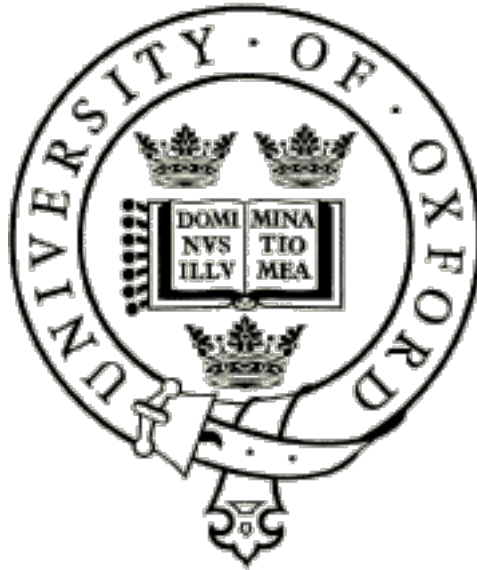


Elucidating the molecular mechanism that determines
the specific localisation of *gurken* mRNA during
Drosophila development



St Hugh's College, University of Oxford
Department of Biochemistry
BBSRC CASE Studentship

A thesis submitted for the degree of Doctor of Philosophy

Kirsty Gill
Supervisor: Professor Ilan Davis
49, 077 words
Hilary Term 2017

Elucidating the molecular mechanism that determines the specific localisation of *gurken* mRNA during *Drosophila* development

Kirsty Gill, St Hugh's, DPhil in Biochemistry, Trinity Term 2016

Abstract

mRNA localisation is a widely used mechanism for achieving temporal-spatial restriction of protein expression and is essential during development to establish cell polarity. mRNA localisation is particularly well studied in the *Drosophila* egg chamber where *gurken* mRNA is localised to the dorsal-anterior corner of the oocyte in a Dynein-dependent process that establishes the anterior-posterior and dorsal-ventral axes of the future embryo. An RNA stem-loop called the *gurken* localisation signal is necessary and sufficient to drive *gurken* localisation through interactions with a specific complement of protein factors. However, the exact RNA sequence and structural features required to promote each stage of *gurken* localisation are unknown. Using a live-cell injection assay I have dissected regions of the mRNA signal that are responsible for driving *gurken* transport and anchoring through their association with Egalitarian, Me31B and Squid proteins. I show the structure of an AU-rich stem and a purine stack are essential for *gurken* transport, and demonstrate that the size of the internal loop between these stems is important. These features of the localisation signal are essential for recruitment of Egalitarian, which links the mRNA to the Dynein transport machinery. I also show that these mRNA sequence and structural elements are present in several other Dynein-transported mRNAs. The bulge at the distal end of the *gurken* localisation signal is important for anchoring *grk* at the dorsal-anterior of the oocyte, possibly through Squid binding, and the proximal third of the signal is essential for recruitment of the translation component Me31B. These studies indicate that the role of the *gurken* localisation signal in controlling *gurken* transport, anchoring and translation can be mapped to distinct regions of the signal and provide insights into how the signal carries out these numerous functions at a molecular level. Determining the molecular interactions involved in mRNA localisation improves our understanding of how specificity is generated to direct different mRNAs to distinct regions of the cell to restrict protein expression.

Declaration

I hereby declare that this thesis is my own work and the experiments described in the following pages were performed by myself, unless otherwise stated. The experiments were carried out in the Department of Biochemistry, University of Oxford. This thesis has not been submitted for any other degree or professional qualification.

The following parts of this thesis were completed by others:

1. Figure 3.6 (A-A'') *in situ* hybridisation images taken from the Dresden Ovary Table (Jambor et al. 2015)
2. Figure 4.20 (mutants M12 and M16 were designed, made and tested by previous members of the Davis Lab: Veronique Van de Bor and Graeme Ball)
3. Figure 4.24 (mutant M3 was designed, made and tested by previous members of the Davis Lab: Veronique Van de Bor and Graeme Ball)
4. Figure 5.1 (*gurken in situ* hybridisation in wild type and *sqd1* mutant egg chambers published in Neuman-Silberberg and Schupbach, 1993)
5. Figure 5.11 (Squid SELEX results published in Blanchette et al. 2009)

Kirsty Gill

08/02/2017

Acknowledgements

I would like to thank my supervisor, Ilan Davis, for his support, encouragement and advice over the last four years. Many thanks are also due to Richard Parton for his time and patience in teaching me new techniques, for his help and support whilst writing this thesis, and for many scientific discussions over the years. I am also thankful to our collaborators Pavel Tomancak, Helena Jambor, Andres Ramos, Russell Hamilton for sharing their data and expertise with us, to Timothy Weil and Simon Bullock for their advice and discussions, and to the BBSRC for funding.

I am extremely grateful to all members of the Davis and Castello labs, past and present, for creating such a friendly environment and supplying exceedingly good cake on a regular basis. I especially want to thank Fran for all her help with biochemistry, and Darragh for always fixing problems and making sure the lab runs smoothly. Thank you to Tamsin, Alex and Lu for always giving me support through the difficult times.

Finally, I would like to thank my husband and my family for their patience, love and continued support throughout the PhD, and for reminding me of the important things in life, whilst encouraging me to never give up on my dreams.

List of Abbreviations

3'UTR	3'-untranslated region
4E-BPs	eIF4E binding proteins
5'UTR	5'-untranslated region
AGO2	Argonaute 2
ASH1	Asymmetric synthesis of HO
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
<i>bcd</i>	<i>bicoid</i>
BicC	Bicaudal-C
BicD	Bicaudal-D
BLE	<i>Bicoid</i> localisation element
BRE	Bruno response element
CamKII α	Calcium/calmodulin-dependent protein kinase alpha
<i>capu</i>	<i>cappuccino</i>
cDNA	Complementary deoxyribonucleic acid
CLIP	Cross-Linking and ImmunoPrecipitation
CPEB	Cytoplasmic polyadenylation element binding protein
CPSF	Cleavage and polyadenylation specificity factor
DA	Dorsal-anterior
Dcp1	Decapping protein 1
Dcp2	Decapping protein 2
DICE	Differential control elements
Dlc	Dynein light chain
DM1	Myotonic dystrophy type 1
DNA	Deoxyribonucleic acid
dsRBD	Double-Stranded RNA-Binding Domain

dsRBP	Double-stranded RNA-binding protein
dsRNA	Double-stranded RNA
EGFR	Epidermal growth factor receptor
Egl	Egalitarian
eIF2	Eukaryotic initiation factor 2
eIF4E	Eukaryotic translation initiation factor 4E
eIF4G	Eukaryotic translation initiation factor 4G
EJC	Exon-junction complex
EM	Electron Microscopy
EMSA	Electrophoretic mobility shift assay
Enc	Encore
ePAB	Embryonic Poly(A)-binding protein
ER	Endoplasmic reticulum
FISH	Fluorescent <i>in situ</i> hybridisation
FMR1	Fragile X mental retardation 1
FMRP	Fragile X mental retardation protein
FRAP	Fluorescent recovery after photobleaching
FXS	Fragile X syndrome
G2LS	G2 localisation signal
GFP	Green fluorescent protein
Gld2	Germline development 2
Glo	Glorund
GLS	<i>gurken</i> localisation signal
<i>grk</i>	<i>gurken</i>
GRNA chromatography	GST-RNA affinity chromatography
GSC	Germline stem cell
<i>hb</i>	<i>hunchback</i>
Heph	Hephaestus
Hfp	Half-pint

HIV	Human Immunodeficiency Virus
hnRNP	Heterogeneous nuclear ribonucleoprotein
Hrb27C	Heterogeneous nuclear ribonucleoprotein at 27C
iCLIP	Individual-nucleotide resolution ultraviolet CLIP
ILS	/ Factor localisation signal
IMP	IGF-II mRNA binding proteins
ISH	<i>in situ</i> hybridisation
JLS	<i>Jockey</i> localisation signal
KH domain	K homology domain
L	Litre
LB broth	Luri Betani broth
LTR	Long terminal repeat
<i>mael</i>	<i>maelstrom</i>
MBP	Myelin basic protein
MCP	MS2 coat protein
MDa	Megadalton
Me31B	Maternal expression at 31B
MFE	Minimum free energy
miRNA	micro-RNA
ml	mililitre
mM	milimolar
mRNA	messenger RNA
MTOC	Microtubule organising centre
MyoV	Myosin V
MZT	Maternal to zygotic transition
NGD	No-go decay
NLS	Nuclear localisation signal
nm	nanometer
NMD	Non-sense mediated decay

NMR	Nuclear magnetic resonance
<i>nos</i>	<i>nanos</i>
NPC	Nuclear pore complex
NSD	Non-stop decay
Orb	Oo18 RNA-binding protein
<i>OrR</i>	<i>Oregon-R</i>
ORF1p	Open reading frame 1 protein
ORF2p	Open reading frame 2 protein
<i>osk</i>	<i>oskar</i>
Otu	Ovarian tumour
P bodies	Processing bodies
PABP	Poly(A)-binding protein
PAGE	Polyacrylamide gel electrophoresis
Paip1	Polyadenylate-binding protein-interacting protein 1
PAP	Poly(A) polymerase
PAR-1	Protease-activated receptor 1
PAR-CLIP	Photo-Activatable Ribonucleoside-CLIP
PARN	Poly(A) ribonuclease
PBS	Phosphate Buffer Saline
<i>pgc</i>	<i>polar granule component</i>
<i>pros</i>	<i>prospero</i>
PTB	Polypyrimidine tract-binding protein
RBP	RNA-binding protein
RIP-Chip	RNP immunoprecipitation-microarray
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RRM	RNA recognition motif
RT	Room temperature

SELEX	Systematic Evolution of Ligands by Exponential Enrichment
siRNA	short interfering RNA
SMA	Spinal muscular atrophy
SOC media	Super Optimal Broth with Catabolite repression media
SOLE	Spliced <i>oskar</i> localization element
<i>spn-F</i>	<i>spindle-F</i>
Squid	Sqd
ssRNA	single-stranded RNA
Stau	Staufen
SWH	Salvador Warts Hippo
Syncrip (Syp)	Synaptotagmin-binding, cytoplasmic, RNA-interacting protein
TAE	Tris-acetate-EDTA
TGF- α	Transforming growth factor alpha
Top	Torpedo
Tral	Trailerhitch
UV	Ultraviolet
VA	Ventral-anterior
<i>wg</i>	<i>wingless</i>
Xrn1	5'-3' exoribonuclease 1
ZBP1	Z-DNA binding protein 1
ZO-1	Zonula occuldens 1
μ l	Microlitre
μ m	Micrometer

Table of Contents

Abstract	ii
Declaration	iii
Acknowledgements	iv
List of Abbreviations	v-ix
Table of Contents	x-xvi
List of Figures	xvi-xx
List of Tables	xxi

Chapter 1: Introduction

The importance and role of mRNA localisation	1
mRNA localisation	1
Functions of mRNA localisation	2
mRNA localisation in disease	4
Mechanisms of mRNA localisation	5
<i>Trans</i> -acting proteins	5-8
<i>Cis</i> -acting regulatory elements	8-9
Types of mRNA localisation	9-10
Active transport and anchoring of mRNAs	10
The cytoskeleton and motor proteins	10-13
The role of motor proteins in mRNA localisation	13-14
Anchoring of localised mRNAs	15-16
Regulation of translation during mRNA localisation	16
Translational silencing during transport of localised mRNAs	16-18
Translational activation of localised mRNAs at their target site	18-19
The role of degradation in mRNA localisation	19-20
<i>grk</i> mRNA localisation	21
<i>Drosophila</i> oogenesis	21
Specification of the <i>Drosophila</i> oocyte	21-25
Development of the <i>Drosophila</i> egg chamber	25-26
<i>grk</i> localisation in the <i>Drosophila</i> oocyte	27-30
The mechanism of <i>grk</i> localisation	30-31

The <i>grk</i> localisation signal (GLS)	31-32
<i>Trans</i> -acting factors involved in <i>grk</i> localisation and translation	32-33
The regulation of <i>grk</i> translation	34
The / Factor	34-37
mRNA localisation in the <i>Drosophila</i> embryo	37
Studying mRNA localisation	
Visualising localised mRNAs	37-38
Characterising mRNA localisation signals	38-39
Studying mRNA-protein interactions	39
The structure of RNAs	39-42
Identifying RNA-binding sites in RNAs	42-45
Key outstanding questions in the field	45
Aims of this thesis	46-48
 <u>Chapter 2: Materials and Methods</u>	
Molecular Biology	49
Solutions and reagents	49
Propagation and preparation of plasmid DNA	50
Cloning	50
DNA sequencing and analysis	50-51
Gel electrophoresis	51
TAE-agarose gel electrophoresis	51
SDS-polyacrylamide gel electrophoresis (SDS-PAGE)	51
Western Blotting	51-52
Preparation of templates for <i>in vitro</i> transcription	52
<i>Drosophila</i> protocols	53
Flystocks	53
Preparation of ovaries for oocyte injection and live imaging	53
Microinjection of oocytes	53
Preparation of ovarian lysate for Western Blotting	53
Preparation of embryos for microinjection	54
Microinjection of embryos	54

Microscopy	54
Widefield live cell imaging	54
Bioinformatics	55

Chapter 3: A rapid assay for a structure-function analysis of the GLS

Introduction	56-57
Newly available genomes should allow further studies of GLS conservation	57
Current injection assays are inadequate for high throughput screening of subtle RNA localisation defects	58-60
Aims of this chapter and experimental rationale	61
Results	62
The GLS primary sequence is highly conserved among <i>Drosophilids</i>	62-65
The predicted GLS secondary structure is highly conserved among <i>Drosophilids</i>	65-68
The GLS primary sequence is conserved in some insects outside of the <i>Drosophilidae</i> family	68-72
The GLS secondary structure is conserved in some insects outside of the <i>Drosophilidae</i> family	72-75
Optimising the injection assay for efficient GLS mutant screening	75-79
Subtle localisation defects can be identified by quantifying fluorescent RNA intensities in defined regions throughout the oocyte	80-89
Discussion	89
The GLS is highly conserved among the <i>Drosophilidae</i> family and in a selection of other Dipterans.	89-90
Current injection assays can be optimised to increase throughput of mutant testing	91
Localisation patterns can be quantitated to assess subtle localisation defects of injected RNAs	91-92
Future work	92-93
Final Conclusions	93

Chapter 4: Regions of the GLS required for *grk* RNA transport

Introduction	94-95
<i>grk</i> RNA is transported in the oocyte along microtubules using the molecular motor Dynein	95-96
Newly identified GLS-like signals should allow effective design of GLS mutations to disrupt transport	96

Recruitment of GLS binding proteins to mutant GLS RNAs can be assessed in living embryos	96-97
Aims of this chapter and experimental rationale	98
Results	99
The most divergent <i>Drosophilid</i> GLS-like signals drive <i>grk</i> transport in the <i>D. melanogaster</i> oocyte	99-100
Several GLS-like signals from Dipterans drive <i>grk</i> transport in the <i>D. melanogaster</i> oocyte	101-106
The conservation of the GLS can be used to design sequence modifications to disrupt <i>grk</i> transport	106-110
The proximal third of the GLS is not essential for <i>grk</i> transport in the oocyte.	110-116
GLS A'-form conformation is unlikely to be required for <i>grk</i> transport	116-120
The length of stem 4 is essential to maintain <i>grk</i> transport	120-126
The AU-richness of stem 4 is essential for <i>grk</i> transport	126-130
The sequence of stem 4 is not essential for driving <i>grk</i> transport but can increase efficiency	130-135
The location and accessibility of the distal stem is essential to maintain <i>grk</i> transport	135-139
The orientation of internal loop 2 is essential for driving <i>grk</i> transport	139-147
The internal bulges are not required to drive <i>grk</i> transport	147-150
Discussion	150
Most of the novel GLS-like signals identified are able to drive <i>grk</i> RNA transport	150
The conservation of the GLS can be used to identify regions important for function	150-153
A model for the recognition of the GLS by the transport machinery	153-154
Future work	156
Final Conclusions	156

Chapter 5: Regions of the GLS required for the DA localisation and anchoring of *grk* RNA

Introduction	157
<i>grk</i> RNA is anchored at the DA of the oocyte	157-158
The role of Sqd during mid-oogenesis	158-159
Aims of this chapter and experimental rationale	160
Results	161
Transport of injected <i>grk</i> RNA to the oocyte anterior is unaffected in a <i>sqd1</i> mutant	161

Injected and MS2-labelled <i>grk</i> RNA forms a strong anterior ring in a <i>sqd1</i> mutant	161-167
<i>grk</i> RNA forms large clusters at the anterior and DA of <i>sqd1</i> oocytes	167-171
<i>grk</i> RNA clusters are statically anchored at the anterior of <i>sqd1</i> oocytes	171-173
<i>grk</i> RNA interdigitates with Me31B at the anterior of <i>sqd1</i> oocytes	174-176
Sqd is enriched at the oocyte anterior and is not recruited to injected <i>grk</i> RNA	176-177
<i>grk</i> RNA injected near the oocyte nucleus rescues the <i>sqd1</i> phenotype	178-179
SELEX data can infer which regions of the GLS may be important for Sqd binding	180
Previous GLS transport mutants can be used to inform anchoring mutant design	180-183
A selection of GLS-like signals from other species show aberrant anchoring in the oocyte	183-184
The GLS-like signal from <i>C. capitata</i> forms large clusters that interdigitate with Me31B at the anterior of the oocyte	184-188
The distal third of the <i>C. capitata</i> GLS-like signal is responsible for aberrant <i>grk</i> localisation and anchoring	188-190
The accessibility produced by bulge 2 is important for restricting <i>grk</i> to the DA of the oocyte	190-194
Discussion	194
<i>grk</i> RNA is localised to both the DA and the anterior of <i>sqd1</i> oocytes	194-196
<i>grk</i> RNA is associated with P bodies at the DA and anterior of <i>sqd1</i> oocytes	196-197
Bulge 2 is required to efficiently restrict <i>grk</i> anchoring to the DA of the oocyte	198-199
Future work	199
Final Conclusions	199-200

Chapter 6: Regions of the GLS required for *grk* RNA association with P bodies

Introduction	201
<i>grk</i> mRNA is translationally silenced during its transport to the DA of the oocyte	201
Translation of <i>grk</i> mRNA occurs at the edge of P bodies	202
P bodies disassemble following egg activation	202-203
Aims of this chapter and experimental rationale	204
Results	205
Injected <i>grk</i> RNA recruits Me31B::GFP to the apical cortex in the embryo	205-207
Me31B is associated with <i>grk</i> RNA at the apical cortex but not in transport particles	207-210

<i>grk</i> RNA co-localises with Me31B at the apical cortex of the embryo	210-211
<i>C. capitata</i> is able to recruit Me31B in the embryo	213
The sequence of stem 4 is not required for Me31B recruitment	214-217
Stem 1 is essential for Me31B recruitment	217-219
The sequence of the internal loops are not essential for Me31B recruitment	219
Bulge 2 is not required for Me31B recruitment	219-223
Discussion	223
<i>grk</i> RNA is actively transported to the apical cortex of the embryo where it recruits Me31B	223-224
Me31B accumulation in the embryo may represent P body-like structures required for the translational regulation of <i>grk</i> RNA	224-225
The proximal third of the GLS is essential for the association between <i>grk</i> RNA and Me31B	225-226
Future work	226-227
Final Conclusions	227

Chapter 7: Discussion and Future Perspectives

GLS-like signals from other species can be used to predict regions of the GLS that are important for <i>grk</i> transport	228
An AU-rich stem, internal loop and purine stack motif may direct recruitment of the key adapter Egl	229-231
The role of the GLS in directing <i>grk</i> anchoring and translation can be mapped to distinct regions of the signal	231-234
Re-interpreting the mechanism of Sqd in <i>grk</i> anchoring and translation	234-237
Translational control of <i>grk</i> in P bodies	237-238
Improving our understanding of how specificity is generated during RNA transport, anchoring and translation	238-240
Final Conclusions	240-242

Chapter 8: References

243-292

Appendix

Appendix A: GLS protein interactors	293-294
Appendix B: Fly strains	294
Appendix C: Imaging parameters and assessing RNA localisation in the oocyte	294-296
Appendix D: Imaging parameters and assessing RNA localisation in the embryo	296

List of Figures

<u>Figure 1.1.</u> mRNA localisation is a widely used mechanism for restricting protein expression	3
<u>Figure 1.2.</u> The stages of mRNA localisation	7
<u>Figure 1.3.</u> Mechanisms of mRNA localisation	11
<u>Figure 1.4</u> Stages of <i>Drosophila</i> oogenesis	23
<u>Figure 1.5</u> Oocyte specification in the <i>Drosophila</i> germarium	24
<u>Figure 1.6</u> mRNA localisation in the <i>Drosophila</i> oocyte	28
<u>Figure 1.7.</u> The GLS recruits several <i>trans</i> -acting factors in the <i>Drosophila</i> ovary and shares conserved features with other DA localising signals	36
<u>Figure 1.8.</u> RNA secondary structure motifs	41
<u>Figure 3.1.</u> Aspirations for a new high-throughput injection assay to screen for subtle localisation defects in mutant RNAs	60
<u>Figure 3.2.</u> The sequence of the GLS is highly conserved among the <i>Drosophilidae</i> family	64
<u>Figure 3.3.</u> The secondary structure of the GLS is highly conserved among the <i>Drosophilidae</i> family	68
<u>Figure 3.4.</u> The GLS shows some primary sequence conservation in other species outside of the <i>Drosophilidae</i> family	71
<u>Figure 3.5.</u> Several structures within the GLS are highly conserved in species outside of the <i>Drosophilidae</i> family	74
<u>Figure 3.6.</u> Endogenous <i>grk</i> <i>in situ</i> hybridisation and MS2-MCP tagged <i>grk</i> show very similar localisation patterns throughout mid-oogenesis	76
<u>Figure 3.7.</u> Injection produces endogenous-like <i>grk</i> RNA localisation and does not induce a stress response	79
<u>Figure 3.8.</u> Injected <i>grk</i> RNA rapidly moves from the centre of the oocyte to the anterior before becoming enriched at the DA	81

Figure 3.9. Injected <i>grk</i> RNA rapidly moves from the site of injection to the anterior of the oocyte before it becomes most highly enriched at the DA cap	83
Figure 3.10. A <i>scrambled</i> version of the GLS shows a high degree of diversity from the wild type GLS in both primary and secondary structures	84
Figure 3.11. Full length <i>grk</i> RNA containing a <i>scrambled</i> form of the GLS is unable to localise in the oocyte	86
Figure 3.12. Injected <i>scrambled</i> RNA does not localise in the oocyte and is diffuse throughout the cytoplasm	87
Figure 3.13. Injected <i>scrambled</i> RNA travels from the centre of the oocyte and is uniformly dispersed throughout the cytoplasm	88
Figure 4.1. The <i>Drosophilid</i> GLS-like signals with the most divergent secondary structures are transported to the anterior of the oocyte.	100
Figure 4.2. The GLS-like signals from three Dipterans form secondary structures that vary at the internal loops, bulges and hairpin loop.	102
Figure 4.3. Novel GLS-like signals from species outside of the <i>Drosophilidae</i> family are able to drive <i>grk</i> transport in the oocyte.	103
Figure 4.4. The GLS-like signal from <i>A. aegypti</i> is the most divergent identified and is incapable of driving <i>grk</i> localisation in the oocyte.	105
Figure 4.5. The stem regions of the GLS are the most highly conserved in both sequence and secondary structure.	108
Figure 4.6. Stem 1 and internal loop 1 are not essential for <i>grk</i> transport in the oocyte but increase efficiency.	111
Figure 4.7. <i>grk</i> RNA localises apically in the syncytial blastoderm embryo and recruits the transport component Egl-GFP.	113
Figure 4.8. M1-7 reduces <i>grk</i> localisation in the embryo and only recruits low levels of Egl to the apical cortex.	115
Figure 4.9. The sequence and orientation of the purine stack within stem 3 is essential for <i>grk</i> transport in the oocyte.	118
Figure 4.10. Changing the sequence of the purine stack in stem 3 disrupts transport of <i>grk</i> RNA due to an inability to recruit Egl.	120
Figure 4.11. The length of distal stem 4 is essential for <i>grk</i> transport in the oocyte.	121

Figure 4.12. Increasing the length of distal stem 4 reduces *grk* transport in the 123
oocyte.

Figure 4.13. Changing the length of distal stem 4 greatly reduces *grk* transport in the 124
embryo due to an inability to recruit the transport machinery component Egl.

Figure 4.14. The AU-content of the distal stem is essential for driving *grk* transport in 127
the oocyte

Figure 4.15. A GC-rich distal stem fails to drive *grk* localisation in the embryo due to 129
an inability to recruit the transport machinery component Egl.

Figure 4.16. Sequence modifications designed to test whether the sequence of distal 131
stem 4 is important for *grk* localisation.

Figure 4.17. Sequence modifications within distal stem 4 have little effect on *grk* 132
transport in the oocyte, but mutation of the distal U-A base pair greatly reduces *grk*
transport.

Figure 4.18. Increasing the stability of distal stem 4 abolishes *grk* transport in the 136
oocyte and embryo due to an inability to recruit the transport machinery component
Egl.

Figure 4.19. Moving the position of distal stem 4 reduces *grk* localisation in the 138
oocyte and the embryo due to poor recruitment of the transport machinery
component Egl.

Figure 4.20. The sequence of internal loop 2 is not essential for *grk* transport but 141
both the size and the orientation of the kink produced is.

Figure 4.21. Changing the orientation of internal loop 2 abolishes *grk* enrichment at 142
the anterior of the oocyte.

Figure 4.22. Reducing the size of internal loop 2 by removing a uridine reduces *grk* 144
transport in the embryo and recruitment of the transport machinery.

Figure 4.23. Switching the orientation of the kink produced by internal loop 2 146
abolishes *grk* transport in the embryo and recruitment of the transport machinery.

Figure 4.24. The presence of bulge 1 and bulge 2 is not required for efficient *grk* 148
transport in the oocyte.

Figure 4.25. Schematic model for the recognition of the GLS by the transport 155
machinery.

<u>Figure 5.1.</u> Injected <i>grk</i> RNA is rapidly transported to the anterior of the <i>sqd1/Df</i> oocyte.	162
<u>Figure 5.2.</u> Injected <i>grk</i> RNA enriches at the DA of <i>OrR</i> and <i>sqd1/Df</i> egg chambers.	164
<u>Figure 5.3.</u> MS2-labelled <i>grk</i> RNA forms a strong anterior ring yet still associates with the DA in <i>sqd1</i> mutants.	166
<u>Figure 5.4.</u> <i>grk*dendra;sqd1/Df</i> mutant flies lay dorsalisated eggs due to a depletion of Sqd protein.	167
<u>Figure 5.5.</u> <i>grk</i> RNA forms bright puncta at the anterior and DA of <i>sqd1</i> oocytes.	168
<u>Figure 5.6.</u> <i>grk</i> RNA clusters at the anterior of <i>sqd1</i> oocytes are similar in size to those at the DA of wild type oocytes.	170
<u>Figure 5.7.</u> <i>grk</i> clusters at the anterior of <i>sqd1</i> oocytes show low levels of dynamics.	173
<u>Figure 5.8.</u> Injected <i>grk</i> RNA interdigitates with Me31B::GFP at the DA and the anterior of <i>sqd1/Df</i> oocytes.	175
<u>Figure 5.9.</u> Sqd is weakly recruited to the anterior of the oocyte following <i>grk</i> RNA injection.	177
<u>Figure 5.10.</u> Injecting <i>grk</i> RNA near the nucleus rescues the <i>sqd1</i> mislocalisation phenotype.	179
<u>Figure 5.11.</u> Predicted Sqd binding sites within the GLS.	181
<u>Figure 5.12.</u> The two putative Sqd binding sites in the proximal third of the GLS are not essential for <i>grk</i> anchoring.	182
<u>Figure 5.13.</u> GLS-like signals from other species show increased levels of <i>grk</i> at the anterior of the <i>D. melanogaster</i> oocyte following injection.	185
<u>Figure 5.14.</u> The GLS-like signal from <i>C. capitata</i> causes <i>grk</i> RNA to dock at the edge of P bodies in large clusters at the anterior and DA of the oocyte.	187
<u>Figure 5.15.</u> The distal third of the GLS-like signal from <i>C. capitata</i> is responsible for conferring aberrant <i>grk</i> anchoring at the anterior of the <i>D. melanogaster</i> oocyte.	189
<u>Figure 5.16.</u> The accessibility provided by bulge 2 is important for efficiently restricting <i>grk</i> to the DA of the oocyte.	191
<u>Figure 5.17.</u> Bulge 2 is required to restrict <i>grk</i> anchoring to the DA of the oocyte.	193

<u>Figure 6.1.</u> Me31B is recruited to apically localised <i>grk</i> RNA in the embryo.	206
<u>Figure 6.2.</u> <i>grk</i> RNA moves independently of Me31B in transport particles.	208
<u>Figure 6.3.</u> Me31B moves independently of <i>grk</i> RNA in transport particles.	209
<u>Figure 6.4.</u> <i>grk</i> RNA and Me31B colocalise at the apical cortex of the embryo.	211
<u>Figure 6.5.</u> The GLS-like signal from <i>C. capitata</i> recruits Me31B to the apical cortex of the embryo.	213
<u>Figure 6.6.</u> The sequence of stem 4 of the GLS is not required to recruit Me31B to the apical cortex of the embryo.	215
<u>Figure 6.7.</u> Moving stem 4 from the distal stem to the proximal stem of the GLS abolishes Me31B recruitment in the embryo.	216
<u>Figure 6.8.</u> The proximal third of the GLS is essential for Me31B recruitment to the apical cortex of the embryo.	218
<u>Figure 6.9.</u> The sequence of the GLS internal loops are not essential for Me31B recruitment in the embryo.	220
<u>Figure 6.10.</u> The sequence surrounding bulge 2 is important for efficient Me31B recruitment in the embryo.	221
<u>Figure 6.11.</u> Summary of the affect of mutations within the GLS on recruitment of Me31B to the apical cortex of the embryo.	222
<u>Figure 7.1.</u> Model for the role of the GLS in <i>grk</i> localisation from the NCs to the DA of the oocyte.	241
<u>Appendix E.</u> Assessing the quality and integrity of <i>in vitro</i> transcribed <i>grk</i> RNA	297

List of Tables

<u>Table 2.1.</u> Recipe for the buffers used in this thesis.	49
<u>Table 2.2.</u> Oligonucleotide primers used to sequence GLS mutant clones in the full-length <i>grk</i> transcript.	51
<u>Table 2.3.</u> Primary and Secondary antibodies and their dilutions used for Western Blotting.	52
<u>Table 3.1.</u> The GLS is highly conserved among the <i>Drosophilidae</i> family	63
<u>Table 3.2.</u> GLS-like signals can be identified in species outside of the <i>Drosophilidae</i> family	69
<u>Table 4.1.</u> Several regions of the GLS primary sequence are conserved among the newly identified GLS-like signals from other species.	107
<u>Table 4.2.</u> The efficiency of <i>grk</i> transport in the oocyte is greatly reduced by truncating the proximal third of the GLS.	112
<u>Table 4.3.</u> Specific sequence modifications in distal stem 4 reduce <i>grk</i> transport in the oocyte.	133
<u>Table 4.4.</u> Modifications in internal loop 2 reduce <i>grk</i> transport in the oocyte.	143
<u>Table 4.5.</u> Modifications in bulge 2 have no effect on <i>grk</i> transport in the oocyte.	149
<u>Table 5.1.</u> <i>grk</i> clusters at the anterior of <i>sqd1/Df</i> oocytes are significantly bigger than those at the anterior of wild type oocytes.	171
<u>Table 7.1.</u> Features of the GLS shown to be essential for <i>grk</i> transport are conserved in several Egl-localised transcripts.	232
<u>Appendix A.</u> GLS protein interactors	293
<u>Appendix B.</u> Fly strains	294

Chapter 1: Introduction

The importance and role of mRNA localisation

mRNA localisation

Until relatively recently it was thought that the main way of restricting protein expression to particular regions of a cell was through post-translational mechanisms involving cell sorting signals within the protein. However, the post-transcriptional localisation of messenger RNAs (mRNAs) has since been shown to be a global and highly conserved strategy to achieve such protein restriction, with particular importance during development.^{1,2}

Intracellular mRNA localisation results in a non-uniform distribution of transcripts in the cytoplasm, and is thought to be functionally highly significant to the cell. First, localisation of a single mRNA transcript allows multiple rounds of translation in the desired region of the cell, which is not only more energy efficient, but also allows rapid changes in protein levels in response to external stimuli.³ This is particularly important in large cells such as oocytes, where transcripts are translated a long distance from the cell nuclei in which they are transcribed, but is also important in small cells such as yeast.⁴⁻⁷ Second, mRNA localisation ensures that proteins that may be detrimental if expressed ectopically are restricted to regions within the cell where they are required. For example, the ectopic expression of maternal determinants results in disrupted embryonic patterning.⁸ Third, mRNAs encoding proteins that form macromolecular complexes can be localised to the same region of the cell resulting in the co-translation of different subunits and high local protein concentrations that favour complex formation.⁹ In fact, computational modelling suggests that some protein complexes in oocytes and neurons are unable to efficiently form without the aid of mRNA localisation mechanisms.¹⁰ Fourth, multiple rounds of translation in different regions of the cell allows distinct post-translational modifications to be carried out to each subset, which could greatly influence the function of the protein.⁸

Functions of mRNA localisation

mRNA localisation plays an important role in many biological processes during development such as: segregating cell-fate determinants for asymmetrical cell divisions, cell motility, and establishing the morphogen gradients required to compartmentalise the embryo.^{6,7,11-21} Furthermore, mRNA localisation is required in later development for axon and dendritic path finding.²² Examples of localised mRNAs include *β -actin* which localises to the leading edge of lamellipodia in migrating cells, resulting in the assembly of actin filaments required for directed cell motility (Figure 1.1A).²³ *CamKII α* localises to the dendrites of mature mammalian neurons where it regulates the synaptic plasticity associated with learning and memory (Figure 1.1B) ²⁴, and *ASH1* mRNA is targeted to a specific daughter cell resulting in asymmetric cell division and preventing mating-type switching in *Saccharomyces cerevisiae* (Figure 1.1C).²⁵ However, the most widely studied examples of mRNA localisation are those responsible for breaking symmetry during patterning of the presumptive embryo within *Xenopus* and *Drosophila* oocytes.

