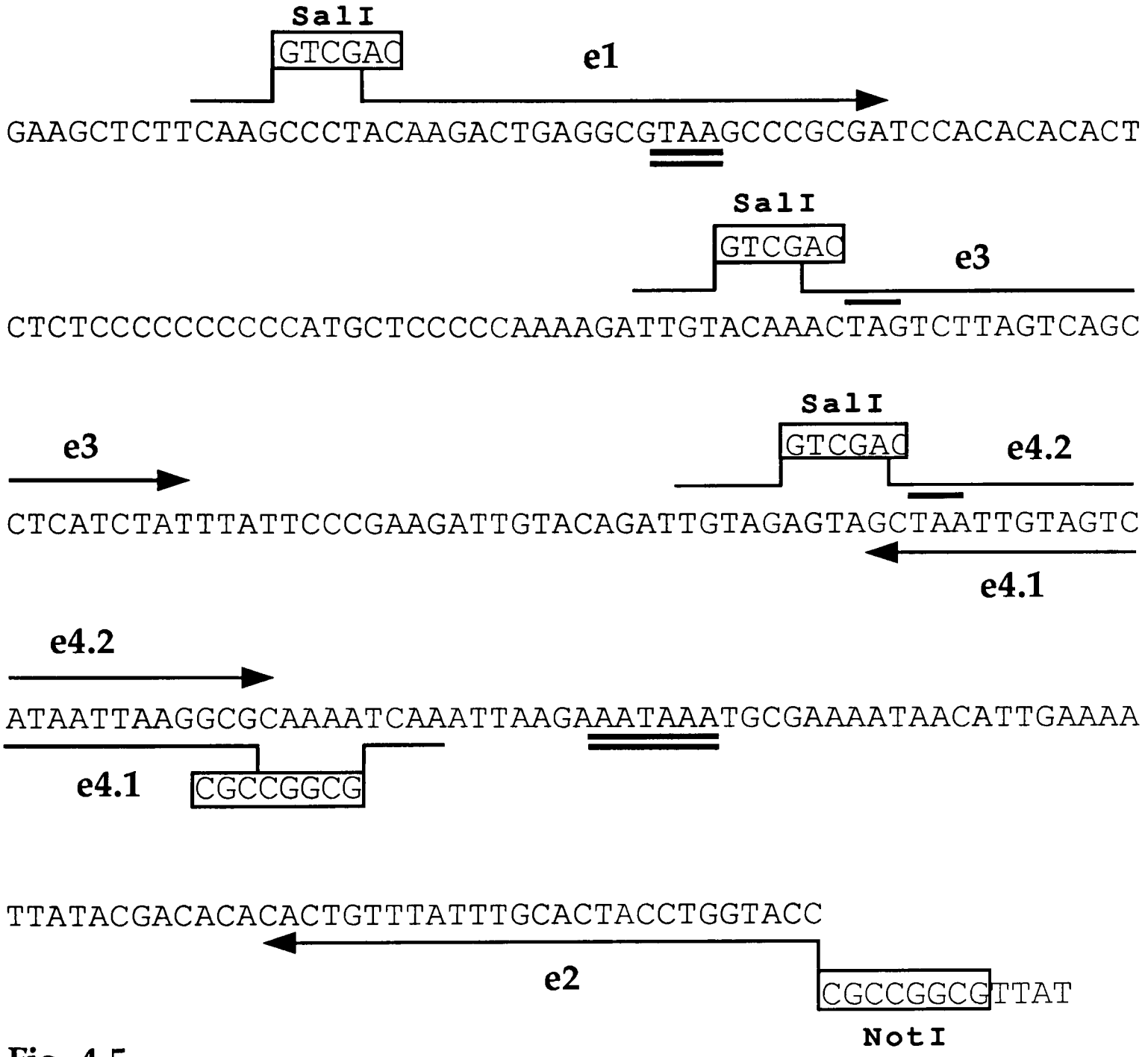


Fig. 4.5

A map of the PCR primers used to generate a deletion series of the 3' untranslated part of *eve*. Light lines with arrows represent parts of the primers which are homologous to *eve* sequences. Sequences in boxes are either SalI or NotI sites, bold lines indicate the polyadenylation signal and termination codons in the primers. Arrows indicate 5'→3'. If the arrows point to the right, the primers contain sense strand, if they point to the left the primers contain the antisense strand. Fig. a) is a map (not to scale) of the relative positions of the different primers. Fig. b) shows the sequence of the primers.

known to contain sequences necessary and sufficient for apical localisation. For this purpose, it was decided to use the SalI to HindIII fragment of *ftz* analysed in the previous chapter.

LTfSH consists of the SalI to HindIII fragment of *ftz* inserted into *LT*. Five independent lines were recovered, three on the second chromosome and two on the third (see Table 4.1). *In situ* hybridisation on two representative lines revealed that *lacZ* mRNA is apically localised (see Fig. 4.7). The result correspond well with *FG2* (chapter 3), showing that the SalI to HindIII fragment of *ftz* directs the apical localisation of *LTfSH* and *FG2*.

LTfSR contains the SalI to EcoRI fragment of *ftz* inserted into *LT*. Four independent lines were recovered, two on the second chromosome and two on the third (see Table 4.1). *In situ* hybridisation on two of these lines revealed that *lacZ* mRNA is localised in the basal periplasm (see Fig. 4.7).

LTf34 contains the SphI to HindIII fragment of *ftz* inserted into *LT*. Four independent lines were recovered, two on the second chromosome and two on the third (see Table 4.1). *In situ* hybridisation on two representative lines revealed that *lacZ* mRNA is unlocalised in a similar way to *ftz/lacC* (see Fig. 4.6). This result was surprising since *LTf34* contains globin sequences and was therefore expected to be basally localised. I have no satisfactory explanation for the result.

It follows from the analysis of the three constructs that the two regions SalI-RI and SphI-HindIII are not sufficient for apical localisation. The two constructs were made before the true position of the polyadenylation signal was known. Unfortunately, the polyadenylation signal turned out to lie in between the two fragments inserted into *LT*. In the previous chapter it was shown that the SalI-SphI fragment is sufficient for localisation. It therefore seems likely that the RI-SphI fragment contains essential sequences for localisation. However, further experiments are necessary to establish whether this region is sufficient for localisation (see Fig. 4.7).

Plasmid	no. of lines	II	III	X	localisation
LT	10	4	5	1	basal
LTfSH	5	3	2		apical
LTfSR	4	2	2		basal
LTf34	4	2	2		unlocalised
LTe12	2		2		unlocalised
LTe14	6	3	3		basal
LTe34	3		2	1	basal
LTe32	9	4	2	3	apical
LTe42	5	2	3		unlocalised
LThsp	2	1			unlocalised
			1		nuclear

Table 4.1

Hybrid constructs containing sequences from the 3' untranslated region of *ftz*, *eve* and *hsp70* inserted into *LT*. The number of independent transformed lines and their pattern of mRNA localisation is shown.

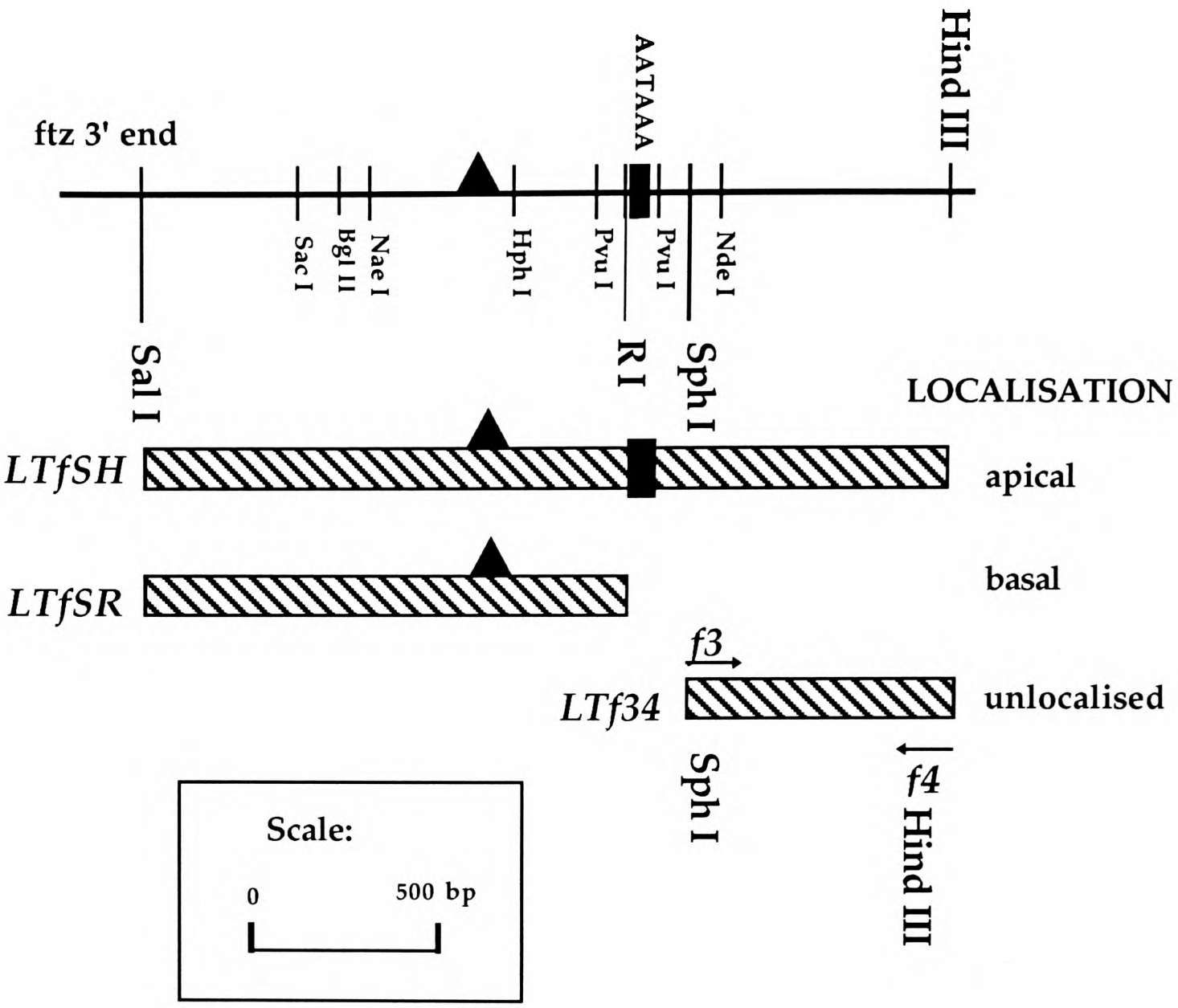
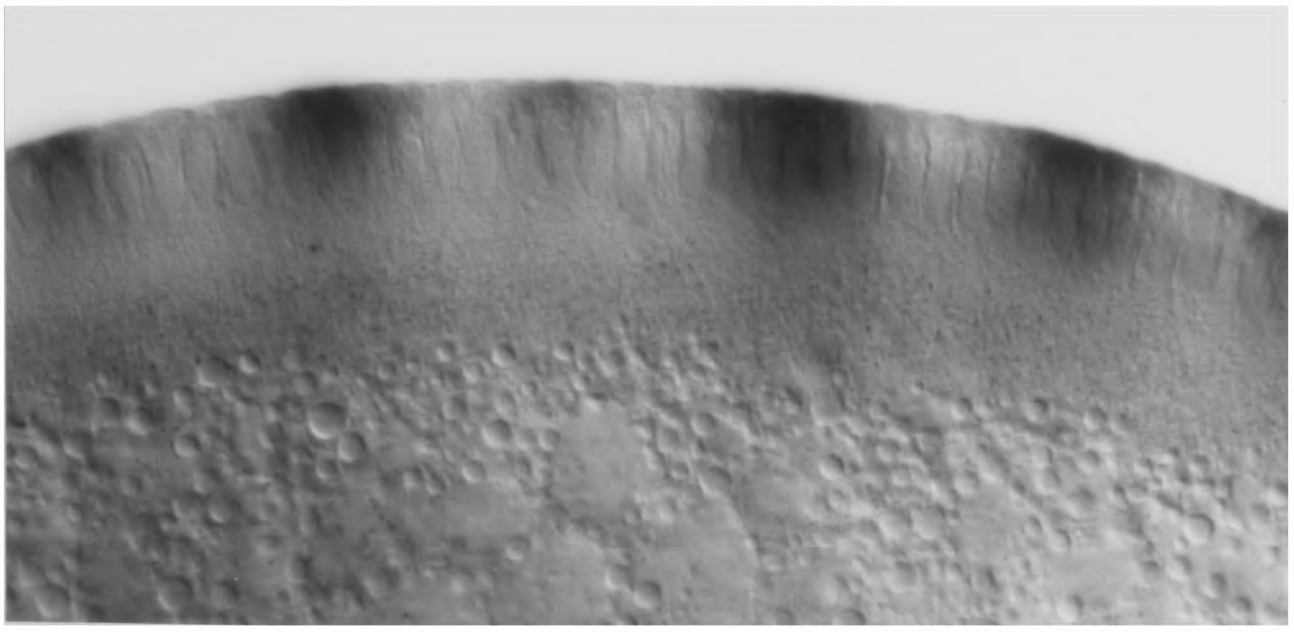


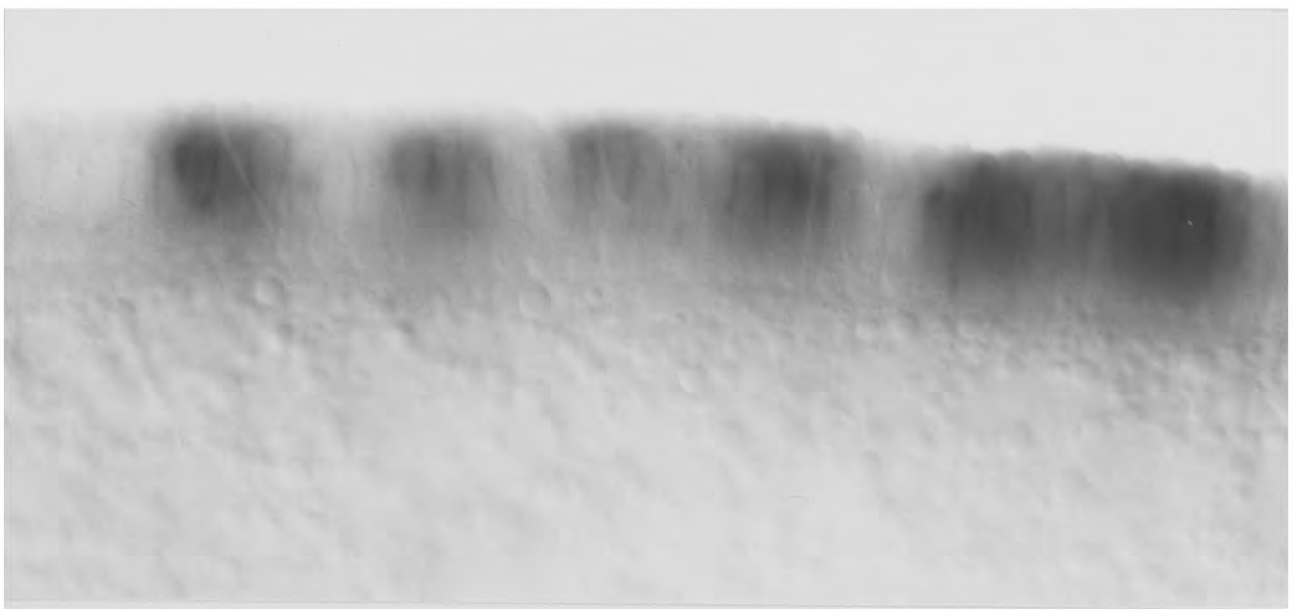
Fig 4.6

Fragments of the *ftz* 3' end inserted into *LT* in order to determine their effect on the localisation of LT. *f3* and *f4* are PCR primers used to produce a small region of ftz and insert it into *LT* (see Fig.4.5). The region sufficient for localisation is a 1.3kb SalI to SphI fragment.

a)



b)



c)

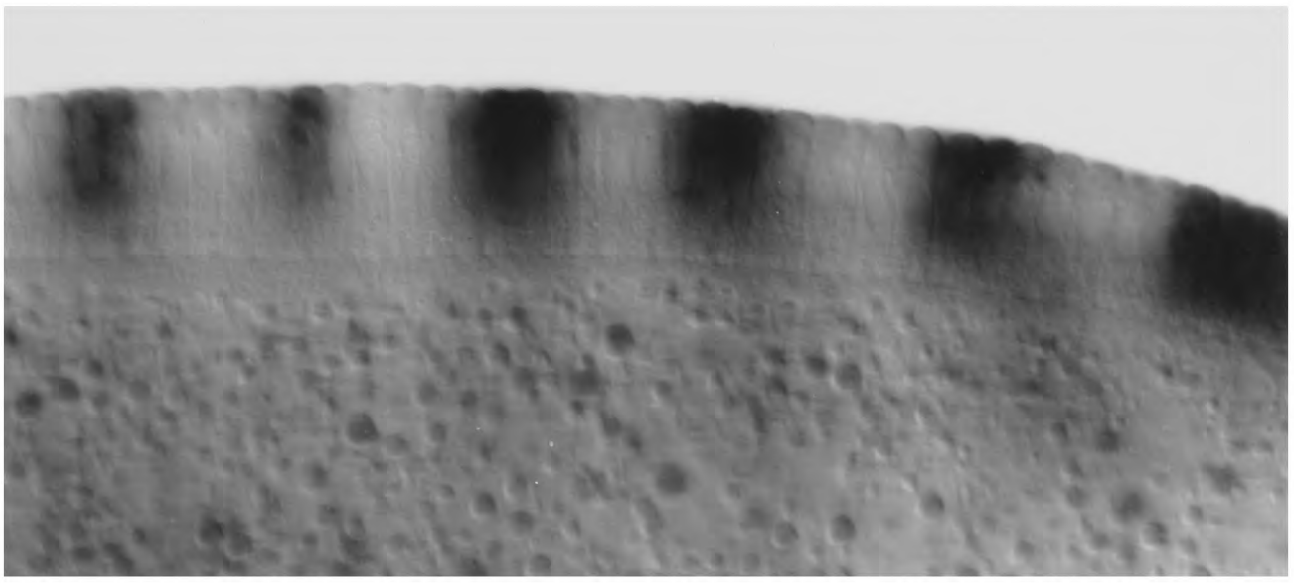


Fig 4.7

The intracellular localisation of the transcripts of several hybrid constructs containing parts of the *ftz* 3' untranslated region inserted into the *LT* vector. The results show that the 3' untranslated part of *ftz* is necessary and sufficient for localisation. Whole mount *in situ* hybridisation was performed using DIG-labelled *lacZ* probe, revealing the message localisation of embryos transformed with three of the constructs shown in Fig. 4.6. a) *LTfSH*, b) *LTfSR* and c) *LTf34*.

A 167bp region of *eve* contains apical localisation sequences

The short 3' untranslated part of *eve* was divided into five overlapping fragments by means of PCR primers so that their apical localisation potential could be tested by insertion into the *LT* vector (see Fig. 4.8).

LTe32 contains a 167bp fragment of the 3' untranslated including the polyadenylation signal and some downstream sequences (see Fig. 4.8). The fragment includes 130bp before the site of polyadenylation (defined in chapter 6) and the AATAAA signal is 102bp from the beginning of the fragment. Nine independent lines were recovered (after injecting *LTe32* DNA into embryos), four on the second chromosome, two on the third and three on the X chromosome (see Table 4.1). Whole-mount *in situ* hybridisation on two representative lines reveals that *LTe32* mRNA is apically localised. *e32* is therefore the smallest region so far defined as containing sequences necessary and sufficient for apical localisation.

e32 was divided into two halves (overlapping by 24bp) by means of the PCR primers, *e4.1* and *e4.2* (see Fig.4.8) to try to narrow down even further the sequences necessary for localisation.

LTe34 contains the 5' half of *e32* and *LTe42* contains the 3' half of *e32* inserted into the *LT* vector. Three independent *LTe34* lines were isolated, two on the second and one on the X chromosome. Five independent *LTe42* lines were isolated, two on the second and three on the third chromosome (see T able 4.1). Two lines of each construct were analysed by *in situ* hybridisation, revealing that *LTe34* mRNA is basally localised and that of *LTe42* is unlocalised. It follows that each overlapping half of *e32* is not sufficient on its own for apical localisation and that the entire *e32* fragment is required for apical localisation.

The *e42* fragment is 104bp in length and contains 67bp before the site of polyadenylation and includes all the sequences necessary for polyadenylation of *eve* (see chapter 6). It follows that the polyadenylation signal is not sufficient to enable apical localisation. However, it remains to be determined whether the localisation sequences can be separated from the polyadenylation signal.

The presence of the *eve* polyadenylation signal in *e42* means that *LTe42* transcripts lack *β-globin* sequences and are not basally localised. On the other hand, *e34* does not include any polyadenylation signal so that *LTe34* transcripts include the *β-globin* 3' end which causes their basal localisation.

Another hybrid construct, *LTe14*, contains a region of *eve* which overlaps with *e32* and also contains the rest of the 3' untranslated part of *eve* (see Fig. 4.8). Six independent lines were isolated after injection into embryos; three on the second and three on the third chromosome (see Table 4.1). *LTe14* transcripts are found in the basal periplasm in a similar manner to *LT* and *LTe34* mRNA (Fig. 4.9). This result supports those described above; *e14* is not sufficient for apical localisation and neither is *e42*. It can also be concluded that additional localisation signals are not present in the upstream half of *e14* that is not included in *e32*.

The last *eve* construct is *LTe12*, consisting of the entire 3' untranslated part of *eve* inserted into *LT*. Two independent lines were isolated after injection into embryos; both lines are on the third chromosome (see Table 4.1). *In situ* hybridisation reveals that in both lines *LTe12* mRNA is unlocalised in a similar manner to *LTe42*. This is surprising as *e12* includes all the sequences present in *e32* which are sufficient for apical localisation. It seems likely that the localisation signal is obscured in *e12* because of sequences present in the upstream part of the fragment. If the apical localisation signal operates on the level of RNA secondary structure then it is easy to imagine

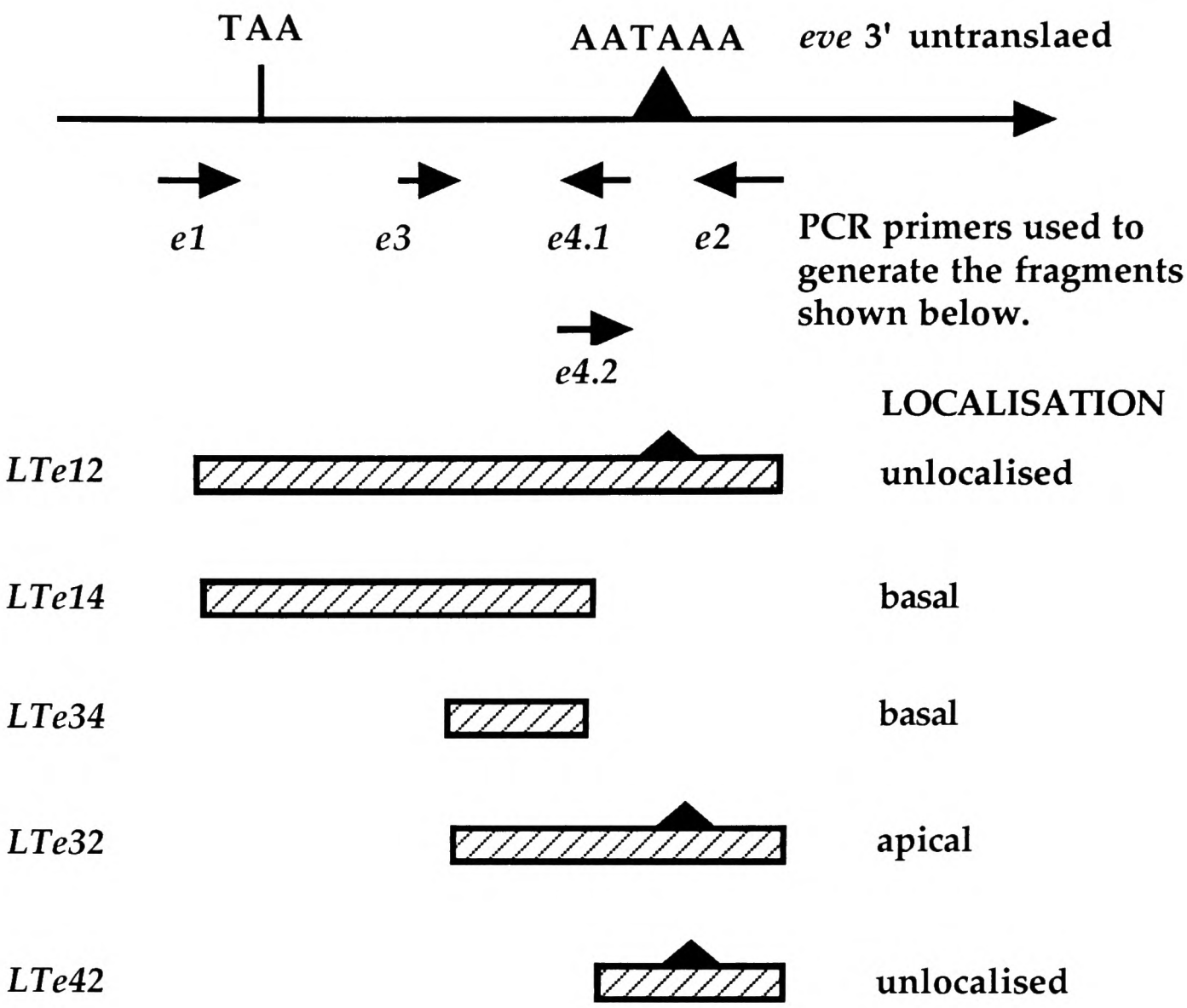
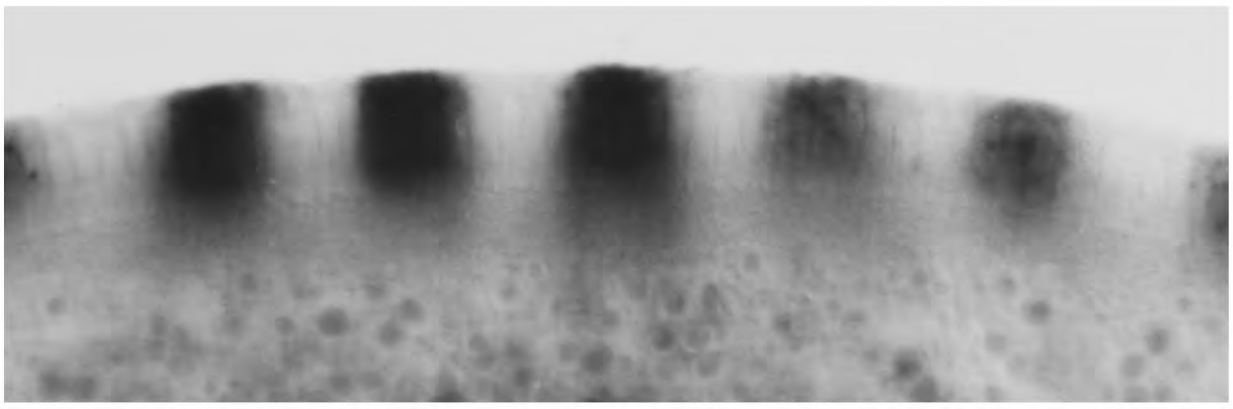


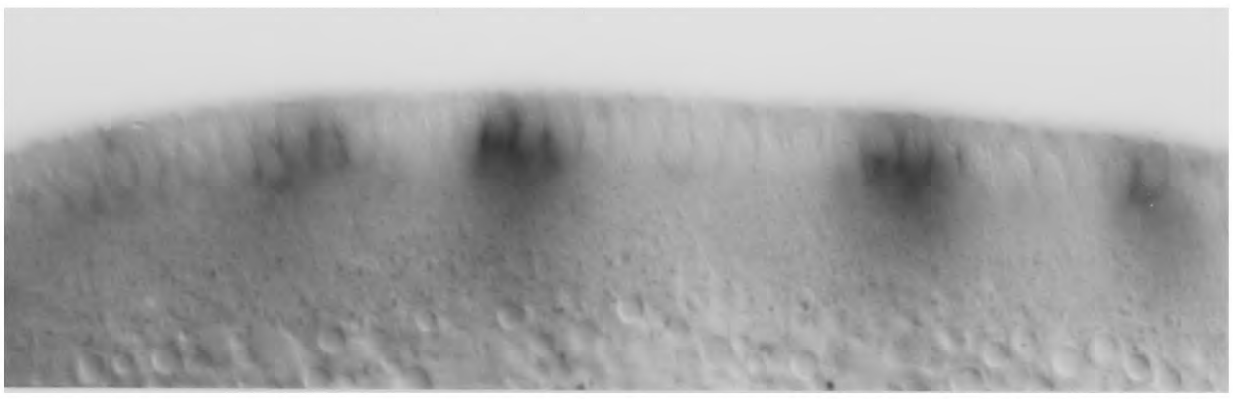
Fig 4.8
The PCR fragments of the *eve* 3' untranslated region inserted into *LT*
The primers used to generate these fragments are indicated as well as the pattern of localisation of the constructs containing these fragments. The position of the translation stop codon (TAA) and polyadenylation signal (AATAAA) are indicated.

Fig 4.9 (Opposite).
Whole-mount *in situ* hybridisation with DIG-labelled *lacZ* probe, revealing the localisation of *lacZ* transcript in embryos transformed with the constructs shown in Fig. 4.9. a) *LTe12*, b) *LTe14*, c) *LTe34*, d) *LTe32* and e) *LTe42*. The results show that a fragment (*e32*) of the 3' untranslated part of *eve* is necessary and sufficient for apical localisation. Two separate fragments (*e34* and *e42*) each containing half of *e32*, are not sufficient for apical localisation.

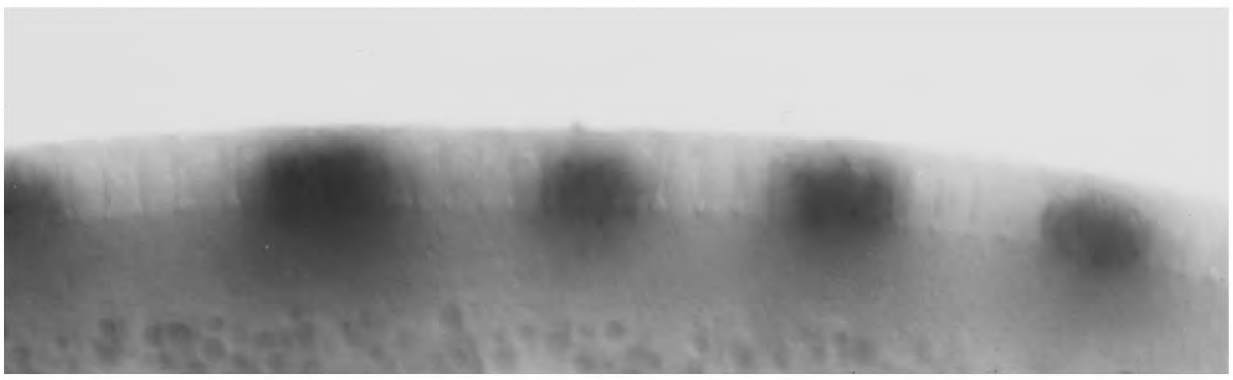
a)



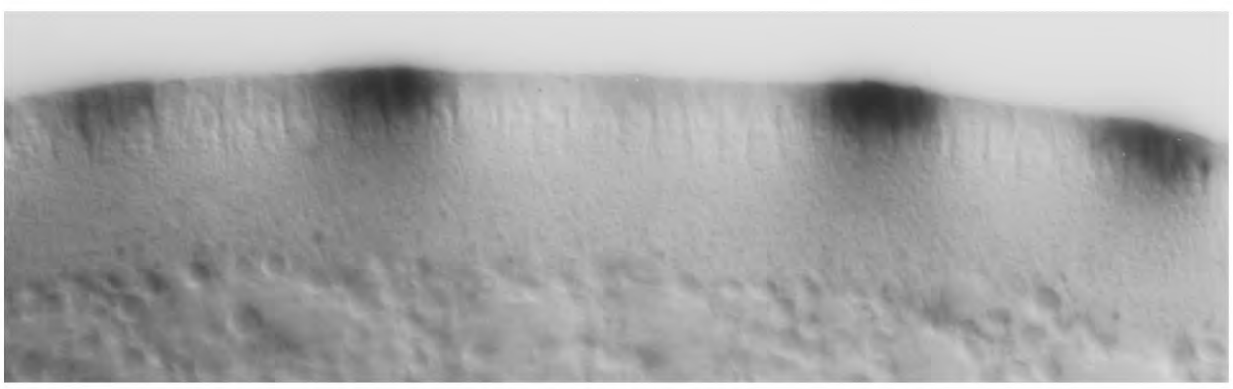
b)



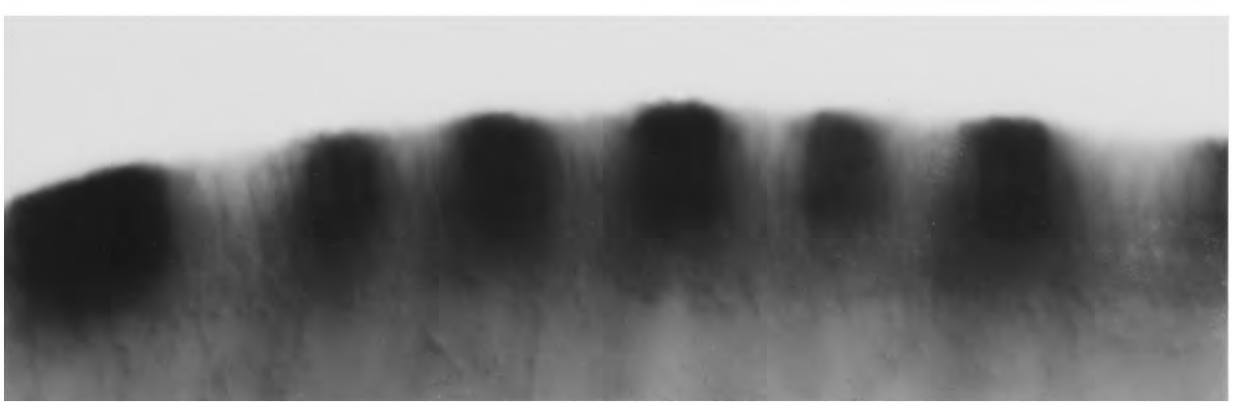
c)



d)



e)



that in combination with *lacZ* sequences, *e12* could fold into a secondary structure which inhibits the formation of the correct RNA folding needed for apical localisation. Alternatively, a mutation which disrupts the localisation signal could have been introduced during PCR synthesis of *e12*.

hsp70 lacks apical localisation sequences

In chapter 3 it was shown that the 3' untranslated end of the *hsp70* gene does not contain any apical localisation sequences. It was important to confirm this result in the *LT* vector to ensure that the results obtained in chapters 3 and 4 are consistent.

LThsp contains the 3' untranslated part of the *hsp70* gene inserted into the *LT* vector (the position of the PCR primers used to generate the fragment is shown in Fig. 4.11). Two independent lines were isolated after injection of the plasmid into embryos, one line on the second chromosome and one line on the third (see Table 4.1). *In situ* hybridisation reveals that one line is unlocalised as expected and that the other line has a very low level of expression which appears to be confined to the nucleus (see Fig. 4.10). The most likely explanation is that the plasmid has integrated in a region which limits its level of expression so that the unlocalised cytoplasmic transcript cannot be detected but very weak nascent RNA transcripts could still be detected (see chapter 7 for a full consideration of punctate nuclear staining). The relevant result is that *LThsp* mRNA is unlocalised in the blastoderm.

DISCUSSION

There are three patterns of transcript localisation shown by hybrid constructs and endogenous genes in the syncytial blastoderm. *LThsp, LTe42* and *LTe12* transcripts are unlocalised, *LTe32* and pair-rule transcripts are apically localised, and *LT, LTe34* and *LTe14* transcripts are basally localised.

