

# Structural studies of RNA localisation signals

Inbal Ringel

University College London

and

Developmental Genetics Laboratory, Cancer Research UK, London

Supervisor: Dr. David Ish-Horowicz FRS

A thesis submitted for the degree of Doctor of Philosophy

University of London

March 2009

I, Inbal Ringel, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

**To my beloved, Tim and Leilani.**

## Preface

I would like to thank David Ish-Horowicz for giving me such a great opportunity to work in his lab. It was always inspiring to listen to and consult with you. I am grateful for the valuable discussion and assistance given to me by members of the Ish-Horowicz lab, especially Mark Wainwright for helping me and teaching me about fly culture. Thank you Barbara, Michael, Chris and Babis for giving me useful comments on parts of my thesis. I would like to thank Daniel, Peter and Alastair from the Light microscopy lab, who helped me with establishing the work on the different microscopes and programmed the tracking and scoring programs. I thank those who have collaborated with me, provided me with reagents and given helpful advice Peter Lukavsky, Trudi Schupbach, Stephen Neidle and Isabel Moraes. Many thanks to Simon Bullock, who taught me the basics of RNA localisation techniques, explained and supported me throughout my studies.

I am eternally grateful for the support of all my family and friends, especially my Mum and Dad, who were supporting me all the way. And last, to my wonderful Tim, who was always there to support me in every aspect, scientifically, technically and emotionally. I cannot express enough my gratitude to you. I could never have done it without you.

Inbal

## Abstract

Asymmetric localisation of cytoplasmic mRNA in the cell appears in a variety of organisms and is important for establishment of spatially differential expression of genes. Localised transcripts typically contain codes (localisation signals), expressed within cis-acting elements that specify subcellular targeting. These signals are recognised by a complex of adaptor and motor proteins that move along the cytoskeleton. Cis-acting elements usually present in the 3'-UTRs of localising transcripts. Different signals do not appear to share either primary or secondary structures that are distinct from non-localising transcripts. Although secondary structure is important, the structural basis of the elements that contribute to the specificity of the localising transcripts is poorly understood.

The presented thesis examines the basis of selective RNA transport, by studying the shortest signal known to drive localisation in *Drosophila Melanogaster*, a 44 nucleotides sequence on the 3' UTR of the fs(1)K10 transcript. This signal is necessary and sufficient for its localisation during embryo development and has a structure of stem loop with two unpaired bases ("bulges"). The components that are important for signal activity are analysed by studying the effect of specific mutations on localisation in the embryo and on the corresponding structure of the RNA. Using NMR structure solution, the importance of the two bulges is demonstrated in wild type and mutated signals and a new structural element of an unusual B-form-like double-stranded RNA helix is revealed. This unusual helix form subsequently been shown to be crucial for the localisation of the K10 transcript. These features might be important as representatives of a general structure that characterise a common group of localising transcripts.

## Table of Contents

Preface.....	4
Abstract .....	5
Table of Contents .....	6
Table of figures .....	11
1 Chapter 1 Introduction .....	12
1.1 RNA localisation is important for establishment of asymmetric cells. .	12
1.2 Localisation of RNA serves an important role in many biological functions.....	13
1.2.1 Localisation of RNA establishes a gradient of proteins .....	13
1.2.2 Differentially localised RNAs can determine cell fate.....	14
1.2.3 Local concentration and local translation of different transcripts can create a local protein complex.....	15
1.2.4 RNA can facilitate protein sorting by localising proteins to the vicinity of organelles .....	16
1.2.5 Human pathologies associated with defective RNA localisation....	19
1.3 Mechanism and Kinetics of RNA localisation .....	22
1.3.1 Diffusion and entrapment of specific RNA.....	22
1.3.2 Degradation of non-localising RNA.....	23
1.3.3 Local synthesis allows local concentration of RNA.....	23
1.3.4 Active directional transport .....	24
1.4 RNA localisation plays a role in different developmental stages in <i>Drosophila</i> .....	25
1.4.1 Early egg chamber and onset of oogenesis (development of the nurse cell and oocyte).....	25
1.4.2 Maternal mRNA is transported from nurse cells to the oocyte .....	25
1.4.3 Stage 7-9 Re-organisation of the cytoskeleton and differential accumulation of mRNA in the oocyte .....	26

1.4.4	Oogenesis Stage 9-14 .....	27
1.4.5	Early zygotic development and asymmetric localisation of mRNA during embryo development.....	27
1.4.6	Asymmetric mRNA localisation also occurs in the neuroblasts .....	29
1.5	Visualisation of RNA .....	29
1.5.1	Fluorescent labelling of RNA.....	31
1.6	The basic participants of the transport Machinery .....	31
1.6.1	Cytoskeletal filaments .....	32
1.6.2	Cytoskeletal motors .....	33
1.6.3	Adaptor proteins .....	36
1.6.4	Transport particles .....	39
1.7	Localisation Signals .....	40
1.7.1	Sequence and structural basis of recognition .....	41
1.7.2	Identifying localisation signals.....	41
1.7.3	Finding localisation signals using sequence similarity .....	42
1.7.4	Finding and comparing LSs based on similar secondary structures ... .....	43
1.7.5	Using related species to identify new motifs.....	44
1.7.6	Examples of localisation signals .....	45
1.8	The K10 Transport Localisation signal (TLS) .....	49
1.8.1	K10 function in the cell .....	49
1.8.2	TLS is one of the shortest signals found in <i>Drosophila</i> .....	50
1.8.3	<i>orb</i> contains a localisation signal that has similar predicted structure to that of K10.....	50
1.9	Understanding the structure of RNA is essential for the understanding of its function.....	51
1.10	Summary .....	52
1.11	Aims of the thesis .....	52
2	Chapter 2 Structural determinants of <i>K10</i> RNA localisation signal.....	59

2.1	Aims .....	59
2.2	Injection assay .....	59
2.3	K10 TLS can be further shortened and still retain localisation activity .... .....	60
2.4	Random Mutagenesis analysis reveals global structural requirement for TLS.....	61
2.5	Single stranded nucleotides play a structural role in recognition of the TLS.....	62
2.5.1	The loop is important for localisation but not its exact size or base identity.....	63
2.5.2	The presence of bulges is important for the TLS structure, but their identity is not crucial .....	63
2.5.3	Further analysis of the function of bulges in the TLS: The structure of orb localisation signal is a stem loop with one bulge .....	65
2.6	Point mutations in nucleotides in the proximity of the bulges .....	66
2.7	Summary .....	67
3	Chapter 3 Evolutionary divergence of <i>K10</i> and <i>orb</i> localisation signals within <i>Drosophila</i> species.....	78
3.1	Aims .....	78
3.2	A comparative study of homologous K10 Transport/Localisation Sequences (TLSs) reveals conserved and diverged regions of the signal... .....	78
3.3	Conservation and divergence of homologous orb localisation signals .	80
3.4	Summary and discussion .....	82
4	Chapter 4 Effect of mutations in the TLS on embryonic development.....	86
4.1	Aims .....	86
4.2	Introduction .....	86
4.3	K10 null flies do not produce offspring and have a dorsalised phenotype .....	87
4.4	Establishing a line of flies with non-localised transcripts.....	88

4.5	Wild-type transgenes rescue the null K10 phenotype .....	89
4.6	A transgene with the U6G mutation did not rescue K10 null phenotype.. .....	90
4.7	A transgene with a mutation in the distal helix rescued the K10 null phenotype .....	90
4.8	Summary and discussion .....	90
5	Chapter 5 Solution structure of K10 localisation signal .....	95
5.1	Aims .....	95
5.2	Introduction .....	95
5.3	Each bulge adopts a different conformation relative to the helix.....	97
5.4	An unusual B-like form helix plays a role in localisation .....	98
5.5	Summary and discussion .....	99
6	Chapter 6 Using live imaging to study the dynamics of different K10 transcripts. ....	105
6.1	Aims .....	105
6.2	Comparison of kinetics of different mutated K10 TLS transcripts with altered helix forms.....	107
6.3	Summary and discussion .....	108
7	Chapter 7 Discussion .....	114
7.1	The structure of RNA localisation signals are likely to be conserved .....	114
7.2	Conclusions from this work .....	116
7.3	Further questions and future prospects.....	117
8	Chapter 8 Materials and Methods .....	124
8.1	Molecular Biology.....	124
8.1.1	Cloning and Constructs preparation .....	124
8.2	RNA injections into syncytial blastoderm embryos.....	127
8.2.1	In vitro transcription and fluorescent labelling .....	127
8.2.2	Blastoderm injection assay .....	128

8.3	Confocal microscopy.....	129
8.3.1	Assaying mRNA Transport in Time-lapse fluorescence microscopy . .....	130
8.4	Fly culture.....	131
8.4.1	Alleles.....	131
8.4.2	Transgenic flies for rescue assay.....	131
8.4.3	Egg hatching assay .....	132
8.5	Solutions and buffers.....	132
8.6	Primers List .....	134
9	References.....	140

## Table of figures

FIGURE 1_1 EXAMPLES FOR LOCALISED RNA IN DIFFERENT CELLS .....	54
FIGURE 1_2 DIFFERENT MECHANISMS FOR MRNA LOCALISATION .....	55
FIGURE 1_3 STAGES OF DEVELOPMENT OF DROSOPHILA AND LOCALISED RNAs IN EACH STAGE.....	56
FIGURE 1_4 SUGGESTED MODEL FOR THE MOTOR COMPLEX THAT DRIVE MRNA LOCALISATION.....	57
FIGURE 1_5 LOCALISATION OF <i>K10</i> MRNA IN <i>DROSOPHILA</i> . .....	58
FIGURE 2_1 INJECTION ASSAY AND REPRESENTATIVE EXAMPLES OF SCORING METHOD. ....	71
FIGURE 2_2 REPRESENTATIVE MUTATIONS IN THE <i>K10</i> TLS. ....	72
FIGURE 2_3 RANDOM MUTAGENESIS OF THE TLS SEQUENCE IMPLIES THAT THE DISTAL BULGE IS DISPENSABLE FOR TRANSPORT.....	73
FIGURE 2_4 THE BULGES SERVE A STRUCTURAL ROLE.....	74
FIGURE 2_5 HYBRIDS OF <i>K10</i> TLS AND <i>ORB</i> LS HAVE A DIFFERENT LOCALISATION PATTERN. ....	75
TABLE 2_1 SUMMARY OF POINT MUTATIONS IN THE TLS.....	77
FIGURE 3_1 PREDICTED TLS SECONDARY STRUCTURE IS CONSERVED. ....	84
FIGURE 3_2 COMPARISON OF TLS LOCALISATION SIGNALS FROM DIFFERENT <i>DROSOPHILA</i> SPECIES .....	85
FIGURE 4_1 EGG SHELL PRODUCED BY WILD-TYPE AND <i>K10</i> <sup>LMOO</sup> FEMALES.....	93
TABLE 4_1 ANALYSIS OF EMBRYO DEVELOPMENT FOR THE PROGENY OF THE RESCUE FLIES.....	94
FIGURE 5_1 NMR SOLUTION STRUCTURE OF THE BULGES OF TLS. ....	101
FIGURE 5_2 NMR SOLUTION SHOWS AN UNUSUAL B-LIKE DOUBLE HELIX. ....	102
FIGURE 5_3 VIEW FROM ABOVE TOWARDS THE LOWER (A) AND UPPER (B) HELIX OF <i>K10</i> -WT RNA .....	103
FIGURE 5_4 MUTANTS IN THE UPPER AND LOWER STEMS OF <i>K10</i> AND <i>ORB</i> .....	104
FIGURE 6_1 SEMI-AUTOMATIC TRACKING OF THE PARTICLES. ....	111
FIGURE 6_2 KINETIC PARAMETERS OF TRANSCRIPT TRANSPORTS.....	112
TABLE 6_1 LOCALISATION OF INJECTED WILD-TYPE <i>K10</i> TLS COMPARED WITH LOCALISATION OF OTHER TRANSCRIPTS.....	113
FIGURE 7_2 A-FORM AND B-FORM HELIX .....	123
SUPPLEMENTARY FIGURE 2_1 LOCALISATION SCORING BY QUANTILE RATIO PROGRAM (QR).....	137
APPENDIX FIGURE 5_1 CRYSTALLIZATION OF RNA.....	138

# 1 Chapter 1 Introduction

One central question in biology is how a single cell becomes a multicellular organism, composed of many different cell types. Asymmetrical localisation of proteins is fundamental to this process, both for asymmetrical cell division and cell polarity. The localisation of specific mRNAs is one mechanism used to create an asymmetric expression of proteins. In this introduction, I will explain the importance of this process, giving examples of different organisms in which RNA localisation serves different functions, and review what is currently known about the mechanisms of RNA localisation, and about the sequence and structural specificity of localisation elements. Finally, I will explain why the localisation signal of *female sterile(1)K10 (K10)* RNA was chosen for a study of its structure, and how studying its structure can contribute to a better understanding of RNA structural elements and RNA-protein recognition.

## **1.1 RNA localisation is important for establishment of asymmetric cells.**

Targeting proteins to different regions of the cell, both to specific sites where their function is required, and away from sites where their expression would cause harm is a fundamental process in every cell. Different mechanisms used by cells to localise proteins include signal peptides, intragenic export signals and different post-translation modulators (Silhavy et al., 1983). One of these mechanisms is the localisation of the RNA, rather than the protein itself, and it is abundant amongst all eukaryotes, from unicellular to multicellular organisms.

The first reports of RNA localisation in eukaryotic cells appeared 40 years ago (Leontis et al., 2002; Westhof and Fritsch, 2000), but it was only in the early 1980s that the notion of RNA subcellular localisation as a key mechanism for polarisation of cells came into view. The first cytoplasmically localised RNAs were identified in *Xenopus* oocytes in a screen for RNAs enriched in either the vegetal (*Vg* RNAs) or animal (*An* RNAs) hemisphere of the oocyte (Rebagliati et al., 1985; Weeks et al., 1985). Shortly afterwards, *bicoid (bcd)* RNA localisation at the anterior pole of the *Drosophila* oocyte and early embryo was demonstrated (Berleth et al., 1988; Frigerio et al., 1986). Studies of *bcd* showed that the localisation *per se* of the RNA is important for the appropriate development of the embryo (Driever and Nusslein-Volhard, 1988a, b). Since then,

subcellular localisation of RNA has emerged as a key mechanism through which cells become polarised. Many more RNAs were found in different organisms, including *Xenopus* oocytes (Mowry and Melton, 1992), Budding yeast (Gonsalvez et al., 2005), zebrafish (Yoon et al., 1997), mosquitoes (Juhn et al., 2008), hydrozoan jellyfish *Clytia hemisphaerica* (Momose et al., 2008; Momose and Houliston, 2007), and mammalian cells like fibroblasts (Lawrence and Singer, 1986) and neuronal cells (Job and Eberwine, 2001).

## **1.2 Localisation of RNA serves an important role in many biological functions**

There are several possible reasons why localisation of the mRNA is advantageous over protein localisation within the cell (St Johnston, 2005). First, it is more energy efficient, as translation of one molecule of mRNA can produce many molecules of protein. Second, localisation of RNA, rather than of protein, prevents the activity of the protein in parts of the cell where it is not needed, or can be harmful. The motor complex driving the localisation may include factors that regulate translation that enable the RNA to be translated only once it reached its final destination. Finally, the timing and spatial control of translation is executed locally, and allows fast expression of local proteins. This enables a rapid response to local requirements of the cell, and allows an independent expression of the gene in different parts of the cell.

### **1.2.1 Localisation of RNA establishes a gradient of proteins**

In the development of the *Drosophila* embryo, different genes act in concert to establish and position axis determinants and to create an asymmetric expression of proteins in the cell. This asymmetric distribution of proteins would later contribute to establish a multicellular fly. Examples of these proteins are Bcd, and Oskar, which RNAs are localised to the anterior and posterior of the *Drosophila* embryo, respectively.

Bcd was the first protein to be shown to act as a morphogen, in which a protein is spread from a localised source and forms a concentration gradient that patterns tissue development. (Driever and Nusslein-Volhard, 1988a, b; Ephrussi and St Johnston, 2004). *bcd* mRNA is tightly localised to the anterior cytoplasm of the *Drosophila* oocyte and embryo (Figure1\_1A). After fertilisation, the translated protein diffuses posteriorly and forms a gradient along the anterior-posterior axis of the embryo. Bcd is a DNA-binding transcription factor. Thus, a gradient of Bcd acts differentially on

different genes, the so-called "gap" and "pair-rule" genes, to position their transcription along the anterior-posterior axis (Macdonald and Struhl, 1986; Struhl et al., 1989).

Similarly, localisation of other mRNAs positions the proteins responsible for posterior axis determinations in their specific location, and, in turn, affects other genes to be expressed locally. The transport of *oskar* (*osk*) mRNA to the posterior of the embryo initiates formation of pole plasm (Ephrussi et al., 1991), where other localised mRNAs are essential for later development of the primordial germ cells (Mahowald, 2001). One of the posteriorly localised mRNAs is *nanos* (*nos*) mRNA, which, upon translation, gives rise to a second morphogen gradient that defines posterior structures of the embryo (Wang and Lehmann, 1991). The dorsoventral axes of the embryo are determined by a third localised mRNA, *gurken* (*grk*) (Nilson and Schupbach, 1999).

### 1.2.2 Differentially localised RNAs can determine cell fate

In budding yeast, *ASH1* mRNA is localised to the distal tip of the daughter cell (Takizawa et al., 1997) (Figure1\_1B) and consequentially, Ash1 protein is expressed in this area. Ash1 is responsible for the suppression of mating-type switching in the daughter cell by repressing the transcription of the HO endonuclease gene. Misexpression of *ASH1* mRNA throughout the cell results in equal distribution of the protein in the nuclei of mother and daughter cells leading to both cells acquiring an identical fate (Bobola et al., 1996). Although more mRNAs were found to localise to the bud or daughter cell tip by a common machinery (Gonsalvez et al., 2005; Shepard et al., 2003), for most of them, targeting to the bud tip is not necessary in order to achieve asymmetric protein distribution (Shepard et al., 2003).

Other cell fate determinants appear during the development of the *Xenopus*, whereby different transcripts encode determinants for differentiation of developing tissues. The *Xenopus* oocyte is divided into a vegetal and an animal pole. Transcripts are localised to the vegetal pole and are later inherited by the progeny cells that stem from this area (King et al., 2005; Kloc and Etkin, 2005). For example, *VegT* mRNA, which encodes a T-box transcription factor, is localised to the vegetal pole and is required for mesendodermal development (Zhang et al., 1998a). Localisation of another mRNA, *Vg1*, has also been implicated in the specification of both mesodermal and endodermal tissues (Melton, 1987; Zhang et al., 1998a).

Another example of cell fate determinants are transcripts that are localised during the development of neuroblasts in the *Drosophila* embryo. Neuroblasts undergo asymmetric divisions to create two different kinds of cells, a ganglion mother cell

(GMC) and a new neuroblast. The fate of the daughter cells is dictated by segregation of cell fate determinants, some of them are also transported as mRNAs (Figure 1\_1C). *Insc* mRNA is transported to the apical cortex in a dynein-dependent process. Its product, Inscuteable (*Insc*) protein, is part of an apically sorted protein complex involved in targeting the determinant proteins (*Numb* and *Prospero*) basally within the cell, (Broadus et al., 1998; Yu et al., 2006). *prospero* mRNA is sorted, during mitosis, to the cell cortex of the basal side, which will subsequently become part of the GMC. In the GMC, *Prospero* protein acts as a transcription factor and determines cell fate. Localisation of *prospero* is mediated by the double-stranded RNA-binding protein *Staufen* and the adaptor *Miranda* that tethers *Staufen* to the cell cortex (Ikeshima-Kataoka et al., 1997; Matsuzaki et al., 1998; Schuldt et al., 1998; Shen et al., 1998).

*prospero* and *inscutable* localisation is not essential for the correct segregation of the daughter cells. Symmetric distribution of *prospero* mRNA, as a result of loss of *Staufen* function does not block asymmetric distribution of *Prospero* protein, as *Miranda* serves to localise *Prospero* protein as well as its mRNA (Schuldt et al., 1998). Failure to localise the *Insc* mRNA results in only a modest defect of asymmetric cell division (Hughes et al., 2004). Hence, localisation of *insc* might serve to facilitate rather than achieve asymmetric protein distribution.

### **1.2.3 Local concentration and local translation of different transcripts can create a local protein complex**

Localised mRNAs can serve as a template for the spatially restricted synthesis of proteins in polarised somatic cells. The local translation is important to maintain cellular asymmetry and to facilitate the compartmentalised assembly of multifactor complexes, by translating all the proteins found within the complex at the same time and place. For example, vertebrate *β-actin* mRNA is localised near the leading edge of lamellipodia in several motile cell types like fibroblasts, myoblasts, and epithelial cells (Figure 1\_1D), where actin polymerisation is necessary for forward protrusion (Kislauskis et al., 1997; Shestakova et al., 2001). Lamellipodia extensions are actin-rich, and polymerising actin filaments provide the protrusive force for their extension during cell motility. A *cis*-acting element on the *β-actin* 3' UTR is important for RNA localisation, and so is Zipcode Binding Protein-1 (ZBP1), which also inhibits the translation of the protein until the message reaches its final destination (Huttelmaier et al., 2005; Ross et al., 1997). The consensus sequence to *β-actin* zipcode element also appears in other mRNAs. These RNAs encode motility-related proteins that localise to the same site

(Mingle et al., 2005). Their proteins form seven subunits of the Actin-Related Protein 2/3 (ARP2/3) complex, which localises to lamellipodia in fibroblasts. The complex is essential for the nucleation of actin filament assembly (Machesky and Gould, 1999; Mingle et al., 2005). As a consequence, the protein  $\beta$ -actin is translated primarily at the sites of mRNA localisation (Rodriguez et al., 2006). Loss of  $\beta$ -actin RNA localisation results in a change in the location of actin polymerisation, and loss of cell morphology and motility (Kislauskis et al., 1997; Shestakova et al., 2001).

The local concentration of RNA enables different monomers of actin filaments to be expressed together in a small, restricted area. Furthermore, the expression of isoforms, which would interfere with the assembly of the correct complex, is prevented, so that only the correct isoforms are localised and translated in the site of translation. ZBP1 interacts preferably with  *$\beta$ -actin* mRNA and not with other actin isoforms (Condeelis and Singer, 2005), thus enabling a restricted expression. In addition, while  $\beta$ -actin is localised to the leading lamellae in differentiating myoblasts, the mRNAs encoding the  $\alpha$ -cardiac actin isoform, which would interfere with lamellae development is associated with a perinuclear compartment (Kislauskis et al., 1993).

Another example of isoform differentiation is the case of Creatine Kinase (CK). CK protein plays an important role in ATP/ADP level regulation, and its spatial expression is important when there are fluctuations in ATP levels in different tissues. Dimerisation of two cytosolic forms of CK, brain (B) and muscle (M), forms the active enzyme. The different dimer combinations are tissue specific, and the specific RNAs co-localise with the enzyme (Schafer and Perriard, 1988). mRNA for the M form is localised at the cell periphery of mouse myoblasts, and the mRNA for the B form is localised in the perinuclear region (Wilson et al., 1995), suggesting that the localisation of the mRNAs for the cytoplasmic isoforms of CK may be involved in the localisation of the enzymes.

#### **1.2.4 RNA can facilitate protein sorting by localising proteins to the vicinity of organelles**

There is ever increasing evidence to support the model that localisation of mRNAs encoding proteins destined for organelles like the endoplasmatic reticulum (ER), mitochondria or nucleus, is important. Localisation of the RNA allows more efficient intake into organelles than protein sorting alone. It also permits a refined targeting of the protein expression.

#### 1.2.4.1 Perinuclear localisation of RNA can serve to enhance nuclear transport

Localisation of mRNAs encoding nuclear proteins to the periphery of the nucleus can facilitate nuclear import. The mRNA for the genes  *$\alpha$ -actin*, *slow troponin C* (*sTnC*), and *slow troponin I* (*sTnI*), which encode different polypeptide partners of the thin filament, localises to the perinuclear cytoplasm of cultured muscle cells (Reddy et al., 2005). The mRNA of transcription factors c-MYC, as well as c-FOS, Metallothionein-1 (MT-1) and cellular retinoic acid binding protein I (CRABPI) accumulate at the nuclear periphery and associate with the perinuclear cytoskeleton (Figure 1\_1E) (Dalglish et al., 2001; Dalglish et al., 1999; Levadoux-Martin et al., 2001; Levadoux-Martin et al., 2006; Levadoux et al., 1999; Mahon et al., 1997; Nury et al., 2005; Veyrune et al., 1996; Veyrune et al., 1997).

Localised concentrations of specific proteins may result from corresponding localisation of their respective mRNAs. Localisation of RNA targets the protein products to the area where they need to be expressed. This is probably the case with *actin*, *Vimentin* and *tubulin*, transcripts encoding cytoskeletal proteins in mammalian cells that are important in different parts of the cytoplasm. *Vimentin* is the subunit of the intermediate filaments (IF) network (Franke et al., 1982). *Vimentin* mRNA exhibits a perinuclear localisation in fibroblasts and myotubes (Bermano et al., 2001; Lawrence and Singer, 1986), which depends on a sequence in its 3'UTR (Bermano et al., 2001). *Vimentin* mRNA localisation is important for optimal filament formation. In contrast, *tubulin* mRNA is concentrated in the peripheral cytoplasm, and *actin* mRNA is distributed to cell extremities, generally in lamellipodia (Lawrence and Singer, 1986).

Metallothionein-1 (MT-1) perinuclear localisation (Mahon et al., 1997) demonstrates a link between the localisation of the RNA to the function of the protein in a protective role against DNA damage and apoptosis induced by external stress (Levadoux-Martin et al., 2001). MT-1 protein is imported into the nucleus upon G1/S phase transition of the cell cycle. Cells with non-localising *MT-1* show lower survival rates following exposure to oxidative stress and chemical agents. Furthermore, these cells showed less DNA damage repair than cells transfected with the full gene (i.e. localised *MT-1*) in response to either hydrogen peroxide or mutagen treatment, and apoptosis was less frequent than in cells without localised MT-1 after exposure to UV light or mutagenesis. Localisation depends on a short sequence in the 3'UTR (Chabanon et al., 2004; Hesketh, 2004; Hesketh et al., 1998; Mahon et al., 1997; Nury

et al., 2005) that is recognised by Elongation factor 1alpha (Mickleburgh et al., 2004; Mickleburgh et al., 2006).

There might be a link between perinuclear localisation of mRNA and its translation. Prevention of the initiation of translation is correlated with the abolishment of localisation activity of the *c-myc* mRNA in mammalian cells, but once the mRNA has been localised, further translation is not required, as prevention of global translation does not have an effect on localisation (Dalglish et al., 1999).

#### **1.2.4.2 Localisation of mRNA to the Endoplasmic Reticulum directs protein into specific sub compartments**

In eukaryotes, mRNAs encoding secreted and integral membrane proteins are targeted to the Endoplasmic Reticulum (ER) to facilitate translation and protein translocation into the ER lumen. In developing rice seeds, different RNAs are localised to different compartments in the ER (Figure 1\_1F). Unlike most plants, which accumulate a single major class of storage proteins, developing rice seeds synthesise two different seed storage proteins, prolamine and globulin-like glutelin. These are stored within different compartments of the endomembrane system. The prolamine and glutelin mRNAs are localised to the ER-derived protein bodies containing prolamines and to cisternal ER, respectively (Choi et al., 2000). RNA localisation dictates the initial site of storage protein synthesis on specific subdomains of the cortical ER. Both prolamine and glutelin mRNAs are targeted via separate RNA-based mechanisms from their site of transcription in the nucleus to distinct subdomains of the ER. The multiple pathways prevent non-productive interactions between different classes of storage proteins that would otherwise disrupt protein sorting.

Key factors involved with RNA localisation can also associate with ER membranes. Yeast She2p, *Xenopus* Vg1 RNA-binding protein and *Drosophila* Staufen, have been found to co-localise with the ER (Allison et al., 2004; Deshler et al., 1997; Herpers and Rabouille, 2004; Schmid et al., 2006). The mechanisms driving their localisation and the reason why components of the RNA localisation machinery are directed to the ER are not yet clear. One suggestion is that association with the ER enables regulation of protein synthesis in areas of new growth (for example during cell division in yeast), or enable confined spatial responses to environmental stimuli (for example in neurons, during synaptic remodelling or in cases of neuronal injury).

### **1.2.4.3 Localisation of mRNAs to the mitochondria might facilitate co-translational import into the mitochondria**

There is some evidence that sorting of a subset of nuclear-encoded proteins to mitochondria involves mRNA localisation. In yeast, *ATP2* mRNA, which encodes a subunit of the mitochondrial ATP synthase, is localised close to the mitochondria. Mislocalisation of the RNA causes severe respiratory deficiency, which might indicate a link between the localisation of the RNA and the function of the protein itself (Margeot et al., 2002).

Interestingly, global analyses of mRNA localisation patterns in yeast, indicate that subcellular localisation of transcripts is associated with the different origins of the genes, and so mRNAs that are of putative bacterial origin are mainly localised and translated on polysomes that are associated with the mitochondrion, whereas those of eukaryotic origin are generally localised and translated on free cytosolic polysomes (Marc et al., 2002; Sylvestre et al., 2003).

### **1.2.5 Human pathologies associated with defective RNA localisation**

RNA localisation plays an important role in the development and formation of the nervous system, as well as the function of the mature nervous system. However, although many diseases are associated with defects in RNA mechanism (Ranum and Day, 2004), few are known to be caused directly from loss of RNA localisation.

A possible explanation for the small number of known RNA-localisation-related diseases is that the localisation process is essential in very early stages of development of the embryo. Loss of function of one of the components in the process might result in lethality of the embryo. Hence, an adult phenotype would not be observed. Indeed, loss of RNA localisation in other model organisms, would lead to infertile mothers, or embryos not being able to hatch. Another explanation could be the fact that RNA components, like many other human genes are redundant, so that the loss of one gene in the pathway could be rescued by the remaining components. Furthermore, not all aberrations in RNA localisation might result in an obvious clinical phenotype, as the process sometimes serves only to fine-tune a protein's localisation and expression, and abolishing localisation would therefore only have weak effects on activity. Aberration in RNA localisation might also be overlooked when searching for genetic defects in patients, as the translated gene might seem to be intact.

There are a few examples of an indirect association, where loss-of-function of one component of the RNA transport machinery or the misexpression of specific mRNAs causes abnormalities during development (Reviewed in (Dahm and Macchi, 2007) and (Bassell and Kelic, 2004)). Defects in axonal and dendritic outgrowth and in dendritic spine development have been correlated with different neurological disorders, such as FXS (Fragile X mental retardation Syndrome), SCA (Spino-Cerebellar Ataxia) and SMA (Spinal Muscular Atrophy).

### **1.2.5.1 Fragile X mental retardation syndrome**

Fragile X syndrome (FXS) is the leading cause of inherited mental retardation. FXS is an X-linked disease caused by loss of expression of the *FMR1* gene, encoding the FMRP (fragile X mental retardation protein) (Bagni and Greenough, 2005; Gantois and Kooy, 2002; Kaytor and Orr, 2001; Pieretti et al., 1991; Verkerk et al., 1991). In addition to impairment of higher-cognitive functions, FXS patients show a variety of physical and other mental abnormalities. Pathological studies from the brains of patients and from *Fmr1* knockout mice show abnormal neuronal dendritic spine morphology, implicating FMRP in synapse formation and function. Acute suppression of FMRP and target mRNA transport in neurons resulted in altered filopodia-spine morphology that mimicked the FXS phenotype (DICTENBERG et al., 2008).

Recent findings link the impairment of stimulus-induced dendritic mRNA transport in a mouse model of FXS to altered developmental morphologic plasticity (Bassell and Warren, 2008; DICTENBERG et al., 2008). DICTENBERG et al. report a function for the FMRP protein in the rapid, activity-regulated transport of mRNAs important for synaptogenesis and neuronal plasticity. mRNAs from *Fmr1* knock-out mice were deficient in dendritic localisation in neurons, and single mRNA particle dynamics in live neurons revealed diminished kinesin. Translocation of FMRP and cognate mRNAs involves the C-terminus of FMRP and kinesin light chain, and *Fmr1* knockout brain cells showed reduced kinesin-associated mRNAs

FMRP is a selective RNA-binding protein that regulates the local translation of a subset of mRNAs at synapses (Miyashiro et al., 2003) by inhibiting translation initiation (Napoli et al., 2008). In the absence of FMRP, excess and dysregulated mRNA translation leads to altered synaptic function. FMRP contains four types of RNA-binding domains; an arginine/glycine-rich region (Darnell et al., 2001; Zalfa et al., 2007), two ribonucleoprotein K homology domains, (Darnell et al., 2005a; Darnell et al., 2005b) and a domain in the N-terminus containing a Tudor motif which binds the

neuronal *BC1* and *BC200* (Brain Cytoplasmic 1 and 200) RNAs (Gabus et al., 2004; Zalfa et al., 2005). The binding of FMRP to *BC1* RNA is believed to recruit mRNAs complementary to *BC1* and maintain the FMRP mRNP (messenger ribonucleoprotein) particle in a translationally silent status (Veneri et al., 2004; Zalfa et al., 2003). Microarray (Brown et al., 2001) and APRA (antibody-positioned RNA amplification) (Miyashiro et al., 2003) analyses have identified hundreds of potential FMRP cargo RNAs. These include *CaMKII $\alpha$*  (Hou et al., 2006; Zalfa et al., 2003), and *Trailer Hitch* in *Drosophila* (Monzo et al., 2006)

Other evidence that link FMRP to RNA localisation pathways is the interaction of FMRP with known *trans*-acting factors, such as Pura $\alpha$  (Ohashi et al., 2002), Staufen1 (Brendel et al., 2004; Ohashi et al., 2002) and ZBP1 (Rackham and Brown, 2004). FMRP also interacts with motor proteins, for example kinesin (Kanai et al., 2004) and regulation of the actin cytoskeleton is under dFMRP control (Reeve et al., 2005).

### **1.2.5.2 Spinal Muscular Atrophy (SMA)**

Spinal muscular atrophy (SMA), the most common inherited cause of infant death, is a neurodegenerative disease that affects motor neurons (Frugier et al., 2002). SMA has been linked to recessive (often deletion) of the survival of motor neuron protein gene (SMN1) that results in the expression of a truncated and unstable isoform lacking the carboxy-terminal exon-7 (Lefebvre et al., 1995; Lefebvre et al., 1997).

SMN is an assembly factor that promotes high fidelity and specific interaction between RNA binding proteins and their target sequences (Paushkin et al., 2002). For example, it binds directly to the mRNA binding protein hnRNP-R (Rossoll et al., 2003; Rossoll et al., 2002) that associates with  *$\beta$ -actin* mRNA and enhances its localisation.

Live cell imaging of EGFP-SMN granules demonstrated rapid, bi-directional, and cytoskeletal-dependent movements within neurites and growth cones of developing neurons. As motor neurons have unusually long axons, their localised mRNPs could be more susceptible to low levels of SMN. Motor neurons cultured from an SMA mouse model have shorter axons and smaller growth cones, which suggest an inefficiency of SMN associated RNPs in motor neuron axons in SMA. These growth cones also have reduced levels of  *$\beta$ -actin* mRNA and protein (Rossoll et al., 2003).

### **1.2.5.3 Spinocerebellar Ataxia (SCA)**

SCAs are a class of inherited neurodegenerative diseases; among them is SCA8. SCA8 is caused by mutations in CUG repeats in the *ATXN8OS* (Ataxin 8 Opposite

Strand) gene (Koob et al., 1999; Ranum and Cooper, 2006). The *ATXN8OS* gene encodes a non-coding RNA that is localised in neurites (Mutsuddi et al., 2004). The RNA is associated with Staufen, a conserved RNA-binding protein that is known to participate in RNA localisation pathways, *ATXN8OS* RNA recruits endogenous Staufen protein to the site of injection when injected into *Drosophila* early embryos via a region that contains the non-expanded CUG-repeat site. The localisation of the RNA in neurites, and its association with Staufen via CUG repeats could suggest a link between the symptoms of SCA8 and defects in the localisation of *ATXN8OS* RNA.

### **1.3 Mechanism and Kinetics of RNA localisation**

Different models have been proposed for the mechanism that leads to localisation of RNA. To determine which one of the modes of localisation is employed to localise a specific molecule, one can study the transported particles. The speed and kinetics of the particles can determine whether it is diffusive or a fast active movement. The proteins that bind, directly and indirectly, to the RNA, can also suggest the type of localisation for example, whether localisation depends on motor proteins, or on proteins that protect from degradation. Lastly, ATP uptake can also provide valuable information, as active transport consumes high energy. (reviewed in (Lipshitz and Smibert, 2000) and (Palacios, 2007)). The following models for localisation include random diffusion of the molecules, sorting by local protection, local synthesis of the RNA and directional transport (Figure 1\_2).

#### **1.3.1 Diffusion and entrapment of specific RNA**

It is thought that some transcripts randomly diffuse through the cytoplasm of cells but become localised once they encounter an anchor that traps them at the site of localisation (Figure 1\_2A). Posterior localisation of *nos* in late oogenesis occurs at a stage when the cytoskeleton shows no anterior–posterior (AP) polarity, and *nos* transcripts move in the oocyte cytoplasm together with cytoplasmic flow. Localisation of *nos* in that stage depends on the entrapment of the transcripts in specific sites. Indeed, most of *nos* molecules are not localised, and only a small fraction of the diffused transcripts, around 4%, is trapped by an actin-dependent association with the germ plasm (Bergsten and Gavis, 1999; Forrest and Gavis, 2003). The entrapment mechanism, together with local and timely translation of *nos* (Rangan et al., 2009), ensures that Nanos activity is properly restricted to the posterior pole of the embryo. Nuclei entering the germ plasm capture the *nos* RNA, and during subsequent nuclear

divisions in some cells, RNA is partitioned asymmetrically such that nuclei that receive only small amounts of *nos* adopt somatic fates.

The same mechanism has been proposed for the vertebrate RNAs encoding *nos*-related proteins, Xcat2 and Xdazl. These transcripts are localised to the mitochondrial cloud in *Xenopus* oocyte in the vegetal pole during early stages of oogenesis (Chang et al., 2004). Similar to Nanos, these proteins are important for germ cell lineage.

### 1.3.2 Degradation of non-localising RNA

The degradation of non-localising RNAs is another mechanism used to promote protein localisation within cells. Transcripts are produced and then rapidly degraded where they are not associated with a localised anchor. However, there are only few examples of the use of this mechanism, probably because it is very “wasteful”, as it requires transcription of much more RNA than needed. One example is the *hsp83* mRNA which is uniformly distributed in the early stages of the fertilised *Drosophila* egg (Bashirullah et al., 2001). As nuclear division advances, *hsp83* is degraded in the cytoplasm, but not in the pole plasm where the RNA is protected (Ding et al., 1993) (Figure 1\_2B). Transcript degradation is regulated by the RNA-binding protein SMAUG (SMG) through a sequence called Hsp83 Instability Element (HIE), in the open reading frame (ORF) (Semotok et al., 2008). SMG destabilises maternal *hsp83* mRNA by recruiting the CCR4/Not deadenylase complex to trigger decay (Semotok et al., 2005; Tadros et al., 2007).

Another example for protection and degradation of the transcripts is the posterior localisation of *nos* in earlier stages of the embryo. *nos* RNA in the posterior pole cytoplasm is higher than elsewhere in the embryo due its protection against being targeted by Smaug for deadenylation and degradation (Zaessinger et al., 2006).

### 1.3.3 Local synthesis allows local concentration of RNA

In some systems, transcripts are exported from the nucleus, diffuse locally and are trapped at the site of final localisation. This usually occurs in multinucleate/syncytial cells. In myofibres, the localised expression of the mRNAs encoding for the  $\beta$  and  $\delta$  subunit of the acetylcholine receptor (AChR) are exported from the nuclei close to the synapse at the neuromuscular junction (NMJ). The local accumulation allows the fast and local expression of the receptor at the synapses. The expression of AChR subunit genes is regulated by locally acting factors from the nerve and by muscle activity (Fromm and Rhode, 2004; Meier et al., 1998).

### 1.3.4 Active directional transport

Many of the best-characterised examples of mRNA localisation occur by active transport along the cytoskeleton. This mechanism requires a functional cytoskeleton and the activity of motor proteins that move mRNAs along cytoskeletal filaments (Tekotte and Davis, 2002). The localising mRNA initially assembles with mRNA-binding proteins that recognise mRNA localisation signals within the transcript. The resulting ribonucleoprotein (RNP) complex is then associated with cytoskeletal motors (myosin, dynein, kinesin), and directed transport occurs along the cytoskeletal fibres. Finally, mRNAs are anchored by specific anchorage complexes, which prevent diffusion away from the site of localisation (Delanoue and Davis, 2005). Examples of mRNAs that localise via active and directed transport are plentiful (Bashirullah et al., 1998; Kloc et al., 2002; Martin and Ephrussi, 2009; St Johnston, 2005). Active directed transport is determined by different parameters:

- The movement of actively transported localisation particles is faster than that of diffusing particles. If a cargo moves directionally at a speed that is too fast to be explained by cytoplasmic flow (i.e. faster than  $60 \text{ nm sec}^{-1}$ ), it is presumably being transported by a motor. Studies tracking the behaviour of specific localised mRNA in living cells indicate that their movement is not due to diffusion by measuring the speed and directionality of the transcripts (Bratu et al., 2003; Bullock et al., 2006; Clark et al., 2007; Fusco et al., 2003).
- Transport proteins consume energy in order to move along the cytoskeletal fibres and depend on active hydrolysis of ATP (Schnitzer et al., 2000), hence in theory, ATP uptake can be an indicator for an active localisation, which can be detected by using ATPase inhibitors (Weil et al., 2006). This however, is harder to define accurately since other modes of localisation, as well as other general functions in the cell also utilise energy.
- The best evidence for active transport is the dependency on and co-localisation with motor proteins and cytoskeletal fibres. mRNAs have been shown to be transported along actin filaments or microtubules, and by members of all three main families of motor proteins — the myosins, dyneins and kinesins. Agents that trigger disassembly of the cytoskeleton (e.g. which depolymerises microtubules) (Edgar et al., 1987), or inhibit molecular motors (e.g. antibodies against dyneins) (McGrail and Hays, 1997; Sharp et al., 2000) block this RNA transport, demonstrating that it is dependent on active transport machinery.

## **1.4 RNA localisation plays a role in different developmental stages in *Drosophila***

Formation of the embryonic axis in *Drosophila* is a direct consequence of asymmetric expression of proteins throughout oogenesis. Studies in *Drosophila* during oogenesis have provided much information on the mechanisms and function of mRNA localisation. Different mRNAs are localised in different tissues by similar proteins using a conserved mechanism (Karlin-Mcginness et al., 1996). The dynein motor complex, which includes the proteins Egalitarian (Egl) and Bicaudal D (BicD), is used to transport maternal RNA from the nurse cells into the oocyte, to drive apical transport of transcripts in the syncytial blastoderm embryo (Bullock and Ish-Horowicz, 2001) and to transport mRNAs in the neuroblast (Hughes et al., 2004). This section describes the dynamics and patterns of subcellular distribution of cytoplasmically localised RNAs in *Drosophila*, in order to provide a cellular and developmental context for the process.

### **1.4.1 Early egg chamber and onset of oogenesis (development of the nurse cell and oocyte)**

Each of the two *Drosophila* ovaries has about 16 ovarioles, each arranged so that oocyte development proceeds from anterior to posterior. At the anterior tip of the ovariole is the germarium, in which the oogonial stem cells divide asymmetrically to produce a stem cell and a committed cell, the cystoblast. Cystoblasts undergo synchronous rounds of incomplete division to form a cyst of 16 cells interconnected by actin-rich ring canals and cytoplasmic bridges. Only 1 of the 16 cystocytes becomes the oocyte, and the remaining 15 become nurse cells (Figure 1\_3A). Each 16-cell germarial cyst becomes surrounded by somatically derived follicle cells to form an egg chamber. Although the oocyte is translationally active, it is transcriptionally inactive through most of oogenesis, until stage 10 (King and Burnett, 1959) and nurse cells provide it with maternal proteins and RNA required for development. Many of the supplied molecules will also be needed during early zygotic development. (Bashirullah et al., 1998; Riechmann and Ephrussi, 2001).

### **1.4.2 Maternal mRNA is transported from nurse cells to the oocyte**

Inside the pro-oocyte lies the only Microtubule Organizing Center (MTOC) in the 16-cell complex. The microtubule-based cytoskeleton that connects the 16 cells through the ring canals is polarised so that the minus end is anchored in the MTOC and faces the oocyte. At this stage, the MTOC, is at the posterior end of the oocyte

(reviewed in (Cooley and Theurkauf, 1994)). This gives the first indication of the importance of a polarised microtubule network for mRNA transport and localisation during oogenesis.

During stages 1-6 of oogenesis, nurse cells transfer proteins and RNAs into the oocyte and as a consequence, the oocyte cytoplasm grows to 40 times its original size. At these stages, maternal mRNAs are transported to and accumulate within the oocyte including, *bcd* (St Johnston et al., 1989), *Bicaudal-D (BicD)* (Suter et al., 1989), *Bicaudal-C (BicC)* (Mahone et al., 1995), *egalitarian (egl)* (Mach and Lehmann, 1997), *gurken (grk)* (Neuman-Silberberg and Schupbach, 1993), *female sterile (1) K10 (K10)* (Cheung et al., 1992; Serano et al., 1994), *oo18 RNA-binding protein (orb)* (Lantz et al., 1992) and *oskar (osk)* (Ephrussi et al., 1991; Kim-Ha et al., 1991). At these stages, mRNAs appear throughout the oocyte cytoplasm, although slightly higher concentrations are present at the posterior cortex, the site of the MTOC.