Specific maternally derived mRNAs are asymmetrically localised to different regions of the cytoplasm in ascidian, amphibian, and insect oocytes and embryos, and their translation is essential for embryonic patterning.²⁶ For example, several transcripts are targeted to the vegetal pole of the *Xenopus* oocyte, including *Vg1* and *VegT*, to specify the germ cells and pattern the mesoderm and endoderm (Figure 1.1D).^{14,17,27,28} In the *Drosophila* oocyte, several maternal transcripts, including *gurken* (*grk*), *bicoid* (*bcd*), *oskar* (*osk*), and *nanos* (*nos*), localise to distinct regions of the *Drosophila* oocyte, breaking the symmetry and establishing the body axes of the embryo (Figure 1.1E).²⁹

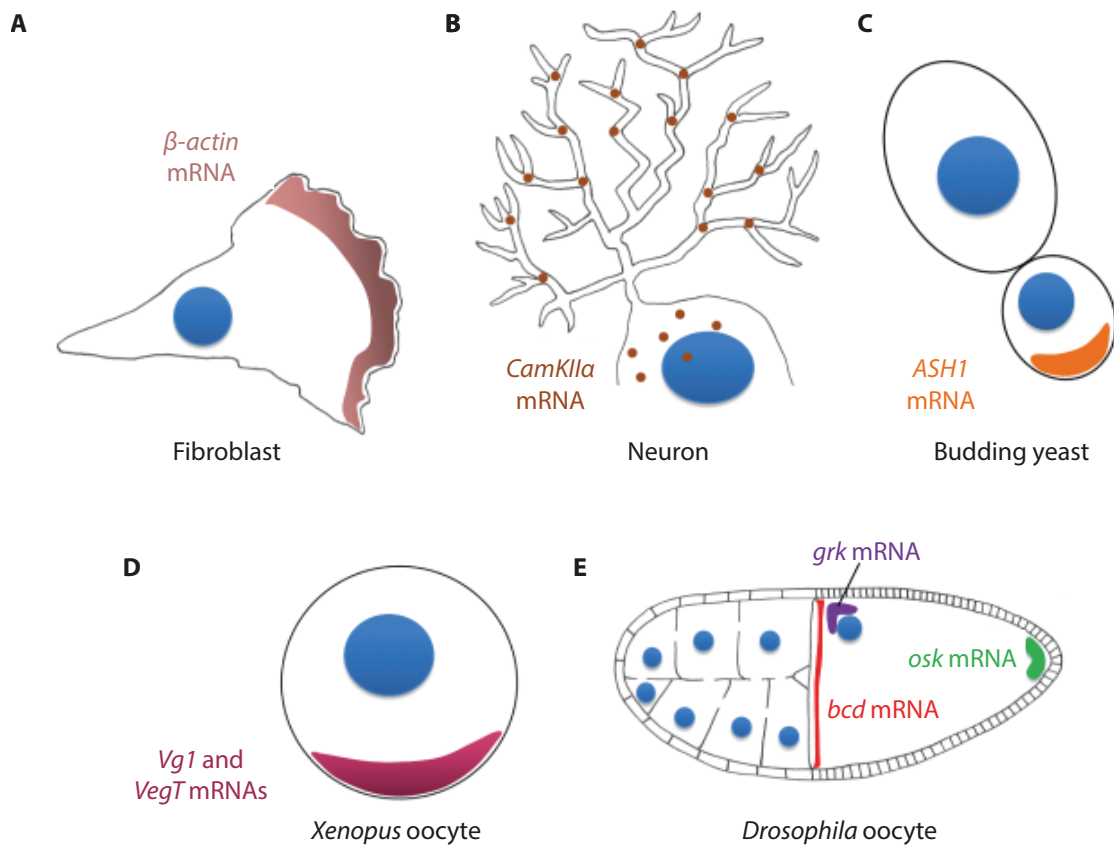


Figure 1.1. mRNA localisation is a widely used mechanism for restricting protein expression (Adapted from Cody et al. 2013). (A) In migratory cells such as fibroblasts, β -actin mRNA localises to lamellipodia at the leading edge. Localised translation of the mRNA results in the assembly of actin filaments required for cell motility. (B) In neurons, *CamKII α* specifically localises to dendrites where its localised translation is required for synaptic plasticity associated with learning and memory. (C) *ASH1* mRNA localises to the daughter cell during mating-type switching in *Saccharomyces cerevisiae*, regulating mother-daughter cell fates. (D) In *Xenopus* oocytes, the endoderm and mesoderm germ layers are specified through localisation of *Vg1* and *VegT* mRNAs to the vegetal pole. (E) In the *Drosophila* oocyte, *grk* (violet), *bcd* (red), and *osk* (green) mRNAs localise to distinct regions of the cell where they establish the anterior-posterior and dorsal-ventral axes of the future embryo.

mRNA localisation in disease

The importance of mRNA localisation is highlighted in several diseases that have been linked to defective mRNA localisation and translation (for a review see Ref ²⁶). Due to the importance of mRNA localisation in establishing cellular polarity, and the fact that one of the first stages of cancer is the disruption of normal cell polarity leading to tissue disorganisation, it is likely that disrupting this process could be a precursor for tumourigenesis (for a review see Ref ³⁰). For example, the apical localisation of *ZO-1* mRNA is required for targeting the protein to tight junctions, maintaining epithelial cell polarity in mice, and deregulation of this process is linked to a poor prognosis in breast cancer patients.³¹ In addition, several RNA-binding proteins (RBPs) have been linked to cancer, including the IMP family which may be involved in degrading the extracellular matrix, enabling cancer cell invasion.³²

Several muscle and neurological disorders have been associated with RNAopathies, including Spinal muscular atrophy (SMA) and Myotonic dystrophy type 1 (DM1). The most studied example is fragile X mental retardation syndrome (FXS) which is the most common form of inherited mental retardation. FXS is caused by CGG trinucleotide repeat expansions in the 5' untranslated region (5'UTR) of the *FMR1* gene, resulting in transcriptional silencing and loss of function of the encoded FMRP protein. FMRP is an RBP involved in controlling synaptic plasticity in response to external cues through mRNA trafficking and translational regulation in dendrites.³³⁻³⁵

mRNA localisation has also been implicated in the establishment of viral infection in both plants and mammalian cells. In plants, several mRNAs have been shown to be transported to chloroplasts within the cell, and into the plant vascular system through plasmodesmata.^{36,37} These intercellular trafficking routes are hijacked by some plant viruses, enabling the distribution of viral nucleic acids (for a review see Ref ³⁸). Subcellular targeting of viral RNAs, such as human immunodeficiency virus (HIV) type 1, has also been shown to occur in mammalian cells.^{39,40}

Mechanisms of mRNA localisation

mRNAs destined for a specific location within the cell are transcribed in the nucleus. Following transcription, nascent transcripts are spliced to remove introns, forming coding mRNAs. However, some transcripts retain a selection of introns to regulate mRNA levels through non-sense mediated decay (NMD), or through retention in the nucleus.⁴¹ After splicing, a cohort of *trans*-acting protein co-factors bind to mRNAs in the nucleus, packaging them into ribonucleoprotein (RNP) complexes (Figure 1.2).^{42,43} For example, the exon-junction complex (EJC) binds to mRNAs at the junction between two exons and is involved in the translation and localisation of the mRNA.⁴⁴⁻⁴⁷

Following the packaging of mRNAs into RNP complexes, they diffuse through the interchromatin tracks to the nuclear pore complex (NPC) where they interact with additional co-factors. For example, RNA helicases remodel the RNP complex and carry out quality control checks before allowing export into the cytoplasm.⁴⁸⁻⁵¹ In the cytoplasm, RNP complexes can oligomerise via protein-protein interactions to form higher-order mRNA complexes called RNA granules. These RNA granules are often composed of several different species of mRNA, along with translational regulatory proteins and ribosomal subunits (Figure 1.2). The formation of RNA granules is thought to increase the efficiency of mRNA transport⁵², but has been shown to not be essential for the transport of individual mRNAs.⁵³

***Trans*-acting proteins**

The eukaryotic genome encodes hundreds of RBPs with diverse functions in the co- and post-transcriptional regulation of RNA. Each protein in turn has hundreds of RNA targets. RNPs are composed of several RBPs, also known as *trans*-acting factors, that can interact directly with the cargo mRNA via *cis*-acting signals in the mRNA, or indirectly via protein-protein interactions. These highly conserved *trans*-acting factors are responsible for controlling each stage of mRNA localisation from post-transcriptional processing, export from

the nucleus, translational silencing, coupling to the transport machinery, specifying the target localisation site, and anchoring at the final destination (for a recent review see Ref ⁵²). Many *trans*-acting factors have been identified to be involved in mRNA localisation and RNA-binding through genetic screens, biochemical purification of RNA granules, and affinity purification with *cis*-acting localisation signals. For example, a range of proteins were identified bound to the *grk* localisation signal (GLS) of *grk* mRNA from *Drosophila* ovary extracts ⁵⁴, and novel RBPs are continually being discovered.^{55,56}

Remodelling of the RNP complex is thought to occur at different stages of the localisation process, including when the mRNA reaches its target, altering the anchoring and translational state of the mRNA, and allowing expression of the protein (Figure 1.2).²⁹ Despite the continual remodelling of *trans*-acting proteins throughout the life cycle of the mRNA, studies have suggested that the early protein interactions in the nucleus, such as those co-transcriptionally associated with the mRNA, can determine the future fate of the mRNA. For example, in the *Drosophila* egg chamber, splicing of *osk* forms a *cis*-acting regulatory element, called the spliced *oskar* localisation element (SOLE), that is required for the deposition of the EJC. Both the SOLE and the EJC are required for targeting the transcript to its final destination in the oocyte ⁵⁷, showing that protein factors that bind in the nucleus can direct transcripts to their target sites in the cytoplasm.

Although *trans*-acting factors have been shown to drive the localisation of their target mRNAs to specific destinations, many of these proteins are able to bind a variety of transcripts destined for distinct regions of the cell. For example, Staufen (Stau) is a highly conserved double-stranded RBP (dsRBP) that interacts with both *bcd* and *osk* mRNA, transcripts which are transported to the anterior and posterior of the *Drosophila* oocyte, respectively.^{20,58,59} As a result, it is likely that the collective affect of the entire cohort of proteins bound to a given mRNA determines the specificity of mRNA localisation within the cell, rather than the binding of single proteins. Therefore, despite the vast number of RBPs that have been identified,

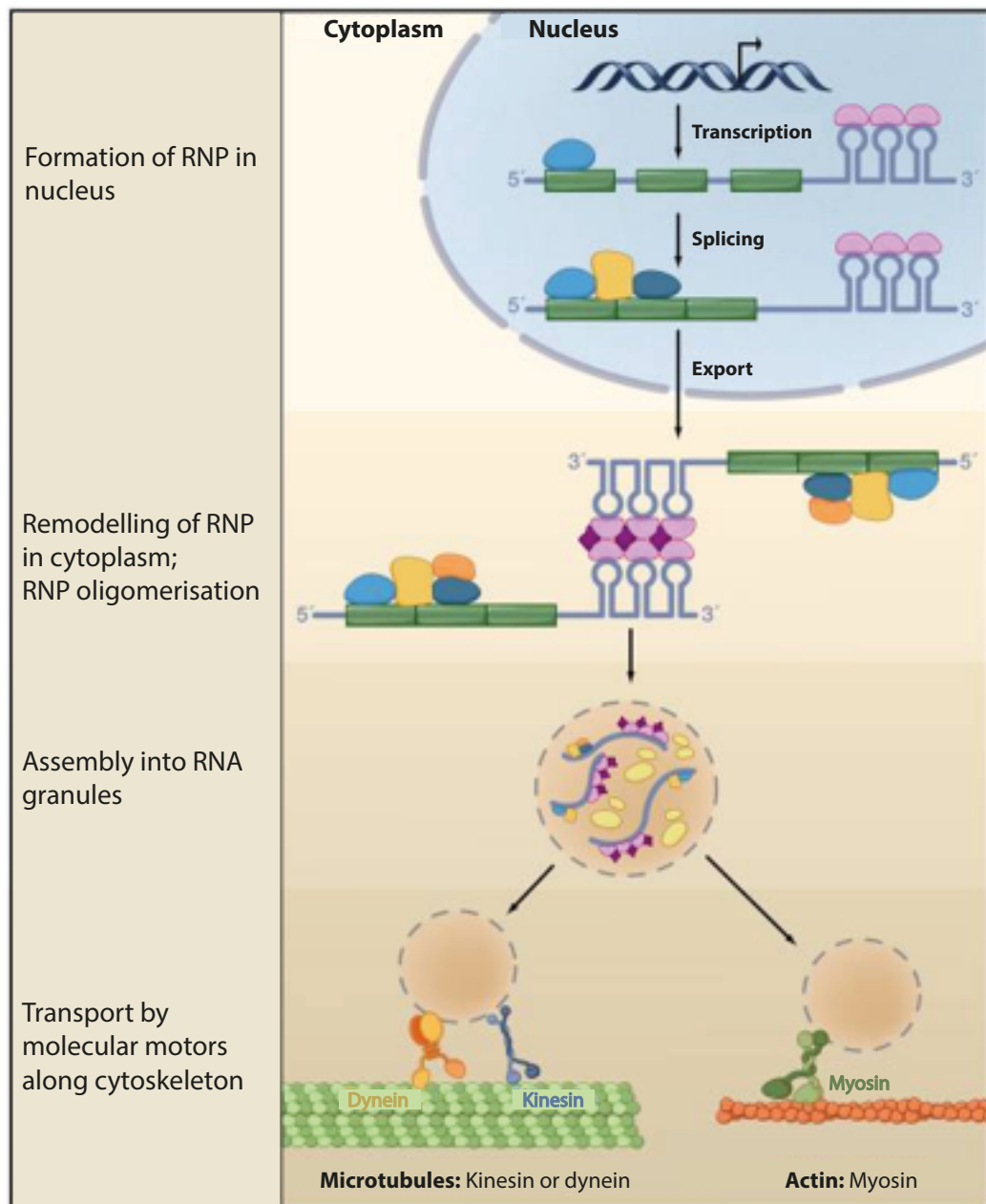


Figure 1.2. The stages of mRNA localisation (Taken from Martin and Ephrussi, 2009). In the nucleus, pre-mRNAs containing exons (green), introns and 5' and 3'UTRs (grey), are bound by RBPs (blue and purple). Splicing of the pre-mRNA coincides with the addition of more RBPs (gold and dark blue), resulting in the formation of an RNP complex. Remodelling of the RNP occurs after export into the cytoplasm, with additional proteins being added (orange and dark purple). Oligomerisation between different RNPs via protein-protein interactions in some instances generates RNA granules containing different RNA species, along with translational regulatory and ribosomal subunits (yellow). Association of the RNA granules with motor proteins results in their transport to their final destination along the cytoskeletal network.

deciphering the exact combination of proteins that drive a given mRNA to its final destination is not trivial and remains an active area of research in the field.⁵⁴⁻⁵⁶

***cis*-acting regulatory elements**

Localised mRNAs contain *cis*-acting regulatory elements that are essential for their biological functions and for directing their localisation within the cell. *Cis*-acting elements are most commonly found in the 3' untranslated region (3'UTR) of an mRNA where their complex stem-loop structures do not have deleterious effects on translation.²⁹ These regulatory “zipcodes”, or mRNA localisation signals, are composed of a variety of different primary, secondary and tertiary mRNA structures. These structures, along with the sequence of the mRNA, provide enough diversity to drive specific interactions with a cohort of *trans*-acting proteins, directing the mRNA to its final destination within the cell. In addition, alternative splicing is able to reshuffle *cis*-acting signals, altering the composition of the proteins binding in the RNP complex, and thus changing the final destination of the mRNA.^{60,61}

Many localisation signals have been identified and characterised across several species, showing that they vary vastly in terms of both length and structure. For example, the length of a localisation signal can vary from the 11-nucleotide mRNA trafficking signal in the 3'UTR of *myelin basic protein (MBP)* mRNA⁶², to the 625-nucleotide region of the *bcd* 3'UTR.⁶³ Some mRNAs contain multiple zipcodes which can either function together or redundantly. Such modular localisation elements have been identified in *bcd* mRNA which contains several *bcd* localisation elements (BLEs) required for targeting of the transcript in the *Drosophila* egg chamber. In addition, each element is required for a distinct step of *bcd* transport, with BLE1 driving transport from the nurse cells (NCs) into the oocyte, and other elements being required for later stages of transport and anchoring of the transcript.⁶⁴⁻⁶⁶ Clustering of localisation elements has also been observed in *Vg1* mRNA in *Xenopus*⁶⁷ and *ASH1* mRNA in yeast⁶⁸, and may facilitate the direct and indirect binding of *trans*-acting factors to the mRNA.

In contrast to those mRNAs containing multiple zipcodes, there is some evidence to suggest that a subset of localised mRNAs do not contain a localisation signal at all. Instead, these mRNAs are thought to hitchhike on those that contain a localisation signal, sharing the same RNP complex, and resulting in the localisation of both transcripts. For example, the *I* factor retrotransposon localises to the dorsal-anterior (DA) of the *Drosophila* oocyte through the *I* factor localisation signal (ILS) but contains an additional stem-loop called the A loop responsible for stimulating RNA:RNA interactions. These RNA:RNA interactions were shown to facilitate the formation of large RNP particles, and stimulate the localisation of RNAs containing other localisation signals, and potentially those without a localisation signal.⁶⁹ RNA:RNA interactions have also been observed for *osk* mRNA in the *Drosophila* oocyte, with oligomerisation of unlocalised *osk* transcripts resulting in the formation of large silencing particles, repressing ectopic translation.^{70,71} However, whether RNA:RNA interactions are a common mechanism for the efficient localisation of mRNAs destined for the same region of the cell, or for driving the localisation of transcripts that lack a localisation signal, is yet to be determined.

Types of mRNA localisation

There are several ways in which an mRNA can be asymmetrically distributed, including local degradation or stabilisation of transcripts, diffusion and entrapment and active transport followed by anchoring (Figure 1.3). In addition, there is some evidence to suggest that a selection of mRNAs are localised by vectorial transport from the nucleus to a specific target, although this is thought to be rare and will not be discussed. The first mechanism of mRNA localisation involves stabilisation and protection of mRNAs from degradation at their final destination, while the remaining mRNA in the rest of the cell is degraded (Figure 1.3A). For example, *D. melanogaster nos* mRNA contains two distinct *cis*-acting regulatory sequences in its 3'UTR; one targets the mRNA for degradation in all regions of the egg except the posterior, and the other protects and stabilises the mRNA at the posterior of the embryo.⁷²

The second mechanism of mRNA localisation involves diffusion and entrapment by an

anchor (Figure 1.3B). Several mRNAs that localise to the *D. melanogaster* pole plasm diffuse throughout the cytoplasm until they become trapped by a localised anchor.⁷³ For example, only 4% of *nos* mRNA is localised and stabilised in the pole plasm of the egg following laying, whilst the remaining mRNA distributed throughout the cytoplasm is translationally repressed and rapidly degraded.^{72,74-76}

The third, and most studied, mechanism of mRNA localisation is via active transport of mRNAs along cytoskeletal elements using molecular motors, followed by retainment of the mRNA at it's final destination by a static anchor (Figure 1.3C). The role of active transport in mRNA localisation will be discussed further in the next section. Together these mechanisms of mRNA localisation ensure the correct temporal-spatial expression of proteins in a way that is essential for maintaining normal cellular function. Despite the wealth of research in this area, it is unclear whether the variety of mechanisms described work independently or together, or whether they show redundancy to ensure the correct localisation of specific transcripts.

Active transport and anchoring of mRNAs

The cytoskeleton and motor proteins

Active transport of mRNAs requires the action of motor proteins which are able to move cargoes along cytoskeletal networks using chemical energy. There are three known families of cytoskeletal motor proteins; Myosins, Kinesins and Dyneins. The type-V Myosin (MyoV) family are required for intracellular transport using the actin cytoskeleton^{77,78} whereas the Kinesin motor proteins are the major plus end directed microtubule motors.⁷⁹⁻⁸¹ The molecular action of motor activity is detailed in recent reviews.⁸²

Dynein was originally identified as an ATPase in *Tetrahymena pyriformis* cilia⁸³ and a cytoplasmic form was subsequently isolated and shown to drive transport towards the minus ends of microtubules.⁸⁴⁻⁸⁶ Cytoplasmic Dynein is a large, 1.2 MDa protein required for many

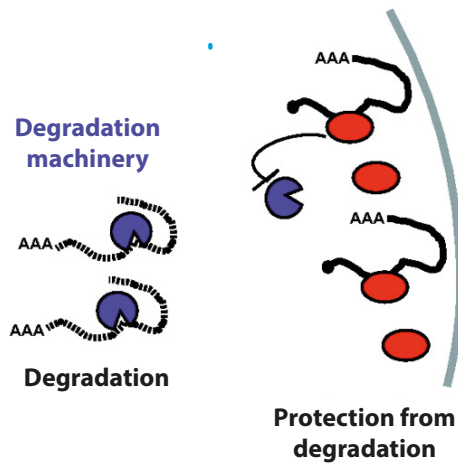
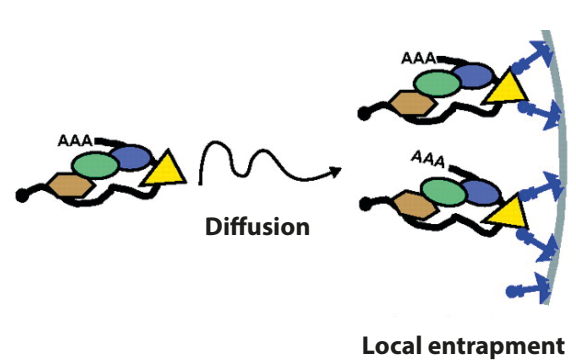
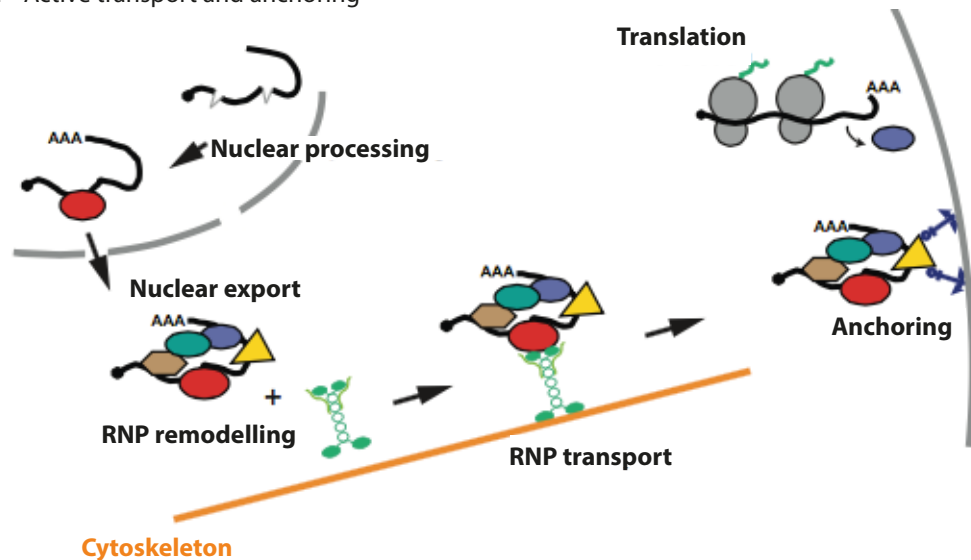
A Local protection from degradation**B** Diffusion and entrapment**C** Active transport and anchoring

Figure 1.3. Mechanisms of mRNA localisation. (Adapted from Medioni *et al.* 2012). (A) Transcripts destined for translation at a specific subcellular location can be protected from the global degradation machinery in that region, whilst being degraded in the rest of the cell. (B) Some mRNAs are localised by free diffusion throughout the cell, followed by local entrapment at their target site. (C) The most common mechanism of mRNA localisation involves the active transport of mRNA cargoes in the form of a ribonucleoprotein (RNP) complex. During processing of the mRNAs in the nucleus they are recognised by *trans*-acting protein factors forming RNP complexes that are matured and remodelled throughout the localisation process. Following nuclear export, mRNAs are coupled to molecular motors which transport the transcripts along polarised cytoskeleton networks. At their final destination, mRNAs are anchored and translationally activated.

processes including; the transport of vesicles from the endoplasmic reticulum (ER) to the Golgi, transport of endosomes, lysosomes, phagosomes and mitochondria, and retrograde movement in axons.^{79,87} Dynein is composed of two heavy chains with motor activity, two intermediate chains, and light-intermediate and light chains involved in cargo recognition and Dynein complex assembly.⁸⁸⁻⁹¹ The Dynein heavy chain belongs to the AAA family of ATPases and contains a large globular 'head' domain, a coiled-coil extension (the stalk) that interacts with the microtubule in an ATP-dependent manner, and an extended tail domain (the 'stem') which dimerises with another heavy chain and binds to the light and intermediate subunits.^{88,89,92-95} The interaction of Dynein with microtubules requires the accessory proteins Dynactin⁹⁶⁻⁹⁹, which has roles in cargo binding and motor processivity^{96,100,101}, and Dynamitin which is crucial for the structure of Dynactin.¹⁰² The co-ordinated movement of the two heavy chains in response to ATP hydrolysis allows Dynein to 'walk' along the microtubule in a hand-over-hand fashion.⁹³

Although Kinesins are mainly plus-end directed and Dynein is mainly minus-end directed, transport on motors can be bidirectional. For example, minus-end directed Kinesins are involved in the transport of early endosomes.^{103,104} Bidirectional movement allows motors to avoid obstructions on the cytoskeleton network and correct mistakes in transport, and is likely to occur by the binding of opposing motors at one time in a regulated manner that is currently not well understood. For example, FMR-containing mRNA transport particles have been shown to use both Dynein and Kinesin-1 to move bidirectionally in *Drosophila* S2 cells.¹⁰⁵ However, Dynein has also been shown to be able to reverse direction.^{106, 107} Bidirectional movement of the Dynein motor is exemplified in the *Drosophila* embryo where the transport of apically localised transcripts in both minus- and plus-ended directions is suppressed when Dynein is inhibited, but not in the absence of Kinesin-1 and Kinesin-2.¹⁰⁸

The organisation of the cytoskeleton is highly complex and is essential for targeting transcripts to their destination within the cell. It is generally thought that in order to generate intracellular polarity, the cytoskeleton must be highly polarised.¹⁰⁹⁻¹¹¹ Post-translational

modifications to subsets of stable microtubules have been demonstrated in some cells to be required for directing cell polarity.^{112,113} In other cells, such as the *Drosophila* embryo, microtubules are stable and arranged with strong apical-basal polarity, allowing rapid localisation of mRNAs involved in segmenting the embryo. However, in the *Drosophila* oocyte, *osk* mRNA was shown to move to the posterior of the oocyte in a random biased walk, rather than a direct path.¹¹⁴ Analysis of the global orientation of microtubules using the microtubule plus end marker EB1-GFP showed the microtubule network is highly dynamic with a random distribution of polarity and only a slight bias towards the posterior of the oocyte.¹¹⁵ Such posterior bias was found to be a result of a gradient of microtubule nucleation from the anterolateral cortex to the posterior of the oocyte, with PAR-1 suppressing microtubule nucleation at the posterior. Differences between the cytoskeletal organisation is likely dependent upon its role in specific cell types. For example, in the *Drosophila* oocyte, microtubules are required for the random distribution of organelles and lipid droplets throughout the cytoplasm, gathering and redistributing cargoes to generate polarity, and maintaining the structural integrity of the cell. Therefore, the overall organisation of the cytoskeleton is likely a result of the need to achieve multiple functions within the cell.¹¹⁵

The role of motor proteins in mRNA localisation

During both early and mid-oogenesis in *Drosophila*, active transport is essential for moving a large number of maternal transcripts into the oocyte from the NCs and on to their final destinations to carry out their functions. This process utilises the Dynein motor and the accessory proteins Egalitarian (Egl) and Bicaudal-D (BicD) which link mRNA cargoes to the motor.¹¹⁶⁻¹¹⁹ Upon reaching the oocyte some transcripts, such as *grk* and *bcd*, continue to their final destinations using the same motor.¹²⁰⁻¹²¹ *osk*, on the other hand, switches to the Kinesin motor and is transported to the plus ends of microtubules.¹¹⁴ These highly conserved mechanisms of active mRNA transport are found across many different cell types, such as the *Drosophila* embryo, where pair-rule transcripts are localised to the minus ends of microtubules by the Dynein-BicD-Egl complex,¹¹⁶ and developing oligodendrocytes where

MBP and *CamKII α* are localised to the plus ends of microtubules using Kinesin.¹²²⁻¹²³

Several transcripts are known to be transported along the actin cytoskeleton using the motor protein myosin, including *ASH1* mRNA in yeast, *β -actin* in fibroblasts and *pros* in *Drosophila* neuroblasts.^{6, 7, 11, 124-126} Actin has also been shown to be involved in the long-range transport of vesicles in mouse oocytes.¹²⁷ Moreover, actin is required indirectly for the organisation and maintenance of the microtubule cytoskeleton in the *Drosophila* oocyte.¹²⁸ As a result, several transport mechanisms operate in parallel to localise different mRNAs and are, to some extent, dependent upon each other.⁵²

The number of motors recruited to an RNP complex, their overall activity, and *cis*-acting signals in mRNA cargoes, can all alter the speed and direction of transport.^{53,129,130} For example, lipid droplets transported into the interior of the *Drosophila* embryo along microtubules use both Dynein and Kinesin-1, and between four and six Dynein complexes were found to bind an individual droplet at any one time.¹³¹ In addition, although mRNAs that lack a localisation signal have also been shown to associate with motor proteins and carry out bidirectional movement, evidence suggests that the presence of a localisation signal increases the frequency and length of directed movements *in vitro*, and in mammalian cells and *Drosophila* embryos.^{53,132,108} For example, the bidirectional movement of apical transcripts in the *Drosophila* embryo is switched to minus-end directed movements by *cis*-acting signals in the transcript which control the number of Dynein molecules recruited to the transcript via Egl and BicD.¹⁰⁸ Moreover, transport particles in the *Drosophila* oocyte were shown to contain more than one motor during mRNA transport,¹³³ and *cis*-acting mRNA signals were shown to be capable of increasing the number of Dynein-Dynactin motors recruited to the RNP complex, enhancing transport of the mRNA.⁵³ Finally, the recruitment of accessory proteins can increase the efficiency of motor protein motility or association with the cytoskeleton.^{134,135}

Anchoring of localised mRNAs

Once mRNAs reach their target destination they must be maintained there. mRNAs can be transported to their anchor directly, or they can be trapped as they diffuse past an anchor.¹³⁶ Anchoring requires the microtubule or actin cytoskeleton and often utilises motor proteins, closely linking transport and anchoring. For example, both microtubules and actin are used in the *Xenopus* oocyte to maintain *Vg1* mRNA at the vegetal pole.¹³⁷ Moreover, following diffusion in the *Drosophila* oocyte, *nos* mRNA is entrapped at the posterior pole by an actin-based anchor.⁷³ In addition, *osk* mRNA is anchored via actin at the posterior pole of the *Drosophila* oocyte requiring; the F-actin-binding protein Bifocal, tropomyosin II, Cap, and Dmoesin which cross-links F-actin to the posterior pole membrane.^{138,139,140}

In many cases, the microtubule network is required for anchoring mRNAs. For example, *grk* mRNA is anchored at the DA of the *Drosophila* oocyte during mid-oogenesis in a process that requires both the microtubule cytoskeleton and the motor protein Dynein.¹³³ Dynein is also required for the anchoring of several mRNAs at the apical cortex of the *Drosophila* embryo, including *wingless* (*wg*) and the pair-rule transcripts responsible for establishing accurate segmentation of the embryo.^{141,142}

The anchoring and translation of some mRNA transcripts can be a pre-requisite for the anchoring of others. For example, capture of *ASH1* transcripts in the bud tip of yeast requires translation of *ASH1* mRNA.¹⁴³ In addition, anchoring of *osk* mRNA in *Drosophila* is dependent upon local expression of the long form of Osk protein which induces the formation of actin projections from the posterior of the oocyte.^{144,145} RNAs have also been shown to be involved in the anchoring of some transcripts. This is exemplified in the *Xenopus* oocyte where early expressed small non-coding *XIsirt* RNAs and late-pathway *VegT* mRNA are essential for organising the cytoskeleton for the anchoring of other late-pathway RNAs, such as *Vg1*, to the vegetal cortex of the oocyte.¹⁴⁶⁻¹⁴⁸ This is likely to occur via complementary regions in *Vg1* and *XIsirts* which may mediate association of the RNAs.¹⁴⁷

As many mRNAs are transported by the same motor complexes and accessory proteins but their transport terminates in a range of different locations within the cell, motor proteins alone cannot provide the required specificity for targeted mRNA transport. Instead, specificity is likely the result of complex interactions between motor proteins, the organisation of their cytoskeletal networks, and a range of accessory binding proteins. Therefore, despite the current understanding of motor protein function in both mRNA localisation and other biological process, it is still unclear how specificity is generated for individual transcripts.

Regulation of translation during mRNA localisation

In eukaryotes, canonical translation is composed of four major phases: initiation, elongation, termination, and recycling (for a recent review of translation see Ref ¹⁴⁹). Several mechanisms exist to control the global levels of translation within the cell, targeting different steps of the translation process. For example, eIF4E-binding proteins (4E-BPs) negatively regulate translation by binding to the translation initiation factor eIF4E (eukaryotic initiation factor 4E) ^{150,151}, and phosphorylation of eIF2 prevents it from providing the initial methionine amino acid required for translation initiation.¹⁵² Finally, the number of ribosomes available for translation in the cell can also modulate protein synthesis levels. Alongside global control of protein expression exists control mechanisms for individual transcripts.

Translational silencing during transport of localised mRNAs

The major function of mRNA localisation is to achieve localised expression of transcripts at their site of function. As a result, translation of transcripts must be finely tuned both spatially and temporally. Possible ways of achieving this include controlled degradation of mis-expressed protein, or translational silencing of the mRNA until it is anchored at its target destination, and the timing is desirable for expression of the protein.¹³⁶ Evidence for these mechanisms comes from the fact that in wild type *Drosophila* oocytes, un-localised *osk* mRNA is associated with polysomes but is translationally repressed until the mRNA is localised at the posterior pole.¹⁵³⁻¹⁵⁷

If translational repression of the transcript is responsible for achieving control over protein expression, *trans*-acting factors originating from the cell nucleus must be involved to prevent default translation of polyadenylated transcripts. EJC components have been shown to be present in transport RNPs in a number of systems, and are only removed from the mRNA during the first round of translation.^{158,159} In addition, several other *trans*-acting factors have been shown to be required for translational silencing of mRNA transcripts in the *Drosophila* oocyte. For example, Bruno binds to Bruno-response elements (BREs) in the 3'UTR of both *grk* and *osk* mRNA and has been implicated in the translational regulation of both transcripts^{160,161}. Furthermore, additional *trans*-acting factors have been shown to be involved in the translational regulation of *grk* and *osk* including the hnRNPs Hrb27C (Hrp48) and Squid (Sqd, Hrp40).¹⁶²⁻¹⁶⁴ Finally, the recruitment of translational repressors to transport RNPs has been demonstrated in a number of cases.^{71,165-167}

In addition to the use of RBPs in controlling translation, mRNA transcripts may themselves mediate these regulatory mechanisms. For example, sequences within the mRNA may promote the binding of a *trans*-acting factor that masks the binding site of a translational activator, or physically restrict access to the translational machinery. Alternatively, the binding of RBPs to differentiated control elements (DICE) in the 3' UTR of a selection of mRNAs can cause circularisation of the mRNA in such a way that prevents the association of the 43S complex with the 60S ribosomal subunit, thus repressing translation.¹⁶⁸

Polyadenylation and deadenylation is a widely used mechanism for controlling translation. Biochemical studies in *Xenopus* oocytes have shown that cytoplasmic polyadenylation element binding protein (CPEB) is responsible for the translational control of several mRNAs including *cyclin A1*, *cyclin B1* and *cyclin B2*.¹⁶⁹⁻¹⁷³ CPEB and the cleavage and polyadenylation specificity factor (CPSF) bind consensus motifs in the 3'UTR of target transcripts in the nucleus¹⁷⁴⁻¹⁷⁹ resulting in cleavage of the transcript before a poly(A) tail is added by poly(A) polymerase (PAP).¹⁸⁰⁻¹⁸³ The lengthening and shortening of the poly(A) tail following nuclear export is dependent upon the CPEB-associated factors Germline

development 2 (Gld2) and poly(A) ribonuclease (PARN), respectively.^{182,184} During this time, translation initiation can be prevented by ePAB which binds to both CPEB and eIF4E, generating a closed, repressive loop in the mRNA.^{150,185,186}

The translational regulation of *osk* mRNA has been extensively studied in the *Drosophila* oocyte. Biochemical studies have shown that the ovarian protein Cup and eIF4E are required for inhibiting translation initiation of *osk*. Cup binds to eIF4E at the 5' UTR, using the same conserved sequence found in the translation initiation factors eIF4G and 4E-BP, and to Bruno at the 3' UTR.^{166, 187} As a result, Cup is likely to inhibit *osk* translation initiation through circularisation of the mRNA during early oogenesis. Following its localisation to the posterior of the oocyte in late oogenesis, Cup dissociates from eIF4E enabling the initiation of translation. Cup has also been shown to be required for the correct localisation of *osk* mRNA, indicating that regulatory mechanisms of both mRNA localisation and translational regulation may be linked.^{166, 187, 188}

Non-coding mRNAs have been shown to play a role in translational regulation, with micro-RNAs (miRNAs) controlling translation in several tissues including the *Drosophila* embryo, dendrites, and synapses.¹⁸⁹⁻¹⁹¹ For example, phosphorylation of FMR1 is linked to a miRNA and argonaut 2 (AGO2)-dependent pathway for controlling gene expression at the *Drosophila* synapse.¹⁹¹ miRNA-mediated silencing of mRNAs may utilise RBPs to enhance translational inhibition,¹⁹² indicating a link between several different regulatory mechanisms that may be transcript-specific.

Translational activation of localised mRNAs at their target site

Translational repression of mRNAs must be alleviated once the transcripts reach their final destination in order to allow protein expression. This de-repression of translation can occur immediately upon anchoring of the mRNA, or can be delayed until the correct temporal cues are received. For example, *bcd* mRNA is targeted to the core of Processing bodies (P bodies) in *Drosophila* oocytes where the translational repressors Maternal expression at 31B

(Me31B) and Bruno are enriched. *bcd* remains translationally silenced in the core of P bodies until egg activation when the P body components disassemble, releasing *bcd* from its repression.¹⁹³⁻¹⁹⁵ P bodies are non-membrane bound and electron dense structures originally described in *Drosophila* NCs as sponge bodies.¹⁹⁶ Their involvement in both the translational regulation and degradation of mRNA transcripts has been shown in yeast,¹⁹⁷ mammalian cells^{198,199} and in *Drosophila* neurons and oocytes.^{193, 200}

There are many lines of evidence to suggest that translational activation of transcripts can be achieved by local proteins that compete for binding to translational repressors at the target site. In the *Drosophila* oocyte, Osk has been shown to interact with the translational repressor Smaug, relieving the Smaug-dependent translational repression of *nos* mRNA.²⁰¹ Moreover, local post-translational modifications of translational repressors has been shown to activate translation in the case of β -actin, where phosphorylation of ZBP1 by Src kinase in specific regions of the cell decreases the binding affinity of ZBP1 for the mRNA, stimulating translation.²⁰²

Trans-acting factors bound to the mRNA can regulate the poly(A) tail length which must be elongated to recruit PABP before translation can occur²⁰³. Poly(A)-binding protein (PABP) at the 3'-end of the mRNA interacts with eIF4G which, in turn, is associated with the 5'-cap binding protein eIF4E, promoting circularisation of the mRNA in such a way that stimulates translation initiation.^{171, 173} In addition, Paip1 binds both PABP and eIF4A at the 5'-cap, increasing the efficiency of the assembly of the translation initiation complex and so aiding translational activation.^{204, 205}

The role of degradation in mRNA localisation

To ensure the tight temporal regulation of mRNAs, their degradation must be finely tuned, with mislocalised transcripts being promptly removed to avoid translation at inappropriate regions of the cell where they may be detrimental. In addition, localised mRNAs must be

protected from such degradation to ensure they are translated at their target site. For example, *nos* and *polar granule component (pgc)* mRNAs are locally protected at the posterior of the *Drosophila* oocyte while transcripts throughout the rest of the cell are degraded.^{72,201,206,207} Finally, localised transcripts that are no longer required must also be degraded.

The most common mechanism for mRNA degradation is through the controlled deadenylation and decapping of mRNAs, followed by exonuclease activity in the 5' to 3' direction.²⁰⁸ Decapping of the mRNA is carried out by the Dcp1/Dcp2 enzymes and degradation by the Xrn1 exonuclease.²⁰⁹⁻²¹⁵ In addition, exonuclease activity has been described in the 5' to 3' direction during translation, indicating that mRNA degradation can occur before the termination of translation.²¹⁶ Degradation of mRNAs through this exonuclease mechanism has been associated with P bodies in yeast.²¹⁷ However, a direct role of P bodies in the degradation of developmental transcripts, such as *bcd* and *grk*, in the *Drosophila* oocyte has not yet been demonstrated.⁵² Furthermore, three mechanisms exist to remove mRNAs from the cell that do not complete translation correctly; non-sense mediated decay (NMD) detects and removes mRNAs with premature stop codons²¹⁸⁻²²⁰, no-go decay (NGD) removes mRNAs that are stalled in translation elongation^{221,222}, and non-stop decay (NSD) removes transcripts without a translation stop codon.^{223,224}

Further mechanisms for controlling mRNA degradation involve the action of miRNAs and short interfering RNAs (siRNAs). For example, during the maternal to zygotic transition (MZT) in both *Drosophila* and zebrafish embryos, maternal mRNAs are deadenylated in a miRNA-mediated process, allowing clearance of the transcripts from the embryo, and the zygotic transcriptome is activated.²²⁵⁻²³⁰ Finally, promoter sequences in the DNA itself can, surprisingly, affect the decay rates of its encoded mRNA once it reaches the cytoplasm in yeast²³¹.

grk mRNA localisation

Drosophila are widely used as model organisms to study various biological processes such as cell cycle regulation, cell polarity and differentiation, endocytosis, exocytosis, morphogenesis and cancer metastasis. This is due to the high degree of experimental advantages including the vast range of techniques available for research, and the ease of culturing and imaging the specimens. Moreover, much of the *Drosophila* genome is highly conserved between flies and humans. This is exemplified by the fact that almost 75% of disease-related genes in humans have functional orthologs in flies.^{232,233} In addition, *Drosophila* are a particularly useful model organism for carrying out covariance searches due to the availability of the full genome sequences for 12 *Drosophilid* species.