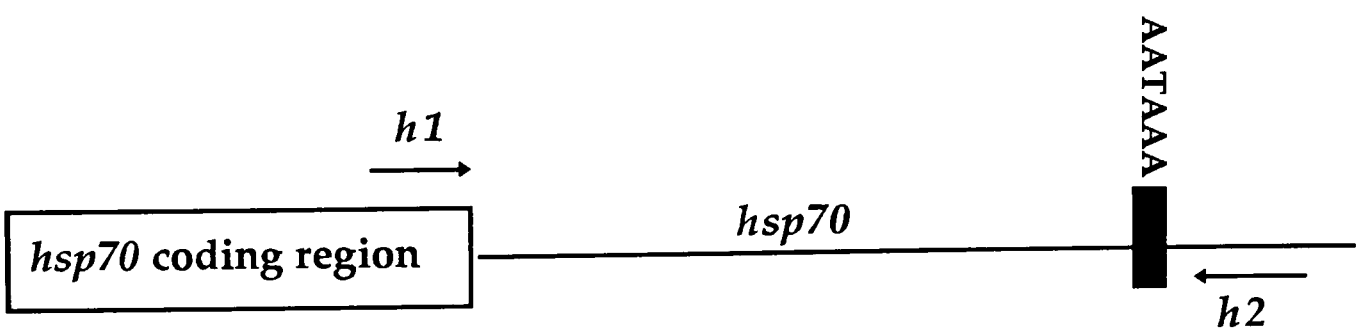
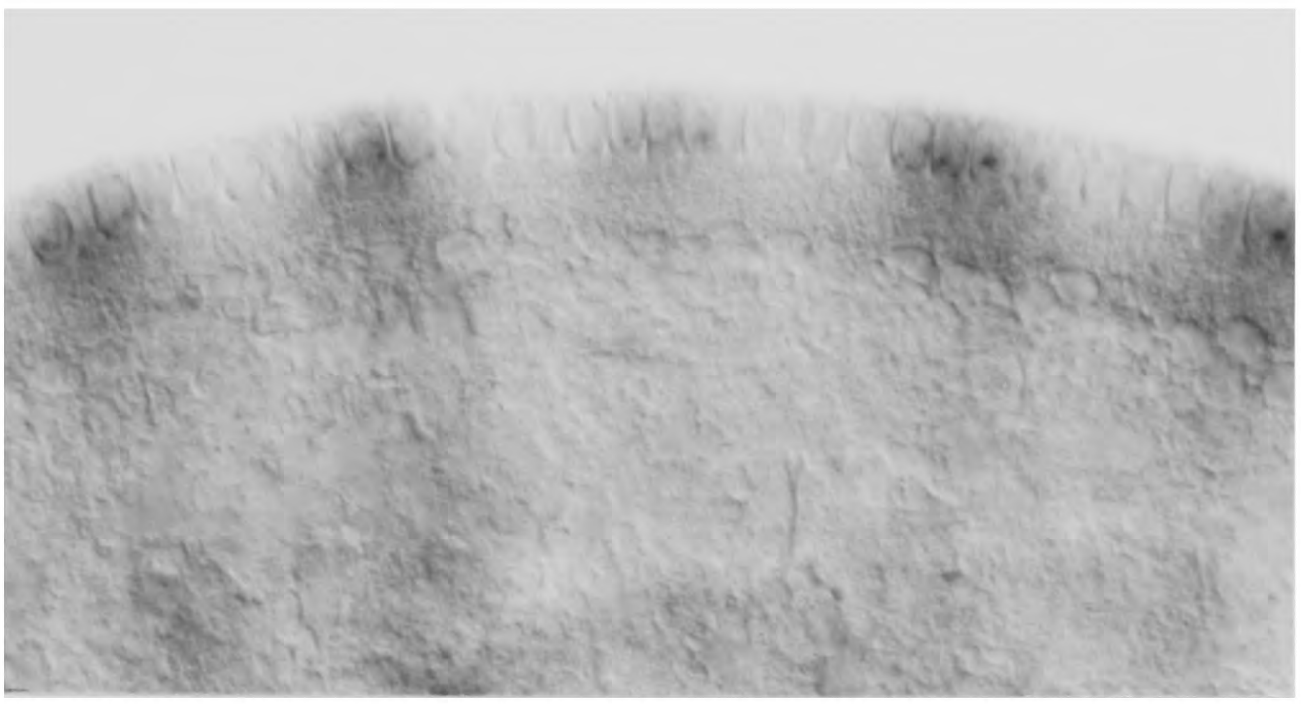


Fig. 4.11

A map (not to scale) of the PCR primers used to generate the 3' untranslated part of *hsp70*. Arrows indicate 5'→3'. If the arrows point to the right, the primers contain the sense strand, if they point to the left the primers contain the antisense strand.

a)



b)

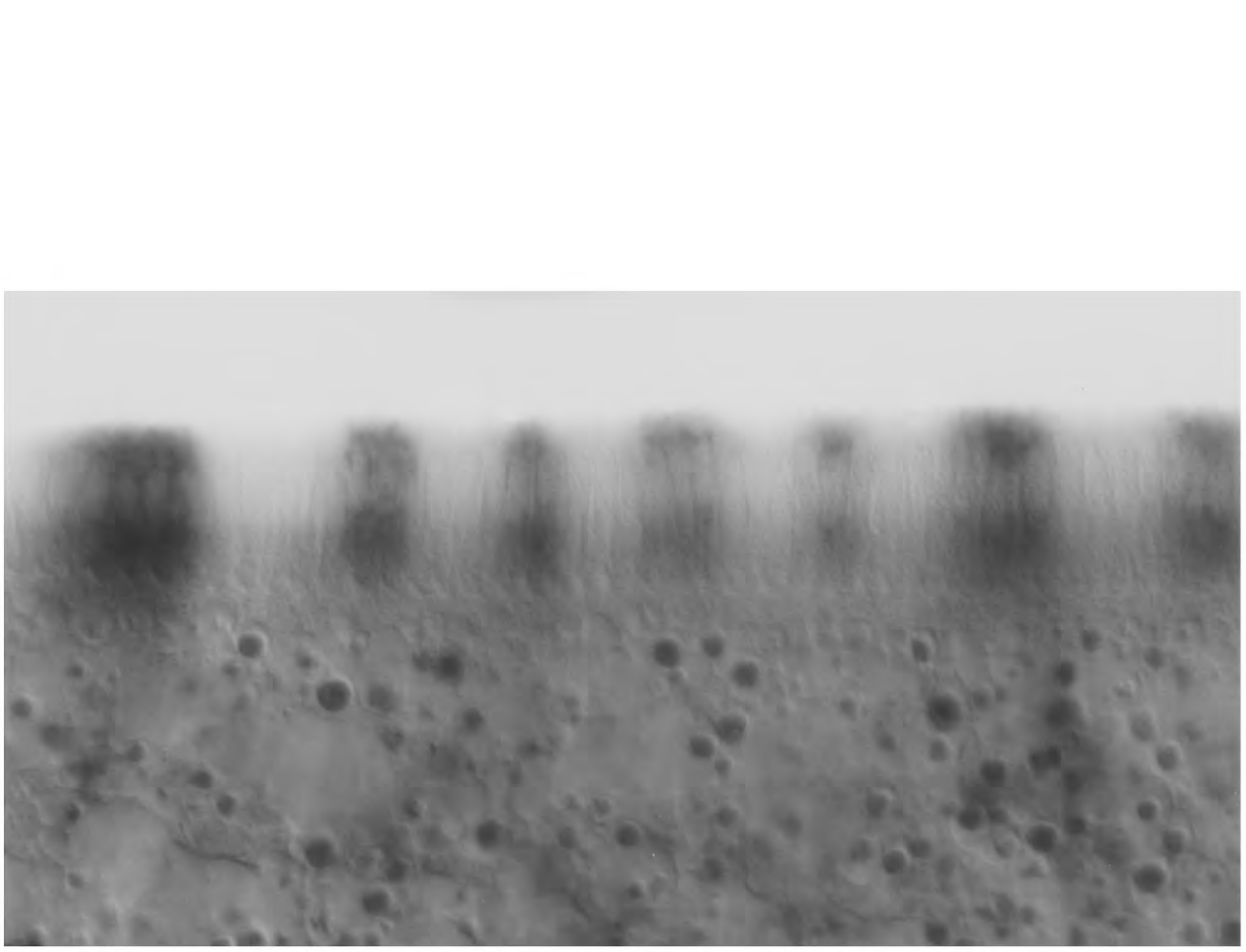


Fig. 4.10

The intracellular localisation of *LThsp* (the hsp-70 3' untranslated end inserted into *LT*), showing that like *ftz/lacC* and *h/lacZ/hsp-27, LThsp* is unlocalised . Two lines are described which give different results. a) Line A shows weak punctate message in the nucleus. b) Line B shows strong unlocalised message.

The most likely interpretation is that an unlocalised distribution of mRNA is the default state found when a transcript lacks any localisation signals. *LT* contains a basal localisation signal in its *β-globin* 3' end which direct its transcript to the basal periplasm. *e32* includes the *eve* apical localisation signal and polyadenylation signal; so the *LTe32* transcript lacks *β-globin* sequences and is apically localised.

The apical localisation sequences present in pair-rule mRNA could interact in a variety of ways with other macromolecules in the blastoderm to achieve apical localisation of pair-rule transcripts. The localisation mechanism may involve more than one distinct step, as is the case for *bcd* localisation in the oocyte (St Johnston *et al.*, 1989). These steps could involve any of the stages of mRNA transport in cells including the site of transcription in the nucleus, transport in the nucleoplasm, export across the nuclear membrane and transport in the cytoplasm. *A priori*, it is difficult to predict the nature of the interaction with other molecules and the mechanism for localisation.

In general the localisation signal could interact with proteins such as the cytoskeletal or with other RNA molecules. The sequence required for localisation may be small and could operate on the level of primary base pair sequence. Alternatively the localisation signal could be quite large and function on the level of RNA secondary structure. The data presented above suggests that localisation cannot be narrowed to a region smaller than approximately 100bp and is therefore likely to act via RNA secondary structure (see chapter 9 and appendix B).

It is interesting to note that both the *eve* localisation sequence and the putative region of *ftz* necessary for localisation contain the polyadenylation signal. *LTe42* contains an *eve* polyadenylation signal but is unlocalised. It is therefore clear that the polyadenylation signal is not sufficient for apical

localisation. However, it remains to be seen whether the polyadenylation signal includes sequences which are also required for apical localisation (see chapter 6).

In chapter 1, a number of alternative models for apical localisation were proposed. The models included processes in the nucleus and the cytoplasm. In the next chapter, the possibility of a cytoplasmic mechanism for apical localisation is investigated experimentally. It is concluded that the mechanism of localisation involves a nuclear process.

CHAPTER 5: APICAL LOCALISATION DOES NOT INVOLVE A CYTOPLASMIC MECHANISM

INTRODUCTION:

At the end of chapter 1, I outlined a number of possible models for the mechanism of apical localisation. One class of these models involves cytoplasmic mechanisms which alter the distribution of transcripts. In this chapter the merits of such cytoplasmic mechanisms are considered, in particular the possibility that some components needed in the localisation of *bcd* mRNA to the anterior pole are shared in common with components involved in the apical localisation of pair-rule mRNA.

There are a number of cytoplasmic mechanisms which could account for apical localisation. The first involves selective mRNA instability in the basal periplasm, so that pair-rule mRNA remains stable in the apical periplasm. The second involves random diffusion in the cytoplasm and selective binding to factors in the apical periplasm. The third mechanism invokes active cytoplasmic transport from the basal periplasm to the apical periplasm (see chapter 1, Fig. 1.5). Let us consider what evidence supports each of the three proposed cytoplasmic mechanisms.

When cycloheximide is injected into blastoderm embryos, protein synthesis is arrested and pair-rule transcripts are stabilised. The normal half-life of *ftz* mRNA in the syncytial blastoderm is approximately 5min (Edgar *et al.*, 1986). Under such circumstances *ftz* continues to be transcribed and accumulates to much higher levels than in wild-type embryos. The high level of transcript which accumulates is exclusively apical in its localisation (Edgar *et al.*, 1987). For differential mRNA stability to account for apical localisation, pair-rule mRNA would have to be substantially less stable (than the 5min half-life) in the basal periplasm, a possibility which seems

very remote. It follows that differential instability in the basal periplasm is unlikely to be the mechanism for apical localisation.

The other two possible cytoplasmic mechanisms involve the movement of mRNA from the basal to the apical periplasm. In the 14th interphase, the blastoderm nuclei are closely packed in the peripheral periplasm, with very little cytoplasm between them. It is difficult to imagine how an extensive cytoplasmic transport system could be located in the small cytosolic space between the nuclei. Furthermore, experiments described in chapter 8 suggest that the apical and basal periplasms are separated by a diffusion barrier. RNA is probably bound to the cytoskeleton so its diffusion in the blastoderm cytoplasm is extremely slow (see appendix A). Therefore, random diffusion and binding in the apical periplasm is unlikely to be the mechanism for localising a very unstable transcript.

While apical localisation is unlikely to involve a cytoplasmic mechanism, cytoplasmic transport of mRNA has been well documented in other systems. Probably the best understood case is the localisation of *bcd* mRNA in *Drosophila*. *bcd* mRNA is found localised in a tight crescent at the anterior tip of the blastoderm embryo, a localisation which originates during oogenesis in the mother and involves four distinct stages. In the first stage, *bcd* mRNA is found localised in a ring in the stage 6 oocyte. In the second stage, mRNA is also localised in the apical cytoplasm of each nurse cell. In the third stage, the mRNA becomes localised to the anterior cortex of the oocyte as the nurse cells contract. This process is thought to involve the transport of *bcd* mRNA from the nurse cells to the oocyte. During the fourth phase of localisation the mRNA becomes localised to a broad region at the dorsal side of the anterior pole (St Johnston *et al.*, 1989).

bcd localisation is disrupted by mutations in 3 maternal genes *exuperantia (exu), swallow (sww)* and *staufen(stau)* (Berleth *et al.*, 1988; Frohnhöfer and

Nüsslein-Volhard, 1987). *sww* and *stau* disrupt the localisation process at stages three and four respectively. The disruption caused in *exu* oocytes is a little more complicated; stages 2, 3 and 4 are completely disrupted but stage 1 is partially disrupted. It is suggested that the available *exu* alleles are only partial loss of function mutations and a true null would cause complete disruption at stage 1 of *bcd* localisation. It follows that stages 2, 3 and 4 are disrupted because they are contingent on stage 1 (St Johnston *et al.*, 1989). It is likely that the *exu* product is required for stage 1 of *bcd* localisation and that stage two of the *bcd* localisation does not require the products of *exu*, *sww* or *stau*.

A priori, it appears unlikely that *bcd* mRNA, which is maternal, could be localised by the same mechanism as pair-rule mRNAs which are zygotically transcribed. While the localisation of *bcd* mRNA is a long range process in the oocyte, pair-rule mRNA is localised in the syncytial blastoderm stage and is a short range phenomenon. However, it is possible that the pair-rule genes are localised at least in part by similar mechanisms to *bcd* localisation so that the two processes share the components used in one or more of the stages of *bcd* localisation. Alternatively, the two processes might share similar cytoplasmic components which are involved in different mechanisms of localisation.

This chapter investigates the possibility of a mechanistic connection between *bcd* localisation and the apical localisation of pair-rule transcripts.

STRATEGY

The *bcd* 3' untranslated region is implicated at least in some stages of *bcd* localisation in the oocyte (Macdonald and Struhl, 1988). The entire 3' untranslated part of *bcd* was inserted into the *LT* vector (forming *LTbcd*), in order to test whether sequences responsible for *bcd* localisation could also

include apical localisation sequences. The intracellular distributions of *LTbcd*, *ftz*, *h* and *eve* mRNAs were examined in *sww*, *exu*, *stau* embryos to determine if *LTbcd* and pair-rule transcript localisations depend on similar processes to those that localise endogenous *bcd* mRNA to the anterior pole.

RESULTS

Pair-rule mRNA is apically localised when first detected

In situ hybridisation using radioactive probes on wax tissue sections has revealed that pair-rule mRNA is apically localised when the genes are first detected (Ingham and Gergen, 1988; Macdonald *et al.*, 1986). Whole-mount *in situ* hybridisation using *ftz*, *h* and *eve* DIG-labelled probes confirmed these results (see Fig. 5.2).

If apical localisation involves a cytoplasmic mechanism, pair-rule transcripts would be expected to be initially unlocalised when they are first transcribed. A subsequent cytoplasmic mechanism would lead to the redistribution of basal mRNA to the apical periplasm. The fact that pair-rule mRNA is apically localised when they are first detected argues against such a cytoplasmic mechanism.

A possible mechanistic connection between *bcd* and pair-rule mRNA localisation.

LTbcd was constructed by inserting the entire 3' untranslated part of *bcd* into the *LT* vector. The *bcd* 3' end was synthesised using PCR primers with SalI and NotI sites so that it included the polyadenylation signal as well as the 630bp fragment known to be required for *bcd* localisation to the anterior pole (Fig. 5.1). The construct was injected into embryos and four independent lines were isolated: two on the second chromosome and two on the third. Two representative lines were analysed by whole-mount *in*

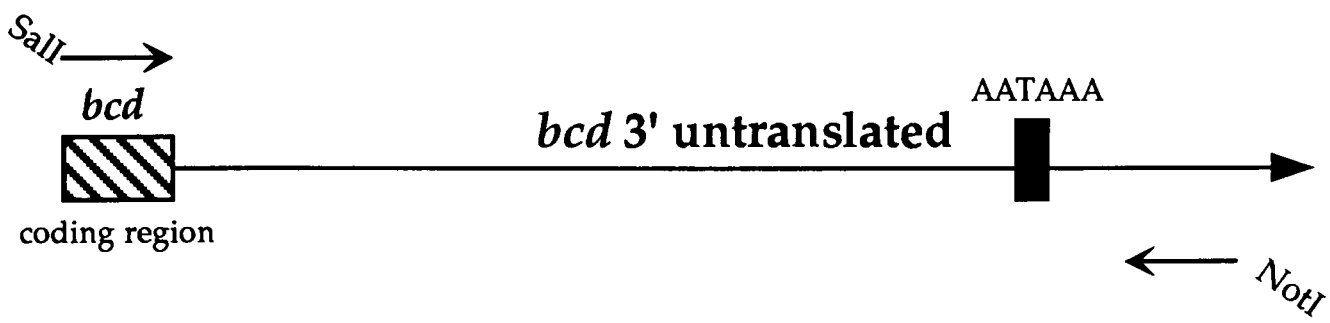
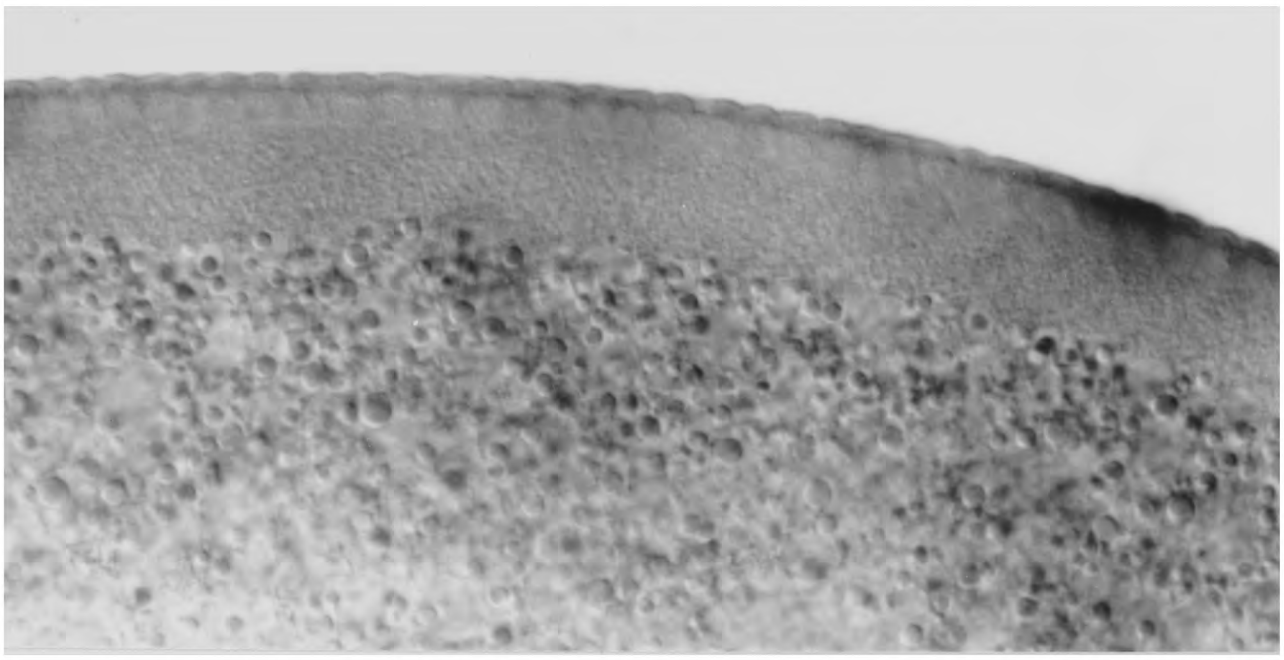


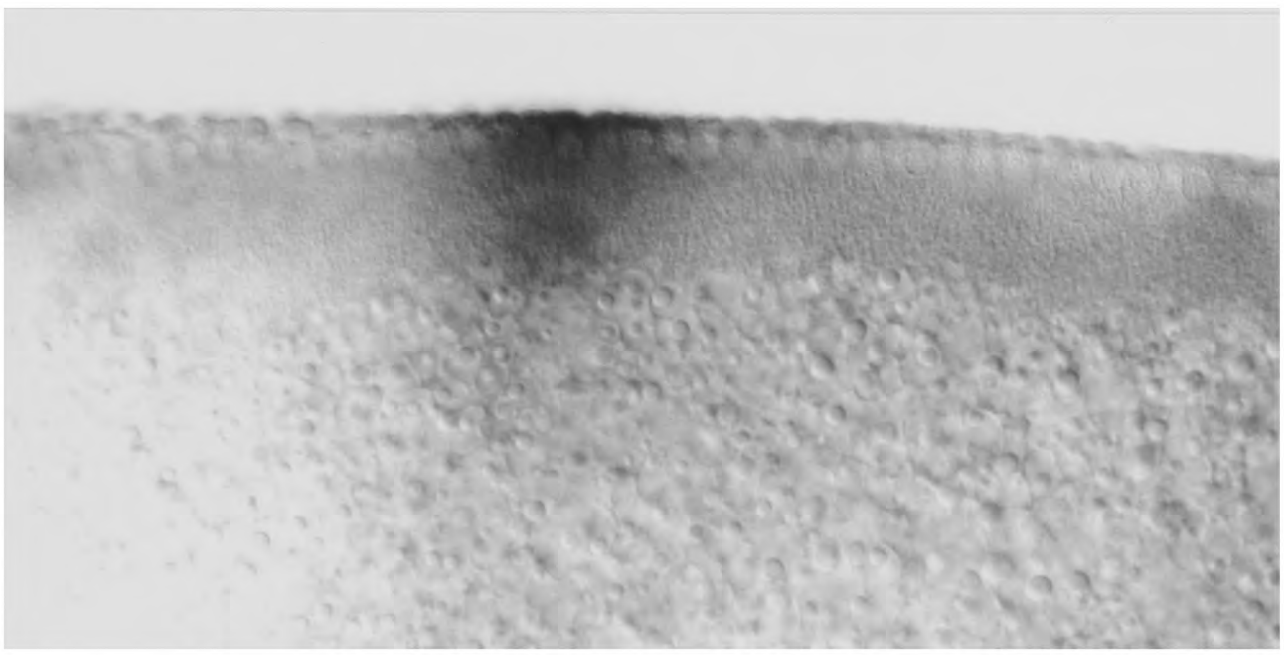
Fig. 5.1

The 3' untranslated part of *bcd* synthesised as a SalI to NotI fragment using PCR primers and inserted into the SalI and NotI sites of *LT*.

a)



b)



c)

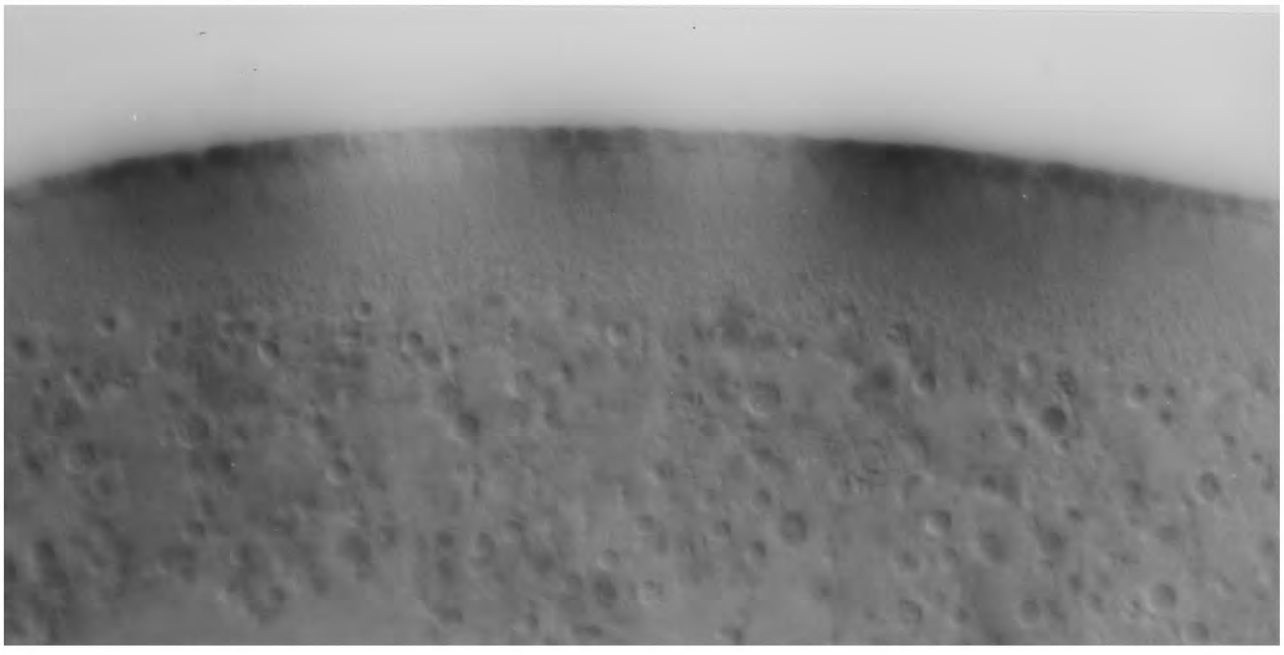


Fig 5.2:

Apically localised pair-rule transcripts in stage 13/14 wild-type blastoderm embryos. The results show that pair-rule transcripts are found exclusively in the apical periplasm when they are first detected. Whole mount *in situ* hybridisation with DIG-labelled probes: a) *ftz,* b) *eve* and c) *h*.

situ hybridisation revealing that *LTbcd* mRNA is found in the apical periplasm (see Fig. 5.3). It can be concluded that the 3' untranslated part of *bcd* includes sequences which are sufficient for apical localisation.

To test if the apical localisation of *LTbcd* involves similar processes to pair-rule transcript localisation, homozygous *LTbcd* males were mated to virgin *sww* or *sta exu* females. The progeny of the cross were analysed using a *lacZ* DIG-labelled probe. *LTbcd* mRNA was found to be apically localised to the apical periplasm in both mutant backgrounds (see Fig. 5.3), indicating that the apical localisation of *LTbcd* mRNA is not due to processes mediated by the products of the genes *exu*, *sww* or *sta*.

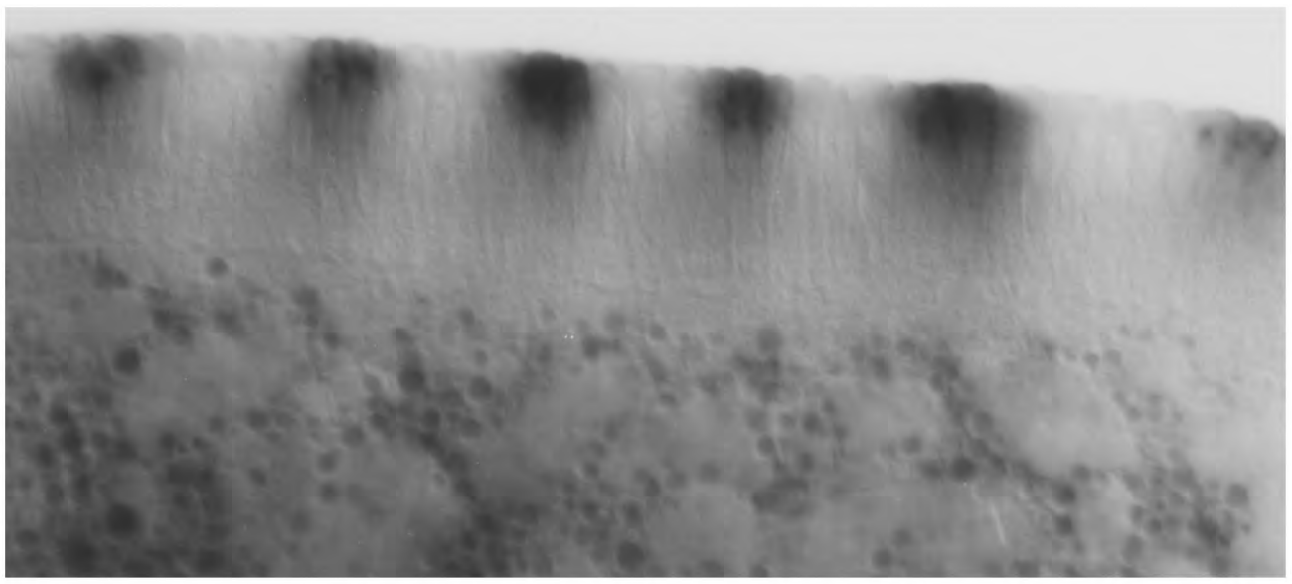
The distribution of pair-rule transcripts was examined by *in situ* hybridisation using *ftz, eve* and *h* probes in *sww* and *sta exu* mutant embryos. The results show that pair-rule transcripts are apically localised in the two mutant embryos (see Fig. 5.4 and 5.5). It follows that the gene products of the three maternal effect genes *exu, stau* and *sww* are not needed for endogenous pair-rule mRNA localisation. It is also likely that *LTbcd* and the pair-rule genes contain similar apical localisation sequences.

DISCUSSION

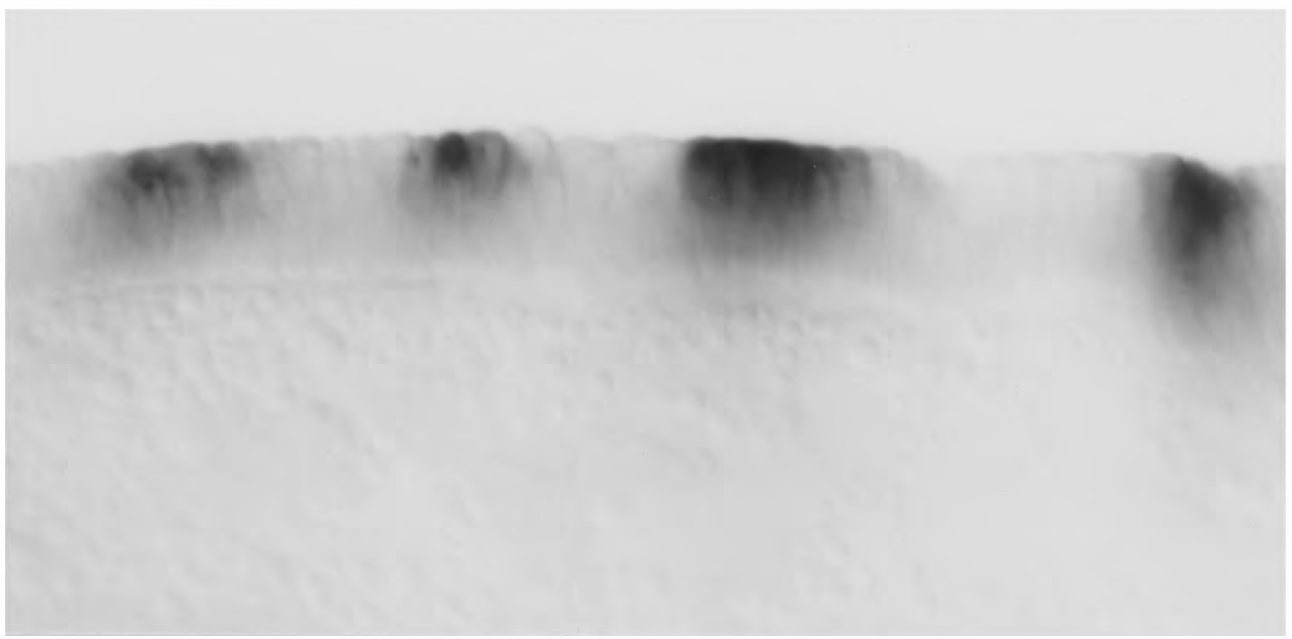
In the second stage of *bcd* localisation in the oocyte, mRNA is found localised on the apical side of the nurse cell nuclei in a manner resembling the apical localisation of pair-rule transcripts in the blastoderm. The results show that *LTbcd* is apically localised in wild-type and in *exu*, *sww* and *stau* embryos (see Fig. 5.3). It is possible that pair-rule transcript localisation is achieved, at least in part, by a mechanism which is shared with the second step of *bcd* mRNA localisation.

LTbcd mRNA is apically localised, but it has been shown that *bcd* mRNA is unlocalised at the syncytial blastoderm, being present in both the apical and

a)



b)



c)

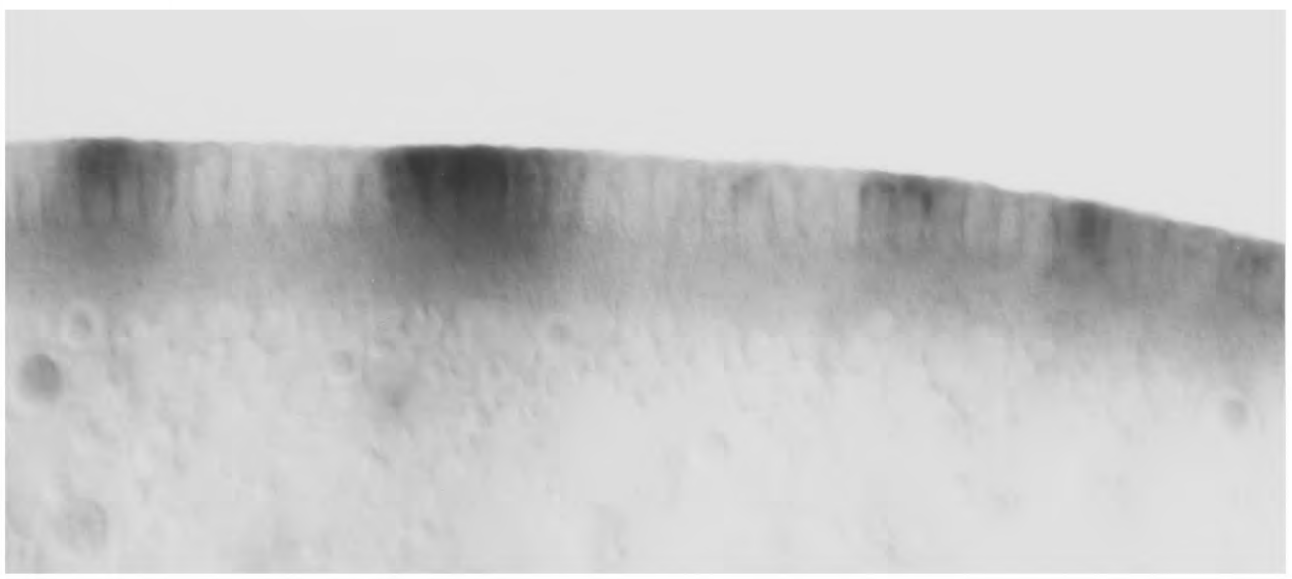
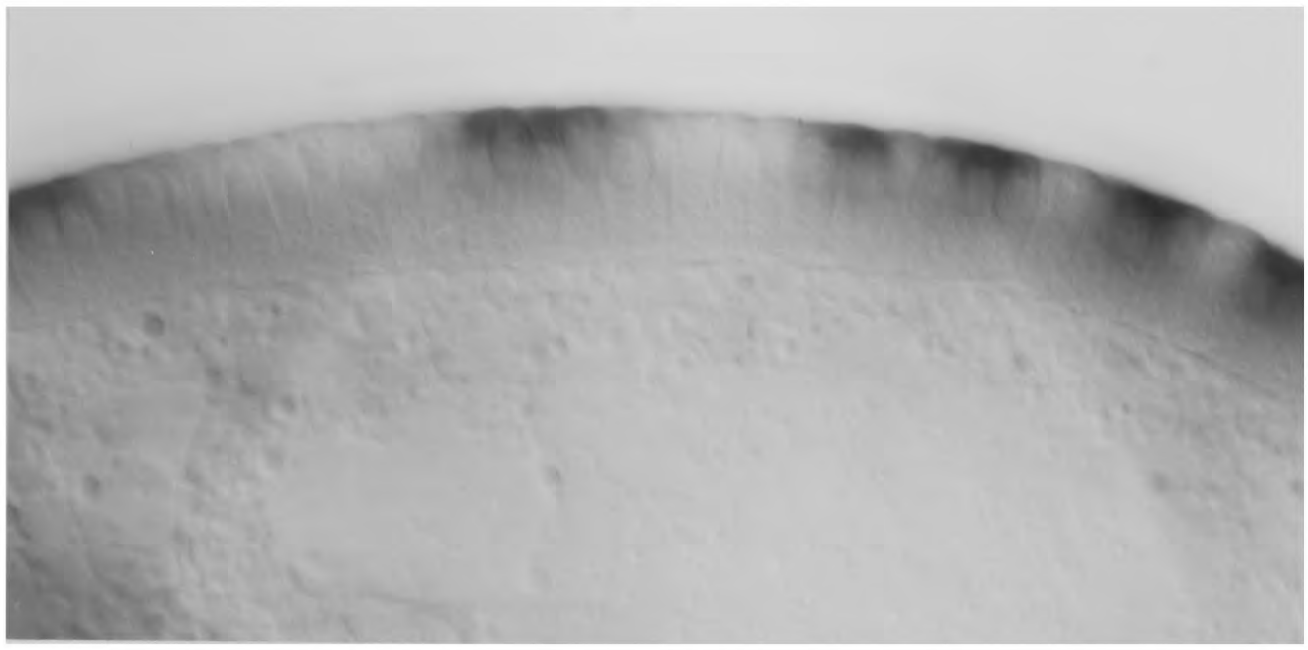


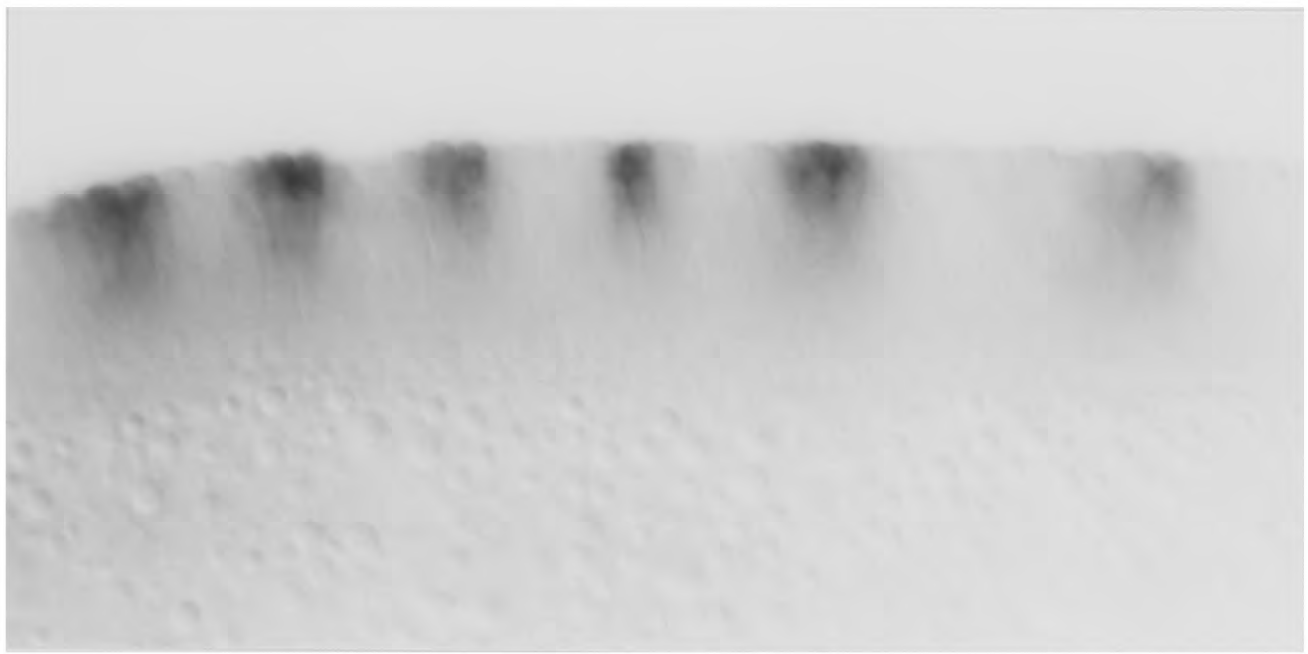
Fig 5.3:

***In situ* hybridisation with a DIG-labelled *lacZ* probe, revealing the intracellular localisation of *LTbcd* mRNA. a) *LTbcd* in wild type, b) *LTbcd* in *swa* embryos, c) *LTbcd* in *sta exu* embryos . The results show that maternal effect mutations which affect *bcd* localisation do not affect the localisation of *LTbcd* mRNA.**

a)



b)



c)

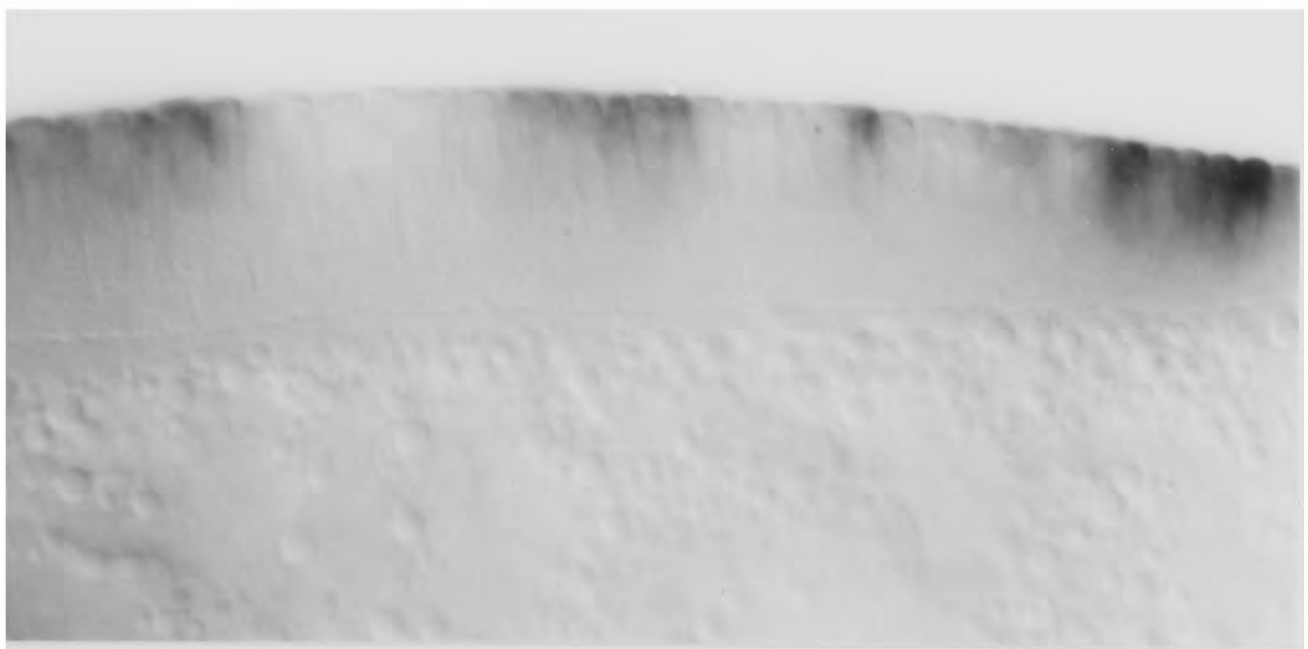
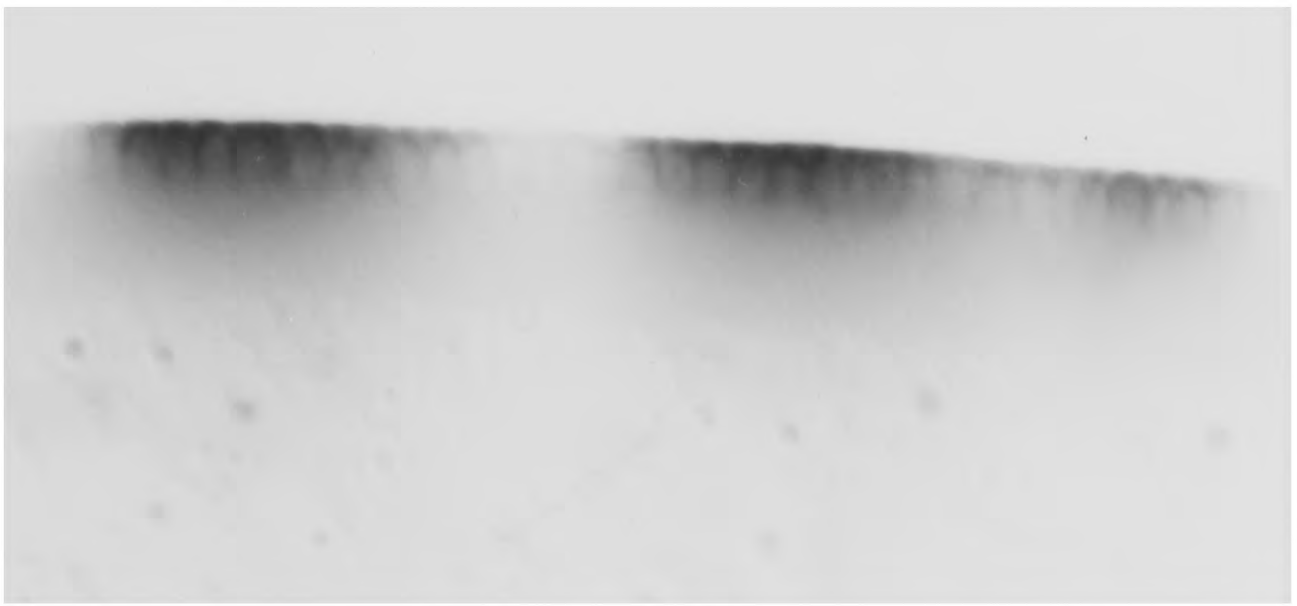


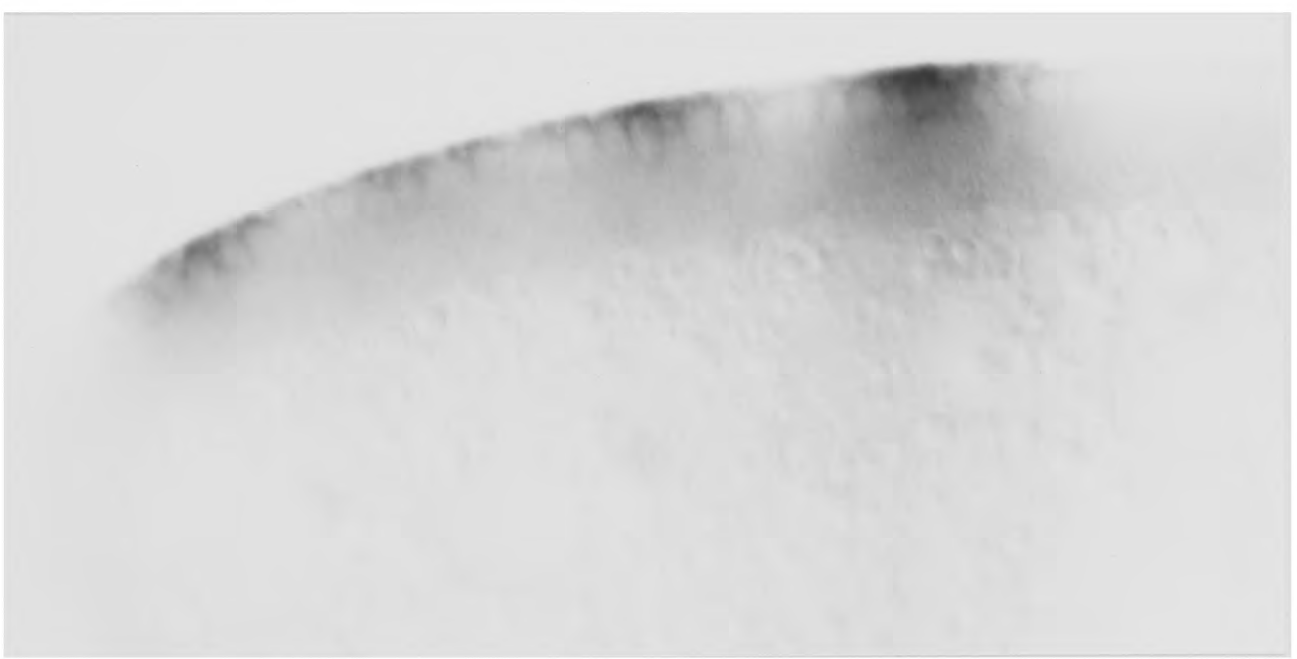
Fig 5.4:

***In situ* hybridisation using DIG-labelled *lacZ* probe revealing the intracellular localisation of a) *ftz*, b) *eve* and c) *h* mRNA in *swa* embryos in which *bcd* mRNA localisation is disrupted. The results show that pair-rule transcripts remain localised in *swa* embryos.**

a)



b)



c)

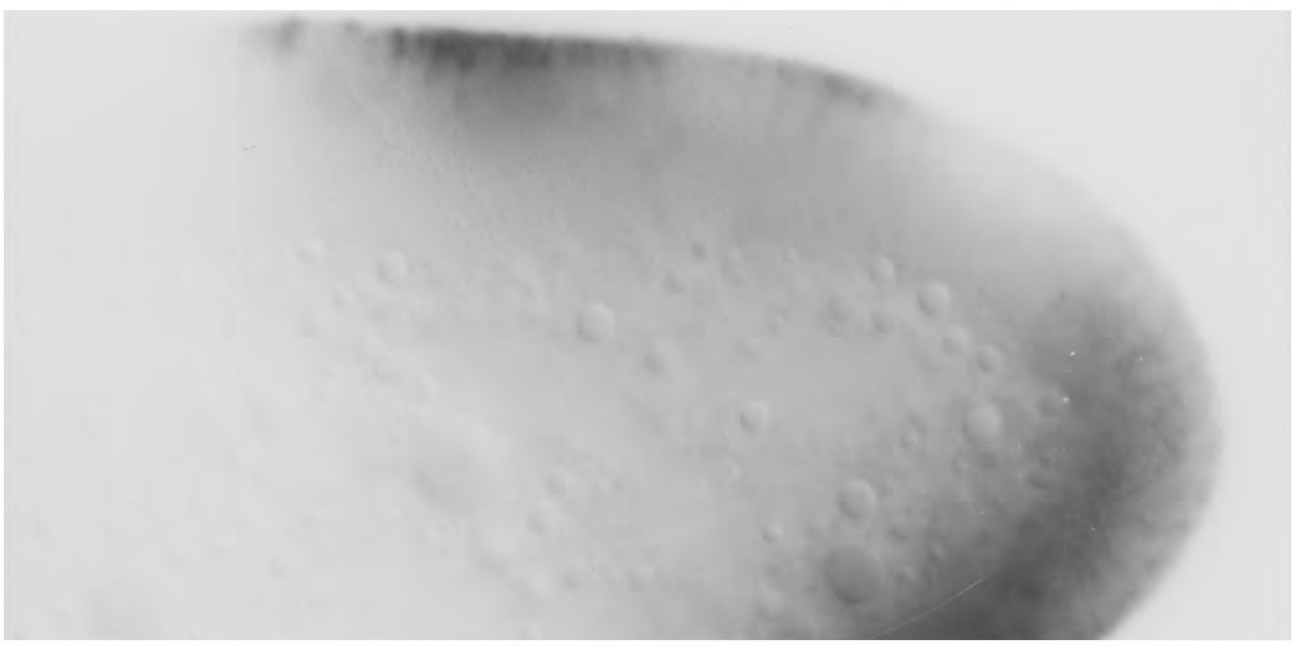


Fig 5.5:

***In situ* hybridisation using DIG-labelled probes, revealing the intracellular localisation of a) *ftz*, b) *eve* and c) *h* mRNA in *sta exu* embryos in which *bcd* localisation is disrupted. The results show that pair-rule transcripts remain localised in *sta exu* embryos.**

basal periplasm in the anterior pole of the egg (Berleth *et al.*, 1988; St Johnston *et al.*, 1989). Why do the apical localisation sequences found in the 3' end of *bcd* not cause apical localisation of wild-type *bcd* mRNA? It could be argued that the when *bcd* transcript is transported from the nurse cells to the oocyte, it binds to some component in the periplasm, thus overriding and preventing apical localisation. However, this cannot be the case since *LTbcd* mRNA contains all the sequences required for such binding and yet *LTbcd* is apically localised.

The most likely explanation for the fact that *bcd* mRNA is unlocalised despite the presence of apical localisation sequences in its 3' end, is that the process of apical localisation can only happen when an mRNA is zygotically transcribed. *LTbcd* is zygotically transcribed, thus enabling the apical localisation of its transcript. *bcd* mRNA is maternally transcribed, so it is apically localised in the nurse cells but cannot become apically localised in the blastoderm embryo. It follows that the mechanism of apical localisation involves processes which take place in the nucleus rather than the cytoplasm.

In the next two chapters the involvement of nuclear mechanisms in apical pair-rule localisation is considered.

CHAPTER 6: IS THE POLYADENYLATION SIGNAL REQUIRED FOR APICAL LOCALISATION ?

INTRODUCTION

In the previous chapter, it was concluded that the mechanism for localisation is unlikely to involve a cytoplasmic process. Indeed it was suggested that localisation involves a nuclear mechanism that leads to the export of pair-rule mRNA to the apical side of the nucleus.

All the apical localisation sequences defined in the previous chapters include a polyadenylation signal. However, it was shown that the polyadenylation signal is not sufficient for localisation and other sequences in the 3' end of the gene are necessary for localisation (see chapter 4). In this chapter, I consider whether the sequences required for apical localisation are separable from the polyadenylation signal.

The addition of a poly(A) tail to a transcript involves two distinct steps: the cleavage of the RNA at a specific site 15-30bp downstream of the AAUAAA signal and the addition of 100-200 adenosine residues to the 3' end (reviewed in Humphrey and Proudfoot, 1988). In addition to the AAUAAA signal, GU and U-rich regions must be present 20-25bp downstream of the AAUAAA. The quality of the downstream sequences and the distance between the two signals determine the efficiency of polyadenylation, but the RNA sequences between the two elements are unimportant (Levitt *et al.*, 1989).

Only the polyadenylation site of *h* has been determined precisely by cDNA cloning; the *h* poly(A) tail is situated 24bp downstream of an AATAAA polyadenylation signal (Rushlow *et al.*, 1989). The polyadenylation signal of *eve* has been determined by RNA protection assays (Macdonald *et al.*, 1986) and potential *ftz* polyadenylation sites were deduced by examination of the

sequence (Laughon and Scott, 1984). There are two major putative sites separated by approximately 190bp. The upstream site is a cluster of three imperfect signals. The downstream site includes a perfect AATAAA (see Fig.6.1). I therefore mapped the polyadenylation site of both *ftz* and *eve* and made a construct with a modified polyadenylation signal to test whether the *eve* polyadenylation signal is necessary for transcript localisation.

STRATEGY

Instead of the usual lengthy procedure of constructing cDNA libraries, I modified existing PCR methods to map the polyadenylation sites of *eve* and *ftz*.

The importance of the polyadenylation signal for localisation was tested by introducing a two base pair mutation in the AATAAA of *e32*, the fragment of *eve* which is necessary and sufficient for localisation (see chapter 4). The mutated fragment was inserted into *LT* and the transcript of the resulting fusion construct was examined to see if the disruption of the polyadenylation process altered transcript localisation.

RESULTS

Mapping the polyadenylation sites of *eve* and *ftz*

cDNA was made from wild type embryonic poly(A)+ RNA using a poly(T) primer with an extra 5' G-C rich region including a NotI site. The cDNA was used as a template for a PCR reaction using the same poly(T) primer and internal genomic primers *e1* and *f1* for *eve* and *ftz* respectively (see materials and methods for further experimental details and chapter 4 for details of *e1* and *f1*). The first PCR cycle synthesised the second strand of the cDNA and subsequent cycles amplified the DNA. In theory, this should amplify cDNA of *eve* and *ftz* transcripts only. In practice, each reaction

gives a smear plus a single predominant band of the expected size, plus several minor bands of other sizes. The major bands found are a 250bp product in the case of *eve* and a 540bp band in the case of *ftz*, which correspond well with the predicted minimum sizes of 240bp and 515bp for *eve* and *ftz* respectively.

Bands of the predicted sizes were inserted into the Bluescript vector. Four *ftz* and seven *eve* clones containing inserts of the correct size were sequenced from double stranded miniprep DNA (see chapter 2).

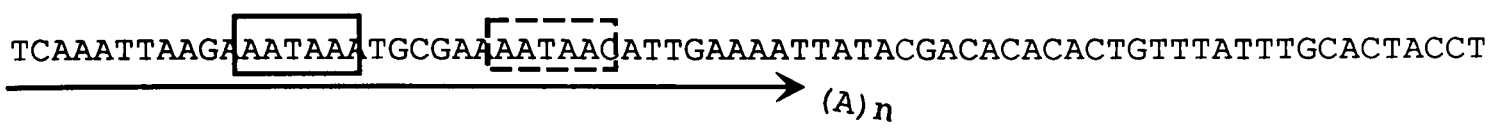
The four *ftz* clones were found to correspond to the cluster of imperfect upstream sites for polyadenylation (just downstream of the RI site). The *ftz* clones fell into three groups all within 11bp of DNA. Two of the clones were derived from RNA molecules which polyadenylated after a CAA (see Fig.6.1), a situation which is commonly found in other genes (Rushlow *et al.*, 1989). The distance between the polyadenylation signals and the site of polyadenylation was found to vary between 22bp and 29bp, corresponding well with other known genes (Levitt *et al.*, 1989 and see Fig.6.1).

In the case of *eve* only one clone was found which corresponded to the predicted polyadenylation site, the other six clones were unidentified. The polyadenylation site found was 21bp downstream of the AATAAA site, again corresponding well with the consensus worked out from other cases (see Fig.6.1 for details of the sites of polyadenylation of *ftz* and *eve*).

Mutating the polyadenylation site of *eve*

The *eve* AATAAA was mutated to test the involvement of polyadenylation in apical localisation. The polyadenylation signal was altered to AATGCA, thus creating an SphI site. The mutation was introduced in the *eve* fragment *e32* which was shown in chapter 4 to be necessary and sufficient for localisation. A mismatch in the primers used for PCR was amplified

a) *eve* Poly A addition site



b) *ftz* Poly A addition site

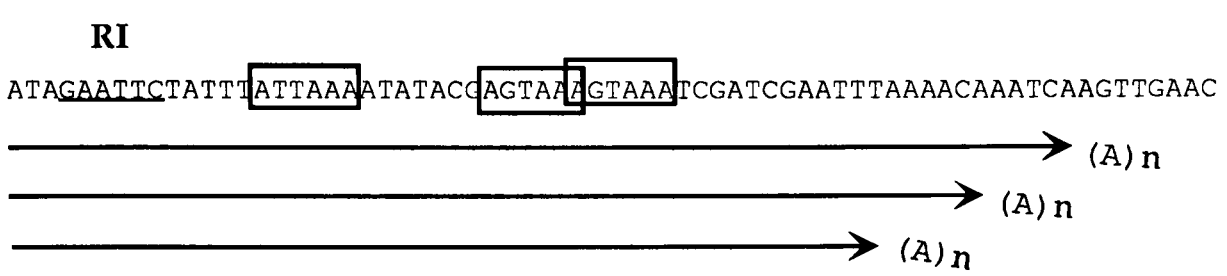


Fig 6.1: Mapping the polyadenylation sites of *ftz* and *eve* .

cDNA was made from wild-type embryonic RNA. *ftz* and *eve* cDNA was selectively amplified using the PCR reaction (see chapter 2, for further details) and cloned into the Bluescript vector. *ftz* and *eve* inserts were sequenced to reveal potential sites of polyA addition. Boxes indicate putative polyadenylation sites, arrows are the regions found before a poly A tail [indicated as (A)n] and important restriction enzyme sites are underlined. The dashed box in the *eve* sequence shows an imperfect putative polyadenylation site.

during a PCR reaction so that it was incorporated into the final product which was vastly in excess of the original DNA template (see chapter 2).

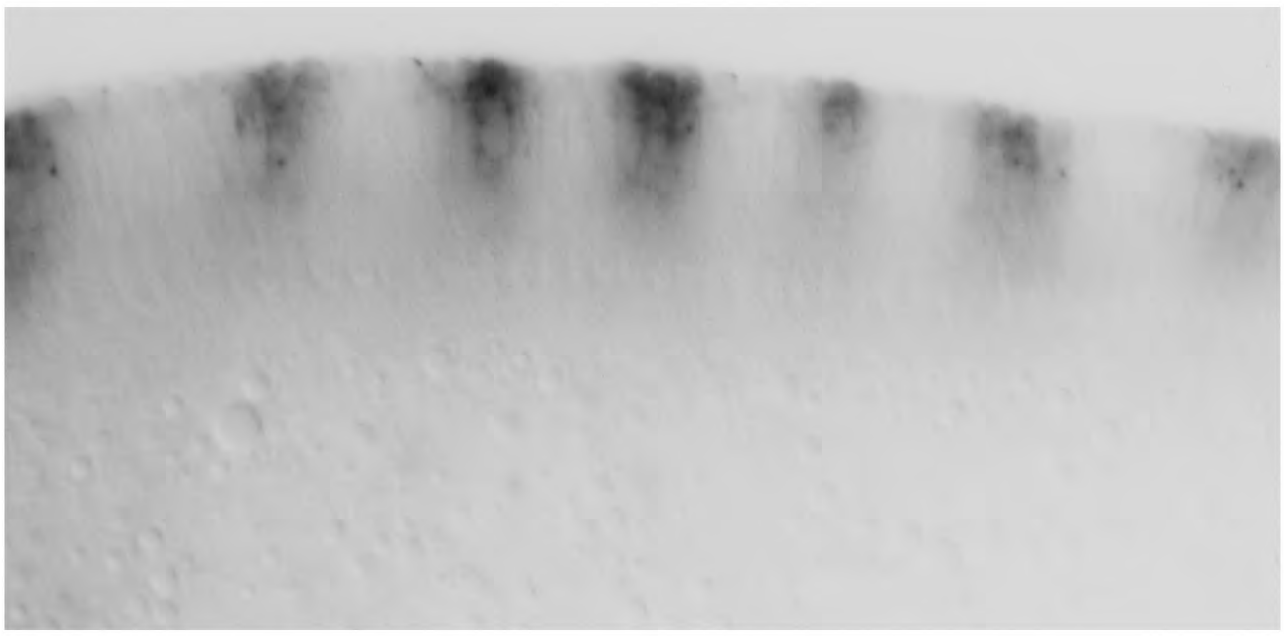
LTeΔ was constructed by inserting the mutated *e32* into the *LT* vector. Before injecting the plasmid into flies, the presence of a mutation in the polyadenylation signal was determined by sequencing the insert in Bluescript. The sequence confirmed that the polyadenylation signal was indeed mutated to AATGCA and included an SphI site.

The plasmid was injected into flies and seven independent permanent lines were recovered, four lines on the second chromosome and three on the third. *In situ* hybridisation was performed with a DIG-labelled *LacZ* probe on the resulting transformed flies *(LTeΔ)*. A preliminary mixing experiment with all the lines revealed that the *LTeΔ* transcript is localised to the apical periplasm. The experiment was repeated on two of the individual lines. In both cases the transcript was restricted to the apical periplasm (see Fig. 6.2).

The results appear to suggest that polyadenylation at the *eve* AATAAA is not necessary for localisation. However, it was important to demonstrate that mutating the AATAAA prevented polyadenylation form taking place at the normal site.

In situ hybridisation with a *globin* probe was used to determine the site of polyadenylation. A *globin* probe was first used on a number of *LT* based hybrid constructs described in the previous chapters (see Fig.6.5). When probed with *globin, LT* gives a similar but weaker pattern of basal localisation to that described in chapter 4, when a *lacZ* probe was used. The signal was weaker with the *globin* probe because the DNA fragment from which the probe was made is 330bp long, compared with the *lacZ* fragment which is 3kb long. The experiment was technically more difficult as it

a)



b)

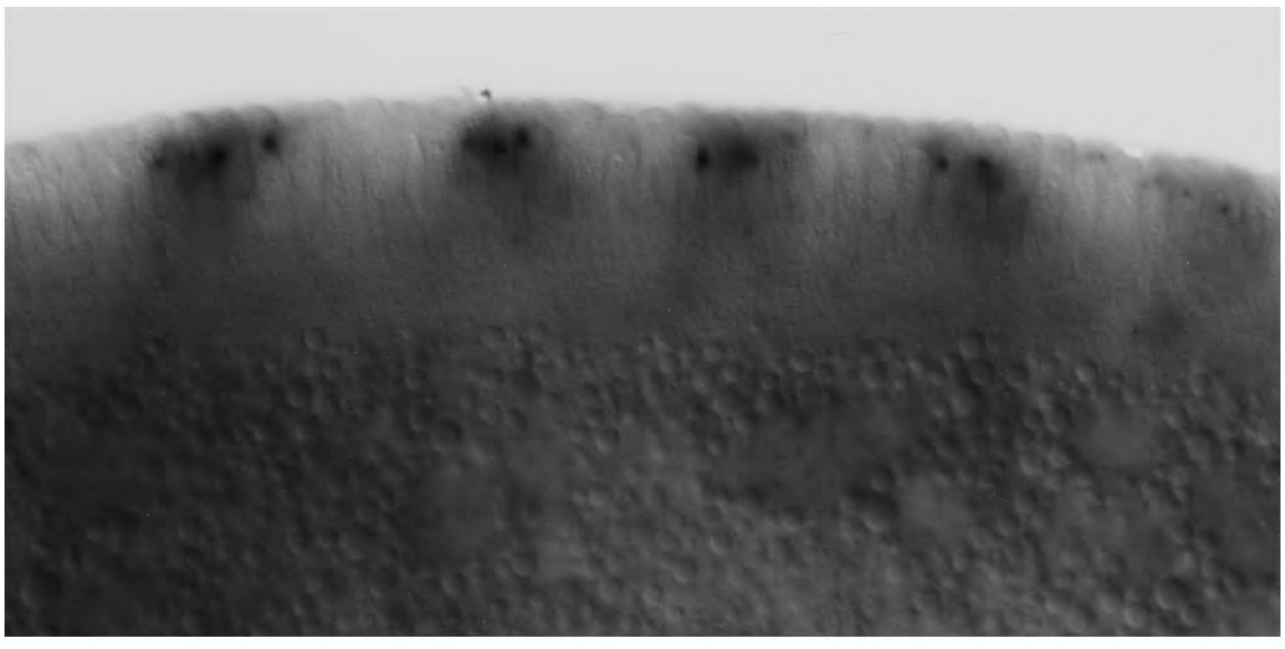


Fig 6.2:

Whole mount *in situ* hybridisation using DIG-labelled probes, revealing that mutating the polyadenylation signal does not affect apical localisation. a) *LTeΔ* with *lacZ* probe, b) *LTeΔ* with *globin* probe.

required long overnight incubations in staining solution (the usual incubation time was one hour-see chapter 2). Nevertheless, the data were interpretable, despite higher background staining than with a *lacZ* probe (see Fig. 6.3). It follows that transcription continues into *globin* sequences in *LT*.

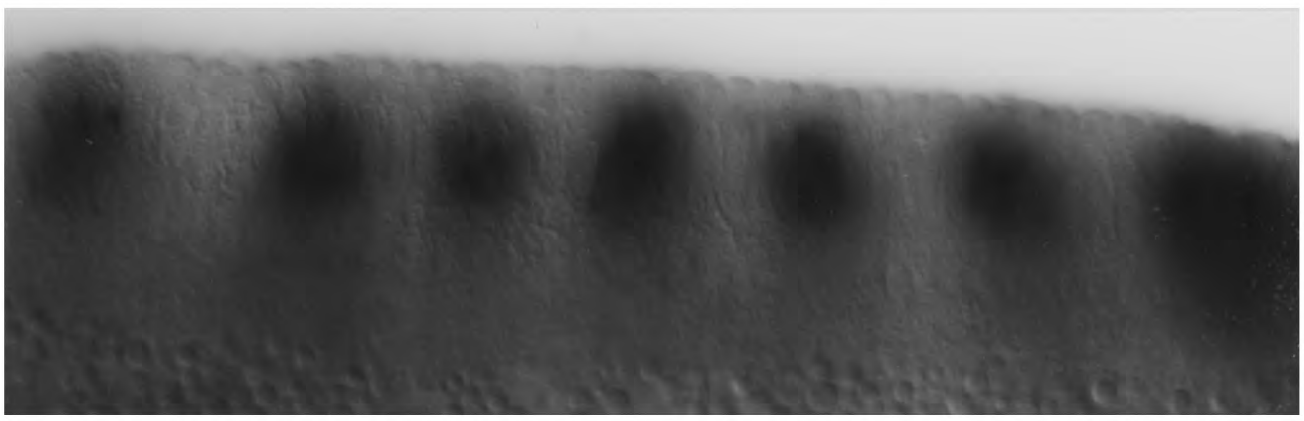
LTe34 gave a similar result to that found with *LT*, suggesting that polyadenylation occurs at the *globin* poly(A) site since the fragment of *eve* used in the construct lacks any polyadenylation signal. In contrast, *LTe32* and *LTe42* probed with a *globin* probe show no signal in the cytoplasm. Both constructs contain the *eve* polyadenylation signal; the lack of cytoplasmic transcript is due to the transcript terminating before the *globin* sequences. However, these two constructs did have very low levels of transcripts in the nucleus when probed with a *globin* probe. These were likely to be preprocessed nascent transcripts (see chapter 7). It is possible to conclude that the *globin* probe can be used to determine whether a fragment inserted into the *LT* vector contains a functional polyadenylation signal.

In situ hybridisation using a *globin* probe on *LTeΔ* embryos fails to detect cytoplasmic transcripts, although punctate staining is visible in the nuclei (see Fig. 6.2). This result is found in both the lines analysed above. The punctate nuclear staining probably represents unprocessed nascent transcripts (see chapter 7); the hybrid gene is likely to be polyadenylated in *eve* sequences and not at the expected *globin* polyadenylation signal. Therefore, it is impossible to conclude that the sequences necessary for polyadenylation can be removed without affecting apical transcript localisation. However, the results suggest that the exact site of polyadenylation within the *eve* sequences is unimportant, since a second site is probably used instead of the mutated site.

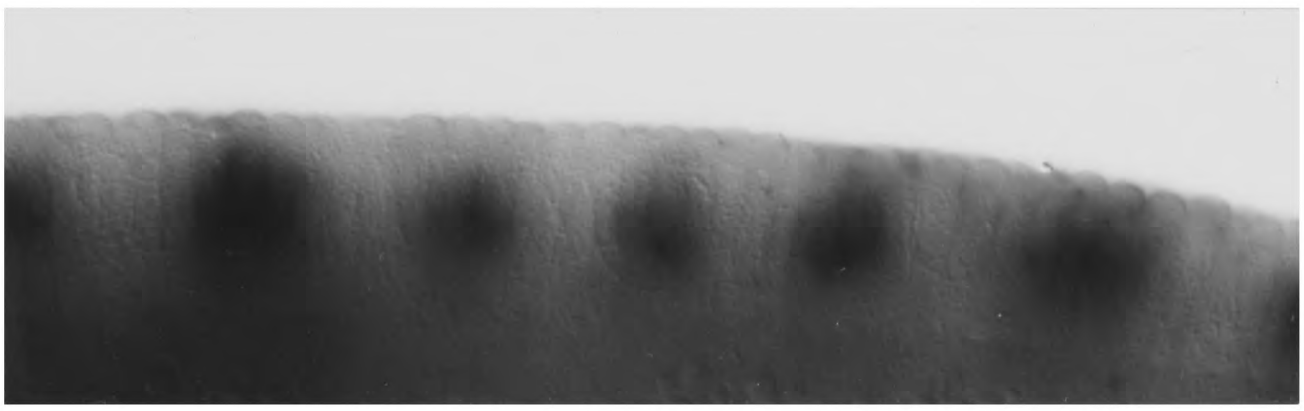
Fig. 6.3

Whole mount *in situ* hybridisation, using DIG-labelled *globin* probes, revealing punctate nuclear transcripts and weak cytoplasmic staining in a) *LT*, b) *LTe34*, c) *LTe32*, d) *LTe42* e) *LTfSR*. (See chapter 4 for a description of the constructs).

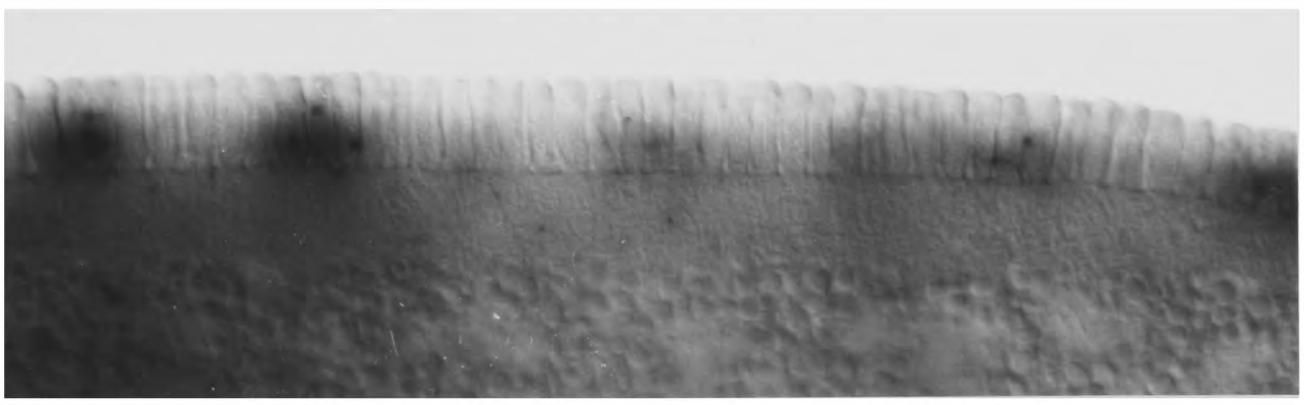
a)



b)

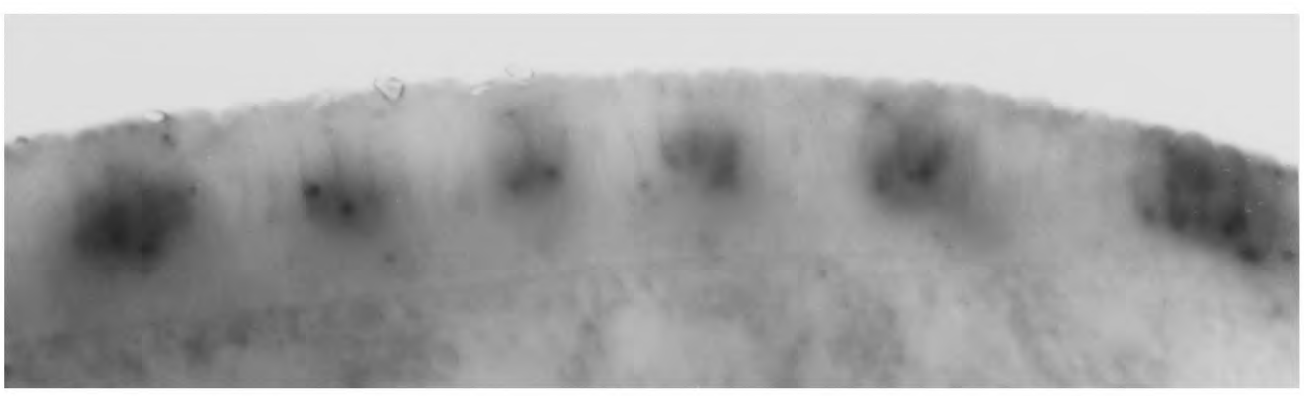


c)



d)

e)



DISCUSSION

The mapping of the polyadenylation sites of both *ftz* and *eve* shown in Fig. 6.1 correspond well with the previously published putative polyadenylation signals (see above). In the case of *ftz* there are three potential imperfect AATAAA sequences and three sites 22-29bp downstream of these at which polyadenylation takes place. Two out of these three occur at CA sequences which are known to be a common site of poly(A) tail addition. In the case of *eve* the polyadenylation site is 21bp downstream of a perfect polyadenylation signal. Thus it seems likely that *eve* and *ftz* are indeed polyadenylated at the sites indicated in Fig. 6.1.

None of the sequences contained extra adenosine residues which were not present in the poly(T) primer used. This can probably be explained by the fact that progressive PCR reactions favour shorter and shorter products so that eventually the shortest possible product is the dominant species. Under such circumstances the PCR reactions would give fairly strong single bands as opposed to a range of sizes corresponding to fragments with different length poly(A) tails. Strong bands were indeed found and were of the expected sizes; 250bp for *eve* and 540bp for *ftz* (the predicted sizes were 240bp and 515bp respectively).

In both cases I obtained some sequences which did not correspond to *ftz* or *eve* sequences and were presumably due to mispriming of the poly(T) PCR primer. An exceptionally A-T rich primer is expected to misprime with far greater ease than a primer of the normal G-C composition because A and T only form two hydrogen bonds whereas G and C are linked by three bonds. It is also possible that the poly(T) primer primed to the poly(A) tail of a completely unrelated gene and the other primer was mispriming in that

gene. The resulting product would give a sequence completely unrelated to either *ftz* or *eve*.

The analysis of a variety of fusion constructs using a *globin* probe indicates that *in situ* hybridisation with this probe is a reliable indicator of whether the transcript extends into the downstream *globin* sequences in *LT*. *LTe42* and *LTe32* are very strongly expressing constructs when assayed using a *lacZ* probe, but a *globin* probe reveals very little if any mRNA in the cytoplasm. In contrast, *LT* and *LTe34* are both quite weakly expressing constructs but give a substantial signal when probed with a *globin* probe. It follows that the *globin* probe is sufficiently sensitive and can be used to test which polyadenylation signal is being used when the *eve* poly(A) signal is mutated although it does not distinguish between polyadenylation sites within the insert.