### **1.4.3 Stage 7-9 Re-organisation of the cytoskeleton and differential accumulation of mRNA in the oocyte**

During stage 6, signalling between follicle cells and oocyte establishes anterior-posterior polarity within the oocyte. Signalling between the follicle cells and the oocyte also results in a reorganisation of the cytoskeleton, so that microtubule minus ends are towards the anterior cortex of the oocyte, and the MTOC disappears from the posterior of the oocyte (Gonzalez-Reyes et al., 1995; Roth and Schupbach, 1994; Theurkauf et al., 1992). As a consequence of this change in cytoskeletal organisation, mRNAs are also re-distributed (Figure 1\_3B). Transcripts that were initially enriched at the posterior cortex localise in a crescent pattern at the anterior margin of the oocyte. These include *K10* (Cheung et al., 1992) as well as *bcd* (Berleth et al., 1988), *BicD* RNA, and its protein product (Suter et al., 1989), *BicC* (Mahone et al., 1995), *egl* RNA and protein (Mach and Lehmann, 1997), *grk* (Neuman-Silberberg and Schupbach, 1993), *orb* (Lantz et al., 1992), *osk* (Ephrussi et al., 1991; Kim-Ha et al., 1991) and *nos* (Wang and Lehmann, 1991). At the same time, *grk* re-localises to the dorso-anterior corner around the oocyte nucleus (Neuman-Silberberg and Schupbach, 1993). *osk* RNA that was previously localised to the anterior of the oocyte through minus-end-directed microtubule motors, must dissociate from these motors and associate with plus-end-directed motors in order to be translocated to the posterior pole (Ephrussi et al., 1991; Kim-Ha et al., 1991).

#### 1.4.4 Oogenesis Stage 9-14

During stages 9-12 of oogenesis, large quantities of material are transferred from the nurse cells to the oocyte. This bulk cytoplasmic transport is termed “nurse cell dumping”. At stage 10b, the polarised MT array disassembles, and short subcortical MTs direct vigorous ooplasmic streaming (Gutzeit, 1986), a process that depends on the actin nucleators Cappuccino and Spire (Dahlgaard et al., 2007). *nos* transcripts move with the cytoplasmic stream and localise at the posterior by anchoring to the germ plasm (see 1.1.3.1). Streaming persists until stage 12.

In the final two stages of oogenesis, the subcortical MT arrays are replaced by short, randomly oriented MT filaments. *bcd*, *osk*, and *grk* RNAs are anchored to the oocyte cortex, to prevent perturbation to their localisation after fertilization (Berleth et al., 1988; Glotzer et al., 1997; St Johnston et al., 1989), but other localised transcripts such as *BicD* (Suter et al., 1989), *BicC* (Mahone et al., 1995), *orb* (Lantz et al., 1992) and *egl* (Mach and Lehmann, 1997) become uniformly distributed throughout the oocyte cytoplasm. *osk* mRNAs localise posteriorly (Ephrussi et al., 1991; Kim-Ha et al., 1991; Kwon et al., 2002).

#### 1.4.5 Early zygotic development and asymmetric localisation of mRNA during embryo development

The source of the polarity of the early embryo is in the mature oocyte, which, at this stage, has distinct anterior and posterior ends. Embryogenesis is activated by fertilization and followed by 13 rapid synchronous mitotic divisions (cycles), which occur about every 10 minutes, without cytokinesis. Nuclear division is not accompanied by cell division. By cycle 10, nuclei have migrated to the periphery of the embryo, but cell membranes between them have not completely formed. This is termed the syncytial blastoderm embryo. The syncytial blastoderm has around 6000 nuclei sharing a common, maternally inherited cytoplasm (Lodish et al., 2000; Sullivan and Theurkauf, 1995). From cycle 10 onwards, the cytoplasm progressively clears of yolk and the embryo can be divided into several distinct domains along the apico-basal axis. The nuclei are arranged just beneath the plasma membrane and a thin layer of cytoplasm, the apical cytoplasm. The basal cytoplasm and a yolk mass are below the nuclei, within the embryo. Microtubules are found in the cytoplasm, where their minus-ends are nucleated by apically located centrosomes and their plus-ends are found in the basal cytoplasm or the yolk.

71% of the transcripts in early *Drosophila* embryos are specifically localised to different compartments in the cell (Lecuyer et al., 2007). There are two main kinds of localisation. One belongs to pair-rule genes that are expressed in stripes along the embryo (Figure 1\_3B), and the second kind represents different genes that are localised to the anterior and posterior axis (Figure 1\_3C).

The pair-rule genes initially establish a repeating, segmented pattern in the embryo. They encode transcription factors that are expressed in partially overlapping stripes along the A-P axis of the syncytial blastoderm embryo (Jackle et al., 1992; Pankratz and Jackle, 1990). mRNA transcripts for almost all the pair-rule genes, including *hairy (h)* (Ingham et al., 1985), *fushi-tarazu (ftz)* (Hafen et al., 1984) and *even-skipped (eve)* (Macdonald et al., 1986), accumulate specifically in the apical cytoplasm above the layer of peripheral nuclei. Studies have found that pair-rule mRNAs target their protein products apically, in close proximity to the nuclei. In turn, this augments nuclear pair-rule protein levels and modulates their activity, increasing the reliability of the embryonic segmentation process (Bullock et al., 2004).

mRNA transcripts of other zygotically expressed genes are also found localised to the apical cytoplasm in blastoderm embryos and epithelial cells. *crumbs (crb)* (Tepass et al., 1990), which encodes an apical transmembrane protein, is initially expressed in blastoderm embryos just prior to cellularisation. *crb* mRNA transcripts are found localised to the apical cytoplasm in blastoderm embryos and in epithelial cells following cellularisation. Crb protein forms part of an apically localised complex that controls epithelial polarity (reviewed in (Tepass et al., 2001)); *crb* mutants have disrupted epithelial organisation together with cell death in these tissues, which leads to the absence of larval cuticle.

Patterning of the anterior-posterior (A-P) axis of the *Drosophila* embryo involves a cascade of zygotically expressed genes that control the segmental pattern of the embryo (reviewed in (Akam, 1987; Spradling, 1993)). A-P patterning is initiated by the anterior localisation of *bcd* mRNA, and the posterior localisation of *nos* and *osk* mRNAs, which is already evident early in oogenesis. (Ephrussi et al., 1991; Kim-Ha et al., 1991; St Johnston et al., 1989). After fertilisation, *bcd* and *nos* mRNAs are translated at the sites of mRNA localisation. *osk* is translated as soon as it localises posteriorly. It is required to make pole plasm and for anchorage of other RNAs, including *nos*. Bcd and Nos proteins then diffuse through the embryo, establishing gradients of Bcd and Nos transcription factor activity along the A-P axis (see section 1.1.1.1.) (Ephrussi and St Johnston, 2004). The highest concentration of Bcd protein

activity anteriorly results in the formation of head structures, whilst Nos activity at the posterior pole specifies abdominal structures. Similarly, Osk protein activity at the posterior of the embryo is required for formation of posterior structures, such as the abdominal region and the primordial germ cells (reviewed in (Bashirullah et al., 1998; St Johnston and Nusslein-Volhard, 1992)).

Localisation of the RNA is regulated by many factors. At this stage, these factors include the Exuperantia and Swallow proteins that regulate *bcd* localization; Cappuccino, Spire, or Stauf, that regulate the localisation of *osk*; and Notch Delta and germ-line Protein Kinase A that are important in localisation of both *oskar* and *bcd* (Lane and Kalderon, 1994; Ruohola et al., 1991).

#### **1.4.6 Asymmetric mRNA localisation also occurs in the neuroblasts**

Neuroblasts are the stem cell-like precursors of the *Drosophila* central nervous system. They undergo asymmetric divisions along the apico-basal axis to produce daughter cells with different fates. *inscuteable* (*insc*) is a key regulator of polarity in the neuroblast. Insc protein and *insc* mRNA transcripts localise apically in these cells (Hughes et al., 2004; Knirr et al., 1997; Li et al., 1997). *insc* transcripts also localise apically in epithelial cells of the procephalic neurogenic region (PNR) (Hughes et al., 2004; Knoblich et al., 1999; Foe, 1989), which are located in the head region of the embryo and are involved in formation of neuroblasts in the optical lobe (Egger et al., 2007) and giving rise to the brain (Campos-Ortega, 1997).

Apical localisation of *bazooka* (*baz*), *miranda* (*mira*) and *prospero* (*pros*) mRNA transcripts in neuroblasts has also been described. *pros* transcripts are subsequently relocated to the basal side of the neuroblast during mitosis. The translocation is required for sufficient Pros activity in the basal daughter cell upon cytokinesis (Broadus and Doe, 1997; Schuldt et al., 1998; Shen et al., 1998).

### **1.5 Visualisation of RNA**

Visualisation of RNA is important for the study of RNA movement in cells (reviewed in (Rodriguez et al., 2007)). It is important, for example, for the understanding of the mechanisms of the RNA movement, finding new localised transcripts, or identifying associated factors. One common method is Fluorescence *in situ* hybridization (FISH) to native or reporter RNA in fixed cells. This method, although very useful, cannot follow the localisation of transcripts in living cells. Another technique for labelling endogenous mRNA localisation complexes *in vivo* is to

generate transgenic lines that express a green fluorescent protein (GFP)-tagged version of an RNA-binding protein that specifically associates with the mRNA, and follow the RNA-protein complex. A disadvantage of this approach is that most RNA-binding proteins associate with more than one mRNA species, and it is therefore necessary to show that any moving particles that are observed contain the RNA of interest.

To solve potential problems caused by the promiscuity of endogenous RNA-binding proteins, another method was developed, that consists on inserting multiple copies of the recognition site for a heterologous, sequence-specific RNA-binding protein into the mRNA of interest, and to co-express this protein as a GFP fusion protein. This idea is the basis of the MS2-GFP system. A reporter RNA, containing MS2 stem-loop aptamers can be tracked by co-expressing an MS2-coat protein-GFP fusion that would bind to the aptamers. To increase binding efficiency, and sensitivity, up to 24 binding site can be attached. This strategy is effective in both yeast and mammalian cells (Bertrand et al., 1998; Fusco et al., 2003), and allows visualisation of fully processed mRNA in multiple cell types.

A recent method to label endogenous mRNAs directly in live cells and follow mRNAs whose localisation requires nuclear events, such as pre-mRNA splicing, is molecular beacons (Tyagi et al., 1998; Tyagi and Kramer, 1996). This technique involves injecting an oligonucleotide probe (“molecular beacon”) complementary to the native RNA. The probe consists of sequence that can fold into a stem-loop. The loop region is complimentary to the target sequence. At the 5' end of the stem a dye is attached that fluoresces in the presence of a complementary target, while at the 3' end a quencher is attached (non fluorescent). In the unbound form, the self-complementary sequences at the 5' and 3' ends of the probe induce it to fold in a hairpin. The quencher dye is then close enough to the fluorophore to prevent it from emitting light, and fluorescence is quenched. Upon binding to the target RNA, the hairpin probe unfolds, and the quencher is removed from the vicinity of the fluorophore, resulting in a fluorescent signal that allows tracking of the target molecule. A major problem in this technique is when the molecular beacons overlap with zip codes, they could mask recognition by the transport machinery and preclude identification of *bona fide* zip codes. Also, binding of the transport machinery to localised transcripts may prevent hybridisation of the RNA to the beacon probe, resulting in a loss of signal. Even when bound, the beacon provides low sensitivity, as it only introduces a single fluorophore into the mRNA.

In the Cross-Linking and Immuno-Precipitation (CLIP) technique (Ule et al., 2003), nucleic acids and proteins are first cross-linked, followed by cell lysis and RNase treatment of nucleic acids. The mRNP protein of interest is then immuno-precipitated, and the bound RNA is isolated, cloned, and sequenced (Ule et al., 2005). This method has the advantage of isolating RNA–protein complexes under physiological conditions. This technique could potentially be useful for isolating zip codes from systems in which the RNA-binding protein is undefined, because mRNP complexes are cross-linked prior to purification, so the identity of the RNA binding protein does not need to be known in advance. In principle, any component of the mRNP can be used for immuno-precipitation.

### **1.5.1 Fluorescent labelling of RNA**

Another method which is used to track *in vitro* transcripts is to label them *in vitro* with fluorescent dyes, inject the transcripts into a cell and use confocal microscopy to study their movement (Cha et al., 2001; Forrest and Gavis, 2003; Glotzer et al., 1997; MacDougall et al., 2003). This is a fast and relatively easy technique that allows an immediate *in vivo* assay of transcripts. To use this technique the DNA template is transcribed with a mix of ribonucleotides that contains a small fraction of nucleotides that are modified with a fluorescein molecule. Direct labelling of the mRNA ensures that the observed fluorescence signal only comes from the mRNA of interest, providing specificity.

There are a few problems that can arise with this method. One is that the mRNP complex that forms when the mRNA is injected into the cytoplasm, has not been exposed to the nucleus, and may lack nuclear factors important for localisation. Also the cell that can be damaged during the injection, and so this technique does not suit small cells such as yeast and bacteria, although it works well with larger ones like the *Xenopus* oocyte and the *Drosophila* blastoderm embryo and oocyte (Clark et al., 2007). Although the method does not follow endogenous transcripts, it is useful for studying and comparing mutated and wild-type mRNAs, and so I have used it in my work to track different transcripts of *K10*.

## **1.6 The basic participants of the transport Machinery**

The current model proposes that the RNA transport machinery consists of RNA cargo, adaptor proteins that can bind to localisation signals and to motor proteins, and the different motor proteins that can move along cytoplasmic fibres (Figure 1\_4). The

combination of the many proteins and subunits and the cargo create RNP complexes that are large and modular (Kress et al., 2004; Wilhelm et al., 2000). Similar transport machinery components can be used to transport organelles, vesicles, pathogens and macromolecules within the cytoplasm and some mRNA targeting factors are conserved between cell types and organisms (see discussion).

### **1.6.1 Cytoskeletal filaments**

The major filaments in the cytoskeleton of most cells include microtubules, actin filaments, and intermediate filaments. Cytoskeletal elements play important roles in maintaining asymmetry in the oocyte-nurse cell cysts, within the oocyte itself, and in the embryo. Inhibitors of microtubule assembly disrupt oocyte determination and growth (Koch and Spitzer, 1983), ooplasmic streaming (Gutzeit, 1986), anterior localisation of mRNAs (Pokrywka and Stephenson, 1991) and anchoring of localised RNA (Delanoue and Davis, 2005). Cytochalasins, which inhibit actin assembly, block bulk transfer of cytoplasm from the nurse cells to the oocyte (Gutzeit, 1991).

#### **1.6.1.1 Microtubules**

Microtubules (MT) are long, cylindrical polymers of  $\alpha/\beta$ -tubulin that can stretch to lengths of up to 25  $\mu\text{m}$ . Microtubules serve as structural components within cells and are involved in many cellular processes including mitosis, cytokinesis, and transport of molecules and organelles over a wide range of distances. (Greber and Way, 2006; Pokrywka and Stephenson, 1995; Quintin et al., 2008). In most cases, MTs are involved in long-range localisation, as opposed to actin networks that are involved in short-distance transport. The MT polymer has a plus-end, the faster growing side, and a slower growing minus-end. This creates a polarity that is important for the establishment of selective transport to one end. In cells, MTs minus ends are nucleated from MTOCs. Microtubules are central to mRNA localisation in oligodendrocytes (Carson et al., 1998), neurons (Job and Eberwine, 2001), *Xenopus* and *Drosophila* oocytes and embryos (Tekotte and Davis, 2002).

In *Drosophila*, at the early stages of oogenesis, MTs in the egg chamber run from the nurse cells into the oocyte so that the minus end lie at the posterior. Signalling events that include Grk signalling to the follicle cells, and the follicle cells signalling back to the oocyte, lead to a reorganisation of the cytoskeletal scaffold of the oocyte. The MTs then extend their plus end toward the posterior pole, and so different RNAs are re-localised accordingly (see section 1.2.5). For example, *K10* transcripts, which

localise to the minus end of MTs, concentrate at the posterior of the oocyte in the early stages and at the anterior cortex after the re-organization of the MT polarity (Cheung et al., 1992). *grk* transcripts move from the posterior of the oocyte into the dorso-anterior corner, close to the nucleus, which also uses the cytoskeleton, to migrate from the posterior cortex to the dorso-anterior corner (Starr, 2007). *nos* transcripts, which do not depend on MTs, stay at the posterior pole (Ephrussi et al., 1991), as they are degraded at the rest of the oocyte.

In the syncytial blastoderm embryo, the MTs are arranged so that the minus ends are nucleated by apically located centrosomes, while their plus end is close to the basal end of the cytoplasm (Foe and Alberts, 1983; Theurkauf et al., 1992) (Figure 1\_3). mRNA transcripts for almost all pair-rule genes accumulate specifically in the apical cytoplasm above the layer of peripheral nuclei (Davis and Ish-Horowicz, 1991). MT (but not actin) filaments are also required for apical anchoring of RNAs in their final destination in the blastoderm embryo (Delanoue and Davis, 2005).

### **1.6.1.2 Actin**

Actin is one of the most highly conserved proteins in nature, differing by no more than 20% in species as diverse as algae and humans. Actin is a thin filament, composed of two intertwined chains. Actin participates in many important cellular functions, including muscle contraction, cell motility and cytokinesis, vesicle and organelle movement, cell signalling, and the establishment and maintenance of cell junctions and cell shape. (Matus, 2000; Montell, 1999; Pantaloni et al., 2001; Pollard and Borisy, 2003). In budding yeast, actin plays the main role in establishing polarity. It appears either as cables, bundles of filamentous actin (F-actin) that run through the cell parallel to the growth axis, or in cortical actin patches, which are dynamic structures found at invaginations of the plasma membrane.

### **1.6.2 Cytoskeletal motors**

Three classes of molecular motors are known today to drive most of the active transport in cells - kinesin, dynein and myosin (Vale and Milligan, 2000). The kinesin and dynein families move toward the plus-end and minus-end of microtubules, respectively, while myosin motors move along actin filaments. The cargo has an important role in regulating the activity of the motors by recruiting accessory factors

(Bullock et al., 2006), and localisation signals on mRNA increases the probability of RNA to bind to a motor complex (Fusco et al., 2003).

The general concept is that short-range transport usually depends on actin filaments and unconventional myosins, while long-range transport depends on microtubule based motors, but there are examples of filament intersections and sharing of the cargo by the different kinds of motors (reviewed in (Ross et al., 2008)). Current models for regulation of cargo transport propose that different motors can work together, and the direction of the movement is determined by the affinity of the motors and the filaments, and by the cargo itself (Gross et al., 2002; Kural et al., 2005; Muller et al., 2008; Zimyanin et al., 2008). For example, kinesin-2, dynein and myosinV can cooperate during reversals of pigment organelles (melanosome) transport in *Xenopus* melanocytes (Levi et al., 2006; Reilein et al., 2003). In cultured *Drosophila* S2 cells, kinesin and dynein worked together to transport the peroxisome (Kural et al., 2005). Kinesin and myosin share a common core structure and use similar conformational change strategy. The different motors are hence optimised for performing distinct biological functions (Vale and Milligan, 2000).

### 1.6.2.1 Myosins

Myosins are molecular motors that move along actin filaments and have several cellular roles in transport. Different myosin motors mediate transport of different organelles, including endosomes, mRNPs in neurons, vesicles and mRNA (Desnos et al., 2007). In the *Drosophila* embryo, non-muscle myosin II regulates asymmetric cell division by excluding determinants from the apical cortex (Barros et al., 2003).

A well-studied example of myosin dependent transport is the transport of *ASH1* mRNA along actin filaments in yeast, which is dependent on Myo4, a type-V myosin (also called She1) (Jansen et al., 1996; Long et al., 1997; Munchow et al., 1999; Takizawa et al., 1997; Takizawa and Vale, 2000). Labelled *ASH1* mRNA moves at speeds of 200–440 nm sec<sup>-1</sup> into the bud tip, which is consistent with the characterised dynamics of myosins. Myosin motors also transports *β-actin* mRNA (Sundell and Singer, 1991) and myosin II is responsible for the localisation of the ARP2/3 complex along actin filaments to the leading edge of motile fibroblasts protrusions (Mingle et al., 2005). A probable adaptor protein is the RNA-binding protein ZBP1. ZBP1 binds to two repeats in the *β-actin* localisation element, co-localises with myosin and is observed to move in this complex with the dynamics predicted for myosin motors (Kislauskis et al., 1994; Oleynikov and Singer, 1998; Ross et al., 1997).

### 1.6.2.2 Kinesins

Kinesins are a family of motor proteins that bind tightly to and move along microtubules. Kinesin has been implicated in the localisation of *Myelin Basic Protein (MBP)* mRNA. Fluorescently labelled *MBP* mRNAs injected into cells were the first transcripts to be visualised in living cells. The particles containing the transcript move from the oligodendrocyte cell body to myelinating processes (Ainger et al., 1993). Calcium/calmodulin-dependent kinase-2 $\alpha$  mRNA (*CaMKII $\alpha$* ) was found to also move in a kinesin-dependent manner to the myelinating processes of dendrites of hippocampal neurons. In dendrites, the dsRNA-binding protein Staufen, known to participate in RNA localisation events, co-fractionates in transport particles with *CaMKII $\alpha$*  mRNA as well as other localising transcripts, like the non-coding RNA BC1 (Mallardo et al., 2003).

In *Drosophila*, *oskar (osk)* mRNA localisation in the oocyte is guided by the kinesin-1 motor (Brendza et al., 2000; Clark et al., 1997). The mechanism is not clear as it is hard to analyse an injected *osk* as localisation requires the splicing of its first intron. Recent work using live tracking of labelled transcripts in the oocyte showed that the *osk* mRNA is actively transported along microtubules in all directions, with a slight posterior bias that is sufficient to localise the mRNA toward the posterior (Zimyanin et al., 2008).

Lipid droplets movement, mediated by kinesin-1, is one of the few known examples for a cargo that moves to the plus end in *Drosophila* embryos. It is assumed that cargos can simultaneously engage multiple kinesin molecules (Kural et al., 2005; Levi et al., 2006) and that the number of motors bound to the cargo increase its ability to localise. However, careful *in vivo* measurements of the lipid droplets movement in *Drosophila* embryo showed that, although few motors can bind to the cargo, its transport properties (travel distances and higher velocities) are largely unaffected by variation in motor number (Shubeita et al., 2008). Apparently, higher-order regulatory mechanisms rather than motor number dominate cargo transport *in vivo*. Indeed, co-regulation of the plus-end kinesin-1 motor and the minus-end dynein is essential during lipid droplet movement. Dynein-dependent minus-end motility is abolished in embryos lacking kinesin-1, and the average number of active dyneins per droplet is reduced by almost half in *Khc/+* embryos that lack one copy of the *Kinesin heavy chain* gene (Shubeita et al., 2008).

### 1.6.2.3 Dynein

Dyneins are exceptionally large multimeric protein complexes that are minus-end-directed microtubule motors, and participate in a wide range of cellular transport processes. There are two classes of dyneins; axonemal and cytoplasmic. Axonemal dyneins are responsible for the movements along microtubules that drive the beating of cilia and flagella. Cytoplasmic dyneins are responsible for transport of vesicles, proteins, and organelles along microtubules as well as the movement of chromosomes (Barton and Goldstein, 1996) and positioning the mitotic spindles during cell division (Gepner et al., 1996; McGrail and Hays, 1997).

Dynein consists of heavy chain homodimers with a “head” that holds the ATPase activity and two projections that extend from the head, and connect it to the microtubule via repeated cycles of coordinated association and dissociation. This mode of attachment ensures processivity of the movement, as the dynein molecule is constantly bound to the cytoplasmic microtubules. The intermediate and light chain subunits of the dynein attach it to its cargo (Bullock, 2007).

Cytoplasmic dyneins are also important for most minus-ended microtubule-related transports. For example, in *Drosophila*, movement of maternal transcripts from nurse cells into the oocyte, localisation of transcripts inside the oocyte (MacDougall et al., 2003), localisation of pair-rule genes in the embryo (Wilkie and Davis, 2001), and anchorage of localised transcripts at their destination (Delanoue and Davis, 2005), are all mediated by dynein motors.

### 1.6.3 Adaptor proteins

Very few proteins that bind specifically to localising RNAs have been identified so far (Dreyfuss et al., 2002; Glisovic et al., 2008). Two proteins in the *Xenopus* oocyte are known to directly bind localisation motifs. VgRBP60, a heterogenous nuclear ribonucleoprotein 1 (hnRNP1) homologue, recognises the VM1 binding motif (UUUCUA and related sequences) in *VgI* localisation element. A mutation in the VM1 motif abolishes both the protein binding and the localisation of the mRNA. Vg1RBP (Vera) binds the E2 motif (UUCAC and related sequences) in *VgI* (Bubunencko et al., 2002; Cote et al., 1999; Deshler et al., 1997; Gautreau et al., 1997; Kwon et al., 2002).

Staufen is a dsRNA binding protein that has five dsRNA-binding domains (dsRBDs) (Ramos et al., 2000). It interacts with several localised transcripts in *Drosophila* oocytes. The dsRBD3 domain of Staufen interacts with *bcd* mRNA and is required for its proper localisation, while the other dsRBDs are required for other

functions, such as translational regulation. Interestingly, the recognition by Staufen is based on the structure, rather than the sequence of the *bcd* mRNA. Staufen binds to the sugar-phosphate backbone of 12nt double helix, without interacting with the bases (Ramos et al., 1999; Ramos et al., 2000). Localisation of Staufen is not specific to a region, as it is required for anterior, as well as posterior localisation in the oocyte and in the embryo (St Johnston et al., 1991). For example, it is required for posterior localisation of *osk* (Micklemeier et al., 2000). Other interacting adaptor proteins probably determine the target site of localisation.

Staufen is also involved in other RNA localising events (reviewed in (Roegiers and Jan, 2000)). Stau localises to the vegetal hemisphere of the *Xenopus* oocyte and interacts with *vg1* (Allison et al., 2004; Yoon and Mowry, 2004). The mammalian Staufen homolog has several conserved dsRBDs and forms granules that co-localise with ribonuclear particles and are transported to the distal dendrite during neuronal maturation (Mallardo et al., 2003; Roegiers and Jan, 2000). In polarised intestinal epithelial cells, Stau is localised to the apical region that contain the RNA localisation machinery (Gautrey et al., 2005) and in cultured mammalian hippocampal neurons Stau is localised to the somato-dendritic domain, along with ribonucleoprotein particles that are known to contain mRNA (Kiebler et al., 1999). This implicates Stau in localisation of mRNA into the dendrites of mammalian neurons.

Staufen interaction with other proteins that are associated with the localisation machinery is also conserved. Barentsz (Btz), which is essential for the localisation of *oskar* mRNA to the posterior pole of the *Drosophila* oocyte, binds mammalian Staufen in an RNA-dependent manner. Ectopic mammalian Btz interact with *Drosophila* Staufen and reach the posterior pole in the wild-type oocyte (Macchi et al., 2003). This data suggest that the mRNA transport machinery is conserved during evolution.

*ASH1*-Myo4 is the only RNA-protein motor complex in which the adaptors are well characterised. She2, a novel RNA-binding protein recognises four localisation elements on the *ASH1* mRNA, and binds as a dimer (Bohl et al., 2000; Long et al., 2000). Upon binding, the affinity of She2 for the C-terminus of She3 is increased, and She3, in turn binds to the motor Myo4 through its N-terminus. Localisation depends on the association of the motor with the mRNA, as the tethering of the RNA to She3 bypasses the requirement for She2 (Kruse et al., 2002). This pathway is shared by other localising transcripts. A genome-wide analysis revealed that at least 23 transcripts are recognised by the same proteins (Gonsalvez et al., 2005; Jambhekar et al., 2005; Shepard et al., 2003).

Zipcode-binding protein (ZBP1), a nuclear-acquired factor, contains four KH domains and one RBD. It binds to *β-actin* mRNA at the site of transcription through a 54nt localisation signal on the 3'UTR and prevents its premature translation in the cytoplasm by blocking translation initiation (Ross et al., 1997). Another nuclear protein, ZBP2, facilitates the binding of ZBP1 to *β-actin* (Pan et al., 2007). The interaction is essential for proper *β-actin* mRNA localisation to the lamella region in several asymmetric cell types (Eom et al., 2003; Tiruchinapalli et al., 2003). The ZBP1-RNA complex then moves into the cytoplasm (Oleynikov and Singer, 2003). This example highlights how nuclear-acquired factors are important in cytoplasmic mRNA metabolism. The nascent RNAs are 'marked' in the nucleus for transport and localisation by the binding of specific proteins in the nucleus and during mRNA processing and export.

### 1.6.3.1 Bicaudal-D and Egalitarian

BicD and Egl are dynein-associated co-factors. Egl is only found in *Drosophila*, BicD appears in other organisms. The distribution of Egl and BicD and Dynein heavy chain (Dhc), is similar during *Drosophila* oogenesis and in blastoderm embryos, and they probably function together in specifying oocyte identity, as mutations in each of these components result in abolishment of the oocyte (Gepner et al., 1996; McGrail and Hays, 1997; Swan et al., 1999). *grk* RNA requires BicD and Dynein for its transport towards the ring canals, where it accumulates before moving into the oocyte (Clark et al., 2007; Delanoue et al., 2007). *bcd* and *osk* transcripts are also delivered to the oocyte by the same mechanism, which is distinct from cytoplasmic flow (Clark et al., 2007).

Egl or BicD do not include any known RNA-binding motifs, and it is not clear, whether either of them bind directly to the recruited RNA or whether there is another unknown adaptor that mediates the binding. The Egl/ BicD complex links specific RNAs to dynein and microtubules in the embryo (Bullock and Ish-Horowicz, 2001). The same machinery also operates to localise *inscuteable* transcripts in neuroblasts (Hughes et al., 2004; Li et al., 1997).

Components of the blastoderm localisation machinery also function in RNA transport from nurse cells into the early oocyte. Egl and BicD are required both for oocyte differentiation and for specific RNA accumulation in the oocyte (Ran et al., 1994; Schupbach and Wieschaus, 1991). During oogenesis, these two proteins form a complex together and co-localise at the minus ends of microtubules (Oh and Steward, 2001).

Both proteins are thought to act as adaptors and can interact with the dynein motor (Hoogenraad et al., 2001; Matanis et al., 2002; Navarro et al., 2004) and with each other (Mach and Lehmann, 1997). The interaction of Egl with dynein light chain (Dlc) is through a domain distinct from that binding BicD (Navarro et al., 2004), which might allow the simultaneous binding of both molecules to the Dynein motor. Point mutations specifically disrupting Egl-Dlc association also disrupt microtubule-dependant trafficking both to and within the oocyte, resulting in a loss of oocyte fate, maintenance and polarity (Navarro et al., 2004). Localising transcripts are anchored in a manner that requires dynein, Egl or BicD (Delanoue and Davis, 2005).

#### 1.6.4 Transport particles

The transport protein complex and the mRNA transcripts form large particles, called 'mRNA granules' which are often visible using microscopy and have been observed in several cell types (Bassell et al., 1999). Formation of these structures is associated with the transport itself, indicating it has an essential role in the process. Additional RNA sequences may be required to promote the assembly of a localisation-competent ribonucleoprotein (RNP) (Czapinski and Mattaj, 2006). In yeast, directionally moving particles were revealed using the MS2-GFP system to track localisation of *ASH1* mRNA. These particles include the mRNA and the different proteins that are involved in the transport (Beach et al., 1999; Bertrand et al., 1998).

In oligodendrocytes, mRNA granules are formed irrespective of the mRNA species, but only particles with mRNAs containing a proper localisation signal, for example *MBP* mRNA, are transported into the processes of the oligodendrocyte (Ainger et al., 1993). These granules, formed early in the perikaryon, are very big. Their diameter is around 0.7  $\mu\text{m}$ , and they contain not only the *MBP* mRNA and mRNA-binding proteins, but also components of the translation machinery and even ribosomes (Barbarese et al., 1995).

Finally, in *Drosophila*, during early and late stages of the oocyte development, particles containing mRNA and proteins move from nurse cells into the oocyte, and within the oocyte. These particles have been visualised *in vivo* using a GFP tag fused to the Exuperantia (Exu) protein. Exu is known to associate with *bcd* localisation (Wang and Hazelrigg, 1994) and the *bcd* transcript itself (Mische et al., 2007) during localisation. Despite the fact that *bcd* and *osk* mRNAs localise to opposite sides of the oocyte, a biochemical assay to isolate a complex containing the Exu protein, showed that a large RNP complex contains these two mRNAs, together with several more

proteins. That suggests a more general role for Exu in localising mRNAs both within nurse cells and the developing oocyte (Wilhelm et al., 2000).

Injected and endogenous localising transcripts in the syncytial blastoderm embryo form particles that move directionally in the cell (Wilkie and Davis, 2001). Injected particles are bigger than endogenous ones, and probably contain many more transcripts and associated proteins than the endogenous particles. Furthermore, these particles can behave differently when they include a mixture of wild type and mutated transcripts, as they are slower than the particles that contain wild-type transcripts only. Since injected wt transcripts are localised much faster than mutant transcripts, even when injected in a mixture, this suggests that a significant proportion of wild-type RNA becomes enriched apically by efficient transport in small cargoes that are not detected by tracking them in a spinning disc microscopy (Bullock and Ish-Horowicz, 2001). Particles do not necessary represent an active transport complex. Injected transcripts that do not have a localisation signal would assemble into particles but the particles but would not move directionally to the apical site (Bullock et al., 2006).

The efficiency of the transport of the particles has been shown to be dependent on the amount of transport proteins inside. The injection of the wild-type *h* localisation signal was compared to the injection of the more efficiently localising construct *h<sup>SL1x3</sup>*, based on the amount of Egl and BicD assembled on individual mRNA particles in transit. The concentration of Egl and BicD in many particles containing *h<sup>SL1x3</sup>* mRNA, but not in those containing *h* mRNAs, were above cytoplasmic level, suggesting that the level of Egl and BicD influences the attributes of localisation (Bullock et al., 2006).

## **1.7 Localisation Signals**

Localisation of transcripts usually depends on recognition of *cis*-acting elements within the RNA. These elements, also named “zip codes” or “localisation signals” (LS), usually reside in the 3’UTR of the transcript and are responsible for the interaction of the RNA with the transport machinery. The 3’ UTR usually controls RNA localisation and initiates distinct temporal patterns of translation of the localised RNAs (Rangan et al., 2009). A few known signals are located instead within the coding region, for example, the RNA encoding *Drosophila* synaptotagmin-like protein, Bite Size (Serano and Rubin, 2003) and the yeast *ASH1* localisation signal (Chartrand et al., 1999). Other signals appear in the 5’UTR. For example the signal on the *Drosophila yemanuclein-alpha* transcript (Capri et al., 1997), the *grk* localisation signal, GLS (Van De Bor et al., 2005) and the *Xenopus Xnif* signal (Claussen et al., 2004).

### 1.7.1 Sequence and structural basis of recognition

The sequence and structural basis of the RNA-protein recognition is unclear. Despite the fact that many transcripts are known to localise and that localisation is seen in such variety of different cells, few localisation signals are currently known. Most characterised signals do not share sequence homology, nor do they share an obvious unusual structural similarity, although there is evidence that secondary structures are important in the recognition of the transport complex. Such is the case for signals in *K10* (Serano and Cohen, 1995b) *c-myc* (Chabanon et al., 2005), *bcd* (MacDonald, 1990), *grk* and *I factor* (Van De Bor et al., 2005) and many more. Tertiary structure of the RNA can be important for recognition and enhanced binding of RNAs to proteins (Fernandez-Miragall et al., 2006; Hermann and Patel, 1999).

In some cases, compensatory mutations are not enough to restore activity. *K10* Transport/Localisation Signal (TLS) distal stem is composed of five U-A base pairs. Changing the U-A stretch to A-A abolishes localisation, as the stem is disrupted. Changing the U-A into A-U, restores the stem structure (but not the primary sequence), but does not retain full activity (Bullock and Ish-Horowicz, 2001; Serano and Cohen, 1995b). Hence, there may be additional elements that are important in the recognition of the stem. These elements can be specific sequence recognition of one or more nucleotides, or a more complex interaction between the stem and other parts of the signal. There could also be another structural element, which is not revealed by the folding prediction.

There is other evidence for the significance of tertiary structures. For example, localisation of *bicoid* (*bcd*) depends on intermolecular loop-loop interaction (Ferrandon et al., 1997). The *cis*-element responsible for dimerisation is part of stem-loop domain (domain III) that contain two essential complementary 6-nucleotide sequences in a hairpin loop (LIIIb) and an interior loop (LIIIa) (Wagner et al., 2004). Localisation of *h* transcripts also depends on oligomerisation of two stem loops, SL1 and SL2. Multimerisation enhances the activity of SL1 but not of SL2a (Bullock et al., 2003).

### 1.7.2 Identifying localisation signals

Finding the elements in the RNA that are responsible for localisation has not been easy for several reasons (reviewed in (Jambhekar and Derisi, 2007)). First, only few proteins are known to bind directly to localising transcripts. Many attempts at identifying specific RNAs that will bind a protein from the motor complex have resulted in unspecific RNA binding. One successful screen was conducted in yeast, where the

protein directly interacting with the RNA interacting proteins was known. The She2p family binds to localised transcripts, such as *ASH1* (Bohl et al., 2000).

Using libraries of partially randomised *ASH1* localisation elements, a screen was conducted to find more transcripts that bind to She2p. The RNAs isolated lacked primary sequence similarity but contained a similar loop-stem-loop structure with a highly conserved CGA triplet in one loop and a single conserved cytosine in the other loop. Mutating these conserved nucleotides or the stem separating them resulted in the loss of She2p binding and in the delocalisation of a reporter mRNA. These findings defined a motif that was further used to search for novel localisation signals in 23 other known bud-localised RNAs. Four were found, but of these, only two localised *in vivo* (Olivier et al., 2005). This method of identifying novel signals was not very efficient, as it did not find known and validated signals. Analysis of additional She-complex-dependent localisation signals revealed that variations in the sequence and structure could be tolerated in some contexts (Jambhekar et al., 2005). Nevertheless, it is a good example of how knowing the binding protein can greatly contribute to finding new localisation signals, and their common motifs.

Another reason for the infrequent discoveries of new LSs, is that many RNAs are localised by a multistep process, with each step being governed by distinct complexes that are potentially coupled to the preceding and/or subsequent step. Therefore, multiple elements are required together to affect localisation, precluding the identification of minimal sub elements.

However, the main reason for the difficulty in identifying new LSs, is that many of them are recognised by the transport machinery on the basis of their structure, rather than sequence. It is very difficult to search and compare new structural elements, and there are very few motifs that are known to appear in localisation signals. In addition, solving an RNA structure is technically challenging; few RNA structures have been solved for comparative studies and predicted secondary structures do not reflect tertiary structure (Leontis et al., 2002).

### **1.7.3 Finding localisation signals using sequence similarity**

There are very few examples of proteins recognising an RNA sequence as a localisation signal, making the sequence comparison strategy generally impractical. Sequence-based recognition allows comparison and searching for similar sequences in the genome with sequence only search methods such as the BLAST (Altschul et al., 1990). One example of a sequence-dependent motif is in the dendritic localisation of the

alpha subunit of Ca<sup>2+</sup>/calmodulin-dependent protein kinase II (alphaCaMKII) mRNA in CNS neurons requiring its 3' untranslated region (3'UTR). One of the two elements, a 30-nucleotide signal mediates dendritic translocation. A homologous sequence found in the 3'UTR of neurogranin was shown to be important as well for its localisation to dendrites (Mori et al., 2000).

Another sequence-based element is the CAC motif. Clusters of short CAC-containing motifs characterise the localisation elements in many mRNAs localised to the vegetal cortex of *Xenopus* oocytes. This sequence was identified by a computer program, called REPFIND, which recognises repeated motifs in localised RNAs. A search for this signal in GenBank resulted in the identification of new localised mRNAs. CAC-rich elements are also found in ascidians and other vertebrates, indicating that these *cis*-regulatory elements are conserved in chordates. Interestingly, biochemical evidence shows that distinct CAC-containing motifs have different functions in the localisation process. Thus, clusters of CAC-containing motifs are a ubiquitous signal for RNA localisation and might signal localisation in a variety of pathways through slight variations in sequence composition (Betley et al., 2002).

The *XNIF*, an RNA which is localised to the vegetal pole of *Xenopus* oocytes by early and late pathways, contains 16 copies of CAC motif in its localisation signal. A critical number of such repeats seem to be required for accumulation in the mitochondrial cloud along the early pathway, but additional repeats seem to be required for localisation along the late pathway (Claussen et al., 2004). CAC repeats also appear in the localisation signal for rat metallothionein-1 mRNA (Nury et al., 2005), although it is not the only factor important for localisation of the RNA, as the secondary structure of the signal might also play a role.

#### **1.7.4 Finding and comparing LSs based on similar secondary structures**

In most cases, localisation signals are defined by their higher-order secondary and tertiary structures. For example, short stem-loop structures direct the apical localisation of *hairy* mRNA in the fly embryo, and the transport of *K10* and *orb* RNA from nurse cell to oocyte and inside the oocyte. Mutational analysis indicates that both the sequence and the structure of the double-stranded stems are important (Bullock et al., 2003; Serano and Cohen, 1995b).

However, the presence of a double stranded structure does not define a localisation signal. Essentially every sequence of RNA can fold into a secondary

structure and yet, only few localise. A bioinformatics search, aimed to identify non-coding RNAs based on the stability of their secondary structure, concluded that the predicted stability of most non-coding RNA secondary structures is important. However, it is not sufficiently different from the stability of a random sequence and so cannot be useful as a general gene finding approach (Rivas and Eddy, 2000) indicating that specific secondary and/or tertiary structure elements must be important for the function of these RNAs. In such cases, searching for a common motif, or for new LSs cannot be based on simple sequence homology search tools, and must use structural prediction tools.

One of the problems with RNA structure predictions tools is that they do not always reflect the nature of RNA folding *in vivo* and should only be regarded as approximations of the secondary structure. Most predictions are based on Watson–Crick base pairing and the G:U wobble pair that is specific to RNA, but RNA molecules exhibit complex structures in which some of the bases engage in non-Watson-Crick base pairing, base triples, G-quartets and pseudoknots (Leontis et al., 2002; Westhof and Fritsch, 2000). In addition, these methods do not include 3D information such as details of base stacking, backbone hydrogen bonding and tertiary interactions.

Nevertheless, in many cases the computer prediction is close to the folded conformation, as studied with biochemical methods (Brunel and Ehresmann, 2004; Wagner et al., 2004). Currently, the most successful approach to predicting RNA motifs begins by predicting the secondary structure by means of either energy minimization using nearest neighbour thermodynamic rules, as implemented in mFOLD and the Vienna RNA package (Hofacker, 2003), or covariation analysis, or a combination of both (Hofacker, 2003; Hofacker et al., 2002; Zuker, 2003). In this work, I used the most commonly applied Minimum Free Energy (MFE) method, mFOLD, which searches for the lowest free energy fold in the molecule by maximizing the number of base pairs within a structure.

### **1.7.5 Using related species to identify new motifs**

A functional RNA element that depends on its secondary structure, rather than its primary sequence should retain an evolutionarily conserved RNA secondary structure in related species. The growing number of genome sequences available makes it an easier, but still not a straightforward assignment. Good sequence alignments are required for accurate structure prediction, which makes it difficult to identify short elements buried in long UTRs or coding regions.

Algorithms such as FOLDALIGN (Gorodkin et al., 1997) and CONSTRUCT (Luck et al., 1999) have been successfully used to identify functional elements in HuR target mRNAs and non-coding telomerase RNAs (Dandjinou et al., 2004; Romero and Blackburn, 1991), but were not successful in identifying zipcode regions. Manual comparison of *bcd* 3' UTRs from six different species of *Drosophila* yielded globally similar structures (MacDonald, 1990). The details of stem and loop lengths and sequences, however, varied between species, making it difficult to pinpoint putative transport complex binding sites (MacDonald, 1990).

Comparison of the *nos* 3' UTRs from *D. melanogaster* and *D. virilis* revealed several conserved sequence elements both within and outside the four regions to which localisation signals had been mapped, as well as a translational control element (TCE) that showed both sequence and structural conservation (Crucs et al., 2000; Gavis et al., 1996a; Gavis et al., 1996b). Mapping of localisation signals in both *bicoid* and *nos* was initially achieved by testing various 3' UTR fragments for localisation, presumably because the sequence and structural similarities between the available sequences of homologous genes were not sufficient to identify candidate zipcode regions (Gavis et al., 1996a; Gavis et al., 1996b; MacDonald, 1990). The available tools for RNA structure prediction are not sufficiently accurate to reliably identify elements within large data sets of sequences and a more accurate description of structures is needed in order to recognise more putative signal sequences.

### **1.7.6 Examples of localisation signals**

Here are some examples of localisation signals and what is known about their sequence and structure determinants.

#### **1.7.6.1 hairy**

The *h* pair-rule transcript localising signal consists of a 121-nt region in the 3'UTR that is necessary and sufficient to mediate apical transport in the blastoderm embryo. Using a combination of extensive mutational analysis and evolutionary sequence comparison, it was shown that the *h* Localisation Element (HLE) comprises of two partially redundant stem loops (Figure 7\_1).

Both primary sequence and secondary structure are important for localisation. Base-pair identities within the stems are not essential, but can contribute to the efficiency of localisation, suggesting that specificity is mediated by higher-order structure. No protein is known to bind the HLE but the distance between the two stem

loops and their cooperative effect suggests that more than one factor is likely to bind the localisation element. The HLE can recruit Egl and BicD, but it is not clear whether Egl or BicD bind directly to the HLE. Mutations in the HLE affect the kinetics of transport and suggest that the cargo can regulate the transport efficiency of the molecular motor (Bullock et al., 2006; Bullock et al., 2003).

### 1.7.6.2 bicoid

One of the best characterised RNA localisation signals is found within the 3'UTR of the *bcd* transcript. The *bcd* localisation element (BLE) is modular and composed of five different domains (Figure 7\_1). Bcd has a complex pattern of localisation involving several different steps, each of which is directed by a distinct element. Each domain folds into a stem loop and is important for different functions in the process of localisation.

Domain IV/V is both necessary and sufficient for the program of early and late ovarian localisation. BLE1 signal in domain V is sufficient to drive *bcd* early localisation (Figure 7\_1B). Domain III is necessary for the mRNA anchoring and for dimerisation. The terminal and side loops of domain III are complementary and the dimerisation of the signal is initiated via base pairing of these single stranded loops, and is then stabilised by surrounding sequences (Wagner et al., 2004). Domain I contains sites for trans-acting factors exhibiting single stranded RNA binding specificity.

Using a variety of chemical and enzymatic structural probes, a detailed description of the 3' UTR of the *bcd* mRNA and its organisation into the different domains was obtained (Brunel and Ehresmann, 2004). One prominent and unexpected result that emerged was the high degree of flexibility of the different domains relative to each other. This plasticity relies upon the open conformation of the central hinge region interconnecting domains II, III, and IV/V. Domain I is mainly unstructured, but each core domain (II-V) is highly organised and folds into helices interrupted by bulges and interior loops and closed by very exposed apical loops. These elements mostly built specific determinants for trans-acting factors (Brunel and Ehresmann, 2004; Bullock et al., 2003; Driever and Nusslein-Volhard, 1988b; Luk et al., 1994; MacDonald, 1990; Macdonald and Struhl, 1988).