The mechanisms of mRNA localisation have been most widely studied in both the *Drosophila* oocyte and embryo because the early screens for patterning genes highlighted many key genes whose transcripts were shown to be asymmetrically distributed. Later, genome-wide *in situ* hybridisation screens in the *Drosophila* early embryo showed that 71% of 3370 genes investigated encode mRNAs that are non-uniform, and perhaps 10-20% which are specifically asymmetrically localised.² In addition, genome-wide imaging of 5,862 mRNAs in the *Drosophila* ovary showed that mRNAs localise differently in different cellular contexts and their localisation pattern can change over time.¹ During *Drosophila* development, the localisation of maternal mRNAs in the oocyte has been shown to establish the major body axes, and localised transcripts in the embryo are responsible for body segmentation. This thesis will focus on the specific molecular interactions involved in the localisation of the maternal mRNA *grk* during mid-oogenesis in the *Drosophila* oocyte.

***Drosophila* oogenesis**

Specification of the *Drosophila* oocyte

Female *Drosophila* contain two ovaries composed of 16-20 ovarioles, each made up of a

series of 6-7 sequentially developing egg chambers (Figure 1.4).^{234, 257} At the anterior of the ovariole, new egg chambers are produced in region 1 of the germarium, starting as germline stem cells (GSC). A GSC divides asymmetrically, producing a new GSC and a cystoblast which divides four more times with incomplete cytokinesis to produce a 16-cell germline cyst (Figure 1.5). The cells of the cyst are interconnected by actin bridges called ring canals. Two of the 16 cells contain four ring canals and are known as 'pro-oocytes' as one will differentiate into the oocyte.

The selection of the oocyte at the end of region 2a of the germarium is the first major symmetry-breaking step in *Drosophila* development and is marked by the restriction of the synaptonemal complex, which initially forms in both pro-oocytes as they enter meiosis, to the oocyte. Specification of the oocyte is mediated by PAR-1 and the fusome, a cytoskeletal structure which spans the cyst and controls the orientation of cell divisions, restricts meiosis to the oocyte, and controls the formation of a microtubule organising centre (MTOC) which directs the transport of cytoplasmic factors, such as mRNAs and proteins, into the oocyte.²³⁵⁻²⁴³ The rest of the cells in the cyst form supporting NCs which produce all of the maternal mRNAs and other cellular components required for the growth and development of the oocyte, as the oocyte nucleus is transcriptionally silent until after fertilisation.²⁴⁴

In region 3 of the germarium, a single layer of 600-1000 somatic epithelial follicle cells, involved in cell signalling with the oocyte and secretion of the eggshell, surround the cyst²⁴⁵ and the oocyte adheres to the posterior follicle cells via Cadherin signalling.²⁴⁶⁻²⁴⁷ Maturing egg chambers are connected in the ovariole by stalk cells which also signal between the egg chambers.²⁴⁸ These cells are specified by the germline through Delta signalling which instructs the anterior follicle cells to adopt a polar cell fate. The polar cells then signal to neighbouring egg chambers to form stalk cells between the egg chambers.²⁴⁹⁻²⁵¹ Encapsulation of the oocyte in follicle cells defines stage 1 of oogenesis. During this stage, the MTOC is at the anterior of the oocyte with the microtubule plus-ends extending into the NCs through the ring canals.¹¹¹ As the egg chambers mature, the microtubule network

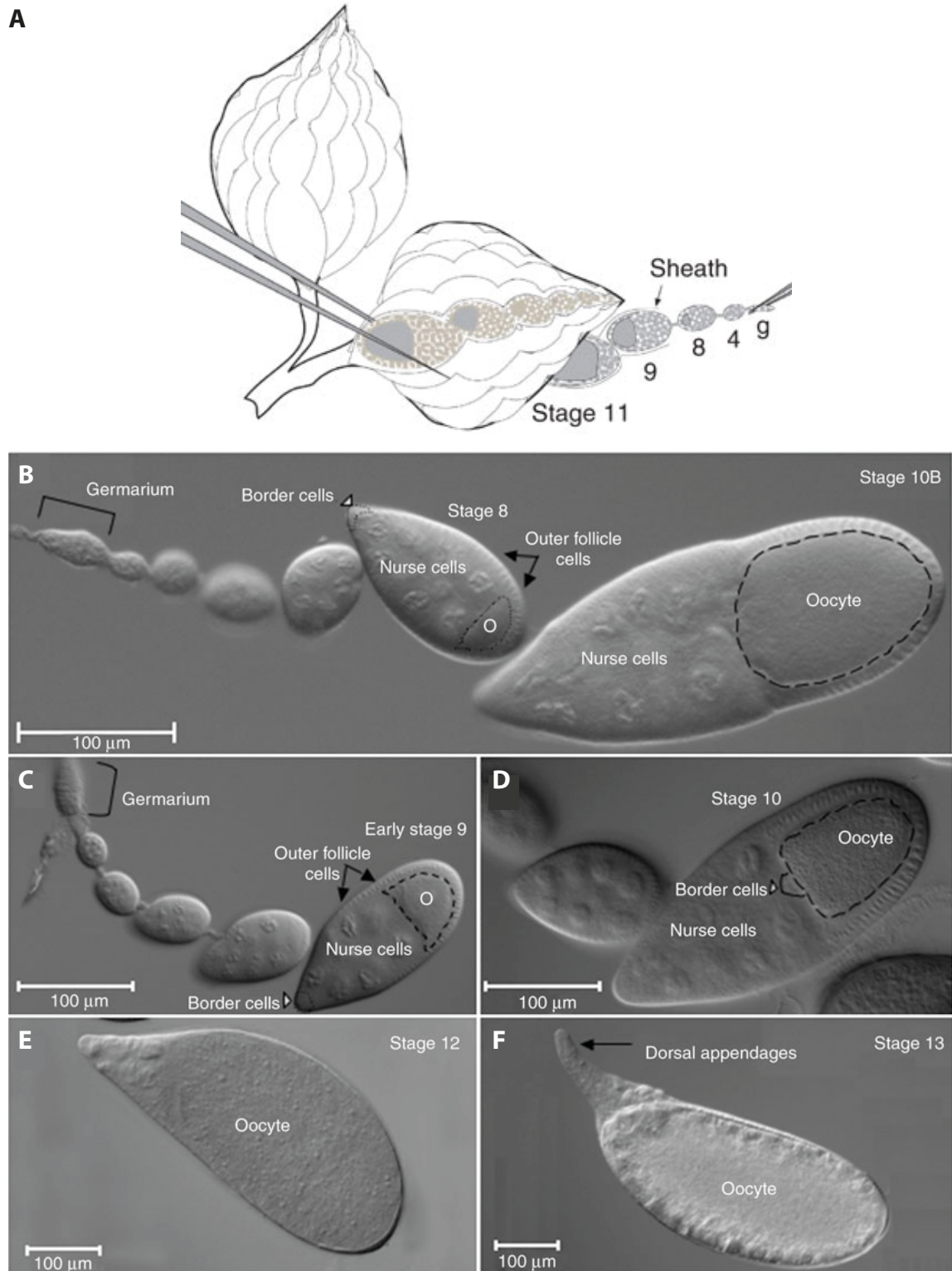


Figure 1.4. Stages of *Drosophila* oogenesis (Adapted from Prasad *et al.*, 2007). (A) Schematic of the *Drosophila* ovaries, each composed of 16-20 ovarioles. Ovarioles can be teased from the ovary and contain egg chambers at different stages of development. (B-F) Normarski images of ovarioles containing different stages of egg chambers ranging from the germarium to stage 13. The anterior is positioned to the left and the oocyte is outlined in stages 8-10B. Arrows point to the follicle cells or dorsal appendages in stage 13, and arrowheads to the border cells. (O = oocyte).

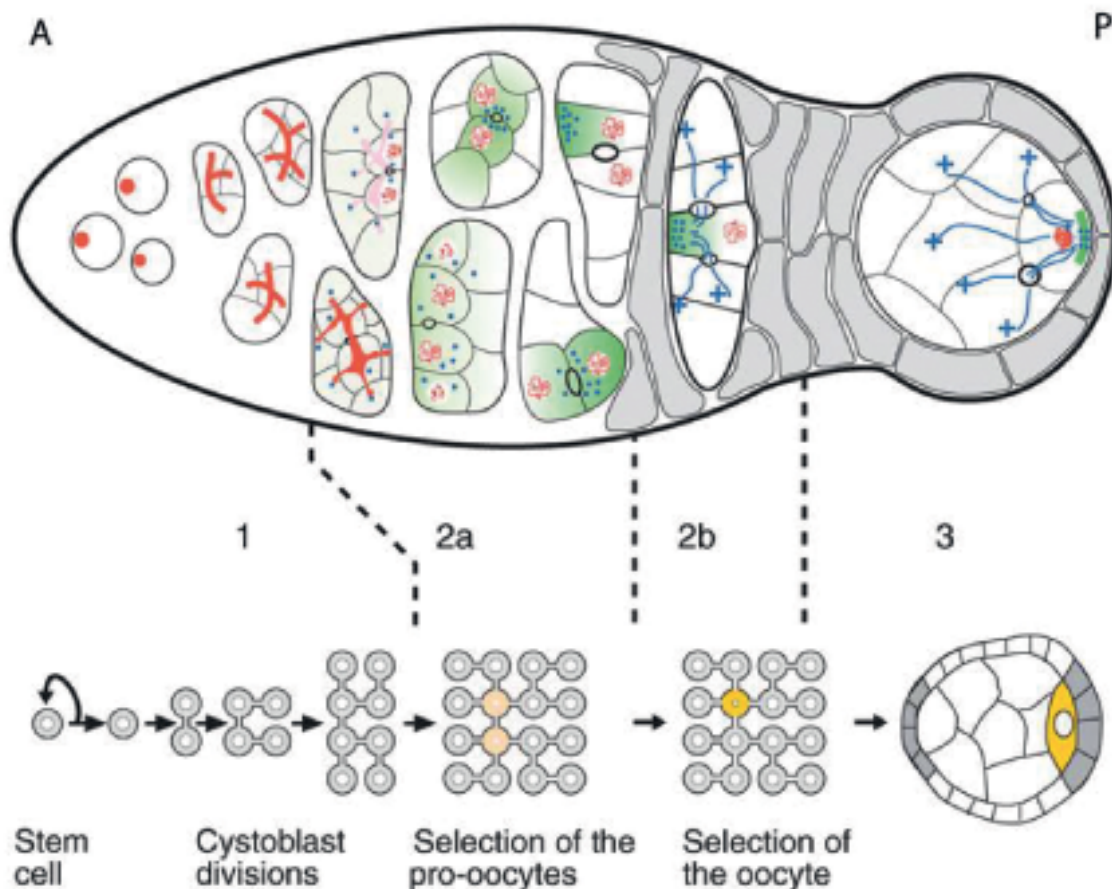


Figure 1.5. Oocyte specification in the *Drosophila* gerarium (Taken from Huynh and St Johnston, 2004). In region 1 at the anterior of the gerarium, germline stem cells (GSCs) divide to produce a cystoblast and a new GSC. Both the stem cells and the cystoblast contain a spectrosome (red circles) which develops into the fusome (red structures) and orients each division of the cyst. The cystoblast divides a further four times with incomplete cytokinesis, producing a 16-cell germline cyst interconnected by ring canals and the fusome. In region 2a, the two cystoblasts with four ring canals are called the pro-oocytes (yellow) as one of them will differentiate into the oocyte. Both pro-oocytes contain the synaptonemal complex (red lines), enter meiosis, and accumulate cytoplasmic proteins, mRNAs, mitochondria (green), and the centrosomes (blue circles) at the anterior of the oocyte. At region 2b, the single cell selected to be the oocyte remains in meiosis, and the minus-ends of microtubules become orientated in the oocyte with the plus ends extending into the nurse cells. The follicle cells (grey) migrate and surround the germline cells. At region 3, the oocyte adheres to the posterior follicle cells and the microtubule network repolarises so the minus-ends are extending towards the posterior of the oocyte. As a result, cytoplasmic components accumulate at the posterior of the oocyte. (A= Anterior, P= Poster).

is repolarised, moving the MTOC and the oocyte nucleus to the posterior of the oocyte. As a result of the microtubule minus ends being positioned at the posterior of the oocyte, several NC-derived cytoplasmic factors, such as axis-determining transcripts, enter the oocyte in a Dynein-BicD-Egl driven process and accumulate at the posterior.^{111,118-119,235,238,241,252-256} Once the egg chambers are produced they round up to form a sphere with the oocyte positioned at the posterior and bud from the germarium. The egg chambers then enter the vitellarium, marking stage 2 of oogenesis, where they continue to develop into a mature egg.

Development of the *Drosophila* egg chamber

Drosophila oogenesis is divided up into 14 morphologically defined stages which take approximately one week to complete (Figure 1.6).^{234,257} Stages 1-6 are referred to as early oogenesis, stages 7-10 as mid oogenesis, and stage 11-14 as late oogenesis. Several egg chamber features are used to determine their stage including egg chamber size, oocyte size/egg chamber ratio and the distribution of follicle cells.

Throughout stages 1-6 both the oocyte and the NCs are continually growing. The NC nuclei increase in size by rounds of endoreplication making them highly polyploid, while the oocyte nucleus moves through meiotic prophase and forms a compact structure called the karyosome. At stage 6, the follicle cells stop proliferating, undergo three rounds of endoreplication and growth, and separate into two distinct groups capable of responding to different signals; terminal follicle cells at the anterior and posterior poles, and mainbody follicle cells.^{248,258-261} At the same time, the symmetry of the oocyte is broken by Grk signalling to the overlying terminal follicle cells at the posterior of the oocyte, the posterior MTOC disassembles, and the microtubule network is reorganised.²⁴⁷ Reorganisation of the microtubules was initially thought to produce a highly polarised network with the minus ends at the oocyte anterior and the plus ends at the posterior.¹⁰⁹⁻¹¹¹ However more recent evidence suggests that a biased random microtubule organisation is instead achieved through PAR-1 signalling, resulting in more plus ends at the posterior as a result of higher levels of

microtubule nucleation at the anterior.¹¹⁵ Nevertheless, the resulting changes in the cytoskeleton are required for the migration of the oocyte nucleus from the posterior of the oocyte to the DA.²⁶²⁻²⁶³

At stage 7, the size of the oocyte increases at a greater rate than the NCs. This is achieved through selective transport of proteins, mRNA and organelles from the supporting NCs into the oocyte, and from the uptake of yolk protein produced by the surrounding follicle cell layer at stage 8. At the beginning of stage 9, the anterior polar cells recruit 4-6 nearby follicle cells to become border cells which migrate posteriorly between the NCs towards the oocyte anterior, and up to the dorsal corner near the nucleus.²⁶⁴ By stage 10 the oocyte has grown so big that it occupies half of the egg chamber.

Stage 10 is split into two sub-stages; 10a and 10b. During stage 10a the mainbody follicle cells migrate posteriorly to cover the oocyte and are converted to columnar cells, forming a secretory epithelium that deposits the eggshell. Approximately 50 anterior terminal cells convert to squamous stretch cells and stretch flat to cover the NCs. Stage 10b is marked by a group of columnar follicle cells, called centripetal cells, migrating inward between the oocyte and the NCs, becoming elongated, and surrounding the oocyte. During this stage, the cytoskeleton is reorganised and the cytoplasm begins “streaming”, mixing the ooplasm. This process aids the formation of the germ plasm at the posterior of the oocyte to establish the germ line.^{128,73,265-266} In addition, the NCs begin to rapidly transport cytoplasmic factors into the oocyte in a non-selective manner in a process termed NC dumping. By stage 13, specialised respiratory structures called dorsal appendages form from a subset of DA follicle cells as a result of earlier *grk*-dependent signalling processes.²⁴⁸ When egg maturation is complete, the NCs and follicle cells apoptose, and the NC nuclei are cleared, marking the start of late oogenesis.^{234,267} By stage 14, the entire egg chamber consists of the large oocyte, and the egg is prepared for activation and depositing.

grk localisation in the *Drosophila* oocyte

In the *Drosophila* oocyte, several maternal transcripts, including *grk*, *bcd*, and *osk*, localise to distinct regions of the cell and are responsible for establishing the major body axes of the future embryo (Figure 1.6). *grk* is involved in establishing both the anterior-posterior and the dorsal-ventral axes of the future embryo through two spatially and temporally distinct rounds of translation during oogenesis. Although *grk* was originally thought to be transcribed in the oocyte nucleus,²⁶⁸ it has since been shown to be produced in the NCs and transported into the oocyte during early oogenesis.²⁶⁹⁻²⁷⁰

Between stages 1-7, a combination of the microtubule organisation, motor proteins and other *trans*-acting factors ensure that several of the cytoplasmic factors transported from the NCs, including *grk*, accumulate at the posterior of the oocyte where the oocyte nucleus resides.^{252,253} In this way, *grk* forms a crescent between the nucleus and the overlying follicle cells, and is translated here during stages 2-6 (Figure 1.6C).^{15,271} *grk* encodes a TGF- α -like secreted protein which signals through the epidermal growth factor receptor (EGFR) known as Torpedo (Top).^{15,272} At stage 6, Grk signalling to the adjacent posterior terminal follicle cells causes them to express different markers and adopt a different cell fate to the rest of the follicle cells with a default anterior fate.^{247,273-274} The posterior follicle cells signal back to the oocyte through an ill-defined mechanism thought to involve the Salvador Warts Hippo (SWH) pathway.²⁷⁵⁻²⁷⁶ This results in the disassembly of the MTOC, repolarisation of the microtubule network, and the formation of a new diffuse MTOC at the anterior of the oocyte, marking stage 7 of oogenesis.¹¹⁰⁻¹¹¹

These same rearrangements of the microtubule network result in the transport of *bcd* mRNA to the anterior of the oocyte by Dynein (Figure 1.6A).²⁷⁷⁻²⁷⁸ *bcd* localisation is required for the formation of a morphogenetic gradient of Bcd protein in the early embryo which specifies the head and thorax.^{279,280} *osk* mRNA is also localised following microtubule reorganisation but is transported to the posterior of the oocyte by Kinesin (Figure 1.6B).²⁸¹ Movement of *osk*

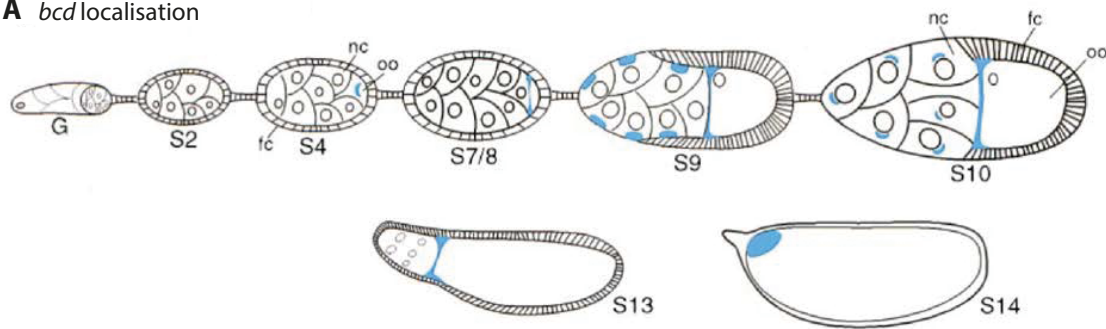
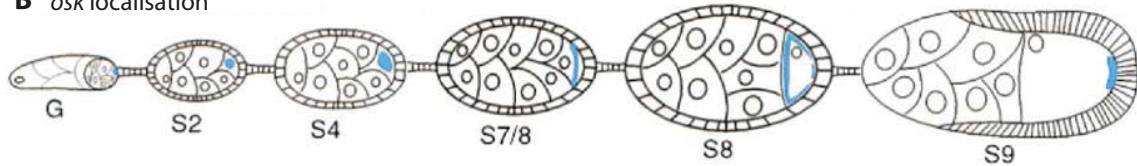
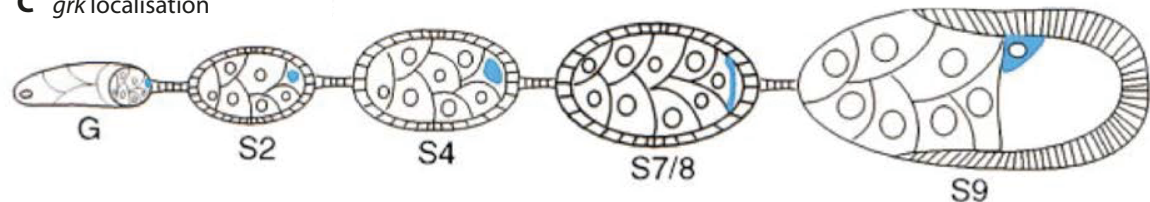
Anterior**Posterior****A** *bcd* localisation**B** *osk* localisation**C** *grk* localisation

Figure 1.5. mRNA localisation in the *Drosophila* oocyte (Adapted from Lasko 1999). (A) During stages 4-6 of oogenesis *bcd* transiently accumulates at the posterior of the oocyte before it moves to the anterior during stages 7-8. At stage 10, *bcd* is highly abundant at the anterior of the oocyte where it is required for head and thorax specification and patterning of the embryo. (B) *osk* mRNA can be found in the oocyte very early in oogenesis, and transiently localises to the centre and the anterior of the oocyte at stages 6-8. From stage 9, *osk* mRNA is relocalised to the posterior of the oocyte where it is required for establishing the germline and posterior structures. (C) Accumulation of *grk* mRNA is seen in the oocyte in the germarium and localises to the posterior of the oocyte from stage 6 where it establishes the anterior-posterior axis. By stage 9, *grk* mRNA is localised around the oocyte nucleus at the dorsal-anterior of the oocyte where it is responsible for establishing the dorsal-ventral axis and dorsal structures. Numbers = stage (S), germarium (G), nurse cells (nc), follicle cells (fc), oocyte (oo).

mRNA is required for germ cell specification, the generation of posterior structures, and recruitment of many proteins and mRNAs to the posterior pole to assemble the germ plasm.^{20,282} Finally, the microtubule rearrangement pushes the oocyte nucleus from the posterior to the presumptive DA corner.^{252,256,263}

As the nucleus migrates, *grk* transiently accumulates at the anterior margin at stage 8, before localising to form a cap above the nucleus at the DA at stages 9-11.^{15,120,283} Local translation and secretion of *grk* during mid-oogenesis results in the second round of Grk signalling to specify a dorsal fate in the overlying mainbody follicular epithelium.^{15,272,274,284-290}

Dorsal follicle cell signalling restricts *pipe* expression to the ventral follicle cells in a largely unknown mechanism involving a 31 base-pair *cis*-regulatory element in *pipe*.²⁹¹⁻²⁹⁵ Pipe forces the follicle cells to adopt a ventral cell fate through Toll signalling, and restricts Dorsal protein localisation to the nucleus.²⁹⁶⁻²⁹⁷ This signalling results in the correct specification and positioning of anterior and dorsal structures, such as the two dorsal appendages that form due to activation of Argos, a repressor responsible for splitting the EGFR signal in two.^{286,298}

Overall, the importance of both rounds of Grk signalling in axis specification can be seen in Grk mutants, or flies containing overexpressed or mislocalised Grk. Mutations preventing the first phase of Grk signalling during early oogenesis disrupt the anterior-posterior patterning of the oocyte due to aberrant microtubule reorganisations at stages 7-8, mislocalisation of *bcd* and *osk* mRNAs, and a failure of the oocyte nucleus to migrate to the anterior.²⁷³⁻²⁷⁴ The importance of the second round of Grk signalling for axis specification is seen in Grk mutants, where dorsal appendages fail to form generating a ventralised phenotype, and in overexpressed or mislocalised mutants where ectopic appendages form generating a dorsalised phenotype.^{15,299-300}

The mechanism of *grk* localisation

Both rounds of *grk* localisation require several components including an intact cytoskeleton, molecular motors, a *cis*-acting localisation signal, and *trans*-acting factors. Experiments using *in vivo* injection of fluorescently labelled *grk* mRNA have shown that *grk* is actively transported during stage 8 to the DA of the oocyte in a two-step process. First, *grk* mRNA is transported from the site of injection to the anterior. Second, *grk* mRNA at the anterior of the oocyte begins to move to the DA and forms a cap above the nucleus. Both steps require microtubules and Dynein but are thought to be distinct, using different mechanisms, and possibly different networks of microtubules.^{120,301} Kinesin-1 is also required for the DA localisation of *grk*, possibly through the recycling of Dynein back to the posterior of the oocyte for subsequent rounds of transport.³⁰²

The role of Dynein in *grk* active transport is exemplified by the fact that over expression of Dynamitin results in the disruption of *grk* localisation,^{91,262} and that the Dynein light chain (Dlc) DDLC1 directly interacts with *grk*.³⁰³ In addition, the highly dynamic *grk* particles found at the posterior during early oogenesis, and the anterior in mid-oogenesis, contain several molecules of the Dynein associated proteins Egl and BicD, and require Dynein for their dynamic behaviour.^{133,283} Egl has been shown to directly interact with the *cis*-acting localisation signals in its cargo mRNAs, including *grk*, linking the mRNA to the Dynein motor.³⁰⁴ Transport of *grk* mRNA from the NCs in early oogenesis requires microtubules and recruitment of the Dynein-BicD-Egl complex.²⁷⁰

As oogenesis progresses, the dynamic *grk* mRNA transport particles at the anterior of the oocyte were reported to become more static and anchored at P bodies specifically at the DA, but not in the NCs or the oocyte cytoplasm.^{133,283} This conversion of transport particles to P bodies was shown to require the switching of Dynein from a processive motor to a static anchor by the hnRNP Sqd.¹³³ However, later studies investigating the role of P bodies in the translational regulation of *grk* found no evidence to suggest that transport particles are involved in the assembly of P bodies.¹⁹³

Despite studies showing the route and mechanism of *grk* active transport, the origin of DA localised *grk* is currently unknown. Possible models include transport of the early posterior-localised *grk* to the anterior during oocyte nucleus migration, or degradation of the early *grk* mRNA pool, and the localisation of new *grk* transcripts from the NCs to the DA for the second round of Grk signalling.²⁹ Moreover, alternative models of *grk* transport have also been described involving multiple rounds of anterior transport and trapping at the DA corner, but these models are less widely accepted.³⁰⁵

The *grk* localisation signal (GLS)

Initial experiments searching for *cis*-acting elements required for *grk* mRNA localisation and translational regulation revealed several regions to be important.^{268,306} For example, the initial accumulation of *grk* mRNA in the oocyte was shown to require sequences within the 5'UTR, whereas regions within the coding region and 3'UTR were shown to be required to restrict *grk* from a transient anterior ring to the DA in mid to late egg chambers.³⁰⁶ In addition, others have proposed that regions in the 5'UTR and coding region are required for stable DA localisation in late oogenesis, and regions in the 3'UTR are required to restrict *grk* to the DA of the oocyte.²⁶⁸

Additional studies using *in vivo* injection assays and GLS-fusion transgenes expressed in the NCs further focussed the mapping of the *grk* localisation element to a 64-nucleotide stem-loop in the coding region of *grk*, overlapping with a region encoding the signal peptide.³⁰⁷ In contrast to the complex array of localisation elements previously mapped for *grk* localisation, the authors showed that this small signal, termed the *grk* localisation signal (GLS), is both necessary and sufficient to drive all steps of *grk* localisation; from the NCs into the oocyte, to the posterior, and finally to the DA of the oocyte (Figure 1.6).³⁰⁷ However, GLS-containing transgenes showed more diffuse localisation at stage 10 than endogenous or injected *grk*, suggesting that stable *grk* localisation during late oogenesis may indeed require other regions of the transcript.³⁰⁷ In addition, later studies suggested that the GLS is only sufficient

to drive *grk* to the anterior of the oocyte, with DA localisation requiring an additional undetermined localisation element.³⁰⁵

The numerous discrepancies throughout the literature regarding the role of different regions of the *grk* transcript in its localisation are likely a result of differences in the experimental procedures used in each case, such as the specific transgenes used, or differences between injected and endogenously expressed RNA. Despite these differences, the GLS is generally accepted to be sufficient for all stages of localisation for injected *grk* RNA, and will be the main focus of this thesis.

mRNA secondary structure folding predictions revealed that the GLS forms a highly stable stem-loop (Figure 1.7B) in several *Drosophila* species. In addition, injection of a 170-nucleotide region of *grk* including the GLS is able to localise to the DA of the *D. virilis* oocyte, suggesting that both the localisation signal and the transport machinery are highly conserved among the *Drosophilids*. Hybrid injection experiments showed that fusing the GLS to either *lacZ* or *hunchback* (*hb*), an mRNA involved in thorax development during embryogenesis that does not localise in the oocyte,³⁰⁸ results in their transport to the DA of the oocyte, whereas removing the GLS from full-length *grk* abolishes its localisation and prevents the transcript from forming transport particles.^{133,307} Transport of the GLS was shown to require Dynein and microtubules, and occurs in the same two-step process previously described for *grk* mRNA localisation to the DA of the oocyte.^{120,133} Moreover, the GLS is able to drive *grk* transport apically in the embryo, as previously shown for full-length *grk* mRNA.^{116,120,307} The function of the GLS is further discussed in subsequent chapters and will be the main focus of this thesis.

Trans-acting factors involved in *grk* localisation and translation

Several *trans*-acting factors that regulate *grk* localisation and translation have been identified through genetic and biochemical studies. Although some proteins have been shown genetically to be required for *grk* localisation due to their roles in maintaining the

cytoskeleton, such as *cappuccino* (*capu*), *spire* and *maelstrom* (*mael*),^{5, 309-311} several others have been shown to affect *grk* localisation more directly. For example, in *sqd* and *K10* mutants, *grk* is mislocalised along the anterior of the oocyte during mid-oogenesis, resulting in ectopic translation of Grk and causing the production of dorsal appendage material around the entire circumference of the egg.^{15, 299, 312-313} The same phenotype is seen when *grk* is overexpressed with *grk* accumulating at, and being ectopically translated along, the anterior of the oocyte.¹⁹³

Biochemical studies have identified several *trans*-acting factors that associate with *grk* mRNA and are required for its localisation, including Sqd, Bicaudal-C (BicC), Hrb27c, Ovarian tumour (Otu), Cup, Glorund (Glo), Syncrip (Syp) and Egl.^{54, 304, 316} In the absence of these proteins, *grk* is mislocalised along the anterior of the oocyte, again causing dorsal patterning defects.^{15, 54, 162, 314, 315, 317} Sqd has been shown to bind to the 3'UTR of *grk* and recruits Bruno, Cup, Hrb27c and Otu to form a complex that represses *grk* translation.^{161, 162, 314, 317} This complex is also likely to contain Glo and Half-pint (Hfp).¹⁶² Activation of *grk* translation has been shown to require Vasa, PABP55B, Encore (Enc), Orb (Oo18 RNA-binding protein), and the spindle genes.^{193, 314, 318-323}

GRNA affinity chromatography was used to determine which *trans*-acting factors interact with the GLS and identified 16 proteins which specifically associate with the *grk* 5'ORF containing the GLS.^{54, 324, 325} Several of the proteins were previously linked to *grk* mRNA localisation and translational regulation, including Sqd, Imp, PTB, PABP, Me31B, and BicC. They also identified the novel RBPs Syp, which they showed associates specifically with both *grk* and *osk* mRNAs, as well as Sqd and Hrb27c, and is essential for the correct localisation and translation of *grk*. A full list of the proteins identified can be found in the Appendix A, and the roles of the *trans*-acting factors Egl, Sqd and Me31B in *grk* localisation and translational regulation are further discussed in Chapters 4, 5 and 6 of this thesis, respectively.

The regulation of *grk* translation

During *grk* transport in the oocyte, translation of the mRNA must be finely regulated to ensure Grk is not ectopically expressed. Translational silencing of both *grk* and *osk* was initially thought to involve the binding of Bruno to a BRE-like sequence in the 3'UTR of the transcript.^{160,161} However, more recent studies have proposed alternative models. For example, Me31B has been implicated in the translational regulation of several maternal mRNAs, including *grk*, during their transport into the oocyte and at the DA of the oocyte.^{193,326}

At the DA of oocytes, *grk* is anchored at the edge of P bodies in a Sqd-dependent manner.^{133, 193} Here, translational activators such as Orb are highly enriched, and ribosomes are easily accessible, resulting in translation of Grk.¹⁹³ Orb (CPEB) is involved in the polyadenylation of *grk* mRNA, through recruitment of the polyadenylase Wispy, specifically at the DA of the oocyte where Orb is hyperphosphorylated by CK2.^{404,411} Orb-dependent lengthening of the *grk* poly(A) tail at the DA is required to enhance translation of the transcript, resulting in localised Grk protein expression. Therefore, in an *orb* mutant Grk levels are reduced and *grk* mRNA is mislocalised along the anterior of the oocyte.^{327,328}

In the NCs, silencing of *grk* translation has been shown to be the result of spatially restricting access to Orb, preventing the polyadenylation of *grk* and its association with P bodies.^{329,404} The role of P bodies in the translational regulation of *grk* is further exemplified by studies showing that over expressed *grk* is targeted to the core of P bodies which is devoid of ribosomes and is enriched in translational inhibitors such as Me31B.¹⁹³ Finally, *grk* accumulates in P bodies in response to starvation,³³⁰ indicating their involvement in the activation and repression of *grk* translation under different cellular conditions.

The / Factor

The / factor is a non-LTR retrotransposon that encodes the two proteins ORF1p (nucleic acid binding) and ORF2p (predicted endonuclease, reverse transcriptase and RNase H activities).

^{331, 332} *I* factor mRNA localises to the DA and the anterior of the oocyte at stages 8 and 9 of oogenesis in a similar manner to a combination of both *grk* and *bcd* mRNA,³³³ and causes the mislocalisation of *grk*, resulting in weakly ventralised egg defects similar to *grk* mutants.
^{307, 334} Localisation of *I* factor to the DA of the oocyte is thought to be required for entry of the mRNA into the nucleus for transposition events to occur.