Mutating the *eve* polyadenylation signal (shown in Fig. 6.1) from AATAAA to AATGCA was expected to abolish polyadenylation at the *eve* insert and thus to lead to a mature transcript containing *globin* sequences. However, no cytoplasmic signal was detected when a *globin* probe was used, suggesting that the *LTeΔ* transcript did not include *globin* sequences. Inspection of the DNA sequences around the normal site of polyadenylation in *eve* indicates that a second imperfect polyadenylation signal is found 12bp downstream of the normal AATAAA (see Fig. 6.1). My preferred explanation is that the imperfect signal replaces the mutated one so that polyadenylation occurs 12bp further downstream.

In situ hybridisation on *LTeΔ* with *lacZ* probes reveals that *LTeΔ* mRNA accumulates at slightly weaker levels than *LTe32* mRNA and is found to localise to the apical periplasm. It follows that the particular *eve* AATAAA signal mutated, is unimportant for localisation, since the construct is probably polyadenylated elsewhere. However, the downstream GT-rich

regions necessary for polyadenylation may be involved in the polyadenylation of both the perfect and the imperfect AATAAA in the *eve* 3' end. Furthermore, the downstream sequences would be more likely to be involved in any localisation signal than the AATAAA itself. Thus the question of the requirement of polyadenylation sequences in the localisation of pair-rule transcripts remains open.

In the long term, one approach would be to remake *LTeΔ* with a deletion of the other putative polyadenylation site. Such an experiment might indicate whether the sequences required for polyadenylation are also necessary for apical localisation.

In the next chapter I describe experiments which test the possible involvement of DNA sequences in determining the topological position of chromatin within the nucleus, and thus the direction of mRNA export from the nucleus. Such a mechanism could explain how pair-rule transcripts become localised to the apical periplasm.

CHAPTER 7: LOCALISATION IS MEDIATED BY TRANSCRIPT-GATING.

INTRODUCTION:

In Chapter 5, it was concluded that pair-rule transcript localisation is unlikely to involve a cytoplasmic mechanism; a nuclear mechanism could account for the available data on apical localisation. *A priori,* the most attractive model for localisation is the export of pair-rule mRNA from the apical side of the nucleus. The transcript would then remain localised in the apical periplasm either because of a diffusion barrier or because of a second step which anchors it to the cytoskeleton. Before discussing more specific models, it is helpful to evaluate different views on the nature of the eukaryotic cell.

There are two conflicting schools of thought describing the nature of the cellular environment (Agutter, 1985). One view is that both the cytoplasm and nucleoplasm consist of an aqueous solution in which macromolecules are free to diffuse, except where a membrane is found. In this 'free-diffusion' model, membranes are the only points at which the movements of molecules are regulated in the cell. A contrasting view of the cell is provided by the 'solid-state model', which proposes that all macromolecules are bound to the cytoskeleton and the nuclear matrix. These skeletal components provide an ordered structure for the movement of different macromolecules in the cytoplasm and nucleoplasm.

The two models have been applied to the question of mRNA movement from the nucleus to the cytoplasm (Agutter, 1985; Clawson *et al.*, 1985) and thus provide a basis for discussing models for apical localisation.

There are three classes of models which could explain how pair-rule mRNA might be exported only from the apical side of the nucleus (see Fig. 1.4 in chapter 1). The first corresponds to the free-diffusion view of the cell. In

this model, RNA is freely diffusible in the nucleus immediately after being transcribed and is only able to exit the nucleus through nuclear pore complexes (NPCs) on the apical side of the nuclear membrane.

The second class of models explaining apical localisation corresponds to the solid-state view of the cytoplasm and nucleoplasm. The 'gene-gating hypothesis' is a theoretical view of nuclear export of mRNA (Blobel, 1985). It states that specific DNA sequences in chromatin direct the export of the transcripts of genes located near these DNA sequences through particular nuclear pores.

The third class of model which I refer to as the 'transcript-gating hypothesis', states that sequences in the mRNA direct the transcript along nuclear tracks to the apical side of the nucleus where they are exported to the apical periplasm. The following is a discussion of experimental work relevant to each of the three models.

RNA is not freely diffusible in the nucleus

A theoretical estimate of the distance diffused by an RNA molecule in the cytoplasm at room temperature points to the fact that even an extremely large globular-shaped protein diffuses very fast (Edgar *et al.*, 1987). A 5kb pair-rule transcript would diffuse 110μm (55 nuclear diameters) in 5min and would therefore spread throughout the nucleus in a very short time (see appendix A). However, there are considerable data suggesting that particular RNA species are found localised to particular parts of the nucleus and are always bound to proteins, so that RNA is not freely diffusing. Indeed, there are many experiments in a variety of systems, briefly described below, which indicate that RNA is bound to an immobile nuclear skeleton.

The degree to which RNA is free to diffuse in the nucleus depends on the structure of the nucleoplasm and the extent to which RNA is bound to

other molecules in the nucleus (reviewed in Gerace and Burke, 1988 and Newport and Forbes, 1987). There is some controversy over the structure of the nucleoplasm, which stems from the conditions used to extract chromatin. Extraction under hypertonic salt conditions indicates that DNA is closely associated with a nuclear skeletal structure, the nuclear scaffold or matrix (Miller *et al.*, 1978). When chromatin spreads are made under hypotonic salt conditions, no nuclear matrix is apparent (Beyer and Miller, 1980).

The two extraction procedures produce different artifacts: under hypertonic conditions nucleosome structure is destroyed but attachment to a nuclear matrix is preserved (Jackson and Cook, 1985). In contrast, under hypotonic conditions, nucleosome structure is maintained but the nuclear matrix is destroyed. However, a third extraction procedure under isotonic salt conditions probably produces fewer artifacts because nuclei are embedded in agarose microbeads, with minimal structural disruption. Under isotonic conditions, a nucleoskeleton is found to be associated with intact chromatin and transcription is specifically associated with the nucleoskeleton, probably via a variety of ribonuclear proteins (RNPs) which are bound to hnRNA (Jackson *et al.*, 1982; Jackson and Cook, 1985; Miller *et al.*, 1978).

Zygotic transcripts are bound to large RNPs while they are being transcribed in the *Drosophila* blastoderm (McKnight and Miller, 1976). Electron microscope analysis of chromatin spreads in 2-4 hour embryos reveals that nascent hnRNA is covered with RNP particles, 240Å in average diameter (Beyer and Miller, 1980). The Balbiani ring (BR) granules in the salivary glands of the dipteran *Chironomous tentans* are large RNP particles attached to nascent hnRNA. Because of their abundance, it has been possible to analyse the 3-D structure of BR granules in detail (Skoglund *et al.*, 1986). It is suggested that RNA remains complexed to large and diverse

proteins after the completion of its transcription. Indeed, polyadenylation and splicing are believed to occur at such complexes (Skoglund *et al.*, 1986). Furthermore, transcripts remain associated with the nucleoskeleton during RNA processing (van Eekelen and van Venrooij, 1981). It follows that RNA is not freely diffusible in the nucleus at any stage, and that the free-diffusion model is inappropriate for apical localisation.

The structure of chromatin

The two remaining classes of models for apical transcript localisation are gene-gating and transcript-gating. Both are based on the notion of a well-structured nucleoplasm in which RNA is not free to diffuse, and are therefore compatible with the evidence described above.

The gene-gating hypothesis (Blobel, 1985) suggests that DNA sequences near genes cause the transcripts of nearby genes to be exported through particular NPCs. This model would appear a rather attractive means of ensuring apical localisation of pair-rule mRNA. DNA sequences required for localisation could direct the transcripts of pair-rule genes to a subset of pores which export mRNA from the apical side of the nucleus.

One strict interpretation of the gene-gating hypothesis suggests that DNA sequences required for localisation interact directly with nuclear pores on the apical side of the nuclear membrane so chromatin containing the pair-rule genes lies in the apical side of the nucleus. However, other interpretations need not require a rigid apical confirmation of chromatin containing pair-rule genes.

Data described below suggest that *Drosophila* chromatin is not sufficiently rigid in its conformation to allow for all apically localised genes to lie next to the apical side of the nuclear membrane. As early as 1885, Rabl suggested a theory on the internal structure of the interphase nucleus as seen under the

microscope. In modern terminology, Rabl stated that chromosomes maintain a polarised orientation within the nucleus (Cremer *et al.*, 1982). Centromeres are aggregated at one pole, and telomeres at the opposite pole, with each chromosome occupying a distinct territory. The 'Rabl orientation' is formed during the previous mitotic division, when the centromeres are all clustered at the chromocentre and the telomeres are clustered in the opposite side of the newly reformed nuclear membrane.

3-D reconstruction from fluorescent optical sections using image intensifying equipment has confirmed Rabl's observations in living interphase cells (Agard and Sedat, 1983). The Rabl orientation has been shown to occur in many different cell types including Chinese hamster cells (Cremer *et al.*, 1982), *Drosophila* salivary gland cells (Mathog *et al.*, 1984) and other larval tissues (Hochstrasser and Sedat, 1987b) and in the syncytial blastoderm (Foe and Alberts, 1985; Hiraoka *et al.*, 1989). In the *Drosophila* blastoderm, the chromocentres are found near the apical side of each peripheral nucleus and are closely associated with the nuclear membrane (Foe and Alberts, 1985).

A detailed analysis of the data shows that the centromeres and telomeres are clustered in opposite sides of the nucleus, and each chromosome arm occupies a mutually exclusive domain in the nucleoplasm (Mathog *et al.*, 1984). The actual orientation of the chromosome arms are different in different individual cells of the same cell-type, as well as in different cell types (Hochstrasser and Sedat, 1987a; Hochstrasser and Sedat, 1987b). While the chromocentres and telomere clusters contact the nuclear membrane at predictable locations at opposite poles, other fixed cytological locations contact the nuclear membrane at variable positions in space (Mathog *et al.*, 1984). The contacts are thought to be mediated by binding to nuclear lamins on the inside of the nuclear membrane (Gerace *et al.*, 1984).

The 3-D reconstruction techniques have been refined to give a glimpse of the pattern of higher-order chromosome folding in living cells. The data suggest that the previous two conflicting views of chromosome folding, namely chromatin loops created by scaffold attachment sites (reviewed in Gasser and Laemmli, 1987) and the hierarchical solenoid model (Bak *et al.*, 1977) are both inaccurate due to artifacts created by chromosome extraction procedures. Instead, chromatin is folded into a complex but characteristic pattern with distinct higher-order structural domains of a variety of sizes which do not form a strict hierarchy of solenoids or loops (Belmont *et al.*, 1987).

The data presented above show that the 3-D confirmation of chromatin in the nucleus is fairly well defined, but is not entirely rigid. It is likely that the conformation of chromatin performs a structural role in preventing chromosomes from becoming tangled. The sites of chromatin contact with the nuclear membrane are unlikely to play a role in regulating transcription since the chromatin of midgut cells does not contact the nuclear membrane at all, although gene expression is regulated normally (Hochstrasser and Sedat, 1987a; Hochstrasser and Sedat, 1987b). In this case, the disruption of the normal chromosome conformation is thought to be due to mechanical stress on the nuclei, and yet all the normal house-keeping and some tissue-specific genes are expressed in the midgut cells.

Chromosome orientation and the cytological position of the pair-rule genes

If chromosome orientation is important in apical localisation, the cytological position of different *Drosophila* genes should determine whether or not they are apically localised. The most strict interpretation of the gene-gating hypothesis predicts that all the pair-rule genes are located at or fairly near the centromeres, because the chromocentre is located at the

apical end and the telomeres at the basal end of the nuclear membrane. The gap genes would lie in the centre of chromosome arms and not at telomeres or centromeres. While *ftz* (84AB) and *opa* (82A-E) lie near the centromere of the 3R (at 81) and *run* (19EF) is very close to the centromere of the X chromosome (20D), other pair-rule genes (*eve, h, prd*) lie in the centre arms (see table 7.1). Three of the gap genes are to be found at the centromeres or telomere, not as predicted in the chromosome arms.

h	66D	3L: 2/3 of distance towards telomere
ftz	84AB	3R: quite near the centromere
eve	46C	2R: 1/3 of distance towards telomere
run	19EF	1: very near the centromere
opa	82A-E	3R: near the centromere
prd	33B	2L: in the middle of the chromosome arm
Kr	60F	2R: at the telomere
hb	85A	3R: near the centromere
kni	77E	3L: quite near the centromere

Table 7.1

The cytological position of several pair-rule and gap genes and their degree of proximity to telomeres or centromeres (derived from Jürgens *et al.*, 1984; Nüsslein-Volhard *et al.*, 1984; Wieschaus *et al.*, 1984).

It follows that gene-gating, in its most strict sense, cannot explain apical localisation of pair-rule mRNA. However, the other less rigid interpretations of the hypothesis are more plausible as they do not require that the pair-rule genes lie near pores on the apical side of the nuclear membrane. One such variation involves DNA sequences interacting with nuclear components which are linked via tracks in the nucleoplasm (Lawrence *et al.*, 1989) to a particular subset of NPCs (see Fig.1.4 in chapter 1).

All the variants of the gene-gating hypothesis have one thing in common: they all involve DNA sequences near or in the genes mediating mRNA export through particular NPCs. They predict that if a gene is located near

DNA sequences responsible for localisation, its transcripts will be found exclusively in the apical periplasm. On the other hand, the transcript-gating hypothesis requires that apical localisation of pair-rule transcripts is mediated by RNA sequences which are independent of chromosome orientation and cytological position. It is thus possible to distinguish between the two hypotheses by establishing whether the mechanism of localisation involves DNA or RNA sequences.

STRATEGY

The gene-gating hypothesis predicts that two adjacent genes under similar transcriptional regulation should encode transcripts with the same intracellular localisation. The transcript-gating hypothesis predicts that the two genes may show different patterns of intracellular localisation according to the sequences in their transcripts. These predictions are tested by examining the intracellular localisation of three *lacZ* enhancer-trap genes and their adjacent *h* gene. It is therefore possible to distinguish between DNA and RNA mediated mechanisms for localisation.

A second test for the involvement of DNA or RNA sequences in apical localisation is provided by inverting a sequence which is sufficient for localisation. If the sequence operates at the level of DNA, its inversion would not prevent it from functioning in apical localisation. However, if the sequence operates at the level of RNA, the inverted orientation would lead to transcription of the antisense strand which would have a very different structure from the sense strand. An inverted RNA sequence would thus not function in apical localisation.

RESULTS

P-*lacZ* enhancer-trap insertions into *h* show that localisation does not depend on DNA sequences.

Enhancer-trap plasmids have been used to define upstream control elements of endogenous genes by inserting the *LacZ* reporter gene upstream of genes (O'Kane and Gehring, 1987). One such plasmid, is *Carnegie20*-based and contains a *P{lac, ry+}A* transposon which is flanked by two P-element inverted repeats. It contains a weak P-element promoter upstream of the *lacZ* reporter gene with a *Drosophila hsp-70* stop codon and polyadenylation signal as well as the *ry* gene for selection of transformed flies according to eye colour. The enhancer-trap *lacZ* gene is essentially transcriptionally inactive unless inserted near enhancer elements (O'Kane and Gehring, 1987).

Several enhancer-trap insertions near the *h* gene have been isolated. I have used these to test whether DNA sequences could be responsible for the localisation of pair-rule mRNA. The *h* enhancer-trap lines used are shown in table 7.2 (see Fig. 7.2 for a description of h^{L43}.

STOCK	ABBREVIATION	SOURCE/REFERENCE
P(lac,ry+)A L43a/TM3	h^{L43}	(Fasano *et al.*, 1988)
P(lac,ry+)A EIII/TM3	h^{EIII}	A. Martinez-Arías (unpublished)
P(lac,ry+)A 58.2C/ P(lac,ry+)A 58.2C	$h^{58.2}$	A. Martinez-Arías (unpublished)

Table 7.2

The three enhancer-trap mutations used to test the gene-gating hypothesis.

The severities of the three *h* mutations were determined from their cuticle phenotypes (Fig. 7.1). The *h* pair-rule cuticular phenotype shows deletions of alternate segment-wide domains consisting of the anterior part of every other even numbered segment and the posterior part of every other odd-numbered segment. The denticle belts are in the anterior part of each segment so in the resulting cuticle the denticle belts of all the even numbered segments are deleted and those of the odd-numbered cuticle bands remain. Stronger phenotypes show additional deletions forming a weak 'lawn' of denticles. In the most severely affected *h* embryos, extensive deletions occur so that an A3-A8 lawn of denticles is formed (Ingham *et al.*, 1985b).

The three *h* enhancer-trap mutations vary in the severity of their cuticle phenotypes. $h^{58.2C}/h^{58.2C}$ is homozygous viable has a weak pair-rule phenotype. Some embryos are normal, some have partial deletions of even numbered denticle belts and some pair-rule cuticles, but no lawn phenotypes are found (see Fig. 7.1). h^{EIII} is a stronger allele. Homozygous lethal embryos show predominantly pair-rule, some weak pair-rule and only very few weak-lawn phenotypes. h^{L43} is the strongest of the three alleles. Homozygous embryos show mainly pair-rule and some weak-lawn phenotypes. The homozygous h^{L43} mutant is slightly stronger than h^{EII} but weaker than *h* null (see Figs. 7.1).

The three lines were analysed by *in situ* hybridisation using *h* and *lacZ* DIG-labelled probes. *lacZ* is expressed in a *h*-like striping pattern in all three lines, indicating that the the *lacZ/hsp-70* gene is regulated by endogenous *h* striping elements (Fig. 7.3). The level of endogenous *h* transcript is substantially reduced in homozygous mutant embryos of the three lines, the degree of reduction being correlated with the severity of the *h* phenotype (see Fig.7.4). The relative levels of transcription of *lacZ* and *h* were

Fig 7.1:

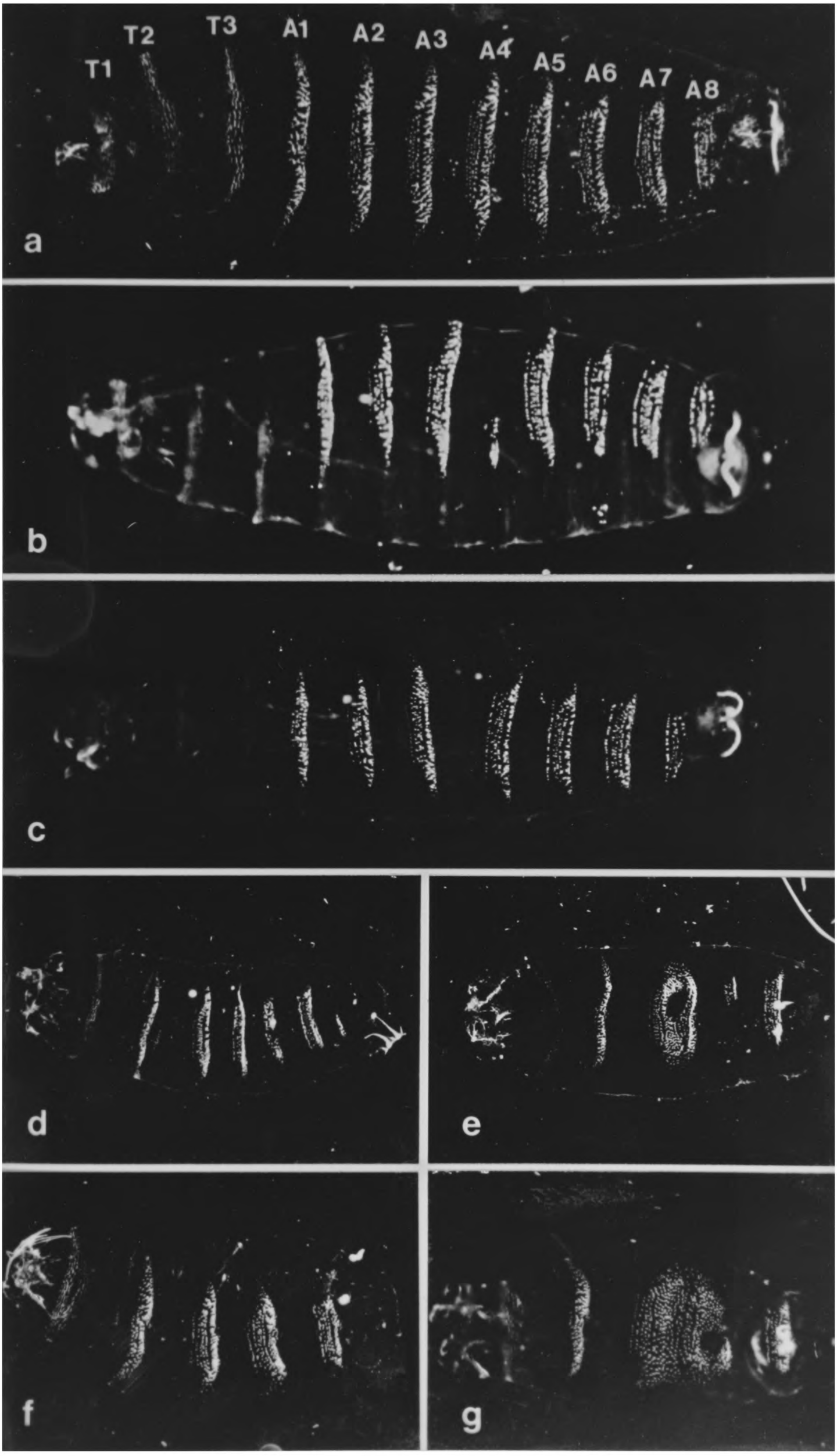
Cuticle phenotype of three enhancer trap *h* mutations compared with wild type. The cuticles are typical of $h^{58.2}$ homozygous viable larvae which have not yet hatched and of h^{AL43} and h^{EIII} homozygous larvae which do not hatch since the mutations are homozygous lethal.

a) The cuticular phenotype of a wild type larva. Thoracic segments are labelled T1-T3, and abdominal segments A1-A8.

b) & c): $h^{58.2}/h^{58.2}$ larvae show occasional segment deletions, characteristic of weak *h* alleles:
b) Very weak pair-rule phenotype with a deletion in the anterior part of A4.
c) Weak pair-rule phenotype with deletions in the anterior part of A4 and some of the anterior of A6.

d) & e): h^{EIII}/h^{EIII} larvae show an intermediate strength pair-rule cuticular phenotype, showing "ideal" or weak pair-rule phenotypes:
d) Slightly weak pair-rule phenotype with the anterior part of A4 and A2 deleted and the anterior part of A6 and A8 reduced.
e) "Ideal" pair-rule phenotype with the anterior part of each even numbered segment and the posterior part of each odd number segment deleted.

f) & g): h^{AL43a}/h^{AL43a} larvae show a strong *h* phenotype, although weaker than a *h* null mutation. "Ideal" pair-rule embryos and weak lawn phenotypes are seen:
f) "Ideal" pair-rule phenotype with the anterior part of each even numbered segment and the posterior part of each odd number segments deleted.
g) Weak lawn phenotype: partial fusions of the posterior parts of the remaining odd numbered segments.


T1
T2
T3
A1
A2
A3
A4
A5
A6
A7
A8
a
b
c
d
e
f
g

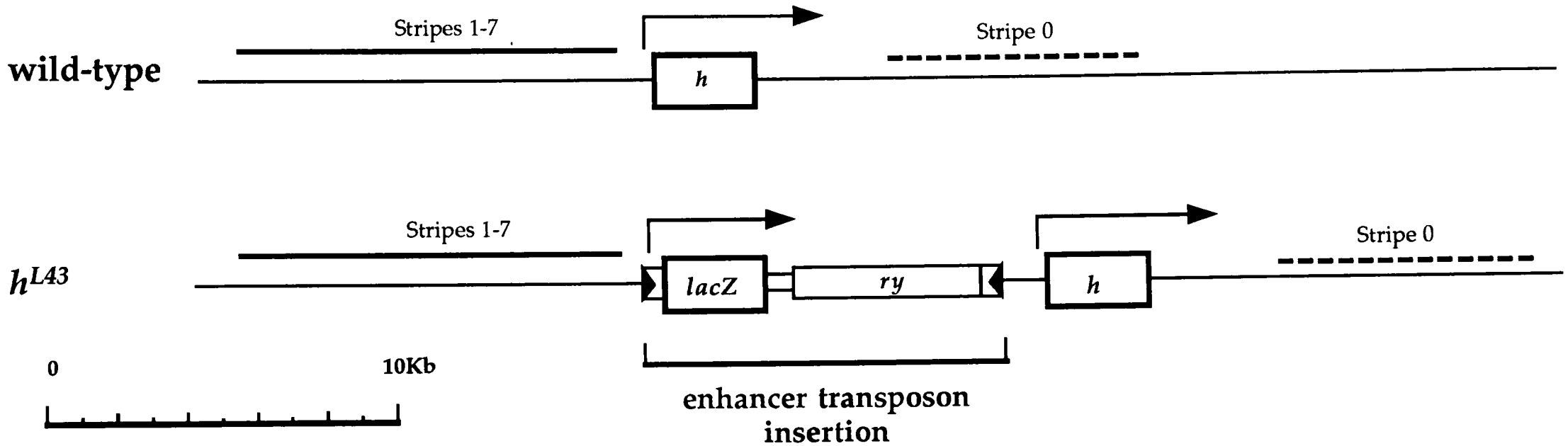


Fig. 7.2

h^{L43} enhancer transposon insertion between the promoter and the enhancer elements for stripes 1-7 of the *h* gene. The position of putative sequences required for the expression of stripe 0 (head patch), is shown by thick dotted lines. The region containing enhancer elements for stripes 1-7 are shown by thick lines. Transcription of the *lacZ* and *h* genes is shown by arrows. The p-element inverted repeats are shown as small black triangles in white boxes.
(Based on Fasano et al, 1988).

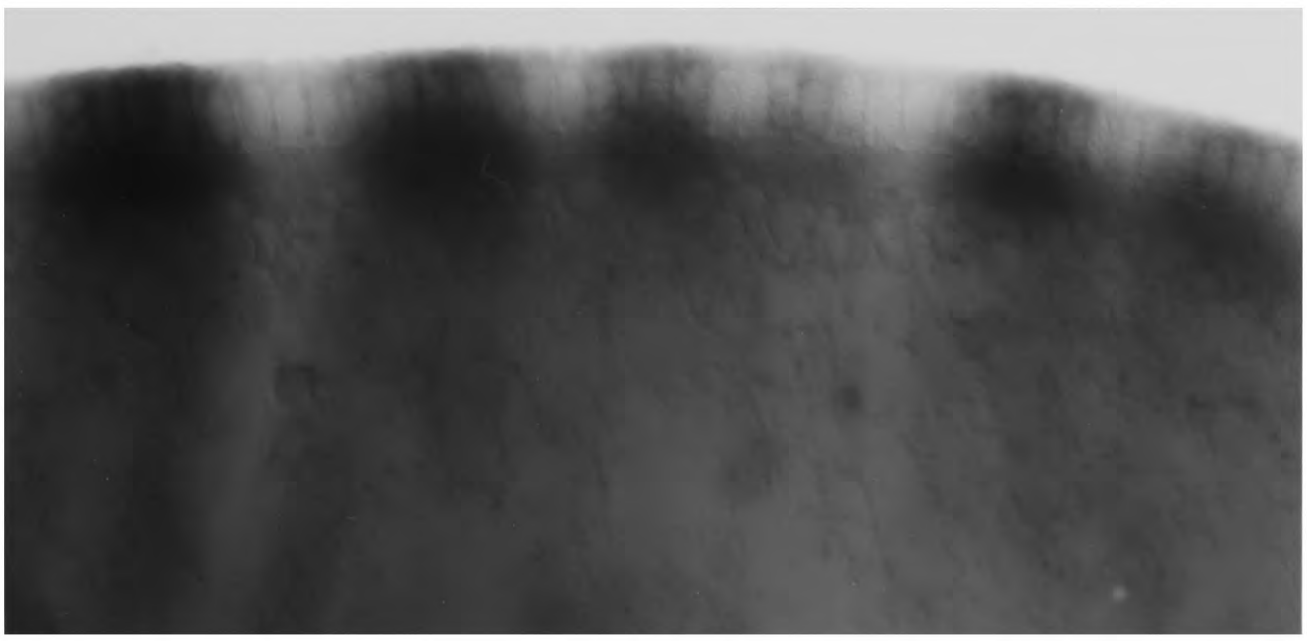
inversely correlated in each line in a manner consistent with the two genes competing for the same trans-acting factors binding to the upstream striping elements, and suggesting that the *lacZ* and *h* genes are under the control of the same upstream striping elements (see Figs. 7.3 and 7.4).

The intracellular localisation of *lacZ* and *h* transcripts was analysed in each of the three lines. In all three lines *lacZ* mRNA is found in both apical and basal periplasms but the *h* transcript is exclusively localised in the apical periplasm of the homozygous mutant embryos. It follows that two genes in close proximity and sharing upstream striping elements show a different pattern of mRNA localisation. As discussed below, these results provide strong evidence against the involvement of DNA-mediated mechanisms of localisation, such as gene-gating, and suggest the involvement of an RNA-mediated signal. The results also argue against a role for cytological position in the localisation of pair-rule transcripts to the apical periplasm.

Inverting the localisation sequences shows that localisation depends on RNA sequences.

The second strategy for determining whether localisation acts via RNA or DNA sequences involves inverting a known localisation sequence. In chapter 4, it was shown that a small region in the 3' end of *eve*, *e32*, was necessary and sufficient for apical localisation. *LTe23INV* was constructed by inserting *e32* in reverse orientation into the *LT* vector forming the hybrid construct. *e32* was generated as described in the previous chapters using the two PCR oligonucleotide primers shown in Fig.7.5. *LTe23INV* was injected into embryos and 3 permanent balanced lines were isolated. Two lines contain third chromosome inserts and one line a second chromosome insert.

a)



b)

c)

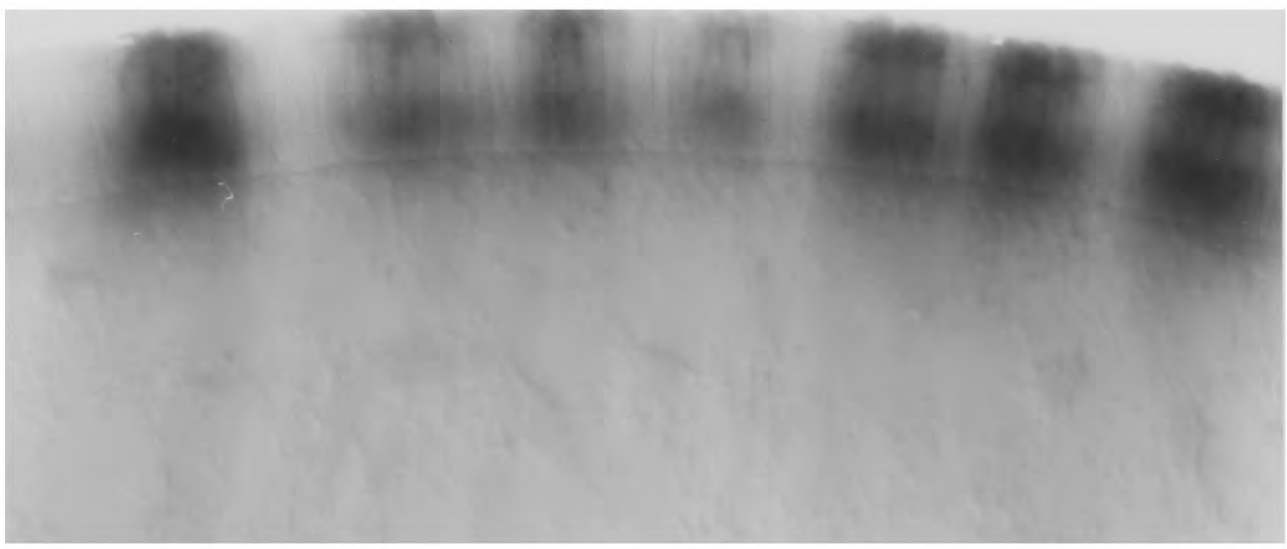
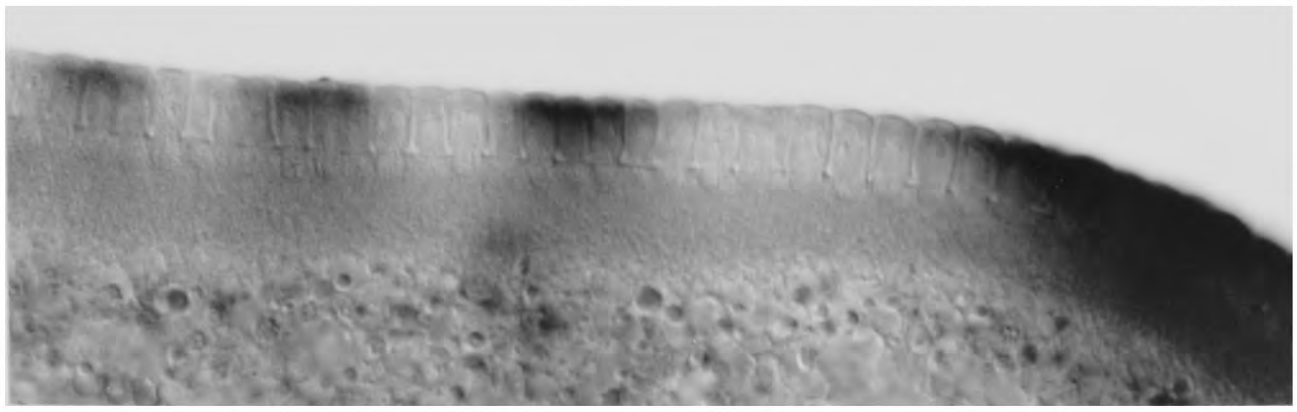
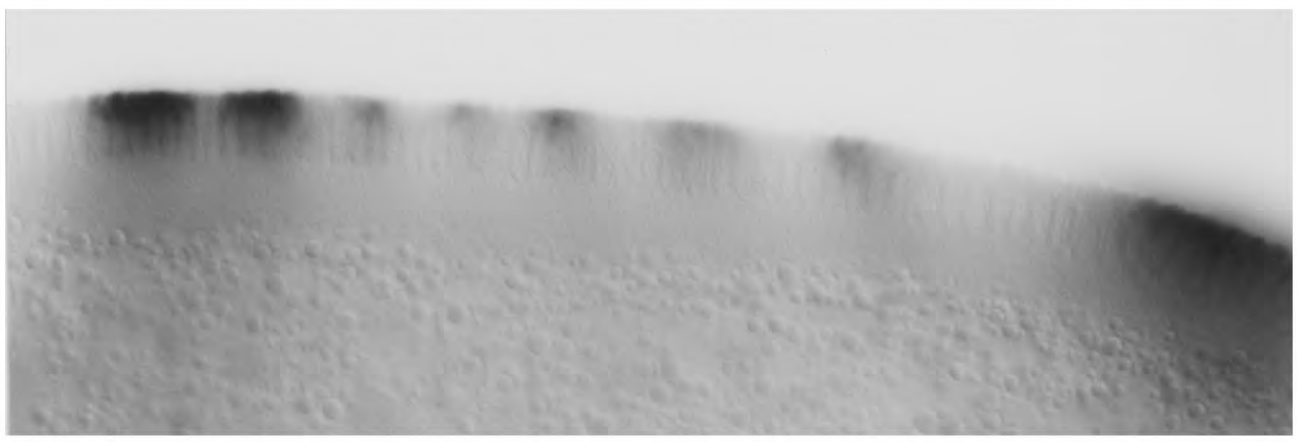


Fig 7.3:
Transcript localisation of the enhancer trap *lacZ* gene. A DIG-lablled *lacZ* probe was used for whole mount *in situ* hybridisation on 2-4h embryo collections. a)$h^{58.2}/h^{58.2}$, b) h^{AL43}/h^{AL43} and c) h^{EIII}/h^{EIII}. The results show that the enhancer trap *lacZ* mRNA is not localised despite the presence of apical localisation sequences in the nearby *h* gene.

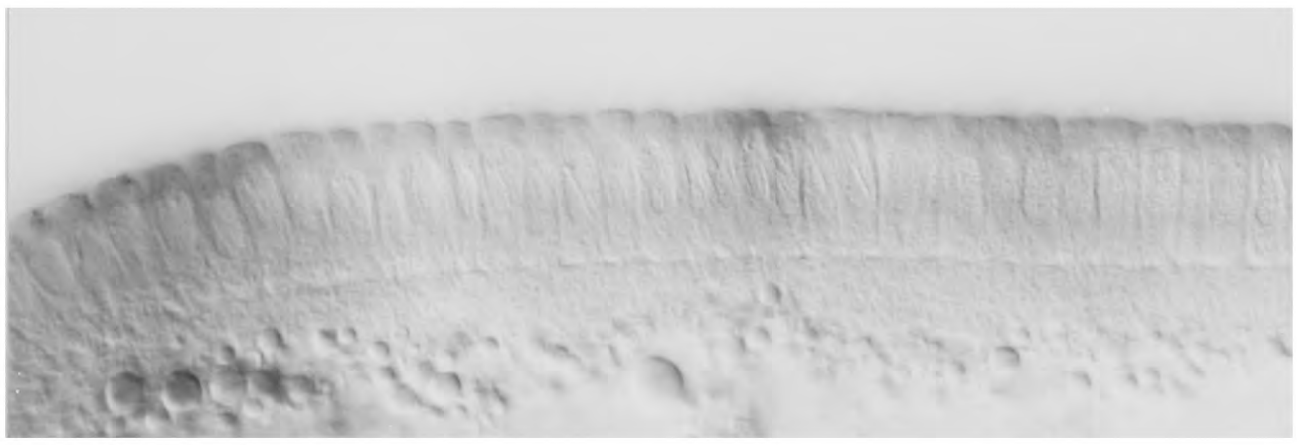
a)



b)



c)



d)

Fig 7.4:
Transcript localisation of the endogenous *h* gene into which the enhancer plasmid inserted. The results show that h transcripts remain localised, despite a reduction in the level of transcription due to the insertion of the enhancer plasmid. The reduced level of expression is correlated with the strength of the mutation (see Fig. 7.1). A DIG-labelled *h* probe was used for whole-mount *in situ* hybridisation on 2-4 hour embryo collections. a) wild-type b) $h^{58.2}/h^{58.2}$, c) h^{AL43}/h^{AL43} and d) h^{EIII}/h^{EIII}.