Manual comparison of 39 predicted secondary structures of *bcd* UTRs from six different species of *Drosophila* yielded globally similar structures (MacDonald, 1990) and conserved activity. For example, the *bcd* mRNA localisation signal originated in *Drosophila pseudoobscura* retains its function in *Drosophila melanogaster*. Thus,

among these *Drosophila* species, there is substantial conservation of components acting in mRNA localisation, and presumably the mechanisms underlying this process (Luk et al., 1994). The *bcd* signal is recognised through the dsRBD3 domain of one or more intact molecules of Stauf protein.

### 1.7.6.3 wingless

Wingless (*wg*) is the founding member of the highly conserved Wnt gene family (reviewed in (Wodarz and Nusse, 1998)). Wnt genes encode secreted glycoproteins that serve as signalling molecules essential during many embryonic patterning processes, including patterning of the *Drosophila* embryo. *wg* is expressed just prior to gastrulation, in one of the epithelial cells in each embryonic segment. The Wg protein is secreted at the apical epithelial cell cortex and signals to nearby cells. Wg signalling is required for correct patterning of the embryo along the A-P axis and *wg* loss-of-function mutations lead to the overproduction of larval denticle. Apical localisation of *wg* transcripts in epithelial cells has been shown to augment Wg signalling in the embryo, although it is not clear how this is achieved (Simmonds et al., 2001)

*wg* RNA is localised apically within cells in the embryonic ectoderm (Simmonds et al., 2001). Previously characterised elements in the wingless 3'UTR, WLE1 and WLE2, were shown to be unnecessary for the apical localisation in an embryo microinjection assay. Another element, referred to as WLE3, is both necessary and sufficient for apical RNA transport, although full, unrestricted activity requires the presence of one of several downstream potentiating elements. WLE3 forms a highly conserved stem-loop structure. Despite these high levels of sequence and predicted structure conservation, however, mutagenesis shows that both sequence and structure varied extensively in the predicted stem-loop, except for few important features. These features include an accessible distal helix sequence motif, which is also found in the predicted structures of other apical localisation elements and might define a consensus localisation motif for apically localised transcripts (dos Santos et al., 2008).

### 1.7.6.4 gurken and I factor

*Drosophila grk* mRNA is localised in a two-step localisation pattern, governed by two zipcodes in the oocytes. In the first step, at stages 1–6, *grk* localises to the posterior of oocytes. This step depends on sequences within the first 35 nucleotides of the ORF termed GLE1 (Thio et al., 2000). In the second step, during stage 8, *grk* re-localises to the anterodorsal corner of the oocyte, and this localisation is mediated by 64

nt downstream of GLE1 (Saunders and Cohen, 1999). The 64 nt element was later shown to be necessary and sufficient for localisation of injected *grk* mRNA in the oocyte and was termed *grk* localisation signal, GLS.

The GLS forms a small consensus RNA stem loop of defined secondary structure, that is conserved among different *Drosophila* orthologs, *D. simulans*, *D. affinis*, *D. erecta*, *D. miranda*, and *D. virilis*, which are thought to have diverged from *D. melanogaster* between 5 to 60 million years ago. Surprisingly, this element also shares a conserved structure (but not sequence) with the *I factor*, a non-LTR retrotransposon. *gurken* and the *I factor* compete *in vivo* for the same localisation machinery, as *I factor* transposition leads to its mRNA accumulating near and within the oocyte nucleus, and disrupts localisation of *grk* and *bcd* mRNA and axis specification. This is an example of a case where knowledge of the structure led to a model for the mechanism of the transposon in the germline, namely that selective transposition is achieved through intracellular mRNA transport in the oocyte, using the host's machinery followed by import into the oocyte nucleus (Van De Bor et al., 2005). Indeed, other retroviruses require MTs and dynein for their transport to the nucleus (Whittaker et al., 2000), suggesting that the transport of their RNA is important in the life cycle of some transposable elements and viruses.

#### **1.7.6.5 c-myc**

*c-myc* mRNA is localised to the perinuclear cytoplasm of fibroblasts and associates with the cytoskeleton in mammalian cells. The localisation signal of *c-myc* was identified in a region between nucleotides 194–280 of its 3'UTR and subsequently narrowed down to a sequence of 45 nucleotides within this area that is sufficient for its localisation (Chabanon et al., 2005). The 45-nucleotide signal is also the binding site for annexin A2 protein. The predicted structure of the signal is a stem loop that includes two internal loops and is conserved between different species. This structure is very similar to the localisation signal of another proto-oncogene, *c-fos*. The three putative helices in the structure contain a stretch of U-A base pairs, of which at least at the distal helix is essential for its localisation. The U-A element is similar to the distal helix of the TLS and the *orb* LS, and might suggest a more general element that is important for other localisation signals

## 1.8 The *K10* Transport Localisation signal (TLS)

### 1.8.1 *K10* function in the cell

*K10* is a putative DNA binding protein that is involved in different biological processes including: dorsal/ventral pattern formation; intracellular mRNA localisation; protein localisation; ovarian follicle cell development; pole plasm mRNA localisation; negative regulation of translation; and oocyte dorsal/ventral axis determination. *K10* expression is confined to the primary oocyte and the protein product is sequestered in the oocyte nucleus (Prost et al., 1988). *K10* is required in the oocyte nucleus to restrict the expression of *grk* to the dorsal part of the anterior oocyte, most likely by interaction with Squid and Bruno proteins (Kelley, 1993; Norvell et al., 1999). Mutation in *K10* causes mis-localisation of *grk* RNA which directs Grk signalling to the follicle cells on all sides of the oocyte (Neuman-Silberberg and Schupbach, 1993). Ventralising mutants *grk* and *torpedo/Egfr* (*top/Egfr*) are both epistatic to the dorsalisating effects of *K10* mutations (Schupbach, 1987). Mutations in *K10* disrupt *grk* mRNA localisation and lead to the production of dorsalised eggs and embryos (Neuman-Silberberg and Schupbach, 1993). Capu and Spir are required for *K10* mRNA localisation from stage 8 of the oocyte (Cheung et al., 1992; Manseau and Schupbach, 1989) but they probably act in parallel with *K10* to regulate *grk* mRNA localisation (Serano and Cohen, 1995a).

*K10* mRNA is synthesised in ovarian nurse cells during the early and middle stages (stages 1–7) of oogenesis and rapidly transported into the oocyte. During stage 8, the transcripts are directed to the oocyte's anterior cortex, where they may persist for many hours or days. *K10* transcripts then localise to the anterior side of the oocyte until fertilisation (Cheung et al., 1992).

Mutant females of maternally acting genes responsible for establishing DV axis produce normally shaped eggs that develop into dorsalised embryos. The *K10* mutation differs from these mutants, because in addition to the dorsalised development of the embryo, it causes a dorsalisation of the egg shape. Embryos produced by homozygous females are completely dorsalised and less than 1% of the eggs are fertilised. The egg chamber in homozygous females first appears abnormal at stage 10, when all the follicle cells surrounding the anterior end of the oocyte begin to thicken and migrate centripetally. The egg chamber is also shorter than normal.

*K10* transcripts also localise apically in neuroblasts when mis-expressed. Ectopic *K10* transcripts can localise apically in the embryonic epithelium, in an

Egl/BicD/dynein-dependent pathway, as localisation of the transcripts is disrupted in *egl* mutant embryos (Hughes et al., 2004).

### **1.8.2 TLS is one of the shortest signals found in *Drosophila***

The *K10* transport localisation signal (TLS) is one of the shortest localisation signals known in *Drosophila* (Figure 1\_5). It comprises of 44 nucleotides, which are sufficient and essential for localisation to and within the oocyte (Cohen et al., 2005; Serano and Cohen, 1995b) and also for recruitment of the Egl and BicD proteins (Bullock and Ish-Horowicz, 2001). *K10* TLS can also mediate the apical mRNA localisation ectopically, when expressed in somatic follicle cells (Karlin-Mcginness et al., 1996) or syncytial blastoderm embryos (Bullock and Ish-Horowicz, 2001), or neuroblasts (Hughes et al., 2004)

The TLS has been extensively characterised by Robert Cohen's laboratory (Cohen et al., 2005; Serano and Cohen, 1995b). They used assays based on whole-mount *in situ* hybridization of ovaries of transgenic flies with mutated *K10* signals to characterise the transport of the transcripts into the oocyte and subsequent anterior localisation. The Cohen lab identified the 44 nt signal that is necessary for localisation, and which can direct the transport of a heterologous mRNA. *K10* TLS is probably recognised by virtue of its structure rather than its sequence and the double stranded stem is crucial for localisation.

However, the single stranded nucleotides appear to be not important as, in their oocyte assay, removing both bulges or changing the sequence of the nucleotides in the loop did not alter localisation activity (Serano and Cohen, 1995a). In further studies they distinguish between mutations that might affect the stereochemistries of the minor and major groove (base pair substitutions, e.g., A:U for G:C) and mutations that only affect the major groove (e.g., A:U for U:A) and show that the later does not change localisation. Hence the interaction must be through the minor groove of the double helix, which is in agreement with the fact that generally, major grooves in RNA helices are very deep, narrow, and not accessible to amino acid side chains (Cohen et al., 2005).

### **1.8.3 *orb* contains a localisation signal that has similar predicted structure to that of *K10***

The Orb protein is required for patterning of antero-posterior and dorso-ventral axis during *Drosophila* oogenesis (Christerson and McKearin, 1994). It functions in the transport of mRNAs coding for two essential proteins that govern polarity in the egg:

Osk (anterior/posterior) and Grk (dorsal/ventral) during the middle phase of oogenesis, stages 8-10 (Lantz et al., 1994). In early oogenesis (stages 2-6) *orb* mRNA is transferred, along with many other maternal transcripts, including *K10*, into the oocyte (Lantz et al., 1992), and it later localises to the anterior cortex of the oocyte (See section 1.3). Orb protein is also enriched in the anterior dorsal region and at the posterior pole of the oocyte. The signal that is responsible for *orb* localisation does not share primary sequence homology with the TLS, but has a predicted stem loop secondary structure that contains some similar features (Lantz and Schedl, 1994; Serano and Cohen, 1995b). Hence, in this work, the *orb* LS was used as a comparison signal to the TLS.

### **1.9 Understanding the structure of RNA is essential for the understanding of its function**

The secondary structure of RNA, and features including double helices, hairpins, internal loops, bulges, or junction loops can be inferred theoretically from nucleotide sequence inspection using a combination of statistics, thermodynamic and comparative sequence analyses. However, for several reasons the accurate prediction of RNA structure is much more difficult than prediction of protein or DNA structure. First, only few RNA structures have been solved to date, hampering comparison or prediction of different motifs based on known ones. Second, RNA molecules fold in much less stringent conditions than DNA, which can lead to multiple conformations of the same molecule. In addition to the classic Watson-Crick nucleotide base-pairs, and RNA specific G-U wobble base, nucleotides in RNA can form two and one-hydrogen bond base pairs, water-mediated base pairs, protonated base pairs and base-phosphate hydrogen bonding interactions (Leontis et al., 2002; Westhof and Fritsch, 2000). Most of these interactions depend upon the local environment (Turner, 1996; Uhlenbeck et al., 1997). Non-Watson-Crick base pairs incorporated into the double helical regions of RNA affect the groove widths and can potentially provide a varied pattern of proton donors and acceptors in the major and minor grooves for recognition by proteins.

Structure prediction computer programs like mFOLD base their predictions on the minimal free energy (MFE) that is required for folding, and only consider canonical base pairing (and wobble base pairing), thus their ability to accurately predict RNA structure is limited. Various experimental strategies can be used to analyse the structure of RNAs while either free or complexed with specific ligands including X-ray crystallography, cryo-electron microscopy, nuclear magnetic resonance spectroscopy,

structure-specific probes, RNA engineering, thermal denaturation and mass spectrometry (Felden, 2007).

## **1.10 Summary**

Asymmetric localisation of cytoplasmic mRNA is important in a vast array of organisms and cell processes. In recent years, our knowledge about RNA localisation has expanded, as many more RNA molecules have been found to localise in a cell. In addition, new information accumulates about the *cis* and *trans* acting elements that drive the transport. However, one of the unsolved questions in the field remains the nature of recognition of cargo and protein complex. With few exceptions, (Bohl et al., 2000), the specific binding protein has not been identified.

Furthermore, although many known localisation elements use the same protein machinery for localisation, there is so far no known minimal element that characterises localisation signals. Hence, it is still unclear how localised elements are different from non-localised sequences and how they are recognised by the localisation machinery.

In this thesis, I examine the structural features of the *K10* localisation signal and show that the structure of the signal is crucial for its activity. I demonstrate the important structural features, including the two bulges and the helix, which adopt an unusual form. These findings suggest a general localisation element that specifies RNA localisation signals in the Dynein-dependent mechanism in *Drosophila*.

## **1.11 Aims of the thesis**

The major aim of my thesis is to understand the mechanism and specificity of the machinery that regulates the dynein-mediated RNA transport in the *Drosophila* embryo. Specifically, I wanted to understand how RNA cargos are recognised by the protein complex. Recognition is most likely based on a combination of the sequence and structure of the localisation signal. Hence, I chose to study the structural elements that contribute to the cargo recognition.

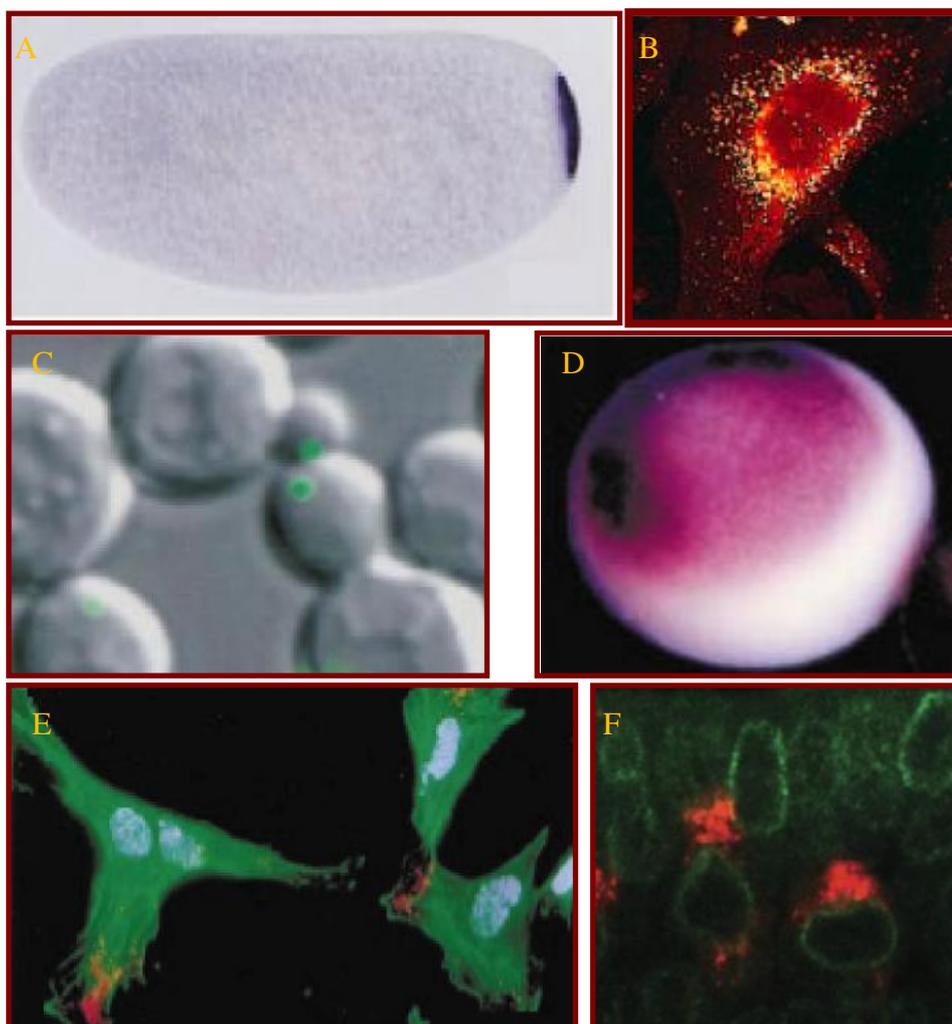
I shall concentrate on activity of the *fs(1)K10* TLS that is crucial for localisation of the *K10* transcripts in the oocyte and of injected transcripts in blastoderm embryos. The TLS would also be compared to other similar localisation signals in *D. melanogaster* and other species. Since the TLS consist of a short sequence, it might represent general minimal element/s of signals that drive localisation.

Specific aims include:

- To identify the sequence and structural elements that contribute to localisation of

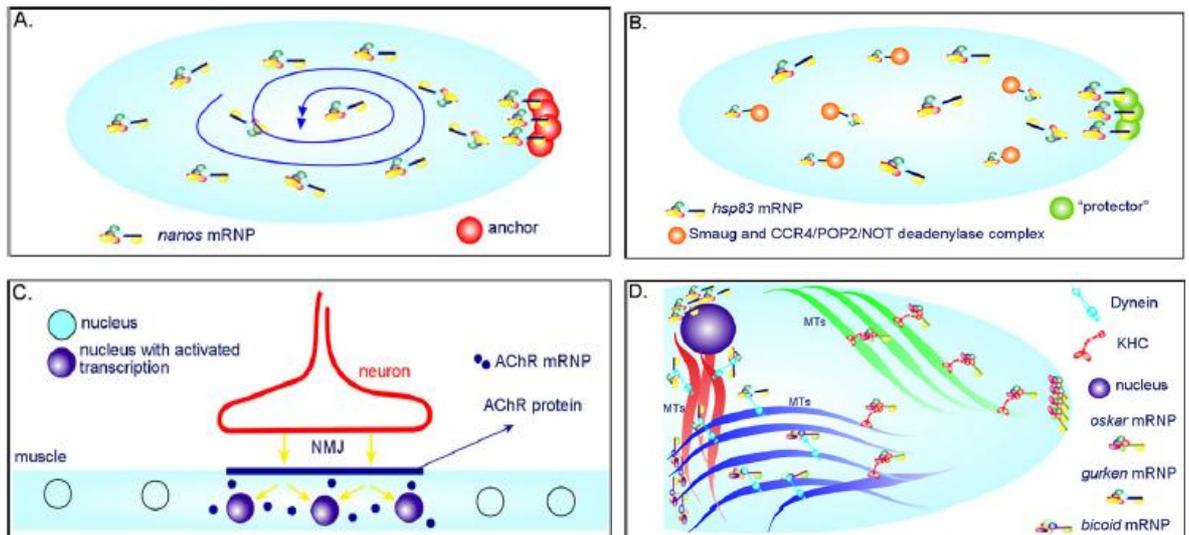
TLS by using a combination of genetic, comparative analysis, and structural methods.

- To establish microscopic techniques for visualising injected transcript transport *in vivo*. This would be used to study the kinetics of particles containing localised transcripts, to understand the effect of different mutations on the movement of the RNA, and compare the relative differences in dynamic of weak and strong mutants.
- To establish lines of flies that contain mutations in the TLS in order to study the effect of the TLS on the development of fly embryo.



**Figure 1\_1** Examples for localised RNA in different cells.

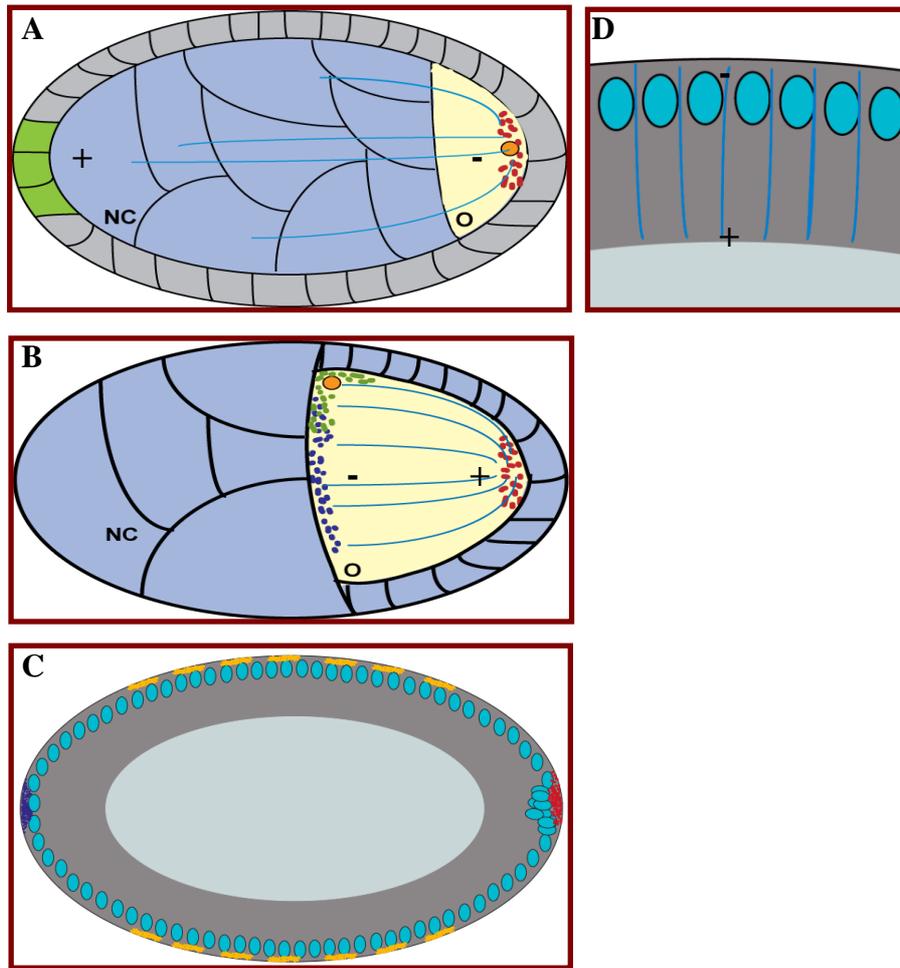
(A) *nanos* encodes a protein that determines posterior structures, and is localised to the posterior of a *Drosophila* blastoderm embryo (Tadros and Lipshitz, 2005). (B) *c-myc* RNA exhibits a perinuclear localisation and association with the cytoskeleton that contributes to the nuclear import of the protein (Veyrune et al., 1996). (C) *ASH1* mRNA is localised to the distal tip of the daughter cell in *Saccharomyces cerevisiae* to determine asymmetric cell division (Long et al., 1997). (D) *Vg1* is localised to the vegetal pole of the *Xenopus* oocyte (Kloc and Etkin, 1994). (E) Vertebrate  $\beta$ -*actin* mRNA is localised near the leading edge in fibroblasts, myoblasts, and epithelial cells where actin polymerisation actively promotes forward protrusion (Kislauskis et al., 1994). (F) *Inscutable* RNA localisation apically to the neuroblast is required for efficient apical targeting of Insc protein, which is required for accurate control of metaphase spindle length, division orientation, and asymmetric cell division (Hughes et al., 2004).



**Figure 1\_2 Different mechanisms for mRNA localisation**

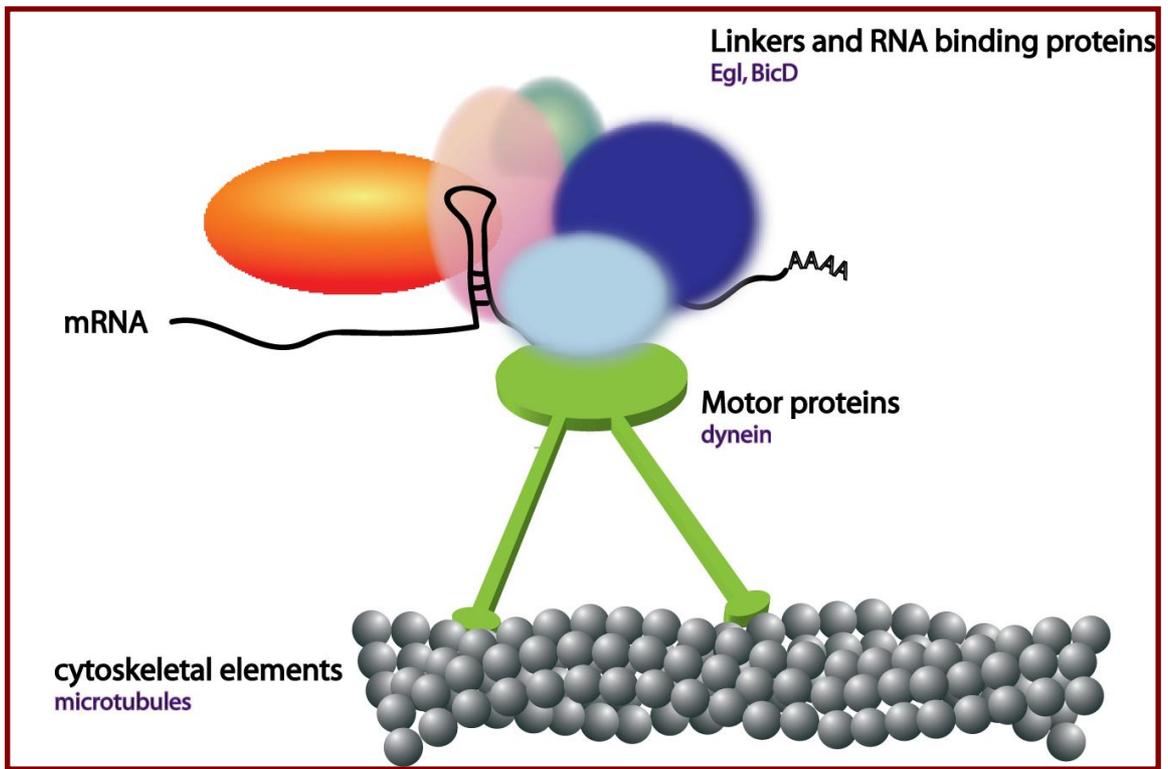
(A) Random cytoplasmic diffusion followed by entrapment: *nanos* mRNA diffusion through the cytoplasm is facilitated by cytoplasmic flows (blue spiral). A specific anchor (red) in the site of localisation (anterior end) is trapping the molecules. (B) Generalised degradation and local protection at the site of localisation: Smaug binds the *hsp83* mRNA and recruit the CCR4/POP2/NOT deadenylase complex for degradation of the transcripts. Transcripts on the site of localisation are protected. (C) Local synthesis: The Acetyl Choline Receptor (AChR) is transcribed at the neuromuscular junction (NMJ), regulated by locally acting factors from the nerve and the muscle (yellow arrows). It is then diffused locally and accumulates at the synapses. (D) Directional transport by motor proteins along cytoskeletal elements: *bcd* and *grk* mRNAs are actively transported by dynein to the anterior pole of the oocyte along an anteriorly nucleated population of MTs (blue). Once *grk* RNA arrives at the anterior end, it shifts to another population of MTs (red) that mediates its dynein-dependent localisation to the anterior-dorsal corner. At the same stage, *oskar* RNA is transported towards the plus ends of the MT at the posterior pole with the kinesin motor protein.

(adapted from (Lipshitz and Smibert, 2000) and (Palacios, 2007)).



**Figure 1\_3 Stages of development of *Drosophila* and localised RNAs in each stage.**

(A) Localisation of maternal transcripts in early oocyte (stage 1-7). Transcripts (red) are transcribed in the nurse cells and are rapidly transported to the pro-oocyte along the MTs (blue stripes). The MTs are polarised and nucleated at their minus end at the posterior of the egg chamber. (B) Localisation of maternal transcripts in the oocyte (stage 7-9). *Drosophila* egg chamber showing the localisation pattern of different mRNAs in the oocyte. *bicoid* (*bcd*) and *K10* mRNA (blue) are localised in an anterior ring, *gurken* (*grk*) mRNA (green) is localised in the dorso-anterior corner near the oocyte nucleus and *oskar* (*osk*) and *nanos* (*nos*) mRNA (red) are localised at the posterior pole. The minus end of the MT (blue) is at the anterior. (C) Localisation of pair rule transcripts and anterior/posterior determinants in the syncytial blastoderm embryo. Fushi taratzo and Wingless are patterned in seven expression stripes, or rings, around the periphery of the embryo. mRNAs (yellow) are apically localised 'above' each nucleus in the expression stripe. Expression of gap genes is determined by the stripy expression of pair rule genes. Posteriorly localised *osk* and *nos* mRNA (blue) and anteriorly localised *bcd* mRNA (red) are important for determination of A-P axis. (D) Representation of the organisation of the MT in the blastoderm embryo. MTs are nucleated and grow with the minus end toward the apical end of the cytoplasm, close to the cell membrane (adapted from (Costa and Schedl, 2001) and (Tekotte and Davis, 2002)). NC nurse cells; O oocyte; MT microtubules.



**Figure 1\_4 suggested model for the motor complex that drive mRNA localisation.**

RNA localisation signal (for example K10 TLS) usually resides in the 3'UTR and folds into stem-loop structure. It is recognised by different proteins that participate in splicing, nuclear export and translation regulation. Following export into the cytoplasm the ribonucleoprotein (RNP) complex associates with adaptor proteins, among them are Egl and BicD, which then bind directly to a motor protein Dynein and move along the cytoskeletal elements to their final destination, near the minus end of the microtubules.



## **2 Chapter 2 Structural determinants of *K10* RNA localisation signal**

### **2.1 Aims**

The aim of the first part of this work is to understand whether there are sequence or structural elements that are important for the TLS localisation, and try to determine, using mutational analysis, the role of the different elements in the signal and their contribution to its activity. The mutational analysis is based on random mutagenesis, extensive point mutations, and specific mutations aimed at changing parts of the signal assumed to be important for its function, based on previous works, or comparison to other localised transcripts and known RNA structures. The effect of the mutation on the signal is measured by the ability of the mutated transcript to localise in blastoderm embryo.

### **2.2 Injection assay**

In the studies contributing to this thesis, the main assay used to determine localisation of a given transcript is the injection assay (Figure 2\_1). It is performed by injecting fluorescently labelled transcripts into the basal cytoplasm of syncytial blastoderm embryo at cycle 14, just before cellularisation, three hours post-fertilisation. At this stage, microtubules are organised in the cytoplasm with the minus ends apical of the nuclei near the periphery of the embryo (Figure 1\_3D), and the plus ends close to or in the yolk, i.e. basally. Injected apically localised transcripts are recognised by the motor machinery and move along the MT to the minus ends. Localisation is stopped by fixing the embryos 5-8 min after the injection.

The localisation of the transcripts is then determined from images generated by a confocal microscope. Localised transcripts are classified into four groups according to their degree of apical localisation. Class I mutants appear unlocalised; class II and III mutants, in which apical accumulation is substantially weakened, and class IV mutants, which localise with similar efficiency compared to the wild type. There was no difference in localisation efficiency whether transcripts were injected individually or in pools. In some instances, a semi-automatic system, Quantile Ratio (QR), was used for determining the degree of localisation (Supplementary figure 2\_1), in an attempt to achieve an automatic quantification that could distinguish subtle differences of

localisation activities, but results were similar to those achieved when localisation was scored by eye (See examples in table 2\_1). Apical localisation of injected transcripts is proposed to be equivalent to the early stages of oogenesis (stages 4-5) in the context of the protein complex that recognises the signal (Bullock and Ish-Horowicz, 2001).

### **2.3 *K10 TLS can be further shortened and still retain localisation activity***

One of the possible determinants of the structural specificity of a localisation signal is its size (Hamilton and Davis, 2007), although the lengths of known localisation signals vary considerably. Very short localisation signals have been reported, such as the *Vgl* and *VegT* localisation signals that contain a series of redundant short motifs, among them the six nucleotides *Vgl* motif1 (VM1) and the five nucleotides E2 motif that are the binding sequences to VGRBP60 and VG1 RBP protein respectively (Bubunencko et al., 2002). Other localisation signals are long and complicated, like the 625 nt *bicoid* Localisation Element (BLE), in which different parts of the structure mediate different stages of localisation, but are all essential for complete and efficient localisation (Macdonald and Struhl, 1988). Nevertheless, it might be possible that the machinery recognise a specific structure depending on its size, within a localisation signal.

The 44 base *K10* Transport/Localisation Signal (TLS), one of the shortest signals known in *Drosophila*, was defined as being both sufficient and required for localisation (Figure 1\_5) (Serano and Cohen, 1995b). In order to examine whether the exact size of 44 bases of the TLS element is crucial for localisation, *in vitro* transcribed fluorescent transcripts containing mutated TLS elements of different lengths were injected into the blastoderm and assayed for their localisation ability. A mutant transcript lacking the C-G base-pair at the base of the stem-loop, (1-44 $\Delta$ CG) was injected into the cytoplasm of *Drosophila* blastoderm embryos. It appeared that this mutation eliminates the localisation ability of the transcript (Figure 2\_2a). Although the loss of localisation ability could be due to reduced size, it is also possible that deleting the C-G pair removes a key recognition base-pair or causes a gross disruption of the stem structure, affecting a required element elsewhere. The last possibility is supported by the mFOLD structure prediction of this mutant, showing a strongly disrupted lower helix (Figure 2\_2a), and by the fact that this G-C base pair is conserved in *K10* in *Drosophila melanogaster* orthologues (see chapter 3).

To distinguish between the three possibilities, a shortened mutant transcript lacking a U-A base-pair in the middle of the lower stem, 2-43 $\Delta$ AU, was created. The RNA folding program predicts that this mutation should not significantly destabilise the stem. Injected 2-43 $\Delta$ AU mutant transcripts localised as wild type (Figure 2\_2b), indicating that the length of the structure can be shorter than 44 nucleotides, and the exact size, by itself, is not a crucial factor in recognition, but that the C-G pair at the bottom of the stem-loop serves as an important stabilising feature in the lower helix. Further reduction of the TLS size by removing more base-pairs from the proximal (lower) stem, disrupted localisation, but removing nucleotides from the loop and from the distal (upper) stem did not affect localisation (See below).

Another indication that signal length is not the only important parameter was the inability of the antisense TLS sequence to localise. The antisense sequence was transcribed using a promotor site on the 3' end rather than the 5' end of the wild type TLS, thus with the same length and a similar CG content and base-pairing. This sequence is predicted to form a stem loop with two bulges on the 5' arm of the stem (Figure 2\_2c). These findings confirm that not any double stranded RNA of 44nt with a similar stem-loop structure can be recognised, and that more elements in the structure contribute to its localisation. A further analysis was performed to identify the key elements of the signal, and to examine whether one element can stand alone as a minimal localisation signal.

## **2.4 Random Mutagenesis analysis reveals global structural requirement for TLS**

In order to roughly define the structure and residues essential for localisation of the K10 TLS, extensive mutagenesis of the signal was performed. Multiple random mutations were made in the middle region of the *K10* stem loop. Mutant TLSs were derived from synthetic oligonucleotides with a 60% mutation rate in the region of the stem that includes both putative bulges (Figure 2\_3). A mutation rate of 60% means that each position is occupied by the wild type nucleotide in 40% of the transcripts, and by each one of the three other nucleotides in 20% of the transcripts. That rate produces an average of six to eight different mutations in each clone.

To maximise throughput in the assay, while ensuring that localised RNA could still be detected within a pool of several non-localising transcripts, mixtures of localised and non-localised transcripts were co-injected. Several ratios of non-localising and localising RNAs were tested, and 3:1 was chosen as suitable for the detection of weakly

localised transcripts. A pool of three mutated clones was mixed and transcribed together with a single fluorescent dye. Three differently coloured dyes were used, resulting in overall nine different transcripts being assayed together.

Of 250 injected mutant transcripts, four retained a wild-type localisation activity (Figure 2\_3). Their predicted structures show similarity to the *K10* wt stem-loop. They all have a stem and a loop, but mutants F4, D4 and H11 contain only one bulge, instead of two, while C9 has one bulge and one internal loop. The predicted structure of the bulges is different from wild type. While the wild type transcript has a bulge that contains one nucleotide, F4, D4 and H11 form a bulge with two nucleotides. C9 forms one internal loop and one bulge with only one nucleotide.

The free energy of the wild type signal is calculated by mFOLD as  $\Delta G = -11.2$  kcal/mole at 22°C. mFOLD searches for the lowest free energy ( $\Delta G$ ) fold in the molecule by maximizing the number of base pairs within a structure. In general, a higher number of base pairs would increase the stability of the structure, corresponding to a lower free energy (see Introduction). Transcripts F4, D4, H11 have a predicted structure that is even more stable than the wild type structure ( $\Delta G = -13.5$ ;  $-14$ ;  $-13.2$  kcal/mole respectively), which suggests that one of the important features in the signal is its stable structure, since unstable structure has more tendency to create different alternative conformations that might not be recognised by the motor complex.

These results are not comprehensive enough to determine a shared structural motif. However, the predicted presence of at least one bulge in each localising transcript suggests that the presence of a bulge is a requirement for localisation. The bulge can contain either one or two nucleotides and its position in the stem can vary. Nevertheless, analysis of additional mutants was needed to understand the role of the bulge.

## ***2.5 Single stranded nucleotides play a structural role in recognition of the TLS***

Double stranded RNA helices have narrower grooves than the double stranded DNA helix, which restricts the spatial binding site, and preclude recognition of the base identity of nucleotides in the helix (Bloomfield et al., 2000; Carlson et al., 2003). Hence, I expected that associated proteins would bind to single stranded nucleotides in an RNA structure, rather than to the double stranded helix.

This hypothesis was strengthened by the observation that random mutagenesis suggests that at least one bulge or internal loop is required for localisation of the transcript. I therefore started investigating the importance of the predicted single

stranded nucleotides in *K10* TLS. The *K10* signal contains two types of single stranded structures: a loop and two bulges (Figure 1\_5). To determine if these elements are important for recognition, different mutations were introduced to disrupt the single-stranded nucleotides.

### **2.5.1 The loop is important for localisation but not its exact size or base identity**

The *K10* signal contains a loop at the end of the helix in bases 18-25 (Figure 1\_5). The requirement for the presence of a loop for localisation became evident from the observation that a transcript that lacks the loop sequence but contains only a helix with two bulges did not localise when injected into the embryo. It was earlier reported that the length of the loop can be reduced to five nucleotides without affecting localisation, but there is an upper size limit of nine nucleotides (Cohen et al., 2005).

I showed that the loop can be further reduced and that the base identity of the loop is not crucial for localisation (Figure 2\_2d). The tetraloop is a common hairpin loop motif that caps many double helices. A sequence of (C)UUCG(G) is a known canonical tetraloop used to stabilise an RNA helix in structural studies requiring stable RNA constructs to perform NMR studies (Antao et al., 1991; Cheong et al., 1990; Ennifar et al., 2000; Woese et al., 1990). Not only does this change reduce the loop size further to four nucleotides, but also changes the nucleotide identity completely. The tetraloop swap did not interfere with localisation (Figure 2\_2d), indicating that there is even more flexibility in the size and sequence of the loop than previously suggested (Serano and Cohen, 1995b). Therefore, the loop appears to be important as a structural element, rather than a sequence recognition element. It allows the formation of a double helix and also helps preserve the overall structure of the *K10* proximal helix. From a conformational point of view, a helix cannot be formed without a loop at the end, as the intermolecular bonds in the sugar-phosphate backbone would not be long enough to extend between two upper base-pairs.

### **2.5.2 The presence of bulges is important for the TLS structure, but their identity is not crucial**

Bulges occur when a base-paired stem is interrupted by single-stranded nucleotides on only one strand (as opposed to an internal loop which can have un-paired nucleotides in both strands). Bulges are important for secondary and tertiary structure formation (Turner, 1996) and can specifically interact with proteins (Dingwall et al.,

1990). In fact they are among the most common non-Watson–Crick features in RNA (Hermann and Patel, 2000; Westhof and Fritsch, 2000).

Previous work has shown that the presence of bulges is important for proper localisation of endogenous *K10* transcript in the oocyte. *LacZ* tagged transcripts, lacking both bulges, become enriched in the oocyte during early stages of oogenesis (stages 4-5), but localisation to the anterior cortex is diffuse and does not persist beyond stage eight of oocyte development (Cohen et al., 2005). To check if that is also the case when *K10* transcripts are injected into the embryo, a similar construct, *d33d37*, lacking the two bulges in the structure, was used to produce mutated transcripts. The *d33d37* transcript did not localise apically in the blastoderm (Figure 2\_4), confirming that the bulges are crucial for localisation.

To explore the significance of the individual bulges, two more mutants were created. Mutant transcript *d33* lacks the upper bulge (cytosine) and *d37* lacks the lower bulge (adenine). Removal of either bulge had only minor effects on localisation. Removal of the upper bulge had a minor effect while deleting the lower bulge had a more pronounced effect, but still less than the cumulative effect of removing both of them (Figure 2\_4, Table 2\_1). This suggests that the two bulges work additively to contribute to full localisation, and that although each one of them can be sufficient for weaker localisation, both bulges have a role in recognition of the signal by the transport complex.

I wanted to understand whether bulge recognition is sequence dependent or if it is only a structure-related property. Transcripts with point mutations in either bulge were generated and their ability to localise was validated. Cytosine in position 33 (C33) was mutated to A, G or U, and adenine in position 37 (A37) was mutated to U, G or C. There was little or no effect of those mutations on localisation activity (Figure 2\_4 and table 2\_1). Therefore, the presence of a bulge, rather than the exact identity of the nucleotide creating the bulge, is important for signal activity.

Having shown that both bulges are important for full localisation, I wanted to understand the spatial relations between them, focussing on the importance of their distance from each other. I therefore generated a mutation that preserves the initial structure, but changes the length of the intervening stem by adding more U-A base pairs between the two bulges. Alteration of the distance between them had no effect on localisation. Thus, there is some degree of flexibility in the length of the stem between the bulges (Figure 2\_2e).

### 2.5.3 Further analysis of the function of bulges in the TLS: The structure of orb localisation signal is a stem loop with one bulge

The motif of stem-loop with two bulges appears in the predicted structure of signals of other *K10* homologues in the *Drosophila* groups (See chapter 3), emphasizing the importance of bulges, as they are evolutionary conserved through 50 million years of divergence. In contrast, the *orb* signal, has a similar localisation pattern to *K10* in the oocyte and in the embryo, and a similar secondary structure, but the signals of *orb* in all *Drosophila* species contain only one bulge.

*orb* RNA, like *K10* transcripts, is transcribed in nurse cells and transported into the oocyte in early stages of oogenesis (stages 2-6). In later stages (7 onward) it is localised anteriorly inside the oocyte from stage 8 to 10b. *orb* transcripts contain a localisation signal that resides in the 3'UTR. This signal has no sequence similarity to the *K10* signal, but the predicted secondary structure, similar to that of the *K10* TLS, folds into a stem-loop structure.

The similarity in predicted structure of the *K10* TLS and the *orb* LS and their similar localisation indicate they could share a common recognition motif. I therefore analysed the *orb* localisation signal. The distal helix is the only part that is identical between the *orb* and *K10* signals, and consists of a stretch of U-A base pairs. The loop is bigger and contains five nucleotides, and the stem is interrupted by only one bulge (Figure 2\_5).

Injection of the *orb* signal into blastoderm results in full localisation (Figure 2\_5). As *orb* contains only one bulge in its lower part, I decided to examine the importance of this bulge and possible consequences of alterations to it. *orbd36*, a transcript with a mutation that created a stem-loop with no bulges was injected, and as predicted, did not localise (Table 2\_1).

Although the mFOLD program predicts very similar structures for the signal responsible for *orb* localisation and *d33*, (the *K10* mutant signal that lacks the upper bulge), *orb* localisation activity is similar to that of *K10* rather than to that of *d33*. In order to understand why these two structures, despite being similar, have different localisation activity, two hybrids were constructed. The hybrids adjoin the two different signals. *OK* contains the lower part of the *orb* signal, and the upper part of the *K10* signal. *KO* contains the lower part of the *K10* signal and the upper part of *orb* (Fig. 2\_5).

Localisation of *OK* is similar to that of a wild type *K10*, while *KO* hybrid loses *K10* localisation activity and is slower than the *d33* mutant (Figure 2\_4). The predicted structure of *OK* has two bulges, while that of the *KO* only has one. A possible explanation for the difference in localisation between the two hybrid mutants would be that there is a tertiary interaction between the lower and the upper part of the stem-loop structure and that the bulges are important to structurally maintain this interaction with the upper part, either by direct contact, or by multimerisation. In *orb*, probably due to subtle changes in secondary structures that cannot be predicted *in silico*, the area around the lower bulge “compensates” for the lack of second bulge. This compensation is not present in the hybrid *KO* (see summary).

## **2.6 Point mutations in nucleotides in the proximity of the bulges**

The hybrid results showed that more factors might affect the shape of the helix than the bulges per se. It might be that the region around the bulges also plays a role in maintaining the shape of the stem-loop. Adjacent nucleotides can change the spatial organisation of their neighbours, and single strands adjacent to helical regions can form tertiary contacts with base-paired nucleotides of the helices (Leontis et al., 2006).

So far, the results pointed out the importance of the structural elements of the *K10* signal, and there was no evidence of a direct recognition of specific nucleotide/s in the structure. To study the effects of the nucleotide identity in the area around the bulges, a set of single point mutations was generated in an area comprised of 12 nucleotides: five in the 5' arm, and seven in the 3' arm (Figure 2\_3). The 12 nucleotides in this segment were mutated to each of the other three possible nucleotides, to generate 36 mutant transcripts.

Most point mutations do not have a significant impact on overall localisation ability (Table 2). Out of 36 injected mutated transcripts, 22 localise well (scored 3 or 4, see scoring system in chapter 8 and in Figure 2\_1), while 10 transcripts localise weakly, and four mutated transcripts appear not to localise at all. The non-localising transcripts are mutated at position six (6U->G and 6U->C) and position eight (8G->A and 8G->C).

The mFOLD program was used to predict the structure of the mutants; these structures were analysed in order to highlight common features within the localising group and within the non-localising group, and to search for differences between the two groups. The exact position of the bulges along the stem does not seem to be important, as mutations, which created a different structure from the wild type but

preserved at least one bulge in position 37, 38, or 39 along the stem, localised well. In most cases, poorly localised transcripts had a predicted structure that changed the lower bulge, usually to a bigger loop. The upper bulge did not appear to have an effect that correlated with localisation.

Point mutations that abolished localisation were found in 6U and 8G. Generally, mutations on the 5' arm had more effect than those of the 3' arm. Based on the predicted structure, I expected that if one nucleotide that is part of the base pair disrupts the structure, its reciprocal base pair would have a similar effect. This is not the case in *K10*. A mutation in uracil in position 6 abolished localisation of the transcripts, but while 6U is predicted to interact with 39A, mutations at 39A do not have such a strong effect on localisation activity. Only changing 39A to T reduced localisation. This might suggest that recognition is specific to nucleotide 6U and that its base identity is important.