Characterisation of the *I* Factor mRNA using an *in vivo* injection assay identified a 58-nucleotide *cis*-acting stem-loop, known as the *I* Factor localisation signal (ILS), that is both necessary and sufficient to drive *I* Factor localisation to the DA of the oocyte (Figure 1.7).³⁰⁷ The predicted secondary structure of the ILS suggests it forms similar structural elements to the GLS. Furthermore, more recent studies have identified additional GLS-like signals in the transposons G2 (G2LS) and *Jockey* (JLS), which are capable of driving mRNA localisation to the anterior or DA of the oocyte (Figure 1.7).³³⁵

Features common to these GLS-like signals may provide consensus recognition motifs for the localisation machinery and *trans*-acting factors that support DA localisation. Use of a selection of the same *trans*-acting factors is exemplified in studies showing that the ILS competes with the GLS and *grk* mRNA for *trans*-acting factors during the two-step localisation process.³⁰⁷ In addition, *Sqd* has been implicated in the anchoring process of the *I* Factor, as the ILS localises along the anterior of the oocyte in a *sqd1* mutant with no sign of a DA cap.³⁰⁷ Despite the similarities between *grk* and *I* factor their localisation speeds differ, with *I* factor localising more efficiently due to the presence of a second stem-loop called the A loop.⁶⁹

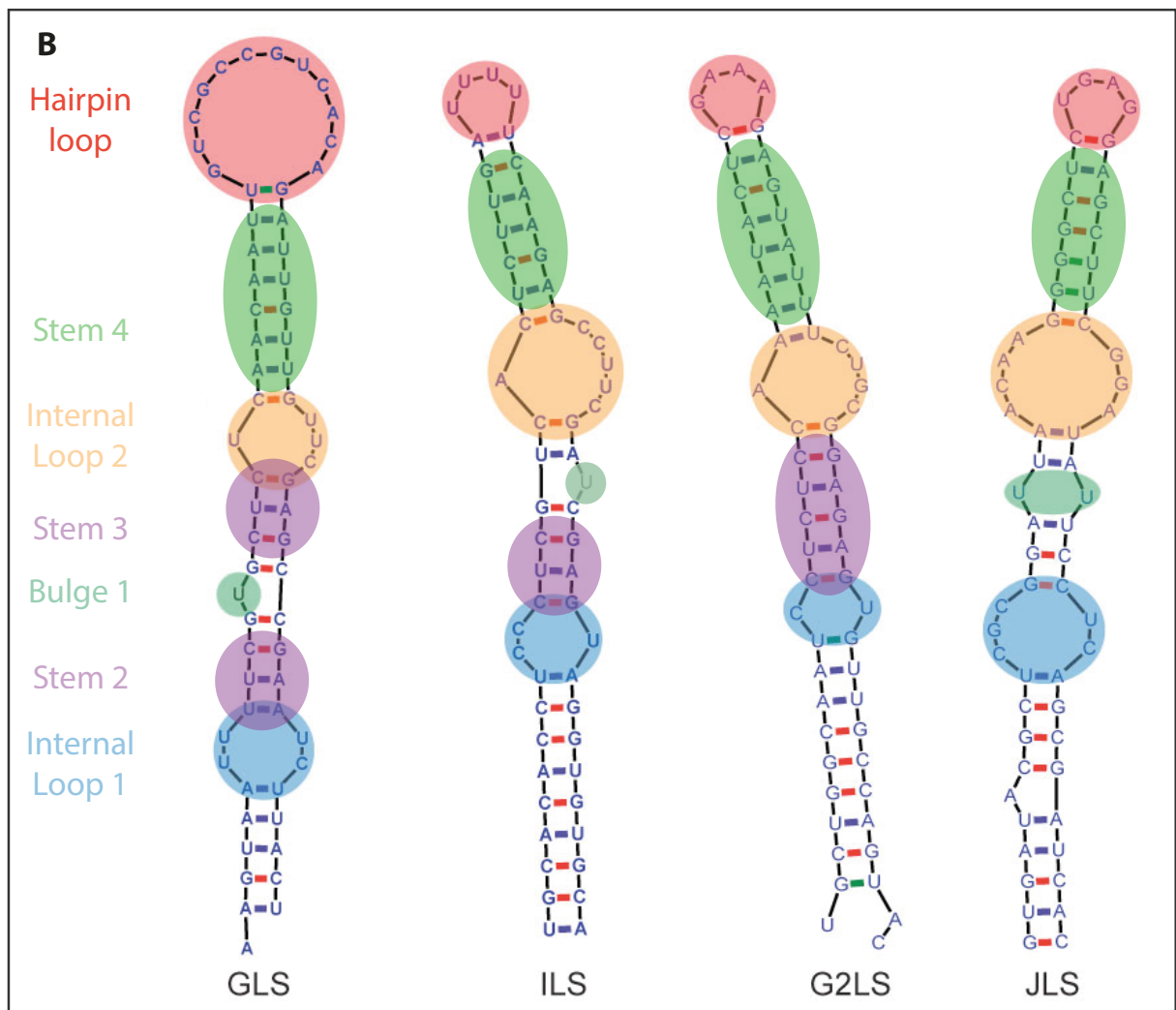
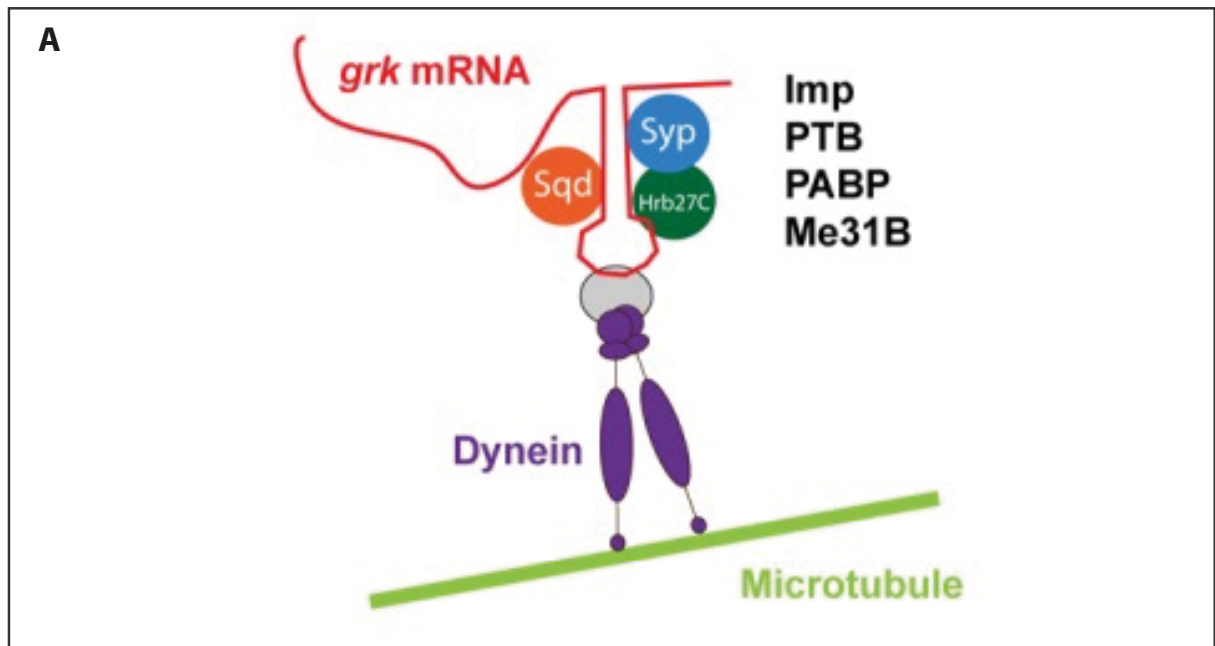


Figure 1.7. The GLS recruits several *trans*-acting factors in the *Drosophila* ovary and shares conserved features with other DA localising signals (Adapted from McDermott et al. 2012 and Hamilton et al. 2009). (A) The *grk* localisation signal (GLS) recruits a plethora of *trans*-acting proteins to *grk* mRNA, including Squid (Sqd), Syncrip (Syp) and Hrb27c, and is localised via the Dynein motor along microtubules. (B) The mFold predicted secondary structure of the GLS reveals several conserved features with the other DA localising signals ILS, G2LS and JLS. Conserved features include a hairpin loop (pink) followed by a long stem 4 (light green), two internal loops (yellow and blue), small stem regions (purple) between the internal loops, and bulges (dark green).

mRNA localisation in the *Drosophila* embryo

Fertilisation of the egg results in 13 rapid synchronous mitotic divisions consisting of DNA synthesis and mitosis. However, cytokinesis does not occur resulting in the formation of a syncytium containing approximately 8000 nuclei. These nuclei migrate peripherally through the maternally inherited cytoplasm and cellularise, changing the syncytial blastoderm embryo into a cellular blastoderm. During this time, developmental control of the embryo is switched from the maternal to the zygotic genome. Several mRNAs have been shown to localise in the embryo, with the most characterised being the pair-rule transcripts responsible for the segmentation of the embryo. These mRNAs localise to the apical cytoplasm above the nuclei in the syncytial blastoderm. Much of the studies investigating mRNA localisation in the embryo utilise an *in vivo* injection assay of fluorescently labelled mRNAs. This technique was used to show that pair-rule transcripts are localised in a process that requires microtubules, Dynein, Egl and BicD. Moreover, mRNAs shown to use this network of transport in the oocyte, such as *grk*, are also able to localise apically in the embryo due to the high conservation of the transport machinery.¹¹⁶

Studying mRNA localisation

Visualising localised mRNAs

There are many ways to visualise mRNA localisation in the cell. Historically this involved using *in situ* hybridisation (ISH) techniques in fixed samples. The ISH technique has been

recently improved to allow single-molecule labelling through the use of multiple probes, increasing the signal.³³⁶ However, although ISH sheds light on where mRNAs are localised at specific time points, it does not provide information about the dynamics of the process.

Several techniques have since been developed to visualise mRNAs in real-time in live cells. First, labelling of specific transcripts can be achieved through the incorporation of fluorescent nucleotides by *in vitro* transcription, followed by microinjection of the RNA. This injection technique is quick and easy and has been used successfully in *Drosophila* to study the localisation of *grk* RNA in the oocyte, and pair-rule transcripts in the embryo.^{116, 337-339} However, some mRNAs that require nuclear processing are unable to localise unless injected into the nucleus. Second, endogenous mRNAs can be labelled through the injection of molecular beacons, but only a single fluorophore is introduced, making the signal weak. Third, the MS2-MCP system involves exogenously expressing MS2-loop-tagged mRNAs which recruit and bind fluorescently-labelled MS2 phage coat protein (MCP).³⁴⁰ Improvements to this technique include the addition of increasing numbers of MS2-loops on the mRNA to increase the occupancy of MCP, and adding a nuclear localisation signal (NLS) to reduce the background MCP signal, therefore improving imaging of the mRNAs. For a recent review on visualising mRNA localisation see Ref ³⁴¹.

Characterising mRNA localisation signals

Characterising localisation signals within an mRNA requires in-depth structure-function analysis through truncations of native mRNA sequences, or the addition of specific regions of the mRNA to reporter transcripts to establish which regions of the mRNA are necessary and sufficient for driving localisation. These studies often involve microinjection of the mRNA to assess for localisation of the transcripts.

β-actin mRNA was one of the first transcripts shown to be sub-cellularly localised when it was identified at the leading edge of chicken embryo fibroblasts.²³ Further characterisation of the mRNA showed it contains a 54-nucleotide *cis*-acting zipcode,

immediately following the termination codon in the 3'UTR,³⁴² which binds ZBP1 (zipcode-binding protein 1, IMP1), an RBP essential for maintaining β -actin mRNA localisation.³⁴³
³⁴⁴ *Trans*-acting factors bind to specific recognition motifs within the *cis*-acting signals of their target mRNAs. Recent work focused on identifying such motifs within β -actin mRNA have demonstrated that the third and fourth KH domains of ZBP1 interact with the bipartite mRNA element composed of a 5' CGGAC sequence, followed by a variable 3' C/A-CA-C/U that must be spaced appropriately. Using this consensus binding site the authors were able to identify other transcripts that could associate with ZBP1.³⁴⁵ However, defining the specific motifs required for the interaction of a given localisation signal and its protein partners is not trivial.

The sequence and predicted secondary structure of characterised localisation signals vary even among those destined for the same location within the cell, making it extremely difficult to define consensus motifs across many mRNA species. An exception is the localisation signal from *Drosophila* K10 which localises to the anterior of the oocyte during mid-oogenesis. K10 contains a 44-nucleotide localisation signal called the transport/localisation element (TLS) which was also found in the 3'UTR of *orb* mRNA, suggesting the two mRNAs utilise the same localisation signal and transport machinery.
³¹² Despite best efforts to characterise other mRNA localisation elements in this way, consensus motifs responsible for protein binding or for driving groups of transcripts to the same destination have yet to be defined in most instances, due to the complexity of the elements and redundancy in the system.

Studying mRNA-protein interactions

The structure of RNAs

Despite being produced as single-stranded molecules, large RNAs form a multitude of secondary and tertiary structures through hydrogen bond base-pair interactions between single-stranded complementary sequences. These structures, along with the primary

sequence, are required to provide specific recognition motifs for *trans*-acting factor protein binding. Watson-Crick base-pairings (e.g. A-U, U-A, G-C, C-G) generate A-conformation secondary structures in the mRNA associated with a deep and narrow major groove, and a shallow minor groove. In this conformation, the discriminatory edges of the bases are buried, reducing the accessibility and recognition of the mRNA by *trans*-acting factors. However, non-canonical base-pairs have also been shown to form (e.g. U-G).³⁴⁶

Due to the limited number of packing arrangements available by base-pair stacking, other structural motifs are required to generate specificity and diversity for protein binding. Base-pairings within the mRNA can generate several features including stems, hairpin loops, internal loops and bulges, and pseudoknots (Figure 1.8).³⁴⁷ Bulges are non-helical motifs consisting of one or more single-stranded nucleotides on the same side of a stem due to their inability to form a base-pair in the double-helix (for a review see Ref ³⁴⁸). As a result, the nucleotide is extruded from the structure which can cause a distortion in the backbone, widening the major groove and increasing the accessibility for *trans*-acting factor binding.

In addition to the numerous secondary structures that an mRNA can form, tertiary structures can be generated by organising the secondary structures, such as stem regions, in three dimensions. It is becoming more clear that the three-dimensional structure of an RNA may be more highly conserved than the secondary structure, and more subject to natural selection. In fact, structural studies have shown that secondary structures, such as internal loops and hairpin loops, can differ in terms of the number of nucleotides involved, but can in fact adopt similar three-dimensional structures, except for differences in the number of single-stranded bases.³⁴⁹

The secondary structure of an mRNA can be predicted fairly accurately from the primary sequence information using thermodynamics and comparative sequence analyses (for a review see Ref ³⁵⁰). For example, UNAFold uses the primary sequence of a single mRNA and predicts the maximum number of base-pairing in the structure before calculating the

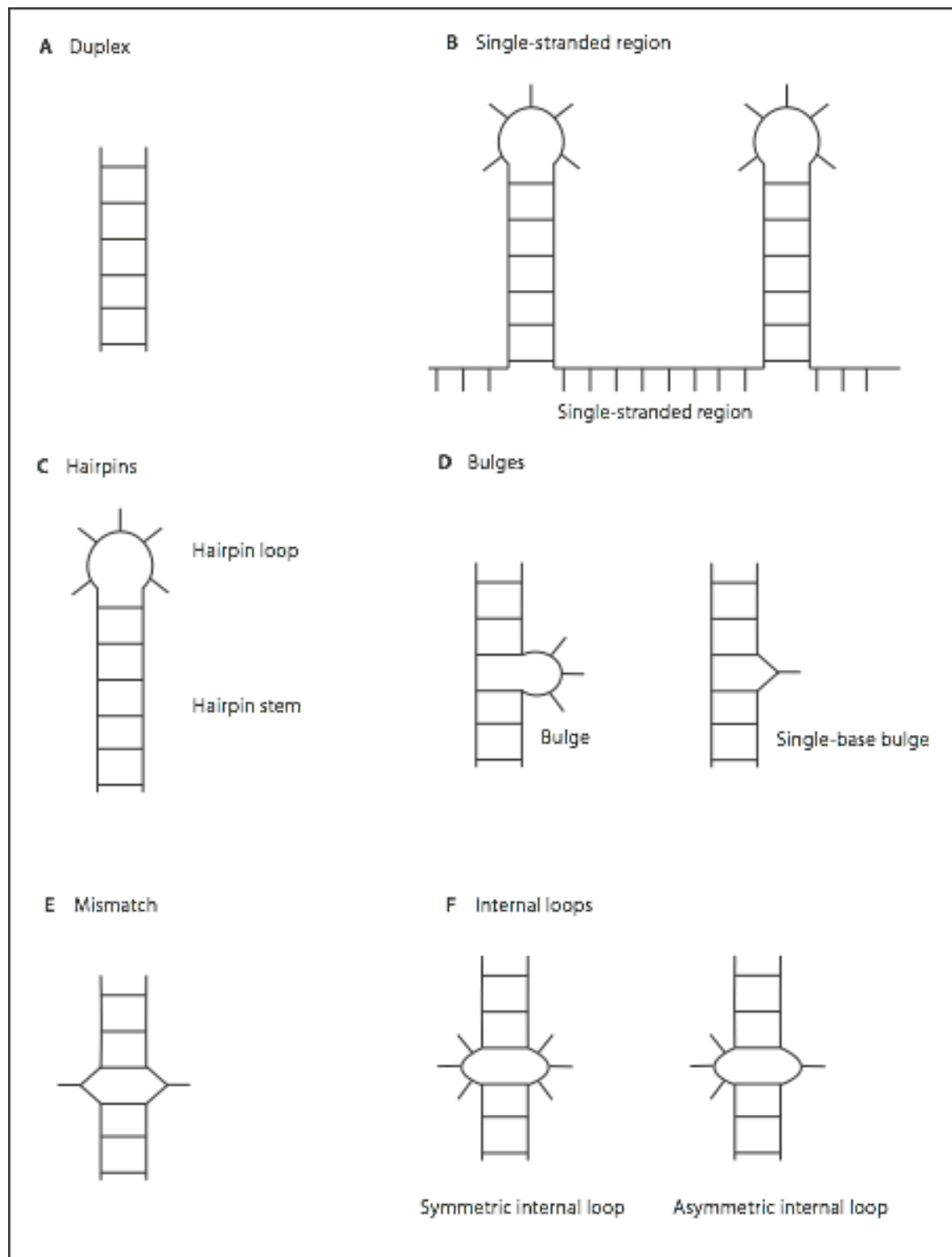


Figure 1.8. RNA secondary structure motifs (Adapted from Chastain and Tinoco, 1991). (A) Double-stranded RNA duplexes can be formed from complimentary single-stranded regions of RNA. (B) Single-stranded RNA. (C) Stem-loops are formed from a hairpin stem and a single-stranded hairpin loop. (D) Bulges are formed from a single non-base paired nucleotide or several looped out on the same side of a stem. (E) Non-complimentary base-pairs within a stem form a mismatch pair. (F) Internal loops are formed from more than one non-base paired nucleotide on both sides of the stem.

minimum free energy (MFE).³⁵¹ However, the accuracy of secondary structure prediction can be improved using software, such as *RNAalifold*, that aligns homologous sequences and utilises covariance information to predict a consensus secondary structure.³⁵²

Identifying RNA-binding sites in RNAs

RBPs often recognise specific sequences in their RNA targets. However, due to the A-form helical structure of double-stranded RNA (dsRNA), these bases are often inaccessible to amino acid side chains. As a result, most RBPs require at least part of their binding sites to fall within or nearby regions of single-stranded RNA (ssRNA). ssRNA interactions can occur in many RNA contexts including; the outside of RNA loops, within a hairpin loop, in an internal loop or bulge, or in more complex loop structures such as multiloops. As RBPs can bind target mRNA at several interfaces, only small disruptions in A-form helical structures may be necessary to expand the major groove enough to allow recognition of double-stranded bases next to disrupted regions. As a result, it is likely that only a small number of bases within a sequence binding motif need to be unpaired.³⁵³

Different RBPs show preferences for different structural contexts of ssRNA. For example, Smaug has been shown to preferentially bind specific sequences when found within a hairpin loop in its mRNA target, including *Drosophila nos*.^{74, 76} The RNA recognition motif (RRM) is one of the most common RNA-binding domains (RBDs) and is found as multiple copies within approximately 1,500 proteins, increasing binding affinity to the RNA.³⁵⁴ RRM recognise approximately four nucleotides in single-stranded regions of target mRNAs (for a review see Ref ³⁵⁵).

The identification of common motifs within a given RBPs target transcripts is extremely challenging as each protein has its own unique sequence and structure specificity for RNA binding, and proteins can use a combination of their RNA-binding domains to further increase specificity. The mRNA targets of many RBPs have been identified using a range

of genome-wide techniques and their sequences have been computationally mined for conserved regions that may act as protein-binding motifs.³⁵⁶⁻³⁶⁰ However, the small consensus motifs returned are often found in other transcripts that do not bind the protein of interest, suggesting they are unable to provide sufficient specificity for protein binding. Instead, specificity for the binding of RBPs can be achieved through the generation of RNA structural motifs.³⁶¹ As RBPs often bind unstructured ssRNA, and the primary sequence can only generate a limited number of motifs, secondary structures have been suggested to provide specificity by sequestering consensus protein binding motifs within double-stranded regions of the RNA that are inaccessible to proteins.³⁵³

Some RBDs recognise their RNA targets mainly by their shape and geometry with no requirement for a specific primary sequence. For example, the double-stranded RNA-binding domain (dsRBD) recognises dsRNA stems in a sequence-independent manner, interacting with the ribose sugars and phosphodiester backbone rather than the bases. Binding specificity in this case can be increased by a variety of RNA structures including stem-loops, internal loops, bulges and helices.³⁵⁴

Most RBPs interact with their target mRNAs in the nucleus of the cell where large RNPs are likely to constrain the secondary structure of the mRNA. Displacement of the proteins must occur during export of the mRNA into the cytoplasm and during translation, suggesting that the RNP is continually remodelled. Remodelling of hnRNPs is further supported by the presence of RNA helicases and RNA chaperones within the complex which may enable RBP interactions with regions of the RNA that would otherwise be inaccessible.³⁶² As a result, the secondary structure of the mRNA is also likely to change throughout its lifetime, making it extremely difficult to predict which regions of the mRNA are accessible for protein binding at any given time. Nonetheless, recent studies have shown that combining the secondary structure predictions and accessibility of a target mRNA, and sequence-specific RNA-binding preferences of a given RBP, can improve predictions of RNA sequence motifs.³⁵³

RNA-protein interactions have been investigated *in vitro* using a number of techniques. For example, electrophoretic mobility shift assays (EMSA) use an RBP to identify interacting RNAs which causes a shift in the migration of the RNA in polyacrylamide gel electrophoresis (PAGE).³⁶³⁻³⁵⁴ Alternatively, the RNA can be used to identify all of the proteins that target it using affinity chromatography. For example, a region of *grk* with and without the GLS was used in GRNA chromatography to identify the repertoire of proteins that bind the GLS from ovary extracts.⁵⁴ Systematic Evolution of Ligands by Exponential Enrichment (SELEX) uses a pool of chemically synthesised DNAs with a region of random sequence flanked on each end by a constant sequence and a T7 RNA polymerase promoter. The DNA is amplified and reverse transcribed to generate an RNA pool. RNAs that bind the protein are separated and eluted and amplified before the cycle is repeated several times to select for those with the highest binding affinity.³⁶⁵ Finally, RNAcompete rapidly characterises the binding specificity of RBPs in terms of both sequence and structure by a single pull down from a pool of RNA sequences and structures using a tagged RBP of interest.³⁶⁰ However, *in vitro* techniques are limited due to the non-physiological conditions which can result in interactions that do not occur *in vivo*.

Current *in vivo* techniques again focus on identifying the RNA targets of a given RBP, or the repertoire of RBPs that interact with an RNA of interest. RNP immunoprecipitation-microarray (RIP-Chip) uses antibodies to isolate a specific RBP and enriches for bound RNA fragments genome-wide.³⁵⁷ To improve the fidelity of the interactions, covalent bonds can be formed between the protein and the RNA using ultraviolet light Cross-Linking and ImmunoPrecipitation (CLIP). Partial proteinase K digestion along with reverse transcription errors at the cross-linked sites allows the prediction of the binding sites within the RNA.³⁶⁶ New variations of CLIP, including Photo-Activatable Ribonucleoside-CLIP (PAR-CLIP) and Individual-nucleotide resolution ultraviolet CLIP (iCLIP), further improve the resolution of the RBP binding site.³⁶⁷⁻³⁶⁸ However, the CLIP technique has a variety of limitations including sequence biases and reduced specificity, and is technically

and computationally challenging.

Key outstanding questions in the field

Despite extensive research into the mechanisms behind mRNA localisation and local translation, there are several key questions that still remain. How are so many transcripts originating from the same nuclei specifically recognised by the transport machinery and segregated to distinct regions of the cell? How do the different cohorts of *trans*-acting factors interact with their target mRNAs, and how are they reorganised and regulated throughout the life time of the transcripts? Which features of the mRNA are responsible for generating specificity for protein binding and targeting? Are these mRNA features universal among transcripts localised by the same transport machinery? How highly conserved are the localisation mechanisms across different tissues and species? How is the translational state of the mRNA finely spatially and temporally controlled to ensure the transcripts are activated at the correct place and time?

Aims of this thesis

The GLS is essential for *grk* localisation to the DA of the oocyte and several *trans*-acting factors have been identified which interact with the GLS, presumably providing specificity for the localisation process. To better understand the role of the GLS in *grk* localisation and translational regulation, and how *trans*-acting factors contribute to this process, I carried out structure-function analysis of the signal, focusing on the following questions:

1. Are GLS-like signals conserved outside of the *Drosophilidae* family?
2. Which features of the GLS are required for *grk* transport and anchoring at the DA of the oocyte?
3. Which features of the GLS are required for the association between *grk* and P bodies for translational regulation of the transcript?

In order to address these questions I have generated a number of GLS mutants, based upon the conservation of the GLS, and tested their ability to be transported and anchored in the *Drosophila* oocyte and embryo using an optimised assay based upon current *in vivo* injection assays.^{120, 116} Using the MS2-labelling system I have further characterised the role of Sqd in the anchoring of *grk* mRNA at the DA of the oocyte. Again using *in vivo* RNA injections in the embryo, I identified regions within the GLS responsible for interacting with the P body component Me31B, suggesting their involvement in the translational regulation of *grk*.

Chapter 2

The materials and protocols used throughout the thesis are described.

Chapter 3

Is the GLS conserved outside of the *Drosophilidae* family and can phylogenetic studies be used to predict which regions of the GLS are required for transport and anchoring of *grk*

mRNA? Can the localisation potential of the GLS be measured quantitatively in terms of driving both mRNA transport and anchoring in the oocyte? First, I show that the GLS is highly conserved among the *Drosophilidae* family, and in some other Dipterans, and use this conservation to predict five regions within the signal that may be important for its function in *grk* localisation. Second, I optimise current *in vivo* injection assays to allow efficient screening of subtle localisation defects in GLS mutant RNAs.

Chapter 4

Are the conserved GLS-like signals functional in driving *grk* transport in the oocyte? Which features of the GLS are required for *grk* transport? Here I use cross-species injections to show that most of the GLS-like signals identified both within and outside of the *Drosophilidae* family are able to localise in the *D. melanogaster* oocyte. Using mutagenesis followed by *in vivo* injections into the oocyte and the embryo, I show that several regions of the GLS are required for recruitment of the transport machinery, and thus *grk* transport including; an AU-rich distal stem, the size of the hairpin loop and internal loop, the sequence of a purine stack in stem 3, and the stability provided by the presence of stem 2.

Chapter 5

What is the role of Sqd in *grk* anchoring? Which features of the GLS are required for *grk* anchoring? Here I use MS2-tagged *grk* mRNA to show that *grk* is enriched at the anterior and the DA of *sqd1* mutant oocytes and shows similar characteristics to anchored *grk* in both locations. I propose a new interpretation for the role of Sqd where it is required for inhibiting inappropriate *grk* anchoring along the anterior of the oocyte. Moreover, I show that bulge 2 of the GLS provides the accessibility required for the association between Sqd and the proximal third of the GLS, leading to correct *grk* anchoring in the oocyte.

Chapter 6

Which features of the GLS are required for *grk* association with P bodies for translational regulation of the transcript? Here I use *in vivo* injection assays in the syncytial blastoderm embryo to show that the bottom third of the GLS is required for the association of *grk* with the P body component Me31B. I propose that *grk* RNA ectopically injected into the embryo is translationally repressed inside P bodies at the apical cortex.

Chapter 2: Materials and Methods

Molecular Biology

Solutions and reagents

All solutions and buffers were prepared as described in 'Molecular cloning - A laboratory manual',³⁶⁹ unless otherwise stated. Luri Betani (LB) broth and Super Optimal Broth with Catabolite repression (SOC) media were prepared and autoclaved by the Department of Biochemistry (University of Oxford) media kitchen along with agar plates with antibiotics, apple juice plates and Tris-acetate-EDTA (TAE) (for ingredients see Table 2.1). Enzymes, such as restriction enzymes and T3 polymerase, were obtained from New England Biolabs (NEB) and Roche. Enzymes were used according to manufacturers instructions.

Buffer	Recipe
LB broth	For 1L: 10g Bacto tryptone (Difco), 5g Bacto yeast extract, 5g NaCl
SOC media	For 1L: 20g Bacto Tryptone, 5g Bacto yeast extract, 0.59g NaCl, 0.186g KCl, 2.03g MgCl ₂ ·6H ₂ O, 2.46g MgSO ₄ ·7H ₂ O, 180ml 20% glucose
LB-agar	For 1L: 16g Bacto tryptone (Difco), 10g Bacto yeast extract, 5g NaCl
Apple juice agar	For 1L: 250ml apple juice, 25g sucrose, 22g agar, 750ml H ₂ O
Agarose (1%)	For 100ml: 1g agarose (Roche), 100ml 1x TAE
10x TAE	0.4M Tris base, 10mM EDTA (pH 8.0), 1.2% (v/v) glacial acetic acid
1 x PBS	For 1L: 10 phosphate buffered saline (Dulbecco A) tablets, 1L MilliQ H ₂ O
Lysis buffer	50mM Tris-HCl pH 8.0, 150mM NaCl, 0.5% NP-40, 10% glycerol, Complete EDTA-free protease inhibitor (Roche)

Table 2.1. Recipe for the buffers used in this thesis.

Propagation and preparation of plasmid DNA

For standard DNA amplification and cloning protocols Stellar Competent Cells (Clontech) were transformed using heat shock techniques. A 50µl aliquot of cells was incubated on ice with 1-50ng of plasmid DNA or 25% of a ligation reaction for 30 minutes. Heat shock of the cells was at 42°C for 45 seconds, followed immediately by ice for 2 minutes. 450µl of warm SOC media was added to the cells before incubation at 37°C for 1 hour to recover. Cells were plated on LB-agar plates containing the appropriate antibiotic (100µg/ml carbenicillin) and cultured over night at 37°C. A single transformed colony was selected and transferred to 3-5ml LB containing the appropriate antibiotic and cultured overnight at 37°C. Isolation of the plasmid was carried out using the QIAprep Spin Miniprep Kit (Qiagen).

Cloning

GLS mutant fragments and GLS-like signals from other species were designed within a surrounding 114-nucleotide region of *grk* cDNA and sent for synthesis as gBlock gene fragments (Integrated DNA Technologies, Inc) or GeneArt Strings DNA Fragments (Thermo Fisher Scientific). To produce full-length *grk*, the mutant fragments were cloned into BsaBI and SphI sites within the full-length pBluescriptKS(+) Grk66 cDNA construct (a gift from T. Schupbach, Ref¹⁵) replacing the wildtype GLS, using In-Fusion HD-Cloning Technology (Clontech) as per the manufacturers protocol. Hairpin loop mutants tested previously as GLS fragments by Ball and Van De Bor were in the pGEM-T Easy vector (Promega).

DNA sequencing and analysis

DNA sequencing was performed by the Source Biosciences Sanger Sequencing Facility. 350ng of plasmid DNA was provided within a total volume of 5µl (70ng/µl). Forward and reverse sequences were analysed for each sample. 16pmols of primers (Eurofins genomics) were provided in a total volume of 5 µl (3.2 pmol/µl) and their sequences are shown in Table 2.2.

Primer	Sequence
Grk66_F2	CCGCGCTTGCTGCTCTTTAG
Grk66_R2	CCAGGTGTCGGTACTGGATC

Table 2.2. Oligonucleotide primers used to sequence GLS mutant clones in the full-length *grk* transcript.

Gel electrophoresis

TAE-agarose gel electrophoresis

Samples were mixed with 1x gel loading dye (New England Biolabs) and loaded onto 1% agarose gels (Table 2.1) containing 0.5µg/ml ethidium bromide. Electrophoresis was carried out in 1x TAE running buffer at a constant voltage of 80V, and nucleic acids were visualised using an enclosed UV transilluminator with associated imaging software.

SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Samples were heated to 90°C in 2x protein LDS sample buffer (with reducing agent added prior to use) and loaded on NuPage Novex 4-12% Bis-Tris gels alongside a pre-stained standard (SeeBlus Plus 2). Electrophoresis was carried out in 1x NuPAGE MOPS buffer at a constant voltage of 180V. All of these reagents were from Thermo Fisher Scientific. Proteins were visualised using Western Blotting.

Western Blotting

Following electrophoresis, proteins were transferred to nitrocellulose membranes (Bio-Rad) using the XCell SureLock Blot Module (Invitrogen) or the Trans-Blot Turbo Transfer System (Bio-Rad) following the manufacturers instructions. The membrane was blocked with 10% non-fat dry milk in 1x PBS with 0.1% Tween20 (Sigma-Aldrich) or with Odyssey Blocking

Buffer (PBS) (LI-COR) for 1 hour. The membrane was incubated with primary antibodies overnight at 4°C in block. The membrane was washed 4x 15 minutes in PBS with 0.1% Tween20 before incubating at room temperature (RT) for 45 minutes with secondary antibodies and repeating the washes. Visualisation of proteins was performed by quantitative infrared imaging (LI-COR Odyssey).

Antibody	Raised in	Dilution	Supplier
Anti-Sqd S 1B11	Mouse	1:2000	DSHB (T. Schupbach)
Anti-actin A2066	Rabbit	1:500	Sigma-Aldrich
Anti-mouse IgG 680CW	Donkey	1:5000	LI-COR
Anti-rabbit IgG	Goat	1:5000	LI-COR

Table 2.3. Primary and secondary antibodies and their dilutions used for Western Blotting. DSHB, Developmental Studies Hybridoma Bank.

Preparation of templates for *in vitro* transcription

5µg plasmid DNA (Grk66 and *hb* gifts from T. Schupbach and Bruce Edgar, respectively) was digested with the appropriate restriction enzyme. Digests were visualised by TAE-agarose electrophoresis and purified using the QIAquick PCR purification kit (Qiagen). Grk66 containing a variety of mutations was transcribed, fluorescently labelled and capped *in vitro* along with *hb* using the protocol from Ref ³³⁷. However, the RNA was not extracted with phenol/CHCl₃ but was instead precipitated with NH₄OAc/EtOH immediately following the removal of unincorporated nucleotides. In addition, transcription was carried out in the presence of 50 U RNasin RNase inhibitor (Promega). Quality checks were carried out at different steps throughout the transcription reaction using agarose gel electrophoresis (see Appendix E), and the concentration and quality of the RNA was measured on a nanodrop.

Drosophila protocols

Flystocks

Flies were raised on standard cornmeal-agar medium at 18°C or 25°C. *Oregon-R* (*OrR*) was used as the wild-type strain. A full list of stocks can be found in Appendix B.

Preparation of ovaries for oocyte injection and live imaging

Newly emerged females were fattened on dried active bakers yeast for 2 days before dissecting their ovaries into halocarbon oil (Series 95, KMZ chemical Ltd). Individual ovarioles were teased from the ovaries on coverslips (Thickness No. 1.5, Thermo Fisher Scientific) in halocarbon oil using a dissecting microscope.³⁷⁰ Egg chambers were imaged for no longer than 1 hour after dissection, unless otherwise stated.

Microinjection of oocytes

Microinjection of oocytes was carried out using Eppendorf femtotip needles. Dextran aqueous stock solution (Invitrogen) was made at 1mM. Unless otherwise stated, fluorescently labelled full-length *grk* or *hb* RNA was injected at a concentration of 140ng/μl in DEPC-treated water. Typically, 10 oocytes were injected per sitting so '0 mins' time points are likely to vary from 0-5 minutes post-injection. For imaging conditions see Appendix C. Hairpin loop mutants were tested previously in the lab by Ball and Van De Bor as GLS stem-loops rather than the full-length *grk* gene.

Preparation of ovarian lysate for Western Blotting

For preparation of fresh ovarian lysate for Western Blotting, 30 ovaries were dissected in cold 1x PBS and flash frozen in liquid nitrogen. 30μl cold lysis buffer was added to the ovaries. Once thawed, ovaries were homogenised on ice and centrifuged at 1000g for 2 minutes before transferring the supernatant to a new tube.

Preparation of embryos for microinjection

Blastoderm embryos were collected from fly cages containing apple juice agar plates every 2 hours for microinjection. The embryos were washed with water in a sieve before being treated with 50% bleach for 5 minutes at RT to remove the chorion. The embryos were thoroughly washed with water and transferred to an apple juice agar plate for alignment with a probe. Heptane glue was applied in a line down the centre of a coverslip (Thickness No. 1.5, Thermo Fisher Scientific) and was used to lift the aligned embryos from the agar plate onto the coverslip. Embryos were dehydrated over silica gel for approximately 10 minutes at RT before being covered with series 700 Halocarbon oil (KMZ Chemical Ltd) to prolong culturing for injection.

Microinjection of embryos

RNA was injected into the basal cytoplasm of the stage 14 blastoderm embryo at 380ng/ μ l in DEPC-treated water using Eppendorf femtotip needles. Typically, 5 embryos were injected per sitting so '0 mins' time points are likely to vary from 0-3 minutes post-injection, during which time the RNA may begin localising. For imaging conditions see Appendix D.

Microscopy

Imaging parameters and analysis are detailed in Appendices C-D.

Widefield live cell imaging

Imaging was performed on a DeltaVision CORE widefield fluorescence deconvolution system (Applied Precision, A GE Healthcare Company) based upon an Olympus IX71 inverted microscope. Images were acquired using a 16 bit Roper Cascade II ECCD camera with standard Chroma filter-sets with either a 20x/0.75NA dry objective or a 100x/1.3NA silicone oil immersion objective with a correction collar to minimise spherical aberration.

Bioinformatics

Conservational studies searching for GLS-like signals were carried out using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Phylogenetic analysis of the GLS-like hits from BLAST was carried out by aligning the sequences using MUSCLE and generating a phylogenetic tree using T-REX.^{371, 372} Sequence logos were generated using WebLogo.³⁷³ RNA secondary structures were predicted using UNAFold.³⁵¹

Chapter 3: A rapid assay for a structure-function analysis of the GLS

Introduction

Although much research has been directed toward understanding how specific RNAs become localised within the *Drosophila* egg chamber, a key outstanding question is how different RNA species originating from the same NCs become sorted and transported to distinct regions of the oocyte. Although these RNAs utilise the same transport machinery and microtubule networks for their transport into the oocyte, they segregate to distinct sub-cellular locations within the oocyte.^{3, 116} While different steps of localisation must be specified by signals in the RNA, it is unclear whether the common and distinct destinations of different mRNAs are dictated by different signals, or can be specified by distinct features within the same signals.

Many *cis*-acting RNA signals that drive the differential localisation patterns of RNAs have been identified and in some cases have been well studied. For example, the *grk* localisation signal (GLS)^{305, 307, 54} and the localisation signals of *ASH1*, *bcd* and *osk* mRNA.^{63, 374-376} The GLS is a 64-nucleotide stem-loop that has been shown to be both necessary and sufficient to drive all stages of *grk* localisation and anchoring at its final destination at the DA of the oocyte.³⁰⁷ It is thought that differences in the complement of *trans*-acting factors that interact with *cis*-acting RNA signals are responsible for establishing the distinct destinations of individual RNAs.³⁷⁷ Although many of the *trans*-acting factors that bind the GLS have been identified,⁵⁴ little is known about how these proteins recognise the RNA and interact with each other on a molecular level to drive *grk* localisation.

In this chapter, I have used bioinformatics, phylogenetic approaches and secondary structure analysis to identify conserved GLS-like signals in order to predict which regions of the GLS are important for recognition by the transport and anchoring machinery. Using this

information has allowed me to formulate hypotheses regarding the features in the GLS required for *grk* localisation, and to design sequence modifications in the signal to test these hypotheses. To test the affect of GLS sequence modifications on *grk* localisation, I have adapted a well established *in vivo* RNA injection assay to allow high-throughput, quantitative screening of GLS mutants for subtle localisation defects.

Newly available genomes should allow further studies of GLS conservation

The GLS is located within the coding region of *grk* and overlaps with a portion of the gene encoding the signal peptide. As there are many evolutionary constraints on the GLS, both at the RNA and the protein level, it is expected that the GLS primary sequence would be highly conserved. Previous studies have identified GLS-like signals within ten species of *Drosophilids* whose genomes were available at the time.^{305, 307} Cross-species injections of fluorescently labelled RNA have shown that the *D. melanogaster* GLS is competent to localise in the *D. virilis* oocyte and vice versa.³⁰⁷ This emphasises the degree of conservation between the localisation signals and the localisation machinery in these two species. Since these studies, several more *Drosophilid* genomes have been sequenced and are available to search for GLS-like signals.