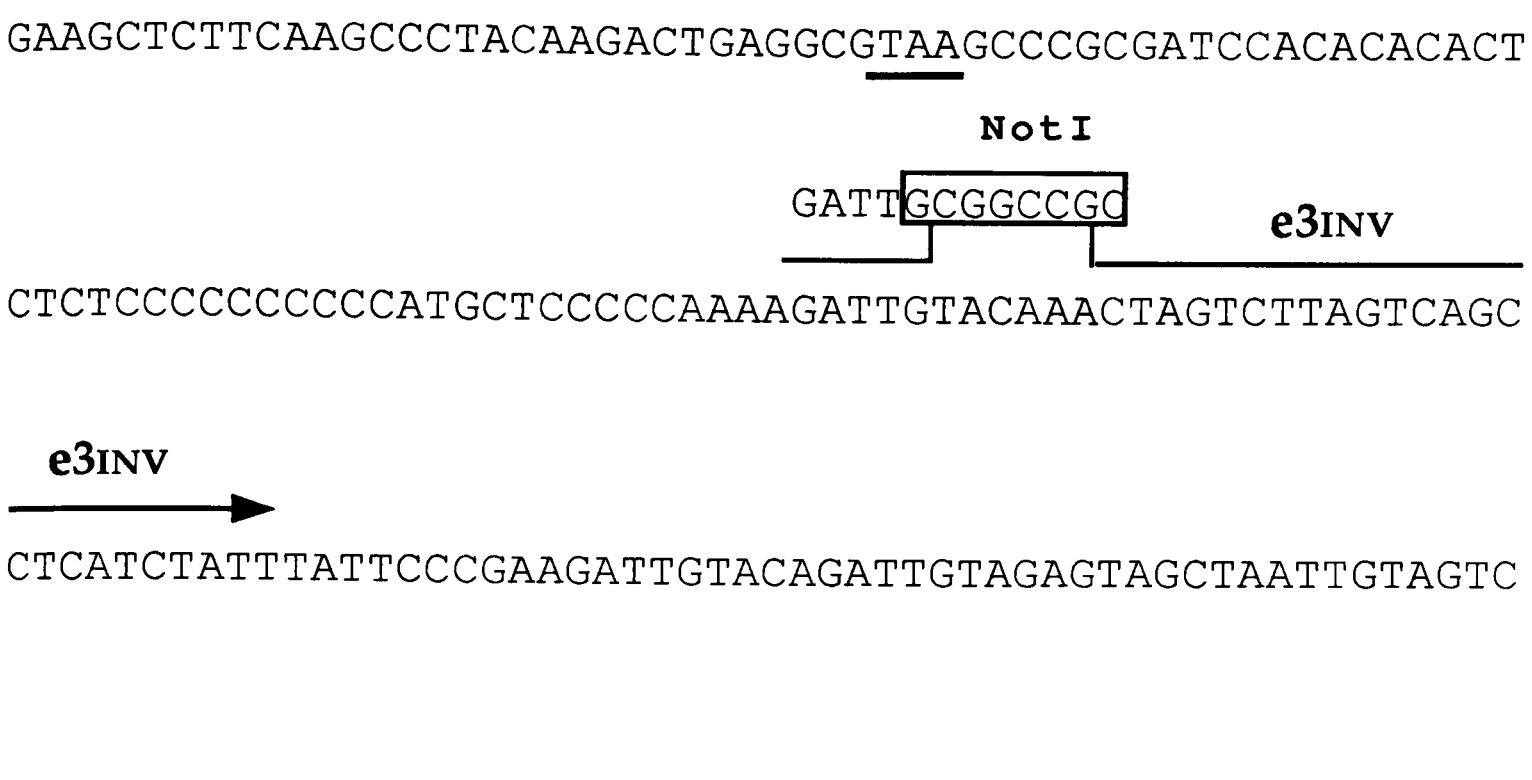
In situ hybridisation using a DIG-labelled *lacZ* probe on two lines, reveals that low levels of unlocalised mRNA are found in the periplasm. Transcript is also found localised in a punctate manner in the nuclei of these embryos (see Fig. 7.6). A DIG-labelled *globin* probe was used for *in situ* hybridisation on the two lines, revealing punctate nuclear transcripts and weak basally localised cytoplasmic mRNA (see Fig. 7.6). It follows that an inverted *e32* fragment cannot specify apical localisation. It can be concluded that apical localisation operates on the level of RNA, and not DNA.

In chapter 4, it was shown that a 1.2kb region of the *ftz* gene is necessary and sufficient for localisation. *LTfHSINV* was constructed by inserting this 1.2kb region of *ftz* in reverse orientation into the *LT* vector (see chapter 4). *LTfHSINV* was injected into flies and two independent permanent lines were isolated, one homozygous viable third chromosome insert and one homozygous lethal X chromosome insert.

In situ hybridisation using DIG-labelled *lacZ* probe reveals punctate staining in discrete regions in the nuclei of both lines (see Fig. 7.6). In what follows, I refer to these regions as punctate nuclear transcripts. These results indicate that most of the *LTfHSINV* transcript is trapped in a discrete site in the nucleus although very low levels of transcript are also detected in the cytoplasm (see Fig. 7.7). The cytoplasmic transcript are unlocalised again indicating that apical localisation is dependent on an RNA-mediated mechanism.

Nascent RNA in the nucleus suggests that transcription occurs at a discrete site in the nucleus

A detailed analysis of the *lacZ in situs* reveals that *LTfHSINV* transcripts are found in discrete punctate regions which are inside the nucleus and not usually in contact with the nuclear membrane (see Fig. 7.7). The results



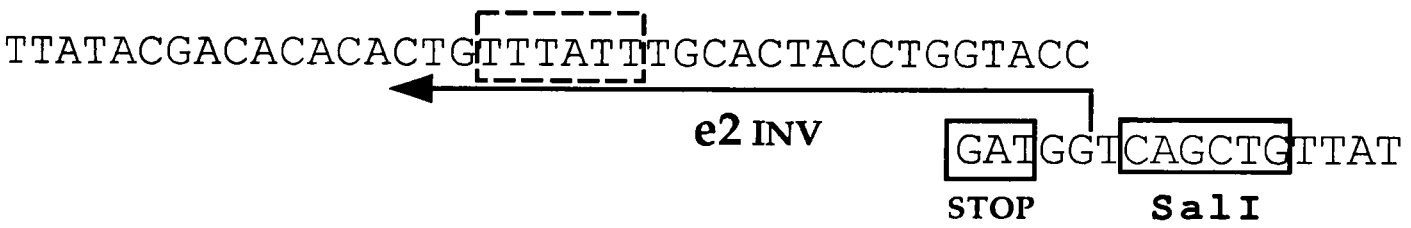
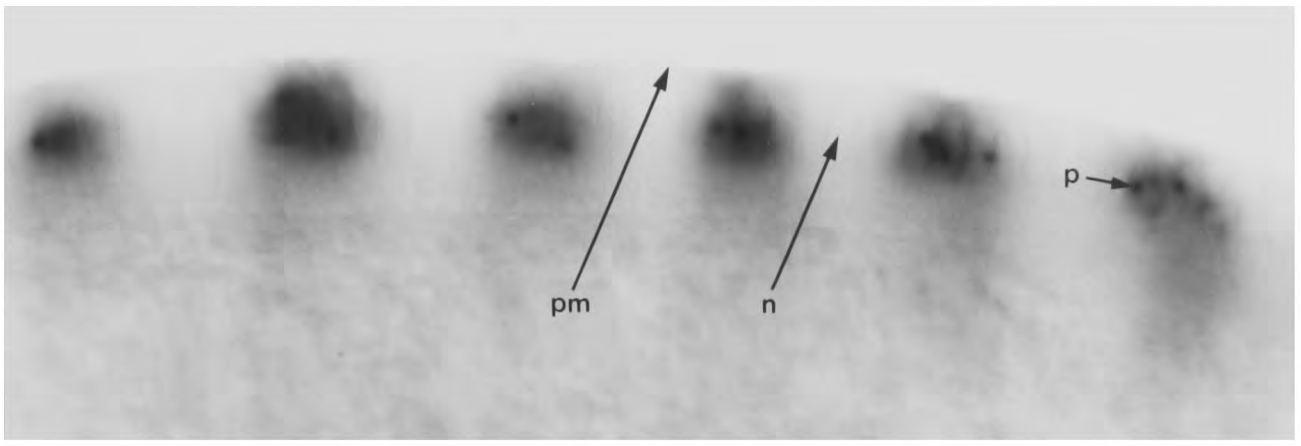


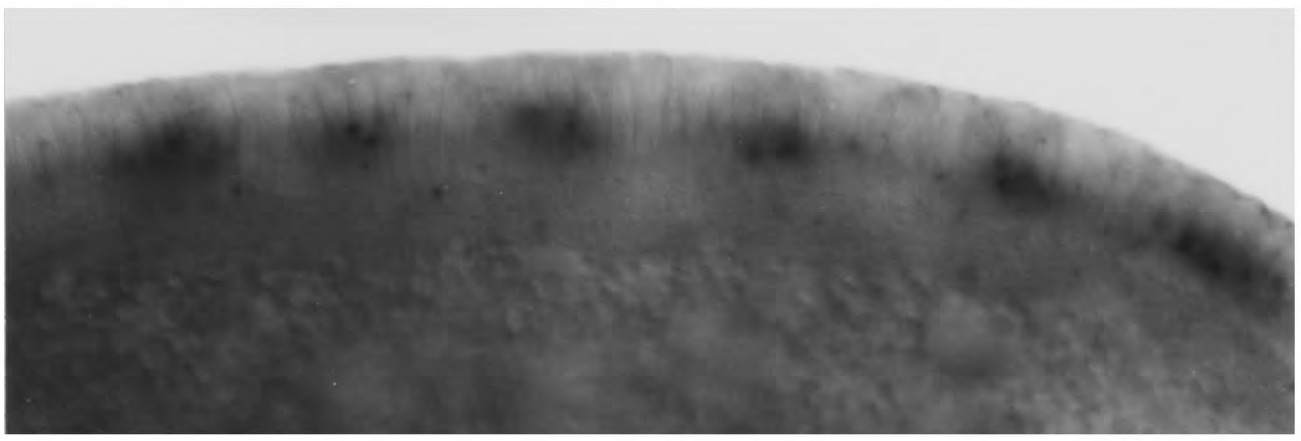
Fig. 7.5

The PCR oligonucleotide (oligo) primers (*e3INV* and *e2INV*) used to generate an *eve32* fragment which is inserted in an inverted orientation into the SalI and NotI sites of *LT*. Stop codons (TAA, TAG) and the polyadenylation signal (AATAAA) are underlined. The box delineated by dotted lines is a putative polyadenylation signal in reverse orientation. The arrows signify regions of the oligo homologous to the sequence. The direction of the arrows indicate whether the sequence of the oligo is that of the sense or antisense strand. The sense strand is shown by arrows pointing to the right, the antisense strands by those pointing to the left.

a)



b)



c)

d)

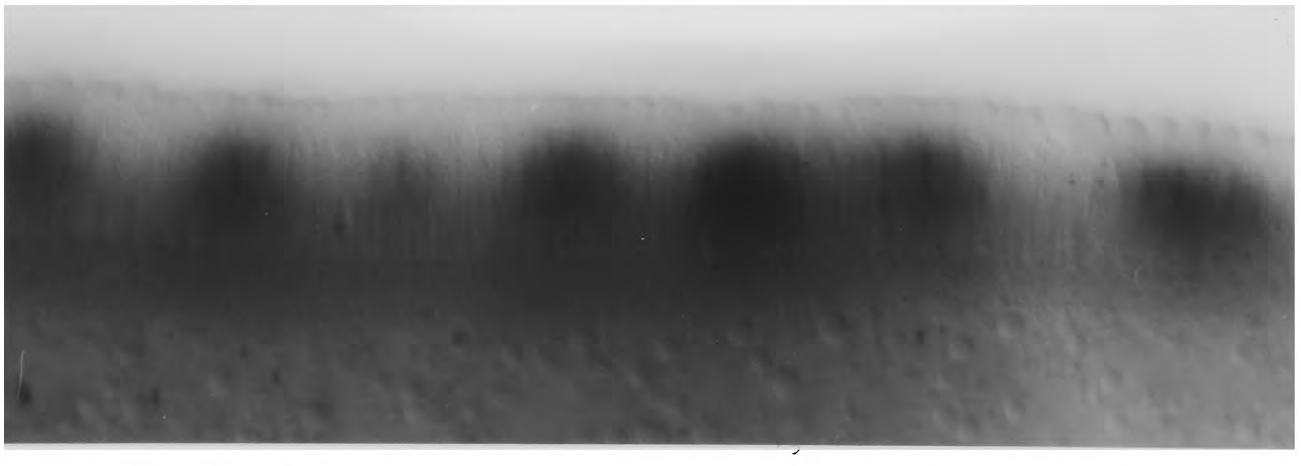
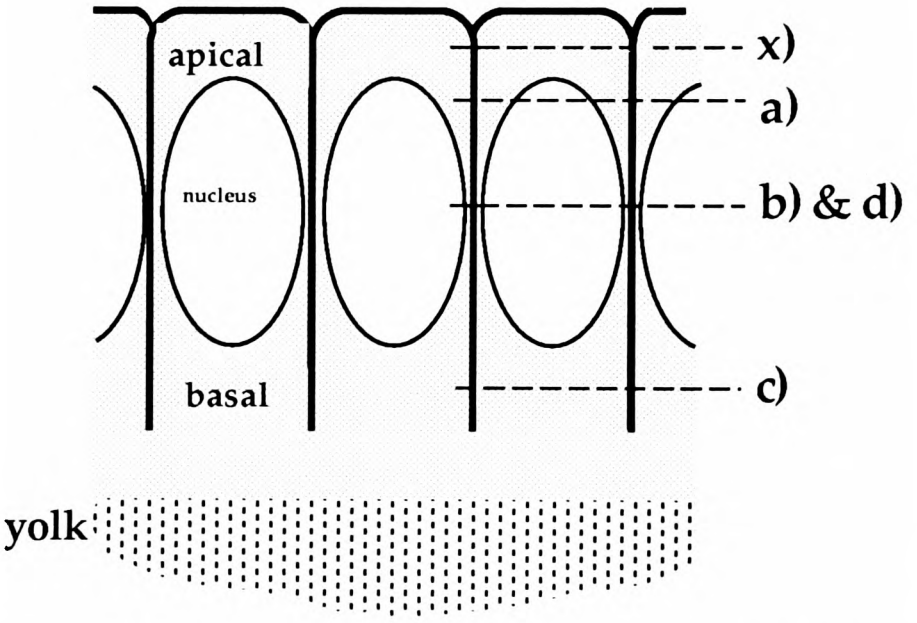


Fig 7.6:
Whole-mount *in situ* hybridisations using DIG-labelled *lacZ* and *globin* probes show the presence of strong punctate nuclear transcripts and very weak cytoplasmic transcripts in constructs containing inverted *eve* and *ftz* localisation sequence (*LTe23INV* & *LTfHSINV* respectively). a) heterozygous *LTfHSINV* embryo probed with a *lacZ* probe, b) *LTfHSINV* with a *globin* probe, c) *LTe23INV* with a *lacZ* probe and d) *LTe23INV* with a *globin* probe. Arrows point to punctate nuclear transcripts (p), nucleus (n), plasma membrane (pm).

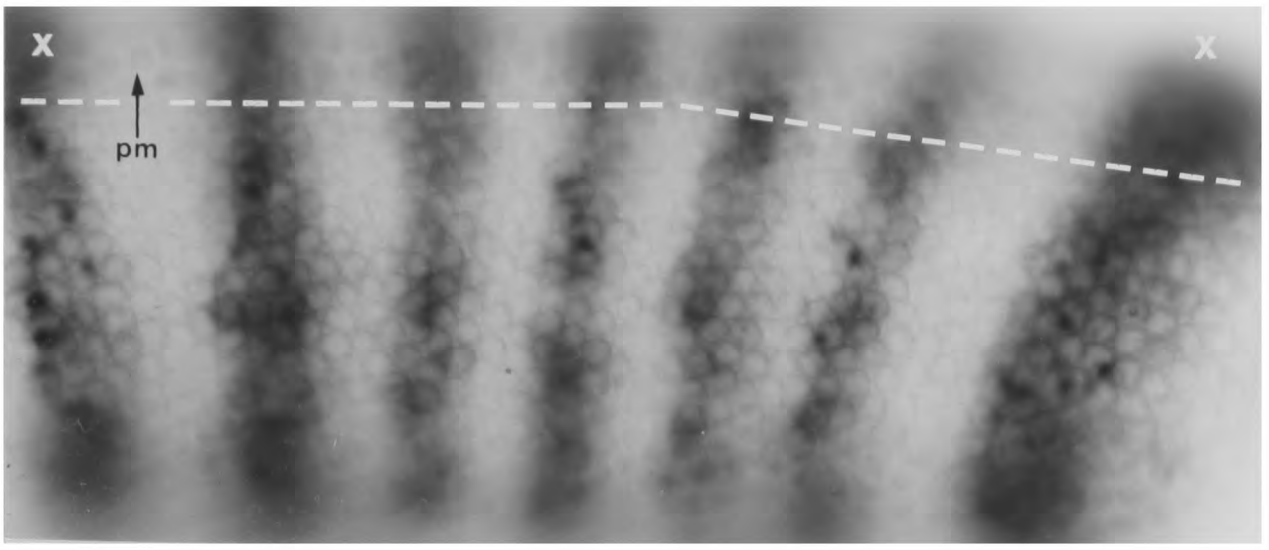
Fig 7.7: (opposite and below)

An oblique view of peripheral nuclei revealing a single punctate region within each nucleus expressing *LTfHSINV*. The punctate region occupies a variable position with respect to the nuclear periphery. The photographs are different optical sections obtained with Nomarski optics of the same heterozygous embryo, shown in Fig. 7.6a, containing a single copy of the construct. a) a section through the apical side of the nucleus. The photograph includes sections through the apical periplasm (x). b) is a section through the centre of the nuclei revealing punctate nuclear message, and c) a section through the basal periplasm below the nuclei. d) an enlarged view of (b): arrows point to nuclear membrane (nm), punctate nuclear transcripts (p) and plasma membrane (pm) of the partially cellularised blastoderm. Nuclear track (t).

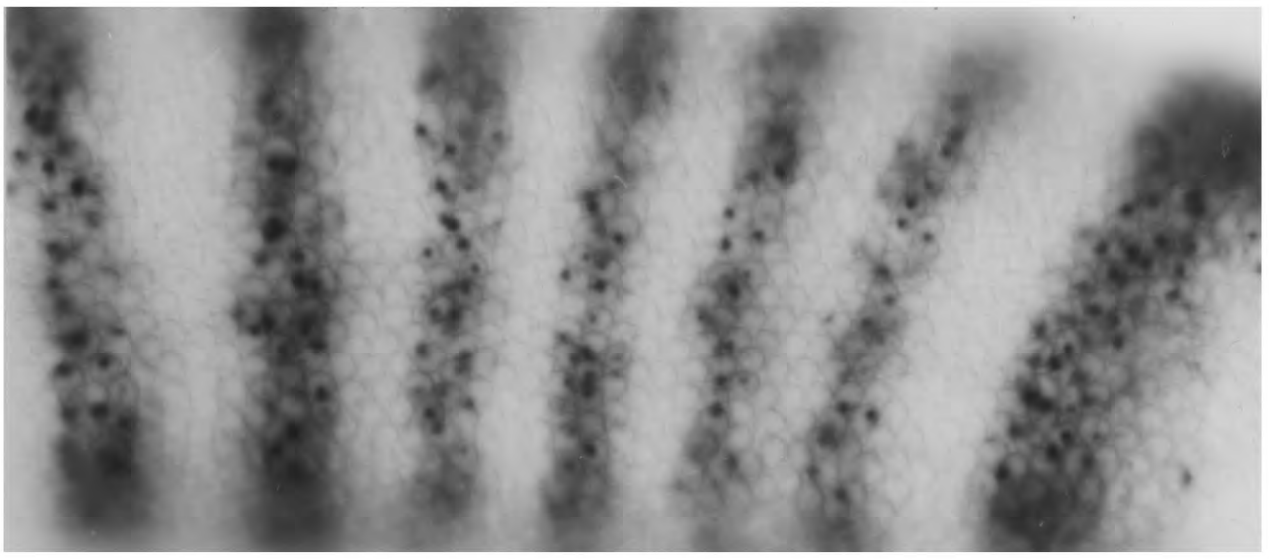


Syncytial blastoderm: position of sections opposite

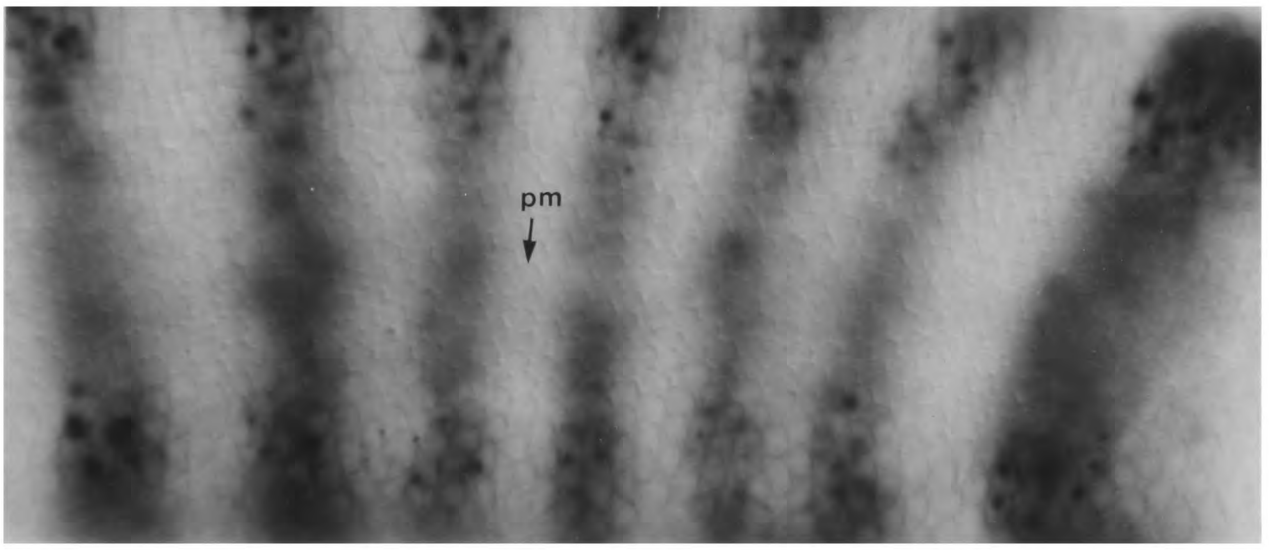
a)



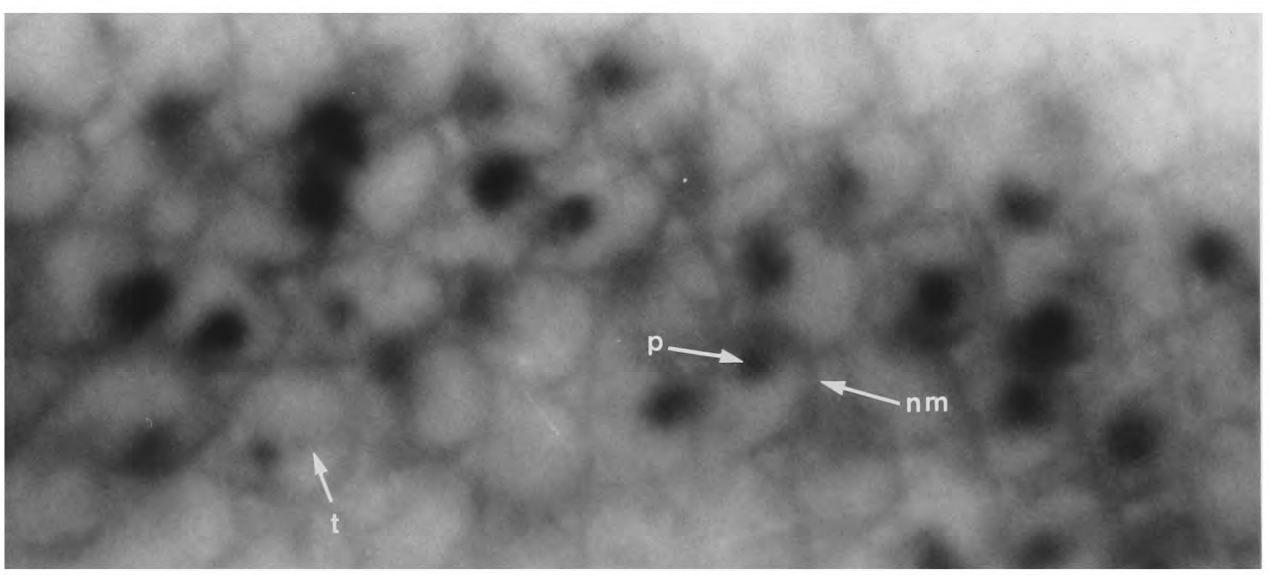
b)



c)



d)



suggest that the hybrid transcript is in some way prevented from reaching the cytoplasm. The punctate nuclear transcripts may be nascent transcripts in the process of being processed and exported.

In situ hybridisation using *globin* probe on *LTfHSINV* reveals weaker punctate nuclear staining (see Fig. 7.6). The signal is expected to be weaker since the DNA fragment used to generate the *globin* probe is 300bp long, compared with the 3kb fragment used to generate the *lacZ* probe.

In situ hybridisation using a *h* intron DIG-labelled probe on wild-type blastoderm embryos reveals pairs of punctate nuclear transcripts as do homozygous embryos for *LTe12 and LTfHSINV* (see Figs. 7.8 & 7.9). These are likely to represent nascent transcripts on the two homologous chromosomes. Similar single punctate regions are found in heterozygous *LTfHSINV* and *LTe23INV* embryos with *lacZ* and *globin* probes (Figs. 7.6). Punctate nuclear staining are also found when *lacZ* and *globin* probes are used on a number of other hybrid constructs (see Figs 6.2 & 6.3).

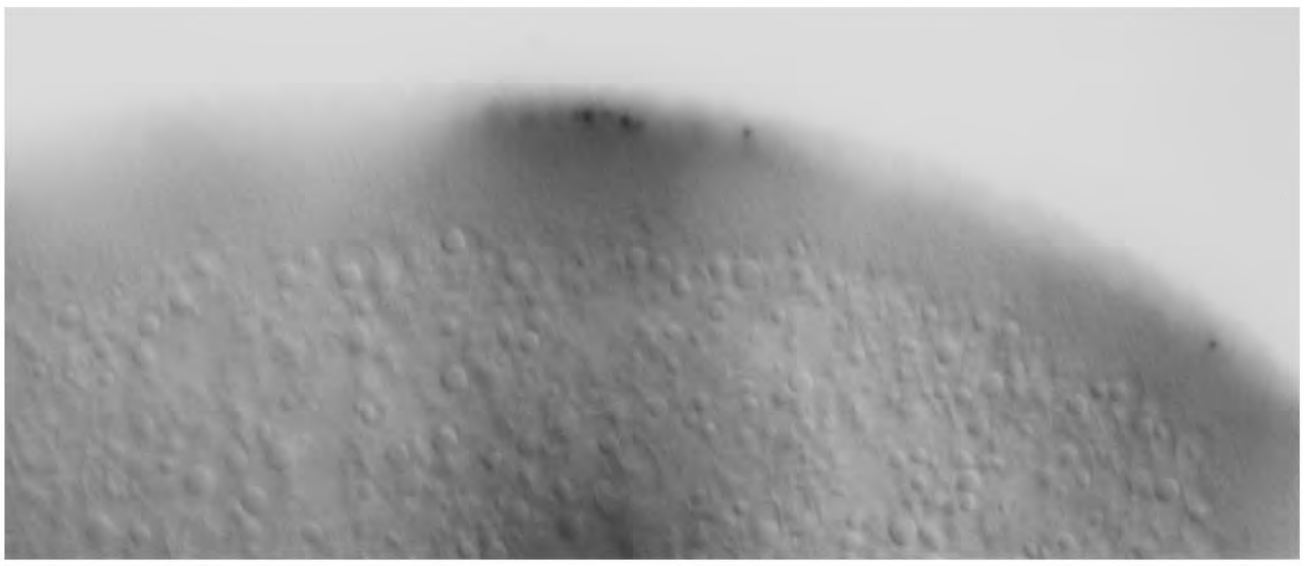
Punctate nuclear transcript is more clearly visible in constructs which have a low level of cytoplasmic staining, but is unrelated to the pattern of intracellular localisation; punctate staining is found both in basally and in apically localised constructs. Furthermore, there is no correlation between the pattern of localisation of a construct and the position of the punctate transcripts in the nucleus; the discrete transcripts are found to be localised in different parts of the nuclei in different constructs (Fig. 7.8).

DISCUSSION

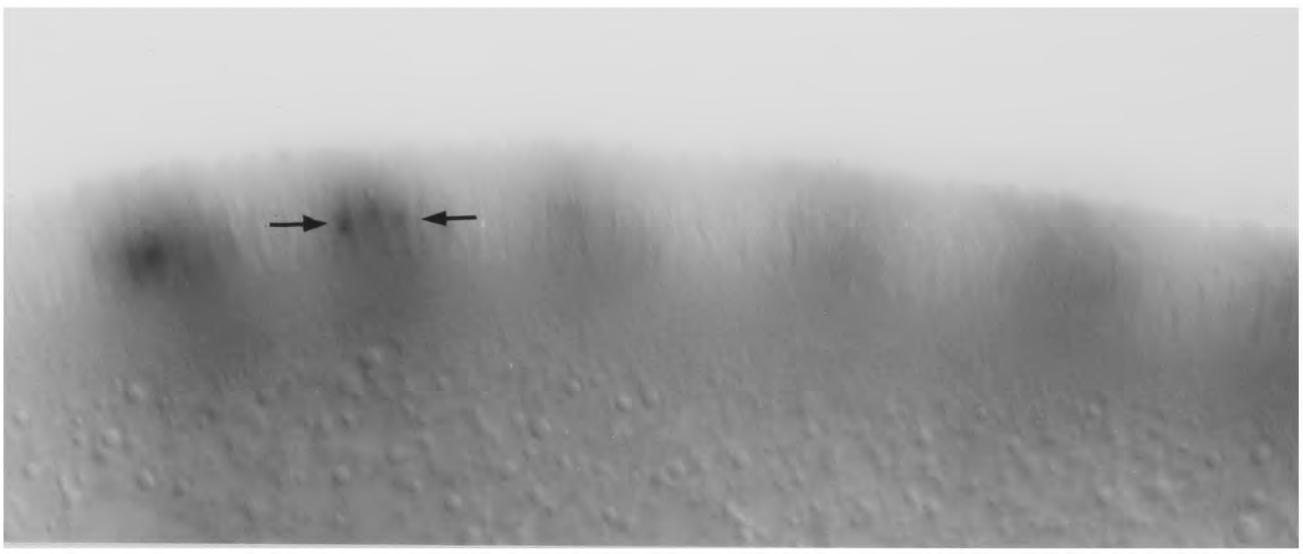
Apical localisation does not depend on cytological location

The chromosomes in peripheral nuclei of the blastoderm embryo appears to adopt the Rabl orientation, in which the chromocentre is situated in the apical side of the nucleus. The strictest version of the gene-gating

a)



b)



c)

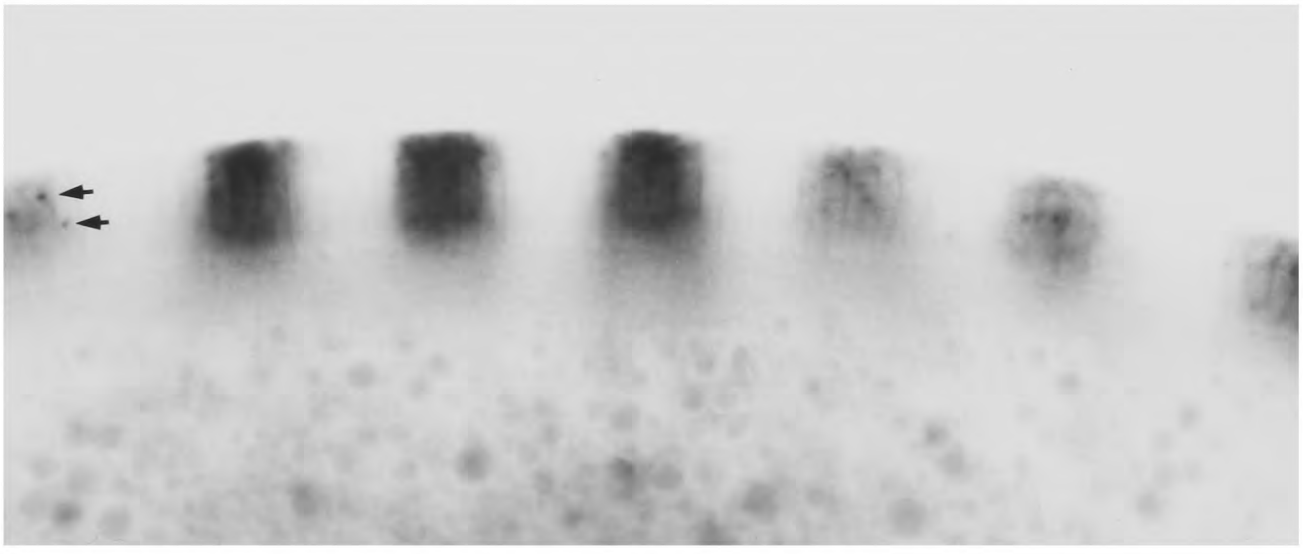
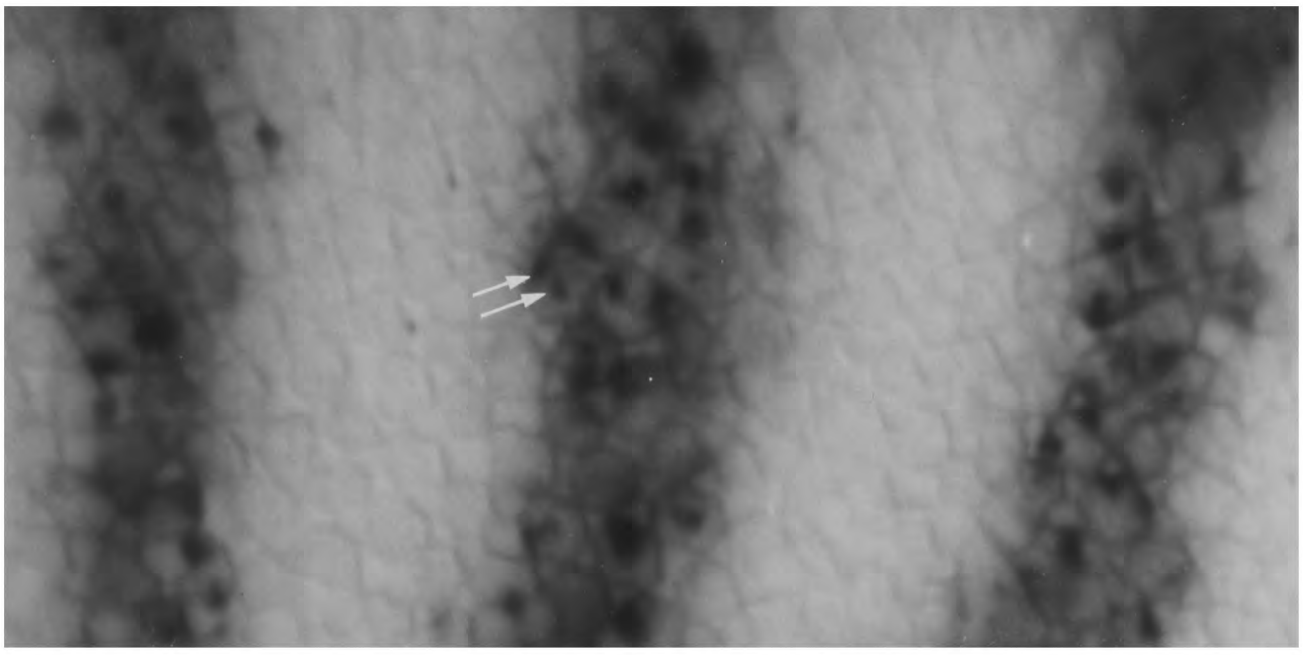
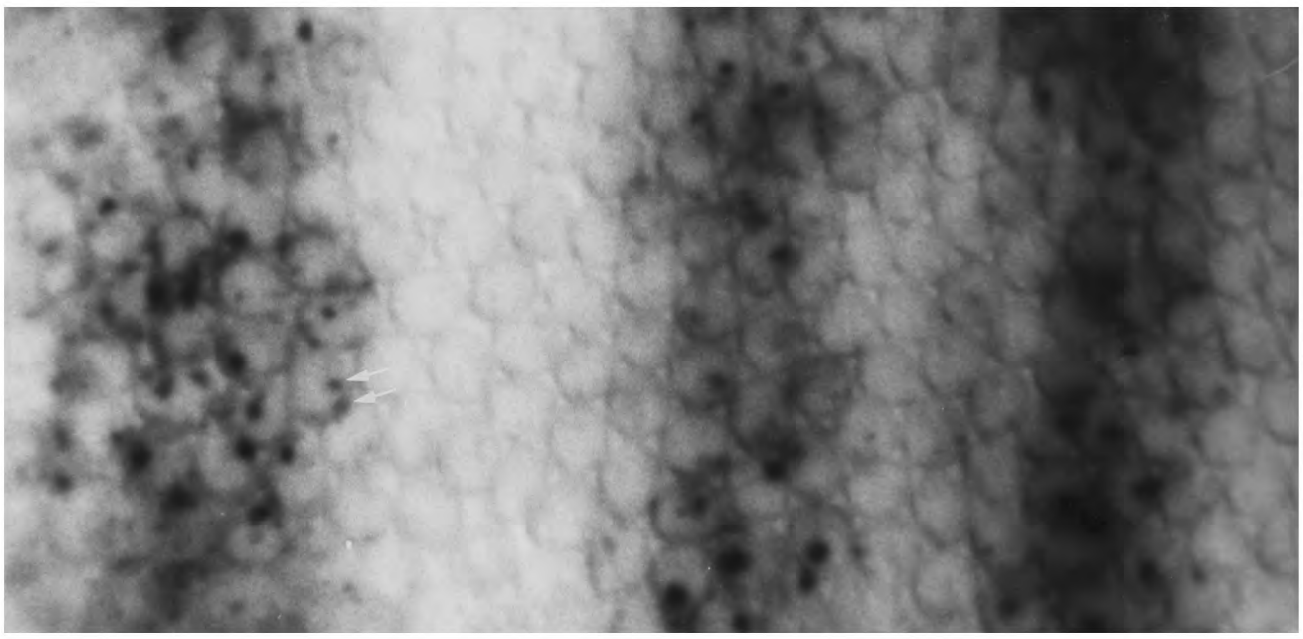


Fig 7.8:
a) & b) *in situ* hybridisation using a *h* intron probe revealing hnRNA pre-processed punctate *h* transcripts. a) at the early syncytial blastoderm when *h* is expressed in a broad region, b) at the cellular blastoderm when *h* is expressed in seven stripes. The punctate transcripts occupy a consistent basal position within the nuclei (see arrows).
c) *in situ* hybridisation using *lacZ* probe on *LTe12* (see chapter 4) revealing punctate nuclear transcripts at a variable position in the nucleus (see arrows).

a)



b)



c)

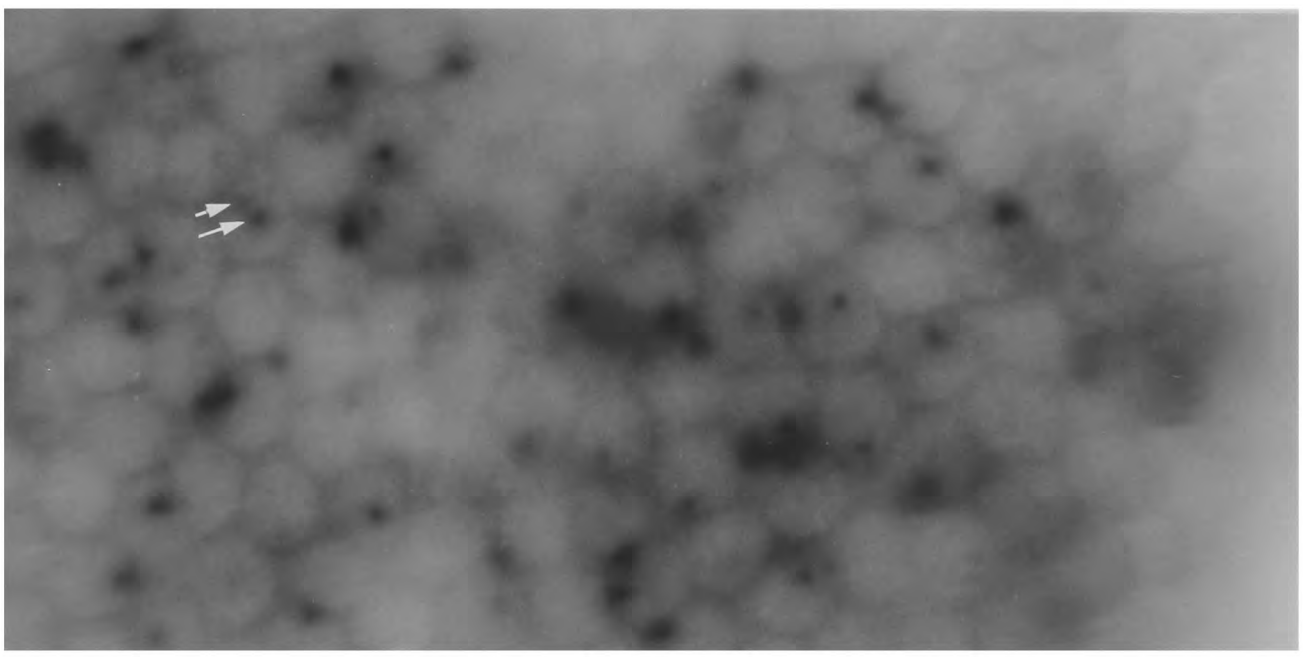


Fig. 7.9
An oblique view of peripheral nuclei, revealing that, in homozygous embryos, two punctate regions are visible within each nucleus (see arrows). a) *LTfHSINV* homozygous embryo probed with *lacZ* . b) *LTe12* homozygous embryo probed with *lacZ*. c) early wild-type syncytial blastoderm with *h* intron probe.

hypothesis therefore predicts that apically localised genes should be situated in the chromocentre so that their transcripts are exported through the apical side of the nuclear membrane.