## **2.7 Summary**

*K10* signal length can be changed for a longer or shorter stem, as long as the change does not affect important elements of the signal. It is possible to reduce the length even further by removing specific base-pairs or reduce the size of the loop. This suggests that it may be possible to isolate each one of the elements, and determine its significance to the overall signal activity. The fact that the length between the bulges can change suggests a separate binding site, perhaps for more than one adaptor protein.

The two bulges play an important role in the structure. Although Cohen and colleagues initially claimed that these elements are not important for localisation in the oocyte (Serano and Cohen, 1995b), a re-evaluation of the results revealed a slight loss of activity of a transcript lacking both bulges (Cohen et al., 2005). A similar mutant transcript lacking the two bulges completely loses activity when injected into the blastoderm embryo (Fig 2\_4). The difference in the results could be due to one of several reasons. First, the assay I used includes injecting labelled transcripts into the cytoplasm of embryo. Localisation of injected RNA in the embryo is a fast process, compared to localisation of endogenous transcripts in oocyte, the assay used in Cohen's lab. The long time that is required to follow the development of an oocyte, might allow "correction" of slower transcripts, i.e. slower transcripts should also arrive to their destination, while in a the injection assay, where localisation is determined after few minutes, slower transcripts would appear as unlocalised.

Another possibility is that anchoring of the transcripts might also have an effect on the overall localisation. Transcripts that arrive at their destination are anchored there by a mechanism that differs from that of the transport machinery (Delanoue and Davis, 2005; Delanoue et al., 2007). The anchorage machinery might be less specific than the transport machinery and so, once transcripts arrive close to their destination (even if not by direct transport) they are anchored, and after a while can accumulate in the correct position.

It is also possible that the difference in the constructs used in this work and in Cohen's work (i.e. the context of the LS) can influence the activity of the transcript by a compensation, or disturbance to the TLS. The construct used in Cohen's lab is called KZK(Cheung et al., 1992), and contains the 5'K10 UTR, part of the K10 coding region flanked by LacZ and the first 50 base pairs of the K10 3'UTR. The sequence of the TLS has BglII-XbaI sites introduced around it. I used a vector that contains the coding part of the 3' coding region and the whole of the UTR with NheI and HindIII sites around the TLS. These differences should not be significant, as it was shown that the TLS, by itself is sufficient and essential for localisation.

An extensive random mutagenesis revealed a group of heavily mutagenised transcripts that are able to localise. The number of localising transcripts was surprisingly small compared to other screens, in our lab and other labs. Only 1.6% of the mutated transcripts had localisation activity. In other screens for localised transcripts, a higher percentage was found. For example, 10% of the endogenous transcripts were found to localise in the oocyte and 71% in the early embryo (Dubowy and Macdonald, 1998; Lecuyer et al., 2007) and 24% of mutant *hairy* transcripts were apically transported when injected to the blastoderm embryo (Bullock et al., 2003).

There could be several reasons for the small number of localising mutant transcripts. One reason is that the sequence/structure of the TLS is so robust, that any few changes can affect its localisation activity. This is not likely to be the case, as many point mutation that were produced in other areas of the signal did not have an effect on its localisation. Most likely, the percentage of mutated nucleotides in my experiment was very high and these mutations changed the structure extensively. Moreover, the mutated area is important for the whole of the structure, as it contains the two bulges that create a structural recognition motif and connect between the two helices.

Although the injected RNAs are transported within particles that contain the motor and adaptor proteins, together with several different transcripts, localised and non-localised transcripts do not have a mutual effect on the overall localisation of each

other (Bullock et al., 2006). This was also confirmed by injecting a group composed of a mixture of known localised and non-localised transcripts before the start of the screen. Hence, it is probably not a case of non-localised transcripts attenuating the localised particles.

The fact that the identity of the bulges is not important for the activity of the signal was confirmed by the site directed mutagenesis. The nucleotides comprising each bulge were mutated to all three other nucleotides, without altering localisation activity of the TLS. Moreover, the identity of the bulges in homologues *K10* localisation signals found in other *Drosophilae* are also not conserved (See chapter 3). Nevertheless, the presence of bulges, as well as of the loop, is crucial for localisation activity of the TLS, as removing them cause elimination of localisation activity.

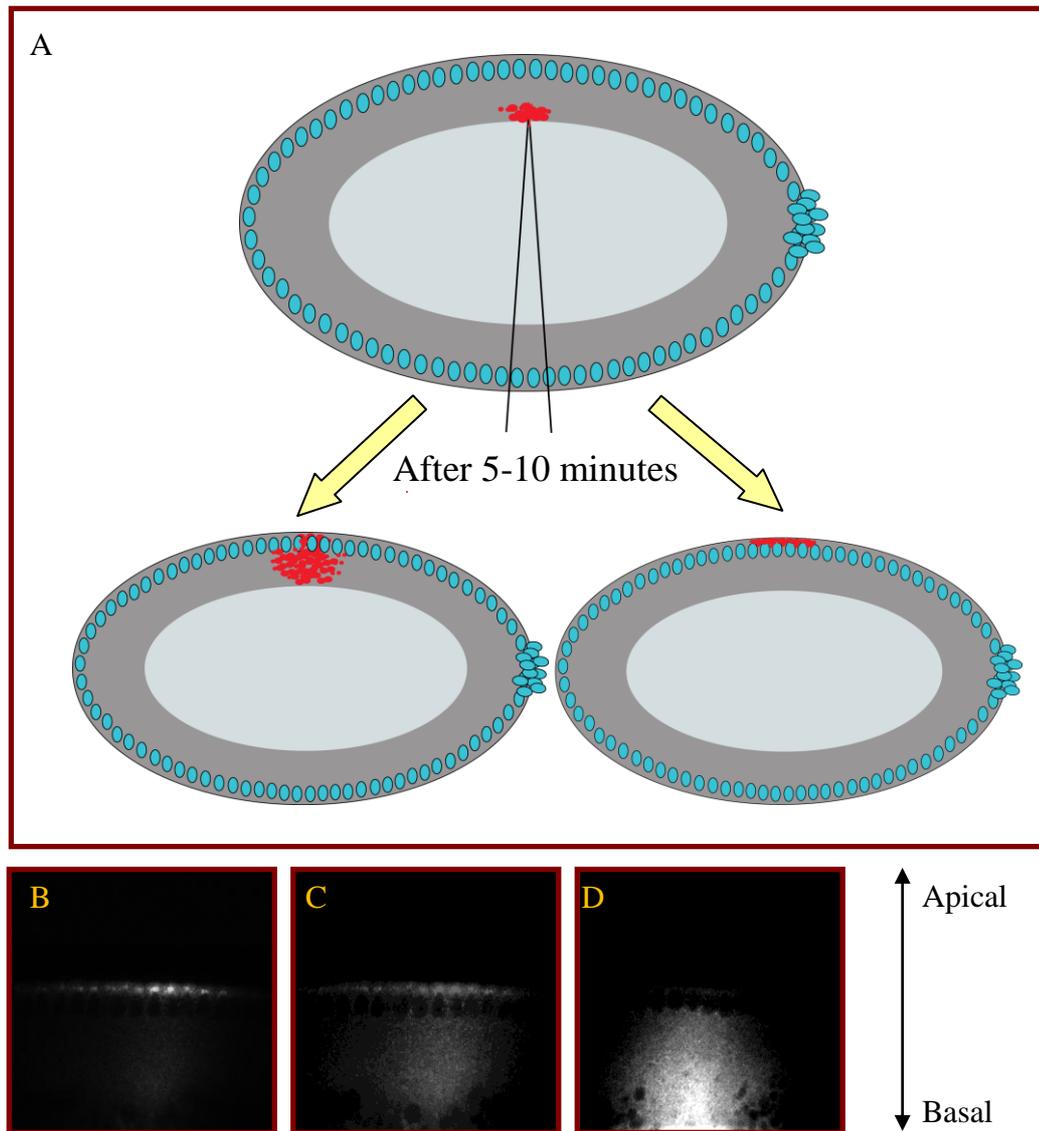
Nucleotides next to the bulges are more sensitive to mutations, especially on the 5' arm of the TLS. Mutations in U6 and G8 had strong effect on localisation activity. The reason is probably due to the fact that bulges act to locally distort the helix structure, and to uncouple the helices. Bases positioned near single stranded nucleotides are more accessible to protein binding or tertiary folding and catalysis (Weeks and Crothers, 1993). Interestingly, reactivity to chemical binding is more pronounced in the 5' rather than in the 3' direction of the helix close to the bulge (Weeks and Crothers, 1993). These bases might also be directly recognised by a protein adaptor

When analysing the localisation ability of the different mutants, I detected a partial correlation between efficiency of localisation and stability of the predicted structure. Seventeen out of the twenty-four strong localisers create a structure with a predicted stability that is the same or close to the wild-type predicted structure. Only one of ten weak localisers has a predicted stable structure, and the non-localisers do not form a stable structure. Computer predictions indicate that RNAs for which the secondary structure is functionally important have lower folding energy than random RNAs of the same length and dinucleotide frequency (Clote et al., 2005). This indicates that the stability of the structured RNA can be important for its activity. However, the correlation is incomplete. For example, a transcript of TLS lacking the two bulges, forms a structure more stable than the wild type TLS, but is not localised.

Another correlation between unlocalised transcripts and their predicted structure is the presence of the lower bulge. Most mutations that affected localisation also changed the predicted structure so that the lower bulge disappeared. This, together with the fact that the *orb* LS contains only the lower bulge, hint to the importance of the lower bulge. Although removing the lower bulge has only a small effect on localisation,

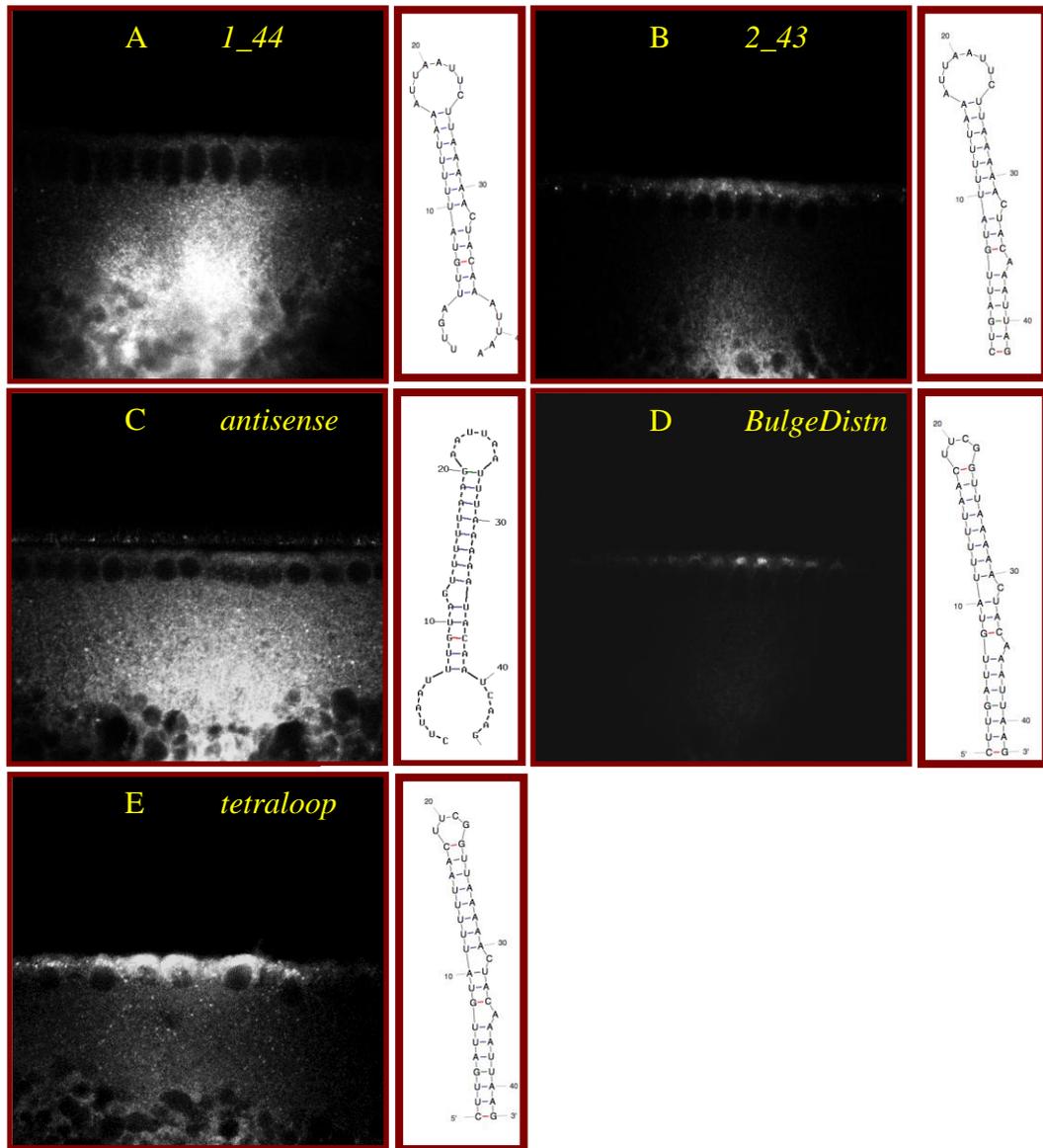
it might act in concert with other elements in the signal, mainly the upper bulge (or alternatively the upper stem in *orb*) and the neighbouring nucleotides, to achieve full localisation

The predicted structure of the efficiently localised mutant transcripts showed that the helical structure still forms, and it has at least one internal loop in the middle. The importance of this motif would need to be confirmed by other means, for example, comparing TLS from different *Drosophila* species (chapter 3), to see whether they contain the same motif, and by structural studies (chapter 5) to verify that this is indeed the 3D structure.



**Figure 2\_1 Injection assay and representative examples of scoring method.**

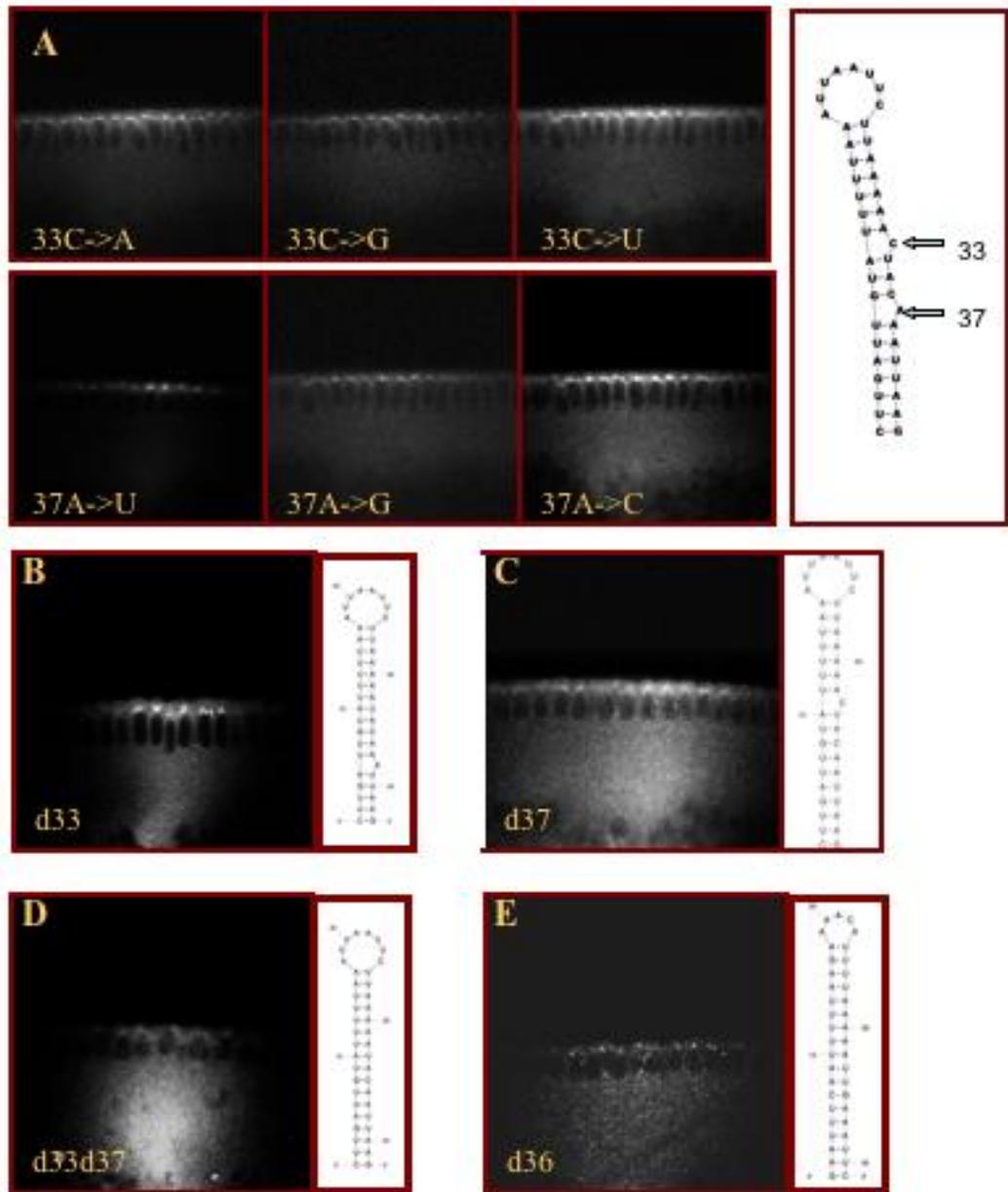
(A) Injection assay. Fluorescently labelled transcripts are injected to the basal end of the cytoplasm of syncytial blastoderm embryo. After 5 to 10 min, embryos are fixed. Localisation of the transcripts is determined using images generated by confocal microscopy. Scoring of the localisation is based on estimating their relative apical accumulation compared with their distribution in the basal cytoplasm between the nuclei and the injection site. (B) Transcripts are defined as localised if they accumulate in the apical end of the cytoplasm. Score = 1 (C) Transcripts are defined as weak localisers when some particles accumulate in the apical end but there are still particles in the cytoplasm. Score = 2. (D) Non-localised transcripts appear diffused around the cytoplasm. Score = 3.



**Figure 2\_2** Representative mutations in the *K10* TLS.

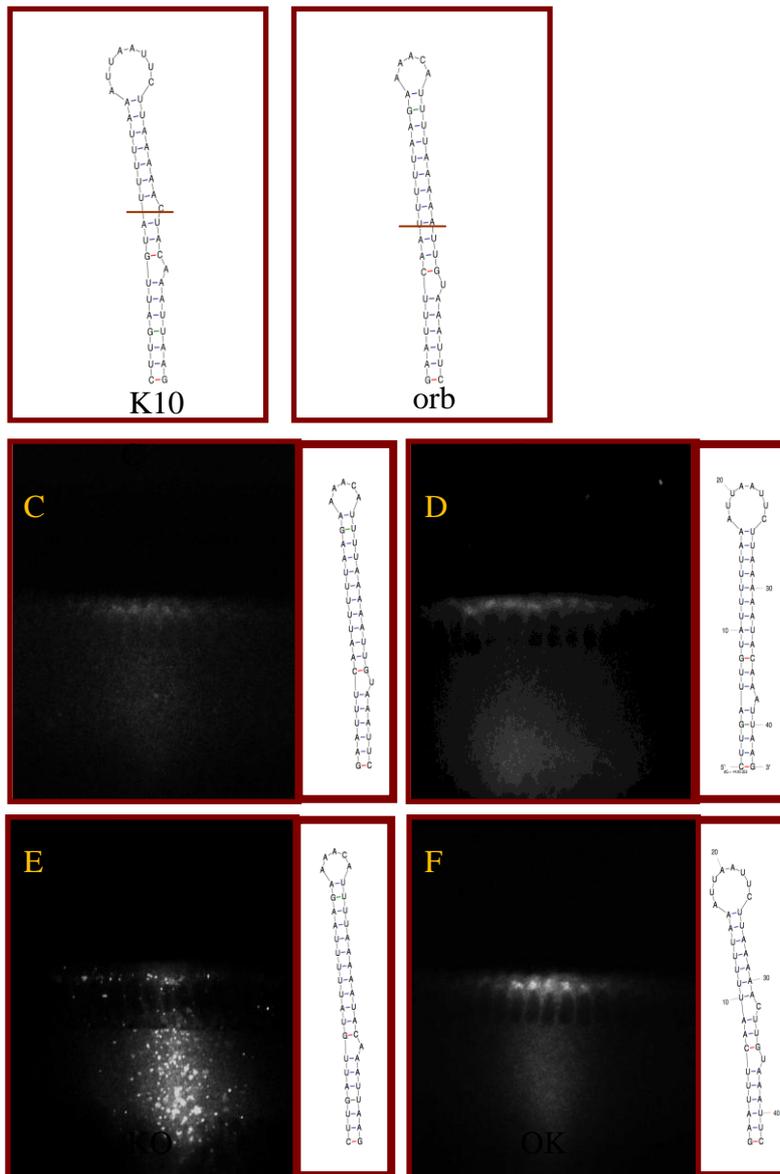
(A) Mutant *1\_44* lacks the G-C base pair at the stem of the signal that serves to stabilise the double stranded stem. Removing it abolishes localisation, probably because it alters the structure of the lower stem. (B) Mutant *2\_43* lacks base pair A-U in the second position, but localises normally. (C) A transcript expressing an antisense TLS fails to localise despite retaining the same base-pairing as a wild-type TLS. (D) Transcript with and added AU base pair between the two bulges localise well. This shows that the exact distance between the two bulges is not crucial for localisation. (E) Replacing the loop sequence with a tetraloop does not affect localisation.





**Figure 2\_4 The bulges serve a structural role.**

(A) Activity is retained when changing the base identity in each bulge. (B) and (C) Deleting either bulge reduces localisation. (D) Removing both bulges abolishes localisation. (E) Removing the only bulge from the *orb* signal abolishes localisation.



**Figure 2\_5 Hybrids of K10 TLS and orb LS have a different localisation pattern.**

Secondary structure of (A) K10 TLS and (B) orb LS shows a similar structure, but orb lacks the upper bulge element. Localisation patterns of (C) orb LS and (D) d33, a TLS mutant that lacks the upper bulge, shows that orb LS is localised in a similar fashion to TLS, but the d33 has reduced localisation ability. Localisation of the K10-orb hybrids (E) KO does not localise, while (F) OK exhibit localisation similar to that of a wild type K10. Red bars represent the point where the hybrids were connected. KO is a hybrid that is constructed from the lower part of K10 and the upper part of orb. OK is a hybrid that is built from the reciprocal parts.

Base	Mutated to...	Localisation	N	QR score	Predicted structure	
					Upper bulge	Lower bulge
U6	A	2	17	0.59	X	V
U6	C	1	17	0.3	X	V
U6	G	1	19	0.4	X	V
U7	A	2	19		X	V
U7	C	3	14	1.8	X	V
U7	G	1	14		X	V
G8	A	1	15		X	V
G8	C	1	16		X	V
G8	T	2	24	0.6	V	V
U9	A	2	16		X	X
U9	C	3			X	X
U9	G	2	11		X	X
A10	C	3	17		V	X
A10	G	4	11		V	V
A10	T	4	19		V	X
C33	A	4	25	1.73	V	X
C33	G	3	15		V	X
C33	T	3	14		V	V
U34	A	3	11		V	X
U34	C	3	12		V	X
U34	G	3	16		V	X
A35	C	2	13		X	X
A35	G	2	17		V	V
A35	T	4			X	X
C36	A	2	29		X	V
C36	G	2	14		X	V
C36	T	4	17		V	V

Base	Mutated to...	Localisation	N	QR score	Predicted structure	
					upper bulge	lower bulge
A37	C	3	17		V	V
A37	G	3	17		V	V
A37	T	4	15		V	V
A38	C	4	11		V	V
A38	G	4	21		V	V
A38	T	4	16		V	V
A39	C	3	17		V	V
A39	G	3	13		X	V
A39	T	2	11		X	V
WT		4	30	2.9	V	V
Scmbd	Scrambled sequence of TLS	1	20	0.2	X	X
Tetraloop	loop	4	18	2.7	V	V

**Table 2\_1 Summary of point mutations in the TLS.**

Localisation was determined by scoring the apical accumulation of injected embryos so that 1=poor localisation and 4=wt localisation. N represents number of embryos that were scored. In some cases scoring was compared to scoring achieved by QR, a computer based analysis, in which the ratio between the brightness of the area around the nuclei (where apically localised particles accumulate) and the brightness of the cytoplasm area, where non-localised particles are concentrated. Higher ratio represents better localisation. The appearance of the bulges in the predicted structure of each one of the mutant signals is shown in the table, as it seems that the appearance of the upper bulge disappear from structure in non-localisers.

## **3 Chapter 3 Evolutionary divergence of *K10* and *orb* localisation signals within *Drosophila* species**

### **3.1 Aims**

The aim of this chapter is to further study the important elements of the *K10* and the *orb* localisation signals by identifying conserved sequence and structural elements. I made use of computer programmes that compare genomic sequences to identify the likely localisation signal regions in different *Drosophila* species. The activity of the signals was tested by injecting them into *D. Melanogaster* embryos. I then analysed their putative secondary structures *in silico*, and compared them. The same procedure was used to analyse the localisation signal of *orb* in different *Drosophila* species. Finally, I used a computer program to establish a consensus structure of all localising transcripts in order to identify conserved sequence and structural elements in the signal.

### **3.2 A comparative study of homologous *K10* Transport/Localisation Sequences (TLSs) reveals conserved and diverged regions of the signal**

Evolution of species, nature's own mutagenesis experiment, enables a comparison of related functional sequences of different organisms. Functionally important bases are likely to be evolutionary conserved, especially in untranslated sequences, which are fast evolving, whereas less important bases might vary. Therefore, a comparison of homologous signals may imply functionally important regions of the localisation signal and complement the study of mutated TLSs described in the previous chapter. In this section, I analyse the variation of the *K10* TLS and the *orb* localisation signal, which occurred during the evolution of *Drosophila* species.

Homologous sequences from 9 *Drosophila* species were identified by BLAST search (Basic Local Alignment Search Tool; [http://www.ncbi.nlm.nih.gov/sutils/genom\\_table.cgi?organism=insects](http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=insects) and <http://flybase.org/blast/>) and aligned using the programme ClustalW (Figure 3\_1 a). Conservation among the sequences is high: *D. melanogaster*, *D. simulans*, *D. sechellia* and *D. erecta*, which are closely related species (Figure 3\_1 B), contain identical TLSs. Identity between *D. melanogaster* TLS and the other homologous sequences ranges from 92 % (*D. willistoni*) to 97 % (*D.*

*yakuba*). Nucleotide changes occur in 9 positions, whereas 36 positions are invariant (marked with asterisks in Figure 3\_1 A).

There was low or no conservation of the TLS sequence in other areas of the 3'UTRs of the different species. Indeed, the only region of the 3'UTRs of most of the species (*D. ananassae*, *D. grimshawi*, *D. virilis*, *D. mojavensis*, *D. persimilis*, and *D. pseudoobscura*) that is similar to that in *D. melanogaster* is the TLS.

To check if the 6 homologous TLSs whose sequences diverge from the *D. melanogaster* TLS are functional and can be recognised by the transport machinery of *D. melanogaster*, they were analysed by injection assays in *D. melanogaster* syncytial embryos (as described in the introduction). In order to exclude influences of 3'UTR sequences outside the TLS, all TLSs were cloned into the *D. melanogaster K10* 3'UTR such that they replace the *D. melanogaster* 44 nt TLS. All homologous TLSs with the exception of (distantly-related) *D. grimshawi* efficiently mediated localisation of the respective transcript to the apical side of the cytoplasm, similar to the wild-type *D. melanogaster K10* signal (Fig. 3\_2 B).

In order to determine how much the localising homologous TLSs vary in their structure, I predicted the consensus secondary structure using the Alifold programme, which calculates the structure of aligned RNA sequences based on the minimal energy model (see the Introduction for more details) (Hofacker et al., 2002), and used the mFOLD software to predict minimum free energy structures of the individual sequences (Zuker, 2003). The resulting structures are very similar to the predicted structure of the *D. melanogaster* TLS (consensus structure: Figure 3\_2 A; individual structures: Figure 3\_2 B).

Several of nucleotide changes between species convert Watson-Crick pairs into wobble base pairs or vice versa. For example, a change of adenine in position 17 (A17) in the *D. melanogaster* TLS to guanine in *D. pseudoobscura*, changes the pair from a Watson-Crick A-U into the wobble base pair G-U. Both base pairs show a similar thermodynamic stability (Varani and McClain, 2000) so that such nucleotide changes do not alter the modelled structure. Similar changes are A29 to G29 in *D. pseudoobscura*, A39 to G39 in *D. ananassae*, and G4 to A4 in *D. mojavensis* and *D. virilis*.

Changes in particularly interesting positions are A18 to G in *D. pseudoobscura*, which locates to the loop, A37 to C in *D. yakuba*, which is an unpaired nucleotide forming bulge I and C33 to U in *D. mojavensis* and *D. virilis*, which is the other unpaired nucleotide of bulge II. The maintenance of localisation of these signals

suggests that the nucleotide identities in loop and bulges are not crucial for the localisation activity of the TLS.

The most drastic sequence change was found in two neighbouring nucleotides, C33 to U33, and U34 to C34 in *D. pseudoobscura*. These two consecutive changes interrupt the base pairs and therefore change the structure. In particular, these changes lengthen the distal helix, and shorten the distance between the bulges, since another A-U base pair was added to the distal helix instead of the helix between the bulges (Figure 3\_2 B). Despite this (and other) changes, the *D. pseudoobscura* localisation signal mediates apical localisation in injection assays, suggesting that the predicted structural changes, i.e. length of the helices, do not affect the structural elements which are required for localisation, or that the structure prediction is inaccurate.

Interestingly, the two G-C base pairs in the lower and middle helices are conserved, probably for keeping the structure together, because G-C forms a stronger base-pair, with three hydrogen bonds. Changes that do not preserve base-pairing, occurred only in the bulges and loops.

The reason why *D. grimshawi* does not localise efficiently could be due to insertion of an A in an important area, the proximal bulge. Another option is that the insertion causes a change of the structure, which again prevents proper recognition or binding of the localisation complex proteins.

In conclusion, comparison of homologous TLSs from various *Drosophila* species reveals a high level of predicted structural conservation. The identity of the nucleotides in the bulges and in the loop, however, is less well conserved, indicating that the specific sequence of bulges and loop are not essential for recognition of the signal. These results corroborate similar conclusions drawn from *D. melanogaster* TLS mutagenesis experiments (see chapter 2). An additional aspect of the predicted structure, which is not conserved and therefore unlikely to be functionally important, is the length of the three double stranded helices.

### **3.3 Conservation and divergence of homologous *orb* localisation signals**

Most known localisation signals of *D. melanogaster* transcripts show no overt similarity in sequence or structure. Exceptions are the localisation signals of *orb* and *K10*, which have similar predicted structures (despite a low level of sequence similarity) (Lantz et al., 1992; Lantz and Schedl, 1994; Serano and Cohen, 1995b). These two signals also share similar localisation patterns both during oogenesis (Serano and

Cohen, 1995b), and in the injection assay in syncytial embryos (Bullock and Ish-Horowicz, 2001), suggesting that they are recognised and localised by the same machinery (also see chapter 2 and Figure 2\_5).

Orb (oo18 RNA-binding protein) is required during oogenesis for the determination of the dorsoventral and anteroposterior axes of egg and embryo (Christerson and McKearin, 1994). Orb also promotes the expression of K10 and its expression is, in turn, negatively regulated by K10 (Chang et al., 2001). Since the *orb* localisation signal seems to have a very similar structure to the *K10* TLS, I tried to generate a general consensus structure based on the comparison of several *orb* and *K10* homologous localisation signals. For this purpose, I analysed *orb* localisation signals from 11 *Drosophila* species, using the same programmes and techniques as in the previous section.

Closely related species *D. melanogaster*, *D. simulans*, *D. sechellia* and *D. yakuba* have identical *orb* localisation signals. Despite a higher degree of variation in the sequences of the other 7 homologous *orb* signals (74 % to 97 % divergence compared to *D. melanogaster*), similar structural conservation was found, except for *D. mojavensis*, which has an extra bulge in the 5' arm of the signal. All of the transcripts (with exception of the *D. mojavensis* transcript) localised in the localisation assay. The extra bulge in the structure of *D. mojavensis* might be the reason why this localisation signal did not localise, as it imposes a big change to the distal helix, which was shown to be important for localisation (see chapter 2).

The *orb* signal is less conserved among species compared to *K10*, which can make the analysis more informative, as it adds more variations. Similar to the *K10* signal, the loop and bulge are variable (See figure 3\_2). But unlike TLS, there are some changes that do not preserve base-pairing. For example, a change of A to C in position 9 in *D. mojavensis*, or a change of A to C in position 33 in *D. ananassae*. Another difference lies in the first base-pair of the helix, which is conserved and important in the *K10* TLS (See chapter 2); in the case of *orb* LS, the G-C base pair was changed to A-U in *D. mojavensis* and *D. virilis* without affecting its localisation.

Combining the results for the *K10* TLS and the *orb* localisation signals, the loop size range varies from eight nucleotides in *K10* of *D. melanogaster* to three nucleotides in *orb* of *D. virilis*. This confirms the previous results showing that the loop is necessary but its sequence composition and size can be changed (chapter 2). It appears that one bulge is enough for full localisation of *orb*, but not of *K10*. The base pairs are also kept in other signals, for example, the change of a U-A to C-G base pairing in position six

and 37 in *D. virilis*. The general consensus structure of *K10* and *orb* elements is more similar to *orb*, as it contains only one bulge, and the loop is smaller.

### 3.4 Summary and discussion

The conserved overall structure of the homologous *K10* and *orb* localisation signals highlights the functional importance of this element, although some regions are less well conserved and presumably are less constrained functionally. Based on the results of the comparative studies in this chapter and on the mutagenesis approach described in the previous chapter (Figure 2\_6), the following features of the localisation signal have been revealed:

a) Important conserved base pairs are C-G or G-C pairs at the base of the structure which presumably stabilise the proximal helix (Figure 2\_2). Removing these base-pairs abolishes localisation of the mutant transcript, and creates a predicted structure with a big loop in the proximal region (see chapter 2). This base pair can be changed under certain conditions. For example, in the *D. virilis orb* signal, the G-C is changed to A-U without affecting localisation.

b) A stretch of U-A at the distal helix is highly conserved, in both the *orb* and *K10* LSs. NMR results (chapter 5) show that this A-U rich double stranded part forms a specific structure that is important for localisation.

c) A stretch of U-A or A-U composing the proximal helix is highly conserved.

d) The single stranded nucleotide identities (loop, bulges) are not conserved, although the presence of at least one bulge and a loop is conserved and is necessary (Figure 2\_4 and 3\_2) for full localisation activity, as the group of *orb* signals, which are all predicted to contain one bulge, localised well. Changes that do not preserve base-pairing occurred only in the bulges and loops.

d) Specific single nucleotides which were found to be important in these signals by mutagenesis (see chapter 2) were also conserved in the homologous LSs of *K10* and *orb*. Uracil in position six and five in *K10* and *orb* respectively impairs localisation if mutated; it is conserved in all sequences. This is also the case with U7, 8G, 35A and 36C (Figure 2\_6)

e) The proximal bulge is important. Removing it reduces localisation (chapter 2), and changing its size by adding another single nucleotide, as in *D. grimshawi* might explain why the *D. grimshawi* TLS is not localising efficiently in *D. melanogaster*. An alternative explanation is that the insertion causes a change of the overall structure, which again prevents proper recognition or binding of the localisation complex proteins.

It would be interesting to check whether endogenous *K10* also localises inefficiently in *D. grimshawi*, and whether it affects the development and cytoarchitecture of the embryo.

It would also be interesting to study whether *K10* TLS from *D. grimshawi* can be localised in *D. grimshawi* and *D. mojavensis* embryos in order to understand the requirements of the machinery for the consensus structure. It is possible that these organisms have a different localisation signal that directs localisation of *K10*, or *orb* transcripts. Another option is that the motor complex in these organisms has different requirements, and even though the *D. melanogaster* motor complex does not recognise these signals in the embryo, these same sequences can be recognised by the motor complex in *D. grimshawi* and *D. mojavensis* embryos, or in their oocytes.

Since RNA can form many different kinds of structures, it is difficult to find and compare secondary RNA structures in the genome. Hence, the number of localisation signals that can be identified using *in silico* predictions is limited. Studying other localisation signals could contribute and enrich our knowledge about the important elements in the structure. First, other, more distant homologues of TLS-in other insects, for example *Anopheles gambiae* (African malaria mosquito), *Apis mellifera* (honey bee), *Nasonia vitripennis* (jewel wasp), *Tribolium castaneum* (red flour beetle) etc. could be studied and compared to the consensus structure. It is not clear whether these organisms are using the same machinery to transport RNA. Some evidence shows that RNA localisation plays an important role in axis formation during the development of the wasp *Nasonia*, and uses similar components of the machinery (Olesnický and Desplan, 2007). *eve* and *h* transcripts are localised in five dipteran species, including the *Anopheles gambiae*, at blastoderm stages in an Egl dependent pathway (Bullock et al., 2004). However, not all of these species' transcripts are localised when injected into the *D. melanogaster* blastoderm embryo. That might suggest a difference in the machinery that is responsible for localisation of mRNA within dipteran species, probably as a result of differences in the embryo cytoarchitectures (Bullock et al., 2004).

Sequence and structure comparison is an important tool to identify conserved features within sequences, and allows the prediction of elements that can be subsequently investigated by biochemical, genetics, or molecular means. The approach has its limitations. The folding of the signal is only inferred theoretically, by prediction of probable stable structure. Ideally, biological function of RNA molecules is best interpreted against a 3D structure. The solution is therefore to study and solve the 3D structure of the elements.

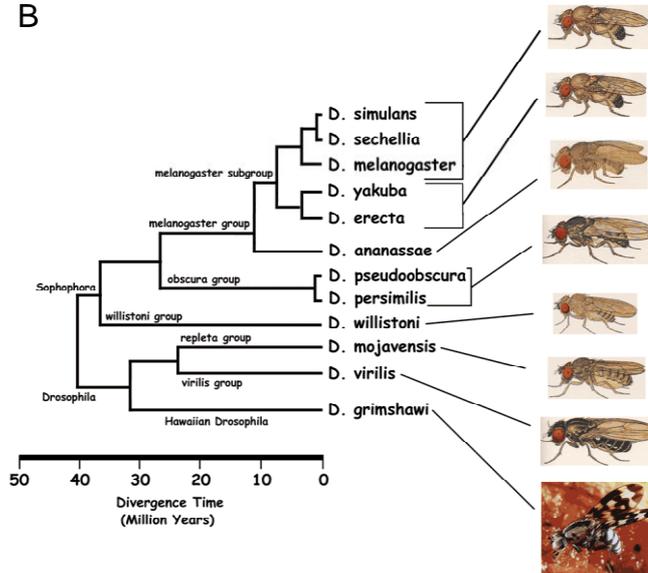
A

```

K10_Dmel      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAA-TTAAG 44
K10_Dsec      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAA-TTAAG 44
K10_Dsim      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAA-TTAAG 44
K10_Dyak      CTTGATTGTATTTTAAATTAATTCTTAAAACTACGAA-TTAAG 44
K10_Dere      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAA-TTAAG 44
K10_Dana      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAG-TTAAG 44
K10_Dpse      CTTGATTGTATTTTAAATTAATTCTTAAAACTACAAA-TTAAG 44
K10_Dmoj      CTTAATTGTATTTTAAATTAATTCTTAAAAATACAAA-TTAAG 44
K10_Dvir      CTTAATTGTATTTTAAATTAATTCTTAAAAATACAAA-TTAAG 44
K10_Dgri      CTTAATTGTATTTTAAATTAATTCTTAAAAATACAAAATTAAG 45
***          *****          *****          ***  ** *  *****

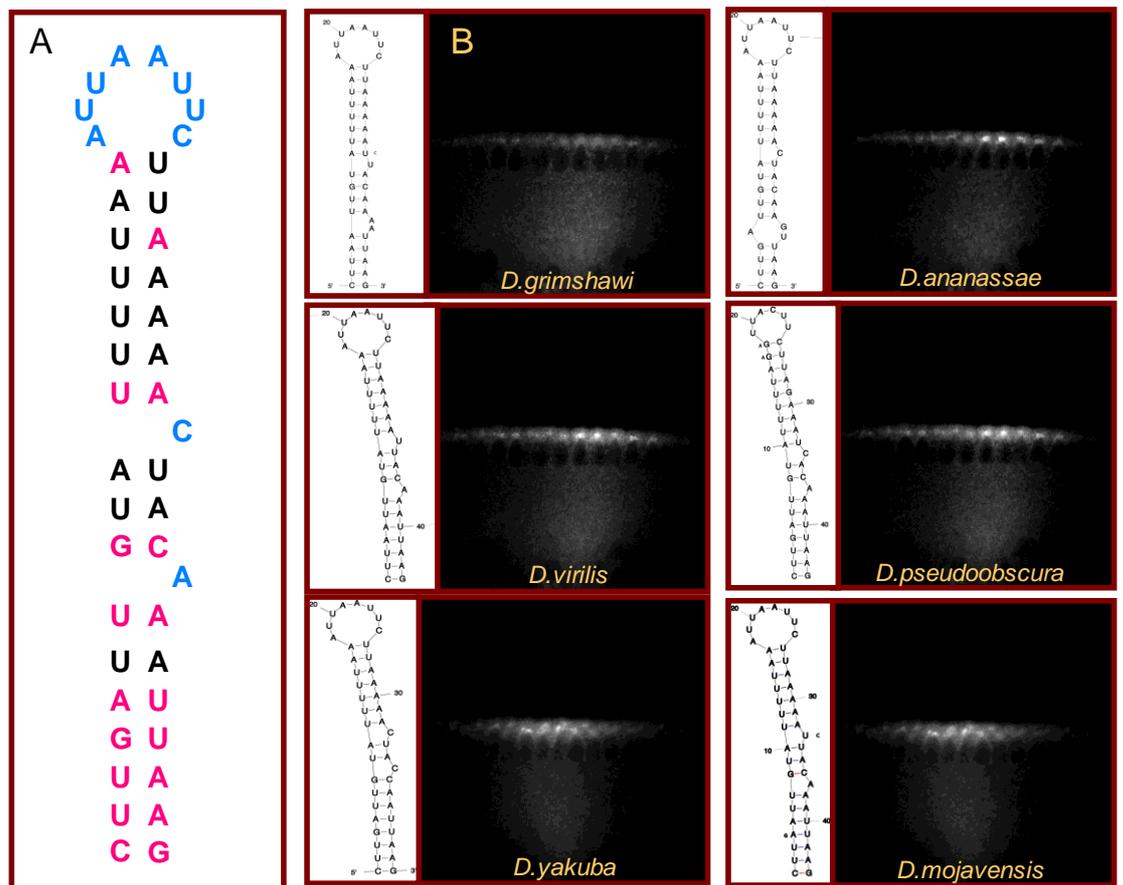
```

B



**Figure 3\_1 Predicted TLS secondary structure is conserved.**

(A) Alignment of *Drosophila K10* 3' UTR sequences with *D. melanogaster* TLS. Sequences are listed, top to bottom, in order of divergence from *D. melanogaster*. (B) Phylogenetic tree of the subset of the *Drosophila* with completely sequenced genomes. Taken from: [http://flybase.org/static\\_pages/species/muller\\_synteny.html](http://flybase.org/static_pages/species/muller_synteny.html)



**Figure 3\_2 Comparison of TLS localisation signals from different *Drosophila* species**

(A) Consensus secondary structure predicted by ALIFOLD. Invariant residues are black, consensus residues are grey and positions with conserved base-pairing are in green. (B) The localisation of TLS transcripts derived from different *Drosophila* species injected into the blastoderm embryo of *D. Melanogaster*. Localisation is weaker in *D.Grimshawi*. To the left of each image are the predicted secondary structures for Drosophila TLS sequences.

## 4 Chapter 4 Effect of mutations in the TLS on embryonic development

### 4.1 Aims

The aim of this chapter is to understand the importance of the TLS in localisation of *K10* transcript in the oocyte, as well as in the development of the oocyte and the embryo. I wanted to examine how different mutations that disrupt localisation might affect the development of the embryo. To evaluate this, I created “rescue” flies. These are flies that do not transcribe the *K10* gene, and were recombined with the full length of the *K10* gene, that contains changes to the TLS. These flies were analysed for their ability to produce viable eggs with a wild type phenotype.

### 4.2 Introduction

*fs(1)K10* belongs to the group of maternal genes that are important for establishing dorsal/ventral polarity of the developing oocyte and embryo. *K10* null female flies lay dorsalised eggs, which either do not hatch, or develop into a dorsalised embryo (Cheung et al., 1992; Forlani et al., 1993; Haenlin et al., 1987; Prost et al., 1988; Schupbach, 1987) (see introduction and figure 4\_1). Localisation of *K10* mRNA in the oocyte has been shown to be important for the function of *K10* protein and mislocalisation of the RNA has a similar phenotype to the null protein (Cheung et al., 1992).

Two phases of *K10* mRNA transport can be recognised during oogenesis. In the first phase, which starts at stage one of oogenesis, *K10* mRNA is synthesised in the nurse cells and immediately transported into the early oocyte. It accumulates in the posterior of the oocyte, until stage seven (See Introduction section 1.2 for description of oogenesis stages). Transport of *K10* mRNA from the nurse cells to the oocyte in early stages is important for the proper localisation and expression of the protein and for the development of the egg and hatching of the embryo. Failure to transport *K10* mRNA to the oocyte in this stage leads to a phenotype similar to that of *K10* null flies (Cheung et al., 1992; Serano and Cohen, 1995b).

The second phase of *K10* mRNA transport starts after the rearrangement of cytoskeletal fibres during stage seven. At this time, microtubule minus ends are

reorganised towards the anterior cortex, and, accordingly, *K10* mRNA is localised to the anterior end of the oocyte. During this phase, mis-localisation of *K10* RNA has no effect on the localisation of K10 protein to the oocyte nucleus or on the protein function in oogenesis (Serano and Cohen, 1995a).