Due to the high levels of conservation of the GLS among *Drosophilids*, it is not feasible to distinguish which parts are conserved because they are important for the RNA to function as a localisation signal, and which are conserved because they are required for protein function. Bioinformatic searches for GLS-like signals in more divergent *Drosophilids* and other insects would provide more information about which regions of the GLS are essential for localisation, and would aid the design of RNA sequence modifications to test their function.

Current injection assays are inadequate for high throughput screening of subtle RNA localisation defects

One way of investigating how localisation signals and their associated proteins act together to drive RNA localisation is through large-scale mutagenesis studies of the RNA. In order to study the effects of RNA mutagenesis on the localisation of a transcript, both spatial and temporal information is required. Many RNA localisation assays involve tagging endogenous RNA transcripts with fluorescent labels. For example, the MS2-MCP system utilises multiple copies of the bacteriophage MS2 stem-loops to tag the 3' or 5'UTR of the RNA.^{73, 340} These stem-loops are recognised by a fluorescently tagged MS2 Coat Protein (MCP) which decorates the RNA allowing real-time visualisation of endogenous transcripts. MS2-tagged *grk* RNA undergoes normal localisation in the oocyte, and is translated correctly, resulting in rescue of the *grk* mutant phenotype.²⁸³ Although this approach is highly biologically relevant and often has few deleterious effects on the specimen, it requires the generation of transgenic animals, which is a very lengthy and costly process making it unfeasible for a high-throughput screen of mutant RNAs.

More high-throughput assays to investigate RNA localisation patterns involve microinjection of *in vitro* transcribed, fluorescently labelled RNA into living cells. These assays have been routinely used to discover the mechanisms of RNA localisation in *Drosophila* embryos^{116, 141, 337, 339, 379} and egg chambers^{69, 120, 133, 270, 338, 339}, and provide a fast method of assessing localisation patterns *in vivo*. However, in order to efficiently screen a large number of GLS mutants for subtle localisation defects, several aspects of the current injection assays would need to be adapted.

First, injecting high concentrations of RNA has the potential of over-loading the transport and anchoring machinery which could create unwanted variations in RNA localisation, masking any subtle GLS mutant phenotypes. Second, current assays assess final localisation patterns at one hour post-injection.^{120, 307} However, injected *grk* has been shown to localise in the oocyte within 40 minutes,³⁰⁷ suggesting that the length of the assay could potentially be

reduced, allowing more high-throughput screening of GLS mutants, and reducing the culturing time of the live specimens. Third, most assays characterising the affects of mutations on RNA localisation involve testing the minimal localisation signal of the RNA, rather than the full-length transcript. Finally, most localisation assays rely on assessing the final pattern of RNA localisation in a qualitative manner that is open to subjectivity. For high-throughput screening of a large number of mutant transcripts, it would be advantageous to develop a faster and more quantitative method for assessing the RNA localisation patterns.

Adaptations to the current injection assays would aspire to design targeted sequence modifications within the localisation signal based upon current knowledge about the sequence, structure, and conservation of the RNA, or any previously defined binding motifs for specific *trans*-acting factors known to interact with the RNA. The resulting mutants would be cloned into the full-length gene, *in vitro* transcribed with fluorescently labelled tags to allow visualisation of the RNA, and injected into the oocyte at low concentrations. The minimum time required for full localisation of the wild type RNA would be determined to limit stress to the specimen and increase throughput, and the localisation pattern would be assessed at multiple time points in a fast and effective manner suitable for the quantitative screening of subtle localisation defects. An example work flow is outlined in Figure 3.1.

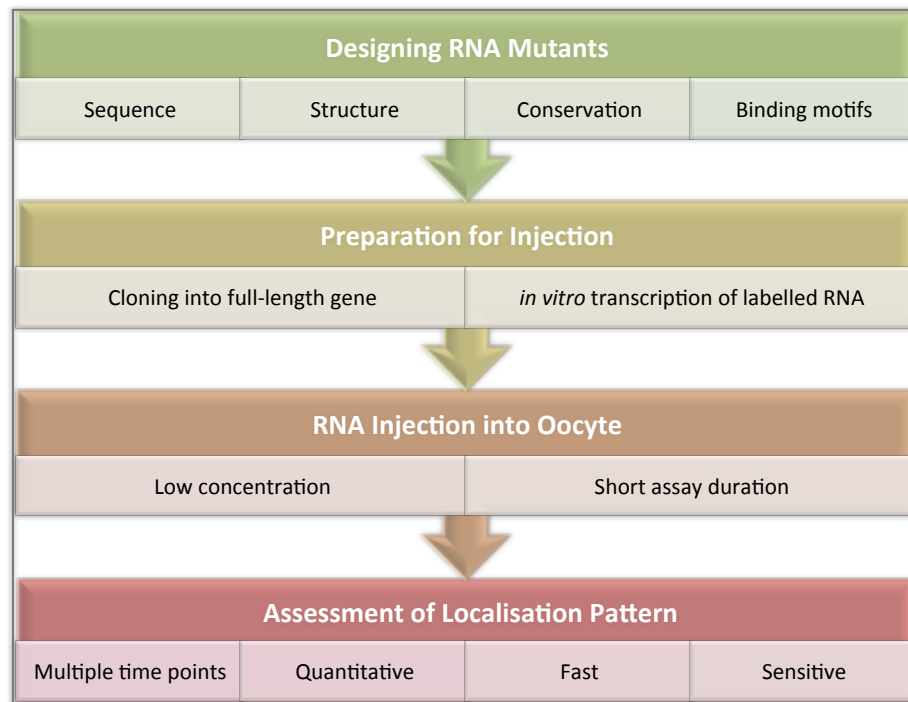


Figure 3.1. Aspirations for a new high-throughput injection assay to screen for subtle localisation defects in mutant RNAs. A high-throughput assay for testing regions of an RNA signal required for localisation would require the design of targeted mutations based upon information about the sequence, structure and conservation of the RNA, as well as any binding motifs mapped for known interacting proteins. Mutant sequences would be inserted into the gene of interest and the full transcript *in vitro* transcribed with fluorescent labels for visualisation of the RNA. Mutant RNAs would then be injected into the oocyte and allowed to localise for the shortest period required for full localisation of the wild type transcript. The localisation pattern of the RNA would be assessed at multiple time points in a fast and quantitative manner that is sensitive enough to detect subtle defects in localisation.

Aims of this chapter and experimental rationale

In this chapter I will use a bioinformatics approach to search for conserved GLS-like signals within and outside of the *Drosophilidae* family. I will adapt the current well established *in vivo* RNA injection assays to allow for high-throughput, fast, and systematic screening of a range of modified GLS sequences. Finally, I will develop a quick and easy method for quantifying RNA localisation patterns. The specific questions addressed are:

1. How highly conserved is the GLS within the *Drosophilidae* family?
2. Is the GLS conserved outside of the *Drosophilidae* family?
3. Can current live RNA injection assays be optimised to:
 - a. reduce the total time of the assay to increase the throughput and limit stress to the specimen?
 - b. quantitatively identify subtle defects in RNA localisation?

In order to target the design of GLS sequence modifications to regions of the signal likely to be important for function, I will search for GLS-like signals in new *Drosophilid* genomes that have been made available since previous publications. In addition, I will extend these phylogenetic studies to the genomes of species outside of the *Drosophilidae* family. This will allow me to understand the extent to which the GLS is conserved throughout the Insecta class. I will use this information in the next chapter to predict which regions of the GLS are essential for each stage of *grk* localisation, and distinguish which are likely to be required for recognition by the transport and anchoring machinery.

In order to optimise current RNA localisation assays I will reduce the concentration of RNA injected into the oocyte and shorten the total length of the assay to increase the throughput, allowing me to systematically screen a large number of GLS modifications. Developing a fast and effective method for quantifying the distribution of injected RNA at different time points throughout the assay will allow the detection of subtle spatial and temporal localisation defects.

Results

The GLS primary sequence is highly conserved among *Drosophilids*

Previous studies have highlighted the conservation of the GLS within ten *Drosophilid* species.^{305, 307} Since then, several new *Drosophilid* genomes have been sequenced. To discover whether these species contain GLS-like signals within their genomes I performed an NCBI BLAST search (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), restricting the scope to the *Drosophilidae* family. The search returned 19 species of *Drosophilids* containing a GLS-like signal with sequence identities varying from 89% for *Drosophila miranda* to 100% for *Drosophila simulans*, *Drosophila sechelia*, and *Drosophila bipectinata*. A summary of the hits and their corresponding percentage sequence identities compared to the *D. melanogaster* sequence can be found in Table 3.1. Although *D. miranda* showed the lowest percentage sequence identity, the GLS-like sequence was only 58-nucleotides long as opposed to the full 64-nucleotides of the GLS. Comparing the 58-nucleotide region of the two signals instead showed a sequence identity of 98%, suggesting it is not as divergent as originally thought.

In order to understand how the GLS sequence has diverged throughout the *Drosophilidae* family, I carried out a phylogenetic analysis on the GLS-like hits from the BLAST search using the program MUSCLE to align the sequences,³⁸⁰ and T-REX to generate the tree (Figure 3.2A).³⁷² The tree showed that the GLS-like signals fall into four major groups. The *D. melanogaster* GLS lies within a group containing *D. bipectinata*, *D. simulans*, and *D. sechellia* indicating that these signals share a common evolution, as predicted from their 100% sequence identities. The rest of the GLS-like signals form additional groups and subgroups based upon their sequence similarities, indicating the common evolutionary lineages of each of the signals.

The overall conservation of the GLS within the *Drosophilidae* family could be used to infer regions of the GLS that may be important for function. Generating a sequence logo from the

Species name	Identities (nucleotides)	% Identity
<i>D. melanogaster</i>	64/64	100%
<i>D. simulans</i>	64/64	100%
<i>D. sechelia</i>	64/64	100%
<i>D. bipectinata</i>	64/64	100%
<i>D. persimilis</i>	63/64	98%
<i>D. erecta</i>	63/64	98%
<i>D. yakuba</i>	63/64	98%
<i>D. ananassae</i>	63/64	98%
<i>D. rhopaloa</i>	63/64	98%
<i>D. elegans</i>	63/64	98%
<i>D. pseudoobscura</i>	62/64	97%
<i>D. ficusphila</i>	62/64	97%
<i>D. mojavensis</i>	61/64	95%
<i>D. grimshawi</i>	61/64	95%
<i>D. kikkawai</i>	61/64	95%
<i>D. takahashii</i>	61/64	95%
<i>D. biarmipes</i>	61/64	95%
<i>D. virilis</i>	60/64	94%
<i>D. willistoni</i>	60/64	94%
<i>D. miranda</i>	57/64	89%

Table 3.1. The GLS is highly conserved among the *Drosophilidae* family. BLAST searches throughout the *Drosophilidae* family identified 19 species containing a sequence similar to the *D. melanogaster* GLS. Three of these species showed 100% sequence identity with the GLS. *D. miranda* showed the lowest percentage sequence identity with the *D. melanogaster* GLS.

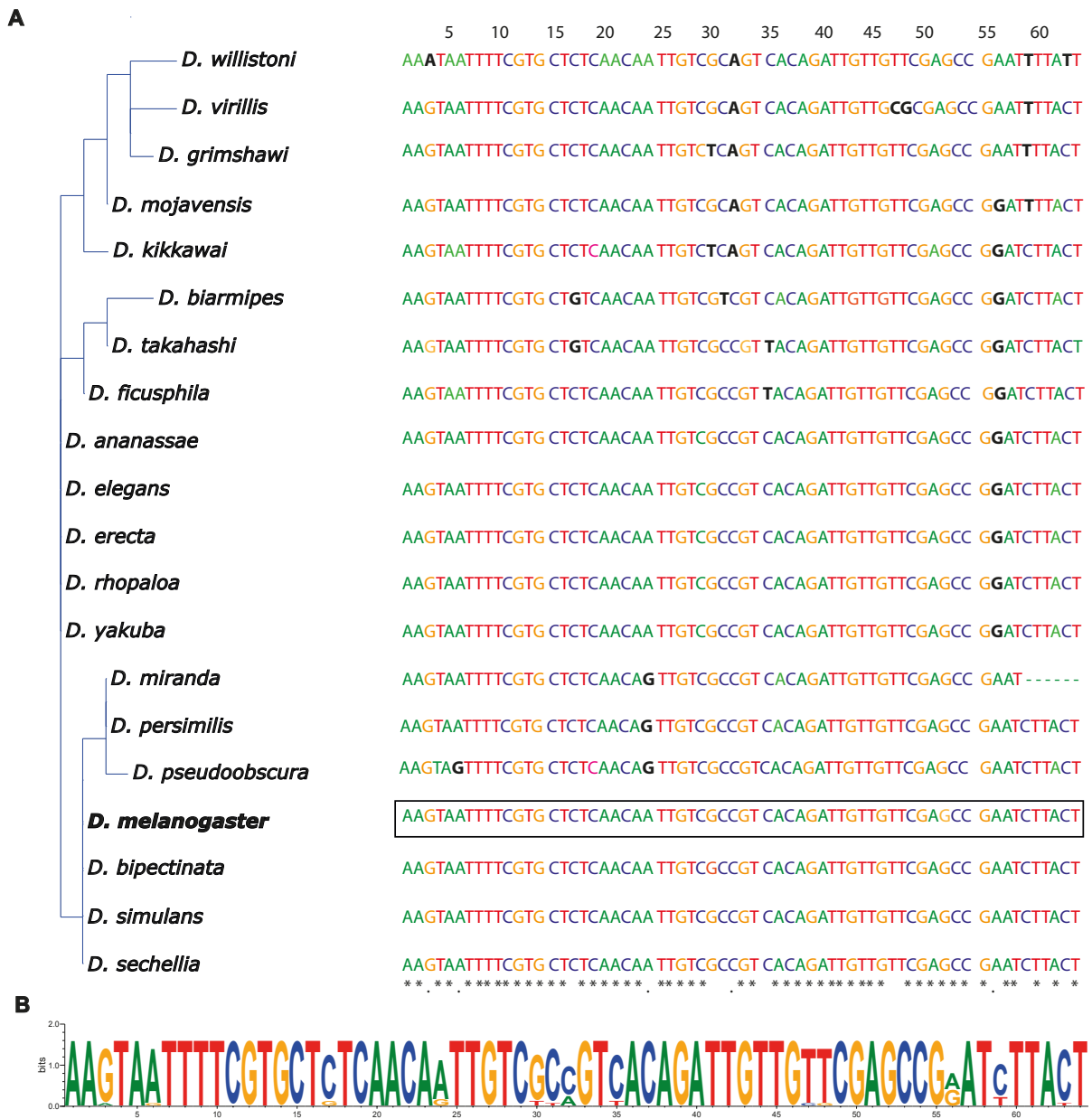


Figure 3.2. The sequence of the GLS is highly conserved among the *Drosophilidae* family. (A) A phylogenetic tree (generated using T-REX; Boc et al. 2012) and corresponding sequence alignments of the GLS-like signals found within different *Drosophilid* species show that the *D. melanogaster* sequence (highlighted in black box) shares the highest similarity with a group composed of *D. bipectinata*, *D. simulans* and *D. sechellia*. ('*' = conserved, '.' = semi-conservative mutations, " " = non-conservative mutations, non-conserved nucleotides highlighted in bold, black letters). (B) A sequence logo of the *Drosophilid* species GLS-like signals, where the size of the letter represents the degree of conservation in each position, shows large regions of high conservation but identifies some regions where the nucleotides are not highly conserved (WebLogo; Crooks et al., 2004).

GLS-like signals using WebLogo (<http://weblogo.berkeley.edu/logo.cgi>) shows that most of the signal is highly conserved, represented by large letters (Figure 3.2B).³⁷³ Regions that stand out as being highly conserved include nucleotides TTTTCGTGCT at positions 7-16, TCAACA at positions 18-23, TTGTC at positions 25-29, ACAGATTGTTG at positions 36-46, CGAGCCG at positions 49-55, and TTA at positions 60-62. However, small regions of the signal do show low levels of conservation, represented by small letters, which could have a substantial effect on the secondary structure of the RNA. This raises the question of whether the secondary structure of the GLS is highly conserved.

The predicted GLS secondary structure is highly conserved among *Drosophilids*

The primary sequence and secondary and tertiary RNA structures of localisation signals are often essential for recognition by *trans*-acting protein factors.³⁸¹ The GLS is predicted by the UNAFold RNA folding program to form a highly-structured stem-loop.³⁵¹ Partial NMR solution structures generated by Graeme Ball and Suzanne McDermott in our lab (unpublished) suggest that the UNAFold predicted secondary structure of the stem regions of the GLS are accurate. To determine the extent to which small regions of non-conserved nucleotides within the GLS affect the secondary structure of the RNA, I generated secondary structure predictions for each of the GLS-like signals identified. The *D. melanogaster* GLS, and the other species containing an identical GLS sequence (*D. simulans*, *D. sechellia* and *D. bipectinata*), form a large stem-loop with a hairpin loop, two internal bulges, and two internal loops (Figure 3.3A). Only 13 nucleotides of the 64-nucleotide GLS (20%) are non-conserved (highlighted in orange circles) in the other *Drosophilid* species. This is likely to reflect the high levels of constraint on this sequence due to its multiple functions at both the RNA and protein-coding level.

Secondary structure predictions of the other GLS-like signals show that the small regions of non-conserved nucleotides have very little effect on the overall secondary structure of the RNA (Figure 3.3B). For example, *D. miranda*, *D. persimilis* and *D. pseudoobscura* all contain

the nucleotide change 24A>G which has no effect on the predicted secondary structure, and results in their inclusion into the same branch of the phylogenetic tree (Figure 3.2A). In addition, *D. pseudoobscura* contains a second nucleotide change 6A>G. Another common nucleotide variance 56A>G is found in half of the species identified and is predicted to have no effect on the predicted secondary structure of the signal (Figure 3.3B). The final single nucleotide change common to many of the species is 59C>U found within the proximal internal loop. Again, this transition has no impact on the predicted secondary structure and is found in *D. willistoni*, *D. virillis*, *D. grimshawi* and *D. mojavensis* (Figure 3.3B), grouping them into the same subgroup in the phylogenetic tree (Figure 3.2A). Nucleotide changes that are less common throughout the species but that have no effect on the predicted secondary structure of their signals include; 30G>U (*D. grimshawi* and *D. kikkawai*), 32C>A (*D. virilis*, *D. mojavensis*, *D. grimshawi*, *D. willistoni*, and *D. kikkawai*), 35C>U (*D. ficusphila*), 47-48 UU>CG (*D. virilis*), and a base-pair change at 3/63 GC>AU (*D. willistoni*) (Figure 3.3B).

Despite these examples of nucleotide changes that have little effect on the predicted secondary structure of the RNA, there are two exceptions. The GLS-like signals from *D. takahashii* and *D. biarmipes* both share a 95% sequence identity with the *D. melanogaster* signal (Table 3.1). Despite this high level of conservation, their predicted secondary structures are significantly different (Figure 3.3B). In both cases, single nucleotide changes in the hairpin loop are predicted to drastically alter the structure of this region. In the case of *D. takahashii*, the nucleotide change 35C>U causes the size of the hairpin loop to increase, whereas the 31C>U transition in *D. biarmipes* results in the constriction of the hairpin loop (Figure 3.3B). Both signals also share the common non-conserved nucleotide substitution 17C>G which results in the constriction of the distal internal loop, and the formation of a new internal bulge at position 50.

Taken together, these results emphasise the fact that although many small changes in primary sequence are predicted to have little effect on the secondary structure of the RNA, some are predicted to induce large structural changes. These secondary structure changes

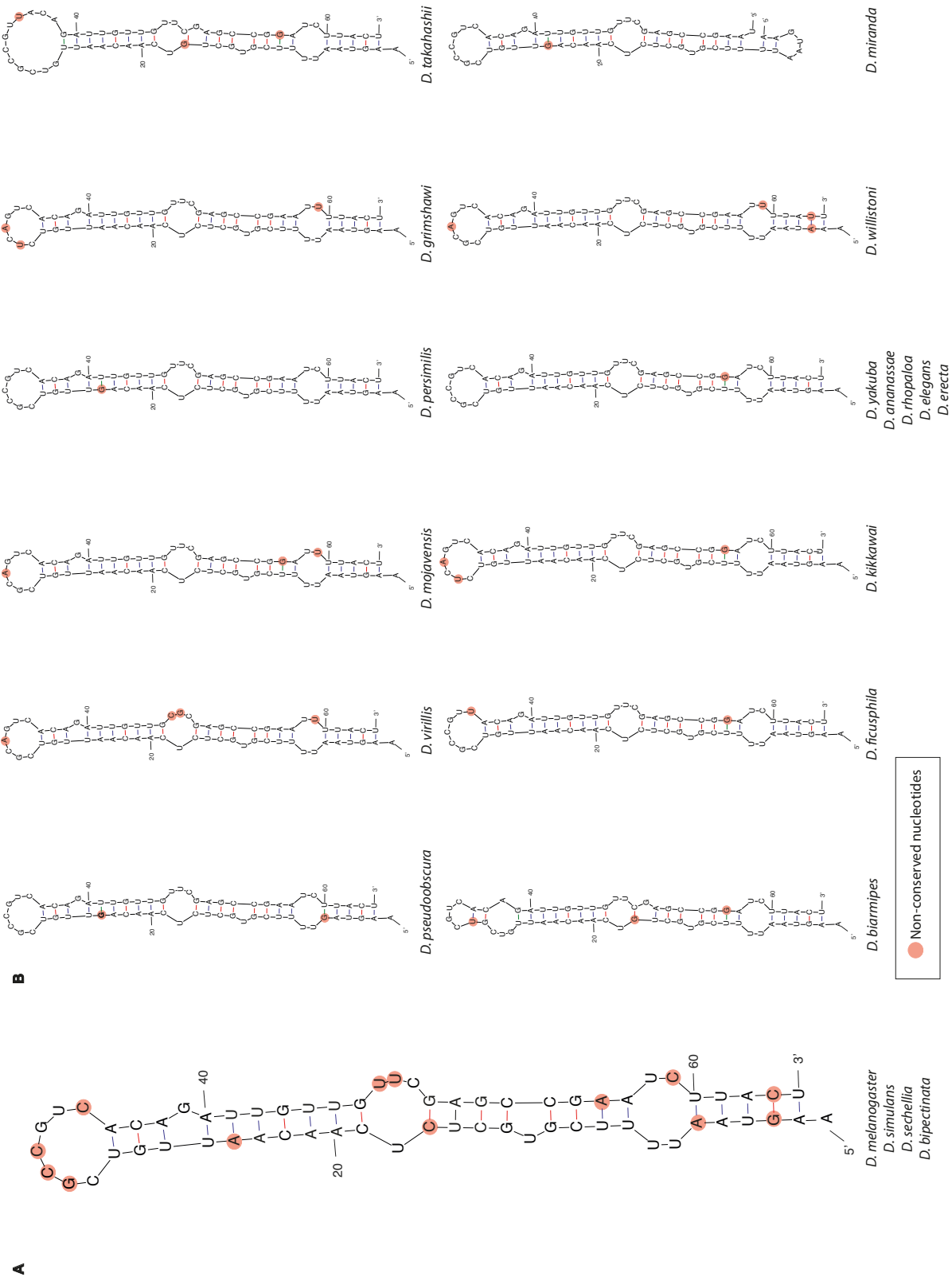


Figure 3.3. The secondary structure of the GLS is highly conserved among the *Drosophilidae* family. (A) UNAFold predicted secondary structure of the GLS from *D. melanogaster*. Nucleotides found to be non-conserved in the other *Drosophilid* species identified are highlighted in orange circles. (B) UNAFold structures of GLS-like signals identified in other *Drosophilid* species suggest the secondary structure is highly conserved. Nucleotides that differ from the *D. melanogaster* GLS are highlighted in orange circles and are mapped onto the *D. melanogaster* signal to summarise the data presented.

have the potential to greatly affect the function of the signal by disrupting critical molecular interactions with *trans*-acting factors.

The GLS primary sequence is conserved in some insects outside of the *Drosophilidae* family

Despite the high degree of conservation found among the *Drosophilids*, conserved GLS-like signals have not been previously found outside of this family. Previous studies searching for Grk-like EGF signalling components outside the Diptera order were unable to find any clear orthologs for *grk* as it is a rapidly evolving gene.³⁸² However, they showed that *Nasonia vitripennis* (wasp) contains a *tgfa*-like gene which localises around the oocyte nucleus in a fashion similar to *D. melanogaster grk* during mid-oogenesis.³⁸² This raises the question of whether GLS-like signals are conserved outside of the *Drosophilidae* family, and whether they are a more universal mechanism for targeting RNAs for axis determination in the oocyte than previously predicted.

In order to identify GLS-like signals outside of the *Drosophilidae* family, I first used the *D. melanogaster* GLS primary sequence to search the entire BLAST database, excluding the *Drosophilidae* family, in order to increase the chance of identifying hits that are less highly conserved. This search returned no positive hits for GLS-like signals. Second, I queried the entire database using the full Grk amino acid sequence. This search returned many hits in which the sequence was highly conserved throughout the protein. However, only some of the hits showed conservation at the region corresponding to the GLS in *D. melanogaster*. Therefore, the results were mined for any containing an alignment within the region

surrounding the signal peptide in the query sequence (amino acids 1-60), thereby limiting the search to the region partially encoded by the GLS in *D. melanogaster*. The top hit returned following the *Drosophilid* species was a predicted Grk-like protein in *Stomoxys calcitrans* (Stable fly). Aligning the coding sequence of this gene with the *D. melanogaster* GLS sequence revealed a GLS-like region with 89% sequence identity. This process was repeated for all of the hits returned and the associated percentage sequence identities compared to the *D. melanogaster* GLS were calculated (Table 3.2).

Species name	Identities (nucleotides)	% Identity
<i>Stomoxys calcitrans</i>	57/64	89%
<i>Musca domestica</i>	54/64	84%
<i>Lucilia cuprina</i>	57/64	89%
<i>Bactrocera cucurbitae</i>	55/64	86%
<i>Bactrocera oleae</i>	55/64	86%
<i>Ceratitis capitata</i>	54/64	84%
<i>Aedes aegypti</i>	42/64	66%

Table 3.2. GLS-like signals can be identified in species outside of the *Drosophilidae* family. BLAST searches using the whole amino acid sequence, the signal peptide region, and the EGF domain from *D. melanogaster* Grk identified seven species containing a sequence similar to Grk. Alignments of the underlying coding sequence of these hits revealed these species contain GLS-like signals with high levels of sequence identity to the *D. melanogaster* GLS.

To determine whether the conservation of the GLS extends further than these original results, I used the Grk signal peptide amino acid sequence as a query for a BLAST search. Any proteins that contain a similar signal peptide to Grk may also contain a GLS-like signal as they have overlapping coding sequences in the *D. melanogaster* gene. Several of the hits overlapped with the previous search using the entire Grk protein, including *M. domestica* and *C. capitata*. However, three new species were returned: *Aedes aegypti* (Yellow fever mosquito), *Danio rerio* (Zebrafish), and *Anopheles darlingi* (American malaria mosquito) with sequence identities at the signal peptide region of 58%, 62% and 71%, respectively. The

underlying coding sequence within this region of the genes were aligned with the *D. melanogaster* GLS. This revealed a region within *A. aegypti* with 66% sequence identity to the GLS, making it the most divergent sequence identified so far (Table 3.2). *D. rerio* and *A. darlingi* were also found to contain GLS-like sequences but with very low sequence identities. UNAFold secondary structure predictions of both of these species revealed stem-loops with very little resemblance to the GLS and so they were removed from consideration. Other BLAST searches completed included using the Grk EGF domain or the *C. capitata* signal peptide to search for even more distant GLS-like signals. However, these searches returned hits with low levels of sequence identity in the region where the GLS is likely to be found, and even less similarity when looking at the predicted secondary structures of these hits.

To compare the new GLS-like signals to the *D. melanogaster* GLS I aligned the sequences and identified the conserved and non-conserved nucleotides. These alignments indicate several regions of the GLS primary sequence are highly conserved in these GLS-like signals (Figure 3.4A). Alignment of the *A. aegypti* GLS-like region reveals that only the core of the signal is conserved, and that a region of 8-9 nucleotides is absent from each end of the returned BLAST sequences. *B. cucurbitae* also contains a sequence that is truncated compared to the GLS but only by a single nucleotide at the beginning of the signal. The rest of the species identified contain full-length GLS-like signals with regions of high conservation throughout the sequence, as reflected in the sequence logo (Figure 3.4A). These regions of high conservation include nucleotides 4-11 TAATTTTC, 13-16 TGCT, 18-20 TCA, 45-49 TGTTTC, 51-53 AGC, and 60-62 TTA (Figure 3.4A).

Phylogenetic analysis of the 20 *Drosophilid* species and the seven newly identified GLS-like signals from species outside the *Drosophilidae* family was carried out using the T-REX program to understand how the GLS has diverged in other Dipterans.³⁷² The results show that the GLS-like signal from *S. calcitrans* is closest in sequence to the *Drosophilid* family of signals, and is found within a subgroup containing the *L. cuprina* and *M. domestica* signals (Figure 3.4B). A second subgroup is formed by signals originating from the species *B. oleae*,

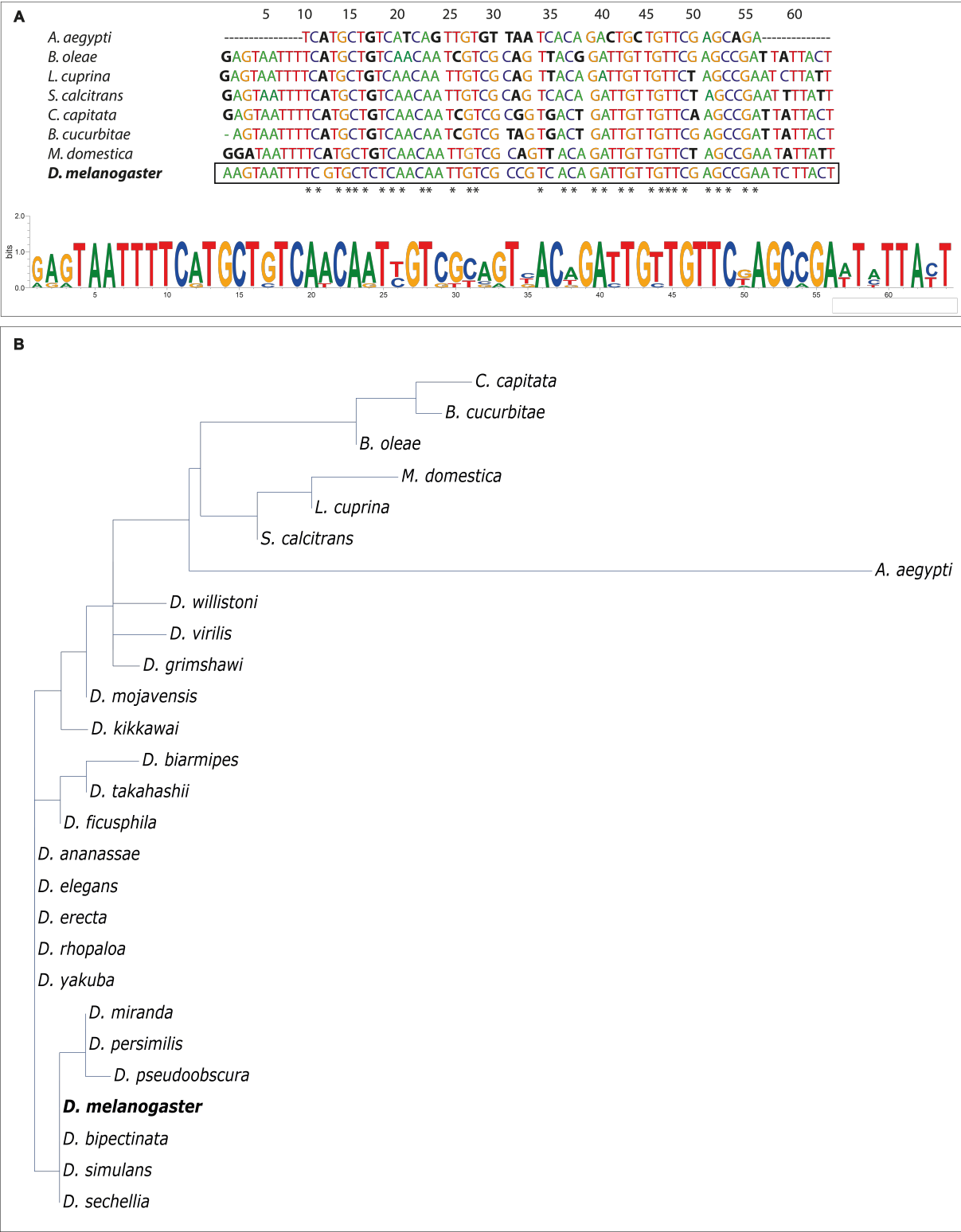


Figure 3.4. The GLS shows some primary sequence conservation in other species outside of the *Drosophilidae* family. (A) Sequence alignments of the GLS-like signals found within Dipterans outside of the *Drosophilidae* family against the *D. melanogaster* GLS sequence (highlighted in bold black box). ('*' = conserved, '.' = semi-conservative mutations, " " = non-conservative mutations, non-conserved nucleotides highlighted in bold, black letters). The corresponding sequence logo (generated using WebLogo; Crooks *et al.*, 2004) highlights regions of the sequences that are highly conserved (large letters) and those that are not well conserved (small letters). (B) A phylogenetic tree (generated using T-REX; Boc *et al.*, 2012) showing the likely pattern of divergence of the GLS through evolution. *S. calcitrans* contains a GLS-like signal that is the most closely related to the *Drosophilid* sequences whereas *A. aegypti* contains the most divergent sequence identified.

B. cucurbitae, and *C. capitata*, indicating that the primary sequences of these signals have diverged from a common ancestor. Finally, the tree shows that the *A. aegypti* GLS-like signal is the most divergent, as predicted from its low sequence identity and truncations in the signal (Figure 3.4B). These searches reveal, for the first time, new GLS-like sequences in seven species outside of the *Drosophilidae* family. From this, I conclude that the primary sequence of the GLS is indeed conserved in other species within the Diptera order.

The GLS secondary structure is conserved in some insects outside of the *Drosophilidae* family

Having identified several highly conserved GLS-like sequences within a range of Dipterans, I wanted to test whether the regions of conservation map to structural regions within the signal that could be important for function. I determined whether the new signals have the potential to form secondary structures resembling the GLS by folding each RNA using UNAFold.³⁵¹ I then mapped the non-conserved nucleotides (orange circles) onto each species, and summarised the results from all of these species (excluding *A. aegypti* due to its large truncations) by mapping them onto the *D. melanogaster* GLS structure (Figure 3.5). This revealed several regions of the GLS that are highly conserved in both primary sequence and predicted structure, such as a long AU-rich stem below the hairpin loop, and a U-rich distal internal loop. In addition, although some regions show low levels of sequence conservation

they still generate a similar structure to the *D. melanogaster* GLS. For example, all of the species, except for *A. aegypti* due to its large truncations, contain a proximal U-rich internal loop, albeit in a slightly shifted position compared to the GLS due to differences in the primary sequence. Several of the species maintain an internal bulge at the distal end of the signal, but the sequence of this region varies. All of the species, except for *A. aegypti*, contain a new internal loop replacing the proximal bulge at position 13 in the GLS.

The regions of lowest sequence conservation are the proximal third of the GLS and the hairpin loop (Figure 3.5). Although the stem structure and internal loop in the proximal third of the signal is maintained, the predicted size of the hairpin loop varies drastically between the Dipterans, as previously seen for the *Drosophilid* species (Figure 3.3B). Overall, the conserved structures identified may be particularly important for molecular interactions with *trans*-acting factors, and for the function of the GLS. Non-conserved regions, on the other hand, may provide insights into sections that are dispensable for function, aiding the design of GLS mutants to test their function. However, these signals would first need to be tested for their ability to localise in the oocyte using an *in vivo* injection assay.

Conclusions The GLS primary and secondary structure is highly conserved among the *Drosophilidae* family and a selection of other Dipterans. Analysis of the conserved nucleotides and structures will allow intelligent design of mutations in regions likely to be important for the function of the GLS.