The cytological position of the apically localised pair-rule genes, and unlocalised gap genes, argues against the importance of cytological position in apical localisation (see the introduction to this chapter). The data I present in this chapter also argue against the involvement of cytological position in apical localisation.

At least two and often more lines were examined for each construct described in this thesis; the lines were always consistent in their pattern of localisation. These included lines containing apically localised constructs, indicating that cytological position is unimportant for apical localisation.

The mechanism for apical localisation depends on transcript encoded sequences

The gene-gating hypothesis predicts that pair-rule localisation is mediated by DNA sequences, which direct the genes to the apical side of the nuclear membrane so that their transcripts are exported through the apical side of the nuclear membrane. However, there are a number of less rigid interpretations of the hypothesis which are compatible with the data presented above, as they do not require that the pair-rule genes lie near pores on the apical side of the nuclear membrane. One such interpretation involves movement of mRNA along nuclear tracks from the site of transcription in the nuclear interior to pores in the apical side of the nuclear membrane. All such interpretations involve a DNA-mediated mechanism for apical localisation.

Three independent enhancer-trap lines, each containing two genes which share the endogenous *h* promoter elements, were analysed by *in situ*

hybridisation. The two genes are not only in the same cytological position, but are also likely to be in the same 'chromatin domain'. In each line the two adjacent genes show a different pattern of localisation. Endogenous *h* transcript remains localised in the apical periplasm, although its level is greatly reduced by the mutation. In contrast, the transcript of the enhancer-trap *lacZ* reporter gene is found unlocalised in a similar manner to *h/lacZ/hsp-27, LThsp* (see chapter 4) and *ftz/lacC* (see chapter 3).

A similar situation to the enhancer-trap experiment is fortuitously created by a large internal deletion, *Df(3R)LIN,* in the ANT-C. The deletion leads to the juxtaposition of the *zen* and *ftz* genes (Rushlow and Levine, 1988). The fusion results in both the *zen* and the *ftz* transcripts being under the control of *ftz* striping elements, with the two genes being transcribed in opposite directions. The *ftz* transcript is localised to the apical periplasm like wild-type *ftz* transcripts, but *zen* mRNA is found in unlocalised (Rushlow and Levine, 1988).

If the gene-gating hypothesis or any other DNA-mediated mechanism for apical localisation is correct, any gene lying close to a pair-rule gene should also become apically localised because of its close proximity to apical localisation sequences. In the *h* enhancer-trap mutants and in the *ftz/zen* mutation, unlocalised genes are in close association with a pair-rule gene whose transcript is localised. However, both the enhancer-trap gene and the *ftz/zen* fusion gene are unlocalised while the endogenous pair-rule genes remain localised. It follows that the apical localisation of pair-rule transcripts does not involve a DNA-mediated mechanism.

A DNA-mediated mechanism for apical localisation predicts that a localisation sequence should function in reverse orientation. However, an inverted localisation signal would lead to the transcription of an antisense RNA which would be unable to function in an RNA-mediated mechanism

for apical localisation. When the fragments of *eve* and *ftz*, necessary and sufficient for localisation, are inserted in reverse orientation into the *LT* vector, the resulting hybrid transcript is unlocalised (see Fig. 7.6). It can be concluded from the enhancer-trap experiments and the inversion of the localisation signal that apical localisation is mediated by a transcript encoded mechanism.

Punctate preprocessed pair-rule transcripts are found in the nuclear interior

Punctate nuclear staining was found in a number of different constructs analysed. This intriguing staining was most prominent in constructs whose cytoplasmic signal was weak, but no correlation was established between the pattern of transcript localisation in the cytoplasm and the presence or absence of the punctate nuclear staining.

In the constructs examined, the punctate staining typically occurs in the nuclear interior and not on the nuclear periphery as proposed by the gene-gating hypothesis. However, the precise site in the nucleus where the discrete staining is found appears to vary between different lines and is not related to the pattern of cellular transcript localisation of the construct (see Fig. 7.8).

Two punctate transcripts were also found in wild-type embryos hybridised with a *h* intron probe, suggesting that the punctate signal represents unprocessed nascent transcripts in the nuclear interior (see Fig. 7.8). When heterozygous embryos containing one copy of a hybrid construct were analysed, only one punctate region is found in the nuclear interior. However, a pair of punctate regions is seen in the nuclei of homozygous embryos containing two copies of a hybrid gene.

Similar punctate nuclear transcripts have also been observed with fluorescently labelled probes for an endogenous gene in a human cell line

(Lawrence *et al.*, 1989). Nuclear Epstein-Barr Virus (EBV) transcript has been observed in an infected human cell line (Lawrence *et al.*, 1989). The viral transcripts are trapped in the nucleus and occur in "curvilinear tracks". Conditions that only detect DNA sequences confirm that the integrated EBV genes are found in the nuclear interior, in a region corresponding to the origins of the tracks. It is speculated that these nuclear tracks represent the path of the viral transcript from its site of transcription within the nucleus to the nuclear membrane (Lawrence *et al.*, 1989).

It follows that the presence of nuclear transcripts in the nuclear interior is a general phenomenon and is not due to experimental artifacts in the system described in this thesis. Like the EBV transcript, the punctate nuclear RNA which I observe is found in the nuclear interior. While I fail to detect strong nuclear tracks, extremely weak lines are visible in *LTfHSINV* nuclei which could represent curvilinear tracks (marked as t in Fig. 7.7). A probable explanation is that the EBV transcripts are prevented from being exported from the nucleus so a non-physiological build-up of transcript occurs along the track. In normal circumstances the transcript may not be easily detected, as it may be rapidly transported along the tracks to the nuclear membrane.

The transcript-gating hypothesis

In what follows, I propose an alternative hypothesis to the gene-gating hypothesis which is compatible with the data presented so far, and makes a variety of testable predictions.

The transcript-gating hypothesis proposes a transcript mediated mechanism for the transport and export of transcripts from the nucleus of any eukaryotic cell. It states that RNA sequences, usually present in the 3' untranslated part of mature mRNA, direct the transport of transcripts along

tracks in the nucleoplasm to pores in the nuclear membrane. Only transcripts containing these sequences can be exported out of the nucleus. The precise position of transcription of a gene within a nucleus is unimportant, but which pores are used for the export of its transcript is important for certain classes of genes.

When applied to intracellular transcript localisation in the *Drosophila* blastoderm the hypothesis is interpreted as follows. Different classes of zygotic genes are exported through different types of pores in the nuclear membrane. Pair-rule transcripts are exported from pores on the apical side of the nuclear membrane, transcripts containing the *globin* 3' untranslated end are exported through pores in the basal side of the nuclear membrane and gap gene transcripts are exported from pores ubiquitously distributed in the nuclear membrane. After nuclear export, the transcripts could either be anchored to a cytoplasmic component, or could be prevented from further movement by a diffusion barrier. Alternatively the transcripts may be actively transported in the cytoplasm. In Chapter 9, I discuss some experimental strategies which could be used to test predictions made by the transcript-gating hypothesis.

The thesis is concerned with three basic questions: defining sequences necessary and sufficient for localisation, the mechanism for localisation and the function of localisation. So far, only the first two questions have been addressed in any detail. In the next chapter, I consider whether the intracellular localisation of transcripts may have an affect on the intracellular localisation of the protein they encode.

CHAPTER 8: THE FUNCTION OF APICAL TRANSCRIPT LOCALISATION.

INTRODUCTION

The existence of three distinct types of subcellular transcript distributions which correspond well with different classes of segmentation genes suggests that the specific localisation of pair-rule mRNA may be functionally important. Different segmentation genes encode different types of proteins. Although not formally demonstrated in some cases, the pair-rule genes, gap genes and some segment polarity genes, namely *en* & *gsb*, are all nuclear transcription factors. *ptc* encodes a transmembrane protein and *wg* encodes a secreted protein (see chapter 1). The trafficking of cytoplasmic proteins in unicellular eukaryotic cells is well understood. However, the *Drosophila* blastoderm embryo is a giant syncytium in which protein trafficking may be very different from the unicellular cases.

There are two main pathways for proteins in the cell (Alberts *et al.*, 1989, chapter 8). The first is taken by secreted proteins, Golgi proteins and endosomal proteins, and involves ribosomes translating the proteins in the rough endoplasmic reticulum before transport of the proteins through the Golgi apparatus (Kelly, 1985). The second pathway is taken by cytosolic, mitochondrial and nuclear proteins: they are translated on free ribosomes before passage to the appropriate subcompartment.

The uptake of nuclear proteins involves two stages. The first is the recognition of a nuclear localisation signal (nuclear tag) and binding to the outside of the nuclear membrane. The second step is ATP dependent transport of the protein through the NPCs (Richardson *et al.*, 1988). It has been suggested that the first step may involve initial binding to a cytoplasmic protein (Breeuwer and Goldfarb, 1990).

The general features of intracellular protein trafficking are common to all eukaryotic cells, so the function of mRNA localisation in other systems may shed light on the function of intracellular localisation in *Drosophila*.

The myoblast cell is a giant syncytium, containing hundreds of nuclei distributed evenly in the cytoplasm and therefore is similar, in this respect, to the syncytial blastoderm embryo. The mRNA transcribed from each nucleus remains in the cytoplasm that is close to that nucleus (Hall and Ralston, 1989) and the proteins they encode are in many cases restricted to the region of cytoplasm where they are required to function (Merlie and Sanes, 1985; Ralston and Hall, 1989). Asymmetric intracellular transcript localisation is also found in fibroblasts which are very small cells in comparison to myoblasts. actin mRNA is localised near the lamellipodia of fibroblasts (Lawrence and Singer, 1986), where actin protein is translated and polymerised to facilitate changes in cell shape (see chapter 1).

The silkmoth blastoderm embryo, *Hyalophora cecropia*, contains maternal actin mRNA associated with cortical cytoskeleton in a region equivalent to the apical periplasm of the *Drosophila* blastoderm. It has been suggested that the mRNA is translated at the site of localisation of the actin protein in a similar manner as the fibroblast cells (Kastern *et al.*, 1990).

In myoblasts, fibroblasts and the silkmoth blastoderm embryo, mRNA is found in the subcellular compartment where the protein encoded by the mRNA is required to function. This provides a functional explanation for the asymmetric subcellular localisation of mRNA encoding cytoplasmic protein.

A similar situation may occur in *Drosophila*, with transcripts being localised at subcellular sites where the cytoplasmic proteins encoded by the transcripts function. One such example is that of *crumbs* mRNA which is apically

localised in blastoderm cells and in epithelium. *crumbs* encodes an EGF-like transmembrane protein which is exclusively localised to the apical membrane of epithelial cells and of gastrulating embryos (Tepass *et al.*, 1990). Another example is provided by *serendipity alpha (ser-α),* whose transcript is localised to the apical periplasm (Schweisguth *et al.*, 1989). The *ser-α* protein is an apically localised membrane-associated protein (Schweisguth *et al.*, 1990).

In the *Drosophila* blastoderm, gap gene mRNAs are found unlocalised within their expression domain. Pair-rule mRNA is found exclusively in the apical periplasm within each of the seven stripes (see chapter 1). Both the pair-rule genes and the gap genes encode nuclear proteins so the distribution of their transcripts cannot be related in the simple way described above to the cytoplasmic distribution of their proteins. However, there is one aspect of the distribution of the proteins of gap and pair-rule genes which is dramatically different.

The protein expression domains of the pair-rule genes correspond fairly precisely with their transcription domains, because little if any diffusion takes place. In contrast, gap gene protein expression is wider than the mRNA and deletion domains. It has been suggested that gap gene proteins can diffuse substantial distances away from their site of translation.

A candidate function for the apical localisation of pair-rule transcripts is to limit the diffusion of the pair-rule proteins. If pair-rule proteins were translated only in the apical periplasm, they would only be able to enter the nuclei which transcribed the transcript encoding the protein. On the other hand, gap gene proteins are translated in all parts of the periplasm and are therefore able to diffuse substantial distances in the basal periplasm before entering the nucleus.

The most satisfactory way to test the functional importance of apical localisation would be to make a construct which expresses an unlocalised pair-rule transcript. This would involve replacing the 3' end of an otherwise functional pair-rule gene with sequences lacking an apical localisation signal. However, this approach requires a detailed knowledge of the size and nature of the localisation signal and only became conceivable after the completion of the experiments described in chapters 4,5 & 6. I therefore adopted an indirect test to ask if cytoplasmic proteins could be targeted to different parts of the periplasm by altering the distribution of the transcripts encoding them.

STRATEGY

If apical localisation of mRNA limits the diffusion of the protein it encodes, then the pattern of mRNA localisation in the various constructs described in the previous chapters should affect the pattern of *lacZ* protein localisation within the cytoplasm. Constructs whose mRNA is localised to the apical periplasm should give rise to apically localised *lacZ* protein; constructs whose mRNA is unlocalised should encode unlocalised *lacZ* protein. I determined the intracellular distribution of cytoplasmic *lacZ* protein in a number of constructs described in previous chapters, in order to settle this question.

RESULTS:

FG2 codes for a nuclear *lacZ/ftz* fusion protein

FG2 has the upstream region of *ftz* controlling the expression of the *lacZ* gene so that the *lacZ* transcript and protein are found in a seven striped pattern characteristic of *ftz*. In chapter 3 it was shown that FG2 transcript is localised to the apical periplasm (see chapter 3).

Antibody staining with an anti *lacZ* polyclonal antibody (see chapter 2 for details) reveals that the *lacZ* protein is predominantly found in the nuclei of late blastoderm embryos (see Fig. 8.1). One explanation could be that apical mRNA localisation directs the localisation of the *FG2* protein to the nucleus. However, I sequenced the fusion point between the two genes, revealing that two thirds of the *ftz* coding region is fused in frame with the *lacZ* protein (see Fig. 8.1). A presumptive *ftz* nuclear tag is present in the 3' end of the *ftz* gene and is included in the *ftz/lacZ* fusion. The most likely explanation for the nuclear localisation of the fusion protein is the presence of a nuclear tag and not the apical localisation of *FG2* mRNA. Therefore, I examined the pattern of protein distribution in other fusion constructs encoding the *lacZ* protein and lacking a nuclear tag.

Unlocalised transcripts lead to unlocalised cytoplasmic proteins

ftz/lacC is similar to *FG2* except that it has an *hsp70* instead of a *ftz* 3' end. In chapter 3 it was shown that *ftz/lacC* transcript is unlocalised.

Unlike the *FG2* fusion protein, *ftz/lacC* encodes the normal *lacZ* gene with no additional amino acids fused to its 3' end. Staining with an antibody against *lacZ* protein reveals that, like the transcript, *lacZ* is unlocalised in the periplasm of the cellular blastoderm (see Fig. 8.2). The protein cannot be detected in early stage 14 blastoderm embryos, when *ftz* protein is first detected. A probable explanation is that the bacterial *lacZ* gene contains a different pattern of codon usage than a typical *Drosophila* gene and is likely to be translated more slowly. Also, the protein is distributed all over the cytoplasm and is therefore more difficult to detect than the nuclear *ftz* protein.

I confirmed the results by sectioning the whole-mount embryos after embedding in methacrylate resin. The sections indeed show that *lacZ*

a)

CAT TAC CAG TTG GTC TGG TGT CGG GGA TCC GTC GAC TAC TTG GAC GTC TAC TCG CCC CAG

BAMHI SALI

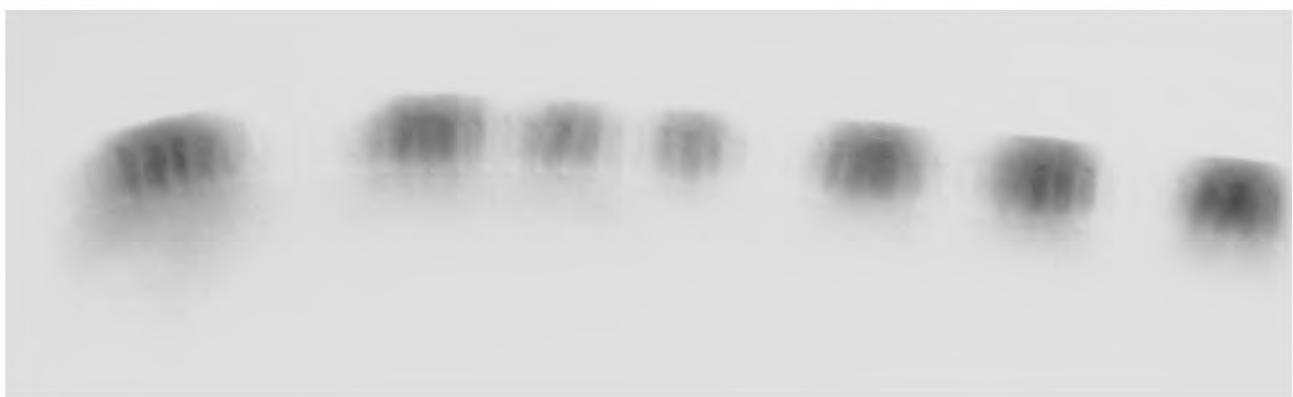
3' end of lacZ ftz open reading frame

Fig. 8.1 (above and opposite)

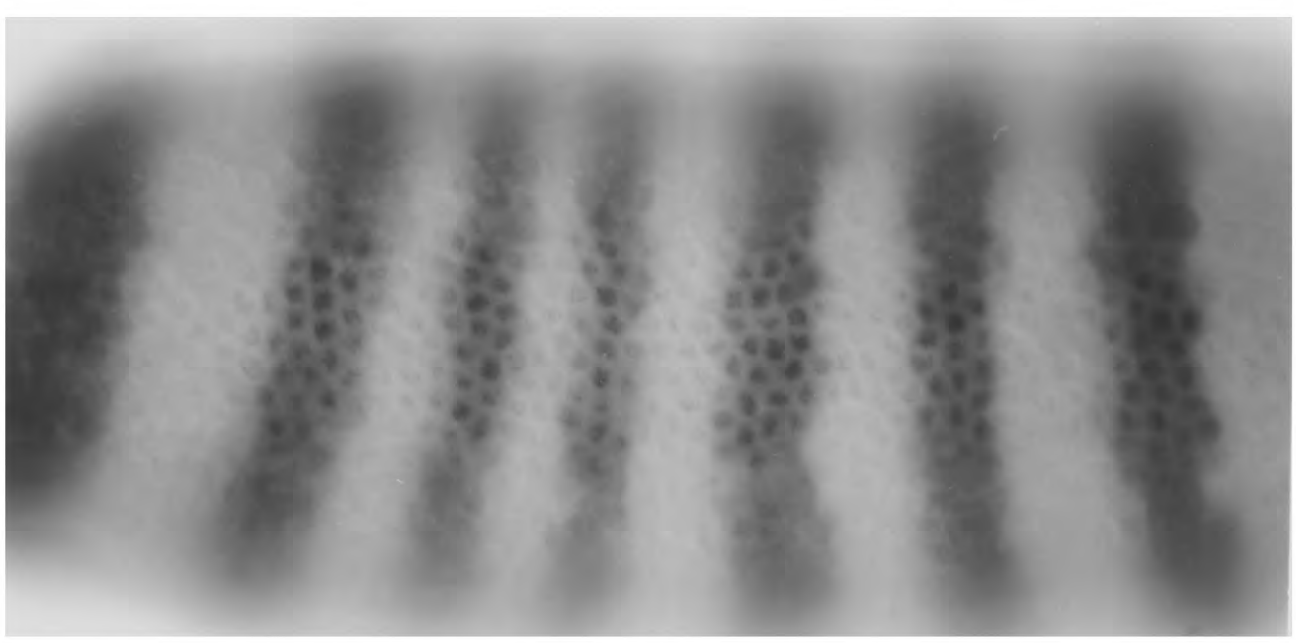
a) The sequence of the *FG2* construct showing that *ftz* is fused in frame with *lacZ*. The *lacZ* gene is missing the termination codons in the 3' end and is fused to *ftz* at the SalI site. The fusion protein contains the *ftz* nuclear tag.

b), c), d) & e) Whole mount localisation of *lacZ/ftz* fusion protein of *FG2* and the *ftz* protein showing nuclear localisation of proteins. b) & c) *ftz* protein seen in side view and glancing section respectively. d) and e) *lacZ/ftz* protein seen in side view and glancing section respectively.

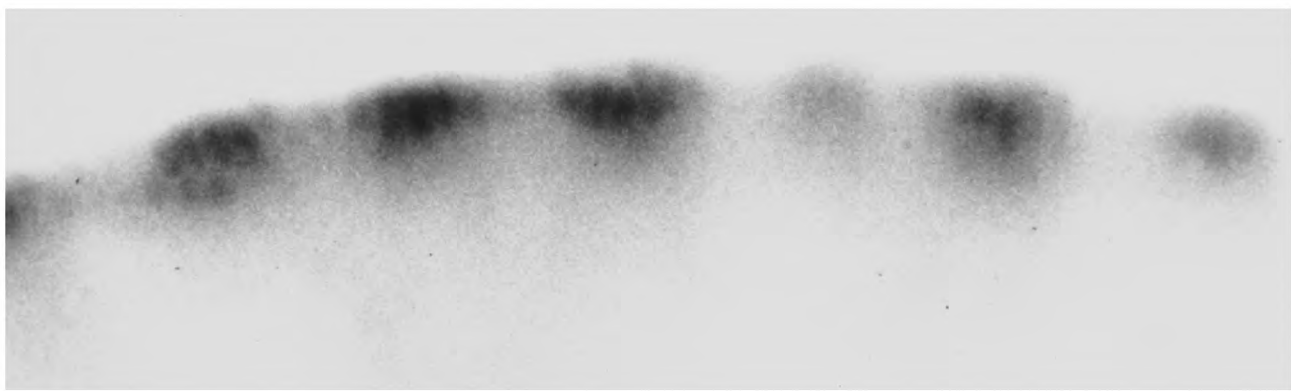
b)



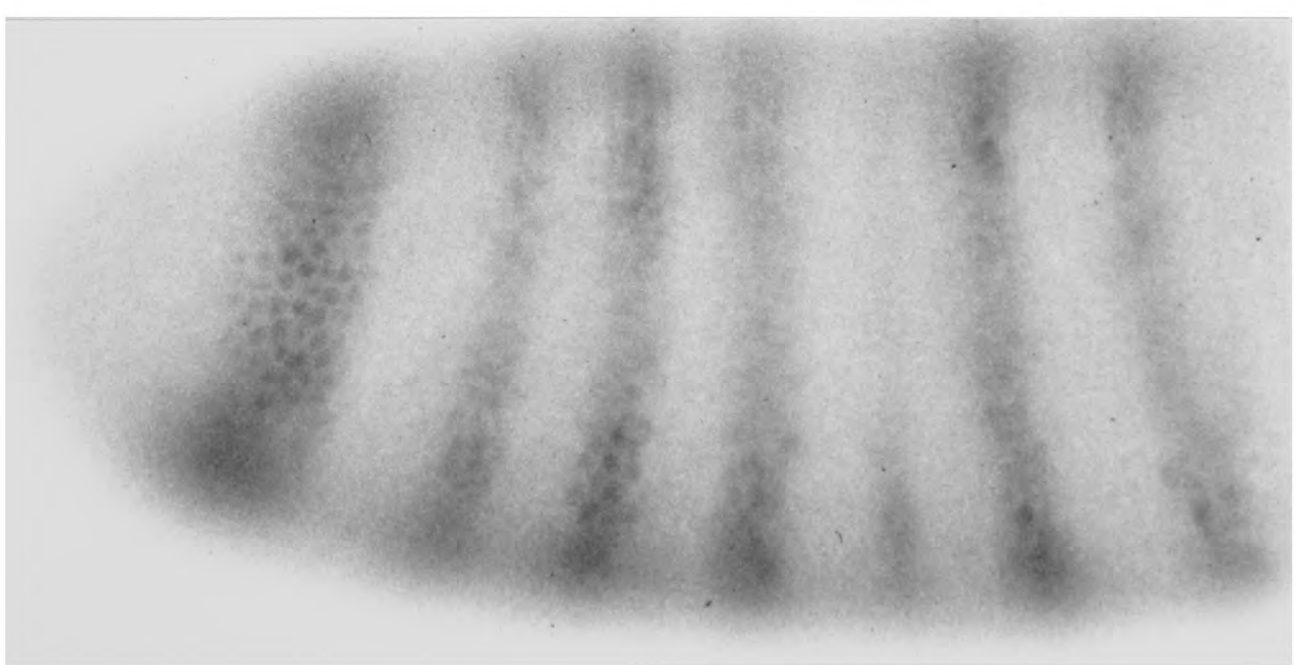
c)



d)

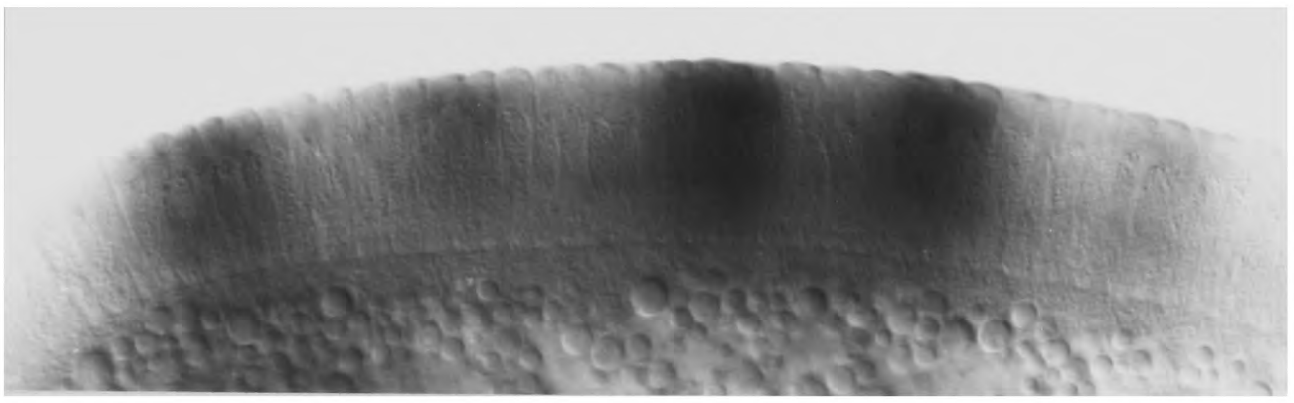


e)



a)

b)



c)

d)

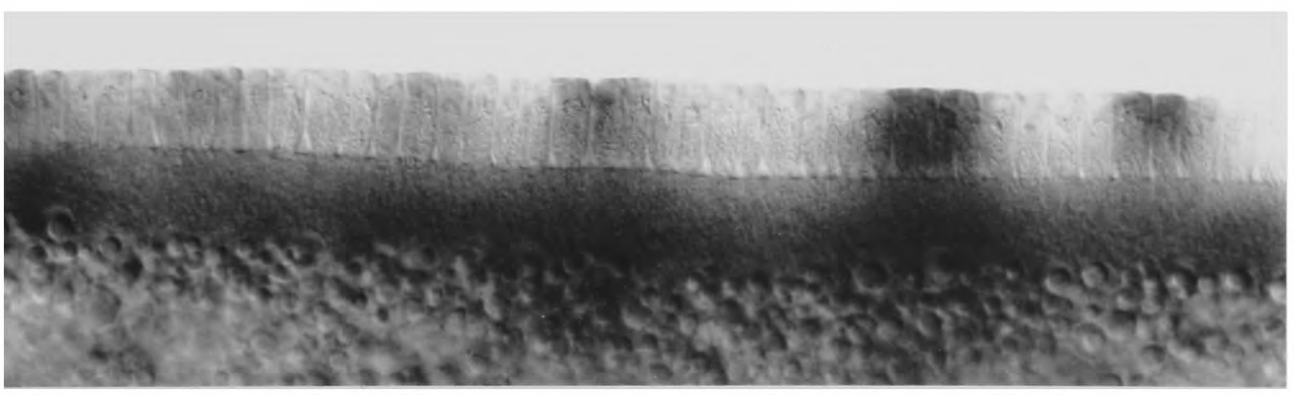


Fig. 8.2

Whole mount localisation of cytoplasmic *lacZ* protein of fusion constructs expressed at the late blastoderm stage, suggesting that constructs with apically localised mRNA, express apically localised cytoplasmic protein. a) *LTe32*, b) *ftz/lacC*, c) *LTfSH* and d) *LTbcd*. (see chapters 3, 4 and 5 for details of the constructs)

protein is unlocalised in *ftz/lacC* periplasm (see Fig. 8.3). The result indicates that β-galactosidase is not inherently localised to any particular part of the cell: it is not nuclear and does not localise to any particular part of the cytoplasm.

Apically localised transcripts lead to apically localised proteins

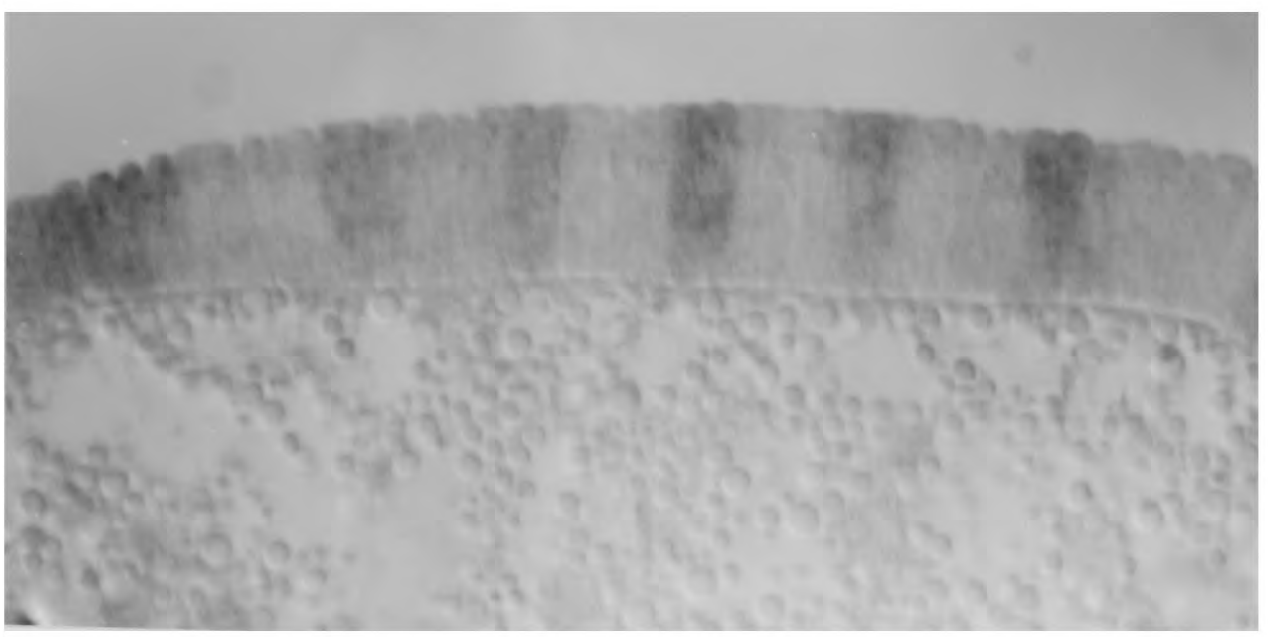
LTfSH contains the same SalI to HindIII fragment of the 3' end of *ftz* fused with *lacZ* in *FG2* inserted into the *LT* vector. However, *LTfSH* contains a termination codon after the *lacZ* gene so that the fusion protein does not contain a nuclear tag from the 3' of the *ftz* protein. In chapter 4, it was shown that *LTfSH* mRNA is found predominantly in the apical periplasm.

Staining *LTfSH* embryos with antibody against *lacZ* reveals that the protein is found in the apical periplasm of late stage 14 blastoderm cells (see Fig. 8.2). The *LTfSH* protein cannot be detected in early stage 14 blastoderm embryos.

Two other constructs previously analysed by *in situ* hybridisation, whose transcripts are apically localised were analysed with antibody against *lacZ*. The constructs were *LTe32* which contains a small region of the 3' untranslated end of the *eve* gene necessary and sufficient for apical localisation inserted into the *LT* vector (see chapter 4) and *LTbcd* which contains the 3' untranslated end of the *bcd* gene inserted into the *LT* vector (see chapter 5).

Staining *LTe32* and *LTbcd* embryos with antibody against *lacZ* reveals that the *lacZ* protein is found predominantly in the apical periplasm of late stage 14 embryos (see Fig. 8.2). Sections of whole-mount embryos stained with the antibody confirm that *LTe32* protein is apically localised (see Fig. 8.3). Like the other constructs, *LTe32* and *LTbcd* protein cannot be detected in early stage 14 blastoderm embryos.

a)



b)

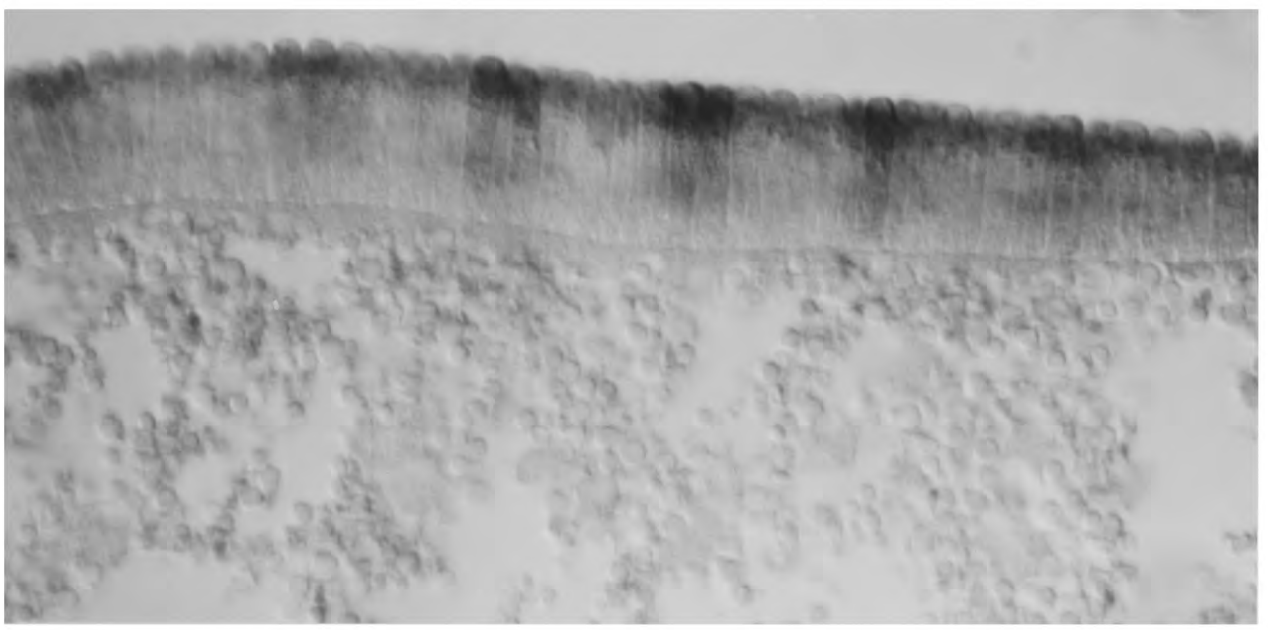


Fig. 8.3

The localisation of the *lacZ* protein in a) *ftz/lacC* and b) *LTe32*, by methacrylate sections of late blastoderm embryos (see chapter 5 for details of the construct). The results confirm those seen in Fig 8.1, in whole-mount embryos.