Transport of maternal RNA from the nurse cells to the oocyte is analogous to the movement of injected transcripts in the syncytial blastoderm embryo in that similar proteins are engaged in the process and the same sequences of localisation signals are recognised (Bullock and Ish-Horowicz, 2001). In both processes, movement of *K10* RNA is towards the minus end of microtubules (Pokrywka and Stephenson, 1995; Theurkauf et al., 1992) and is mediated by a dynein motor, (Clark et al., 2007; Wilkie and Davis, 2001) and Egl and BicD as accessory proteins (Bullock and Ish-Horowicz, 2001). The prediction is that transcripts that do not localise when injected into the blastoderm should not be transported into the oocyte from the nurse cells.

Cohen et al. used transgenic flies with mutated *K10* TLS to study the importance of the TLS for the function of the K10 protein. They demonstrated the importance of the integrity of the TLS for localisation of *K10* mRNA during oogenesis by presenting a correlation between the severity of the mutation in the TLS and the phenotype of the embryo (Cohen et al., 2005).

Therefore, I wanted to see whether mutations in the TLS that affect localisation of injected transcripts in the embryo cytoplasm would also be defective during oogenesis, abolishing transport from the nurse cells to oocyte to perturb development. To pursue this, transgenic fly lines with mutations in the TLS signal were generated in flies homozygous for *K10*<sup>LM00</sup>, a null mutant that produces no detectable *K10* RNA (Cheung et al., 1992). Homozygous *K10*<sup>LM00</sup> females are sterile and lay dorsalised eggs that do not hatch. The transgenic females thereby produce the mutant *K10* RNAs in a null *K10* background so the extent of rescue produced by the various mutations may be assayed by determining hatch rate and the extent of dorsalisation of the embryos.

### **4.3 *K10* null flies do not produce offspring and have a dorsalised phenotype**

First, to make sure that the *K10*<sup>LM00</sup> flies produce a null phenotype, the number of embryos hatched from these flies was counted. As *K10* is a maternal RNA, and the phenotype appears in the oocyte, the fertility of the mothers was assayed by recording the percentage of hatched eggs, and analysing their shape. *K10*<sup>LM00</sup>/*K10*<sup>LM00</sup> mothers

produced no wild type embryos (Figure 4\_1 and table 4\_1), and none of the eggs hatched (N=24). However,  $K10^{LM00}/FM7$  produced wild type shaped eggs (Figure 4\_1). 67% (N=39) of these eggs hatched and none of them exhibited the null phenotype. Therefore, homozygous, but not heterozygous flies for  $K10^{LM00}$  produce offspring with a  $K10$  phenotype.

#### **4.4 Establishing a line of flies with non-localised transcripts**

To check the TLS transcript in its full-length context, I used as a source a genomic clone, BAC R48022, from which I extracted a 5.5 Kb sequence containing the entire gene, complete with the flanking genomic sequences of 537 nucleotides and 853 nucleotides at the 5' and 3' ends respectively (BAC5.5HindIII). Constructs, which contained the 3'UTR of  $K10$  with either wild-type or mutated signals (previously used for injection assays), were used to introduce different mutations in the TLS into the full length genomic sequence, using the "Recombineering" technique (Figure 4\_1).

The "Recombineering" (recombination-mediated genetic engineering) method is based on homologous recombination systems in *Escherichia coli* using recombination proteins from the  $\lambda$  bacteriophage. These systems mediate efficient homologous recombination between transfected linear DNA molecule (containing a mutant TLS) and a DNA molecule already present in the bacteria (in this case BAC5.5HindIII) via short sequences of mutual homology. The Recombineering technique is faster and more efficient than other cloning techniques and obviates the need for conveniently positioned restriction sites (Copeland et al., 2001; Court et al., 2002; Zhang et al., 1998b) (See material and methods).

Using this technique, four different mutations in the 3'UTR were generated and introduced into a vector containing the genomic  $K10$  gene:

- (i) Wild-type TLS (WT). The WT TLS was included as a control, to provide the same context as the rest of the transgenic constructs.
- (ii) U2G2, a mutation in the distal helix that changes two base pairs of UA to GC (Figure 5\_5). This mutation was chosen as it reduces localisation in transcripts injected into a blastoderm and represents mutations that affect the structure of the TLS (See Chapter 5).

- (iii) U6G a point mutation in position 6U to G (Figure 2\_6). This mutation was chosen as it abolishes localisation of injected transcripts and represents mutations with an effect on the sequence identity.
- (iv) Scmb, a scrambled sequence of *K10* TLS that was used as negative control (Figure 1\_7).

These constructs were sequenced and sub-cloned into a vector containing an *attB* site in order to use the *attB/attP* system of site-specific recombination to insert a transgene into fly lines with *attP* sites present at predetermined locations in the *Drosophila* genome. A range of *attP* sites are available at precisely mapped intergenic locations in the fly genome (Bischof et al., 2007). Germline expression of phiC31 integrase promotes recombination, leading to the integration of the exogenous gene of interest into the pole cells and thereby propagation to future generations. Each transgene is thus incorporated at exactly the same site in the genome, allowing them to be compared directly, without the need to compare independent isolates (Bischof and Basler, 2008; Bischof et al., 2007).

The constructs were injected into *attP* flies in which the vasa promoter drives recombinase expression in the germline. Previous reports had indicated that this technique was very efficient, giving transformants at a frequency of 16-55% (Bischof et al., 2007), however the injections of the *K10* transgene did not produce transformants very efficiently. Despite repeating the injections several times, there was a very low number of transformants amongst the progeny, of 0.75%-5%. The WT transgene injections produced one transformant fly (from 29 G1). The U6G transgene produced 2/37 transformants. The U2G2 transgene produced one transformant out of 133, and the scrambled sequence, did not produce any transformants (after four rounds of injections and 218 G1 survivors). The progenitors of the injected flies were crossed with the *K10<sup>LM00</sup>* flies, to produce a transgene fly in a null *K10* background (See table 4). The genotypes of all transgenic flies were verified using PCR and sequencing to ensure that both the correct exogenous sequence was integrated, and that the fly did not contain any endogenous *K10* sequence.

#### **4.5 Wild-type transgenes rescue the null *K10* phenotype**

The wild type transgene in the null background was used as a positive control for the transgene activity. The wild type transgene rescued the null mutation and produced normal progeny. 100% of the eggs had wild-type phenotype, and 94% of them hatched

(N=46). As expected, heterozygote wild type embryos also showed a rescue phenotype, with 100% wt phenotype for both lines and 93% or 75% hatching eggs (N=43, 44) for line 8.2 and line 8.3, respectively. Hence, one copy of the wild-type construct can rescue the null phenotype.

#### ***4.6 A transgene with the U6G mutation did not rescue K10 null phenotype***

A mutation exchanging uracil in position 6 to guanine (U6G) abolishes localisation of injected transcripts in embryos (Table 1\_1). Transgenes with one copy of mutation U6G produced a null phenotype in 37% and 20% of the eggs and had a very low percentage of hatching eggs (25% and 0%) in lines G6\_17 and G6\_8 respectively. Unexpectedly, two copies of the mutation U6G rescued the phenotype of the eggs, whereby 100% of them looked like wild type, although the percentage of hatching eggs was still considerably lower than wild type eggs (48% and 25%, N=31 and 59).

#### ***4.7 A transgene with a mutation in the distal helix rescued the K10 null phenotype***

A mutation in the proximal helix that exchanged base pairs in position 12, 14 and 29, 30 to G and C respectively (U2G2) reduced localisation of injected transcripts in blastoderm embryos and hence, it was not expected to rescue the K10 null phenotype. Nevertheless, in flies carrying a single copy of the transgene, 100% of the progeny had a wild-type phenotype, and produced normal eggs (N=42).

#### ***4.8 Summary and discussion***

In these experiments, the effect of mutations in the *K10* TLS on the activity of the *K10* gene and the viability of the embryo was examined. These experiments measured whether mutations that have an effect on transcripts movement in the blastoderm embryo, also have an effect on transcript localisation in the oocyte (of the transgenic flies). The difference between the transport in the oocyte and in the embryo lies in the timing and distances of the transport process. While injected transcripts travel a short distance from the basal to the apical edge of the cytoplasm, in around five minutes, the transport from nurse cells to oocyte is a much longer process and takes few hours. Hence, weak mutations in the transcripts can be compensated for, if there is enough time.

In this case, the effect of a mutation on the movement of injected transcripts would be much stronger than the effect of the mutation in the oocyte, and so even slower, or less effective transcripts, eventually manage to reach their destination within the oocyte. In previous work, mutated transcripts abolished localisation and altered the development of the embryo (Cohen et al., 2005). Weak mutations in *K10* TLS that do not localise well in the oocyte had a weak effect on the final development of the egg from the oocyte (Cohen et al., 2005).

Cloning the transgene included using the *attP* system that leads to integration at a specific site in the genome. Hence, the fact that there was not always more than one line for each transgene should not be a problem. Unlike conventional transformation techniques, there is less need to compare results of multiple lines of transgenic flies to avoid different effects due to site of insertion in the genome. Using the *attP/attB* system, every transgene is inserted into the same specific site in the genome. The lack of transformants may have reflected a problem with either the genotype of the recipient flies or the injected DNA. To test these possibilities, I injected a shorter sequence, unrelated to *K10* gene, in the same *attB* vector, into the *yw, attP/attP* (3R 86F) background. This injection worked much better with a higher number of transformants (10%) amongst the progeny. Hence, it is most likely that an element in the *K10* genomic sequence attenuates the transgenic process. Nevertheless, a single copy of a wild type *K10* transgene was able to fully rescue the activity of the gene in a *K10* null background, attesting that the transgenic procedure and the rescue experiment have worked.

The U6G mutation abolishes localisation of transcripts injected into syncytial embryos (Figure 2\_6). Indeed, the U6G variant did not fully rescue the null phenotype. Mothers carrying one copy of the transgene produce only a small percentage of hatched eggs. The observation that WT TLS, but not TLS with U6G mutation, fully rescued the null flies confirms that the ability of the transcript to localise, when injected into a blastoderm embryo, is associated with its ability to be transported from nurse cells to oocyte as well as its correct expression and proper development of the cell in which it localised.

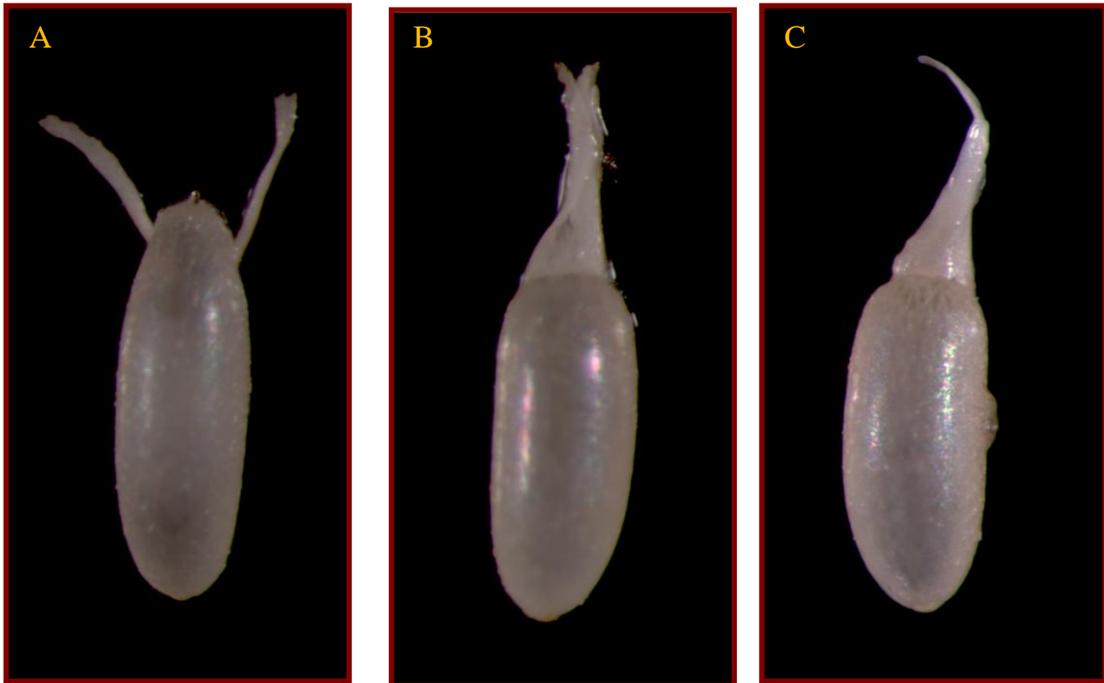
However, this variant did have a partial rescue of the *K10* null phenotype, as some of the unhatched eggs showed no evidence of dorsalisation. Further rescue towards wild type is observed from mothers carrying two copies of the transgene, with all of the embryos exhibiting a normal shape, although many still do not hatch. The reason for the

partial activity of the U6G mutation may be due to the nature of the binding of the transport machinery to the signal itself. The mutation most likely affects the structure of the signal, but the adaptor protein/s may still bind weakly or temporarily to the signal. This interaction is not enough to retrieve complete activity when there is only one copy of the gene, but once there are two copies of the gene and more transcripts are produced, there is a higher chance for these transcripts to be bound and transported by the machinery. My observation that un-localised *K10* transcripts cannot fully rescue the *K10* null phenotype supports previous results that show that localisation of *K10* RNA plays an important role in localisation of K10 protein and for the development of the fly embryo (Forlani et al., 1993; Serano et al., 1995).

The full rescue phenotype of the U2G2 variants was also surprising, as mutation U2G2 interrupts localisation of injected transcripts in blastoderm embryos, hence was expected to produce transgenic flies with mutant phenotype. One explanation for the rescuing effect of the U2G2 mutants would be that although this mutation affected the transport of injected transcripts in the embryo, it is not severe enough to affect transport of transcripts from the nurse cells to the oocyte.

Another possibility is that the differences in rescuing abilities might derive from differences in the identity or available concentration of the recognition factors. Injection of transcripts into the embryo supplies an excess of RNA to the cytoplasm, which decreases the ratio of RNA-protein and might reduce the availability of the localisation factors for the injected transcripts. Hence, overall localisation of injected transcripts would be reduced compared to endogenous transcripts in the oocyte.

Had time permitted, the best way to study the reason for the differences between the mutants would have been to examine the mRNA localisation of the transgenes by *in situ* hybridisation. This would have given a better indication as to whether the mRNA was localised, and at which stages of development. I made preliminary attempts at *in situ* hybridisation to detect transcripts made from the transgenes in ovaries, however the probe I used showed very high background, impeding analysis.



**Figure 4\_1 Eggshell produced by wild-type and K10<sup>LMOO</sup> females.**

(A) Wild type. (B,C) Strong dorsalis ed eggs produced by K10<sup>LMOO</sup> mother. Most of the eggs have ventrally fused dorsal appendage. The appendages are enlarged and completely encircle the anterior end of the egg. Less than 1% contain embryos. Anterior is up.

	Wild type phenotype	Percentage of wt	Null phenotype	Percentage of null phenotype	Number of hatched eggs	Percentage of hatched eggs	N
LM <sup>oo</sup> /LM <sup>oo</sup>	0	0	24	100	0	0	24
LM <sup>oo</sup> /FM7	39	100	0	0	26	66.67	39
WT3/WT3	46	100	0	0	43	93.48	46
WT2/TM3	43	100	0	0	40	93.02	43
WT3/TM3	44	100	0	0	33	75	44
G6_17/G6	31	100	0	0	15	48.39	31
G6_17/TM3	31	63.26	18	36.73	12	24.49	49
G6_8/G6	59	100	0		15	25.42	59
G6_8/TM3	21	80.77	5	19.23	0	0	26
U2G2/TM3	42	100	0	0	38	90.48	42

**Table 4\_1 Analysis of embryo development for the progeny of the rescue flies.**

LM<sup>oo</sup> /LM<sup>oo</sup> are flies with null K10. WT3/WT3 are flies with rescue constructs containing wild-type TLS on null K10 background. G6\_17/G6 and G6\_8/G6 are flies with rescue construct that contain mutation in position 6 in the K10 TLS. U2G2 are flies with rescue construct that contain mutation in the upper helix of the K10 TLS

## 5 Chapter 5 Solution structure of K10 localisation signal

### 5.1 Aims

This chapter addresses how structural studies have contributed to our knowledge about the *K10* TLS. This chapter mainly reports the work of Peter Lukavsky using NMR to solve the solution structure of the wild-type *K10* TLS and three TLS mutants. It correlates the structural data, specifically an unusual helix described in these results, to the localisation activity of the TLS.

### 5.2 Introduction

Several tools can be utilised to study the structure of RNA localisation signals, including RNA engineering, X-ray crystallography, single-particle cryo-electron microscopy, nuclear magnetic resonance spectroscopy, structure-specific chemicals and enzymatic probes, thermal denaturation, and mass spectrometry (Felden, 2007).

RNA engineering (mutational analysis) consists of introducing mutations into the sequence of areas expected to be important for the function. By employing different assays, the effect of these mutations on the ability of the transcript to localise can then be studied. Another approach is to use different enzymatic and chemical probes which bind preferably to paired or non-paired nucleotides (Brunel and Romby, 2000). Examples include the studies on the *c-myc* (Chabanon et al., 2005) and *bicoid* localisation signals (Brunel and Ehresmann, 2004; Wagner et al., 2004). These studies are complemented by the use of bioinformatics tools able to make structure predictions based on sequence conservation and minimal energy calculations (Zuker, 2003).

As work on the structure of the *K10* TLS progressed, it became apparent that the tools I had used so far were not sufficient to fully understand its structure. These tools accurately predicted the base-pairing and the interactions between nucleotides, providing a faithful two-dimensional description of the TLS. However, they were not able to fully explain the specificity of the TLS and which elements in the signal distinguish the TLS from any non-localising transcript. For example, the bulges seemed to have structural importance, but it was not clear how exactly their structure influences the overall signal and what role they play in the recognition by the motor machinery.

To address these questions, I tried to study the structure of the TLS by using X-

ray crystallography in collaboration with Stephen Neidle's lab (Biomolecular Structure Unit, School of Pharmacy, London). To obtain milligram quantities of clean RNA, we used *in vitro* transcription and gel purification. To ensure conformational homogeneity we denatured RNA samples and then renatured them with urea and refolded the sample by gradual dialysis into buffer lacking urea. The crystals were grown during a slow and controlled precipitation from the aqueous solution under non-denaturing conditions using the hanging-drop vapor diffusion technique (Supplementary figure 5\_1). This method is based on applying different concentrations of sample onto an inverted glass cover slide and allowing it to equilibrate against a much larger volume of reservoir solution through diffusion of the vapor inside a sealed chamber (Supplementary figure 5\_1) (Ke and Doudna, 2004).

Many factors affect the inherent ability of an RNA sample to form crystals, including purity, conformational homogeneity, molecular surface area, and available sites for intermolecular contacts, and structural dynamics (Ke and Doudna, 2004). Since the settings in which RNA crystals form are varied, we generated matrix sets of different conditions in order to find the optimal ones for the TLS RNA crystallization. Parameters that were used included the identity and concentration of divalent metal ions in the resuspension buffer, the concentration of the transcripts in the solution, the temperature of the drops, and the mother liquid. The times that take for crystal to form can last between 24 hours to nine months. Twenty trays, each contains 36 wells of different crystallization parameters were used, but none of them produced any crystals.

I then started a collaboration with Dr Peter Lukavsky (MRC Laboratory of Molecular Biology, Cambridge), to determine the three-dimensional structure of the signal by Nuclear Magnetic Resonance spectroscopy (NMR). High-resolution studies of RNA motifs by NMR have provided many insights into the unique structural characteristics of different RNAs (Doudna, 2000; Tzakos et al., 2006) and are especially useful in describing short oligonucleotide sequences (Felden, 2007). Applying NMR techniques to determine the *K10* RNA structure was not trivial: the structure is relatively long, and there are many repetitive base-pairs (for example the distal helix contains five consecutive U-A base pairs). This can lead to a substantial resonance overlap (Allain and Varani, 1997). The three-dimensional structure of RNA molecules larger than 30 nucleotides is normally analysed by "breaking" the whole structure into smaller parts. However, solving the TLS structure as a whole sequence would provide a more accurate description of its important elements.

Despite this high redundancy in sequence and secondary structure, Dr Lukavsky

solved the structure to a very high resolution, with a defined final ensemble structure (r.m.s. deviation of 1.15 Å) and high number (115) of angular restraints derived from residual dipolar couplings (Lukavsky and Puglisi, 2005) (See attached DVD).

The structural solution shows that the basic base-pairing is similar to that predicted by mFOLD, whereby the TLS forms a stem-loop that contains the lower and upper helices, and two single nucleotide bulges (Figure 5\_2A). Nevertheless, it has provided a comprehensive description of the three-dimensional structure of the helices and the bulges that is significant to our further understanding of the localising elements. The stem helix is distorted, partly due to the bulge bases, and also by the unusual conformation of the upper and lower helices resulting from a stack of Purine bases in the lower and the upper helices.

### ***5.3 Each bulge adopts a different conformation relative to the helix***

The solution structure reveals that the unpaired bases adopt different conformations relative to the helix (Figure 5\_1A). They also distort the plane of the base-pairing interaction of their adjacent nucleotides. The upper bulge, a cytosine in position 33 (C33) protrudes from the helix, such that the base moiety lies in the major groove (Figure 5\_1B). In contrast, the lower bulge, an adenine in position 37, is stacked between the neighboring base pairs. In this way, it distorts and opens the helix, and has more impact on the nearby nucleotides.

As mentioned before (chapter 2), individual deletion of the bulges reduced the efficiency of apical K10 localisation and deletion of both C33 and A37 rendered the signal inactive. Moreover, any nucleotide at both position 33 and 37 supported full apical K10 transport. Therefore, it is unlikely that specific functional groups introduced by the bulged nucleotide bases are directly recognised by the transport machinery. Instead, the presence of nucleotides in these positions is presumably important because it determines the relative orientation of the widened major grooves in the lower and upper helix; the widest openings of the major grooves are orientated along the helix at an angle of ~ 90 degrees relative to each other in the wt TLS. The effect of mutations in the upper bulge was not as pronounced as the effect of mutations in the lower one, probably as a result of the different conformation each adopts relative to the helix, as A37 is stacked inside the helix and alters its neighboring base pair alignment (see chapter 2 and table 2\_1).

## **5.4 An unusual B-like form helix plays a role in localisation**

One novel feature revealed by the NMR solution structure of the *K10* TLS is a widened major groove at both the upper and lower helical stems, often observed in regions of non-Watson-Crick base pairs. The helices formed differ from the A-form helix of a usual double stranded RNA in the orientation of the base pairs and the structure of the major groove. Wild type RNA helices form positive inclination angle of the base-pairs, relative to the helical axis, and the major grooves are deep and narrow (Figure 5\_2). The unusual structure resembles that of B-like form helix, which is typical of DNA, where the base pairs align perpendicular along the axis. The inclination angle is close to zero and the major groove is wide, shallow, and generally more accessible for ligand interaction than the major groove of A-form RNA helix.

A probable reason for the unusual structure is the interactions between the stack of seven purines (adenines 28-32) on one side of the strand, and a cross-strand stacking between bases of A16 and A28 that positions each of the adenine H-2 protons above the other base moiety (Figure 5\_3). Similarly, in the lower helix, a continuous stack of five purine bases on one side of the helix, also cause an unwinding of the groove. The four U-A base-pairs create the usual Watson-Crick contacts, but the G-U base pairing is distorted and does not adopt the wobble conformation usually seen in conventional helical regions (Figure 5\_3).

In previous work, it was shown that disrupting the upper helix by changing the 3' strand from 4 adenines to 4 uracils to completely break the base-pairing of the helix, also abolishes localisation (Bullock and Ish-Horowicz, 2001; Serano and Cohen, 1995b). Altering the identity of the nucleotides while keeping the helix base-pairing, by converting the 5 U-A base pairs with A-U base-pairs, reduced but did not stop localisation (Bullock and Ish-Horowicz, 2001). Hence, the helix is important for localisation.

To test whether there is any correlation between the presence of the B-like form helix and the activity of the signal, I designed specific mutations that retain base pairing and so are prone to create a helix, but which might disrupt the B-like form of the structure. I then tested their mobility by injecting them into embryos. Two mutations that disrupt one base pair in the middle of the helix (13U-30A) and change it to A-U or G-C, (13\_30AU and 13\_30GC, respectively) localised well, similar to a wild type localisation (Figure 5\_4). However, a mutation that changed two base-pairs in the upper

helix, 12U-31A, and 14U-29A to G-C or C-G (U2G2 and U2C2, respectively), reduced localisation activity (Figure 5\_4).

A mutated transcript that localised well, *13\_30GC*, and a second one with altered localisation activity, U2G2, were processed for NMR structure studies by Dr Peter Lukavsky, in order to try to see if there was a difference between their structures and a wild-type structure. The NMR results showed that despite the change to G-C, the unique B-like helix form was still part of the structure. In contrast, the structure of U2G2 revealed that, although the transcript retains an upper helix, it now formed a typical RNA A-helix instead of B-like form. These two findings demonstrate a correlation between the localisation activity and the form of the helix.

Interestingly a one base-pair mutation in the *orb* helix, that changed base pair 12U-30A to A-U (12\_30AU) also causes a reduction in transcript localisation, as does changing the second base pair from UA to CG (Figure 5\_4). Hence, both upper stem and lower stem helices contribute to the localisation activity of the *K10* signal.

## **5.5 Summary and discussion**

The NMR structural determination of the TLS provides a three dimensional description of the structure of the helix. It explains the specificity of the structure of the proximal bulge that is situated inside the helix and so distorts its neighbouring nucleotides and provides an explanation why this specific bulge is conserved, as well as the nucleotides around. The NMR also reveals an unusual helix, not previously documented in a natural RNA.

Like DNA, RNA folds into helical structures. Examples can be found in tRNA (Blanquet et al., 2000), small RNAs (Rana, 2007), and mRNA (Svoboda and Di Cara, 2006). The RNA double helix differs from that of the DNA double helix by the presence of ribose, rather than deoxyribose, in the sugar phosphate backbone of the molecule. The addition of a hydroxyl group at the C2 position in the ribose sugar is responsible for the A-form geometry in double stranded RNA (Figure 5\_2). Because the RNA A-form double helix contains a major groove that is too narrow and deep for proteins to access, usually the minor groove becomes more important for protein interactions with RNA helices. Also, many proteins that interact with specific RNA sequences commonly bind single-stranded RNA segments (Antson, 2000).

The formation of an unusual helix is very surprising, since both double helical regions are composed of Watson-Crick base pairs and a G-U base pair that should maintain regular A-form geometry within an RNA helix. The reason for the specific B-

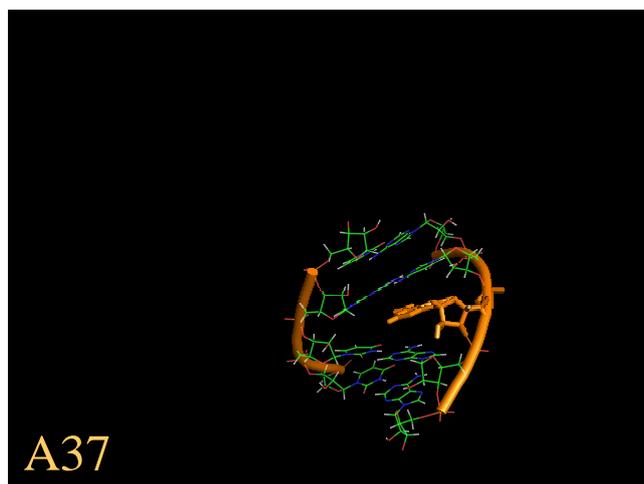
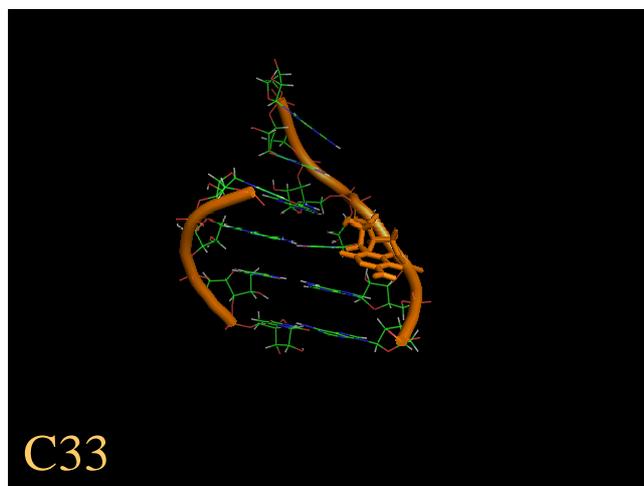
form fold is a continuous stacking interaction of purine bases, in each one of the stems, which unwind the helix to form a wider major groove, and reduce the inclination angles such that they are more reminiscent of those in B form DNA than A-form RNA.

The upper helix, contains an uninterrupted stack of five adenines and a cross-strand interaction between A16 and A28, which positions the adenine H-2 protons of A16 above the base moiety of A28 and vice versa. This creates a continuous chain of seven adenines. In the lower helix, there is also a stack of purine bases on one side of the helix, and a distortion of the G-U base pairing. As a result, the G-U base pair does not adopt a wobble conformation with two imino-carbonyl hydrogen bonds as usually seen in conventional RNA helical regions.

This B-like helix appears to be important for the function of the signal, as mutations that retain the helical structure but change its form to an A-helix reduced localisation. B-like form helix is not, by itself, sufficient for localisation of the transcript: In previous work of Cohen et al., mutants bearing only the distal helix with a deletion of the proximal region reduced localisation of lacZ reporter transcripts but did not abolish it completely (Cohen et al., 2005).

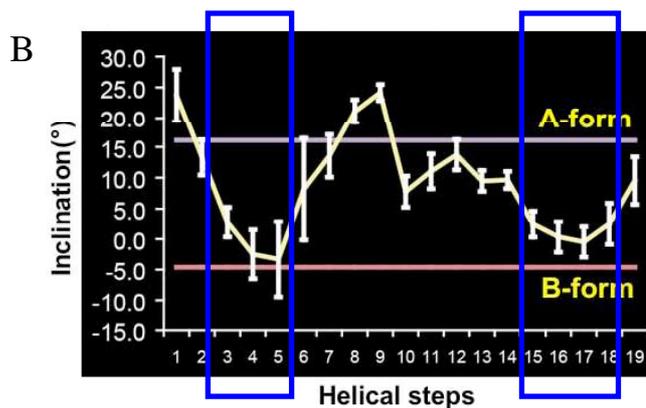
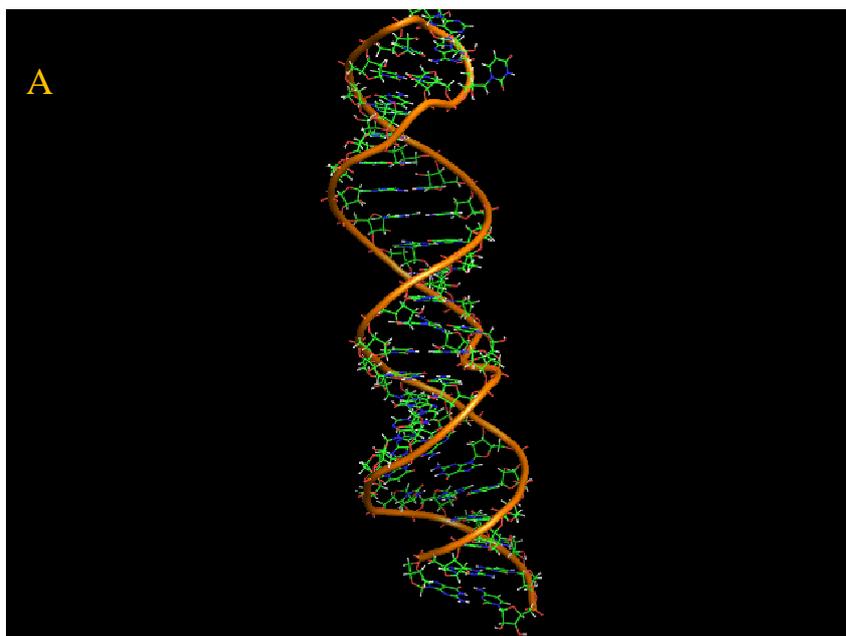
Nevertheless, the position of the two unusual stems that create specific alignment of the major groove is important and correlate with the efficiency of transport. Further experiments made in Simon Bullock's lab (MRC Laboratory of Molecular Biology, Cambridge) demonstrate the role of the lower stem in localisation of the transcript. *A-low-2GC-up*, a transcript with a combination of the UGUG mutation (a A-helix) that distorts the upper helix, and another mutation sourced from BC1 RNA (a known A-helix) which distorts the lower helix, formed a stem loop that did not localise (Suppl. Figure 5\_2). We envisage that UGUG partially decreases access for protein interactions within the upper helix, but that this only becomes limiting for signal activity when the affinity for the lower helix is also reduced, as in *A-low-2GC-up*.

These results suggest a model in which the two widened major grooves are major factors in increasing the affinity of the signal for the recognition machinery, with sequence specific recognition playing little part. The groove widths in the TLS are sufficiently increased to readily accommodate positively charged protein loops or beta-hairpins from proteins that link the TLS to the dynein motor complexes.



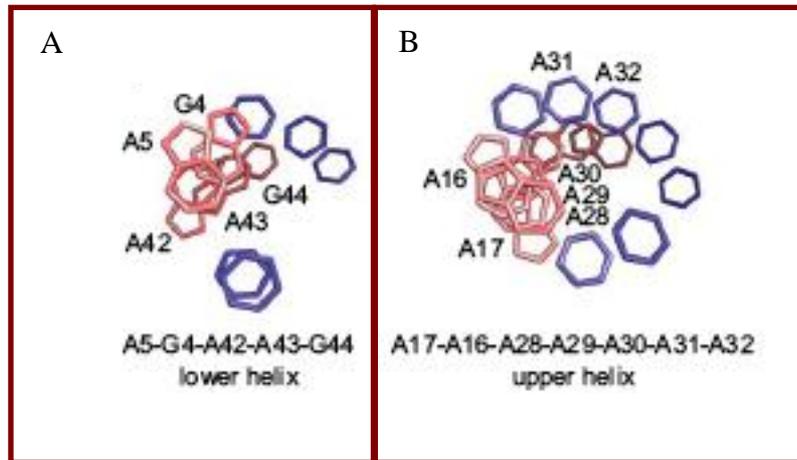
**Figure 5\_1 NMR solution structure of the bulges of TLS.**

The bulges adopt different conformations relative to the stem axis. Cytosin 33 (C33), the upper bulge, resides in the major groove and protrudes from the stem structure, maintaining the helical twist between the adjacent base pairs, while adenine 37 (A37), the lower bulge, is stacked in between the middle and the lower helix and increases the helical twist and change the orientation of the neighbouring bases. The bulges and the helix backbone are indicated in orange.



**Figure 5\_2 NMR solution shows an unusual B-like double helix.**

(A) The nucleotides are arranged perpendicular to the helix axis, which is usually typical to DNA helix. The major groove is much wider than the normal A form RNA double helix. (B) Representative graph that shows inclinations of base pairs along the helix. Lower inclination entails a B-form helix. The X axis show the base pairs of the stem loop, Base pairs 3-5 and 15-18 (inside the blue frames) create the lower and upper stem respectively.



**Figure 5\_3** View from above towards the lower (A) and upper (B) helix of *K10-wt* RNA

The lower helix displays a stack of five purine bases: A5, G4, A42, A43 and A44. The upper helix contains seven adenine bases: A17, A16, A28, A29, A30, A31 and A32. These continuous base-base stacking is giving rise to widened major grooves in a B-form-like inclination angles. Pyrimidine bases are blue and purine bases are pink; ribose-phosphate backbone and chemical groups on the bases are omitted for clarity. **This figure was made by PJ Lukavsky.**



## 6 Chapter 6 Using live imaging to study the dynamics of different K10 transcripts.

### 6.1 Aims

The aim of this chapter was to study the kinetics of injected transcripts in the cytoplasm of live blastoderm embryos. Tracking of particles that contain localised or non-localised transcripts, allows comparison of their dynamics, in principle providing a better evaluation of the effect of weak and strong mutations on localisation, and a more detailed understanding of the localisation machinery that regulates the dynein mediated, minus end MT-directed RNA transport.

Assaying transcript localisation in fixed embryos is fast and efficient. However, it describes the end result of a dynamic process, and it does not represent the process of localisation itself, for example the kinetics of the particles, the direction of the movement before the transcripts are anchored and the subtle differences between weaker mutations that affect localisation. In order to understand in more detail the localisation of the TLS, I used *in vivo* time-lapse microscopy to characterise the movement of wild type and weakly localising mutant transcripts in the *Drosophila* blastoderm. The dynamics of *K10* TLS localisation was compared with localisation of mutant TLS transcripts, and a vector control.

Injected wild-type *K10* TLS transcripts are transported as particles to the apical edge of the embryo, similar to the localisation of other transcripts such as *ftz*, *runt*, *wg* and *h* (Bullock et al., 2003; Wilkie and Davis, 2001) and of other cargos that are recognised by the dynein motor system (Gross et al., 2000; Shah et al., 2000; Welte et al., 1998). The similarity is manifested not only in the particle formation but also by the potential bidirectional movement of the motor complex.

Particle formation starts immediately after injection. Their appearance is evident at the beginning of the films (i.e. within 45 sec), although not all particles start to move immediately. It is likely that there are three separate events; first, the assembly of many individual RNAs into a particle, then the activation of movement of the complex, probably as a result of binding to the motor proteins and to the microtubules and, finally, the docking of particles once they reach the apical side of the cytoplasm (Bullock et al., 2003).

By tracking individual particles, it is possible to follow the detailed kinetics of localisation of specific injected transcripts. Embryos were visualised with an UltraVIEW LCI confocal scanner using an Ultrapix (Perkin Elmer) camera at a lapse interval of 1.18s. Image acquisition from a chosen focal plane started around 60s after injection.

Each frame was captured as a Tiff image, and a Mathematica program was used to identify individual particles in each individual image, which were then verified by eye to ensure correct identification of the particle on the bright background. The programme then determined a movement of the particle between consecutive frames. Active transport was defined by at least three successive displacements in a persistent direction. Tracking of a particle began when it first displayed these characteristics of active transport.

The resulting pattern (Figure 6\_1) can be used to calculate different parameters of the kinetics of the particles. Mutant transcripts 6C (mutation U6 to C) and 8A (mutation 8G to A) and *pBluescript* vector were injected and their motion compared with that of wt transcripts. These mutants showed an abolishment of localisation when injected into blastoderm and analysed in fixed embryo (chapter 2, and table 2\_1) and were chosen in order to see how differences in their kinetics compared with those of wt transcripts and transcripts containing no localisation signal.

Movement of particles can be divided to three different types: apically directed, basally directed and paused movement. There is also movement in the z-axis, i.e. through the plane of focus, but technical limitations make it hard to follow these latter particles. Indeed some particles “disappear” after a few runs, and they were not taken into consideration when tracking the movement.

It is rare for *K10* wild-type particles to move apically in a single persistent motion. Usually, they move intermittently, alternating between apically directed motion (on average, 30% of the time) (Figure 6\_2B) and paused states (50% of the time) (Figure 6\_2C). There is little movement in a basal direction (20% of the time) (Figure 6\_2D). Mutant 6C moved apically 20% of the time (Figure 6\_2B), basally 10% of the time (Figure 6\_2C) and was paused for 70% of the time (Figure 6\_2D). Therefore, the forces that are active in the transport represent a bi-directional motor system, in which transcripts move to both directions and the net transport is unidirectional and is probably determined by the ability of the motor complex to attach to the cargo. Weak attachment between the motor protein and the RNA cargo can be a result of mutations that alter the structure and/or change the specific binding site.

Comparison of the kinetics of wild-type *K10* localisation with kinetics of weaker localisers reveals that they differ not only in their directionality and by the fact that they stay on the complex for a longer time, but also by their speed. Localisation of wild type *K10* TLS takes around three minutes (see attached Movies 6\_1). The mean velocity of transport was  $76\pm 6$  nm/s with a maximum speed of 118.4 nm/s (see Table 6\_1). Mutant transcripts 8A and 6C were also transported as particles in a predominantly apical direction, but completion of localisation took longer than wild type particles, as they paused more often (see Movie 6\_2 on attached DVD). While wild-type containing particles moved at a speed of  $76\pm 6$  nm/s, those containing 6C and 8A had a speed of  $22\pm 3$  and  $13\pm 3$  nm/s, respectively (see Table 6\_1).

Once transcripts have reached the apical site, they are anchored there in a dynein-dependent way (Delanoue and Davis, 2005). It is indeed apparent in the movies that wild-type, as well as mutant transcripts, ceased their movement once at their apical destination. Mutant particles appeared to be retained normally at their apical destination and weak localisers still anchor apically (movie 6\_1). A possible explanation for the fact that even non-localising transcripts can be anchored, is that the contact of the anchorage system, which consists of MT and dynein but not BicD and Egl, to the RNA, does not require a strong binding at the recognition site. Hence even weak localisers are being recognised, once they reached the apical site.

## **6.2 Comparison of kinetics of different mutated *K10* TLS transcripts with altered helix forms**

The structures of the different mutated versions of the TLS, as solved by NMR, showed that different mutations in the distal helix could adopt different conformations. Changing the base pair in position 13 and 30 from A-U to G-C (*13\_30GC*) did not change the helical conformation, and maintained the specific B-form like helix, that is seen in the wild type TLS. Four mutations in the flanking base pairs 12-31 and 14-29 from A-U to G-C (*UGUG*) lead to higher inclination angles and reduced major groove widths compared with wild-type, albeit still distinguishable from regular DNA A-form geometry (see chapter 5). The *UGUG* mutant had reduced localisation activity in fixed embryos, but *13\_30GC* mutant did not seem to have such an effect. To study the effect of these mutations on the kinetics of the particles, and to see whether the changes to the helix caused subtle differences in transport, the mutant transcripts were injected into embryos, and their movement was analysed.

The time-lapse movies revealed that the mutations to the distal stem affected the speed of apical accumulation of the *UGUG* transcripts and had a small, statistically insignificant effect on the *13\_30GC* transcripts. This effect was pronounced in the speed and directionality of the mutant. The rate of apical accumulation of wild-type is  $76 \pm 6$  nm/s, compared with a speed of 66 nm/s for particles containing mutant transcript *13\_30GC*, and inefficient speed of 39 nm/s for particles containing mutant transcript *UGUG*. A randomised transcript (Scrambled; Sc) based on the *K10* TLS sequence was used as a comparison. The mutation *UGUG* reduces localisation but its kinetics was more efficient than the Sc negative control (0.026 nm/s). The mutants also affected other factors, for example, the average distance a particle moved in one lag of apical direction was reduced from 0.76 nm in wild-type to 0.618nm in *UGUG*, and the average time on a direct track distance was reduced from 0.23 s to 0.18 s, respectively. These data demonstrate that a shift from the B towards the A helix form impairs apical transport.

### **6.3 Summary and discussion**

Kinetics of different injected transcripts in the cytoplasm of syncytial blastoderm embryos were studied using semi-automatic tracking that outlined the course of the run and the speed. The results show that different mutations in the localisation signal have an effect on the transport. Strong mutations like 6C and 8A reduced the speed of transport and its directionality. A similar, but less pronounced effect was found for *UGUG* mutant transcripts and the weaker mutant *13\_30GC*.

The differences in kinetics of the particles correlate with how much the mutation distorts the structure of the signal. Mutations like *UGUG* distort the upper helix of the TLS signal. Injected transcripts with this mutation were slower than injected transcripts with mutation *13\_30GC*, which does not change the specific B-like form of the upper helix.

The trend in which mutant transcripts exhibited slower and less directional movement than wild-type, or weak mutations was persistent. However, not all results had statistically significant differences (see table 6\_1). In some cases, a higher number of particles would improve the statistical analysis and differences might become significant. The reason for analysing relatively few particles was mainly that difficulties in setting the microscope up prevented making of more movies. In addition, in order to avoid false positive read-out, the sensitivity of the tracking program was reduced so that only bright particles that are clearly different from the background would be taken into

consideration. This significantly reduced the amount of particles that were tracked and analysed. Hence, in order to gain a more accurate account on the particles kinetics, more movies had to be made. Nevertheless, studying the properties of the overall transport, rather than comparison of localised or non-localised transcripts in fixed embryos, revealed fine differences among transcripts, and allowed a better understanding of subtle mutations that affect localisation.

There could be several reasons for poor transport of specific mutant transcripts. The main reason is probably disruption of the binding of the transcript particle to adaptor or to motor protein and impaired formation of an active transport complex. Although non-localised transcripts, like the Scrambled signal, form particles immediately when injected into the embryo, these particles are mainly stationary and do not start a directional transport; hence, they do not create an active transport complex.

It is also possible that these mutations interrupt the binding of the transport complex to the microtubule fibres. A way to distinguish between these two options would be to follow particles with labelled proteins (For example Egl and BicD which are known to be components of the transport particles (Bullock and Ish-Horowicz, 2001)) and labelled microtubules, (which would probably require a better image resolution). It might be also interesting to look at *in vitro* interactions between the isolated components, and compare interactions of wild type and mutant RNAs to microtubules (in the presence of the other components of the complex).