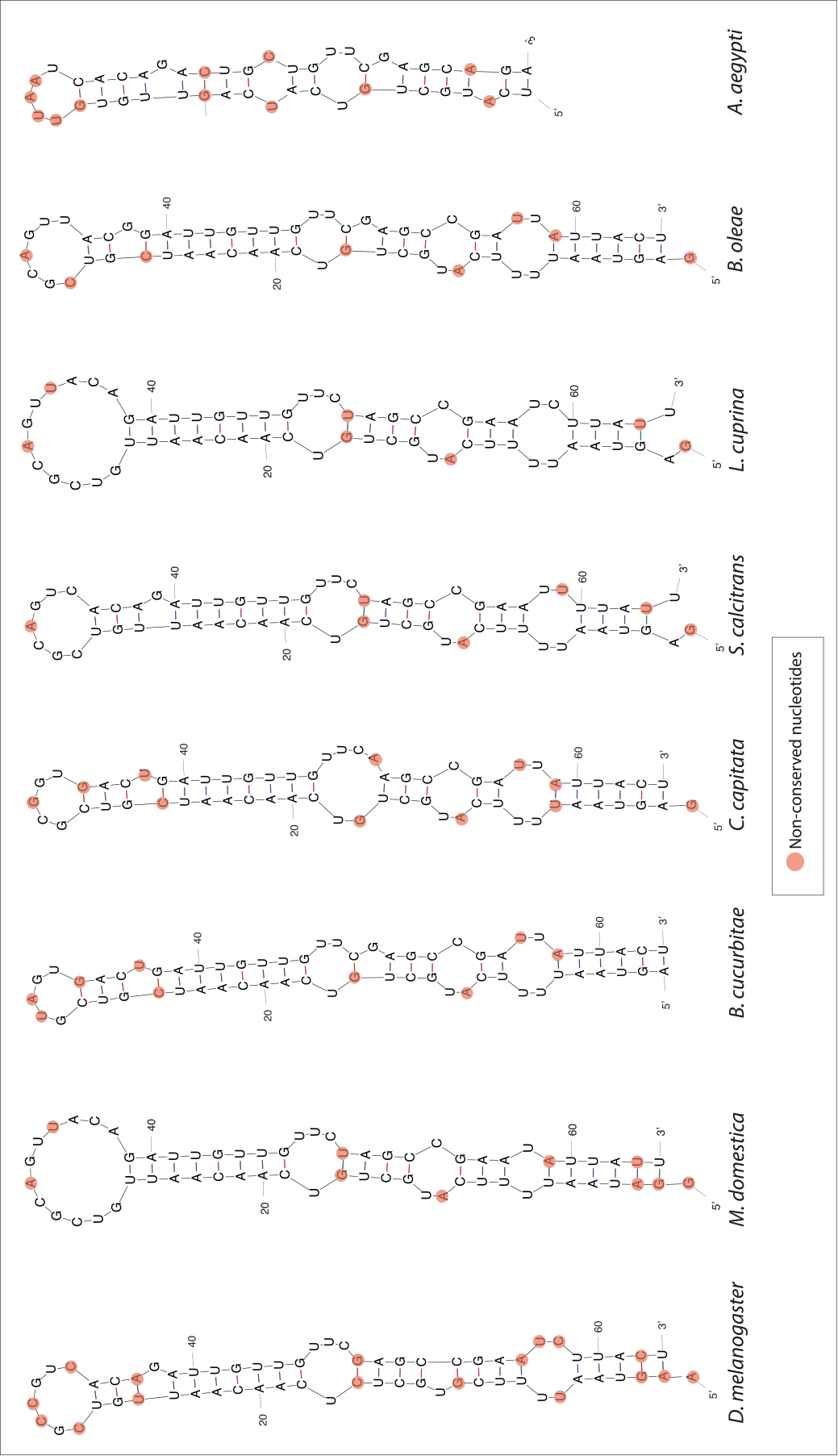


Figure 3.5. Several structures within the GLS are highly conserved in species outside of the *Drosophilidae* family. UNAFold predicted secondary structures of the GLS from *D. melanogaster* and other Dipterans found to have a GLS-like sequence reveal several regions that are structurally conserved even when the primary sequence is altered. Conserved regions include the two internal loops and a long stem below the hairpin. The size of the hairpin differs between the species and many contain an extra internal loop in place of the proximal internal bulge found in the *D. melanogaster* signal. *A. aegypti* contains a stem-loop with several structural similarities to the GLS but with a large truncation in the proximal third of the signal. Nucleotides that differ from the *D. melanogaster* GLS are highlighted in orange circles and have been mapped onto the *D. melanogaster* signal (excluding *A. aegypti* due to its large truncations) to summarise the data presented here.

Optimising the injection assay for efficient GLS mutant screening

Current *in vivo* RNA injection assays assess the pattern of RNA localisation at one hour post-injection.^{120, 307} In order to utilise this assay to efficiently screen for subtle defects in the localisation patterns of GLS mutants, we must determine the stages of oogenesis that are important for *grk* transport and anchoring at the DA corner, and the minimum assay length required to achieve complete localisation.

The different stages of *grk* mRNA localisation have been previously studied using the MS2-MCP system.²⁸³ In line with these studies, I show that MS2-MCP-GFP-labelled *grk* RNA (*grk*GFP*) localises the same as published *in situ* hybridisations of endogenous *grk* (Dresden Ovary Table resource).¹ At stage 7, *grk* RNA forms a ring around the cortex of the oocyte in both endogenous *in situ*'s and MS2-tagged RNA (Figure 3.6A and 3.6B, respectively). During stages 8-9, *grk* forms a tight cap above the nucleus at the DA (Figure 3.6A'-A'', 3.6B'-B'') and a weak ring around the anterior of the oocyte (data not shown).

RNA injection assays used to investigate *grk* transport and anchoring at the DA of the oocyte currently focus on injecting stage mid-9 oocytes.^{120, 307} However, studies addressing the dynamics of *grk* localisation using the MS2-MCP system have shown that highly dynamic *grk* RNA particles become localised to the DA of the oocyte during stages 7 to early 8 of oogenesis.²⁸³ They show that as the RNA becomes restricted to the DA during stages late 8

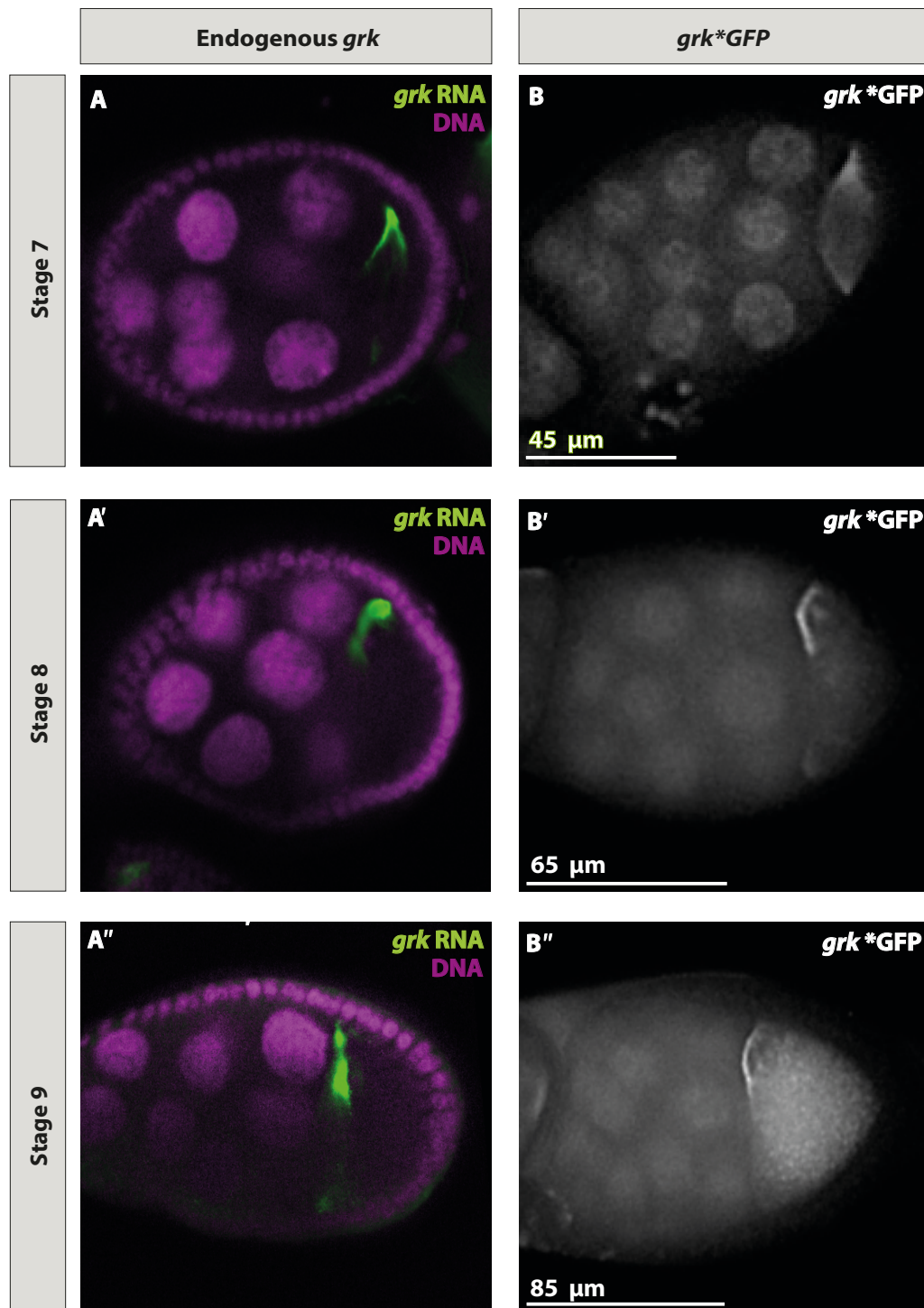


Figure 3.6. Endogenous *grk* *in situ* hybridisation and MS2-MCP tagged *grk* show very similar localisation patterns throughout mid-oogenesis. (A-A'') *grk* *in situ* hybridisation (Jambor *et al.*, 2015) and (B-B'') *grk*-MS2 MCP-GFP at different stages throughout oogenesis. At stage 7 both endogenous (A) and MS2-labelled (B, n=17) *grk* RNA is localised around the cortex while the oocyte nucleus is migrating from the posterior to the presumptive DA of the oocyte. During stages 8-9 endogenous (A'-A'') and MS2-labelled (B'-B'', n=34) *grk* RNA is localised to the DA of the oocyte forming a cap above the oocyte nucleus.

to early 9, it becomes less dynamic as a result of the anchoring process.²⁸³ By stage late 9 most *grk* RNA particles are static and fully anchored at the DA of the oocyte. As a result, I will implement my injection assay on stage late 8 to early 9 egg chambers to ensure the mutants are tested at the most biologically relevant time, where endogenous *grk* RNA is being actively transported and anchored.

Most injection assays use the minimum localisation signal identified to generate mutants and test for RNA localisation. The minimum 64nt stem-loop GLS has been previously shown to be both necessary and sufficient for *grk* localisation to the DA of the oocyte.³⁰⁷ Despite this, when designing GLS mutants I used the full length *grk* transcript as a backbone to ensure that the mutants were tested in their native context. Previous published assays have injected *grk* RNA at a concentration of 250-500ng/ul.^{120, 307} Injection of highly concentrated RNA (480ng/ul) has been shown to result in *grk* RNA localising to the core of P bodies where it is translationally repressed, as opposed to its usual destination at the edge of P bodies seen for endogenous transcripts, and when injected at a much lower concentration of approximately 120ng/ul.¹⁹³ To reduce any potential localisation defects associated with overloading the cellular machinery with high levels of injected RNA, I injected *grk* at the reduced concentration of 140ng/ul. I have shown that full-length *grk* RNA injected at this lower concentration is still able to localise to the DA of the oocyte, and the pattern of localisation can be clearly visualised (Figure 3.7D'; n=17).

Current injection assays in our lab culture the egg chambers in highly oxygenated halocarbon oil which increases the length of time you can keep the samples viable with low levels of stress.^{120, 307} In order to determine the degree of stress experienced by the egg chambers, I injected fluorescently labelled *grk* RNA into Me31B::GFP expressing oocytes. Me31B is a component of P bodies and stress granules and accumulates in large puncta in stressed specimens, making it a good indicator of stress.³²⁶ Levels of stress were measured in egg chambers at different time points following injection, or in the absence of injection, by the appearance of large, bright Me31B puncta. In un-injected stage 9 oocytes, Me31B::GFP

forms low levels of small, bright puncta in the cytoplasm of the oocyte (Figure 3.7A-A'). These Me31B::GFP puncta become enlarged and brighter over time with stress bodies beginning to form within the cytoplasm at 40 minutes (Figure 3.7B-B'), and increasing in size and number by 80 minutes (Figure 3.7C-C'; n=16). These results suggest, in agreement with previous literature, that the viability of the oocyte is maintained for up to 40 minutes of culturing.³⁷⁰

To test whether injection greatly increases stress to the specimen over time, Me31B::GFP egg chambers were injected with *grk* RNA and assessed for stress at different time points (n=17). At 40 minutes post-injection, the egg chamber shows a similar level of stress granule accumulation to the un-injected sample (Figure 3.7D, 3.7B, respectively), indicating that injection itself does not cause increased stress to the specimen. Within 40 minutes, *grk* RNA is also fully localised to form a strong cap above the nucleus at the DA corner of the oocyte (Figure 3.7D'-D''). This suggests that 40 minutes is adequate time in which to leave the RNA to localise fully, whilst maintaining the viability of the tissue.

Conclusions Fluorescently labelled full-length *grk* RNA localises the same as MS2-tagged and wild type *grk* mRNA when injected at low concentrations into stage late 8 to early 9 oocytes and allowed to localise for 40 minutes.

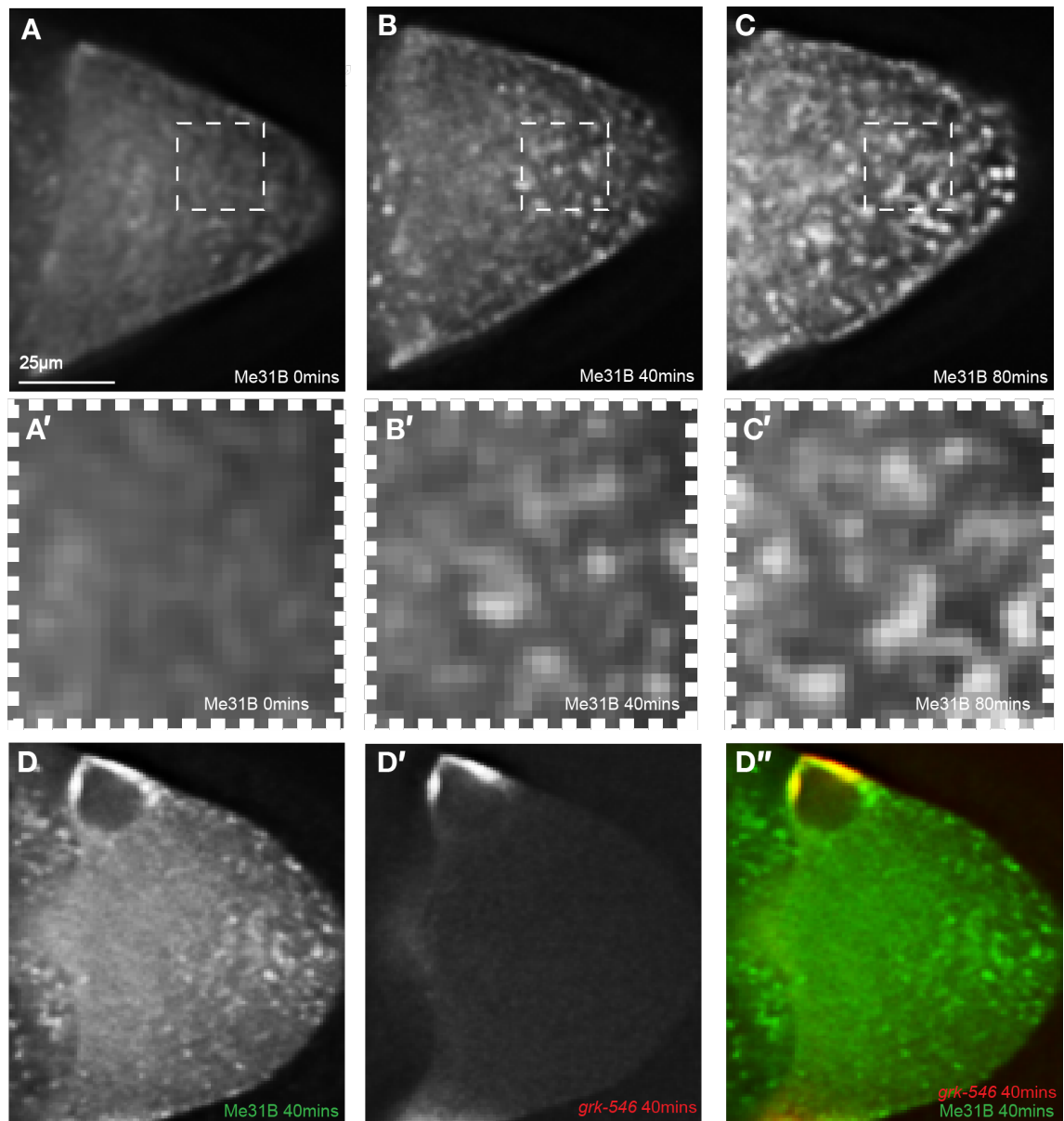


Figure 3.7. Injection produces endogenous-like *grk* RNA localisation and does not induce a stress response. (A-C) Stress granules accumulate over time in stage 9 Me31B::GFP egg chambers imaged every 10 minutes for 80 minutes ($n=16$). At 40 minutes the stress granules become larger and brighter indicating a reduction in the health of the egg chamber. Region highlighted in white square enlarged in A'-C'. (D-D'') Injection of *grk* RNA labelled with Alexa 546 (red) shows no additional damage to Me31B::GFP egg chambers, and the RNA localises in an endogenous-like manner within 40mins following injection ($n=17$). A-C, A'-C' and D are equally contrasted and D' and D'' are contrasted to best show RNA.

Subtle localisation defects can be identified by quantifying fluorescent RNA intensities in defined regions throughout the oocyte

grk mRNA has been shown to localise in a two step process: first, towards the anterior of the oocyte where it forms a ring of dynamic particles, followed by enrichment at the DA via an anchoring mechanism that is not well understood.^{120, 133, 283} Some RNAs, such as *bcd*, have been shown to localise to the anterior of the oocyte without undergoing such anchoring at the DA.²⁷⁹ This raises the question of whether transport and anchoring represent two distinct processes during *grk* localisation that could be controlled by different regions of the GLS. One way of testing whether distinct regions of the GLS are responsible for driving *grk* transport and anchoring is through mutagenesis of the RNA and assessing subtle defects in its localisation pattern. However, this requires a quantitative method to assess each step of localisation to be devised.

To test for transport of an injected RNA to the anterior of the oocyte, the mean fluorescent intensity of the RNA was measured using a line profile of 20x75 pixels (white box) through the centre of the oocyte every 10 minutes for a total of 40 minutes (Figure 3.8A-E). Immediately following injection into the centre of the oocyte, *grk* RNA forms a bright region at the injection site, creating a large single peak of fluorescence (Figure 3.8A-A'). At 10-20 minutes post-injection the RNA moves away from the site of injection towards the anterior (Figure 3.8B, 3.8C), reducing the size of the fluorescent intensity peak at the centre of the oocyte, and creating a smaller second peak at the anterior (~15 pixels, Figure 3.8B', 3.8C'). Over time, *grk* continues to enrich at the anterior (Figure 3.8D, 3.8E), and the intensity peak corresponding to the injection site reduces further until just above baseline at 40 minutes, while the second peak corresponding to the anterior of the oocyte continues to increase (Figure 3.8D',E'). A schematic of the pattern of *grk* localisation in the assay is outlined and was found to be typical for most of the injected oocytes (n=54) (Figure 3.8A''-E'').

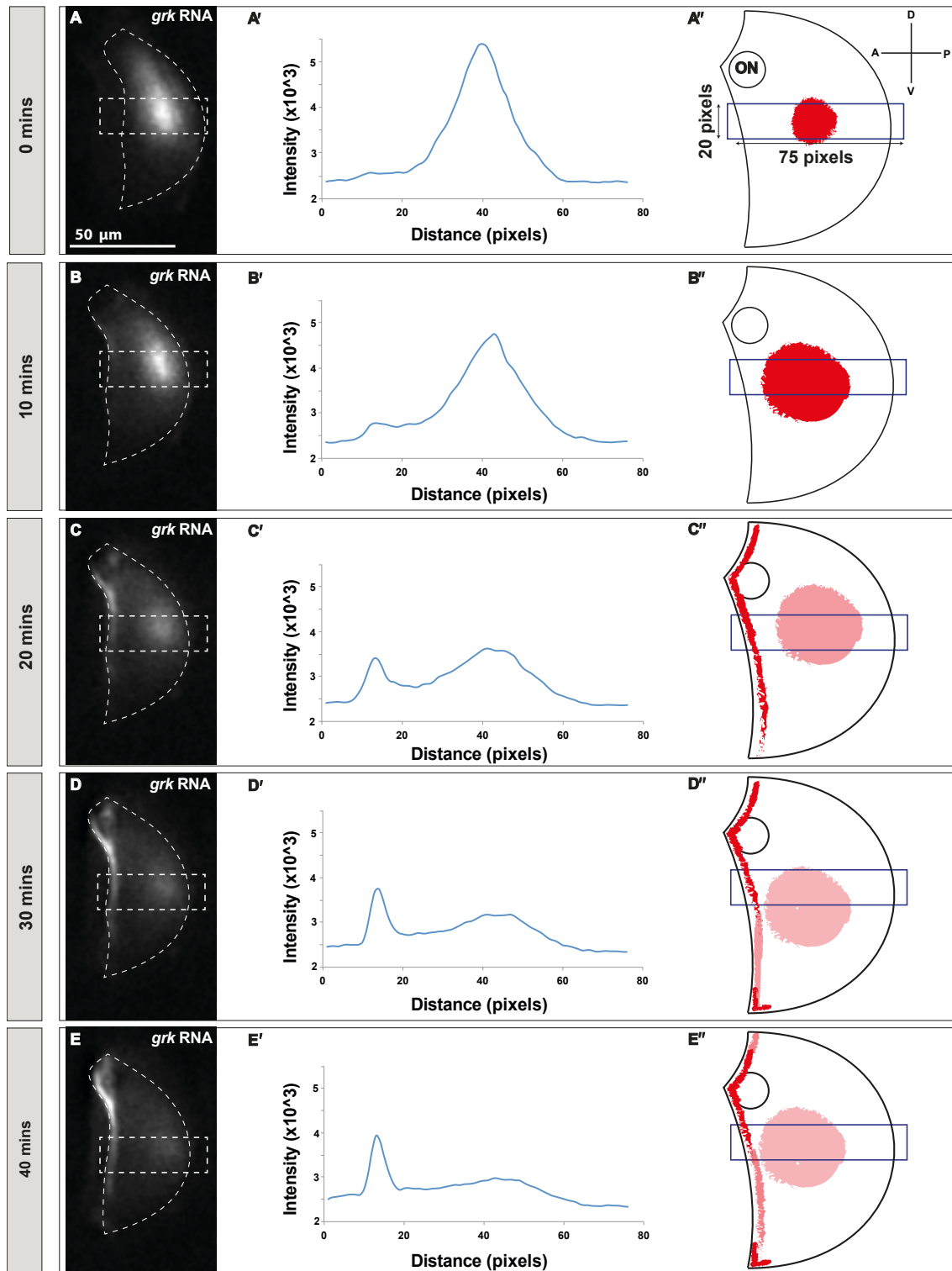


Figure 3.8. Injected *grk* RNA rapidly moves from the centre of the oocyte to the anterior before becoming enriched at the DA. (A-E) Immediately following injection *grk* RNA is seen as a small, bright region within the middle of the oocyte. Between 10 and 20 minutes *grk* RNA moves towards the anterior of the oocyte before it becomes further restricted to the DA by 40 minutes post-injection. (A'-E') Quantitation of fluorescent intensities of *grk* RNA through a cross-section of the oocyte (dashed rectangular region) over the 40 minutes following injection. (A''-E'') Schematic representation of the pattern of *grk* localisation following injection into the stage 8 oocyte. (n=54, ON = oocyte nucleus).

In order to decipher which regions of the GLS are required for *grk* anchoring I needed to quantify differences between anterior localisation (RNA that is not anchored) and DA localisation (RNA that is anchored). To do this, mean fluorescent intensity readings were taken at four regions within the oocyte, outlined in Figure 3.9A. The mean background value was subtracted to give the mean fluorescent intensity for each region within the oocyte every 10 minutes for a total of 40 minutes following injection. The mean fluorescent intensity for each region was then normalised to that of the centre (site of injection) at 0 minutes.

At 0 minutes post-injection, the normalised percentage RNA localisation shows that most of the total fluorescent RNA signal is at the injection site in the centre of the oocyte, and only a very low percentage is found at the anterior and the DA (Figure 3.9B). Between 0 and 20 minutes post injection, the fluorescent intensity decreases at the centre of the oocyte as *grk* moves towards the anterior (Figure 3.9B-C). After approximately 25 minutes, the fluorescent intensity of RNA at the DA begins to increase above the signal at the anterior, suggesting the RNA is enriching and anchoring at the DA corner (Figure 3.9B). *grk* becomes further enriched at the DA over time, leaving a weak signal around the anterior of the oocyte (Figure 3.9B-C), as previously shown for wild type and MS2-tagged *grk* RNA. This pattern of *grk* localisation was seen in 81% of oocytes injected (n=54) and is also demonstrated by calculating the relative distributions of the RNA in each region measured (Figure 3.9C).

To test whether these quantitative analyses can be used to distinguish between a positive localising RNA and one with localisation defects, it was necessary to design a negative control mutant RNA that was similar to the GLS, but with a disrupted localisation pattern. In order to design such a mutant, Russell Hamilton (a former postdoctoral fellow in the Davis Lab) generated a shuffled version of the GLS in which the same sequence composition is maintained but in a pseudo-random order, disrupting both the primary sequence and the predicted secondary structure of the GLS stem-loop. The resulting *scrambled* RNA shares only 33% sequence identity with the GLS (Figure 3.10A), and is predicted to form a vastly different secondary structure to the wild type GLS (Figure 3.10B-C).

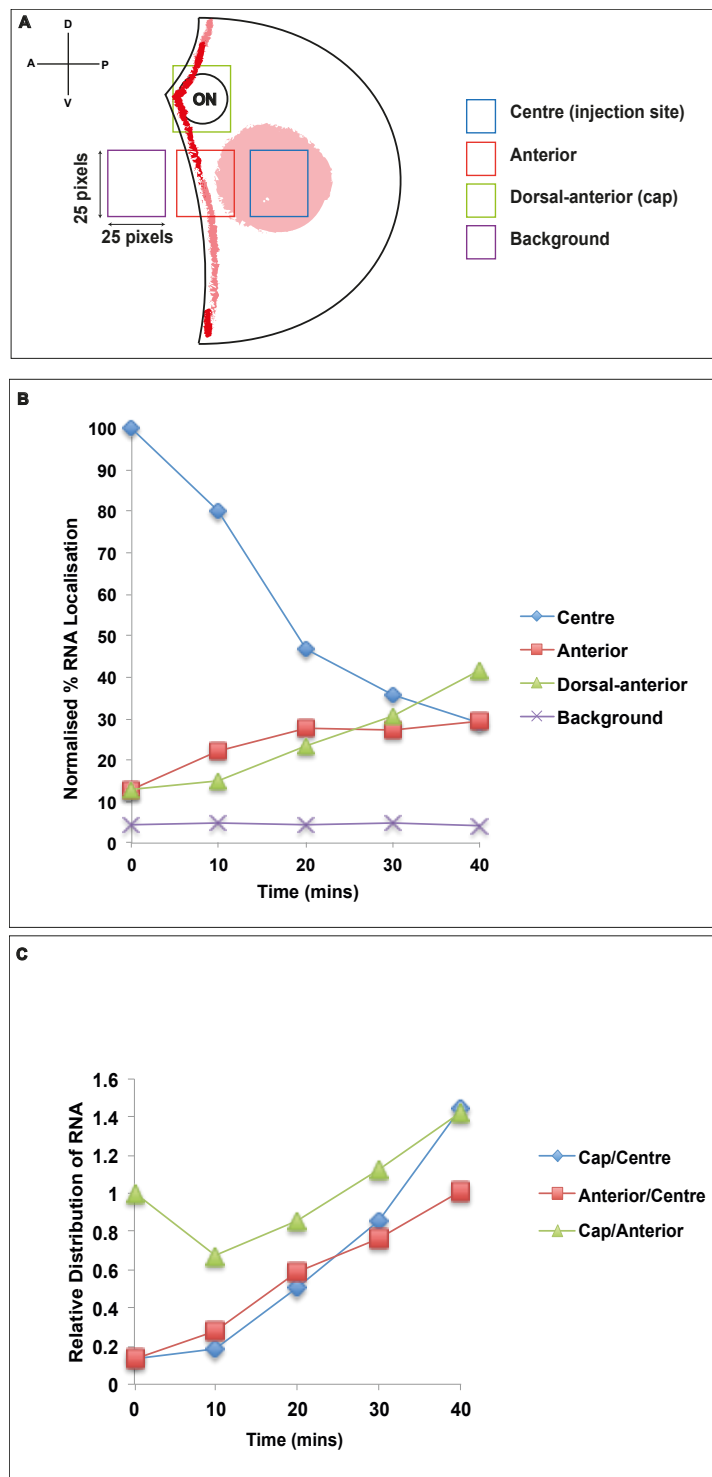


Figure 3.9. Injected *grk* RNA rapidly moves from the site of injection to the anterior of the oocyte before it becomes most highly enriched at the DA cap. (A) Schematic representation of fluorescent intensities measured for injected *grk* RNA for 40 minutes after injection. (B) Following injection, fluorescent *grk* RNA intensity decreases at the site of injection (centre of oocyte) and begins increasing at the anterior before it becomes enriched at the DA. (D) Injected *grk* RNA becomes more enriched at the DA over the anterior, the centre, and the background level of the oocyte over time. (Typical for $n=5$, ON = oocyte nucleus).

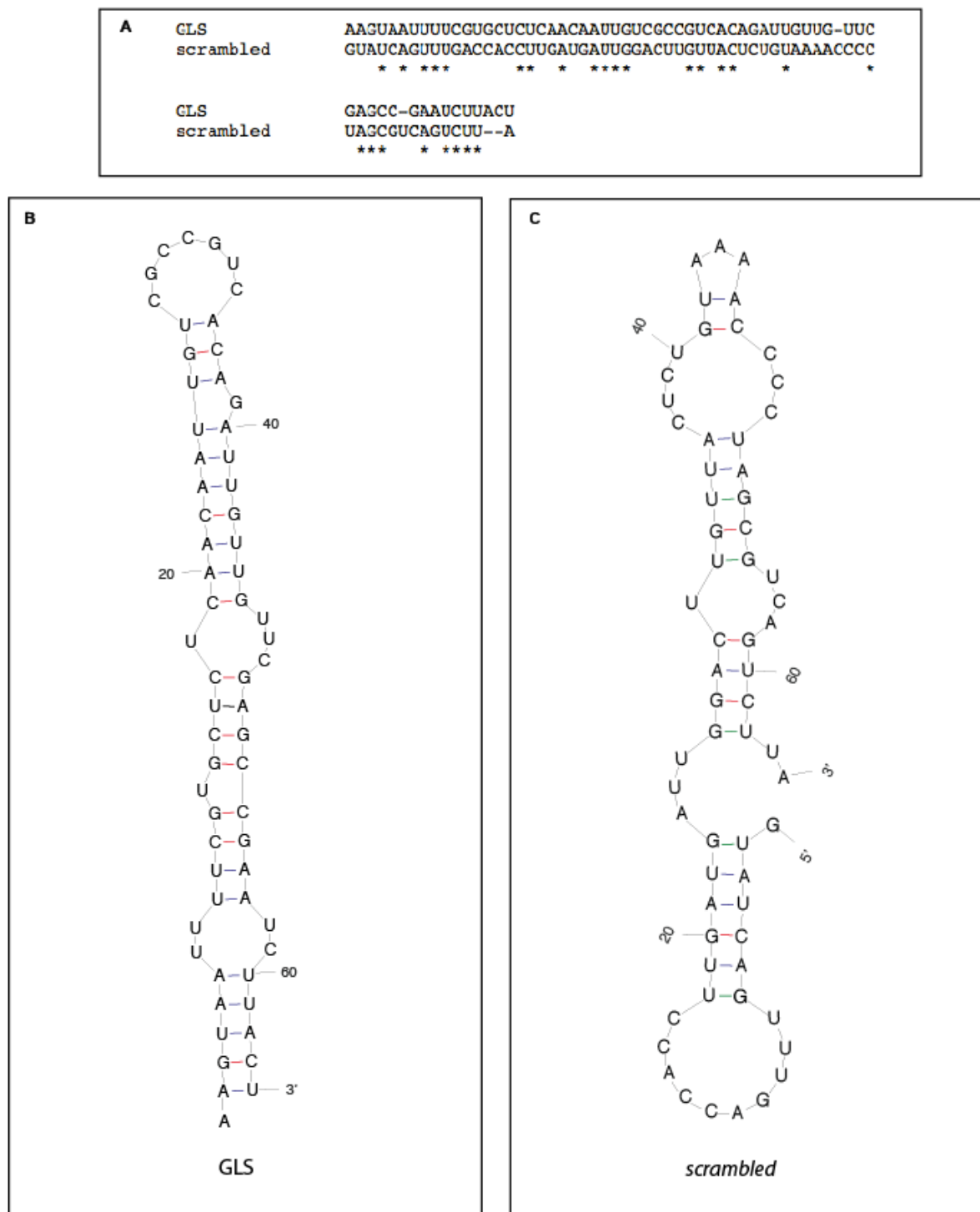


Figure 3.10. A *scrambled* version of the GLS shows a high degree of diversity from the wild type GLS in both primary and secondary structure. (A) Alignment of the wild type GLS with a *scrambled* version designed as a putative negative control for localisation. UNAFold predicted secondary structures of the wild type GLS (B) and the new *scrambled* RNA (C) indicate the secondary structure has been vastly altered.

To test whether the *scrambled* GLS mutant was a sufficient negative control for *grk* localisation, I replaced the GLS with *scrambled* and assessed its ability to drive full-length *grk* localisation following injection in the oocyte. Within 10 minutes of injection (Figure 3.11B',C'), *scrambled* and the non-localising control dye Dextran disperse uniformly in all directions from the bright injection site in the centre of the oocyte (Figure 3.11A-C), whereas wild type *grk* travels in the direction of the anterior (Figure 3.11A-A'). Differences between the localisation patterns of both Dextran (n=3) and *scrambled* RNA (n=16) with *grk* RNA (n=54) becomes clearer over time with *grk* RNA becoming more and more enriched at the anterior, followed by the DA, and the negative controls becoming more diffuse throughout the cytoplasm with time (Figure 3.11A''-C'''), resulting in an even distribution of signal all over the oocyte after 40 minutes.

In order to quantitate the localisation pattern of the new *scrambled* negative control, mean fluorescent intensities were measured from a line profile through the middle of the oocyte over time. From 0 minutes to 40 minutes post-injection *scrambled* RNA did not localise to the DA of the oocyte and instead was diffuse throughout the cytoplasm (Figure 3.12A-E). This is represented in the fluorescent intensity profile by a large peak in the centre of the graph at 0 minutes which widens and flattens over time, corresponding to the RNA become more diffuse throughout the oocyte (Figure 3.12A'-E'). The pattern of localisation for *scrambled* RNA is presented schematically (Figure 3.12A''-E'').

To determine whether these quantitative methods could be used to distinguish transport mutants from anchoring mutants the fluorescent intensities of the *scrambled* RNA injection was assessed for enrichment at the DA. The results show that the mean fluorescent intensities measured in each region of the oocyte, outlined in Figures 3.13A and 3.13B, are distinctly different from injection of wild type *grk* RNA. In contrast to wild type *grk* RNA, which enriches first at the anterior and then at the DA cap (Figure 3.13A'), the mean fluorescent intensity of *scrambled* decreases at the centre and begins increasing at both the anterior and the DA of the oocyte at the same time (Figure 3.13B'). This indicates that *scrambled* RNA

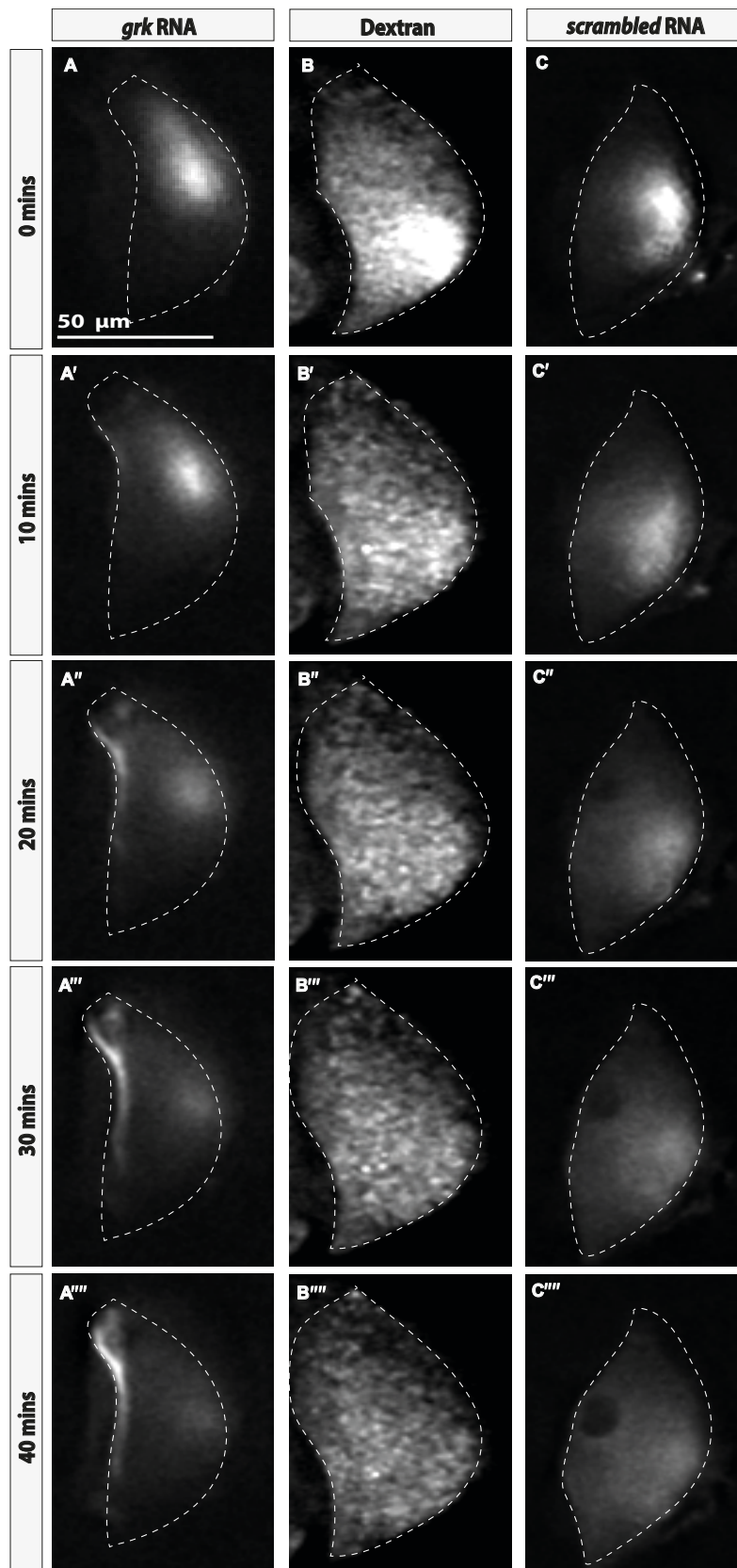


Figure 3.11. Full length *grk* RNA containing a *scrambled* form of the GLS is unable to localise in the oocyte. (A-A''') injected wild type *grk* RNA localises to the dorsal-anterior of the oocyte within 40 minutes. (B-B''') Injected dextran dye diffuses throughout the cytoplasm following injection. (C-C''') *grk* RNA containing a scrambled version of the GLS is unable to localise in the oocyte and diffuses throughout the cytoplasm.