Thus, constructs with apically localised transcripts encode apically localised cytoplasmic protein. In contrast, *ftz/lacC* transcript and cytoplasmic protein are unlocalised in the blastoderm embryo. The results suggest that a diffusion barrier exists between the apical and basal periplasm which prevents cytosolic proteins from moving from the apical to the basal periplasm. It follows that the nuclear proteins encoded by apically localised pair-rule transcripts are more likely to enter the nucleus adjacent to the apical periplasm than protein encoded by unlocalised genes.

DISCUSSION:

The currently accepted view of protein trafficking within cells states that the presence of a nuclear tag determines whether a protein enters the nucleus. A nuclear protein is translated in the cytoplasm and is believed to diffuse randomly until it encounters an NPC or cytoplasmic factor involved in its uptake. In normal cellular circumstances, the intracellular localisation of a gene encoding a nuclear protein is unimportant since the protein is imported into the single nucleus. However, the *Drosophila* blastoderm is a syncytium containing thousands of peripheral nuclei sharing the same cytoplasm. Under these unusual circumstances, the intracellular localisation of a transcript can influence the extent of diffusion of the nuclear protein it encodes.

Both the gap genes and the pair-rule genes encode nuclear proteins of approximately the same molecular weight, but there is a marked difference in the degree of diffusion of the two proteins in the blastoderm embryo. The pattern of pair-rule protein expression accurately resembles their pattern of transcription and the regions deleted in mutant embryos (reviewed in Ingham, 1988). However, gap gene proteins are detected in a substantially broader region than their pattern of transcription and are

expected to be as broad as the pattern of cuticle deletions in mutant embryos (Gaul and Jäckle, 1989; Pankratz *et al.*, 1989; Stanojevic *et al.*, 1989).

In appendix A, I estimate that a large protein diffusing in free cytoplasm would diffuse several nuclear diameters in a very short space of time. It follows that something must prevent the pair-rule proteins from diffusing beyond their transcriptional domain, while allowing gap proteins to diffuse. It is significant that gap gene transcripts are all unlocalised, while the pair-rule genes are apically localised in the syncytial blastoderm. Although the correlation between the pattern of intracellular transcript localisation and extent of protein diffusion may be fortuitous, it seems more likely to be functionally significance.

By late stage 14, the nuclei have elongated and membranes have moved to divide the syncytium into distinct regions of cytoplasm associated with each nucleus. At this late blastoderm stage, membranes prevent the lateral diffusion of protein. However, pair-rule protein can be detected earlier, when protein is able to diffuse laterally in the basal periplasm. Proteins are probably prevented from diffusing laterally in the apical periplasm by the presence of lateral constrictions of the plasma membrane. Pair-rule proteins are therefore only able to enter the adjacent nucleus, a phenomenon which I refer to as protein-targeting (see Fig. 8.4).

Gap genes are first detected at an earlier stage than the pair-rule genes; their proteins are first detected at cell cycle 10-11, 2-3 cycles before the onset of cellularisation. Gap gene transcripts are unlocalised so that gap proteins translated in the basal periplasm are able to diffuse away from the nuclei that expressed them.

The data presented in this chapter strongly support the idea of protein-targeting. *FG2* mRNA is found predominantly in the apical cytoplasm and the protein is found in the nucleus because of the presence of a nuclear tag

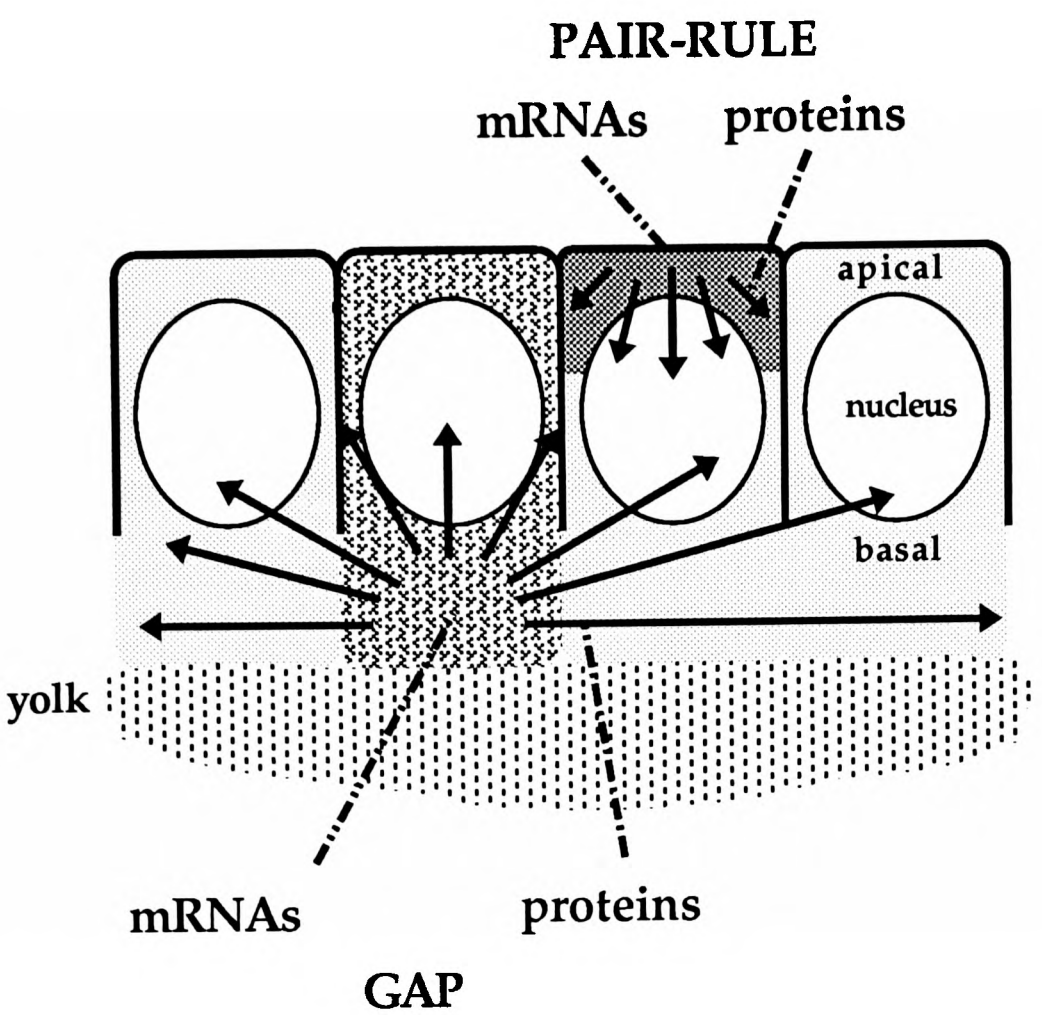


Fig 8.4: Protein targetting

**A model explaining how the apical localisation of pair-rule transcrtipts targets the pair-rule proteins into the adjacent nucleus.
Gap gene transcripts are unlocalised, leading to a graded distribution of nuclear proteins. The arrows represent protein diffusion in the cytoplasm before entry into the nucleus. The apical and basal periplasms are labelled as apical and basal, respectively.**

in the fusion protein. The transcript of another construct, *h/lacZ/hsp27* is unlocalised (see chapter 3) and yet the *lacZ* protein it encodes is found in the nucleus because of the presence of a nuclear tag (Riddihough personal communication). It follows that the apical localisation of pair-rule genes does not determine whether a protein is nuclear or cytoplasmic.

In other constructs described above, the proteins are found in the cytoplasm because they lack nuclear tags. In these constructs the patterns of localisation of the mRNAs influence the intracellular distributions of the cytoplasmic proteins they encode. In the case of constructs whose mRNAs are localised in the apical periplasm, the proteins are also found in the apical cytoplasm. In constructs whose mRNA is unlocalised, the protein is also found in both subcompartments.

To test directly the function of apical localisation, a construct must be made which is unlocalised but encodes a functional pair-rule protein. The construct would have to be expressed in an identical pattern to the wild type pair-rule gene. The ability of the construct to rescue pair-rule function and the degree of diffusion of protein away from the transcriptional domain would then be analysed to determine the function of apical localisation. This approach is discussed in further detail in the next chapter.

CHAPTER 9: DISCUSSION AND FUTURE PERSPECTIVES

I have defined the sequences required for apical localisation of pair-rule transcripts in the syncytial blastoderm. The mechanism of localisation was found to involve RNA-mediated rather than DNA-mediated processes; transcript-gating was proposed as the mechanism for localisation. Finally, it was shown that apically localised transcripts encoding cytoplasmic protein give rise to apically localised proteins. It is proposed that the function of apical transcript localisation is to limit the diffusion of pair-rule proteins before entry into the nucleus (protein-targeting).

Defining the localisation signal:

Apical localisation signals lie in the 3' untranslated part of *eve*, *ftz*, *h* & *α1-tubulin* (chapter 3), while the same region of *β-globin* includes a basal localisation signal (chapter 4). In other well-characterised cases of asymmetric transcript localisation, the 3' untranslated part of the gene also contains localisation signals. Such genes include *bcd* (Macdonald and Struhl, 1988) and *Vg1* (Yisraeli *et al.*, 1989), whose transcripts are maternally transcribed and are localised in one part of the oocyte.

124bp in the *eve* 3' end contains apical localisation signals (chapter 4). The *eve* localisation signal could not be divided into two overlapping halves, without disrupting its ability to confer localisation. Further experiments would have to be performed to determine the minimum size of the *eve* localisation signal. Small deletions from either ends of *e32* would have to be tested for their localisation potential.

The *eve* apical localisation signal is smaller than the 630bp region proposed as the *bcd* localisation signal (Macdonald and Struhl, 1988). On the other hand, the results indicate that the localisation signal is likely to be larger than the 6-20bp postulated as RNA-recognition sites of the components of

the polyadenylation machinery (Levitt *et al.*, 1989). Instead the localisation signal appears to be a region of RNA approximately 100bp in length and probably acts by forming a particular secondary structure.

The RNA secondary structure of *e32* was predicted using both sub-optimal and optimal folding algorithms (see appendix A). The Optimal folding algorithm gives the secondary structure with the lowest free energy and represents an RNA species not bound to other molecules which alter its folding. Sub-optimal folding gives a more realistic secondary structure including a number of possible motifs of higher energy than the optimum. Both algorithms were used to predict the secondary structure of *e32* and of small regions near the polyadenylation signal of *ftz* and *h*. However, specific secondary structural motifs found for the three pair-rule genes were also present in *hb*.

The comparisons of secondary structures could be done with far greater confidence if *ftz* and *h* localisation signals were more precisely defined. 1.2kb fragment of the 3' end of *ftz* and a 700bp fragment of the 3'end of *h* contain an apical localisation signal. The first experiments that should be done to extend those described in this thesis should define a much smaller region of *ftz* and *h* which include apical localisation signals. This could be done by inserting small fragments of the 3' untranslated part of *ftz* and *h* into the *LT* vector to assay their localisation potential in a similar manner to that described in chapter 4. The secondary structure of the small regions of *ftz* and *h* could then be compared with *e32* and the three fragments could be used to isolate molecules involved in apical localisation (see below).

RNA secondary structure predictions are only valuable in that they provide a number of structures which can be tested experimentally. Constructs could be made with mutated localisation sequences to test if disruptions of the predicted secondary structure affects localisation.

Transcript-gating

Compelling evidence rules out the involvement of transcript stability in apical localisation. Cytoplasmic models involving the movement of RNA from the basal to the apical periplasm (see chapter 1) are also implausible because the apical and basal periplasms are separated by a diffusion barrier (see chapters 5 and 8).

LTbcd includes the 3' untranslated part of *bcd* required for the localisation of *bcd* mRNA inserted into the *LT* vector. *LTbcd* transcript is apically localised (see chapter 5). However, endogenous *bcd* transcript is not localised in the apical periplasm, despite the presence of the same sequences which cause the localisation of *LTbcd* (Berleth *et al.*, 1988; St Johnston *et al.*, 1989). The most likely explanation for these observations is that *bcd* mRNA is maternally transcribed but *LTbcd* is transcribed zygotically; the mechanism of pair-rule localisation involves processes which must occur in the nucleus.

Hybrid constructs could be made in order to show conclusively that the mechanism of localisation is nuclear. Two constructs could be made which contain different upstream regions controlling the expression of identical transcripts encoding the *lacZ* protein. The upstream region of the first construct would cause its zygotic expression, and that of the second construct would lead to maternal expression at the same stage. If apical localisation depends on a cytoplasmic mechanism, both the maternal and zygotic constructs would localise. If apical localisation depends on a nuclear mechanism, only the zygotic construct would be localised.

In chapter 7, it was shown that apical localisation could not occur by a mechanism involving DNA sequences or chromosome orientation in the nucleus (e.g. gene-gating). Transcript-gating was proposed as an alternative

mechanism for apical localisation. Transcript-gating suggests that pair-rule transcripts are transported along tracks in the nucleoplasm to apical pores in the nuclear membrane.

The transcript-gating hypothesis predicts that transcripts are transported along tracks in the nucleoplasm which could be visualised under condition which prevent the export of mRNA from the nucleus. It also predicts that different RNA molecules are exported through different NPCs at different locations on the nuclear membrane. Nuclear tracks for pair-rule transcripts are expected to lead to the apical side of the nuclear membrane. Apical localisation is expected to involve the binding of RNA to proteins in the nuclear matrix. These predictions could be tested by investigating the polarity of RNA tracks in the nuclei and by analysis of the nuclear distribution of factors required for apical localisation.

The cellular components involved in localisation

In chapter 7, it was shown that apical localisation of pair-rule genes is mediated by sequences in their transcripts rather than by sequences in the DNA. It seems likely that apical localisation is mediated by the interaction of a cellular protein or protein complex with pair-rule transcripts. Apical localisation could in theory be mediated by a different protein or protein complex in the case of each pair-rule gene. However, it is more productive to assume that the same proteins mediate localisation of all pair-rule transcripts.

It would be interesting to isolate RNA binding proteins which interact with *e32*. However, such experiments are technically difficult, since many non-specific RNA-binding proteins would be isolated together with molecules involved in apical localisation. If small localisation signals were defined in

ftz and *h* as well as in *eve*, all three sequences could be used in parallel to identify specific proteins which bind to all three RNA sequences.

Protein-targeting

To test directly the function of apical localisation, a construct could be made with an unlocalised transcript encoding a functional pair-rule protein. The construct would have to be expressed in an identical pattern to the wild type pair-rule gene. The ability of the construct to rescue pair-rule function and the degree of diffusion of protein away from the transcriptional domain would then be analysed to determine the function of apical localisation.

The protein-targeting model predicts that an unlocalised pair-rule transcript would lead to protein diffusion outside the normal pair-rule stripes. Under such circumstances, the protein stripes would be broader than the mRNA domains of expression and pattern defects would occur in the interstripe regions.

A complimentary construct could be made to determine the effect of apical localisation of a gap gene on its function. The construct would contain the upstream control region and the entire coding region of one of the gap genes, but the 3' end would be replaced by an apical localisation signal. The protein-targeting model predicts pattern defects at the edges of the expression domain of the gap gene because its protein would not diffuse beyond the transcriptional domain.

Asymmetric intracellular transcript localisation is not restricted to the *Drosophila* blastoderm (see chapter 1). It is possible that transcript-gating is responsible for the intracellular localisation of transcripts in other unicellular and syncytial eukaryotic cells.

Eukaryotic genes could be transcribed in discrete locations in the nuclear interior and transported along tracks in the nucleus to the nuclear

membrane; different categories of transcripts could contain different RNA sequences which determine their export through different parts of the nuclear membrane. The subcellular distribution of protein is proposed as a means of targeting proteins to their site of function.

Further experiments may lead to a better understanding of the mechanism and function for intracellular RNA distribution in *Drosophila* and in animal cells in general.

APPENDIX A: DIFFUSION ESTIMATES

A number of models for apical localisation have been presented in this thesis which rely on the free-diffusion in the nucleus or cytoplasm. To evaluate how realistic such models are, it is important to estimate the rate of free-diffusion of molecules in the cell.

RNA

It has been empirically estimated (Edgar *et al.*, 1987) that the diffusion coefficient (D) of a globular shaped molecule of molecular weight W in water at room temperature is:

$D=(2.23 \times 10^{-5} / W^{1/3})\ cm^2/sec$

The molecular weight of the *ftz* transcript can be estimated as the number of base pairs × the average molecular weight of a base pair. For a 5000bp *ftz* transcript and an average molecular weight for a base of 300, its molecular weight would be 1.5×10^6. The diffusion coefficient of such a molecule is thus estimated at approximately $2 \times 10^{-7}\ cm^2/sec$.

The cellular environment has been estimated to reduce the diffusion coefficient by a factor of 4 in comparison to water (Luby-Phelps *et al.*, 1986). It follows that $D= 5 \times 10^{-8}\ cm^2/sec$ in a cellular environment.

Now, the distance (L) diffused by a labile molecule (of half-life $T_{1/2}$) before degradation reduces its concentration by two-fold, is estimated as (Edgar *et al.*, 1987):

$L= \sqrt{(D \times T_{1/2})}$

ftz has a half-life of approximately 600 sec (Edgar *et al.*, 1986) so that we can estimate the approximate distance it diffuses before its concentration is reduced by two fold as:

L= 55μm, about 14 nuclear diameters.

Even if this estimate is wrong by a large factor, because RNA does not behave like a globular molecule, it would still mean that free RNA would diffuse to all parts of a nucleus before it substantially degrades. However, there is a large volume of data (described below) suggesting that particular RNA species are found localised to particular parts of the nucleus. mRNA is therefore likely to be bound to some components which prevent it from freely diffusing. Indeed, it has been suggested that mRNA is bound to the nuclear skeleton (see chapter 7).

The estimate of the diffusion rate of *ftz* suggests that in the unicellular syncytial blastoderm, pair-rule mRNA would diffuse away from the striped domain of transcription. However, it is clear from the data available, that pair-rule mRNA does not diffuse substantial distances in the cytoplasm. It follows that pair-rule mRNA is prevented from diffusing in the cytoplasm by binding to some component such as the cytoskeleton.

Diffusion of nuclear proteins before their entry into the nucleus

The molecular weight of the *ftz* protein is estimated at 6500 Da (Laughon and Scott, 1984) so its diffusion coefficient is:

$D=2.23 \times 10^{-5}/(6500)^{1/3}=$ approximately $1.2 \times 10^{-6} cm^2/sec$ in water. The cellular environment has been estimated to reduce the diffusion coefficient by a factor of 4 in comparison to water (Luby-Phelps *et al.*, 1986). The diffusion coefficient for *ftz* protein in the cellular environment can therefore be estimated at $3 \times 10^{-7} cm^2/sec$.

The half-life of the *ftz* protein has been estimated to be 300sec (Edgar *et al.*, 1987), so the distance the protein would diffuse before degradation reduces its concentration two-fold is:

$L=\sqrt{(D \times T_{1/2})}= \sqrt{(3 \times 10^{-7} \times 300)}\mu m=$ approximately 120μm, or 24 nuclear diameters.

The lack of *ftz* protein in the interstripes (see chapter 8) is due to transcriptional repression and is not due to post-translational regulation of protein degradation (Edgar *et al.*, 1986). Since *ftz* stripes are maintained for much longer than its half-life, it follows that something must prevent the *ftz* protein from diffusing beyond the boundaries of *ftz* transcription. While membranes probably prevent diffusion at late stage 14 embryos, the early and mid stage 14 blastoderm does not contain membranes in the basal periplasm.

In chapter 8, it is argued that restricting the distribution of pair-rule mRNA in the apical periplasm prevents the lateral diffusion of the protein it encodes. Under such circumstances, pair-rule proteins enter the adjacent nuclei.

APPENDIX B: RNA SEQUENCE ANALYSIS

In chapter 7, I showed that apical localisation is achieved by a transcript-mediated mechanism. The RNA sequences required for apical localisation may bind to a protein or to another RNA sequence in order to facilitate localisation. Such interactions could either involve the recognition of primary RNA sequence or the recognition of a secondary structural motif in the 3' untranslated ends of pair-rule mRNAs.

If the same factors are involved in the apical localisation of all pair-rule transcripts, the pair-rule genes may share common sequences in their 3' ends. I decided to compare the primary and secondary structure of the 3' end of the pair-rule genes described in this the previous chapters.

Primary sequence comparisons between pair-rule 3' ends

In chapter 4, I demonstrated that 167bp in the 3' untranslated part of *eve* are required for apical localisation. Only 124bp of this sequence are found in the *eve* transcript, the remainder of the sequence being downstream of the polyadenylation site. This 124bp region (*e32*) was used in sequence comparisons with the entire 3' untranslated parts of *ftz*, *h* and *prd*. A number of small regions of *e32* were found to be partly homologous with each of these genes. However, none of the homologies were found between more than one pair of genes. Moreover, the 3' ends of all the pair-rule genes are very A-T rich so that the small sequence homologies found are likely to be fortuitous and are not discussed further.

The secondary structure of the RNA localisation signal

The RNA secondary structure of *e32* was predicted using both sub-optimal and optimal folding algorithms (Jaeger *et al.*, 1989a; Jaeger *et al.*, 1989b; Zuker, 1989). The Optimal folding algorithm gives the secondary structure with the lowest free energy and thus represents the secondary structure of

an RNA molecule which is not bound to any molecules which alter the folding. Sub-optimal folding gives a more realistic folding including a number of possible secondary structural motifs of higher energy than the optimum. Both algorithms are more likely to give a meaningful folding with a small region rather than the whole of the 3' untranslated.

Both algorithms were used to predict the folding of *e32* (see Fig. B.1 & B.2) and the secondary structure was compared with that of similar sized regions near the polyadenylation sites of *ftz* and *h* (Fig. B.3). Similar secondary structural motifs for the three pair-rule genes were chosen from a number of possible sub-optimal foldings suggested by the algorithm. The results were extremely tentative since the localisation sequences of *ftz* and *h* were arrived at by extrapolation from *e32* and not from experimental data. The regions folded were all AT-rich and lead to unstable secondary structure with fairly high energy.

The folding which was found to be similar in the cases of *e32* and the predicted regions of *h* and *ftz* were very different from those defined for *prd*, another apically localised pair-rule gene (see Fig. B.4). Moreover, a similar secondary structure was found in the case of *hb*, which lacks an apical localisation signal (see Fig. 5b). *Kr* which also lacks a localisation signal, does not contain similar secondary structure motifs to those found in *e32* (see Fig. 5a). It follows that the the predicted secondary structure of *e32* does not necessarily represent the RNA motif which functions in apical localisation. Further molecular work to determine the localisation sequences of *ftz* and *h* is required before their secondary structure could be predicted with any degree of confidence.

Both the algorithms used did not allow pseudo-knots to form, so it is possible that the correct secondary structure would be missed with these algorithms. Any future predictions could be generated with algorithms

which include pseudo-knots as these are currently being developed (reviewed in Abrahams *et al.*, 1990).

RNA secondary structure predictions are only valuable in that they provide a number of structures which can be tested experimentally. Constructs can be made, which include mutated localisation sequences, to test if the disruption of the predicted stems affects localisation. Specific sequences occurring in the single stranded loops could also be tested for their importance in localisation.

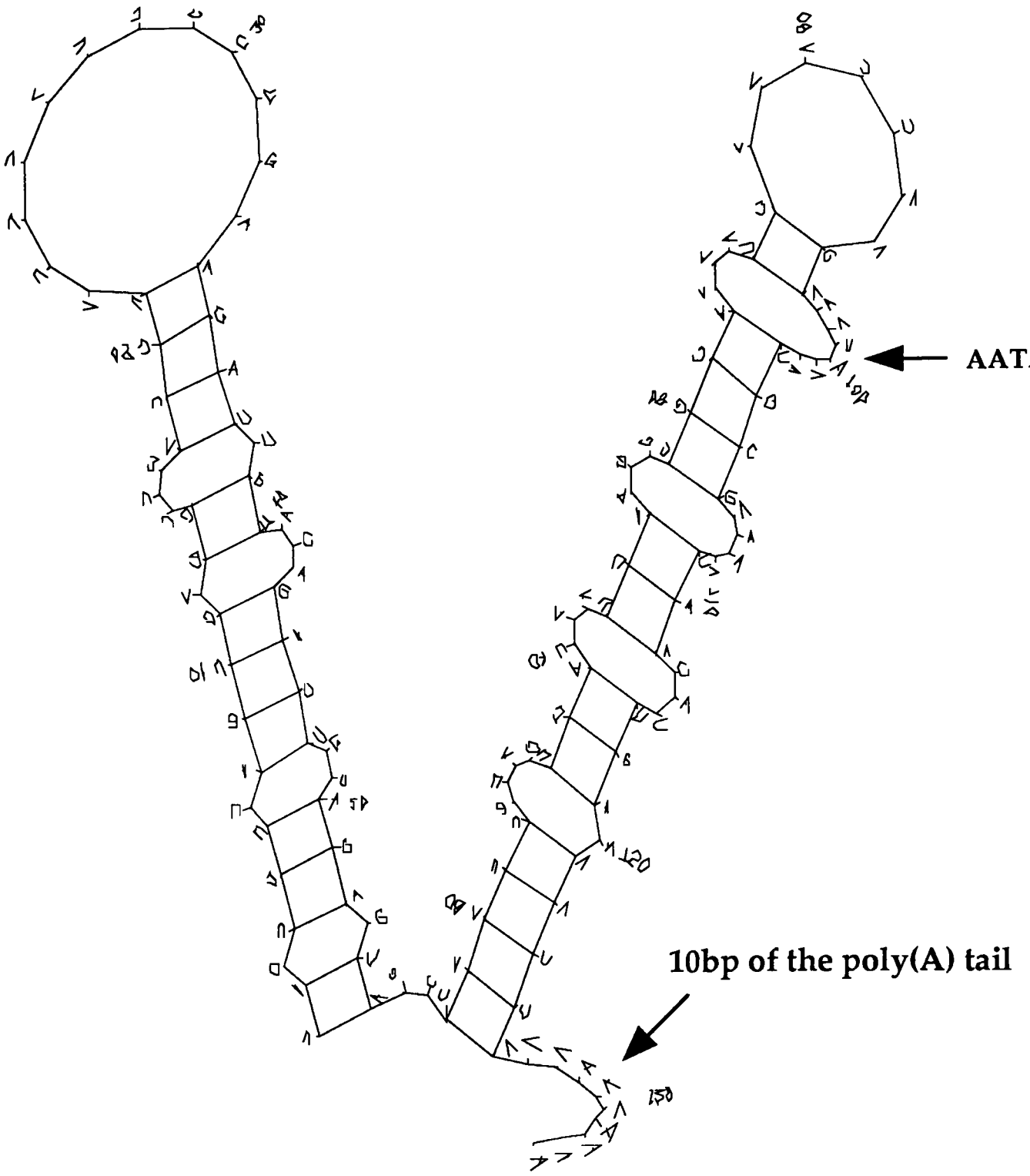


Fig. B.1 *eve32* length=134bp, energy=-10.0
(minimum free-energy is in kcal/mol)

A prediction of the secondary structure of the *e32* fragment of the *eve* gene using the Zuker RNA folding algorithm on a VAX minicomputer. The fragment includes 10bp of the poly(A) tail, but the predicted stem-loop regions were unaltered when the 10 Adesidues were deleted. Base pairing is shown by parallel lines which form stems. Two stem-loops are seen, with the right hand loop includes the polyadenylation signal and is also found in a suboptimal prediction of secondary structure shown in Fig. B. 4.

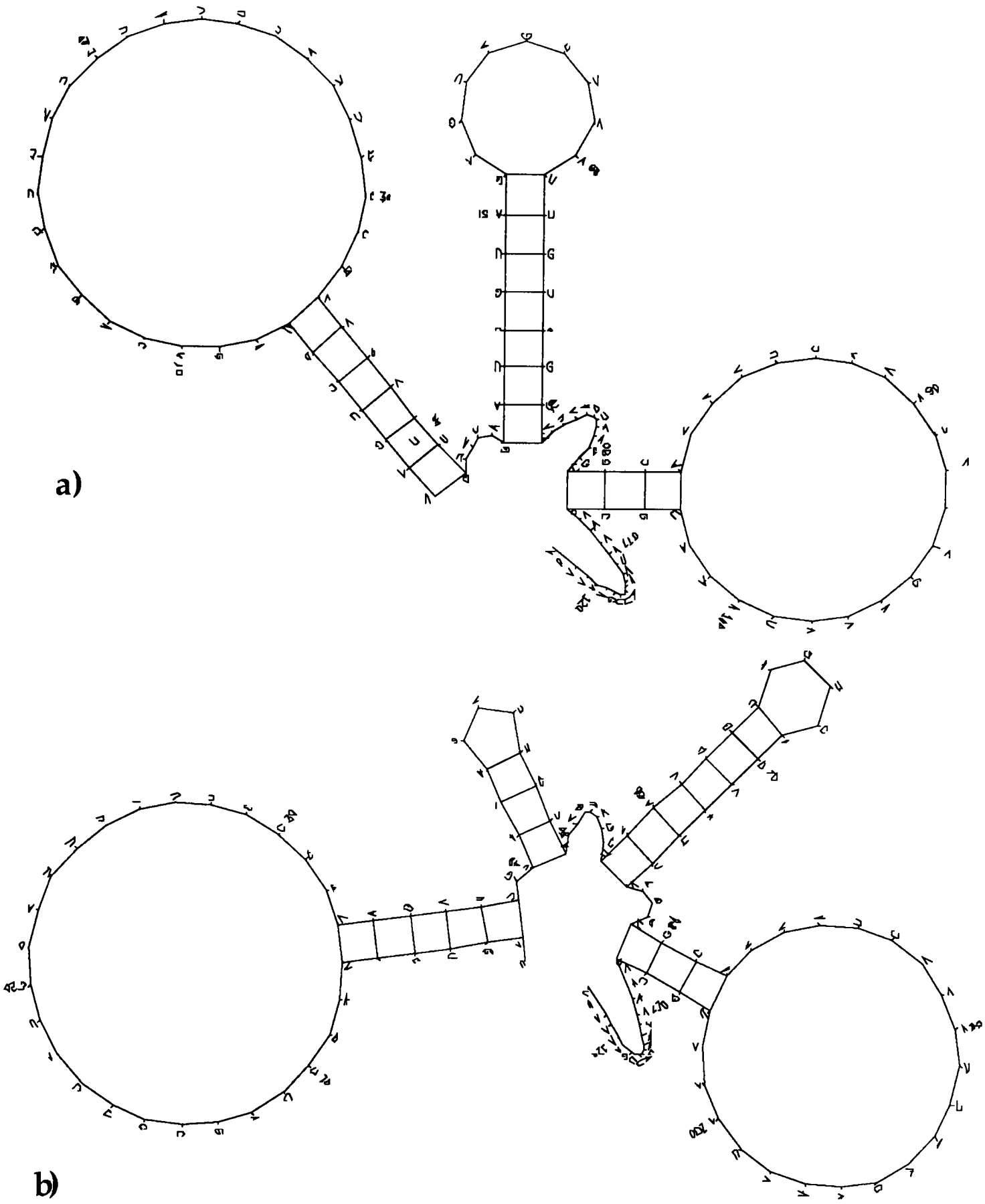
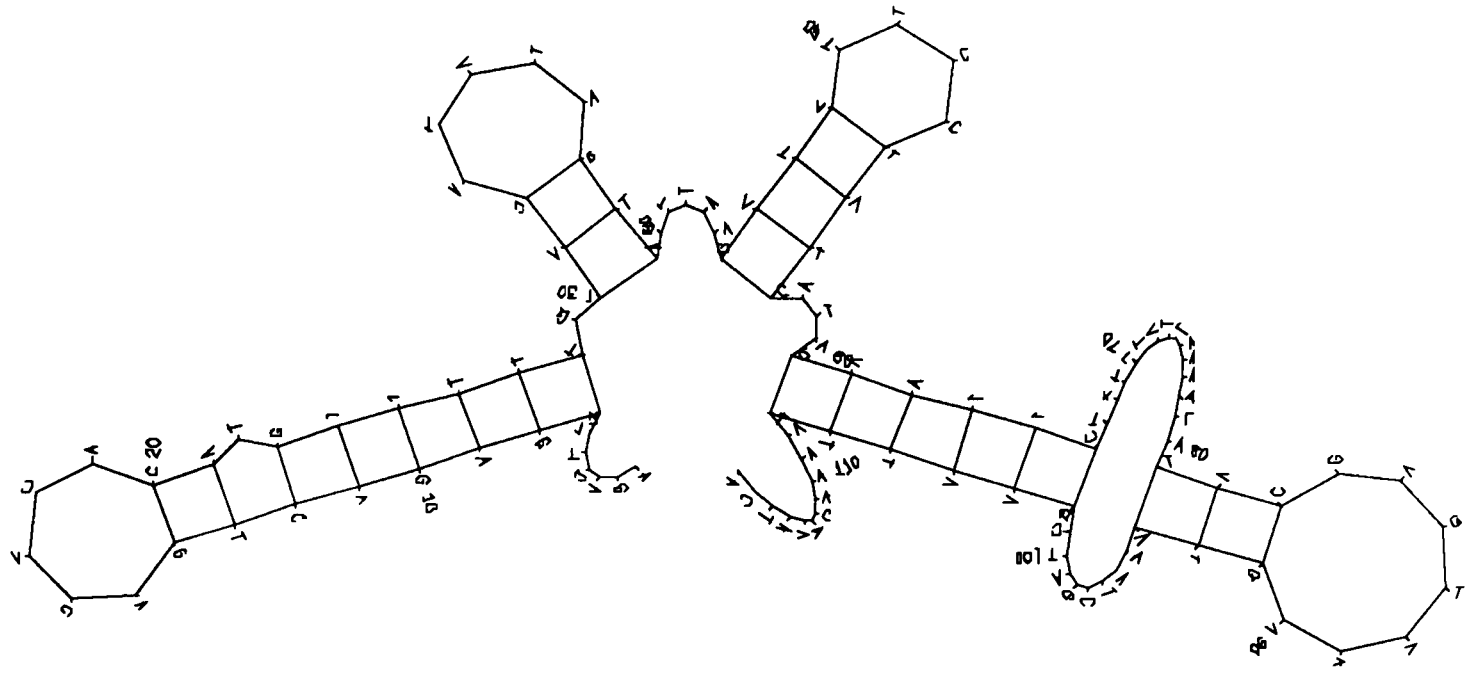
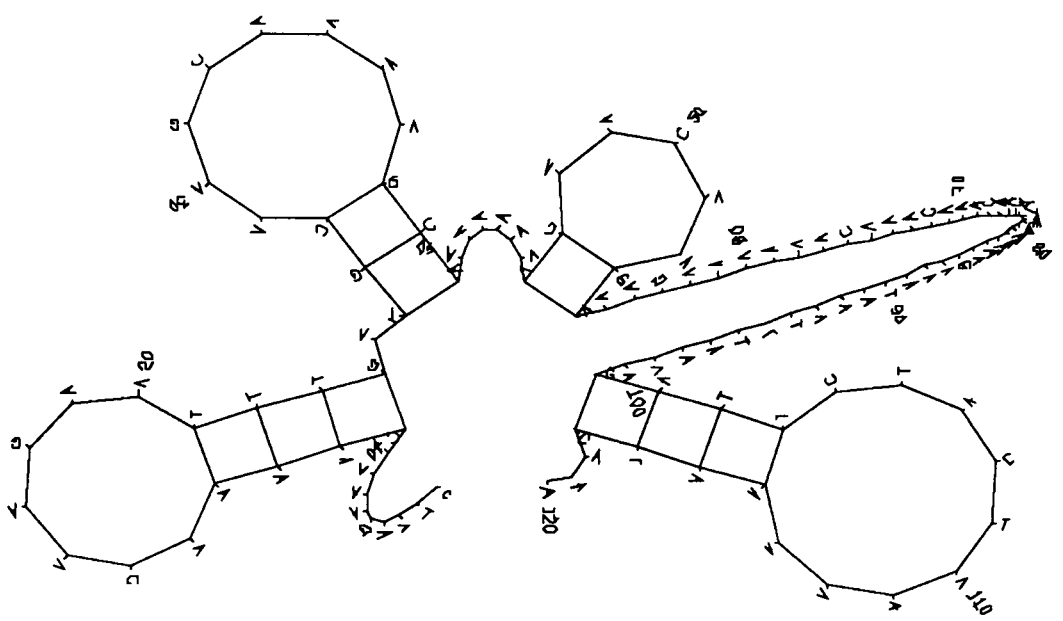


Fig. B.2 *e32* RNA, length= 124bp, energy= -7.0

A prediction of the secondary structure of the *e32* fragment of the *eve* gene using the Zuker RNA folding algorithm on a VAX minicomputer. The fragment includes 10bp of the poly(A) tail. Base pairing is shown by parallel lines which form stems. Two stem loops are seen, with the right hand loop includes the polyadenylation signal and is also found in a suboptimal prediction of secondary structure shown in Fig. B.4.



a) ***ftz*** **mRNA sub-optimal folding, length= 118bp, energy= -11.5**



b) *h* **mRNA sub-optimal folding, length= 120bp energy= -4.5**

Fig. B.3

A prediction of the secondary structure of a 120bp region upstream of the polyadenylation sites of *ftz* and *h*. There are four stem-loop structures similar to those seen in eve (see Fig. B.3b).