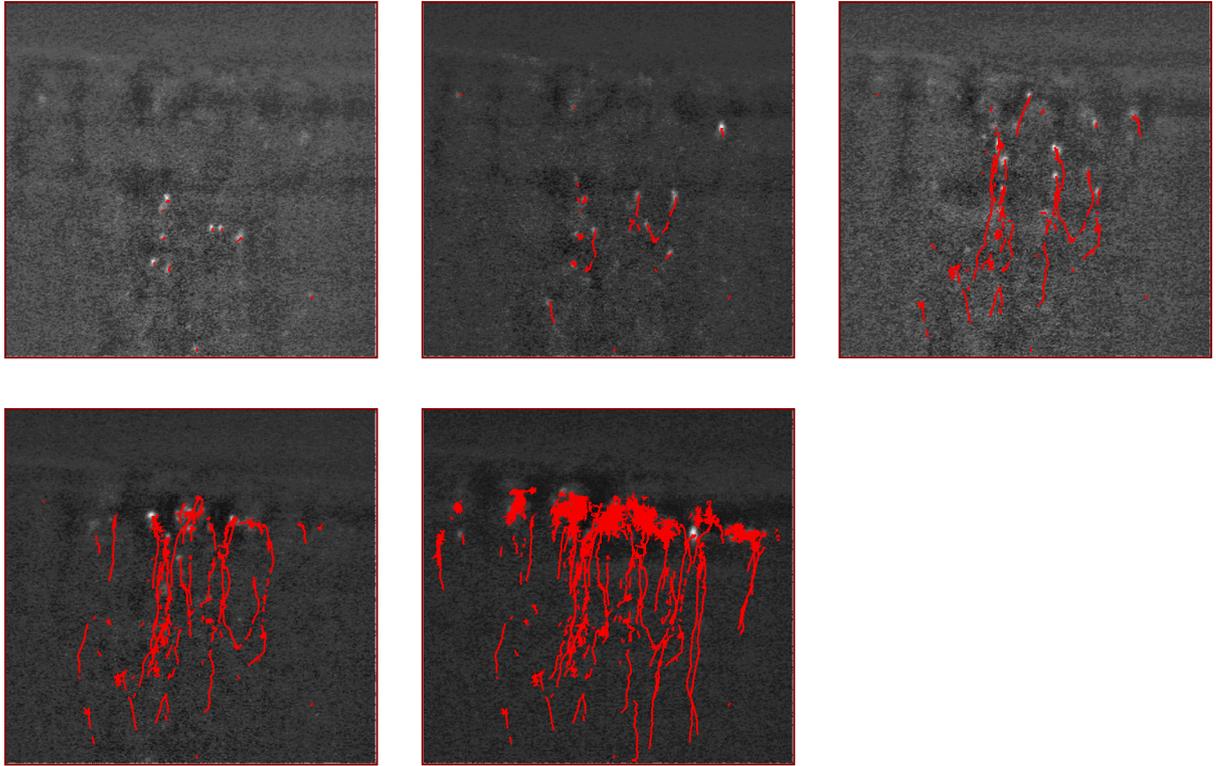
Another possible reason for poor localisation is that, after assembly of the motor complex, mutated transcripts can have a problem in sustaining the transport for a long distance. This is evident when comparing the processivity of the different transcript movements. Wild-type transcripts move more time in a directional mode towards the apical end, and have fewer pauses, compared with mutant transcripts (Figure 6\_2)

Transport might also be inefficient because of the effect of the cargo on the motor activity, so that weak localisers reduce the speed of the motors. Indeed the speed of mutants is reduced, probably due to the fact that there are more pauses and oscillatory movements that reduce the overall speed. However, apparent speed depends on imaging frequency. Faster imaging might reveal that apparently slower speeds are actually due to more pauses.

Finally, insufficient anchorage of the transcripts at the apical end might contribute to poor localisation. This is possible but I could not find any evidence for basal motion of apically localised particles. Once mutated transcripts arrive at the apical

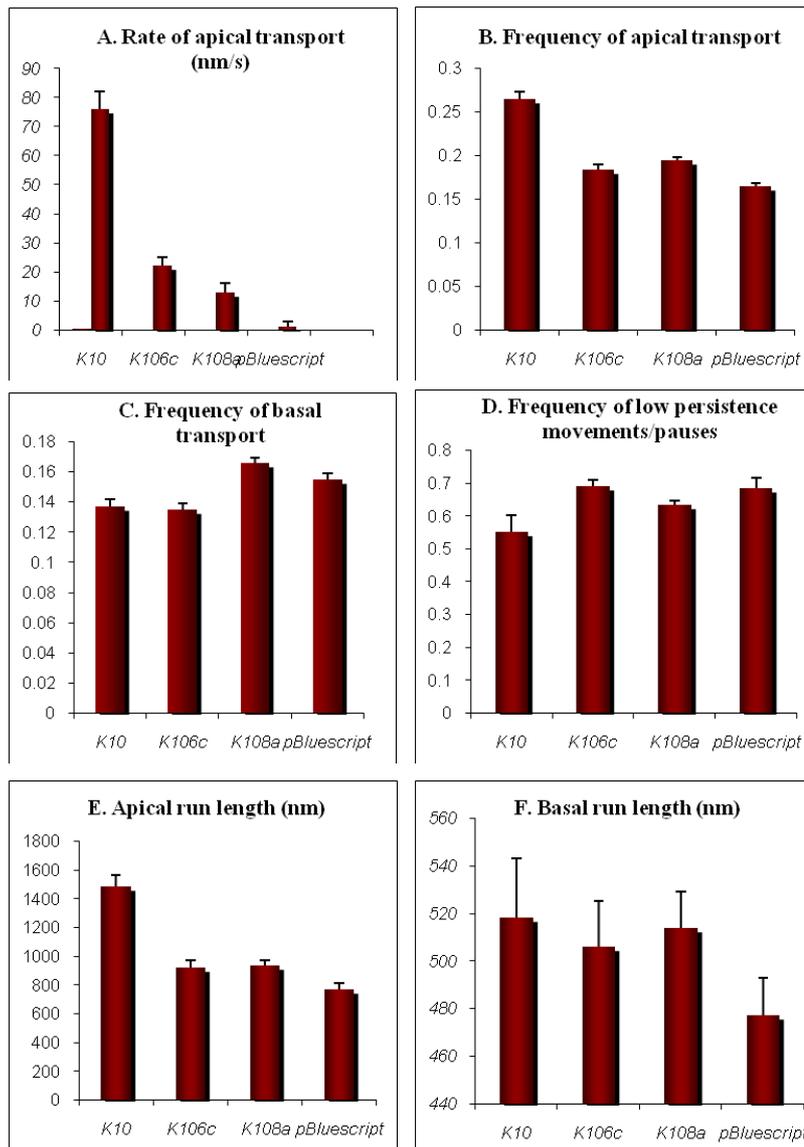
side of the nuclei, they stay there, suggesting that the anchoring system is less sensitive to mutations than the transport system.

The transcripts that were detected move in large particles. These particles are presumably much bigger than the endogenous motor complexes, which look like a haze of smaller, possibly monomeric complexes (Bullock et al., 2003). Nevertheless, the behaviour of these large particles parallels that of bulk injected transcripts and thus are likely to reflect the movement of endogenous (transported) RNA.



**Figure 6\_1 Semi-automatic tracking of the particles.**

Each image represents a single exposure (out of a series of ~20,000) taken by the confocal microscope. The time interval between the different panels is 5 seconds. The tracking program then scans each image for particles that look brighter than the background (depicted by red dots in the first image). The result of mounting all the images shows the full movement and the direction that particles have taken. Kinetics and directionality can then be studied by analysing the different parameters of movement, and compared between different injected embryos.



**Figure 6\_2 Kinetic parameters of transcript transports.**

(A) Rate of apical transport describes the speed of wt *K10* transcripts, which is higher compared with *K106C* and *K1010C* mutants and vector *pBluescript*. Frequency of apical localisation (B), compared with basal movements (C) or pauses. (D) shows that wt *K10* spends more time moving apically than moving basally or pausing. Non-localisers spends more time pausing. The distance of apical movements (E) is, as expected, higher in wt *K10*, but there is no significant difference in the distances the different transcripts travel basally (F).

	<i>K10</i>	<i>K106c</i>	<i>K108a</i>	<i>pBluescript</i>
Number of assayed embryos	5	3	4	7
Number of assayed particles	830	1357	2127	1809
Rate of apical accumulation (nm/s)	76+/-6	22+/-3 (P = 0.10)	13+/-3*	1+/-2**
Frequency of apical transport	0.264+/- 0.009	0.184+/- 0.005*	0.194+/- 0.004*	0.164+/- 0.004**
Frequency of basal transport	0.137+/- 0.005	0.135+/- 0.004	0.166+/- 0.003*	0.155+/- 0.004
Frequency of low persistence movements/pauses	0.553+/- 0.049	0.692+/- 0.018 (P=0.06)	0.635+/- 0.013	0.683+/- 0.033
Apical run length (nm)	1482+/-84	924+/-46	933+/-40	770+/-39
Basal run length (nm)	518+/-25	506+/-19	514+/-15	477+/-16
% of apical localisation	66	58	54	51

**Table 6\_1 Localisation of injected wild-type *K10* TLS compared with localisation of other transcripts.**

*K106C* and *K108A* are two transcripts with point mutations in the TLS that abolish localisation. *pBluescript* is a transcript of a vector that does not contain a localisation signal. Transport is defined as persistent movement (Bullock et al., 2003) with speed over 250nm/s.

\*P<0.05, \*\* P<0.01 and \*\*\* P<0.001 (ANOVA test). Numbers are mean +/- S.E.M.

## 7 Chapter 7 Discussion

RNA molecules can perform many different biological functions including regulation and catalytic activities. They undergo a series of different posttranscriptional events such as splicing, transport and decay. These events are mediated by specific *trans*-acting factors, either binding proteins or microRNAs, and by *cis*-acting sequences.

*cis* acting sequences usually reside in the UTR, which have a fundamental role in the spatial control of gene expression at the post-transcriptional level (Mignone et al., 2002). The interactions of these regulatory sequences with other molecules are highly dependent upon RNA folding and structure. Therefore, studying RNA structure is pivotal to the understanding of the mechanisms and regulations of the processes involved.

### **7.1 The structure of RNA localisation signals are likely to be conserved**

As the number of known localising mRNAs increases with recent systems-level analyses (Blower et al., 2007; Dubowy and Macdonald, 1998; Lecuyer et al., 2007; Mili et al., 2008; Shepard et al., 2003), it becomes more imperative to understand the specificity of the recognition of the localisation signals by the motor machinery.

Two main models represent the nature of signal recognition. One possibility is that each localising RNA has a unique signal, which is specifically recognised, based on its sequence or structure, by the adaptor protein/s. An alternative model is that a group of transcripts contain a common feature that is unique to localising transcripts, which are recognised by specific proteins. To support this model, there seem to be many common factors that are utilised for the localisation of different mRNAs .

Many families of proteins that participate in localisation are found in different cells and organisms. For example, Vera, a protein that binds specifically to, and is involved in the localisation of *Vgl* mRNA in *Xenopus* oocytes is homologous to chicken zip-code-binding protein (ZBP), which binds to a short RNA sequence required for localisation of  *$\beta$ -actin* mRNA in chick embryo fibroblasts. Both proteins contain five RNA-binding domains and putative signals for nuclear localisation and export (Deshler et al., 1998; Havin et al., 1998).

Another example of conserved protein is BicaudalD, which, in *Drosophila*, participates in the localisation of maternal mRNA during oogenesis, localisation of pair

rule gene transcripts in the embryo and transport of *prospero* and *miranda* RNA in the neuroblast (Bullock and Ish-Horowicz, 2001; Hughes et al., 2004). BicD is conserved among other species, and its mammalian homologue forms a complex with Rab6a and dynein-dynactin to function in retrograde Golgi-ER transport (Matanis et al., 2002). Its human homologue, Bicaudal D2, is capable of inducing microtubule minus end-directed movement (Hoogenraad et al., 2003). These observations strongly support the hypothesis that BicD is part of evolutionary conserved localisation machinery.

Some of the known components of the transport machinery can participate in more than one mode of intracellular transport within the same cell (Kural et al., 2005; Muller et al., 2008). For example, Staufen (Stau) participates in both posterior localisation of maternal transcripts including *oskar* (St Johnston et al., 1991) and anterior localisation of *bcd* (Ferrandon et al., 1994) along the microtubules. Stau is also part of a kinesin1 complex that moves to the plus end of the microtubules at the posterior end of the oocyte along with *osk* (but not with *bcd*) (Brendza et al., 2000). In later stages, Stau is required for actin dependent localisation of *prospero* RNA (Broadus et al., 1998).

In addition, localisation of the RNA is associated with other posttranscriptional events and there are proteins that participate in more than one of these processes in the cell. From nuclear binding (Cote et al., 1999), splicing (Goodrich et al., 2004; Hachet and Ephrussi, 2004; Le Hir et al., 2001; Palacios, 2002), nuclear export (Kress et al., 2004), translation repression (Goodrich et al., 2004; Smibert et al., 1999; Yano et al., 2004), degradation (Palacios et al., 2004) or initiation of translation (Palacios et al., 2004).

All these examples of conserved components of the transport machinery support a model in which a limited number of proteins are participating in RNA regulation. The function they serve might change depending on the location and timing of the formation of the complex, the organisation of the cytoskeleton, and the identities of the other proteins and RNAs within the complex. According to this model, different cargoes modulate transport by means of affecting the components of the transport complex, either by changing the number of bound proteins, or the kind of proteins, or the state of the proteins (for example by changing the conformation, dimerisation state, phosphorylation etc.). The overall direction of transport might be modified by interaction between the motor complex and the localisation sequences within the mRNA. For this model to work, signals need to have a conserved element that is recognised by the

limited amount of adaptors. In this work, I tried to identify a conserved element that acts as general recognition site in other localising signals.

## **7.2 Conclusions from this work**

Essential residues of the TLS were defined using mutational screens, site-directed mutagenesis, and comparison of different *K10* localisation signals from other *Drosophila* family members (chapter 2 and 3). The identities of U6, U7, and G8 are important for recognition by the localisation complex as changing their identity impairs localisation. They are all conserved across the species of *Drosophila* that I examined. The identity of the first base pair, C1-G44 is important, but probably not as a recognition site, but rather as structural element, as it forms a strong binding, based on three hydrogen connections, rather than two that can be seen in A-U and G-U pairs. The sequence identity of the loop is not conserved and can be replaced by a tetraloop sequence. The bulges, do not bare sequence identity conservation, and replacing them with other nucleotides does not affect the localisation of the RNA.

Despite the importance of the above sequences, it is apparent from the structural solution and from analysis of further mutations that the basic interactions of base-pairs, as predicted by mFOLD are not enough to explain the specificity of the signal and the recognition of the RNA is based on structural properties rather than specific nucleotides.

Transport of *K10* transcripts was analysed by injecting mutant and wild-type transcripts into a blastoderm embryo and following their accumulation or movement in fixed or live embryos (chapter 6). *K10* moves in an apical direction, and changes to its structure can affect its localisation activity. *K10* movement is reminiscent of other RNAs and organelles that are moving with dynein motor proteins (Bullock et al., 2006; Clark et al., 2007; Gross et al., 2000; Hughes et al., 2004; Jaramillo et al., 2008; Ross et al., 2008; Welte, 2004; Wilkie and Davis, 2001).

Localisation of *K10* mRNA is important for the activity of the protein and for the development of the fly embryo (Cheung et al., 1992; Haenlin et al., 1987). Mutations in the TLS affected the localisation of the transcript in the oocyte of transgenic flies that as a result had misshaped eggs and low percentage of hatching (chapter 4). However, a slight reduction in transport speed, as a result of weak mutations, was not always sufficient to create a phenotype.

Many RNA binding proteins recognise the single stranded nucleotides in an RNA structure. Proteins that bind to a double helix, usually bind to an area that is altered by a bulge or loop, which widens the groove and allows more space for the

protein to bind. Another option is a local distortion of the helix. This might be achieved by an RNA helicase-type factor, (i.e. similar to Vasa) which bends double-stranded RNA and forces local unwinding (Sengoku et al., 2006).

Solving the structure of the *K10* TLS (chapter 5 and figure 5\_2) was crucial for understanding the specificity of the signal. The structure proposed by the NMR results exposed intra-molecular interactions between distant, non base-paired nucleotides, for example between G16 and A28 in the upper helix, and G4 and A42 in the lower helix. These interactions are important because they change the size of the major grooves. The B-like helix form suggests a specificity of a structure of the RNA to the binding protein. An RNA A-form helix creates a deep and narrow groove that does not allow enough space for a protein to recognise specific nucleotides (Figure 7\_2).

Solving the structure of the mutant TLS, in addition to the wild-type signal, gave insights into structural changes that cause mis-localisation. The mutants in the upper and lower stem showed a reduction of localisation. This was correlated with the change of the structure of the helix. Mutant 13\_30GC in the upper helix that preserved the B-like form also did not affect localisation. Mutant U2G2 that disrupts the B-like form of the upper helix also caused a reduction of localisation, although it did not abolish it completely.

A further examination of the structure and localisation, that was done later on, by Simon Bullock in his lab (MRC Laboratory of Molecular Biology, Cambridge, UK), showed that a combination of mutations in both upper and lower helix of *K10* TLS abolishes localisation completely. Furthermore, he showed that replacing the two B-like helices with a different, known A-form helix from *BCI* RNA, also eliminates localisation. Hence, localisation is strongly dependent on the alignment of the two helices, and their specific widened grooves. The results also demonstrate the importance of comparative structural studies of mutant elements, in addition to wild type. Structural analysis of the wild type alone could not have identified the importance of the unexpectedly widened major grooves that are critical for signal activity.

### **7.3 Further questions and future prospects**

The unusual helix found in the *K10* TLS may be a general regulatory feature found in other localisation signals. Inter-strand purine stacks appear in other RNA molecules, for example, in stem II of the hammerhead ribozyme (Pley et al., 1994). The predicted secondary structures of other signals that mediate apical localisation in *Drosophila* embryos suggest that they could contain similar features to the *K10* TLS.

The localisation signal of *orb* lacks the upper bulged nucleotide but preserves the lower and upper stem stacking of base paired purines as found in the *K10* TLS (Cohen et al., 2005). Maybe to compensate for that, the upper U-A base paired helix is extended, which could maintain the relative orientation of the widened major grooves.

Other known signals that form a stem-loop element, such as signals responsible for localisation of *fushi tarazu* (Snee et al., 2005), *hairy* (Bullock et al., 2003), *wingless* (dos Santos et al., 2008) and *c-myc* (Chabanon et al., 2005), contain three or more adjacent purines on the same side of the stem. These stretches are candidates to form stacking interactions and extensive mutagenesis of these elements has revealed that some of these purines are essential for signal activity (Figure 7\_1a). Other elements that are predicted to create a similar structure with two accessible major grooves, are pyrimidine-rich bulges separated by short helical segments. These features appear in a group of localisation signals within the RNA of *bicoid*, (Macdonald and Kerr, 1998), *gurken* and *I-factor* (Van De Bor et al., 2005), *metallothionein-1* (Chabanon et al., 2004; Nury et al., 2005), *slow troponin C* (Reddy et al., 2005) and *vimentin* (Bermano et al., 2001) (Figure 7\_1b) and might also contribute to localisation activity of those signals. It remains to be seen if indeed the mentioned signals form as well unusual B-like helices. This could only be confirmed by structural means like NMR or crystallography.

Specific future studies should include the next questions:

1. Identifying common structural elements for localising signals. Experiments should include analysis of the structure of signals that might share a B-form helix, as suggested above, to determine, for example by NMR or X-ray crystallography. A comparison of their structure and determination if they indeed share the similar features, specifically the unusual B-like helix form, should provide, for the first time, a general structural feature that defines the specificity of localisation signals.

If this is found to be the case, further issues to be addressed will include establishing the identity of the recognition complex, and whether it includes several proteins with related recognition motifs or a single specific binding protein. In light of the mutational and structural findings, it is most likely that the TLS is recognised by dsRNA, rather than ssRNA, binding proteins, as the single stranded bulges are essential only as structural elements. Many RNA-binding proteins recognise secondary structures such as hairpins rather than the nucleotide sequence, in part because the narrow major groove in double-stranded RNA might not allow proteins to come into contact with the sides of the bases. It is not easy to identify the protein that binds the RNA directly *in*

*vivo* to link the cargo to the motor. RNA binding proteins can be non-specific *in vitro*. Methods employed to identify the protein/s that directly bind to the localisation signal on the RNA have primarily been either genetic screens for genes involved in mRNA localisation or affinity purification of proteins that bind the identified localisation elements (Arn et al., 2003; Kataoka et al., 2000; Long et al., 2000; Snee et al., 2005) (Lewis and Mowry, 2007). Identification of the binding partner(s) will help determine specificity and thereby structural predictions.

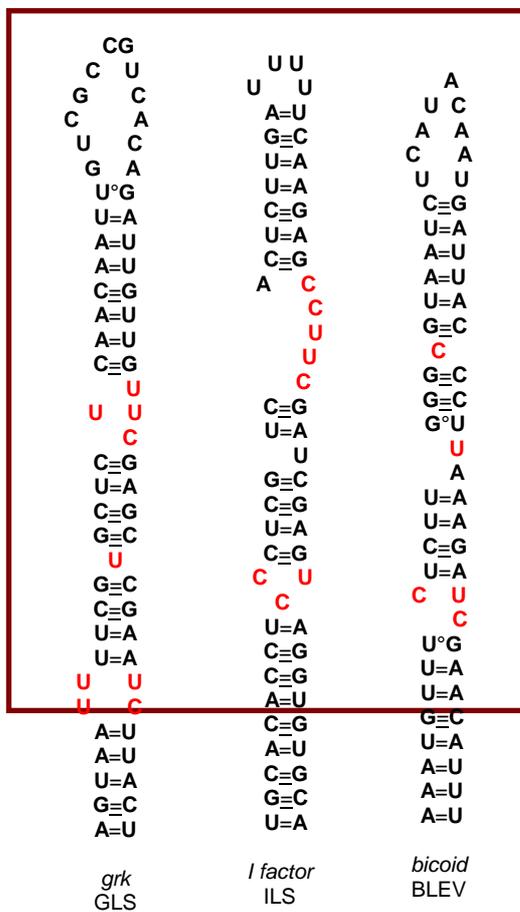
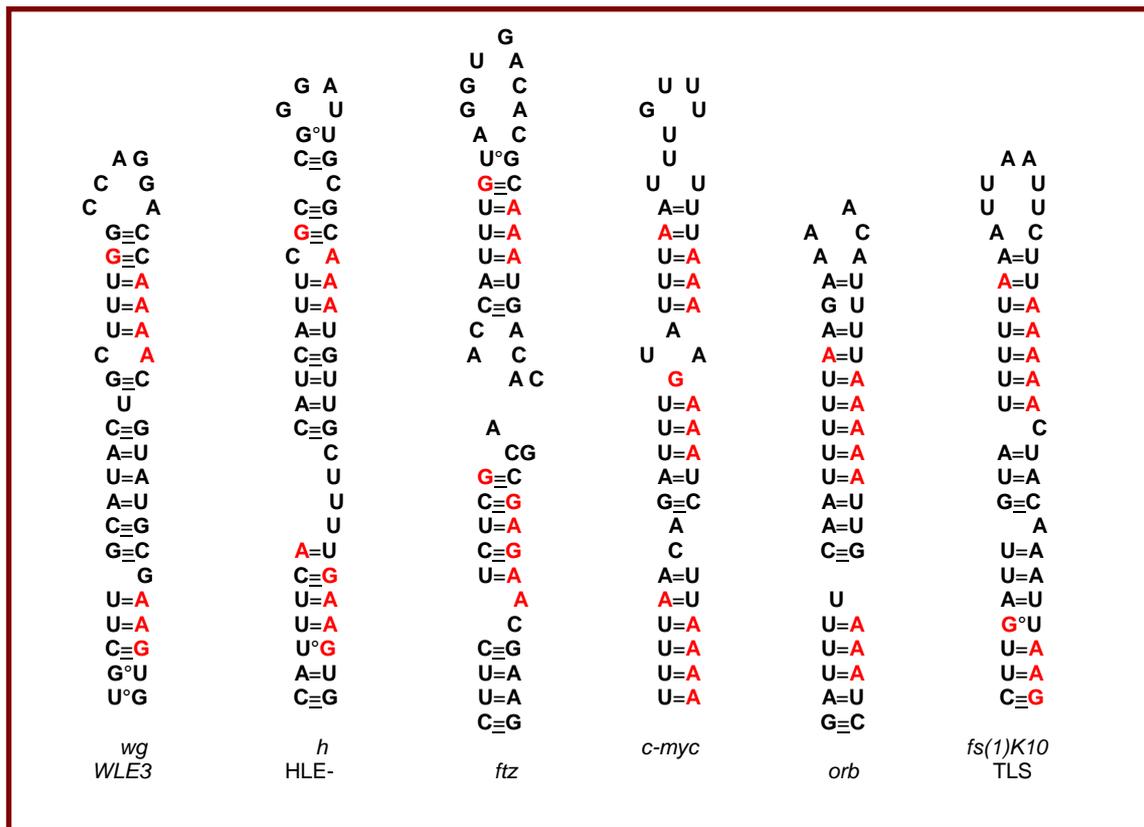
2. It might be possible, using bioinformatics tools (Gasteiger et al., 2003), to predict the structure of protein domains that can bind specifically to B-form helix, and try to infer the candidate protein and then verify their ability to bind the different localisation signals using biochemical tools. Prediction of the structure of the binding site of the RNA, as well as the protein domain (for example in (Auweter and Allain, 2008) and (Stroupe et al., 2006)) might facilitate the search for binding protein.
3. If these signals, or a group of them, share a single structural element, this structural motif will provide improved tools for prediction of *cis*-acting elements by bioinformatics approaches. New tools that allow screening for candidates for similar predicted secondary structure of known signals might help with finding signals based on their structure rather than their sequence. For example, a recent bioinformatics pipeline developed for searching across the genome for small RNA elements that are similar to the secondary structures of particular localisation signals (Hamilton et al., 2009), would benefit from a more refined description of the structural elements.
4. It would also be useful to establish new methods to identify the structure of a stem loop. Solving a structure by NMR is hard, time consuming, and limited to small sequences. There are different enzymatic and chemical probes (RNases, DMS, CMCT) that distinguish between paired and accessible (unpaired) nucleotides. Establishing new ways to distinguish between B and A form helices can further facilitate the study of signals, and allow better understanding of wild-type and mutation structures. Methods could include:
  - Comparing transcripts mobility on native acrylamide gels in order to correlate mobility on the gel with the structural changes (for example, unstable structures might create several conformers that run as multiple bands on a gel, while those that have only one stable conformer would run as one band).

- Identifying a reagent that selectively recognises B-form helices of RNA. Currently, all the known B-form selective molecules are DNA-targeted. They might not recognise the 2' hydroxyl of the ribose sugar even if it is contained within the grooves of a fundamentally B-form helical structure. Nevertheless, there are several candidates for a screen for RNA binding (Chaires, 2003; Dervan, 1986; Escude et al., 2007; Shaw and Arya, 2008), including B-form selective small molecules that are DNA-targeted, molecules that are known to be specific for major/minor groove structures, and different polyamides that have broad selectivity to AT-rich clusters. Comparing the binding of these reagents to B-form RNA helix (wild-type TLS) and A-form RNA helix (mutant U2G2 in TLS) can help in finding the right reagent to establish an easy and fast method to analyse RNA structures.
5. Many additional transcripts are now known to localise within cells (Dubowy and Macdonald, 1998; Lecuyer et al., 2007; Mili et al., 2008; Shepard et al., 2003). It will be interesting to see whether these transcripts reveal a new localisation pathway, or whether the transport machinery is based on limited number of pathways and participants. If this is the case, it is yet to be determined how the modulation of adaptors, motor proteins, and cytoplasmic fibres is functioning, and how the specificity of these transcripts is determined.

Localisation of RNA is now an expanding field, and the main reason being the fact that this is a fundamental process in almost any organism, from unicellular to vertebrates (Czaplinski and Singer, 2006; King et al., 2005; Paquin and Chartrand, 2008; St Johnston, 2005). Localisation of RNA is important in many aspects of cellular functions, including axis formation, cell fate determination, cell motility and neural plasticity. There is also evidence of human diseases associated with aberrant RNA localisation. Hence, understanding the mechanisms of RNA localisation, and the composition of proteins and RNAs that participate in it, can contribute to each of these biological functions. Understanding the structural requirement for localisation signals is a step towards the understanding of RNA-protein recognition and function in the context of transport complexes.

As more knowledge about RNA structure, particularly structures of small RNA motifs, folding and dynamics and structure of RNA–protein complexes, accumulates (Uhlenbeck et al., 1997), the database of RNA structural elements is ever increasing. Hence, it will be easier in the future to compare motifs, find new ones, and even predict functional elements based on known structures (Leontis et al., 2006). New studies into the structure of RNA and its binding proteins will lead to a better understanding of the

mechanisms and structural specificity of RNA signals. Improved imaging methods should allow a better and more precise visualisation of the dynamic of cargo in the cells. These new tools should lead to better understanding of the mechanisms of RNA localisation and other associated activities like transport of other cargoes, translational control, axis determination, neuronal plasticity and more.



**Figure 7\_1 Examples of localization signals that might share structural features.**

Elements were studied by a combination of chemical probes, mutation analysis and computer prediction (mFOLD). The names of the genes and the signal within the gene are indicated below the structure.

A. Shared feature of a continuous stack of purines, marked in red, in different stem-loop signals, is predicted to form an unusual B-like helix. These purine stacks appear in both upper and lower helix.

B. Another candidate to form the B-form helix is the feature of pyrimidine bulges, marked in red, separated by helices.



**Figure 7\_2 A-form and B-form helix**

A-form helix appears in RNA. It is right-handed with 11 nucleotide residues per turn. The plane of the bases is tilted with respect to the helix axis. The major groove is too narrow ( $\sim 3.8\text{\AA}$ ) to accommodate a protein. B-form Helix appears in most DNAs. It is a right-handed helix with 10 nucleotide residues in one helical pitch. The plane of the bases is almost perpendicular to the helix axis. The wide angle of the bases tilting creates prominent major and minor grooves ( $\sim 11.4\text{\AA}$ ).

## **8 Chapter 8 Materials and Methods**

### **8.1 Molecular Biology**

#### **8.1.1 Cloning and Constructs preparation**

##### **8.1.1.1 Generating templates for transcription**

The pBluescript II SK+ plasmids containing the entire 3' UTR and a 860-bp portion of the 3' genomic sequence of K10 was adapted with few changes: artificial HindIII and EcoRI or NheI sites were introduced by Site Directed Mutagenesis (SDM). The HindIII site (position 1277 (5' to the TLS)) and the EcoRI or NheI site (position 1565 or 1560 (3' to the HLE) in the K10 3'UTR are flanked by 288nt of an unrelated sequence, (from mouse Axin gene) instead of the 60nt that include the 44nt TLS sequence, to allow a faster recognition of successful 60nt clones.

Annealed oligonucleotides corresponding to the desired TLS derivatives were used to replace the sequence between the HindIII and NheI. The *orb* localisation signal, and the mutations associated with it were produced in exactly the same context of K10. In these, and all other cases, the resulted clones were sequenced to confirm that undesirable mutations had not been introduced.

##### **8.1.1.2 Site Directed Mutagenesis (SDM)**

HindIII, EcoRI and NheI sites were introduced by SDM into the 3'UTR around the TLS. SDM was performed using QuickChange kit 200528 from Stratagene according to the manufacturer's instructions

##### **8.1.1.3 Annealing of oligos for ligation**

5' directed and 3' directed oligos were mixed with TE buffer, heat for 5 minute in 95°C, and cooled down to room temperature. The double stranded product was then extracted using ethanol precipitation and 2µg of it were digested with the appropriate enzymes (usually HindIII and NheI), purified with MERmaid® SPIN kit from Q-BIOgene and resuspended in 10µl nuclease-free water.

#### **8.1.1.4 Enzymatic manipulation of DNA fragments, ligation and transformation of competent bacteria**

Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (NEB) and used for restriction digestion and DNA ligations, respectively, using manufacturer's instructions. Ligation products were then transformed into *DH5 $\alpha$*  competent bacteria. The transformation of *DH5 $\alpha$*  competent bacteria was performed according to the manufacturer's instructions (Invitrogen kit). To screen for the right ligation product, colonies were picked from plates, placed in 5ml LB medium with selective antibiotic and incubated on a shaker at 37°C overnight. Suspensions were centrifuged at 4000rpm in an Eppendorf floor centrifuge (5810), medium was discarded and plasmid DNA was purified using an automated version of the Qiagen mini-prep kit (Cancer Research UK Equipment Park). The plasmids were then cut with the appropriate enzymes and their length was estimated by running them on agarose gel. The clones that had the right insertion were 60nt of length instead of 288nt of the original vector.

#### **8.1.1.5 Random mutagenesis and point mutations**

For the generation of mutagenised *K10* signals, synthetic oligos were used with a 10%, 30% and 60% rate of mutations, based on a protocol for mutagenesis with degenerate oligonucleotides from (Ausubel et al., 1998). Mutation rate of 60% in a part of the signal that includes 12 nucleotides, thereby means that each position is occupied by the wild type nucleotide in 40% of the transcripts, and by each one of the three other nucleotides in 20% of the transcripts. That rate produces an average of six to eight different mutations in each clone. A rate of 10% produced 1-2 mutations in the region. The 10% mutant rate produced most of the clones that had point mutations in the area (see chapter 2).

To maximise the throughput in the assay, while ensuring that localised RNA can be detected within a pool of several non-localising transcripts, mixtures of localised and non-localised transcripts were co-injected. Several ratios of non-localising and localising RNAs were tested. Localisation could still be detected in a mixture of 4:1 non-localised/localised transcripts. However, I decided to use a 3:1 ratio in order to increase the likelihood of detecting transcripts that localised weakly.

A pool of three mutated clones were mixed and transcribed together with a single fluorescent dye. Three differently coloured dyes were used, so overall nine

different transcripts were assessed together. The transcripts were injected into blastoderm embryos and assayed for the ability to localise. If a localising transcript was detected in a pool, the three appropriate clones were transcribed and detected separately to check which one was the localising one and were then sequenced.

#### **8.1.1.6 Recombineering**

The wild type construct used in this study was generated by cutting the genomic sequence of K10 with HindII restriction enzyme from RP98-48022 that contains the BAC R48022 (obtained from <http://bacpac.chori.org/>). A band of 7Kb was extracted from an agarose gel, purified and sub-cloned into a pZERO1 vector. Polymerase Chain Reaction (PCR) with internal probes H1 and H2 was conducted to check which one of the colonies had the right insert and the right clones were checked using different restriction digests and sequencing to validate the sequence. It was then sub-cloned again into attB vector. In order to produce mutations in the signal I used plasmids containing the different mutations in the TLS in the context of 3'UTR. Thus the homology between the attB vector containing the wild type gene and the mutated plasmid is around 400bp in each side.

Recombineering procedures were based on previous protocols (Copeland et al., 2001; Court et al., 2002). The plasmids containing the mutations ("donors") were electroporated using a BIO-Rad gene pulser (0.2cm cuvettes, at 1.8kV, 25 $\mu$ F with pulse controller set to 200 Ohms) into EL 250 competent bacteria. After confirming the right product by PCR and sequencing, the cells were grown to OD 0.6-0.8 in 32°C. They were then induced by a 15 minute heat-shock at 42°C, followed by 10 minutes shaking on ice water and four washes with sterile water. The target vector was linearised (in StuI site, which is 15nt before the TLS) and then electroporated into the bacterial cells. The cells were left to recover by shaking for 2 hours at 32°C and plated on Ampicillin containing agar plates. Since the donor plasmid only contained Zeocin antibiotic resistance, and the target vector is linearised, in theory only plasmids that underwent homologue recombination would grow on the Ampicillin containing agar plates. Colonies were then checked by PCR, restriction enzymes and sequencing to confirm that the right recombination indeed occurred.

#### **8.1.1.7 Polymerase Chain Reaction (PCR)**

Amplification of DNA products for cloning was carried out using a Peltier (PTC-200, DNA Engine) thermal cycler. PCR was performed using the PCR Master

Mix system (Qiagen) according to manufacturer's instructions. For all PCR reactions, the thermal cycle used was: 94°C for 30 seconds, 45-55°C for 30 seconds and 72°C for 1 minute, for 30 cycles. Before the first cycle, the PCR reaction mix was incubated at 96°C for 5 minutes and cooled to 10°C after the final cycle.

The TOPO TA cloning kit (Invitrogen) was used for cloning of PCR fragments into the pCR 2.1-TOPO vector, according to manufacturer's instructions. Plasmids were then transformed into *DH5α* competent bacteria.

### **8.1.1.8 Sequencing**

All DNA Sequencing was performed on an ABI 3730 DNA Analyzer (Applied Biosystems) by the CRUK equipment park. Sequencing reaction was performed in a solution containing 200ng template DNA, 3.2pmol sequencing primer, and 8.0 μl BigDye Terminator Ready Reaction mix (Applied Biosystems), made up to a final volume of 20μl. Sequencing was carried out in either a single Thermo-Tube thin-walled tubes (ABgene), or in a 96 well plates (ABgene), when used for the mutant screens. DNA products were purified using DyeEx from Qiagene and ethanol participated at room temperature and air dried before loading onto the polyacrylamide sequencing gel.

## **8.2 RNA injections into syncytial blastoderm embryos**

### **8.2.1 In vitro transcription and fluorescent labelling**

10μg of plasmid DNA was linearised using NotI restriction enzyme, to produce a template for sense RNA synthesis. Template DNA was purified by phenol/chloroform extraction, precipitated with 0.3M NaOAc/EtOH and resuspended in 10μl nuclease-free water. Linearisation of template DNA was checked on a 1% TBE agarose gel. The template DNA was transcribed in a solution containing 0.4mM ATP, 0.4mM CTP, 0.36mM UTP, 0.04mM Cy3-, Cy5- (Perkin Elmer) or Alexa-488 (Molecular Probes) UTP, 0.12mM GTP, 0.3mM 7mG(5')pppG cap analogue (Ambion), and 10U RNase inhibitor (Stratagene), using 30U T7 polymerase (Stratagene) for sense or T3 polymerase (Roche) for anti-sense transcripts, and 2.5μl 10x transcription buffer (Roche). The reaction mixture was made up to 25μl using nuclease free water (Ambion). For time-lapse microscopy a 1:4 ratio instead of a 1:10 ratio of Alexa-488-UTP (Molecular Probes) to UTP was used in the transcription reaction.

The transcription reaction was performed at 37°C for 3 hours and was then treated with 10U DNase I (Stratagene) for 15 minutes at 37°C to remove template

DNA. RNA was extracted with phenol/chloroform and spun through a mini Quick Spin G50 column (Roche) to remove unincorporated nucleotides. RNA was precipitated with 0.3M NH<sub>4</sub>OAC/EtOH and resuspended in 2µl nuclease-free water (Ambion). The final concentration of RNA was typically around 1µg/µl and the fluorescent transcript contains 1 fluorochrome / 250 nucleotides. RNAs were stored at -20°C. The efficiency of RNA synthesis was checked by running 1/10th of the transcription reaction on a 1% TBE agarose gel after 10 min incubation in 75°C in RNA loading buffer (Molecular Probes). Gel tanks were washed with 10% SDS, prior to pouring of the agarose gel, in order to remove nucleases.

### **8.2.2 Blastoderm injection assay**

Wild-type flies (around one week old) were caged and induced to lay embryos on apple juice agar plates (produced by CRUK research services), by placing a little fresh yeast mixture at the centre of the plate. Cages were kept in a closed box, in the dark, at 25°C. To synchronise egg lays, a prelay was performed for 1 hour. Eggs from the prelay were discarded, fresh apple juice plates and yeast were put onto cages and a second 30-minute egg lay was performed. These embryos were then aged appropriately at 25°C, so that RNA was injected into mitotic cycle 13-14 blastoderm embryos (2.5 – 3 hours after egg lay at 25°C).

For preparation of embryos before injection, embryos were removed from the apple juice plate using a wet paintbrush and washed with water in a wire basket. After washing, embryos were soaked in bleach (Sodium hypochlorite; Anachem) for 1 min to remove the chorion and rinsed with water to wash off the bleach. Embryos were then lined up in rows on an apple juice plate so that dorsal sides were facing the same direction.

One side of a coverslip (9mm x 35mm) was covered in glue. Glue was prepared by dissolving the glue from packing tape, with 5ml n-heptane (AnalaR) in a 25ml glass bottle, placed on a tilting roller for about 30 minutes. Aligned embryos were picked up off the apple juice plate by gently sticking them to the glued coverslip. This was followed by dehydration for 10 min in a box containing Silica Gel. Embryos were covered with 10S voltaeff oil (Atachem) prior to RNA injection.

Glass injection needles were prepared by pulling capillary tubes on a Narashige needle puller, and broken on the edge of a glass slide, to give a tapered end of 1-2µm in diameter. RNA was pipetted into the glass injection needle, which was placed into a Leitz needle holder on a Narashige micromanipulator. Typically around 50 blastoderm

embryos were injected in a single experiment. RNA was injected at a concentration of 400ng/μl. All injections were conducted at 22°C. After 5 to 10 minutes, localisation was stopped by transferring the embryos to a fixing solution.

Injected embryos were fixed in n-heptane (AnalaR) saturated with formaldehyde (37% solution; AnalaR) (fix solution), 2 minutes after injection of the last embryo. Voltaleff oil was removed first by rinsing with fix solution until embryos started to detach from the glue. Embryos were then washed off the glue with fix solution into a 1.5ml Eppendorf tube and fixed for 15 minutes.

Following fixation, embryos were rinsed in heptane and dropped onto a glass slide using a plastic Pasteur pipette. After evaporation of heptane, embryos were stuck to another glass slide with double sided tape. Embryos were covered with PBST buffer and hand peeled with a fine syringe needle to remove the vitelline membrane. For observation of injected RNAs, embryos were immediately mounted.

Embryos were mounted on glass slides. A piece of insulation tape was put onto the slide, into which a square hole was cut, which formed a chamber for the embryos. Any remaining buffer was aspirated. The samples were then covered with Citifluor (Citifluor Ltd.) and overlaid with a coverslip. Covering the edges of the coverslip with nail varnish sealed the chamber.

Localisation of the transcripts is based on their relative distribution in the cytoplasm. Localisation was scored without the observer knowing the nature of the sample, and was designated as strong (most of the RNA in the apical cytoplasm, for example pair-rule transcripts), weak localisation (most of the of RNA remains basal but there are some apical caps of RNA) or non-localised (no apical concentration of RNA), (See Figure 2\_1).

### **8.3 Confocal microscopy**

Confocal imaging was performed on a Zeiss LSM 510 using a 40X water immersion lens. The standard image size was 1024 x 1024 pixels. Scale bars in all figures were calculated using LSM 510 software. Digital images were processed and arranged using the Adobe Photoshop CS2 software. For time-lapse microscopy, data were collected using a Perkin-Elmer Ultraview RS spinning disk confocal system and an Olympus IX71 inverted microscope, using a 603/1.2 NA UPlanApo water objective. Images were acquired in 3.37 frames/s.

### 8.3.1 Assaying mRNA Transport in Time-lapse fluorescence microscopy

Injection of capped fluorescent mRNAs for *in vivo* analyses was similar to the injected and assayed transcripts in fixed embryos, except for a higher ratio of labelled UTP that was used in preparation (25:75 instead of 10:90, see description above) and the final RNA concentration in the needle was increased from 1 mg/ml to 1.5 mg/ml.

Analysis of particles was based on previous work in our lab (Bullock et al., 2004). Image acquisition from a chosen focal plane started around 60s after injection. Particles were tracked using custom software in Mathematica 4.0 (Wolfram Research), beginning when they first displayed characteristics of active transport (i.e. at least three successive displacements in a persistent direction). All such particles in the observation field were tracked. Tracking ended when particles reached the apical cytoplasm or, more rarely, moved out of the plane of focus.

Images were improved by removing background fluorescence using functions from the Wavelet Explorer add-on package. The Inverse Wavelet Transform was applied to each image transform omitting the residual component to recover the improved image. Fluorescent mRNA particles were automatically detected in the processed image as contiguous clusters of pixels above a threshold. The threshold was automatically derived from a cumulative histogram of the improved image. It was taken as the position where the sequential step increment in cumulative frequency falls below a value of 250, although, in some cases, this value was manually elevated, to prevent false detection of background interference (and by that, reduced the number of particles taking into consideration for tracking). Small objects (<5 pixels or <500 total intensity) were rejected. For each particle, weighted centroids were determined, based on fluorescence intensity. Particles were then tracked, and exported to produce sequences of (x, y) coordinates.

Data were converted to nm/s and rotated manually, based on the predominant direction of particle movement, so that the direction of the apical movement is parallel to the y-axis. Higher y values, hence, represent more apical positions.

Short tracks of less than 12 consecutive frames were excluded from the analysis, as they usually represent particles that move away from the focal plane. Particles on the site of injection, and particles anchored to the apical end of the nuclei were excluded too from the analysis.

Directed transport was defined as displacements of particles in more than 250 nm/s. Apical-directed and basal-directed runs were defined as uninterrupted directed transport events along the y-axes. Run lengths are the net, direct displacements of these events. Persistence of movement defined in numbers between 0 and 1. A value of 0 is for the lowest persistence when in two steps the particle remains static or returns to its original position and persistence of 1 is for movement in a straight line.

For each particle, the mean of all displacements and the mean of directed displacements were calculated. These values were expressed as speed in  $\mu\text{m/s}$  and used to calculate the mean, SEM and maximum velocity for all particles of a given transcript in all injected embryos. Significances of differences between transcripts were tested in Mathematica 4.0 for analysis of all data using ANOVA tests with a nested unbalanced model (Milliken and Johnson, 1992).  $p$  values  $< 0.05$  were considered statistically significant.

## **8.4 Fly culture**

### **8.4.1 Alleles**

- **Wild-type** flies are of the strain Oregon-R.
- **$K10^{\text{LM00}}$**  (gift of Trudi Schupbach). A complete description of all alleles and balancer chromosomes is found at <http://flybase.bio.indiana.edu>. The strong  $K10$  loss-of-function allele,  $K10^{\text{LM00}}$ , was maintained over the X chromosome balancer, FM7. Homozygous  $K10^{\text{LM00}}$  females survive to adulthood but are sterile, laying dorsalised eggs that do not hatch. Hemizygous males ( $K10^{\text{LM00}}/Y$ ) are viable and fertile.
- **attP flies** (a gift from Konrad Basler) have the genotype  $yw$   $hs\text{-}flp, attP/attP$  (3R 86F),  $vas\text{-}phi/vas\text{-}phi$

### **8.4.2 Transgenic flies for rescue assay.**

The different  $K10$  transgenes that have full length coding region and 3'UTR containing the WT TLS, and mutant TLS (U2G2, 6G, and a scrambled sequence) were cloned first into a vector that contains an attB site (see above). These four vectors were then injected into  $yw$   $hs\text{-}flp, attP/attP$  (3R 86F),  $vas\text{-}phi/vas\text{-}phi$  flies. These fly lines have a precisely mapped attP site that allows the insertion of transgenes into predetermined intergenic location throughout the fly genome. The flies also contain an endogenous site-specific recombinase, phiC31 integrase, which eliminates the

necessities of co-injecting integrase mRNA. The flies also have a selection system based on the white gene that enables removing the rare events of non-specific integration. Hence, this system allows a specific integration of the gene into a predetermined location, and a direct comparison between the transgenes (Bischof et al., 2007). Four different constructs were injected into attP-zh14/vas-phi-zh1 embryos. Males that had w<sup>+</sup> were crossed to balancer virgins (yw, sb/TM3 ser y<sup>+</sup>) and the progeny were chosen for Full length K10 gene.