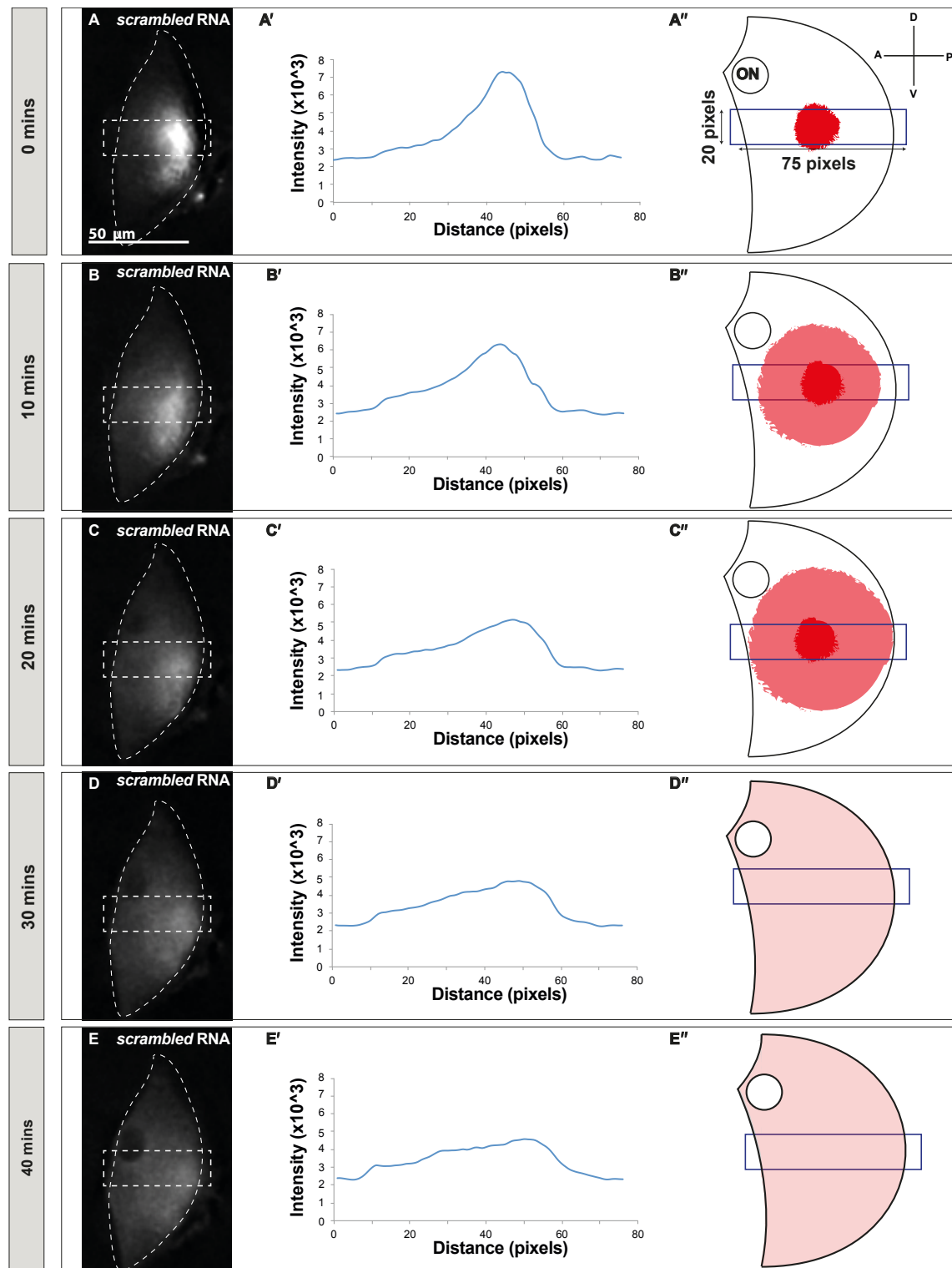


Figure 3.12 Injected *scrambled* RNA does not localise in the oocyte and is diffuse throughout the cytoplasm. (A) *scrambled* RNA is seen as a small, bright region at the middle of the oocyte following injection. Between 10 and 40 minutes (B-E) *scrambled* RNA diffuses throughout the cytoplasm and becomes evenly distribution within the oocyte. (A'-E') Quantitation of fluorescent intensities of *scrambled* RNA through a cross-section of the oocyte (dashed rectangular region) over the 40 minutes following injection. (A''-E'') Schematic representation of the localisation of injected fluorescently labelled *scrambled* RNA. (n=16, ON = oocyte nucleus).

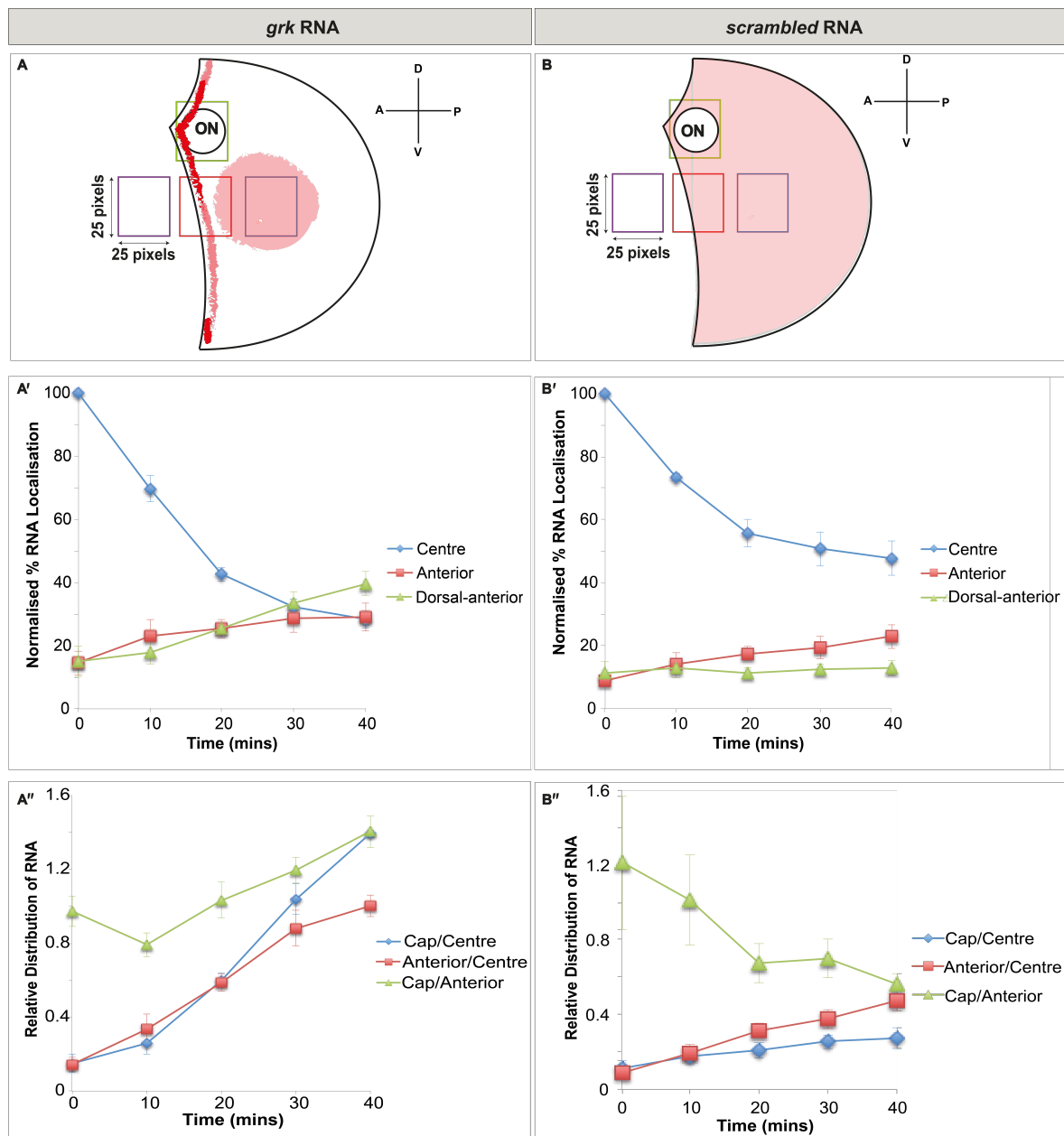


Figure 3.13. Injected scrambled RNA travels from the centre of the oocyte and is uniformly dispersed throughout the cytoplasm. Schematic representation of fluorescent intensities measured at different regions of the oocyte for 40 minutes following the injection of *grk* RNA (A) and *scrambled* RNA (B). These regions include the centre (blue box), the anterior (red box), the dorsal-anterior (cap, green box), and the background (purple box). (A'-B') Percentage RNA localised in each region over time normalised to the centre at 0 minutes post-injection. Following injection, *grk* RNA fluorescent intensity begins increasing at the anterior before it becomes enriched at the dorsal-anterior (A'). The fluorescent intensity of *scrambled* RNA begins to decrease in the centre and increases slightly at the anterior. However, the dorsal-anterior (cap) never reaches higher than the anterior (B'). (A''-B'') Relative distribution of RNA in different regions of the oocyte for 40 minutes post-injection. Injected *grk* RNA becomes more enriched at the dorsal-anterior (cap) than the anterior, the centre, and the background level of the oocyte over time (A''). Injected *scrambled* RNA never becomes more enriched at the dorsal-anterior (cap) than other regions of the oocyte (B''). (n=5 per condition, error bars represent SEM, ON = oocyte nucleus).

does not carry out two distinct steps of localisation. The relative distribution of *scrambled* RNA shows that within 10 minutes following injection, the relative mean RNA intensity at the anterior and the DA is minimal, indicating that both regions have a similar amount of RNA enrichment (Figure 3.13B"). This is in contrast to wild type *grk* RNA which forms a pattern in which the relative mean RNA intensity at the DA (cap) versus the anterior begins at 1 (equal levels in both regions) and steadily increases as the RNA moves from the anterior and enriches at the cap (Figure 3.13A"). By 40 minutes, the region with the highest RNA enrichment is the centre at the injection site, and the signal at the anterior and DA never enriches above the centre, suggesting that the RNA is uniformly distributed, except for some residual RNA left in the centre at the injection site (Figure 3.13B"). In the wild type GLS condition, on the other hand, the RNA intensity at the DA (cap) is higher than at the centre and the anterior at 40 minutes post-injection (Figure 3.13A"). These data indicate that *scrambled* RNA is unable to efficiently carry out the two steps of *grk* localisation and is likely failing to be transported and anchored at the DA.

Discussion

The GLS is highly conserved among the *Drosophilidae* family and in a selection of other Dipterans.

My data reveal new GLS-like signals in several *Drosophilid* species and a selection of other Dipterans. From the conservation of the primary and predicted secondary structure of these signals, we can isolate five regions of the GLS that may be important for *grk* localisation in *Drosophilids*. However, the data suggest that variations in even single nucleotides surrounding these regions of primary sequence conservation can have a large effect on the resulting predicted secondary structure of the RNA. The regions of highest conservation fall within the two long stems located between the internal loops and below the hairpin. However, due to the flexibility of the GLS generated by the large hairpin loop and the number of internal loops, the NMR structure of the full-length GLS could not be obtained (Ball and McDermott, unpublished).

The hairpin loop and the two internal loops form the regions of the GLS with the lowest sequence conservation, but they are still present in varying sizes in all of the *Drosophilid* species. However, *D. miranda* lacked the proximal internal loop due to its truncated sequence, suggesting this region is not essential for the function of the GLS. This is supported by injection experiments carried out previously in the lab showing that a 49nt truncated form of the GLS similar to the *D. miranda* signal, known as M1, could localise normally to the DA of the oocyte, albeit with a slightly lower efficiency (Van De Bor unpublished, Chapter 4 of this thesis). Surprisingly, species with the lowest level of sequence conservation did not always correspond with those with the most divergent predicted secondary structures. For example, *D. virilis* and *D. willistoni* contained four nucleotide variances but were predicted to completely retain the secondary structure of the GLS, whereas *D. takahashii* and *D. biarmipes* only contained three nucleotides that varied but were predicted to fold differently at both the hairpin loop and the distal internal loop. These data confirm the importance of secondary structure predictions in designing mutant GLS RNAs.

Due to the high levels of conservation of the GLS among *Drosophilids*, we could only define large regions as having potential importance for driving *grk* localisation. Bioinformatic searches for more distant GLS-like signals outside of the *Drosophilidae* family revealed similar sequences in a further seven species with varying degrees of primary sequence and secondary structure conservation. The most highly conserved region fell within the distal stem shown to be conserved in the *Drosophilid* signals. However, the highly conserved stems between the internal loops in *Drosophilids* are less well conserved in these signals, forming a new internal loop in all but one of the non-*Drosophilid* species, and suggesting its primary sequence rather than secondary structure may be important for GLS function. *A. aegypti* was the most divergent signal identified making it an interesting candidate to test in the RNA injection assay to see whether it is competent to drive *grk* localisation.

Current injection assays can be optimised to increase throughput of mutant testing

Current *in vivo* RNA injection assays inject large concentrations of the minimal *grk* localisation signal (GLS) into stage mid-9 oocytes and assess localisation patterns at one hour post-injection.^{120, 307} However, previous studies have shown that *grk* RNA has undergone most of its transport and anchoring by stage late 8 to early 9 of oogenesis,²⁸³ and that injection of high concentrations of RNA results in the inappropriate localisation of *grk* RNA to the interior of P bodies where it is translationally repressed.¹⁹³ As I plan to study subtle localisation distinctions between point mutants in the GLS, I decided to reduce the incubation time to avoid variability in the assay because of possible damage to the oocyte in longer incubations, and increase the throughput for testing a large number of mutants.

By assessing damage to the specimen resulting from long culturing times and the process of oocyte injection itself, I demonstrated that un-injected oocytes expressing the P body and stress body marker Me31B showed stress body accumulation after 80 minutes, despite having no damage from injection. However, after only 40 minutes of culturing, the egg chambers showed substantially lower signs of stress in both the presence and absence of injected *grk* RNA. In addition, I showed that low concentrations of full-length injected *grk* RNA could localise correctly to the DA of the oocyte within 40 minutes of injection. I conclude that reduced incubation times minimise the stress to the specimens and increase the throughput, so that the injection assay can be used to assess subtle differences between many GLS mutants more effectively.

Localisation patterns can be quantitated to assess subtle localisation defects of injected RNAs

Previous studies investigating the role of the GLS in *grk* localisation used *hunchback* (*hb*) RNA or full-length *grk* with the GLS deleted (FL Δ GLS) as negative controls for localisation.³⁰⁷ *hb* is an RNA that is distributed evenly throughout the oocyte during oogenesis and, although

it is a good control for assessing non-localised transcripts, it is likely to be modulated by different molecular components to *grk* RNA. The FL Δ GLS, on the other hand, contains the rest of the *grk* gene meaning many of the endogenous control mechanisms for *grk* would still be used. However, removal of a region of the RNA is likely to result in substantial changes to its secondary structure and brings together two sequences within the RNA which would normally not be in contact. This could potentially result in the formation of novel protein binding sites. In order to directly compare GLS mutants in the same biological background as endogenous *grk*, we designed a *scrambled* form of the GLS and tested its ability to drive localisation of full-length *grk* RNA. My data demonstrate that *scrambled* RNA is unable to localise in the oocyte, and can therefore be used as a novel negative control for *grk* localisation assays.

An assay to assess subtle defects in the localisation of GLS mutants requires an objective, quantitative read-out of localisation patterns. Most injection assays involve qualitatively assessing localisation patterns in a manner that is potentially subject to experimenter biases. I have shown that localisation patterns can be quantitated by measuring the fluorescent intensities of the RNA in different regions of the oocyte throughout the time course of the assay. This can be used to identify differences between RNAs that are capable of being transported to the anterior of the oocyte and those that are not. In addition, these methods can recognise differences between RNAs that are able to enrich at the DA of the oocyte and those that cannot. As a result, I am now able to investigate the effect of GLS mutations on the localisation pattern of *grk* in a fast and quantitative manner.

Future Work

Newly identified GLS-like signals should be tested for functionality by their ability to localise in the *D. melanogaster* oocyte. This would provide evidence for the regions of the GLS that are essential for localisation and would aid the design of mutants to further test their functionality. Mutagenesis studies should be directed toward the most highly conserved

regions of the signal, in sequence and/or structure, including the distal stem and the internal loops. These mutations should address the role of the sequence and the structure of each region in *grk* localisation, as the function of the GLS is likely a result of molecular interactions involving a combination of the two. Again, these mutants should be tested for functionality by injection into the oocyte. The RNA injection assay could be further optimised by titrating the concentration of the RNA in order to find the lowest concentration required to maintain adequate assessment of the localisation pattern. Once candidate mutants are identified as transport or anchoring mutants they could be validated *in vivo* by the generation of transgenic flies expressing the mutant form of the GLS. MS2-MCP labelling of these mutants would allow direct visualisation of their localisation patterns in the oocyte. Some of the future work proposed here is addressed in the following chapters.

Final Conclusions

Here I have identified novel GLS-like signals in a number of species and shown that the GLS is highly conserved in both primary sequence and secondary structure throughout the *Drosophilidae* family, and in other Dipterans. Further analysis of these signals will aid the design of sequence modifications in the GLS to investigate its structure-function. To test the effects of large numbers of sequence modifications on the function of the GLS I have adapted current RNA injection assays to increase the throughput and reduce potential variability that could result from prolonged culturing of the specimen. I have shown that fluorescently labelled, full length *grk* RNA can localise like wild type *grk* at 40 minutes post-injection into stage late 8 to early 9 oocytes, whereas a *scrambled* version of the GLS cannot, providing a negative control for assessing RNA transport and anchoring. Finally, I have demonstrated that the two steps of *grk* localisation can be quantitated allowing the detection of mutants with subtle localisation defects.

Chapter 4: Regions of the GLS required for *grk* RNA transport

Introduction

Since the discovery of the GLS, biochemistry studies have identified a number of *trans*-acting factors that interact with the signal that were previously shown to be important for *grk* localisation using genetics.^{15, 54, 133, 162, 299, 304, 315, 383} Several of these proteins have also been shown to function together in mammalian cells, such as the dendrites of hippocampal neurons.^{123, 384, 385} Despite the wealth of knowledge about the GLS protein-interacting partners, little is known about how they recognise and interact with the RNA on a molecular level, except for Egl which has been shown to bind to A'-form RNA conformations in the localisation signal of *fs(1)K10 (K10)*.³⁸⁶

The GLS is predicted to form a highly structured stem-loop containing two internal loops, two internal bulges, several stem regions, and a hairpin loop. NMR studies predict several of the stem regions to form RNA helices in an A-conformation (Ball and McDermott, unpublished) where the RNA bases likely to be required for generating specificity for protein interactions are buried inside the deep and narrow major groove of the helix.³⁴⁸ RNA secondary structures such as internal loops and bulges have been shown to distort the backbone between helical regions in order to widen the major groove, making it deeper and allowing proteins to access the buried bases.³⁴⁸ Therefore, both the sequence and the structure of the GLS may be equally important in generating the specificity required for recognition by *trans*-acting factors.

In this chapter, I have tested the ability of the novel GLS-like signals identified in Chapter 3 to localise within the *D. melanogaster* oocyte using my optimised oocyte injection assay. I have used this information to design sequence and structural modifications in the GLS and tested

them *in vivo* to identify regions that are essential for *grk* transport in the oocyte, and recruitment of transport machinery components in the embryo.

***grk* RNA is transported in the oocyte along microtubules using the molecular motor Dynein**

grk RNA localises to the DA corner of the oocyte during mid-oogenesis to establish the dorsal-ventral axis of the egg and the future embryo.^{15, 387} DA *grk* localisation is a two-step process that requires the Dynein motor, microtubules and the GLS. First, *grk* mRNA is transported to the anterior of the oocyte where it is found as dynamic particles. Second, *grk* mRNA becomes restricted to, and statically anchored at, the DA.^{120, 133}

The GLS drives *grk* localisation by interacting with the transport component Egl and its non-RNA binding partner BicD, and the anchoring RBP Sqd.^{54, 133, 304, 307} BicD and Egl link *grk* to the microtubule network by recruiting the Dynein motor, forming an Egl-BicD-Dynein complex.^{108, 116, 118, 270} Despite its lack of an obvious RNA-binding domain, Egl interacts with the signal sequences of RNA cargoes, such as the GLS, while BicD binds the N-terminal domain (NTD) of Egl via its C-terminal domain (CTD), and Dynein via its NTD.^{304, 388} Interactions between Egl and the Dlc, are also required to produce an active transport complex for *grk* localisation.^{108, 119, 304}

Although several of the proteins involved in *grk* transport and anchoring have been characterised, how these proteins specifically distinguish their cargos from other RNAs is a challenging question that many groups working in the field are trying to address.³⁸⁹ A big step towards understanding how the transport machinery recognises its cargo came from studies showing that Egl interacts with two spatially registered A'-form helices in one of its target RNAs *K10*, directly linking it to the Dynein motor.³⁸⁶ A'-form is generated by stacking interactions of three or more purines on the same side of a canonical, dsRNA helix which results in lower base-pair inclinations and a widened major groove, similar to that seen in B-

form DNA.³⁸⁶ The authors predicted two such A'-form helices to form in the GLS, providing a possible insight into how the molecular machinery recognises *grk*. Designing and testing modifications within these putative Egl binding sites would reveal whether they are essential for recognition of the GLS by the transport machinery.

Newly identified GLS-like signals should allow effective design of GLS mutations to disrupt transport

The GLS is highly conserved among the *Drosophilid* family and in other Dipterans both in terms of primary sequence and predicted secondary structure (Chapter 3, this thesis).^{305, 307} Previous studies using cross-species injections have shown that the *D. melanogaster* GLS is capable of localising normally in the *D. virilis* oocyte and vice versa, demonstrating the high degree of conservation of both the localisation signals and the transport and anchoring machinery.³⁰⁷ The GLS is likely to be highly conserved due to its dual function in directing *grk* localisation and encoding part of the signal peptide, which makes determining the regions important for driving *grk* localisation more challenging. GLS-like signals with low sequence similarities but several common predicted secondary structures have also been identified that drive transcripts to the anterior or DA of the oocyte, including the ILS, G2LS and JLS.^{307, 335} Testing the ability of the highly conserved GLS-like signals to drive *grk* transport through injections into the *D. melanogaster* oocyte would be the first step in determining whether they can be used to infer which regions of the GLS are essential, and to design modifications to test these regions.

Recruitment of GLS binding proteins to mutant GLS RNAs can be assessed in living embryos

The mRNA transport machinery components are often conserved and widely used in several tissues.³⁹⁰ The Egl-BicD-Dynein transport complex is utilised for RNA transport in several tissues throughout development including egg chambers, syncytial blastoderm embryos,

embryonic neuroblasts, and epithelial cells.³⁹⁰ For example, this motor complex has been shown to localise approximately 10% of mRNAs transcribed in the NCs into the oocyte^{270, 391, 392} and many mRNAs, such as the pair-rule genes, to the apical cortex of the embryo.^{16, 116, 337, 393}

Cross-tissue injections of mRNAs originating from the oocyte or the embryo show that these mRNAs are successfully recognised and localised to the anterior or apical cortex, emphasising the high degree of conservation of the transport machinery.¹¹⁶ Moreover, embryo injections have demonstrated that components of the localisation machinery, such as Egl and BicD, are recruited to injected RNA within 1-2 minutes, including those of oocyte origin.¹¹⁶ Due to the large size of embryos and their ease of collection and imaging, injection of *in vitro* transcribed *grk* RNA into embryos is a good assay for addressing which localisation machinery is recruited to the GLS.

Aims of this chapter and experimental rationale

In this chapter I will design sequence modifications within the GLS based upon the conservation of the signal to identify regions that are essential for *grk* transport. The specific questions addressed are:

1. Do the novel GLS-like signals drive *grk* RNA transport?
2. Can information gained from the novel GLS-like signals be used to aid sequence modification design?
3. Which regions of the GLS are required for recruitment of transport machinery components and thus *grk* transport?

In order to design sequence modifications within the GLS I will test whether a selection of the novel GLS-like signals identified in Chapter 3 are recognised by the transport machinery by injecting the RNA into the *D. melanogaster* oocyte and embryo. The sequence and secondary structure predictions of those signals that retain localisation activity will be used to predict regions of the GLS that are important for *grk* transport, and to design sequence modifications within these regions to test the hypothesis. The effects of the sequence modifications on *grk* transport will be investigated by assessing localisation of the RNA in the oocyte, and recruitment of the transport machinery in the embryo.

Results

The most divergent *Drosophilid* GLS-like signals drive *grk* transport in the *D. melanogaster* oocyte

In the previous chapter I identified 19 species of *Drosophilid* containing a signal with high primary sequence and secondary structure similarity to the *D. melanogaster* GLS (Table 3.1; Figure 3.2-3.3). To test whether they can drive transport of full-length *grk* RNA, I chose GLS-like signals from two *Drosophilids*, cloned them into the *D. melanogaster grk* gene, and assessed their localisation following injection as fluorescently labelled RNA into the *D. melanogaster* oocyte. The GLS-like signals from *D. biarmipes* and *D. takahashii* show a high degree of secondary structure divergence in the distal internal loop, whilst being highly conserved in their primary sequence conservation with the *D. melanogaster* GLS (95%, Table 3.1). The *D. takahashii* signal also shows an increase in the size of the hairpin loop, removing bulge 2 at position 39, whereas the hairpin loop in *D. biarmipes* is constricted, generating a tetra-loop and a new internal loop (Figure 4.1A). Both the *D. takahashii* and *D. biarmipes* GLS-like signals efficiently drive *grk* localisation from the centre of the oocyte towards the anterior within 20 minutes of injection in 96% and 97% of the oocytes, respectively, in a similar manner to the *D. melanogaster* GLS (Figure 4.1B-F). For ease of comparison, wild type *grk* control images are shown in multiple figures. These results indicate that the signals maintain the sequence or structural motifs required to be efficiently recognised by the *D. melanogaster* transport machinery.

Conclusions The GLS-like signals with the most divergent predicted secondary structures from the *Drosophilidae* family drive *grk* localisation within the *D. melanogaster* oocyte.

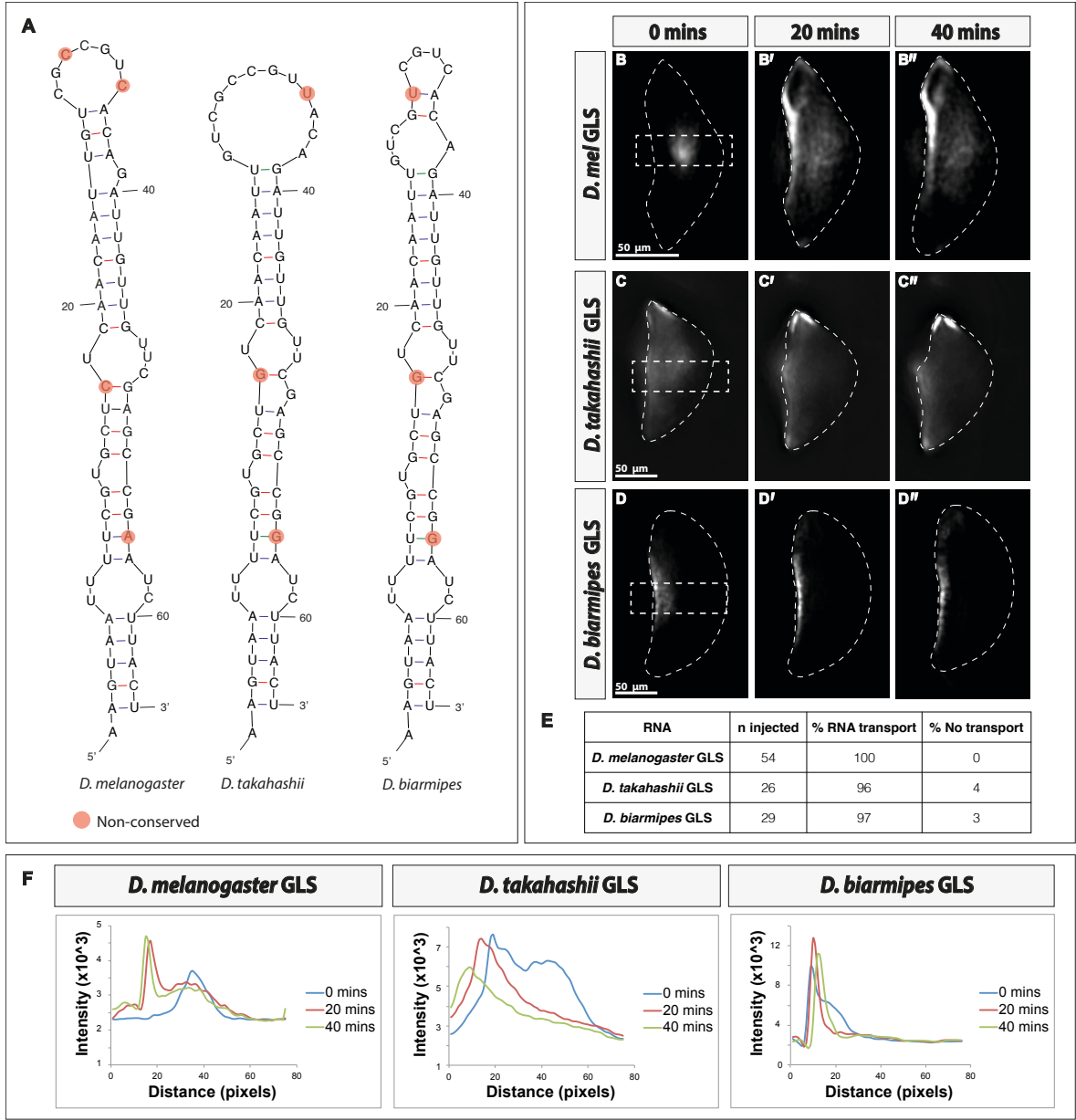


Figure 4.1. The *Drosophilid* GLS-like signals with the most divergent secondary structures are transported to the anterior of the oocyte. (A) UNAFold predicted secondary structures of the GLS signals from *D. melanogaster*, *D. takahashii*, and *D. biarmipes* suggest they are highly conserved in regions except internal loop 2, bulge 2, and the hairpin loop. Non-conserved nucleotides are highlighted in orange circles and have been mapped on to the *D. melanogaster* GLS. (B-E) Injection of these signals into the *D. melanogaster* oocyte shows they all undergo efficient transport to the anterior of the oocyte within 40 minutes of injection. White dashed rectangles show the region of the oocyte used to measure the fluorescent intensity profiles in (F) at 0 mins, 20 mins and 40 mins following injection of RNA.

Several GLS-like signals from Dipterans drive *grk* transport in the *D. melanogaster* oocyte

Bioinformatic searches for GLS-like signals outside of the *Drosophilidae* family returned sequences in seven species within the Diptera order. To investigate whether these signals can drive *grk* transport in the *D. melanogaster* oocyte, three were chosen that vary in secondary structure (Figure 4.2). First, the GLS-like signal from *M. domestica* contains an enlarged hairpin loop, a new internal loop, and a reduced internal loop 1. Second, the signal from *B. cucurbitae* contains a reduced hairpin loop and internal loop 2, and a new internal loop at position 11 in place of bulge 1. Third, the *C. capitata* GLS-like signal contains a reduced hairpin loop, an enlarged internal loop 2, and a new internal loop in place of bulge 1 at position 12. All three species showed transport to the anterior of the oocyte within 20 minutes following injection (Figure 4.3B-F). However, the *M. domestica* signal had a slightly reduced function, being transported to the anterior of 73% of the oocytes injected compared to 100% for the *D. melanogaster* GLS (Figure 4.3A-A'', 4.3E). The transport efficiencies of the signals from *B. cucurbitae* and *C. capitata*, on the other hand, were unaffected by changes in the primary and secondary structures of the RNA.

The most divergent GLS-like signal identified came from *A. aegypti* and contains a poorly conserved proximal stem that is not predicted to form part of the stem-loop (Figure 4.4A, blue box). Injection of a truncated form of the signal without the non-conserved region showed it is unable to support *grk* transport (Figure 4.4C-E) like the wild type GLS (Figure 4.4B-B'', 4.4D-E). Although we cannot rule out the possibility that the non-conserved proximal third of the signal is required for localisation, the data suggests the signal has diverged too far to be recognised by the *D. melanogaster* transport machinery.

Conclusions Novel GLS-like signals from outside the *Drosophilidae* family drive *grk* transport in the *D. melanogaster* oocyte, with the exception of the most divergent signal from *A. aegypti*

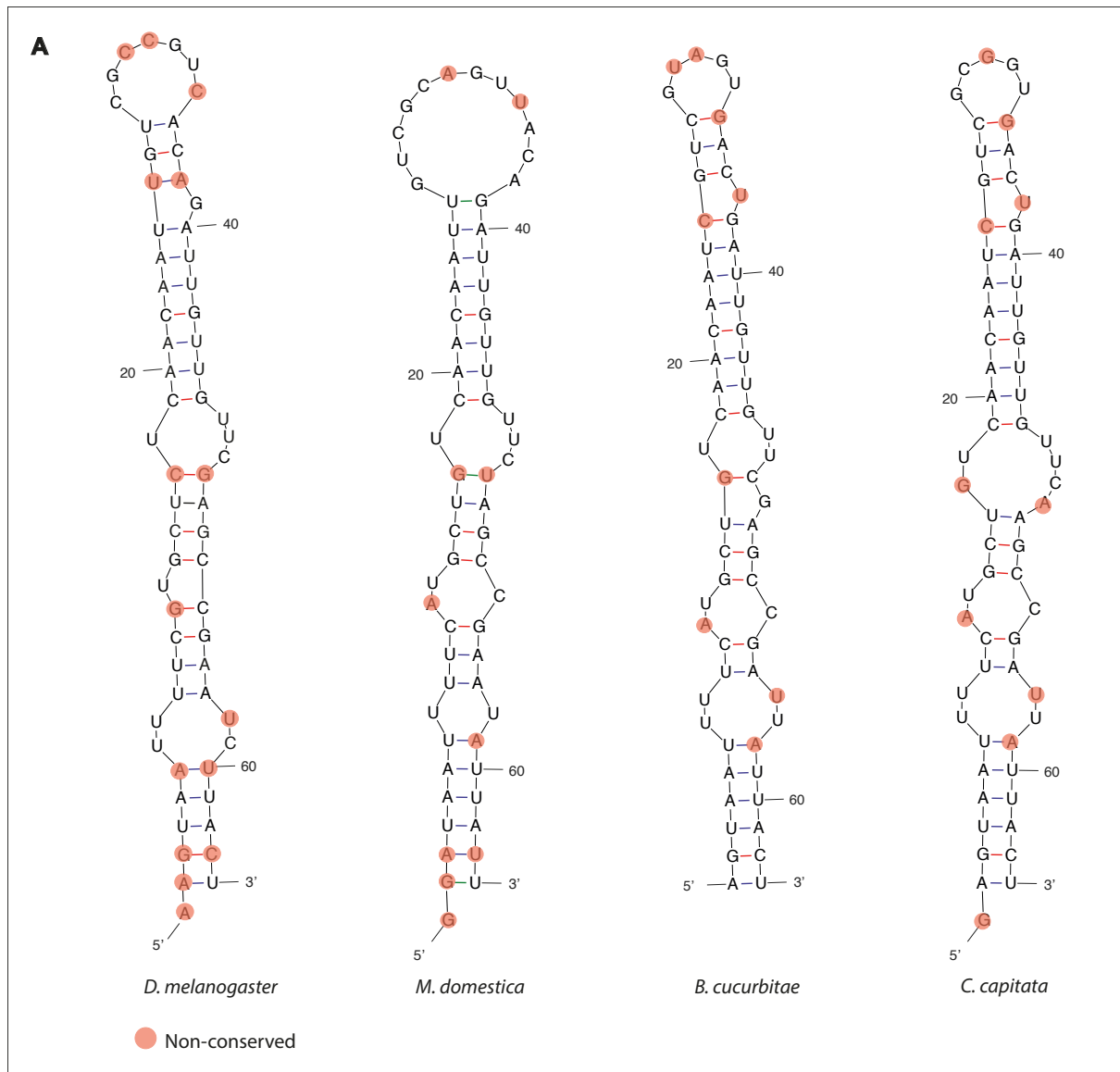


Figure 4.2. The GLS-like signals from three Dipterans form secondary structures that vary at the internal loops, bulges and hairpin loop. (A) UNAFold predicted secondary structures of GLS-like signals found in three different species indicate differences in the size of the hairpin and internal loop 2, as well as an extension of bulge 1 into a new internal loop, and changes in the size of internal loop 1. These differences in structure are a result of a small number of changes in the primary sequence in each signal (highlighted in orange circles). The non-conserved nucleotides from all three species are summarised on the *D. melanogaster* GLS.