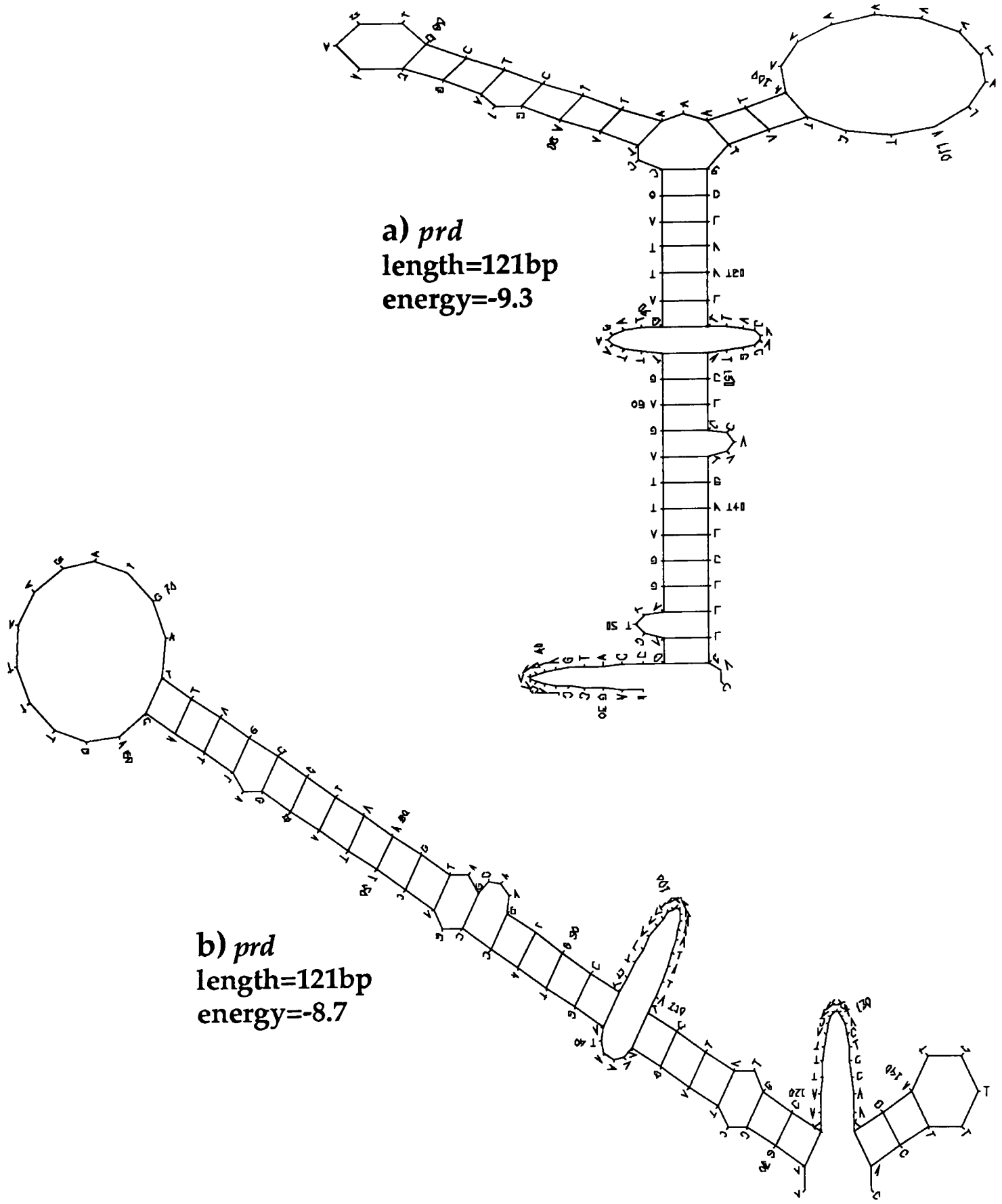
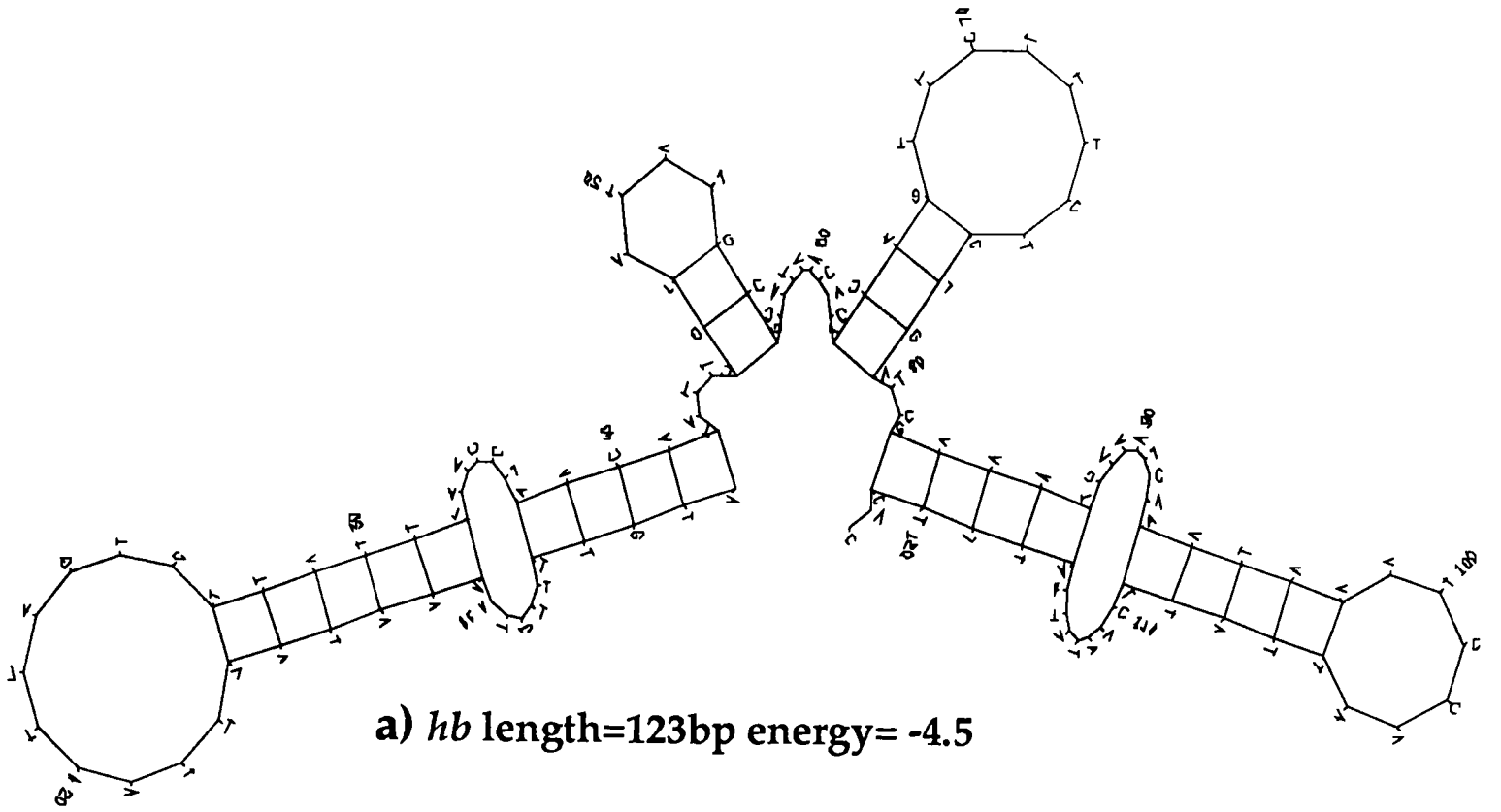
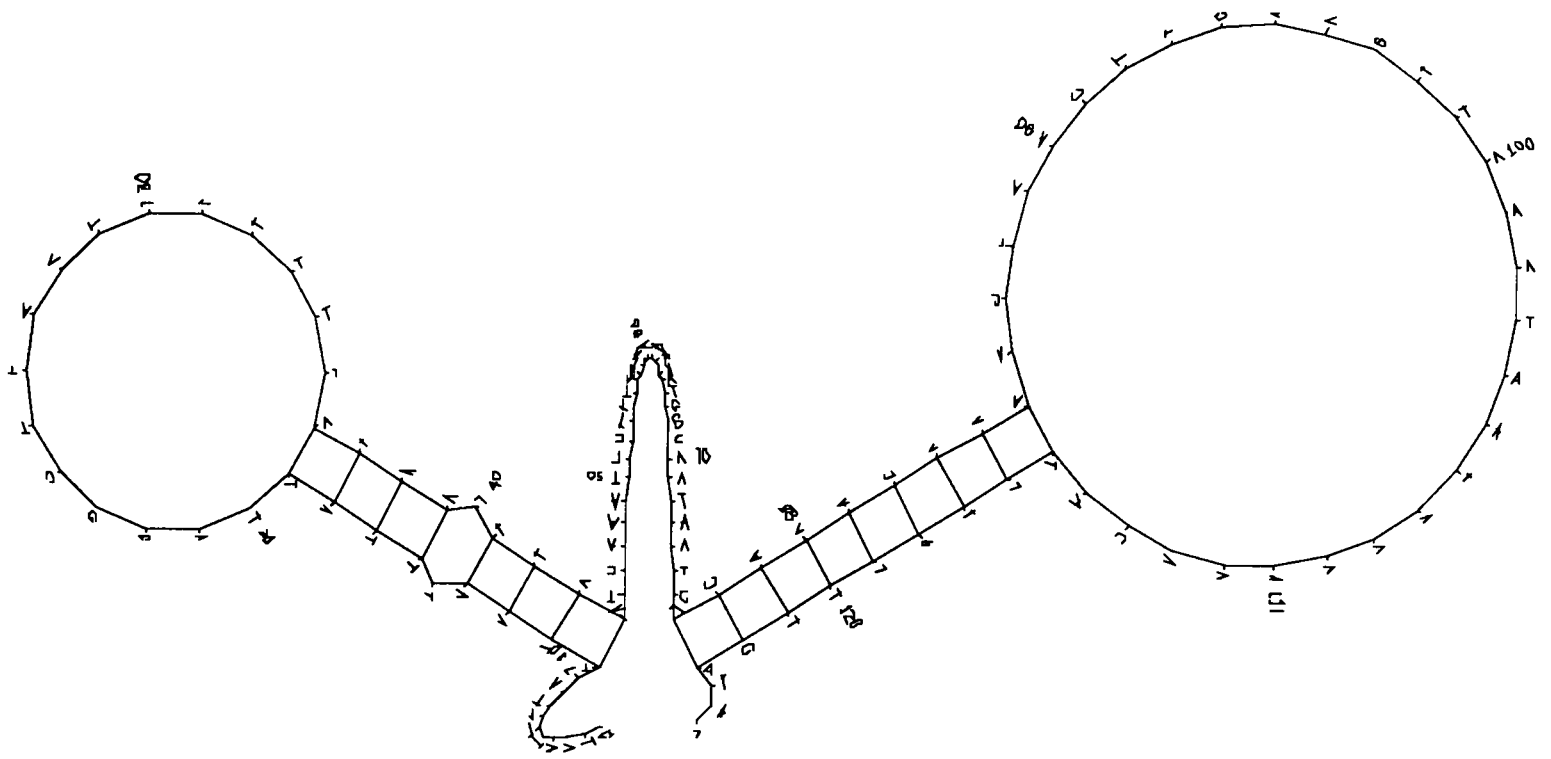


Fig. B.4

A prediction of the secondary structure of a 120bp region upstream of the polyadenylation sites of *prd* using the sub-optimal folding algorithm. There are no stem-loop structures similar to those seen in *eve* (Fig. B.3b) *ftz* or *h* (Fig. B.4).



a) ***hb*** **length=123bp energy= -4.5**



b) ***Kr*** **length=126bp energy= -4.2**

Fig. B.5

The secondary structure of small fragments of the 3' end of two gap genes. The fragments included 123bp and 126bp adjacent to the polyadenylation sites of *hb* and *Kr* respectively. *hb* has a similar structure to those found in *eve, ftz* and *h* but *Kr* displays a conpletely different secondary structure.

BIBLIOGRAPHY

Abrahams, J. P., van den Berg, M., van Batenburg, E. and Pleij, C. (1990). Prediction of RNA secondary structure, including pseudoknotting , by computer simulation. Nuc. Acids Res. *18*, 3035-3044.

Agard, D. A. and Sedat, J. W. (1983). Three-dimensional architecture of a polytene nucleus. Nature *302*, 676-681.

Agutter, P. S. (1985). RNA processing, RNA transport and nuclear structure. In "Nuclear Envelope Structure and RNA Maturation". pp539-559, A.R. Liss, ed., New York.

Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. Development *101*, 1-22.

Akam, M. E. (1985). The distribution of *Ultrabithorax* transcript in *Drosophila* embryos. EMBO J. *4*, 1689-1700.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1989). "Molecular Biology of the Cell". (Garland Publishing Inc, New York & London).

Baas, P. W. (1988). Polarity orientation of microtubules in hippocampal neurons: uniformity in the axon and non-uniformity in the dendrite. P.N.A.S. *85*, 8335-8339.

Bak, A. L., Zeuthen, J. and Crick, F. H. C. (1977). Higher-order structure of human mitotic chromosomes. P.N.A.S. *74*, 1595-1599.

Baker, N. E. (1987). Molecular cloning of sequences from *wingless*, a segment polarity gene in *Drosophila*: the spatial distribution of a transcript in embryos. The EMBO J. *6*, 1765-1773.

Baumgartner, S., Bopp, D., Burri, M. and Noll, M. (1987). Structure of two genes at the *gooseberry* locus related to the *paired* gene and their spatial expression during *Drosophila* embryogenesis. Genes and Dev *1*, 1247-1267.

Belmont, A. S., Sedat, J. W. and Agard, D. A. (1987). A three-dimensional approach to mitotic chromosome structure: evidence for a complex hierarchical organization. J Cell Biol *105*, 77-92.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nüsslein-Volhard, C. (1988). The role of localisation of bicoid RNA in organising the anterior pattern of the *Drosophila* embryo. The EMBO Journal *7*, 1749-1756.

Beyer, A. L. and Miller, O. L. (1980). Ribonucleoprotein structure in nascent hnRNA is non-random and sequence-dependent. Cell *20*, 75-84.

Blobel, G. (1985). Gene-gating: a hypothesis. P.N.A.S. *82*, 8527-8529.

Breeuwer, M. and Goldfarb, D. S. (1990). Facilitated nuclear transport of histone H1 and other small nucleophilic proteins. Cell *60*, 99-1008.

Burn, T. C., Vigoreaux, J. O. and Tobin, S. L. (1989). Alternative *5C actin* transcripts are localized in different patterns during *Drosophila* embryogenesis. Dev. Biol. *131*, 345-355.

Campos-Ortega, J. A. and Hartenstein, V. (1985). "The embryonic development of *Drosophila melanogaster*". (Springer Verlag, Berlin).

Capco, D. G. and Jeffrey, W. R. (1982). Transient localizations of messenger RNA in *Xenopus laevis* oocytes. Dev. Biol. *89*, 1-12.

Clawson, G. A., Feldherr, C. M. and Smuckler, E. A. (1985). Nucleocytoplasmic RNA transport. Molecular and Cellular Biochemistry *67*, 87-100.

Clawson, G. A., Koplitz, M., Castler-Schechter, B. and Smuckler, E. A. (1978). Energy utilization and RNA transport: their interdependence. Biochem *17*, 3747-3753.

Cremer, T., Cremer, C., Baumann, H., Luedtke, E. K., Sperling, K., Teuber, V. and Zorn, C. (1982). Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. Hum. Genet. *60*, 46-56.

Dale, L., Matthews, G., Tabe, L. and Colman, A. (1989). Developmental expression of the protein product of *Vg1*, a localized maternal mRNA in the frog *Xenopus laevis*. EMBO J *8*, 1057-1065.

Davis, L., Banker, G. A. and Steward, O. (1987). Selective dendritic transport of RNA in hippocampal neurons in culture. Nature *330*, 477-479.

De Robertis. (1983). Nucleocytoplasmic segregation of proteins and RNAs. Cell *32*, 1021-1025.

Desplan, C., Theis, J. and O'Farrell, P. H. (1985). The *Drosophila* developmental gene *engrailed* encodes a sequence-specific DNA-binding activity. Nature *318*, 630-635.

Driever, W. and Nüsslein-Volhard, C. (1988a). The *bicoid* protein determines position in the *Drosophila* embryo in a concentration-dependent manner. Cell *54*, 95-104.

Driever, W. and Nüsslein-Volhard, C. (1988b). A gradient of *bicoid* protein in *Drosophila* embryos. Cell *54*, 83-93.

Driever, W. and Nüsslein-Volhard, C. (1989). The *bicoid* protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. Nature *337*, 138-143.

Dworetzky, S. I. and Feldherr, C. M. (1988). Translocation of RNA-coated gold particles through the nuclear pores of oocytes. Cell *106*, 575-584.

Edgar, B. A., Odell, G. M. and Schubiger, G. (1987). Cytoarchitecture and the patterning of *fushi tarazu* expression in the *Drosophila* blastoderm. Genes and Development *1*, 1226-1237.

Edgar, B. A., Weir, M. P., Schubiger, G. and Kornberg, T. (1986). Repression and turnover pattern *fushi tarazu* RNA in the early *Drosophila* embryo. Cell *47*, 747-754.

Fasano, L., Coré, N. and Kerridge, S. (1988). Expression of a reporter gene resembles that of its neighbour: an insertion in the *hairy* gene of *Drosophila*. Roux's Arch Dev. Biol. *197*, 507-512.

Fey, E. G., Ornelles, D. A. and Penman, S. (1986). Association of RNA with the cytoskeleton and the nuclear matrix. J. Cell Sci. Supplement. *5*, 99-119.

Finlay, D. R. and Forbes, D. J. (1990). Reconstitution of biochemically altered nuclear pores: transport can be eliminated and restored. Cell *60*, 17-29.

Finlay, D. R., Newmeyer, D. D., Price, T. M. and Forbes, D. J. (1987). Inhibition of *in vitro* nuclear transport by a lectin that binds to nuclear pores. Cell Biology *104*, 189-200.

Foe, V. (1989). Mitotic domains reveal early commitment of cells in *Drosophila* embryos. Development *107*, 1-22.

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

Foe, V. E. and Alberts, B. M. (1985). Reversible chromosome condensation induced in *Drosophila* embryos by anoxia: visualization of interphase nuclear organization. J. Cell Biol. *100*, 1623-1636.

Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986). Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. Cell *47*, 735-46.

Frohnhöfer, H. G. and Nüsslein-Volhard, C. (1986). Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. Nature *324*, 120-125.

Frohnhöfer, H. G. and Nüsslein-Volhard, C. (1987). Maternal genes required for the anterior localization of *bicoid* activity in the embryos of *Drosophila*. Genes Dev. *1*, 880-890.

Fullilove, S. L. and Jacobson, A. G. (1971). Nuclear Elongation and Cytokinesis in *Drosophila montana*. Developmental Biology *26*, 560-577.

Garcia-Bellido, A., Rippoll, P. and Morata, G. (1973). Developmental compartmentalisation of the wing disc of *Drosophila*. Nature New Biol *245*, 251-253.

Garner, C. C., Tucker, R. P. and Matus, A. (1988). Selective localisation of messenger RNA for cytoskeletal protein MAP2 in dendrites. Nature *336*, 674-677.

Gasser, S. M. and Laemmli, U. K. (1987). A glimpse at chromosome order. TIG *3*, 16-22.

Gaul, U. and Jäckle, H. (1989). Analysis of maternal effect mutant combinations elucidates regulation and function of the overlap of *hunchback* and *Krüppel* gene expression in the *Drosophila* blastoderm embryo.

Gerace, L. and Burke, B. (1988). Functional organisation of the nuclear membrane. Annu. Rev. Cell Biol. *4*, 335-374.

Gerace, L., Comeau, C. and Benson, M. (1984). Organization and modulation of nuclear lamina structure. J. Cell Sci. Suppl. *1*, 137-160.

Gerttula, S., Jin, Y. and Anderson, K. V. (1988). Zygotic expression and activity of the *Drosophila Toll* gene, a gene required maternally for embryonic dorsal-ventral pattern formation. Genetics *119*, 123-133.

Guddat, U., Bakken, A. and Pieler, T. (1990). Protein-mediated nuclear export of RNA: 5S rRNA containing small RNAPs in *Xenopus* oocytes. Cell *60*, 619-628.

Hall, Z. and Ralston, E. (1989). Nuclear domains in muscle cells. Cell *59*, 771-772.

Heaphy, S., Dingwall, C., Ernberg, I., Gait, M., Green, S., Karn, J., Lowe, A., Singh, M. and Skinner, M. (1990). HIV-1 regulator of viral expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element. Cell *60*, 685-693.

Higuchi, R., Krummel, B.and Saiki, R. K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. N.A.R. *16*, 7351-7364.

Hiraoka, Y., Minden, J. S., Swedlow, J. R., Sedat, J. W. and Agard, D. A. (1989). Focal points for chromosome condensation and decondensation revealed by three-dimensional *in vivo* time-lapse microscopy. Nature *342*, 293-296.

Hiromi, Y. and Gehring, W. J. (1987). Regulation and function of the *Drosophila* segmentation gene *fushi tarazu*. Cell *50*, 963-974.

Hiromi, Y., Kuroiwa, A. and Gehring, W. J. (1985). Control elements of the *Drosophila* segmentation gene *fushi tarazu*. Cell *43*, 603-613.

Hochstrasser, M. and Sedat, J. (1987a). Three-dimensional organization of *Drosophila melanogaster* interphase nuclei II. Chromosome spatial organization and gene regulation. J. Cell Biol. *104*, 1471-1483.

Hochstrasser, M. and Sedat, J. (1987b). Three-dimensional organization of *Drosophila melanogaster* interphase nuclei. I. Tissue-specific aspects of poly(T)ene nuclear architecture. J. Cell Biol *104*, 1455-1470.

Hooper, K. L. H., Parkhurst, S. M. and Ish-Horowicz, D. (1989). Spatial control of *hairy* protein expression during embryogenesis. Development *107*, 489-504.

Howard, K. and Ingham, P. (1986). Regulatory interactions between the segmentation genes *fushi tarazu, hairy* and *engrailed* in the *Drosophila* blastoderm. Cell *44*, 949-957.

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

Humphrey, T. and Proudfoot, N. J. (1988). A beginning to the biochemistry of polyadenylation. TIG 4, 243-245.

Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. Nature *335*, 25-34.

Ingham, P. W., Baker, N. E. and Martinez-Arias, A. (1988). Regulation of segment polarity genes in the *Drosophila* blastoderm by *fushi tarazu* and *even-skipped*. Nature *331*, 73-75.

Ingham, P. W. and Gergen, J. P. (1988). Interactions between the pair-rule genes *runt, hairy, even-skipped* and *fushi tarazu* and the establishment of periodic pattern in the *Drosophila* embryo. In "Mechanisms of Segmentation". Development *s104*, pp51-60, V. French, Ingham, P.W., Cooke, J. and J., S., ed.,

Ingham, P. W., Howard, K. R. and Ish-Horowicz, D. (1985a). Transcription pattern of the *Drosophila* segmentation gene hairy. Nature *318*, 439-445.

Ingham, P. W., Pinchin, S. M., Howard, K. R. and Ish-Horowicz, D. (1985b). Genetic analysis of the *hairy* locus in *Drosophila melanogaster*. Genetics *111*, 463-486.

Ish-Horowicz, D. and Gyurkovics, H. (1988). Ectopic segmentation gene expression and metameric regulation in *Drosophila*. In "Mechanisms of Segmentation". *s104*, pp67-73, V. French, Ingham, P.W., Cooke, J. and Smith, J., ed., Development).

Ish-Horowicz, D. and Pinchin, S. M. (1987). Pattern abnormalities induced by ectopic expression of the *Drosophila* gene hairy are associated with repression of *fushi tarazu* transcription. Cell *51*, 405-415.

Jackson, D. A., Canton, A. J., McCready, S. J. and Cook, P. R. (1982). Influenza virus RNA is synthesized at fixed sites in the nucleus. Nature *296*, 366-368.

Jackson, D. A. and Cook, P. R. (1985). Transcription occurs at a nucleoskeleton. EMBO J. *4*, 919-925.

Jaeger, J. A., Turner, D. H. and Zuker, M. (1989a). Improved predictions of secondary structures for RNA. P.N.A.S. *86*, 7706-7710.

Jaeger, J. A., Turner, D. H. and Zuker, M. (1989b). Predicting optimal and sub-optimal secondary structures for RNA. In "Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences". Methods in Enzymology *183*, pp281-306, R.F. Doolittle, ed., (Academic Press)

Jaynes, J. B. and O'Farrell, P. H. (1988). Activation and repression of transcription by homoeodomain-containing proteins that bind a common site. Nature *336*, 744-749.

Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. and Kluding, H.. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*
II. Zygotic loci on the third chromosome. Roux's Arch Dev Biol *193*, 283-295.

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

Kastern, W. H., Watson, C. A. and Berry, S.J. (1990). Maternal messenger RNA distribution in silkmoth eggs
I. Clone EC4B is associated with the cortical cytoskeleton. Development *108*, 495-505.

Kellogg, D. R. and Alberts, B. M. (1988). Behaviour of microtubules and actin filaments in living *Drosophila* embryos. Development *103*, 675-686.

Kellogg, D. R., Field, C. M. and Alberts, B. M. (1989). Identification of microtubule-associated proteins in the centrosome, spindle and kinetochore of the early *Drosophila* embryo. J. Cell Biol. *109*, 2977-92.

Kelly, R. B. (1985). Pathways of protein secretion in eukaryotes. Science *230*, 25-32.

Kilcherr, F., Baumgartner, S., Bopp, D., Frei, E. and Noll, M. (1986). Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. Nature *321*, 493-499.

King, R. (1970). "Oogenesis in *Drosophila*". (Academic Press,

Konings, D. A. M. and Mattaj, I. W. (1987). Mutant U2 snRNAs of *Xenopus* which can form an altered higher-order RNA structure are unable to enter the nucleus. Experimental Cell Research *172*, 329-339.

Kyuhyung, H., Levine, M. S. and Manley, J. (1989). Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. Cell *56*, 573-583.

Laughon, A. and Scott, M. P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature *310*, 25-31.

Lawrence, J. B. and Singer, R. H. (1986). Intracellular localization of messenger RNAs for cytoskeletal proteins. Cell *45*, 407-415.

Lawrence, J. B., Singer, R. H. and Marselle, L. M. (1989). Highly localized tracks of specific transcripts within interphase nuclei visualized by *in situ* hybridization. Cell *57*, 493-502.

Lawrence, P. A., Johnston, P., Macdonald, P. and Struhl, G. (1987). Borders of parasegments in *Drosophila* embryos are delimited by the *fushi tarazu* and even-skipped genes. Nature *328*, 440-442.

Lehmann, R. and Nüsslein-Volhard, C. (1987). *hunchback*, A gene required for segmentation of an anterior and posterior region of the *Drosophila* embryo. Dev. Biol. *119*, 402-417.

Levitt, N., Briggs, D., Gil, A. and Proudfoot, N. J. (1989). Definition of an efficient synthetic poly(A) site. Genes and Development *3*, 1019-1025.

Luby-Phelps, K., Taylor, D. L. and Lanni, F. (1986). Probing the structure of cytoplasm. J. Cell Biol. *102*, 2015-2022.

Macdonald, P. M., Ingham, P. and Struhl, G. (1986). Isolation, structure and expression of even-skipped: a second pair-rule gene of *Drosophila* containing a homeo box. Cell *47*, 721-734.

Macdonald, P. M. and Struhl, G. (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying body pattern. Nature *324*, 537-545.

Macdonald, P. M. and Struhl, G. (1988). Cis-acting sequences responsible for anterior localisation of bicoid mRNA in *Drosophila* embryos. Nature *336*, 595-598.

Mahowald, A. and Kambysellis, M. (1980). Oogenesis. "The genetics and biology of *Drosophila*". Vol *2c*: Ashburner, M. and Wright, T. R. F. Eds. (Academic Press London, NY, San Francisco)

Martinez-Arías, A., Baker, N. E. and Ingham, P. W. (1988). Role of segment polarity genes in the definition and maintenance of cell states in the *Drosophila* embryo. Development *103*, 157-170.

Martinez-Arías, A. and Lawrence, P. A. (1985). Parasegments and compartments in the *Drosophila* embryo. Nature *313*, 639-642.

Mathews, K. A., Miller, D. F. B. and Kaufman, T. C. (1989). Developmental distribution of RNA and protein products of the *Drosophila* α-tubulin gene family. Dev. Biol. *132*, 45-61.

Mathog, D., Hochstrasser, M., Gruenbaum, Y., Saumweber, H. and Sedat, J. (1984). Characteristic folding pattern of polytene chromosomes in *Drosophila* salivary gland nuclei. Nature *308*, 414-421.

Mattaj, I. W. and De Robertis, E. M. (1985). Nuclear segregation of U2 snRNA requires binding of specific snRNA proteins. Cell *40*, 111-118.

McKnight, S. L. and Miller, O. L. (1976). Ultrastructural patterns of RNA synthesis during early embryogenesis of *Drosophila melanogaster*. Cell *8*, 305-319.

Melton, D. A. (1987). Translocation of a localized maternal mRNA to the vegetal pole of *Xenopus* oocytes. Nature *128*,

Merlie, J. P. and Sanes, J. R. (1985). Concentration of acetylcholine receptor mRNA in synaptic regions of adult muscle fibres. Nature *317*, 66-68.

Miller, K., Karr, T., Kellogg, D., Mohr, I., Walter, M. and Alberts, B. (1985). Studies on the cytoplasmic organization of early *Drosophila* embryos. Molecular Biology of Development. *L*: 79-90.

Miller, K. G., Field, C. M. and Alberts, B. M. (1989). Actin binding proteins from *Drosophila* embryos: a complex network of interacting proteins detected by F-actin chromatography. J. Cell Biol. *109*, 2963-2976.

Miller, T. E., Huang, C. and Pogo, A. O. (1978). Rat liver nuclear skeleton and ribonucleoprotein complexes containing hnRNA. J. Cell Biol. *76*, 675-691.

Mirkovitch, J., Mirault, M. E. and Laemmli, U. K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell

Nakano, Y., Guerrero, I., Hidalgo, A., Taylor, A., Whittle, J. R. S. and Ingham, P. W. (1989). A protein with several possible membrane-spanning domains encoded by the *Drosophila* segment polarity gene patched. Nature *341*, 508-513.

Newmeyer, D. D. and Forbes, D. J. (1988). Nuclear import can be separated into distinct steps *in vitro*: nuclear pore binding and translocation. Cell *52*, 641-653.

Newport, J. and Forbes, D. (1987). The nucleus: structure, function and dynamics. Annu. Rev. Biochem. *56*, 535-566.

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

Nüsslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature *287*, 795-801.

Nüsslein-Volhard, C., Wieschaus, E., Kluding, H. and Jürgens, G. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*
I. Zygotic loci on the second chromosome. Roux's Arch Dev Biol *193*, 267-282.

O'Kane, C. J. and Gehring, W. J. (1987). Detection *in situ* of genomic regulatory elements in *Drosophila*. P.N.A.S. *84*, 9123-9127.

Olsen, H., Nelbock, P., Cochrane, A. and Rosen, C. (1990). Secondary structure is the major determinant for interaction of HIV rev protein with RNA. Science *247*, 845-848.

Pankratz, M. J., Hoch, M., Seifert, E. and Jäckle, H. (1989). *Krüppel* requirement for *knirps* enhancement reflects overlapping gap gene activities in the *Drosophila* embryo. Nature *341*, 337-340.

Pankratz, M. J., Seifert, E., Gerwin, N., Billi, B., Nauber, U. and Jäckle, H. (1990). Gradients of *Krüppel* and *knirps* gene products direct pair-rule gene stripe patterning in the posterior region of the *Drosophila* embryo. Cell *61*, 309-317.

Pavlakis, G. N. and Felber, B. K. (1990). Regulation of expression of human immunodeficiency virus. The New Biologist 2, 20-31.

Pesacreta, T. C., Byers, T. J., Dubreil, R., Kiehart, D. P. and Branton, D. (1989). *Drosophila* spectrin: the membrane skeleton during embryogenesis. J. Cell Biol. *108*, 1697-1709.

Ralston, E. and Hall, Z. (1989). Transfer of a protein encoded by a single nucleus to nearby nuclei in multinucleated myotubes. Science *244*, 1066-1069.

Rebagliati, M. R., Weeks, D. L., Harvey, R. P. and Melton, D. A. (1985). Identification and cloning of localized maternal RNAs from *Xenopus* eggs. Cell *42*, 769-777.

Richardson, W. D., Mills, A. D., Dilworth, S. M., Laskey, R. A. and Dingwall, C. (1988). Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. Cell *52*, 655-664.

Rickoll, W. L. (1976). Cytoplasmic continuity between embryonic cells and the primitive yolk sac during early gastrulation in *Drosophila melanogaster*. Developmental Biology. *49*, 304-310.

Ringertz, N. (1986). Computer analysis of the distribution of nuclear antigens: studies on the spatial and functional organization of the interphase nucleus. J. Cell Sci. Suppl. *4*, 11-28.

Rushlow, C., Doyle, H., Hoey, T. and Levine, M. (1987). Molecular characterization of the *zerknüllt* region of the *Antennapedia* complex in *Drosophila*. Genes and Dev. *1*, 1268-1279.

Rushlow, C. and Levine, M. (1988). Combinatorial expression of a *ftz-zen* fusion promoter suggests the occurrence of cis-interactions between genes of the ANT-C. EMBO J *7*, 3479-3485.

Rushlow, C. A., Hogan, A., Pinchin, S. M., Howe, K. R., Lardelli, M. T. and Ish-Horowicz, D. (1989). The *Drosophila hairy* protein acts in both segmentation and bristle patterning and shows homology to *N-myc*. EMBO. J. *8*, 3095-3103.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). "Molecular Cloning, A Laboratory Manual'. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. P. N. A. S. *74*, 5463.

Schweisguth, F. C., Lepesant, J.-A. and Vincent, A. (1990). The *serendipity alpha* gene encodes a membrane associated protein required for the cellularization of the *Drosophila* embryo. Genes and Dev. *4*, 922-931.

Schweisguth, F. C., Yanicostas, F. P., Lepasant, J.-A. and Vincent, A. (1989). Cis-regulatory elements of the *Drosophila* blastoderm-specific *serendipity alpha* gene: ectopic activation in the embryonic PNS is promoted by the deletion of an upstream element. Dev. Biol. *136*, 181-193.

Skoglund, U., Anderson, K., Strandberg, B. and Daneholt, B. (1986). Three-dimensional structure of a specific pre-messenger RNP particle established by electron microscope tomography. Nature *319*, 560-564.

Slack, J. M. W., Darlington, B. G., Gillespie, L. L., Godsave, S. F., Isaacs, H. V. and Paterno, G. D. (1989). The role of fibroblast growth factor in early *Xenopus* development. Development *107 supplement*, 141-148.

Smith, J. C., Price, B. M. J., Van Nimmen, K. and Huylebroeck, D. (1990). Identification of a potent *Xenopus* mesoderm-inducing factor as a homologue of activin A. Nature *345*, 729-731.

St Johnston, D., Driever, W., Berleth, T., Richstein, S. and Nüsslein-Volhard, C. (1989). Multiple steps in the localisation of bicoid RNA to the anterior pole of the *Drosophila* oocyte. Development Suppl *107*, 13-19.

Stanojevic, D., Hoey, T. and Levine, M. (1989). Sequence specific DNA-binding activities of the gap proteins encoded by *hunchback* and *Krüppel* in *Drosophila*. Nature *341*, 331-335.

Steller, H. and Pirrotta, V. (1985). A transposable P vector that confers G418 resistance to *Drosophila* larvae. EMBO J *4*, 167-171.

Stevens, B. J. and Swift, H. (1966). RNA transport from nucleus to cytoplasm in chironomus salivary glands. J. Cell Biol *31*, 55-77.

Struhl, G. (1985). Near-reciprocal phenotypes caused by inactivation or indiscriminate expression of the *Drosophila* segmentation gene *ftz*. Nature *318*, 677-680.

Struhl, G., Struhl, K. and Macdonald, P. M. (1989). The gradient morphogen *bicoid* is a concentration-dependent transcriptional activator. Cell *57*, 1259-1273.

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 (Berl) *98*, 81-85.

Tepass, U., Theres, C. and Knust, E. (1990). *crumbs* encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. Cell *61*, 787-799.

van den Heuvel, M., Nusse, R., Johnston, P. and Lawrence, P. (1989). Distribution of the wingless gene product in *Drosophila* embryos: a protein involved in cell-cell communication. Cell *59*, 739-749.

van Eekelen, C. A. G. and van Venrooij, W. J. (1981). hnRNA and its attachment to a nuclear protein matrix. Journal of Cell Biology *88*, 554-563.

Warn, R. M. (1986). The cytoskeleton of the early *Drosophila* embryo. J.Cell Sci. Suppl. *5*, 311-328.

Warn, R. M. and Magrath, R. (1983). F-actin distribution during the cellularization of the *Drosophila* embryo visualized with FL-phalloidin. Experimental Cell Research *143*, 103-114.

Warn, R. M. and Warn, A. (1986). Microtubule arrays present during the syncytial cellular blastoderm stages of the early *Drosophila*. Experimental Cell Research *163*, 201-210.

Weeks, D., Rebagliati, M., Harvey, R. and Melton, D. (1985). Localized maternal mRNAs in *Xenopus laevis* Eggs. Molecular Biology of Development. *L:* 21-30.

Weeks, D. L. and Melton, D. A. (1987). A maternal mRNA localized to the animal pole of *Xenopus* eggs encodes a subunit of mitochondrial ATPase. Proc. Natl. Acad. Sci. *84,* 2798-2802.

Weiner, A. J., Scott, M. P. and Kaufman, T. C. (1984). A molecular analysis of *fushi tarazu,* a gene in *Drosophila melanogaster* that encodes a product affecting embryonic segment number and cell fate. Cell *37,* 843-851.

Weir, M. P. and Kornberg, T. (1985). Pattern of *engrailed* and *fushi tarazu* transcripts reveal novel intermediate stages in *Drosophila* segmentation. Nature *318,* 433-39.

Wieschaus, E. and Nüsslein-Volhard, C. (1986). Looking at embryos. In "*Drosophila,* a practical approach". pp199-226, D.B. Roberts, ed., (IRL, Oxford).

Wieschaus, E., Nüsslein-Volhard, C. and Jürgens, G. (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*
III. Zygotic loci on the fourth chromosome. Roux's Arch Dev Biol *193,* 296-307.

Yisraeli, J., Sokol, D. and Melton, D. (1990). A two-step for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of *Vg1* mRNA. Development *108,* 289-298.

Yisraeli, J. K. and Melton, D. A. (1988). The maternal mRNA *Vg1* is correctly localized following injection into *Xenopus* oocytes. Nature *336,* 592-595.

Yisraeli, J. K., Sokol, S. and Melton, D. A. (1989). The process of localising a maternal messenger RNA in *Xenopus* oocytes. Development Suppl. *107,* 31-36.

Zasloff, M. (1983). tRNA transport from the nucleus in a eukaryotic cell: carrier-mediated translocation process. Pro. Natl. Acad. Sci. *80,* 6436-6440.

Zuker, M. (1989). On finding all sub-optimal foldings of an RNA molecule. Science *244,* 48-52.