The genotypes of all transgenic flies were verified using PCR and sequencing to ensure that both the correct exogenous sequence was integrated, and that the fly did not contain any endogenous *K10* sequence. All of the injected constructs, including the wild-type, contained two restriction sites next to the TLS, not present in the endogenous *K10* gene. PCR primers (H1\_Bac and H2\_Bac, see primer list below) that are homologous to the restriction enzyme elements were used to recognise the construct, rather than the endogenous gene. The product of the PCR was sequenced to ensure the right genotyping.

#### **8.4.3 Egg hatching assay**

The K10 transgenes were introduced into a homozygous K10<sup>LM00</sup> mutant background using standard genetic crosses. For the rescue assay, five females of genotype K10<sup>LM00</sup>/K10<sup>LM00</sup>; transgene/transgene were mated to three to five wild type males. Approximately 200 eggs were collected on yeasted apple plates and the percentage that hatched was recorded.

### **8.5 Solutions and buffers**

**PBS:** 8 g NaCl, 0.2 g KCl, 1.44 g Na<sub>2</sub>HPO<sub>4</sub> and 0.24 g KH<sub>2</sub>PO<sub>4</sub> are added to 800 ml distilled H<sub>2</sub>O. The pH is adjusted to 7.4 with 1N HCl, and water added to make up to 1l. The solution is autoclaved.

**SOC:** 20 g bacto-tryptone, 5 g bacto-yeast extract and 0.5 g NaCl are added to 950 ml H<sub>2</sub>O and dissolved. 10 ml of a 250mM KCl solution is added, and the pH adjusted to 7.0 with 5 N NaOH. The volume is adjusted to 1 litre with deionised H<sub>2</sub>O, and autoclaved. The solution is cooled, and 20 ml of a sterile solution of 1M glucose added.

**LB Medium:** To 950 ml of deionised H<sub>2</sub>O, 10 g bacto-tryptone, 5 g bacto-yeast extract, and 10 g NaCl are added. The pH is raised to 7.0 with 5 N NaOH, then the

volume increased to 1 litre with deionised H<sub>2</sub>O. To produce LB Agar 15 g/litre of bacto-agar is added. The solution is autoclaved.

**LB plates:** 400 ml of LB Agar was melted in the microwave, and then cooled to 50°C, mixed with a selective antibiotic and poured into petri dishes (Sterilin, 90mm). For Ampicillin plates, 400µl or 50µl of a 100mg/ml ampicillin stock was added to make final concentration of 100µg/ml and 12.5µg/ml, respectively. For Zeocin plates, 200µl of 100mg/ml Zeocin stock is added to a final concentration of 50µg/ml.

**1X TBE (Tris-Borate-EDTA)** gel electrophoresis buffer: 89 mM Tris base, 89 mM Boric acid, 2 mM EDTA.

**PBT:** 0.1% Tween-20 (Sigma) in 1x PBS.

**PBST:** 0.1% Triton-X100 (Sigma) in 1x PBS

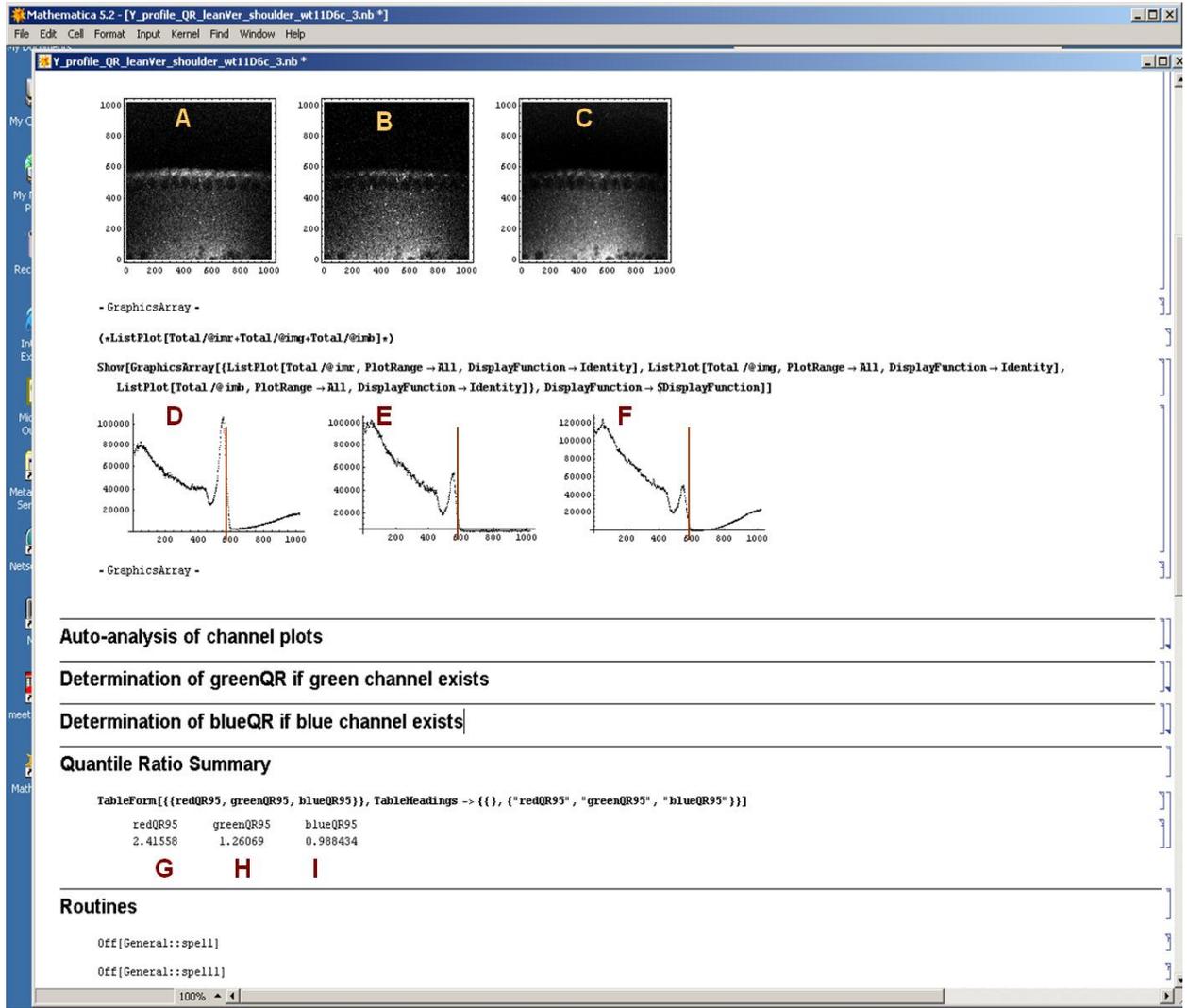
**37% fix solution:** 37% formaldehyde / heptane (1:1), shaken for 1 minute and allowed to settle for 15 minutes.

## 8.6 Primers List

Primer name	Sequence	Used for
H1BAC	GCAAGCACCTCCAGGCTATTGTGG	sequencing
T3	ACAAGAAGAAAAAGAAAGAGGGT	sequencing
s_k10	ATTAACCCTCACTAAAGG	sequencing
M13R	GGCTCTACTATTCCGCTAAGTGG	sequencing
M13F	CAGGAAACAGCTATGAC	sequencing
kry	GTAAAACGACGGCCAG	sequencing
K10@16vF1	ATGCTTGGATTTCACTGGA ACT	sequencing
	GCGCGGTGAAACCCAATTG	sequencing
H2BAC	GGCTGCAGGACGGAAAAGTCCAA	sequencing
F_56	TAAAGGGCAGACATATGAAGCC	sequencing
F_55	GGAAATAGCCAAGAGAAGTTTCGC	sequencing
attB2SEQ	CTATTGAACACCGTTGGCAATTC	sequencing
attB1SEQ	CGCCTTCCACAGCAGCTCTGGC	sequencing
F_66	GTGGTTTATTACTAAGGTATCC	sequencing
R_65	GGTCACCCAACAGCTGAGCGCGC	sequencing
	GTTCCTGGTACACCTCCTCCCGCC	sequencing
	AGCTTACCTTGATTGTATTTTTAAA	
F_HN	TAAATTCCTAAAACTACAAATTA	General primer for cloning mutation.
	AGATCACTCTG	
	CTAGCAGAGTGATCTTAATTTGTA	
R_HN	GTTTTTAAGAATTAATTTAAAAAT	General primer for cloning mutation.
	ACAATCAAGGTA	
	TCCAAGCTTACCTTGATTGTATTT	
	TTAAGAAAACATTTTTAAAAATACA	
5'HyB_K10_orb	AATTAAGATCACTCTGTG	Hybrid of K10 and orb
	TCCGGAATTCAAAAACAAATCCGT	
	AGATGCACACCATCGAGCACACGT	
3'HyB_K10_orb	TCACAGAGTGATCTTAAT	Hybrid of K10 and orb
	TCCAAGCTTACGAATTTCAATTTT	
	TAAATTAATTCTTAAAACTTGTA	
5'HyB_orb_K10	AATTCATCACTCTGTG	Hybrid of K10 and orb
	TCCGGAATTCAAAAACAAATCCGT	
	AGATGCACACCATCGAGCACACGT	
5'HyB_orb_K10	TCACAGAGTGATGAATTTAC	Hybrid of K10 and orb
T7 primer	GACTCTTGCTGTTGCTCGC	sequencing
	CTTAATTGTATTTTTAAATTAATTC	Sequence of TLS signal in different Dosophilea
D.Virilis	TTAAAAATTACAAATTAAG	Sequence of TLS signal in different Dosophilea
	CTTGATTGTATTTTTAAATTAATTC	Sequence of TLS signal in different Dosophilea
D.Yakuba	TTAAAACTACCAATTAAG	Sequence of TLS signal in different Dosophilea
	CTTGATTGTATTTTTAAATTAATTC	Sequence of TLS signal in different Dosophilea
D.ananassae	TTAAAACTACAAATTAAG	Sequence of TLS signal in different Dosophilea
	CTTGATTGTATTTTTAGGTTACTTC	Sequence of TLS signal in different Dosophilea
D.pseudoobscura	TTAGAAATCACAAATTAAG	Sequence of TLS signal in different Dosophilea
	CTTAATTGTATTTTTAAATTAATTC	Sequence of TLS signal in different Dosophilea
D.mojavensis	TTAAAAATTACAAATTAAG	Sequence of TLS signal in different Dosophilea
	AAATTCCAATTTTTAAGTACGATT	Sequence of TLS signal in different Dosophilea
Dv	TAAAAGTTGAGAATTT	Sequence of TLS signal in different Dosophilea
	GAATTTCAATTTTTAAGACATTTTA	Sequence of TLS signal in different Dosophilea
Da	AAAATTGTAAATTC	Sequence of TLS signal in different Dosophilea
	AAATTCCAATTTTTAAATACGAT	Sequence of TLS signal in different Dosophilea
Mj	TTAAAGGTTGTGAATTTT	Sequence of TLS signal in different Dosophilea
	GAATTTCAATTTTTAGAAAGCATT	Sequence of TLS signal in different Dosophilea
Dp	TTAAAAATTGTAAATTC	Sequence of TLS signal in different Dosophilea
Dy	GAATTTCAATTTTTAAGAAAACAT	Sequence of TLS signal in different Dosophilea

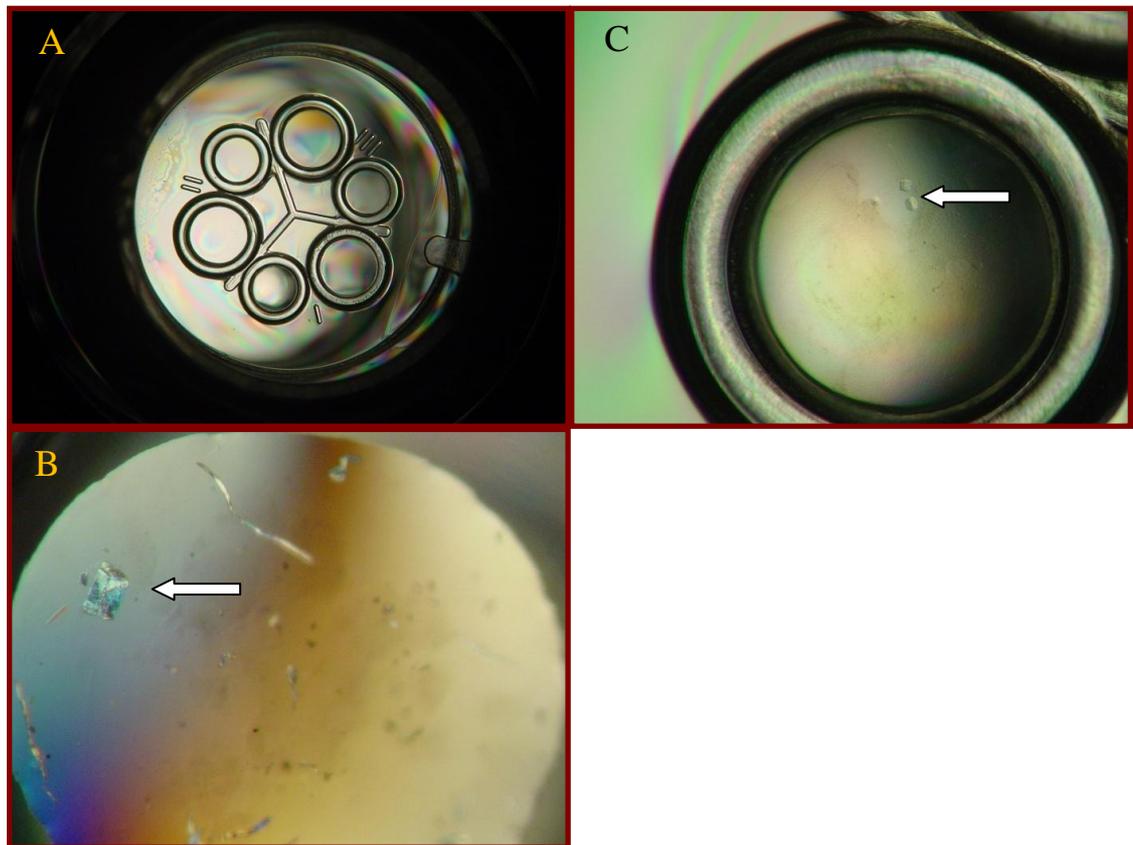
dM	TTTAAAAATTGTAAATTC GAATTTCAATTTTAAAGAAAACAT TTTAAAAATTGTAAATTC	different Dosophilea Sequence of TLS signal in different Dosophilea
short	CTTGATTGTATT::AAATTAATTCTT AA::CTACAAATTAAG	TLS mutants
long	CTTGATTGTTATTTTTAAATTAATT CTTAAAAACTAACAAATTAAG	TLS mutants
d2_43d13_30	CCTGATTGTA:TTTTAAATTAATTC TT:AAAACACTACAAATT:AG	TLS mutants
8c_36g	CTTGATTCTATTTTTAAATTAATTC TTAAAAACTAGAAATTAAG	TLS mutants
13_30UA-AU	CTTGATTGTATTATTAATTAATTC TTAATAACTACAAATTAAG	TLS mutants
13_30UA-CG	CTTGATTGTATTCTTAAATTAATTC TTAAGAACTACAAATTAAG	TLS mutants
13_30UA-GC	CTTGATTGTATTGTTAAATTAATTC TTAACAACTACAAATTAAG	TLS mutants
d_Clamp	CCTGATTGTATTTTTACTTCGGTAA AAACTACAAATTAAG	TLS mutants
Orb_12_30UA-AU	GAATTTCAATTATTAAGAAAACAT TTTAATAATTGTAAATTC	TLS mutants
Orb_12_30UA-CG	GAATTTCAATTCTTAAGAAAACAT TTTAAGAATTGTAAATTC	TLS mutants
Orb_5U-A	GAATATCAATTTTTAAGAAAACAT TTTAAAAATTGTAAATTC	TLS mutants
Orb_5U-C	GAATCTCAATTTTTAAGAAAACAT TTTAAAAATTGTAAATTC	TLS mutants
Orb_5U-G	GAATGTCAATTTTTAAGAAAACAT TTTAAAAATTGTAAATTC	TLS mutants
Orb_d36	GAATTTCAATTTTTAAGAAAACAT TTTAAAAATTGAAATTC	TLS mutants
xClamp	CCCGGCCGTACCCCGCTTCGGCG GGGGCTACAGGCTGGG	TLS mutants
UUUUU->CCCCC	CTTGATTGTACCCCAATTAATT CTTGGGGGCTACAAATTAAG	TLS mutants
UUUUU->UGUGU	CTTGATTGTATGTGTAATTAATTC TTACACACTACAAATTAAG	TLS mutants
UUUUU->UCUCU	CTTGATTGTATCTCTAAATTAATTC TTAGAGACTACAAATTAAG	TLS mutants
BSST7_44XhoBSS_F	CGCGCTAATACGACTCACTATAGG GCTTGATTGTATTTTTAAATTAATT CTTAAAAACTACAAATTAAGCCTC GAG	Introducing BssHI and XhoI restriction sites
BSST7_44XhoBSS_R	CGCGCTCGAGGCTTAATTTGTAGT TTTTAAGAATTAATTTAAAAATAC AATCAAGCCCTATAGTGAGTCGTA TTAG	Introducing BssHI and XhoI restriction sites
scrmbled	TTTATACTCATATATTATTAATGT AATTAATCTAGAACAATG	TLS mutants
Xclamp	CCCGGCCGTACCCCGCTTCGGCG GGGGCTACAGGCTGGGCACA	TLS mutants
H1	GCAAGCACCTCCAGGCTATTGTGG ACAAGAAGAAAAAGAAAGAGGGT TTATCGATACCGTCGACCTCGAG	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H1ext	GGTGGACAAAGTGCAACACGTGA AGATCCTTTCCAAGAAGCAGCGCA AGCACCTCCAGGCTATTGTGG	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H1p+pBS(2stepPCR)	AGAAGAAAAAGAAAGAGGGTTTA TCGATACCGTCGACCTCGAG	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H1_bac	GCAAGCACCTCCAGGCTATTGTGG ACAAGAAGAAAAAGAAAGAGGGT	primers to extend pBS with 40nt homology to BAC (by PCR, for

H2	GGCTGCAGGACGGAAAAGTCCAA TAAAGGGCAGACATATGAAGCCG CTTGATATCGAATTCCTGCAGC CGGGTCATATGCGTGCTTTGGAGC	recombineering). Also used for sequencing. primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H2ext	TGAGGGTGCGGCGGATTGGCTGCA GGACGGAAAAGTCCAATAAAGG	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H2p+pBS(2stepPCR)	AAGGGCAGACATATGAAGCCGCTT GATATCGAATTCCTGCAGC	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering)
H2_BAC	GGCTGCAGGACGGAAAAGTCCAA TAAAGGGCAGACATATGAAGCC CTACAAATTAAGATCACTCTGCTA	primers to extend pBS with 40nt homology to BAC (by PCR, for recombineering). Also used for sequencing.
NheI mutF	GCGTGTGCTCGATGGTGTGC GCACACCATCGAGCACACGCTAGC	creating NheI restriction site by SDM.
NheI_mutR	AGAGTGATCTTAATTTGTAG	creating NheI restriction site by SDM.



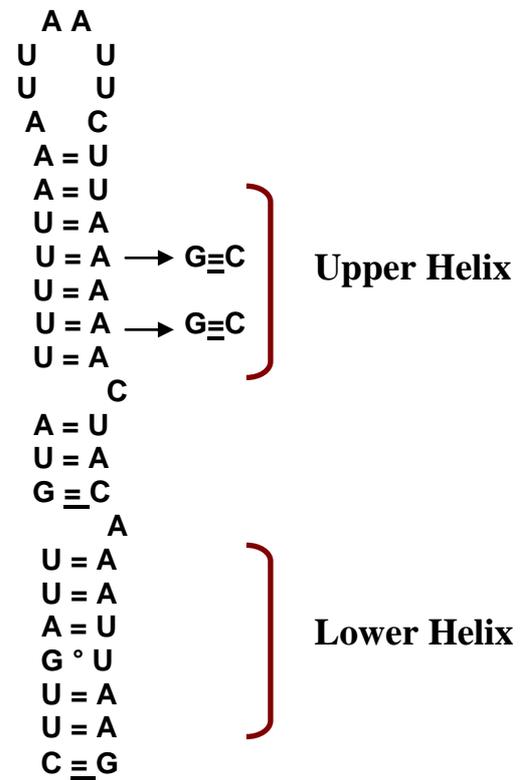
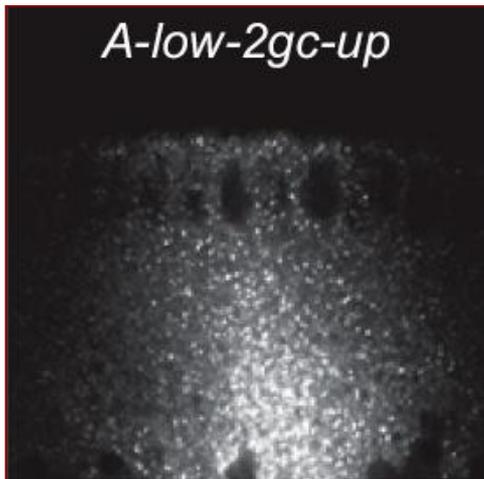
### Supplementary figure 2\_1 Localisation scoring by Quantile Ratio program (QR)

(A-C) An example of different injected transcripts in the cytoplasm of a fly embryo. (A) localised transcripts, (B) weak localisers and (C) non-localised transcripts. The ratio between the particles in the upper, apical part of the nuclei and the rest of the cytoplasm is calculated using the differences in brightness. Graphs (D-F) plot the relative accumulation of the particles along the cytoplasm. Y axis represent the brightness of the area in the image, X axis represent the length of the image from the point of injection at the bottom of the image to the top. The red vertical line denote the pick of the plot, where particles accumulate apically and create higher brightness. The ratio between the area under the high pick and the area under the rest of the graph (G-I) is calculated. High ratio (G) represents apical localisation, while low ratio (H-I) represents a weaker localisation.



**Appendix Figure 5\_1 crystallization of RNA**

(A). Crystallisation was carried out using a “hanging drop” technique. (B) Salt crystals are abundant, and are bigger than RNA or protein crystals (indicated by the arrow). (C) A suspected RNA crystal(indicated by the arrow).



**Supplementary figure 5\_1.**

Injected mutant transcript *A-low-2GC-up* in a blastoderm embryo does not localise. The transcript consist of mutation in the upper helix of the 2 base –pairs (similar to mutant UGUG) that enforce an A-form helix, and an exchange of the lower helix to a fragment taken from *BCI* RNA. This fragment was solved by NMR and it forms another A-form helix.

**This figure describes a result by Simon Bullock.**

## 9 References

- Ainger, K., Avossa, D., Morgan, F., Hill, S.J., Barry, C., Barbarese, E., and Carson, J.H. (1993). Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. *J Cell Biol* 123, 431-441.
- Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* 101, 1-22.
- Allain, F.H., and Varani, G. (1997). How accurately and precisely can RNA structure be determined by NMR? *J Mol Biol* 267, 338-351.
- Allison, R., Czaplinski, K., Git, A., Adegbenro, E., Stennard, F., Houliston, E., and Standart, N. (2004). Two distinct Staufen isoforms in *Xenopus* are vegetally localized during oogenesis. *Rna* 10, 1751-1763.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403-410.
- Antao, V.P., Lai, S.Y., and Tinoco, I., Jr. (1991). A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res* 19, 5901-5905.
- Antson, A.A. (2000). Single-stranded-RNA binding proteins. *Curr Opin Struct Biol* 10, 87-94.
- Arn, E.A., Cha, B.J., Theurkauf, W.E., and Macdonald, P.M. (2003). Recognition of a bicoid mRNA localization signal by a protein complex containing Swallow, Nod, and RNA binding proteins. *Dev Cell* 4, 41-51.
- Ausubel, F.M., Brent, R., E., K., Moore, D.D., J.G, S., J.A, S., and K., S. (1998). *Current Protocols in Molecular Biology* (John Wiley & Sons).
- Auweter, S.D., and Allain, F.H. (2008). Structure-function relationships of the polypyrimidine tract binding protein. *Cell Mol Life Sci* 65, 516-527.
- Bagni, C., and Greenough, W.T. (2005). From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome. *Nat Rev Neurosci* 6, 376-387.
- Barbarese, E., Koppel, D.E., Deutscher, M.P., Smith, C.L., Ainger, K., Morgan, F., and Carson, J.H. (1995). Protein translation components are colocalized in granules in oligodendrocytes. *J Cell Sci* 108 ( Pt 8), 2781-2790.
- Barros, C.S., Phelps, C.B., and Brand, A.H. (2003). *Drosophila* nonmuscle myosin II promotes the asymmetric segregation of cell fate determinants by cortical exclusion rather than active transport. *Dev Cell* 5, 829-840.
- Barton, N.R., and Goldstein, L.S. (1996). Going mobile: microtubule motors and chromosome segregation. *Proc Natl Acad Sci U S A* 93, 1735-1742.

- Bashirullah, A., Cooperstock, R.L., and Lipshitz, H.D. (1998). RNA localization in development. *Annu Rev Biochem* 67, 335-394.
- Bashirullah, A., Cooperstock, R.L., and Lipshitz, H.D. (2001). Spatial and temporal control of RNA stability. *Proc Natl Acad Sci U S A* 98, 7025-7028.
- Bassell, G.J., and Kelic, S. (2004). Binding proteins for mRNA localization and local translation, and their dysfunction in genetic neurological disease. *Curr Opin Neurobiol* 14, 574-581.
- Bassell, G.J., Oleynikov, Y., and Singer, R.H. (1999). The travels of mRNAs through all cells large and small. *Faseb J* 13, 447-454.
- Bassell, G.J., and Warren, S.T. (2008). Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60, 201-214.
- Beach, D.L., Salmon, E.D., and Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. *Curr Biol* 9, 569-578.
- Bergsten, S.E., and Gavis, E.R. (1999). Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. *Development* 126, 659-669.
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., and Nusslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. *Embo J* 7, 1749-1756.
- Bermano, G., Shepherd, R.K., Zehner, Z.E., and Hesketh, J.E. (2001). Perinuclear mRNA localisation by vimentin 3'-untranslated region requires a 100 nucleotide sequence and intermediate filaments. *FEBS Lett* 497, 77-81.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2, 437-445.
- Betley, J.N., Frith, M.C., Graber, J.H., Choo, S., and Deshler, J.O. (2002). A ubiquitous and conserved signal for RNA localization in chordates. *Curr Biol* 12, 1756-1761.
- Bischof, J., and Basler, K. (2008). Recombinases and their use in gene activation, gene inactivation, and transgenesis. *Methods Mol Biol* 420, 175-195.
- Bischof, J., Maeda, R.K., Hediger, M., Karch, F., and Basler, K. (2007). An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc Natl Acad Sci U S A* 104, 3312-3317.
- Blanquet, S., Mechulam, Y., and Schmitt, E. (2000). The many routes of bacterial transfer RNAs after aminoacylation. *Curr Opin Struct Biol* 10, 95-101.
- Bloomfield, V.A., Crothers, D.M., and Tinoco, I. (2000). *Nucleic Acids: Structures, Properties, and Functions* (University Science Books, 2000).

Blower, M.D., Feric, E., Weis, K., and Heald, R. (2007). Genome-wide analysis demonstrates conserved localization of messenger RNAs to mitotic microtubules. *J Cell Biol* 179, 1365-1373.

Bobola, N., Jansen, R.P., Shin, T.H., and Nasmyth, K. (1996). Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Cell* 84, 699-709.

Bohl, F., Kruse, C., Frank, A., Ferring, D., and Jansen, R.P. (2000). She2p, a novel RNA-binding protein tethers ASH1 mRNA to the Myo4p myosin motor via She3p. *Embo J* 19, 5514-5524.

Bratu, D.P., Cha, B.J., Mhlanga, M.M., Kramer, F.R., and Tyagi, S. (2003). Visualizing the distribution and transport of mRNAs in living cells. *Proc Natl Acad Sci U S A* 100, 13308-13313.

Brendel, C., Rehbein, M., Kreienkamp, H.J., Buck, F., Richter, D., and Kindler, S. (2004). Characterization of Staufen 1 ribonucleoprotein complexes. *Biochem J* 384, 239-246.

Brendza, R.P., Serbus, L.R., Duffy, J.B., and Saxton, W.M. (2000). A function for kinesin I in the posterior transport of oskar mRNA and Staufen protein. *Science* 289, 2120-2122.

Broadus, J., and Doe, C.Q. (1997). Extrinsic cues, intrinsic cues and microfilaments regulate asymmetric protein localization in *Drosophila* neuroblasts. *Curr Biol* 7, 827-835.

Broadus, J., Fuerstenberg, S., and Doe, C.Q. (1998). Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. *Nature* 391, 792-795.

Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., *et al.* (2001). Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 107, 477-487.

Brunel, C., and Ehresmann, C. (2004). Secondary structure of the 3' UTR of bicoid mRNA. *Biochimie* 86, 91-104.

Brunel, C., and Romby, P. (2000). Probing RNA structure and RNA-ligand complexes with chemical probes. *Methods Enzymol* 318, 3-21.

Bubunencko, M., Kress, T.L., Vempati, U.D., Mowry, K.L., and King, M.L. (2002). A consensus RNA signal that directs germ layer determinants to the vegetal cortex of *Xenopus* oocytes. *Dev Biol* 248, 82-92.

- Bullock, S.L. (2007). Translocation of mRNAs by molecular motors: think complex? *Semin Cell Dev Biol* 18, 194-201.
- Bullock, S.L., and Ish-Horowicz, D. (2001). Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* 414, 611-616.
- Bullock, S.L., Nicol, A., Gross, S.P., and Zicha, D. (2006). Guidance of bidirectional motor complexes by mRNA cargoes through control of dynein number and activity. *Curr Biol* 16, 1447-1452.
- Bullock, S.L., Stauber, M., Prell, A., Hughes, J.R., Ish-Horowicz, D., and Schmidt-Ott, U. (2004). Differential cytoplasmic mRNA localisation adjusts pair-rule transcription factor activity to cytoarchitecture in dipteran evolution. *Development* 131, 4251-4261.
- Bullock, S.L., Zicha, D., and Ish-Horowicz, D. (2003). The *Drosophila* hairy RNA localization signal modulates the kinetics of cytoplasmic mRNA transport. *Embo J* 22, 2484-2494.
- Campos-Ortega, J.A. (1997). Asymmetric division: dynastic intricacies of neuroblast division. *Curr Biol* 7, R726-728.
- Capri, M., Santoni, M.J., Thomas-Delaage, M., and Ait-Ahmed, O. (1997). Implication of a 5' coding sequence in targeting maternal mRNA to the *Drosophila* oocyte. *Mech Dev* 68, 91-100.
- Carlson, C.B., Stephens, O.M., and Beal, P.A. (2003). Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers* 70, 86-102.
- Carson, J.H., Kwon, S., and Barbarese, E. (1998). RNA trafficking in myelinating cells. *Curr Opin Neurobiol* 8, 607-612.
- Cha, B.J., Koppetsch, B.S., and Theurkauf, W.E. (2001). In vivo analysis of *Drosophila* bicoid mRNA localization reveals a novel microtubule-dependent axis specification pathway. *Cell* 106, 35-46.
- Chabanon, H., Mickleburgh, I., Burtle, B., Pedder, C., and Hesketh, J. (2005). An AU-rich stem-loop structure is a critical feature of the perinuclear localization signal of c-myc mRNA. *Biochem J* 392, 475-483.
- Chabanon, H., Nury, D., Mickleburgh, I., Burtle, B., and Hesketh, J. (2004). Characterization of the cis-acting element directing perinuclear localization of the metallothionein-1 mRNA. *Biochem Soc Trans* 32, 702-704.
- Chaires, J.B. (2003). A competition dialysis assay for the study of structure-selective ligand binding to nucleic acids. *Curr Protoc Nucleic Acid Chem* Chapter 8, Unit 8 3.
- Chang, J.S., Tan, L., Wolf, M.R., and Schedl, P. (2001). Functioning of the *Drosophila* orb gene in gurken mRNA localization and translation. *Development* 128, 3169-3177.

Chang, P., Torres, J., Lewis, R.A., Mowry, K.L., Houliston, E., and King, M.L. (2004). Localization of RNAs to the mitochondrial cloud in *Xenopus* oocytes through entrapment and association with endoplasmic reticulum. *Mol Biol Cell* *15*, 4669-4681.

Chartrand, P., Meng, X.H., Singer, R.H., and Long, R.M. (1999). Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo. *Curr Biol* *9*, 333-336.

Cheong, C., Varani, G., and Tinoco, I., Jr. (1990). Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature* *346*, 680-682.

Cheung, H.K., Serano, T.L., and Cohen, R.S. (1992). Evidence for a highly selective RNA transport system and its role in establishing the dorsoventral axis of the *Drosophila* egg. *Development* *114*, 653-661.

Choi, S.B., Wang, C., Muench, D.G., Ozawa, K., Franceschi, V.R., Wu, Y., and Okita, T.W. (2000). Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. *Nature* *407*, 765-767.

Christerson, L.B., and McKearin, D.M. (1994). orb is required for anteroposterior and dorsoventral patterning during *Drosophila* oogenesis. *Genes Dev* *8*, 614-628.

Clark, A., Meignin, C., and Davis, I. (2007). A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte. *Development* *134*, 1955-1965.

Clark, I.E., Jan, L.Y., and Jan, Y.N. (1997). Reciprocal localization of Nod and kinesin fusion proteins indicates microtubule polarity in the *Drosophila* oocyte, epithelium, neuron and muscle. *Development* *124*, 461-470.

Claussen, M., Horvay, K., and Pieler, T. (2004). Evidence for overlapping, but not identical, protein machineries operating in vegetal RNA localization along early and late pathways in *Xenopus* oocytes. *Development* *131*, 4263-4273.

Clote, P., Ferre, F., Kranakis, E., and Krizanc, D. (2005). Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency. *Rna* *11*, 578-591.

Cohen, R.S., Zhang, S., and Dollar, G.L. (2005). The positional, structural, and sequence requirements of the *Drosophila* TLS RNA localization element. *Rna* *11*, 1017-1029.

Condeelis, J., and Singer, R.H. (2005). How and why does beta-actin mRNA target? *Biol Cell* *97*, 97-110.

Cooley, L., and Theurkauf, W.E. (1994). Cytoskeletal functions during *Drosophila* oogenesis. *Science* *266*, 590-596.

Copeland, N.G., Jenkins, N.A., and Court, D.L. (2001). Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* 2, 769-779.

Costa, A., and Schedl, P. (2001). Conservation signals location. *Nature* 414, 593-595.

Cote, C.A., Gautreau, D., Denegre, J.M., Kress, T.L., Terry, N.A., and Mowry, K.L. (1999). A *Xenopus* protein related to hnRNP I has a role in cytoplasmic RNA localization. *Mol Cell* 4, 431-437.

Court, D.L., Sawitzke, J.A., and Thomason, L.C. (2002). Genetic engineering using homologous recombination. *Annu Rev Genet* 36, 361-388.

Crucs, S., Chatterjee, S., and Gavis, E.R. (2000). Overlapping but distinct RNA elements control repression and activation of nanos translation. *Mol Cell* 5, 457-467.

Czaplinski, K., and Mattaj, I.W. (2006). 40LoVe interacts with Vg1RBP/Vera and hnRNP I in binding the Vg1-localization element. *Rna* 12, 213-222.

Czaplinski, K., and Singer, R.H. (2006). Pathways for mRNA localization in the cytoplasm. *Trends Biochem Sci* 31, 687-693.

Dahlgaard, K., Raposo, A.A., Niccoli, T., and St Johnston, D. (2007). Capu and Spire assemble a cytoplasmic actin mesh that maintains microtubule organization in the *Drosophila* oocyte. *Dev Cell* 13, 539-553.

Dahm, R., and Macchi, P. (2007). Human pathologies associated with defective RNA transport and localization in the nervous system. *Biol Cell* 99, 649-661.

Dalgleish, G., Veyrune, J.L., Blanchard, J.M., and Hesketh, J. (2001). mRNA localization by a 145-nucleotide region of the c-fos 3'-untranslated region. Links to translation but not stability. *J Biol Chem* 276, 13593-13599.

Dalgleish, G.D., Veyrune, J.L., Accornero, N., Blanchard, J.M., and Hesketh, J.E. (1999). Localisation of a reporter transcript by the c-myc 3'-UTR is linked to translation. *Nucleic Acids Res* 27, 4363-4368.

Dandjinou, A.T., Levesque, N., Larose, S., Lucier, J.F., Abou Elela, S., and Wellinger, R.J. (2004). A phylogenetically based secondary structure for the yeast telomerase RNA. *Curr Biol* 14, 1148-1158.

Darnell, J.C., Fraser, C.E., Mostovetsky, O., Stefani, G., Jones, T.A., Eddy, S.R., and Darnell, R.B. (2005a). Kissing complex RNAs mediate interaction between the Fragile-X mental retardation protein KH2 domain and brain polyribosomes. *Genes Dev* 19, 903-918.

Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T., and Darnell, R.B. (2001). Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107, 489-499.

Darnell, J.C., Mostovetsky, O., and Darnell, R.B. (2005b). FMRP RNA targets: identification and validation. *Genes Brain Behav* 4, 341-349.

Davis, I., and Ish-Horowicz, D. (1991). Apical localization of pair-rule transcripts requires 3' sequences and limits protein diffusion in the *Drosophila* blastoderm embryo. *Cell* 67, 927-940.

Delanoue, R., and Davis, I. (2005). Dynein anchors its mRNA cargo after apical transport in the *Drosophila* blastoderm embryo. *Cell* 122, 97-106.

Delanoue, R., Herpers, B., Soetaert, J., Davis, I., and Rabouille, C. (2007). *Drosophila* Squid/hnRNP helps Dynein switch from a *gurken* mRNA transport motor to an ultrastructural static anchor in sponge bodies. *Dev Cell* 13, 523-538.

Dervan, P.B. (1986). Design of sequence-specific DNA-binding molecules. *Science* 232, 464-471.

Deshler, J.O., Highett, M.I., Abramson, T., and Schnapp, B.J. (1998). A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Curr Biol* 8, 489-496.

Deshler, J.O., Highett, M.I., and Schnapp, B.J. (1997). Localization of *Xenopus* Vg1 mRNA by Vera protein and the endoplasmic reticulum. *Science* 276, 1128-1131.

Desnos, C., Huet, S., and Darchen, F. (2007). 'Should I stay or should I go?': myosin V function in organelle trafficking. *Biol Cell* 99, 411-423.

Dicthenberg, J.B., Swanger, S.A., Antar, L.N., Singer, R.H., and Bassell, G.J. (2008). A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Dev Cell* 14, 926-939.

Ding, D., Parkhurst, S.M., Halsell, S.R., and Lipshitz, H.D. (1993). Dynamic Hsp83 RNA localization during *Drosophila* oogenesis and embryogenesis. *Mol Cell Biol* 13, 3773-3781.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M., and Skinner, M.A. (1990). HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *Embo J* 9, 4145-4153.

dos Santos, G., Simmonds, A.J., and Krause, H.M. (2008). A stem-loop structure in the wingless transcript defines a consensus motif for apical RNA transport. *Development* 135, 133-143.

Doudna, J.A. (2000). Structural genomics of RNA. *Nat Struct Biol* 7 *Suppl*, 954-956.

Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* 3, 195-205.

- Driever, W., and Nusslein-Volhard, C. (1988a). The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* 54, 95-104.
- Driever, W., and Nusslein-Volhard, C. (1988b). A gradient of bicoid protein in *Drosophila* embryos. *Cell* 54, 83-93.
- Dubowy, J., and Macdonald, P.M. (1998). Localization of mRNAs to the oocyte is common in *Drosophila* ovaries. *Mech Dev* 70, 193-195.
- Edgar, B.A., Odell, G.M., and Schubiger, G. (1987). Cytoarchitecture and the patterning of fushi tarazu expression in the *Drosophila* blastoderm. *Genes Dev* 1, 1226-1237.
- Egger, B., Boone, J.Q., Stevens, N.R., Brand, A.H., and Doe, C.Q. (2007). Regulation of spindle orientation and neural stem cell fate in the *Drosophila* optic lobe. *Neural Dev* 2, 1.
- Ennifar, E., Nikulin, A., Tishchenko, S., Serganov, A., Nevskaya, N., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. (2000). The crystal structure of UUCG tetraloop. *J Mol Biol* 304, 35-42.
- Eom, T., Antar, L.N., Singer, R.H., and Bassell, G.J. (2003). Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses. *J Neurosci* 23, 10433-10444.
- Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37-50.
- Ephrussi, A., and St Johnston, D. (2004). Seeing is believing: the bicoid morphogen gradient matures. *Cell* 116, 143-152.
- Escude, C., Roulon, T., Lyonnais, S., and Le Cam, E. (2007). Multiple topological labeling for imaging single plasmids. *Anal Biochem* 362, 55-62.
- Felden, B. (2007). RNA structure: experimental analysis. *Curr Opin Microbiol* 10, 286-291.
- Fernandez-Miragall, O., Ramos, R., Ramajo, J., and Martinez-Salas, E. (2006). Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation. *Rna* 12, 223-234.
- Ferrandon, D., Elphick, L., Nusslein-Volhard, C., and St Johnston, D. (1994). Stauf protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell* 79, 1221-1232.
- Ferrandon, D., Koch, I., Westhof, E., and Nusslein-Volhard, C. (1997). RNA-RNA interaction is required for the formation of specific bicoid mRNA 3' UTR-STAU-FEN ribonucleoprotein particles. *Embo J* 16, 1751-1758.

Foe, V.E. (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.

Forlani, S., Ferrandon, D., Saget, O., and Mohier, E. (1993). A regulatory function for K10 in the establishment of dorsoventral polarity in the *Drosophila* egg and embryo. *Mech Dev* *41*, 109-120.

Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*. *Curr Biol* *13*, 1159-1168.

Franke, W.W., Grund, C., Kuhn, C., Jackson, B.W., and Illmensee, K. (1982). Formation of cytoskeletal elements during mouse embryogenesis. III. Primary mesenchymal cells and the first appearance of vimentin filaments. *Differentiation* *23*, 43-59.

Fromm, L., and Rhode, M. (2004). Neuregulin-1 induces expression of Egr-1 and activates acetylcholine receptor transcription through an Egr-1-binding site. *J Mol Biol* *339*, 483-494.

Frugier, T., Nicole, S., Cifuentes-Diaz, C., and Melki, J. (2002). The molecular bases of spinal muscular atrophy. *Curr Opin Genet Dev* *12*, 294-298.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., and Bertrand, E. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol* *13*, 161-167.

Gabus, C., Mazroui, R., Tremblay, S., Khandjian, E.W., and Darlix, J.L. (2004). The fragile X mental retardation protein has nucleic acid chaperone properties. *Nucleic Acids Res* *32*, 2129-2137.

Gantois, I., and Kooy, R.F. (2002). Targeting fragile X. *Genome Biol* *3*, reviews1014.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* *31*, 3784-3788.

Gautreau, D., Cote, C.A., and Mowry, K.L. (1997). Two copies of a subelement from the Vg1 RNA localization sequence are sufficient to direct vegetal localization in *Xenopus* oocytes. *Development* *124*, 5013-5020.

Gautrey, H., McConnell, J., Hall, J., and Hesketh, J. (2005). Polarised distribution of the RNA-binding protein Staufen in differentiated intestinal epithelial cells. *FEBS Lett* 579, 2226-2230.

Gavis, E.R., Curtis, D., and Lehmann, R. (1996a). Identification of cis-acting sequences that control nanos RNA localization. *Dev Biol* 176, 36-50.

Gavis, E.R., Lunsford, L., Bergsten, S.E., and Lehmann, R. (1996b). A conserved 90 nucleotide element mediates translational repression of nanos RNA. *Development* 122, 2791-2800.

Gepner, J., Li, M., Ludmann, S., Kortas, C., Boylan, K., Iyadurai, S.J., McGrail, M., and Hays, T.S. (1996). Cytoplasmic dynein function is essential in *Drosophila melanogaster*. *Genetics* 142, 865-878.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 582, 1977-1986.

Glotzer, J.B., Saffrich, R., Glotzer, M., and Ephrussi, A. (1997). Cytoplasmic flows localize injected oskar RNA in *Drosophila* oocytes. *Curr Biol* 7, 326-337.

Gonsalvez, G.B., Urbinati, C.R., and Long, R.M. (2005). RNA localization in yeast: moving towards a mechanism. *Biol Cell* 97, 75-86.

Gonzalez-Reyes, A., Elliott, H., and St Johnston, D. (1995). Polarization of both major body axes in *Drosophila* by gurken-torpedo signalling. *Nature* 375, 654-658.

Goodrich, J.S., Clouse, K.N., and Schupbach, T. (2004). Hrb27C, Sqd and Otu cooperatively regulate gurken RNA localization and mediate nurse cell chromosome dispersion in *Drosophila* oogenesis. *Development* 131, 1949-1958.

Gorodkin, J., Heyer, L.J., and Stormo, G.D. (1997). Finding the most significant common sequence and structure motifs in a set of RNA sequences. *Nucleic Acids Res* 25, 3724-3732.

Greber, U.F., and Way, M. (2006). A superhighway to virus infection. *Cell* 124, 741-754.

Gross, S.P., Welte, M.A., Block, S.M., and Wieschaus, E.F. (2000). Dynein-mediated cargo transport in vivo. A switch controls travel distance. *J Cell Biol* 148, 945-956.

Gross, S.P., Welte, M.A., Block, S.M., and Wieschaus, E.F. (2002). Coordination of opposite-polarity microtubule motors. *J Cell Biol* 156, 715-724.

Gutzeit, H.O. (1986). The role of microfilaments in cytoplasmic streaming in *Drosophila* follicles. *J Cell Sci* 80, 159-169.

Gutzeit, H.O. (1991). Organization and in vitro activity of microfilament bundles associated with the basement membrane of *Drosophila* follicles. *Acta Histochem Suppl* 41, 201-210.

Hachet, O., and Ephrussi, A. (2004). Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* 428, 959-963.

Haenlin, M., Roos, C., Cassab, A., and Mohier, E. (1987). Oocyte-specific transcription of fs(1)K10: a *Drosophila* gene affecting dorsal-ventral developmental polarity. *Embo J* 6, 801-807.

Hafen, E., Kuroiwa, A., and Gehring, W.J. (1984). Spatial distribution of transcripts from the segmentation gene *fushi tarazu* during *Drosophila* embryonic development. *Cell* 37, 833-841.

Hamilton, R.S., and Davis, I. (2007). RNA localization signals: deciphering the message with bioinformatics. *Semin Cell Dev Biol* 18, 178-185.

Hamilton, R.S., Hartswood, E., Vendra, G., Jones, C., Van De Bor, V., Finnegan, D., and Davis, I. (2009). A bioinformatics search pipeline, RNA2DSearch, identifies RNA localization elements in *Drosophila* retrotransposons. *Rna* 15, 200-207.

Havin, L., Git, A., Elisha, Z., Oberman, F., Yaniv, K., Schwartz, S.P., Standart, N., and Yisraeli, J.K. (1998). RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization. *Genes Dev* 12, 1593-1598.

Hermann, T., and Patel, D.J. (1999). Stitching together RNA tertiary architectures. *J Mol Biol* 294, 829-849.

Hermann, T., and Patel, D.J. (2000). RNA bulges as architectural and recognition motifs. *Structure Fold Des* 8, R47-54.

Herpers, B., and Rabouille, C. (2004). mRNA localization and ER-based protein sorting mechanisms dictate the use of transitional endoplasmic reticulum-golgi units involved in gurken transport in *Drosophila* oocytes. *Mol Biol Cell* 15, 5306-5317.

Hesketh, J. (2004). 3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein. *Biochem Soc Trans* 32, 990-993.

Hesketh, J.E., Vasconcelos, M.H., and Bermano, G. (1998). Regulatory signals in messenger RNA: determinants of nutrient-gene interaction and metabolic compartmentation. *Br J Nutr* 80, 307-321.

Hofacker, I.L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Res* 31, 3429-3431.

Hofacker, I.L., Fekete, M., and Stadler, P.F. (2002). Secondary structure prediction for aligned RNA sequences. *J Mol Biol* 319, 1059-1066.

Hoogenraad, C.C., Akhmanova, A., Howell, S.A., Dortland, B.R., De Zeeuw, C.I., Willemsen, R., Visser, P., Grosveld, F., and Galjart, N. (2001). Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-dynactin pathway by interacting with these complexes. *Embo J* 20, 4041-4054.

Hoogenraad, C.C., Wulf, P., Schiefermeier, N., Stepanova, T., Galjart, N., Small, J.V., Grosveld, F., de Zeeuw, C.I., and Akhmanova, A. (2003). Bicaudal D induces selective dynein-mediated microtubule minus end-directed transport. *Embo J* 22, 6004-6015.

Hou, L., Antion, M.D., Hu, D., Spencer, C.M., Paylor, R., and Klann, E. (2006). Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* 51, 441-454.

Hughes, J.R., Bullock, S.L., and Ish-Horowicz, D. (2004). Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in *Drosophila* neuroblasts. *Curr Biol* 14, 1950-1956.

Huttelmaier, S., Zenklusen, D., Lederer, M., Dichtenberg, J., Lorenz, M., Meng, X., Bassell, G.J., Condeelis, J., and Singer, R.H. (2005). Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* 438, 512-515.

Ikeshima-Kataoka, H., Skeath, J.B., Nabeshima, Y., Doe, C.Q., and Matsuzaki, F. (1997). Miranda directs Prospero to a daughter cell during *Drosophila* asymmetric divisions. *Nature* 390, 625-629.

Ingham, P.W., Pinchin, S.M., Howard, K.R., and Ish-Horowicz, D. (1985). Genetic Analysis of the Hairy Locus in *DROSOPHILA MELANOGASTER*. *Genetics* 111, 463-486.

Jackle, H., Hoch, M., Pankratz, M.J., Gerwin, N., Sauer, F., and Bronner, G. (1992). Transcriptional control by *Drosophila* gap genes. *J Cell Sci Suppl* 16, 39-51.

Jambhekar, A., and Derisi, J.L. (2007). Cis-acting determinants of asymmetric, cytoplasmic RNA transport. *Rna* 13, 625-642.

Jambhekar, A., McDermott, K., Sorber, K., Shepard, K.A., Vale, R.D., Takizawa, P.A., and DeRisi, J.L. (2005). Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 102, 18005-18010.

Jansen, R.P., Dowzer, C., Michaelis, C., Galova, M., and Nasmyth, K. (1996). Mother cell-specific HO expression in budding yeast depends on the unconventional myosin myo4p and other cytoplasmic proteins. *Cell* 84, 687-697.

- Jaramillo, A.M., Weil, T.T., Goodhouse, J., Gavis, E.R., and Schupbach, T. (2008). The dynamics of fluorescently labeled endogenous gurken mRNA in *Drosophila*. *J Cell Sci* *121*, 887-894.
- Job, C., and Eberwine, J. (2001). Localization and translation of mRNA in dendrites and axons. *Nat Rev Neurosci* *2*, 889-898.
- Juhn, J., Marinotti, O., Calvo, E., and James, A.A. (2008). Gene structure and expression of nanos (nos) and oskar (osk) orthologues of the vector mosquito, *Culex quinquefasciatus*. *Insect Mol Biol* *17*, 545-552.
- Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* *43*, 513-525.
- Karlin-Mcginness, M., Serano, T.L., and Cohen, R.S. (1996). Comparative analysis of the kinetics and dynamics of K10, bicoid, and oskar mRNA localization in the *Drosophila* oocyte. *Dev Genet* *19*, 238-248.
- Kataoka, N., Yong, J., Kim, V.N., Velazquez, F., Perkinson, R.A., Wang, F., and Dreyfuss, G. (2000). Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol Cell* *6*, 673-682.
- Kaytor, M.D., and Orr, H.T. (2001). RNA targets of the fragile X protein. *Cell* *107*, 555-557.
- Ke, A., and Doudna, J.A. (2004). Crystallization of RNA and RNA-protein complexes. *Methods* *34*, 408-414.
- Kelley, R.L. (1993). Initial organization of the *Drosophila* dorsoventral axis depends on an RNA-binding protein encoded by the squid gene. *Genes Dev* *7*, 948-960.
- Kiebler, M.A., Hemraj, I., Verkade, P., Kohrmann, M., Fortes, P., Marion, R.M., Ortin, J., and Dotti, C.G. (1999). The mammalian staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: implications for its involvement in mRNA transport. *J Neurosci* *19*, 288-297.
- Kim-Ha, J., Smith, J.L., and Macdonald, P.M. (1991). oskar mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* *66*, 23-35.
- King, M.L., Messitt, T.J., and Mowry, K.L. (2005). Putting RNAs in the right place at the right time: RNA localization in the frog oocyte. *Biol Cell* *97*, 19-33.
- King, R.C., and Burnett, R.G. (1959). Autoradiographic study of uptake of tritiated glycine, thymidine, and uridine by fruit fly ovaries. *Science* *129*, 1674-1675.
- Kislauskis, E.H., Li, Z., Singer, R.H., and Taneja, K.L. (1993). Isoform-specific 3'-untranslated sequences sort alpha-cardiac and beta-cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J Cell Biol* *123*, 165-172.

Kislauskis, E.H., Zhu, X., and Singer, R.H. (1994). Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *J Cell Biol* 127, 441-451.

Kislauskis, E.H., Zhu, X., and Singer, R.H. (1997). beta-Actin messenger RNA localization and protein synthesis augment cell motility. *J Cell Biol* 136, 1263-1270.

Kloc, M., and Etkin, L.D. (1994). Delocalization of Vg1 mRNA from the vegetal cortex in *Xenopus* oocytes after destruction of Xlsirt RNA. *Science* 265, 1101-1103.

Kloc, M., and Etkin, L.D. (2005). RNA localization mechanisms in oocytes. *J Cell Sci* 118, 269-282.

Kloc, M., Zearfoss, N.R., and Etkin, L.D. (2002). Mechanisms of subcellular mRNA localization. *Cell* 108, 533-544.

Knirr, S., Breuer, S., Paululat, A., and Renkawitz-Pohl, R. (1997). Somatic mesoderm differentiation and the development of a subset of pericardial cells depend on the not enough muscles (nem) locus, which contains the inscuteable gene and the intron located gene, skittles. *Mech Dev* 67, 69-81.

Knoblich, J.A., Jan, L.Y., and Jan, Y.N. (1999). Deletion analysis of the *Drosophila* Inscuteable protein reveals domains for cortical localization and asymmetric localization. *Curr Biol* 9, 155-158.

Koch, E.A., and Spitzer, R.H. (1983). Multiple effects of colchicine on oogenesis in *Drosophila*: induced sterility and switch of potential oocyte to nurse-cell developmental pathway. *Cell Tissue Res* 228, 21-32.

Koob, M.D., Moseley, M.L., Schut, L.J., Benzow, K.A., Bird, T.D., Day, J.W., and Ranum, L.P. (1999). An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet* 21, 379-384.

Kress, T.L., Yoon, Y.J., and Mowry, K.L. (2004). Nuclear RNP complex assembly initiates cytoplasmic RNA localization. *J Cell Biol* 165, 203-211.

Kruse, C., Jaedicke, A., Beaudouin, J., Bohl, F., Ferring, D., Guttler, T., Ellenberg, J., and Jansen, R.P. (2002). Ribonucleoprotein-dependent localization of the yeast class V myosin Myo4p. *J Cell Biol* 159, 971-982.

Kural, C., Kim, H., Syed, S., Goshima, G., Gelfand, V.I., and Selvin, P.R. (2005). Kinesin and dynein move a peroxisome in vivo: a tug-of-war or coordinated movement? *Science* 308, 1469-1472.

Kwon, S., Abramson, T., Munro, T.P., John, C.M., Kohrmann, M., and Schnapp, B.J. (2002). UUCAC- and vera-dependent localization of VegT RNA in *Xenopus* oocytes. *Curr Biol* 12, 558-564.

Lane, M.E., and Kalderon, D. (1994). RNA localization along the anteroposterior axis of the *Drosophila* oocyte requires PKA-mediated signal transduction to direct normal microtubule organization. *Genes Dev* 8, 2986-2995.

Lantz, V., Ambrosio, L., and Schedl, P. (1992). The *Drosophila orb* gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. *Development* 115, 75-88.

Lantz, V., Chang, J.S., Horabin, J.I., Bopp, D., and Schedl, P. (1994). The *Drosophila orb* RNA-binding protein is required for the formation of the egg chamber and establishment of polarity. *Genes Dev* 8, 598-613.

Lantz, V., and Schedl, P. (1994). Multiple cis-acting targeting sequences are required for orb mRNA localization during *Drosophila* oogenesis. *Mol Cell Biol* 14, 2235-2242.

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

Le Hir, H., Gatfield, D., Braun, I.C., Forler, D., and Izaurralde, E. (2001). The protein Mago provides a link between splicing and mRNA localization. *EMBO Rep* 2, 1119-1124.

Lecuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174-187.

Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., *et al.* (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80, 155-165.

Lefebvre, S., Burlet, P., Liu, Q., Bertrand, S., Clermont, O., Munnich, A., Dreyfuss, G., and Melki, J. (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16, 265-269.

Leontis, N.B., Lescoute, A., and Westhof, E. (2006). The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* 16, 279-287.

Leontis, N.B., Stombaugh, J., and Westhof, E. (2002). The non-Watson-Crick base pairs and their associated isosteric matrices. *Nucleic Acids Res* 30, 3497-3531.

Levadoux-Martin, M., Hesketh, J.E., Beattie, J.H., and Wallace, H.M. (2001). Influence of metallothionein-1 localization on its function. *Biochem J* 355, 473-479.

Levadoux-Martin, M., Li, Y., Blackburn, A., Chabanon, H., and Hesketh, J.E. (2006). Perinuclear localisation of cellular retinoic acid binding protein I mRNA. *Biochem Biophys Res Commun* 340, 326-331.

- Levadoux, M., Mahon, C., Beattie, J.H., Wallace, H.M., and Hesketh, J.E. (1999). Nuclear import of metallothionein requires its mRNA to be associated with the perinuclear cytoskeleton. *J Biol Chem* 274, 34961-34966.
- Levi, V., Serpinskaya, A.S., Gratton, E., and Gelfand, V. (2006). Organelle transport along microtubules in *Xenopus melanophores*: evidence for cooperation between multiple motors. *Biophys J* 90, 318-327.
- Lewis, R.A., and Mowry, K.L. (2007). Ribonucleoprotein remodeling during RNA localization. *Differentiation* 75, 507-518.
- Li, P., Yang, X., Wasser, M., Cai, Y., and Chia, W. (1997). Inscuteable and Staufen mediate asymmetric localization and segregation of prospero RNA during *Drosophila* neuroblast cell divisions. *Cell* 90, 437-447.
- Lipshitz, H.D., and Smibert, C.A. (2000). Mechanisms of RNA localization and translational regulation. *Curr Opin Genet Dev* 10, 476-488.
- Lodish, H., Berk, A., Zipursky, L.S., Matsudaira, P., Baltimore, D., and Darnell, J. (2000). *Molecular Cell Biology*, Fourth edition edn (W. H. FREEMAN).
- Long, R.M., Gu, W., Lorimer, E., Singer, R.H., and Chartrand, P. (2000). She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA. *Embo J* 19, 6592-6601.
- Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyth, K., and Jansen, R.P. (1997). Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* 277, 383-387.
- Luck, R., Graf, S., and Steger, G. (1999). ConStruct: a tool for thermodynamic controlled prediction of conserved secondary structure. *Nucleic Acids Res* 27, 4208-4217.
- Luk, S.K., Kilpatrick, M., Kerr, K., and Macdonald, P.M. (1994). Components acting in localization of bicoid mRNA are conserved among *Drosophila* species. *Genetics* 137, 521-530.
- Lukavsky, P.J., and Puglisi, J.D. (2005). Structure determination of large biological RNAs. *Methods Enzymol* 394, 399-416.
- Macchi, P., Kroening, S., Palacios, I.M., Baldassa, S., Grunewald, B., Ambrosino, C., Goetze, B., Lupas, A., St Johnston, D., and Kiebler, M. (2003). Barentsz, a new component of the Staufen-containing ribonucleoprotein particles in mammalian cells, interacts with Staufen in an RNA-dependent manner. *J Neurosci* 23, 5778-5788.
- MacDonald, P.M. (1990). bicoid mRNA localization signal: phylogenetic conservation of function and RNA secondary structure. *Development* 110, 161-171.

- 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 Kerr, K. (1998). Mutational analysis of an RNA recognition element that mediates localization of bicoid mRNA. *Mol Cell Biol* *18*, 3788-3795.
- Macdonald, P.M., and Struhl, G. (1986). A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern. *Nature* *324*, 537-545.
- Macdonald, P.M., and Struhl, G. (1988). cis-acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos. *Nature* *336*, 595-598.
- MacDougall, N., Clark, A., MacDougall, E., and Davis, I. (2003). *Drosophila* gurken (TGF $\alpha$ ) mRNA localizes as particles that move within the oocyte in two dynein-dependent steps. *Dev Cell* *4*, 307-319.
- Mach, J.M., and Lehmann, R. (1997). An Egalitarian-BicaudalD complex is essential for oocyte specification and axis determination in *Drosophila*. *Genes Dev* *11*, 423-435.
- Machesky, L.M., and Gould, K.L. (1999). The Arp2/3 complex: a multifunctional actin organizer. *Curr Opin Cell Biol* *11*, 117-121.
- Mahon, P., Partridge, K., Beattie, J.H., Glover, L.A., and Hesketh, J.E. (1997). The 3' untranslated region plays a role in the targeting of metallothionein-I mRNA to the perinuclear cytoplasm and cytoskeletal-bound polysomes. *Biochim Biophys Acta* *1358*, 153-162.
- Mahone, M., Saffman, E.E., and Lasko, P.F. (1995). Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. *Embo J* *14*, 2043-2055.
- Mahowald, A.P. (2001). Assembly of the *Drosophila* germ plasm. *Int Rev Cytol* *203*, 187-213.
- Mallardo, M., Deitinghoff, A., Muller, J., Goetze, B., Macchi, P., Peters, C., and Kiebler, M.A. (2003). Isolation and characterization of Staufen-containing ribonucleoprotein particles from rat brain. *Proc Natl Acad Sci U S A* *100*, 2100-2105.
- Manseau, L.J., and Schupbach, T. (1989). cappuccino and spire: two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the *Drosophila* embryo. *Genes Dev* *3*, 1437-1452.
- Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., and Jacq, C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep* *3*, 159-164.

- Margeot, A., Blugeon, C., Sylvestre, J., Vialette, S., Jacq, C., and Corral-Debrinski, M. (2002). In *Saccharomyces cerevisiae*, ATP2 mRNA sorting to the vicinity of mitochondria is essential for respiratory function. *Embo J* 21, 6893-6904.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719-730.
- Matanis, T., Akhmanova, A., Wulf, P., Del Nery, E., Weide, T., Stepanova, T., Galjart, N., Grosveld, F., Goud, B., De Zeeuw, C.I., *et al.* (2002). Bicaudal-D regulates COPI-independent Golgi-ER transport by recruiting the dynein-dynactin motor complex. *Nat Cell Biol* 4, 986-992.
- Matsuzaki, F., Ohshiro, T., Ikeshima-Kataoka, H., and Izumi, H. (1998). *miranda* localizes *stau1* and *prospero* asymmetrically in mitotic neuroblasts and epithelial cells in early *Drosophila* embryogenesis. *Development* 125, 4089-4098.
- Matus, A. (2000). Actin-based plasticity in dendritic spines. *Science* 290, 754-758.
- McGrail, M., and Hays, T.S. (1997). The microtubule motor cytoplasmic dynein is required for spindle orientation during germline cell divisions and oocyte differentiation in *Drosophila*. *Development* 124, 2409-2419.
- Meier, T., Masciulli, F., Moore, C., Schoumacher, F., Eppenberger, U., Denzer, A.J., Jones, G., and Brenner, H.R. (1998). Agrin can mediate acetylcholine receptor gene expression in muscle by aggregation of muscle-derived neuregulins. *J Cell Biol* 141, 715-726.
- Melton, D.A. (1987). Translocation of a localized maternal mRNA to the vegetal pole of *Xenopus* oocytes. *Nature* 328, 80-82.
- Mickleburgh, I., Burtle, B., Nury, D., Chabanon, H., Chrzanowska-Lightowlers, Z., and Hesketh, J.E. (2004). Isolation and identification of a protein binding to the localization element of Metallothionein-1 mRNA. *Biochem Soc Trans* 32, 705-706.
- Mickleburgh, I., Chabanon, H., Nury, D., Fan, K., Burtle, B., Chrzanowska-Lightowlers, Z., and Hesketh, J. (2006). Elongation factor 1alpha binds to the region of the metallothionein-1 mRNA implicated in perinuclear localization--importance of an internal stem-loop. *Rna* 12, 1397-1407.
- Micklem, D.R., Adams, J., Grunert, S., and St Johnston, D. (2000). Distinct roles of two conserved *Staufen* domains in *oskar* mRNA localization and translation. *Embo J* 19, 1366-1377.
- Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biol* 3, REVIEWS0004.

- Mili, S., Moissoglu, K., and Macara, I.G. (2008). Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* *453*, 115-119.
- Mingle, L.A., Okuhama, N.N., Shi, J., Singer, R.H., Condeelis, J., and Liu, G. (2005). Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. *J Cell Sci* *118*, 2425-2433.
- Mische, S., Li, M., Serr, M., and Hays, T.S. (2007). Direct observation of regulated ribonucleoprotein transport across the nurse cell/oocyte boundary. *Mol Biol Cell* *18*, 2254-2263.
- Miyashiro, K.Y., Beckel-Mitchener, A., Purk, T.P., Becker, K.G., Barret, T., Liu, L., Carbonetto, S., Weiler, I.J., Greenough, W.T., and Eberwine, J. (2003). RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice. *Neuron* *37*, 417-431.
- Momose, T., Derelle, R., and Houliston, E. (2008). A maternally localised Wnt ligand required for axial patterning in the cnidarian *Clytia hemisphaerica*. *Development* *135*, 2105-2113.
- Momose, T., and Houliston, E. (2007). Two oppositely localised frizzled RNAs as axis determinants in a cnidarian embryo. *PLoS Biol* *5*, e70.
- Montell, D.J. (1999). The genetics of cell migration in *Drosophila melanogaster* and *Caenorhabditis elegans* development. *Development* *126*, 3035-3046.
- Monzo, K., Papoulas, O., Cantin, G.T., Wang, Y., Yates, J.R., 3rd, and Sisson, J.C. (2006). Fragile X mental retardation protein controls trailer hitch expression and cleavage furrow formation in *Drosophila* embryos. *Proc Natl Acad Sci U S A* *103*, 18160-18165.
- Mori, Y., Imaizumi, K., Katayama, T., Yoneda, T., and Tohyama, M. (2000). Two cis-acting elements in the 3' untranslated region of alpha-CaMKII regulate its dendritic targeting. *Nat Neurosci* *3*, 1079-1084.
- Mowry, K.L., and Melton, D.A. (1992). Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in *Xenopus* oocytes. *Science* *255*, 991-994.
- Muller, M.J., Klumpp, S., and Lipowsky, R. (2008). Tug-of-war as a cooperative mechanism for bidirectional cargo transport by molecular motors. *Proc Natl Acad Sci U S A* *105*, 4609-4614.
- Munchow, S., Sauter, C., and Jansen, R.P. (1999). Association of the class V myosin Myo4p with a localised messenger RNA in budding yeast depends on She proteins. *J Cell Sci* *112 ( Pt 10)*, 1511-1518.

Mutsuddi, M., Marshall, C.M., Benzow, K.A., Koob, M.D., and Rebay, I. (2004). The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with stauferin in *Drosophila*. *Curr Biol* 14, 302-308.

Napoli, I., Mercaldo, V., Boyl, P.P., Eleuteri, B., Zalfa, F., De Rubeis, S., Di Marino, D., Mohr, E., Massimi, M., Falconi, M., *et al.* (2008). The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* 134, 1042-1054.

Navarro, C., Puthalakath, H., Adams, J.M., Strasser, A., and Lehmann, R. (2004). Egalitarian binds dynein light chain to establish oocyte polarity and maintain oocyte fate. *Nat Cell Biol* 6, 427-435.

Neuman-Silberberg, F.S., and Schupbach, T. (1993). The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell* 75, 165-174.

Nilson, L.A., and Schupbach, T. (1999). EGF receptor signaling in *Drosophila* oogenesis. *Curr Top Dev Biol* 44, 203-243.

Norvell, A., Kelley, R.L., Wehr, K., and Schupbach, T. (1999). Specific isoforms of squid, a *Drosophila* hnRNP, perform distinct roles in *Gurken* localization during oogenesis. *Genes Dev* 13, 864-876.

Nury, D., Chabanon, H., Levadoux-Martin, M., and Hesketh, J. (2005). An eleven nucleotide section of the 3'-untranslated region is required for perinuclear localization of rat metallothionein-1 mRNA. *Biochem J* 387, 419-428.

Oh, J., and Steward, R. (2001). Bicaudal-D is essential for egg chamber formation and cytoskeletal organization in *Drosophila* oogenesis. *Dev Biol* 232, 91-104.

Ohashi, S., Koike, K., Omori, A., Ichinose, S., Ohara, S., Kobayashi, S., Sato, T.A., and Anzai, K. (2002). Identification of mRNA/protein (mRNP) complexes containing Puralpha, mStaufen, fragile X protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor. *J Biol Chem* 277, 37804-37810.

Olesnicky, E.C., and Desplan, C. (2007). Distinct mechanisms for mRNA localization during embryonic axis specification in the wasp *Nasonia*. *Dev Biol* 306, 134-142.

Oleynikov, Y., and Singer, R.H. (1998). RNA localization: different zipcodes, same postman? *Trends Cell Biol* 8, 381-383.

Oleynikov, Y., and Singer, R.H. (2003). Real-time visualization of ZBP1 association with beta-actin mRNA during transcription and localization. *Curr Biol* 13, 199-207.

Olivier, C., Poirier, G., Gendron, P., Boisgontier, A., Major, F., and Chartrand, P. (2005). Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud. *Mol Cell Biol* 25, 4752-4766.

Palacios, I.M. (2002). RNA processing: splicing and the cytoplasmic localisation of mRNA. *Curr Biol* 12, R50-52.

Palacios, I.M. (2007). How does an mRNA find its way? Intracellular localisation of transcripts. *Semin Cell Dev Biol* 18, 163-170.

Palacios, I.M., Gatfield, D., St Johnston, D., and Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature* 427, 753-757.

Pan, F., Huttelmaier, S., Singer, R.H., and Gu, W. (2007). ZBP2 facilitates binding of ZBP1 to beta-actin mRNA during transcription. *Mol Cell Biol* 27, 8340-8351.

Pankratz, M.J., and Jackle, H. (1990). Making stripes in the *Drosophila* embryo. *Trends Genet* 6, 287-292.

Pantaloni, D., Le Clainche, C., and Carlier, M.F. (2001). Mechanism of actin-based motility. *Science* 292, 1502-1506.

Paquin, N., and Chartrand, P. (2008). Local regulation of mRNA translation: new insights from the bud. *Trends Cell Biol* 18, 105-111.

Paushkin, S., Gubitz, A.K., Massenet, S., and Dreyfuss, G. (2002). The SMN complex, an assemblysome of ribonucleoproteins. *Curr Opin Cell Biol* 14, 305-312.

Pieretti, M., Zhang, F.P., Fu, Y.H., Warren, S.T., Oostra, B.A., Caskey, C.T., and Nelson, D.L. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66, 817-822.

Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature* 372, 68-74.

Pokrywka, N.J., and Stephenson, E.C. (1991). Microtubules mediate the localization of bicoid RNA during *Drosophila* oogenesis. *Development* 113, 55-66.

Pokrywka, N.J., and Stephenson, E.C. (1995). Microtubules are a general component of mRNA localization systems in *Drosophila* oocytes. *Dev Biol* 167, 363-370.

Pollard, T.D., and Borisy, G.G. (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453-465.

Prost, E., Deryckere, F., Roos, C., Haenlin, M., Pantesco, V., and Mohier, E. (1988). Role of the oocyte nucleus in determination of the dorsoventral polarity of *Drosophila* as revealed by molecular analysis of the K10 gene. *Genes Dev* 2, 891-900.

- Quintin, S., Gally, C., and Labouesse, M. (2008). Epithelial morphogenesis in embryos: asymmetries, motors and brakes. *Trends Genet* 24, 221-230.
- Rackham, O., and Brown, C.M. (2004). Visualization of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs. *Embo J* 23, 3346-3355.
- Ramos, A., Bayer, P., and Varani, G. (1999). Determination of the structure of the RNA complex of a double-stranded RNA-binding domain from *Drosophila* Staufen protein. *Biopolymers* 52, 181-196.
- Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Staufen double-stranded RNA-binding domain. *Embo J* 19, 997-1009.
- Ran, B., Bopp, R., and Suter, B. (1994). Null alleles reveal novel requirements for Bic-D during *Drosophila* oogenesis and zygotic development. *Development* 120, 1233-1242.
- Rana, T.M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 8, 23-36.
- Rangan, P., DeGennaro, M., Jaime-Bustamante, K., Coux, R.X., Martinho, R.G., and Lehmann, R. (2009). Temporal and spatial control of germ-plasm RNAs. *Curr Biol* 19, 72-77.
- Ranum, L.P., and Cooper, T.A. (2006). RNA-mediated neuromuscular disorders. *Annu Rev Neurosci* 29, 259-277.
- Ranum, L.P., and Day, J.W. (2004). Pathogenic RNA repeats: an expanding role in genetic disease. *Trends Genet* 20, 506-512.
- 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.
- Reddy, K.K., Oitomen, F.M., Patel, G.P., and Bag, J. (2005). Perinuclear localization of slow troponin C m RNA in muscle cells is controlled by a cis-element located at its 3' untranslated region. *Rna* 11, 294-307.
- Reeve, S.P., Bassetto, L., Genova, G.K., Kleyner, Y., Leyssen, M., Jackson, F.R., and Hassan, B.A. (2005). The *Drosophila* fragile X mental retardation protein controls actin dynamics by directly regulating profilin in the brain. *Curr Biol* 15, 1156-1163.
- Reilein, A.R., Serpinskaya, A.S., Karcher, R.L., Dujardin, D.L., Vallee, R.B., and Gelfand, V.I. (2003). Differential regulation of dynein-driven melanosome movement. *Biochem Biophys Res Commun* 309, 652-658.
- Riechmann, V., and Ephrussi, A. (2001). Axis formation during *Drosophila* oogenesis. *Curr Opin Genet Dev* 11, 374-383.

Rivas, E., and Eddy, S.R. (2000). Secondary structure alone is generally not statistically significant for the detection of noncoding RNAs. *Bioinformatics* *16*, 583-605.

Rodriguez, A.J., Condeelis, J., Singer, R.H., and Dichtenberg, J.B. (2007). Imaging mRNA movement from transcription sites to translation sites. *Semin Cell Dev Biol* *18*, 202-208.

Rodriguez, A.J., Shenoy, S.M., Singer, R.H., and Condeelis, J. (2006). Visualization of mRNA translation in living cells. *J Cell Biol* *175*, 67-76.

Roegiers, F., and Jan, Y.N. (2000). Staufen: a common component of mRNA transport in oocytes and neurons? *Trends Cell Biol* *10*, 220-224.

Romero, D.P., and Blackburn, E.H. (1991). A conserved secondary structure for telomerase RNA. *Cell* *67*, 343-353.

Ross, A.F., Oleynikov, Y., Kislauskis, E.H., Taneja, K.L., and Singer, R.H. (1997). Characterization of a beta-actin mRNA zipcode-binding protein. *Mol Cell Biol* *17*, 2158-2165.

Ross, J.L., Ali, M.Y., and Warshaw, D.M. (2008). Cargo transport: molecular motors navigate a complex cytoskeleton. *Curr Opin Cell Biol* *20*, 41-47.

Rossoll, W., Jablonka, S., Andreassi, C., Kroning, A.K., Karle, K., Monani, U.R., and Sendtner, M. (2003). Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *J Cell Biol* *163*, 801-812.

Rossoll, W., Kroning, A.K., Ohndorf, U.M., Steegborn, C., Jablonka, S., and Sendtner, M. (2002). Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Hum Mol Genet* *11*, 93-105.

Roth, S., and Schupbach, T. (1994). The relationship between ovarian and embryonic dorsoventral patterning in *Drosophila*. *Development* *120*, 2245-2257.

Ruohola, H., Bremer, K.A., Baker, D., Swedlow, J.R., Jan, L.Y., and Jan, Y.N. (1991). Role of neurogenic genes in establishment of follicle cell fate and oocyte polarity during oogenesis in *Drosophila*. *Cell* *66*, 433-449.

Saunders, C., and Cohen, R.S. (1999). The role of oocyte transcription, the 5'UTR, and translation repression and derepression in *Drosophila* *gurken* mRNA and protein localization. *Mol Cell* *3*, 43-54.

Schafer, B.W., and Perriard, J.C. (1988). Intracellular targeting of isoproteins in muscle cytoarchitecture. *J Cell Biol* *106*, 1161-1170.

Schmid, M., Jaedicke, A., Du, T.G., and Jansen, R.P. (2006). Coordination of endoplasmic reticulum and mRNA localization to the yeast bud. *Curr Biol* *16*, 1538-1543.

Schnitzer, M.J., Visscher, K., and Block, S.M. (2000). Force production by single kinesin motors. *Nat Cell Biol* *2*, 718-723.

Schuldt, A.J., Adams, J.H., Davidson, C.M., Micklem, D.R., Haseloff, J., St Johnston, D., and Brand, A.H. (1998). Miranda mediates asymmetric protein and RNA localization in the developing nervous system. *Genes Dev* *12*, 1847-1857.

Schupbach, T. (1987). Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in *Drosophila melanogaster*. *Cell* *49*, 699-707.

Schupbach, T., and Wieschaus, E. (1991). Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* *129*, 1119-1136.

Semotok, J.L., Cooperstock, R.L., Pinder, B.D., Vari, H.K., Lipshitz, H.D., and Smibert, C.A. (2005). Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early *Drosophila* embryo. *Curr Biol* *15*, 284-294.

Semotok, J.L., Luo, H., Cooperstock, R.L., Karaiskakis, A., Vari, H.K., Smibert, C.A., and Lipshitz, H.D. (2008). *Drosophila* maternal Hsp83 mRNA destabilization is directed by multiple SMAUG recognition elements in the open reading frame. *Mol Cell Biol* *28*, 6757-6772.

Serano, J., and Rubin, G.M. (2003). The *Drosophila* synaptotagmin-like protein bitesize is required for growth and has mRNA localization sequences within its open reading frame. *Proc Natl Acad Sci U S A* *100*, 13368-13373.

Serano, T.L., Cheung, H.K., Frank, L.H., and Cohen, R.S. (1994). P element transformation vectors for studying *Drosophila melanogaster* oogenesis and early embryogenesis. *Gene* *138*, 181-186.

Serano, T.L., and Cohen, R.S. (1995a). Gratuitous mRNA localization in the *Drosophila* oocyte. *Development* *121*, 3013-3021.

Serano, T.L., and Cohen, R.S. (1995b). A small predicted stem-loop structure mediates oocyte localization of *Drosophila* K10 mRNA. *Development* *121*, 3809-3818.

Serano, T.L., Karlin-McGinness, M., and Cohen, R.S. (1995). The role of fs(1)K10 in the localization of the mRNA of the TGF alpha homolog gurken within the *Drosophila* oocyte. *Mech Dev* *51*, 183-192.

Shah, J.V., Flanagan, L.A., Janmey, P.A., and Leterrier, J.F. (2000). Bidirectional translocation of neurofilaments along microtubules mediated in part by dynein/dynactin. *Mol Biol Cell* *11*, 3495-3508.

Sharp, D.J., Brown, H.M., Kwon, M., Rogers, G.C., Holland, G., and Scholey, J.M. (2000). Functional coordination of three mitotic motors in *Drosophila* embryos. *Mol Biol Cell* *11*, 241-253.

Shaw, N.N., and Arya, D.P. (2008). Recognition of the unique structure of DNA:RNA hybrids. *Biochimie* *90*, 1026-1039.

Shen, C.P., Knoblich, J.A., Chan, Y.M., Jiang, M.M., Jan, L.Y., and Jan, Y.N. (1998). Miranda as a multidomain adapter linking apically localized Inscuteable and basally localized Staufen and Prospero during asymmetric cell division in *Drosophila*. *Genes Dev* *12*, 1837-1846.

Shepard, K.A., Gerber, A.P., Jambhekar, A., Takizawa, P.A., Brown, P.O., Herschlag, D., DeRisi, J.L., and Vale, R.D. (2003). Widespread cytoplasmic mRNA transport in yeast: identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc Natl Acad Sci U S A* *100*, 11429-11434.

Shestakova, E.A., Singer, R.H., and Condeelis, J. (2001). The physiological significance of beta -actin mRNA localization in determining cell polarity and directional motility. *Proc Natl Acad Sci U S A* *98*, 7045-7050.

Shubeita, G.T., Tran, S.L., Xu, J., Vershinin, M., Cermelli, S., Cotton, S.L., Welte, M.A., and Gross, S.P. (2008). Consequences of motor copy number on the intracellular transport of kinesin-1-driven lipid droplets. *Cell* *135*, 1098-1107.

Silhavy, T.J., Benson, S.A., and Emr, S.D. (1983). Mechanisms of protein localization. *Microbiol Rev* *47*, 313-344.

Simmonds, A.J., dosSantos, G., Livne-Bar, I., and Krause, H.M. (2001). Apical localization of wingless transcripts is required for wingless signaling. *Cell* *105*, 197-207.

Smibert, C.A., Lie, Y.S., Shillinglaw, W., Henzel, W.J., and Macdonald, P.M. (1999). Smaug, a novel and conserved protein, contributes to repression of nanos mRNA translation in vitro. *Rna* *5*, 1535-1547.

Snee, M.J., Arn, E.A., Bullock, S.L., and Macdonald, P.M. (2005). Recognition of the bcd mRNA localization signal in *Drosophila* embryos and ovaries. *Mol Cell Biol* *25*, 1501-1510.

Spradling, A. (1993). *The Development of Drosophila melanogaster* (Cold Spring Harbor Laboratory Press, New York,).

St Johnston, D. (2005). Moving messages: the intracellular localization of mRNAs. *Nat Rev Mol Cell Biol* 6, 363-375.

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

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

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

Starr, D.A. (2007). Communication between the cytoskeleton and the nuclear envelope to position the nucleus. *Mol Biosyst* 3, 583-589.

Stroupe, M.E., Tange, T.O., Thomas, D.R., Moore, M.J., and Grigorieff, N. (2006). The three-dimensional architecture of the EJC core. *J Mol Biol* 360, 743-749.

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

Sullivan, W., and Theurkauf, W.E. (1995). The cytoskeleton and morphogenesis of the early *Drosophila* embryo. *Curr Opin Cell Biol* 7, 18-22.

Sundell, C.L., and Singer, R.H. (1991). Requirement of microfilaments in sorting of actin messenger RNA. *Science* 253, 1275-1277.

Suter, B., Romberg, L.M., and Steward, R. (1989). *Bicaudal-D*, a *Drosophila* gene involved in developmental asymmetry: localized transcript accumulation in ovaries and sequence similarity to myosin heavy chain tail domains. *Genes Dev* 3, 1957-1968.

Svoboda, P., and Di Cara, A. (2006). Hairpin RNA: a secondary structure of primary importance. *Cell Mol Life Sci* 63, 901-908.

Swan, A., Nguyen, T., and Suter, B. (1999). *Drosophila* Lissencephaly-1 functions with *Bic-D* and dynein in oocyte determination and nuclear positioning. *Nat Cell Biol* 1, 444-449.

Sylvestre, J., Vialette, S., Corral Debrinski, M., and Jacq, C. (2003). Long mRNAs coding for yeast mitochondrial proteins of prokaryotic origin preferentially localize to the vicinity of mitochondria. *Genome Biol* 4, R44.

Tadros, W., Goldman, A.L., Babak, T., Menzies, F., Vardy, L., Orr-Weaver, T., Hughes, T.R., Westwood, J.T., Smibert, C.A., and Lipshitz, H.D. (2007). *SMAUG* is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase. *Dev Cell* 12, 143-155.

- Tadros, W., and Lipshitz, H.D. (2005). Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in *Drosophila*. *Dev Dyn* 232, 593-608.
- Takizawa, P.A., Sil, A., Swedlow, J.R., Herskowitz, I., and Vale, R.D. (1997). Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* 389, 90-93.
- Takizawa, P.A., and Vale, R.D. (2000). The myosin motor, Myo4p, binds Ash1 mRNA via the adapter protein, She3p. *Proc Natl Acad Sci U S A* 97, 5273-5278.
- Tekotte, H., and Davis, I. (2002). Intracellular mRNA localization: motors move messages. *Trends Genet* 18, 636-642.
- Tepass, U., Tanentzapf, G., Ward, R., and Fehon, R. (2001). Epithelial cell polarity and cell junctions in *Drosophila*. *Annu Rev Genet* 35, 747-784.
- 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.
- Theurkauf, W.E., Smiley, S., Wong, M.L., and Alberts, B.M. (1992). Reorganization of the cytoskeleton during *Drosophila* oogenesis: implications for axis specification and intercellular transport. *Development* 115, 923-936.
- Thio, G.L., Ray, R.P., Barcelo, G., and Schupbach, T. (2000). Localization of gurken RNA in *Drosophila* oogenesis requires elements in the 5' and 3' regions of the transcript. *Dev Biol* 221, 435-446.
- Tiruchinapalli, D.M., Oleynikov, Y., Kelic, S., Shenoy, S.M., Hartley, A., Stanton, P.K., Singer, R.H., and Bassell, G.J. (2003). Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and beta-actin mRNA in dendrites and spines of hippocampal neurons. *J Neurosci* 23, 3251-3261.
- Turner, D.H. (1996). Thermodynamics of base pairing. *Curr Opin Struct Biol* 6, 299-304.
- Tyagi, S., Bratu, D.P., and Kramer, F.R. (1998). Multicolor molecular beacons for allele discrimination. *Nat Biotechnol* 16, 49-53.
- Tyagi, S., and Kramer, F.R. (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 14, 303-308.
- Tzakos, A.G., Grace, C.R., Lukavsky, P.J., and Riek, R. (2006). NMR techniques for very large proteins and rnas in solution. *Annu Rev Biophys Biomol Struct* 35, 319-342.
- Uhlenbeck, O.C., Pardi, A., and Feigon, J. (1997). RNA structure comes of age. *Cell* 90, 833-840.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 37, 376-386.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212-1215.

Vale, R.D., and Milligan, R.A. (2000). The way things move: looking under the hood of molecular motor proteins. *Science* 288, 88-95.

Van De Bor, V., Hartswood, E., Jones, C., Finnegan, D., and Davis, I. (2005). *gurken* and the I factor retrotransposon RNAs share common localization signals and machinery. *Dev Cell* 9, 51-62.

Varani, G., and McClain, W.H. (2000). The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep* 1, 18-23.

Veneri, M., Zalfa, F., and Bagni, C. (2004). FMRP and its target RNAs: fishing for the specificity. *Neuroreport* 15, 2447-2450.

Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P., *et al.* (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65, 905-914.

Veyrune, J.L., Campbell, G.P., Wiseman, J., Blanchard, J.M., and Hesketh, J.E. (1996). A localisation signal in the 3' untranslated region of c-myc mRNA targets c-myc mRNA and beta-globin reporter sequences to the perinuclear cytoplasm and cytoskeletal-bound polysomes. *J Cell Sci* 109 ( Pt 6), 1185-1194.

Veyrune, J.L., Hesketh, J., and Blanchard, J.M. (1997). 3' untranslated regions of c-myc and c-fos mRNAs: multifunctional elements regulating mRNA translation, degradation and subcellular localization. *Prog Mol Subcell Biol* 18, 35-63.

Wagner, C., Ehresmann, C., Ehresmann, B., and Brunel, C. (2004). Mechanism of dimerization of bicoid mRNA: initiation and stabilization. *J Biol Chem* 279, 4560-4569.

Wang, C., and Lehmann, R. (1991). Nanos is the localized posterior determinant in *Drosophila*. *Cell* 66, 637-647.

Wang, S., and Hazelrigg, T. (1994). Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila* oogenesis. *Nature* 369, 400-403.

Weeks, D.L., Rebagliati, M.R., Harvey, R.P., and Melton, D.A. (1985). Localized maternal mRNAs in *Xenopus laevis* eggs. *Cold Spring Harb Symp Quant Biol* 50, 21-30.

- Weeks, K.M., and Crothers, D.M. (1993). Major groove accessibility of RNA. *Science* *261*, 1574-1577.
- Weil, T.T., Forrest, K.M., and Gavis, E.R. (2006). Localization of bicoid mRNA in late oocytes is maintained by continual active transport. *Dev Cell* *11*, 251-262.
- Welte, M.A. (2004). Bidirectional transport along microtubules. *Curr Biol* *14*, R525-537.
- Welte, M.A., Gross, S.P., Postner, M., Block, S.M., and Wieschaus, E.F. (1998). Developmental regulation of vesicle transport in *Drosophila* embryos: forces and kinetics. *Cell* *92*, 547-557.
- Westhof, E., and Fritsch, V. (2000). RNA folding: beyond Watson-Crick pairs. *Structure Fold Des* *8*, R55-65.
- Whittaker, G.R., Kann, M., and Helenius, A. (2000). Viral entry into the nucleus. *Annu Rev Cell Dev Biol* *16*, 627-651.
- Wilhelm, J.E., Mansfield, J., Hom-Booher, N., Wang, S., Turck, C.W., Hazelrigg, T., and Vale, R.D. (2000). Isolation of a ribonucleoprotein complex involved in mRNA localization in *Drosophila* oocytes. *J Cell Biol* *148*, 427-440.
- Wilkie, G.S., and Davis, I. (2001). *Drosophila* wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles. *Cell* *105*, 209-219.
- Wilson, I.A., Brindle, K.M., and Fulton, A.M. (1995). Differential localization of the mRNA of the M and B isoforms of creatine kinase in myoblasts. *Biochem J* *308* ( Pt 2), 599-605.
- Wodarz, A., and Nusse, R. (1998). Mechanisms of Wnt signaling in development. *Annu Rev Cell Dev Biol* *14*, 59-88.
- Woese, C.R., Winker, S., and Gutell, R.R. (1990). Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc Natl Acad Sci U S A* *87*, 8467-8471.
- Yano, T., Lopez de Quinto, S., Matsui, Y., Shevchenko, A., Shevchenko, A., and Ephrussi, A. (2004). Hrp48, a *Drosophila* hnRNPA/B homolog, binds and regulates translation of oskar mRNA. *Dev Cell* *6*, 637-648.
- Yoon, C., Kawakami, K., and Hopkins, N. (1997). Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. *Development* *124*, 3157-3165.
- Yoon, Y.J., and Mowry, K.L. (2004). *Xenopus* Staufen is a component of a ribonucleoprotein complex containing Vg1 RNA and kinesin. *Development* *131*, 3035-3045.

- Yu, F., Kuo, C.T., and Jan, Y.N. (2006). *Drosophila* neuroblast asymmetric cell division: recent advances and implications for stem cell biology. *Neuron* 51, 13-20.
- Zaessinger, S., Busseau, I., and Simonelig, M. (2006). Oskar allows nanos mRNA translation in *Drosophila* embryos by preventing its deadenylation by Smaug/CCR4. *Development* 133, 4573-4583.
- Zalfa, F., Adinolfi, S., Napoli, I., Kuhn-Holsken, E., Urlaub, H., Achsel, T., Pastore, A., and Bagni, C. (2005). Fragile X mental retardation protein (FMRP) binds specifically to the brain cytoplasmic RNAs BC1/BC200 via a novel RNA-binding motif. *J Biol Chem* 280, 33403-33410.
- Zalfa, F., Eleuteri, B., Dickson, K.S., Mercaldo, V., De Rubeis, S., di Penta, A., Tabolacci, E., Chiurazzi, P., Neri, G., Grant, S.G., *et al.* (2007). A new function for the fragile X mental retardation protein in regulation of PSD-95 mRNA stability. *Nat Neurosci* 10, 578-587.
- Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B., and Bagni, C. (2003). The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell* 112, 317-327.
- Zhang, J., Houston, D.W., King, M.L., Payne, C., Wylie, C., and Heasman, J. (1998a). The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell* 94, 515-524.
- Zhang, Y., Buchholz, F., Muirers, J.P., and Stewart, A.F. (1998b). A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* 20, 123-128.
- Zimyanin, V.L., Belaya, K., Pecreaux, J., Gilchrist, M.J., Clark, A., Davis, I., and St Johnston, D. (2008). In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell* 134, 843-853.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406-3415.