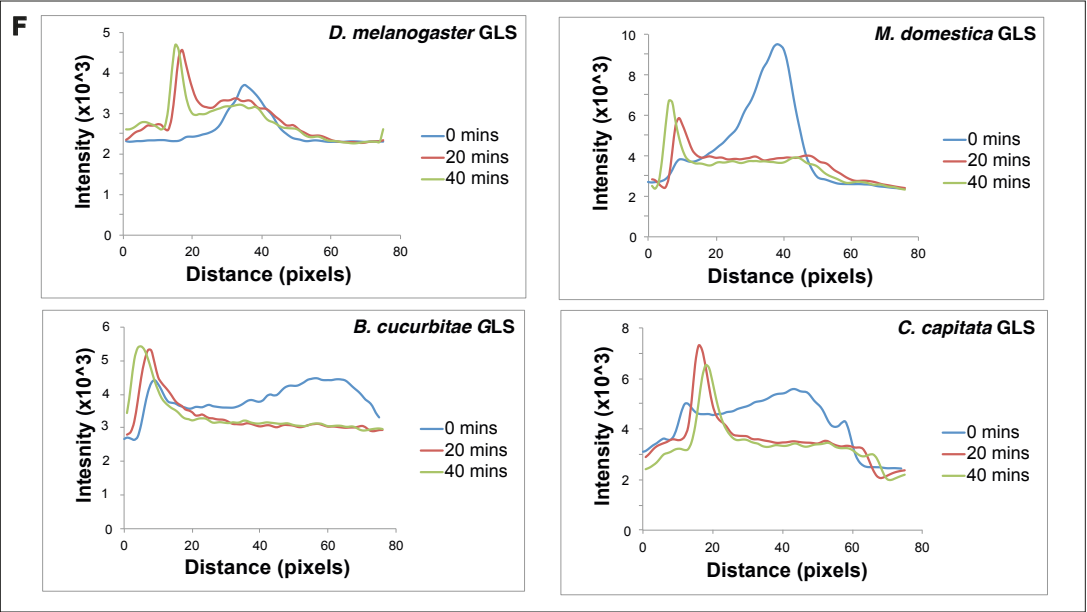
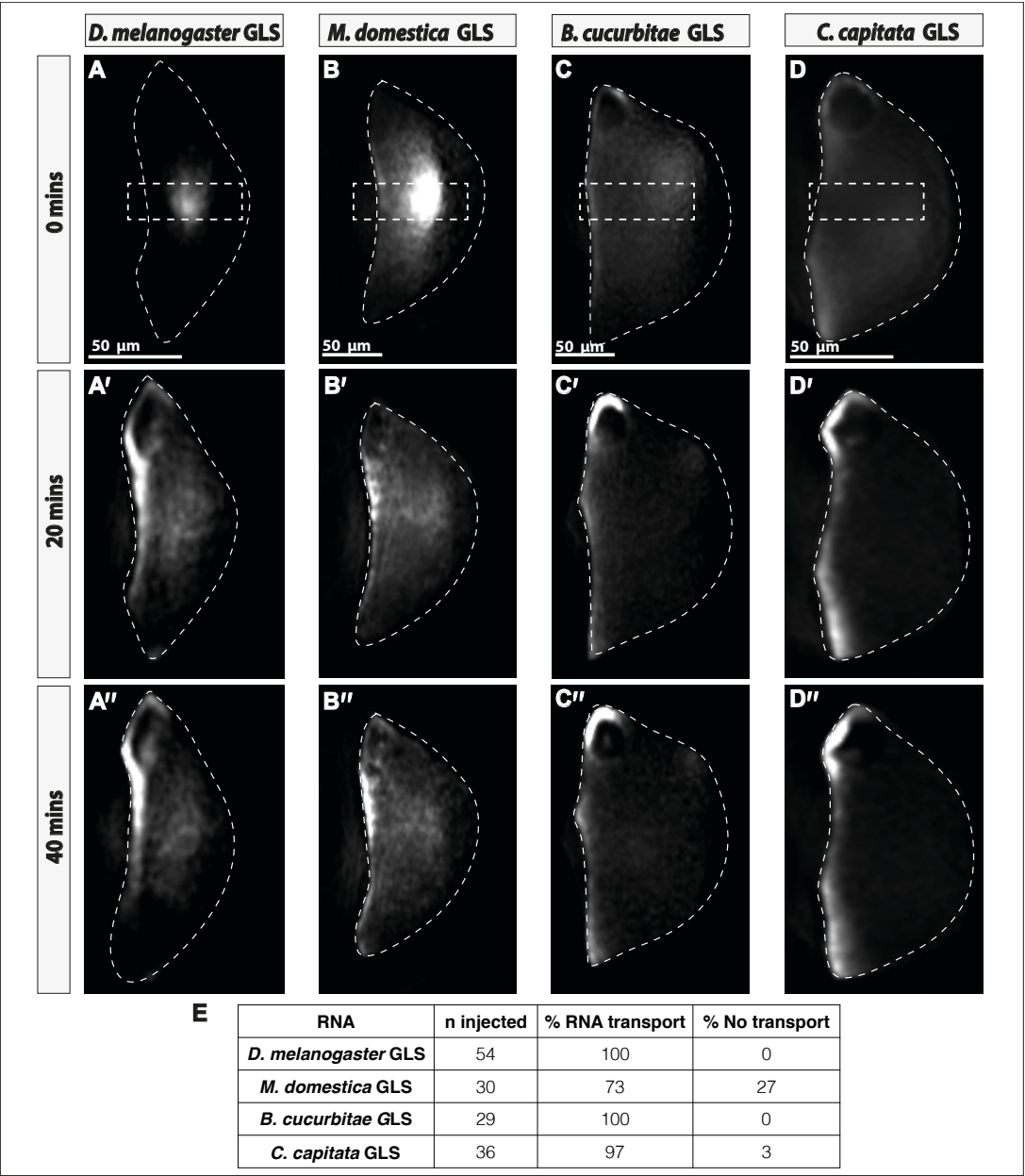


Figure 4.3. Novel GLS-like signals from species outside of the *Drosophilidae* family are able to drive *grk* transport in the oocyte. (A-A'') The GLS drives the transport of *grk* RNA injected into the *D. melanogaster* oocyte. Cross-species injections show that GLS-like signals from *M. domestica* (B-B''), *B. cucurbitae* (C-C'') and *C. capitata* (D-D'') can all efficiently drive *grk* transport when injected into the stage 8-9 *D. melanogaster* oocyte. (E) All of the signals, except from *M. domestica*, can drive RNA transport as efficiently as the *D. melanogaster* GLS, reaching the anterior of the oocyte within 20 minutes of injection (F). White boxes show the region of the oocyte used to measure the fluorescent intensity profiles in (F) at 0 mins, 20 mins and 40 mins following injection of RNA.

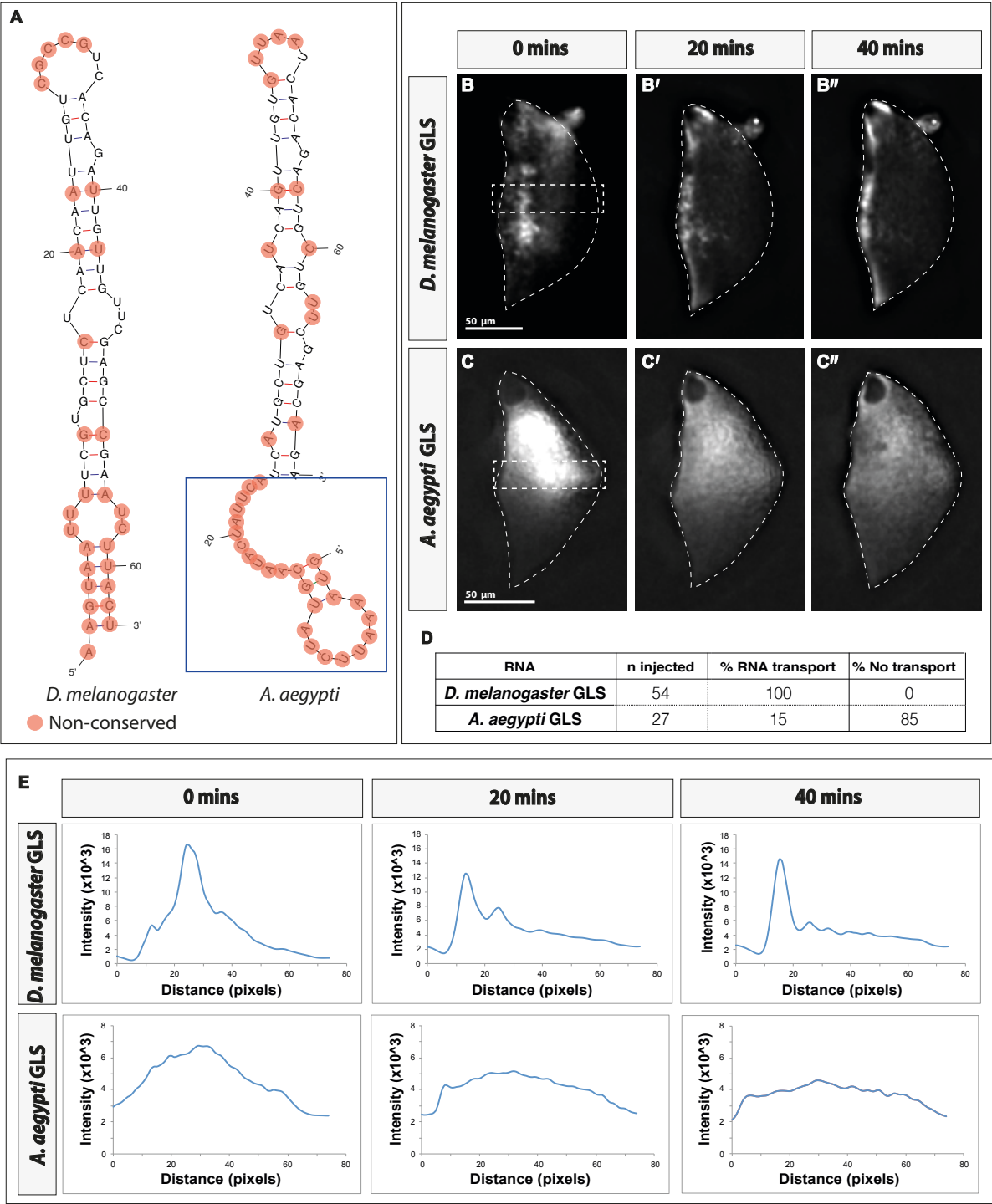


Figure 4.4. The GLS-like signal from *A. aegypti* is the most divergent identified and is incapable of driving *grk* localisation in the oocyte. (A) UNAFold secondary structure prediction of the GLS-like signal from *A. aegypti* shows it contains several highly conserved features, despite its low primary sequence conservation (non-conserved nucleotides highlighted in orange circles and mapped onto the *D. melanogaster* GLS, blue box shows region removed for injections due to low sequence conservation). (B-B'') The GLS from *D. melanogaster* efficiently drives *grk* transport to the anterior of the oocyte within 20 minutes of injection, whereas *grk* containing the *A. aegypti* signal is diffuse throughout the cytoplasm (C'-C'') in 85% of oocytes injected (D). (E) Measuring the fluorescent intensities of the RNA in a cross-section of the oocyte shows that the *D. melanogaster* signal drives *grk* towards the anterior of the oocyte whereas the *A. aegypti* signal cannot. White boxes show the region of the oocyte used to measure the fluorescent intensity profiles in (E) at 0 mins, 20 mins and 40 mins following injection of RNA.

The conservation of the GLS can be used to design sequence modifications to disrupt *grk* transport

Due to the high levels of conservation of the GLS among *Drosophilids*, we could only define large regions as having potential importance for *grk* localisation. Comparing these highly conserved regions (defined as three or more consecutive conserved nucleotides) from the *Drosophilids* (Figure 3.2B) to those from the more divergent Dipterans (Figure 3.4A) reveals several regions that are common to all species identified, except for the truncated signals of *A. aegypti* and *D. miranda* (Table 4.1).

In order to identify conserved regions of the stem-loop secondary structures, I mapped the non-conserved nucleotides from all of the species onto the *D. melanogaster* GLS (Figure 4.5). However, due to its inability to localise in the *D. melanogaster* oocyte, the signal from *A. aegypti* was removed from this analysis along with the truncated *D. miranda* signal. A high number of non-conserved nucleotides fall within the hairpin loop and bulge 2, and are predicted to affect the size of the hairpin and the position of the bulge (Figure 4.5A). As predicted by its poor conservation, the sequence of the hairpin loop has been previously shown to not be required for *grk* transport, but a minimum of 5-nucleotides in the hairpin is essential (Van de Bor and Ball, unpublished).

Conserved sequence	Nucleotide position	Present in <i>Drosophilids</i>	Present in other Diptera
TA <u>ATTTTC</u>	4-11		✓
<u>TTTTCGTGCT</u>	7-16	✓	
<u>IGCT</u>	13-16	✓	✓
<u>ICA</u>	18-20	✓	✓
<u>ICAACA</u>	18-23	✓	
TTGTC	25-29	✓	
ACAGATTGT <u>IG</u>	36-46	✓	
<u>IGTTC</u>	45-49		✓
<u>CGAGCCG</u>	49-55	✓	
<u>AGC</u>	51-53	✓	✓
TTA	60-62	✓	✓

Table 4.1. Several regions of the GLS primary sequence are conserved among the newly identified GLS-like signals from other species. The nucleotide positions of conserved regions found throughout the GLS and whether they are present in species from the *Drosophilidae* family and other Dipterans. Nucleotides that span more than one region are underlined.

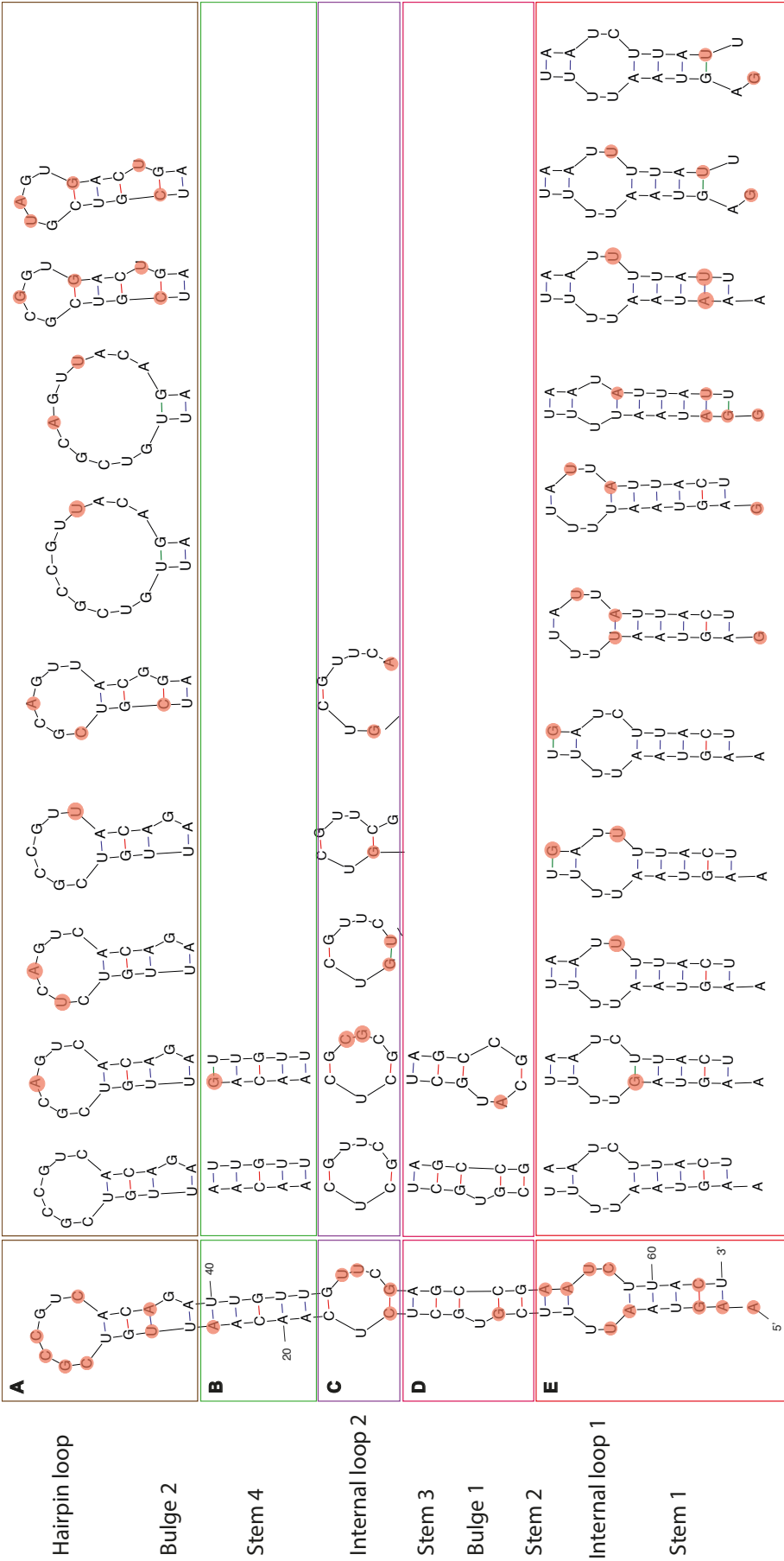


Figure 4.5. The stem regions of the GLS are the most highly conserved in both sequence and secondary structure. (A) The hairpin loop and bulge 2 of the GLS are some of the least conserved regions resulting in a variety of secondary structures predicted by UNAFold. (B) Stem 4 is highly conserved in sequence and secondary structure. (C) Internal loop 2 shows low levels of sequence conservation but is present in all species. (D) Stem 3 is highly conserved and is located above bulge 1 in all species. Bulge 1 is expanded into an internal loop in some cases. Stem 2 is less conserved but is present in all species above internal loop 1, which has low sequence conservation (E). Stem 1 has low sequence conservation but is present in almost all of the signals identified. (Non-conserved nucleotides are highlighted in orange circles and are mapped onto the GLS).

Stem 4 is the most highly conserved region of the signal with only one non-conserved nucleotide which is predicted to have no effect on its secondary structure (Figure 4.5B). This suggests that the primary sequence and the secondary structure of stem 4 may be important for the function of the GLS. Internal loop 2 is poorly conserved and differs in size between the species (Figure 4.5C). Despite this, the internal loop is always present in the signal, suggesting a single-stranded sequence may be important here for recognition by *trans*-acting factors.

Stem 3 is entirely conserved throughout the signals and is always found above a single-stranded region formed by bulge 1 that could provide accessibility for protein binding to conserved sequences within stem 3 (Figure 4.5D). Stem 2 is less conserved (Figure 4.5D-E) than stem 3, suggesting that the sequence of this region may not be important for the function of the signal. However, these sequence changes are mostly predicted to have no effect on the overall secondary structure of stem 2, suggesting it may be important structurally.

The proximal third of the GLS, composed of internal loop 1 and stem 1, is the least conserved region in both sequence and predicted secondary structure, and is completely absent in the *A. aegypti* and *D. miranda* signals. Non-conserved nucleotides in this region cause a variety of different predicted secondary structures altering the size of the internal loop or its position within the stem. Therefore, the sequence of the proximal third of the GLS

may not be essential for function, but the presence of stem 1 and an internal loop may be important.

Conclusions The primary sequence of the GLS is highly conserved in stem 4 and stem 3. Two internal loops, two bulged nucleotides, and a hairpin loop are present in every species but their sequences are not highly conserved.

The proximal third of the GLS is not essential for *grk* transport in the oocyte.

The GLS has been defined as a 64-nucleotide stem-loop that is both necessary and sufficient to drive *grk* localisation.³⁰⁷ However, the proximal third of the GLS is poorly conserved (Figure 4.5E), suggesting it may not be essential for function. In fact, a truncated 49-nucleotide form of the GLS called M1 was shown to localise in the oocyte following injection (Van De Bor, Ball and McDermott, unpublished).

In order to define the minimum localisation signal capable of driving *grk* transport in the oocyte, I designed three further truncations in the GLS: M1-3, M1-5 and M1-7. Despite truncations within these signals, the remaining sequence is predicted to fold in a similar manner to the full-length GLS (Figure 4.6A). Unlike the full-length GLS (Figure 4.6B-B", 4.6G), M1-3 causes *grk* to form a diffuse signal throughout the cytoplasm in 100% of oocytes, indicating that one or more regions removed from this signal are essential for *grk* transport (Figure 4.6C-C", 4.6G; Table 4.2). Restoring bulge 1 to produce M1-5 was unable to rescue *grk* localisation, with 77% of oocytes showing a diffuse signal throughout the cytoplasm following injection (Figure 4.6D-D", 4.6G; Table 4.2), and the remaining 33% showing DA enrichment in addition to the diffuse signal. However, the addition of half of stem 2 in M1-7 reduces the percentage of oocytes showing a diffuse signal throughout to 45%, with 23% showing DA enrichment and a diffuse signal (Figure 4.6E-E", 4.6G; Table 4.2). In addition, 32% of oocytes showed strong DA enrichment with low levels of signal throughout the oocyte indicative of efficient *grk* transport. Finally, M1 directs *grk* to the DA of the oocyte

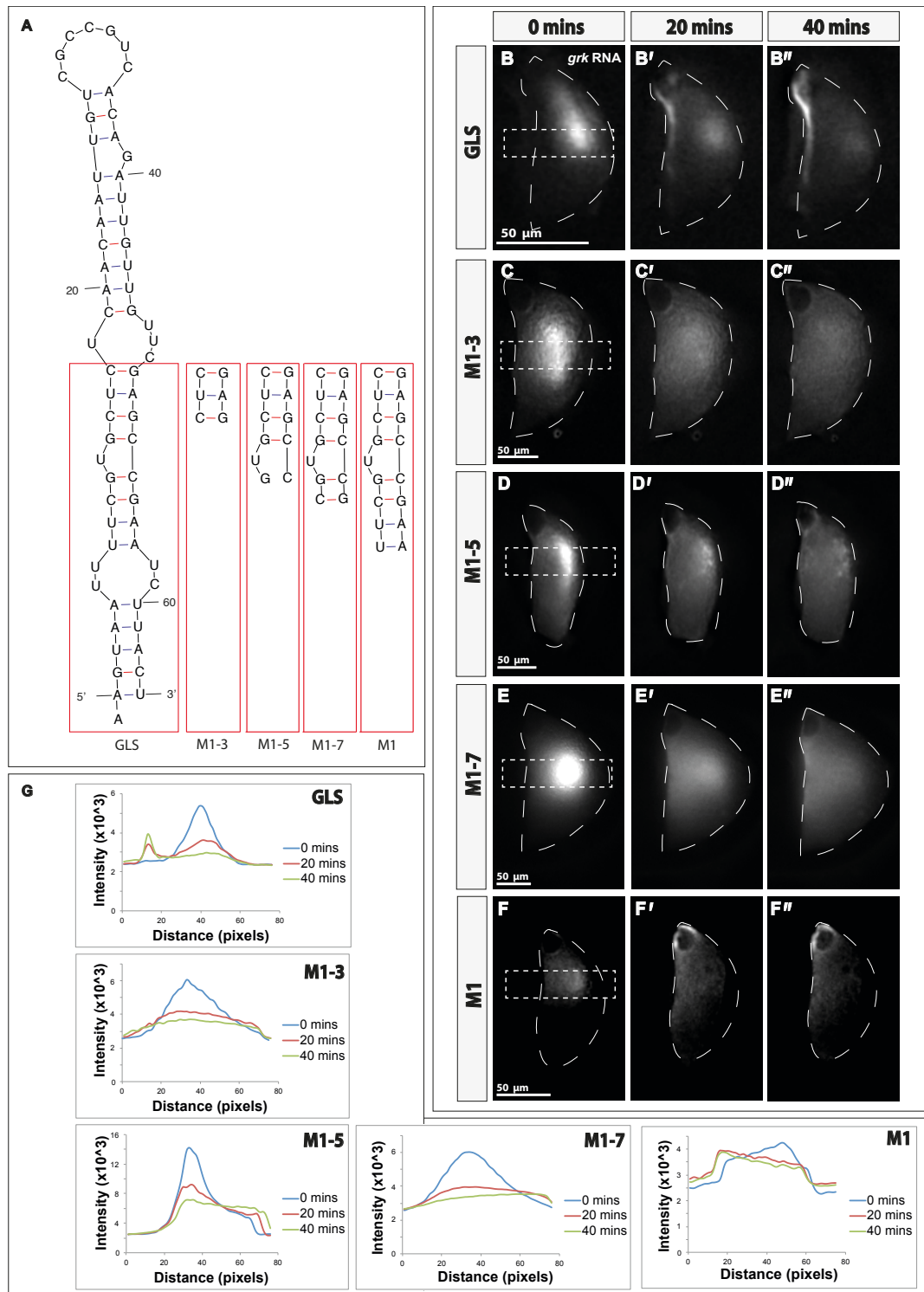


Figure 4.6. Stem 1 and internal loop 1 are not essential for *grk* transport in the oocyte but increase efficiency. (A) UNAFold secondary structure predictions suggest that truncating the GLS has no effect on the secondary structure of the remaining stem-loop. (B-B'') Wild type *grk* RNA is transported to the anterior of the oocyte within 20 minutes of injection and then becomes enriched at the DA ($n=54$). M1-3 (C-C''), M1-5 (D-D'') and M1-7 (E-E'') injected RNA localises throughout the oocyte within 20 minutes in most of the oocytes injected. (F-F'') M1 RNA is transported towards the anterior of the majority of oocytes within 20 minutes of injection and then becomes enriched at the DA by 40 minutes. White boxes show the region of the oocyte used to measure the fluorescent intensity profiles in (G) at 0 mins, 20 mins and 40 mins following injection of RNA.

RNA	n injected	% RNA transport	% No transport
GLS	54	100	0
M1-3	22	0	100
M1-5	31	23	77
M1-7	31	55	45
M1	18	67	33

Table 4.2. The efficiency of *grk* transport in the oocyte is greatly reduced by truncating the proximal third of the GLS. Various truncations to the proximal third of the GLS result in a reduction in the percentage of oocytes showing *grk* transport to the anterior within 40 minutes of injection.

in 67% of cases with additional signal throughout the oocyte in 11%, and 22% showing a diffuse signal and no DA enrichment (Figure 4.6F-F'', 4.6G; Table 4.2). These results indicate that although the addition of bulge 1 and stem 2 rescues *grk* transport in most cases, some features of the proximal third of the GLS are important for the full efficiency of *grk* transport.

Despite M1-5 and M1-7 showing poor *grk* transport in the oocyte, DA enrichment was seen in a large number of cases, suggesting that anchoring is not affected in these mutants. Sequences or structures within these signals may support anchoring of any RNA that reaches the DA via diffusion or non-processive transport. Alternatively, the signal may be capable of supporting very inefficient levels of RNA transport, such that only a small subset of the RNA becomes localised. To distinguish between these possibilities, a well-defined embryo injection assay,¹¹⁶ outlined in Figure 4.7A, was used to assess recruitment of the motor component Egl to the mutants.

Following injection into syncytial blastoderm embryos, *grk* is transported to the apical cortex within 20 minutes (Figure 4.7B-B'', 4.7D-D'', n=10) and recruits Egl-GFP in 70% of embryos (Figure 4.7E-E'', n=10), but shows no bleed-through in *OrR* embryos lacking Egl-GFP (Figure 4.7C-C'', n=3). The non-localising control *hb*, on the other hand, does not enrich apically in

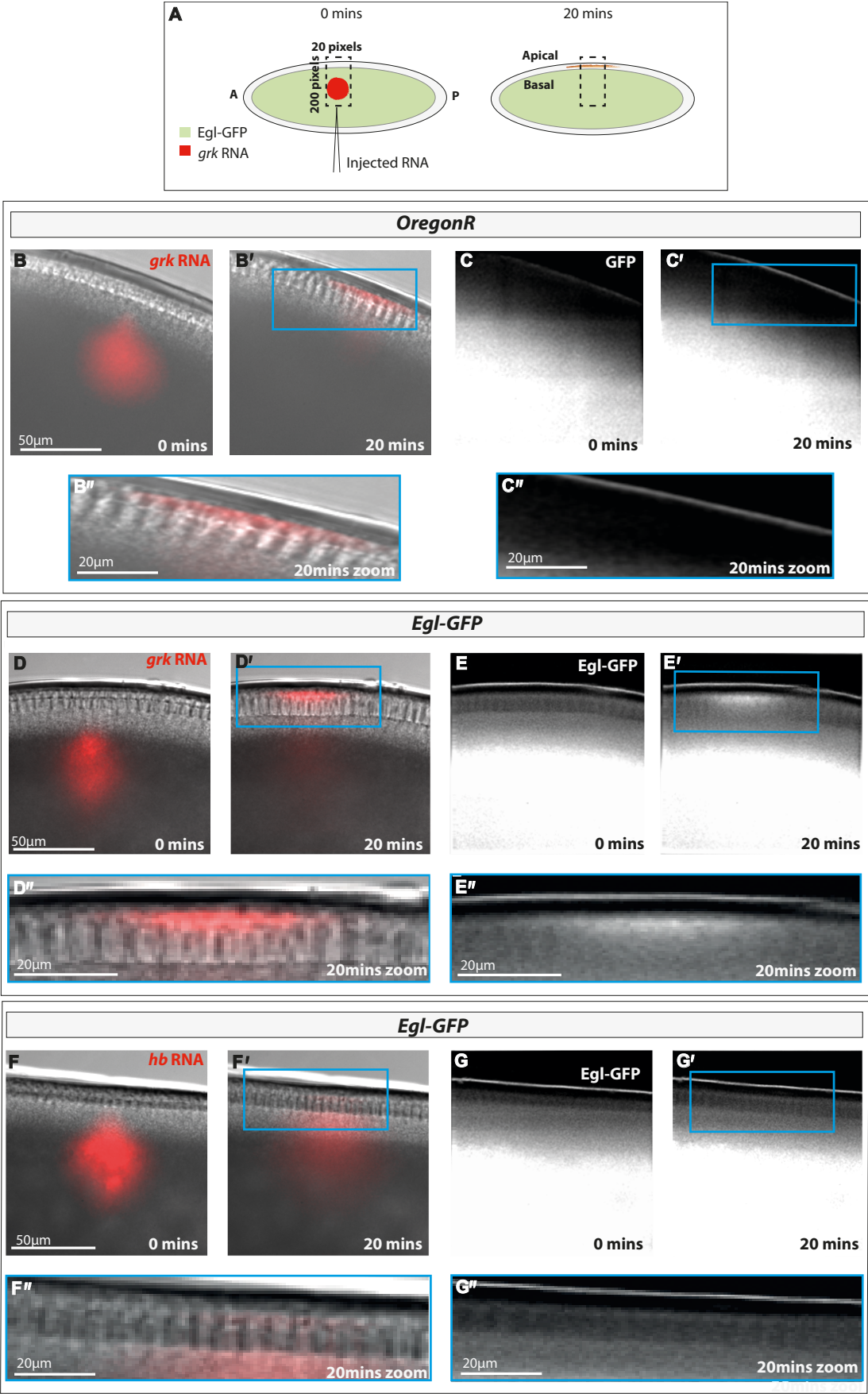


Figure 4.7. *grk* RNA localises apically in the syncytial blastoderm embryo and recruits the transport component Egl-GFP. (A) Schematic summarising the embryo recruitment assay where AlexaFluor labelled RNA (red) is injected into the basal cytoplasm of syncytial blastoderm embryos expressing GFP-tagged *trans*-acting factors, in this case Egl-GFP. Several mRNAs, including *grk*, localise to the apical cortex of the embryo following injection, and recruit members of the transport machinery, such as Egl, within 20 minutes of injection. Black box shows the cross-section of the embryo used to measure fluorescent intensities of both the RNA and *trans*-acting factors to assess localisation and recruitment in the assay to the apical cortex at 0 mins and 20 mins post-injection. (B-B'') *grk* RNA (red) injected into the wild type embryo (DIC) localises to the apical cortex within 20 minutes, but the signal does not bleed through into the green channel (C-C'', greyscale). (D-D'') *grk* RNA (red) injected into embryos (DIC) expressing Egl-GFP (greyscale) localises to the apical cortex and recruits Egl-GFP within 20 minutes (E-E''). (F-F'') *hb* RNA injected into the embryo does not localise apically and cannot recruit Egl-GFP (G-G''). (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment).

86% of embryos (Figure 4.7F-F'', n=7) and does not recruit Egl-GFP in 100% of cases (Figure 4.7G-G'', n=7).

Removal of stem 1 and internal loop 1 in M1-7 (Figure 4.8A) reduces *grk* localisation to the apical cortex, with a large proportion remaining at the injection site (Figure 4.8C'-C'', n=7) compared to wild type *grk* (Figure 4.8B, n=10). In addition, recruitment of Egl-GFP was greatly reduced in 57% of embryos injected (Figure 4.8D-D'', n=7) compared to the wild type GLS (Figure 4.8B', n=7), or completely absent in 43% of embryos (Figure 4.8D-D'', n=7). These results are reflected in the fluorescent intensity plots from a cross-section of the embryo, where colocalised peaks of *grk* RNA and Egl-GFP are present in the region of the graph representing the apical cortex at 20 minutes post-injection (Figure 4.8E). These peaks are reduced for M1-7 injections, but are completely absent in the non-localising control *hb*, suggesting that M1-7 retains some ability to be recognised by the transport machinery. Therefore, the low levels of DA enrichment seen for M1-7 and M1-5 in the oocyte may be explained by inefficient RNA transport rather than diffusion followed by anchoring.

Conclusions The presence of stem 1 and internal loop 1 is essential for maintaining efficient *grk* transport in the oocyte and embryo. Truncated forms of the GLS have inefficient *grk* transport due to reduced recruitment of Egl, a component of the transport machinery.

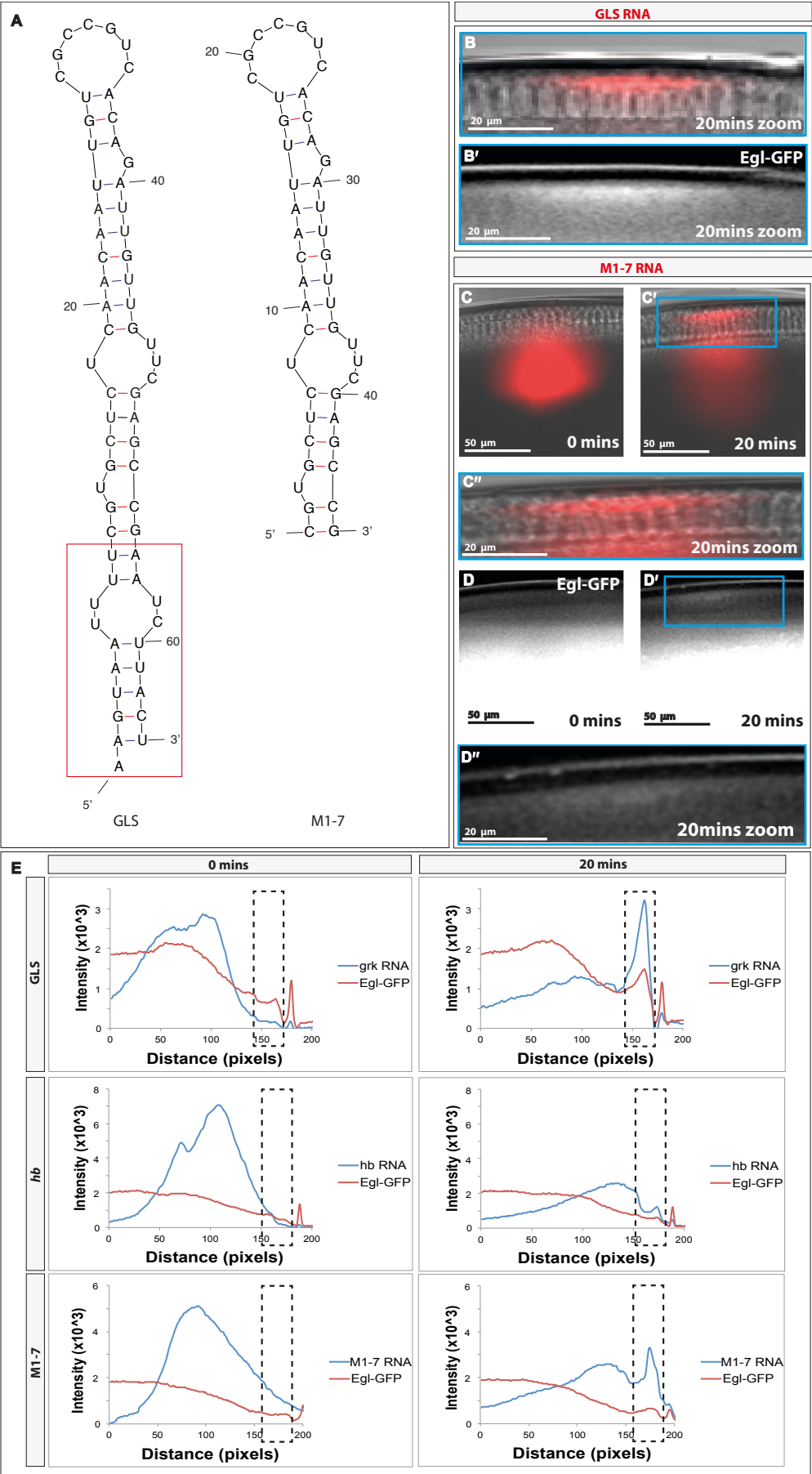


Figure 4.8. M1-7 reduces *grk* localisation in the embryo and only recruits low levels of Egl to the apical cortex. (A) UNAFold secondary structure folding shows that the remaining stem-loop structure is unaffected in the truncated M1-7 signal. (B) The GLS drives *grk* RNA (red) to the apical cortex of the embryo (DIC) and recruits Egl-GFP (greyscale) (B') within 20 minutes of injection. (C-C'') M1-7 enriches at the apical cortex but a large proportion of the signal remains at the injection site. (D-D'') M1-7 recruits much lower levels of Egl-GFP to the apical cortex. (E) Fluorescent intensity profiles show that, when compared to the non-localising control *hb*, M1-7 is able to drive low levels of transport and Egl-GFP recruitment to the apical cortex (black box) in the embryo. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (E) refers to the position through the cross-section of the embryo (see Figure 4.8) from the site of injection to the apical cortex (black box)).

GLS A'-form RNA conformation is unlikely to be required for *grk* transport

Egl has been shown to link the *K10* localisation signal to the dynein motor by recognising A'-form helices within the mRNA. These helices have a widened major groove providing accessibility for the protein and are formed by stacking interactions between three or more purines on the same side of a stem. Two such stems (stem 2 and stem 3) are predicted to form in the GLS and thus may be responsible for creating a recognition site for Egl (Figure 4.9A).³⁸⁶

In order to test whether the two putative A'-form helices generate a recognition site for Egl, mutations were designed in the purine stacks. M10-9 contains mutations in stem 2 (9-11UUC>GGG, 54-56GAA>CCC) that changes the sequence of the purine stack and switches it to the opposite side of the stem in an attempt to abolish A'-form RNA that may be formed by the two purine stacks (Figure 4.9A). Both the wild type GLS (Figure 4.9B-B'', 4.9F, n=54) and M10-9 (Figure 4.9C-C'', 4.9F, n=32) are able to direct *grk* to the anterior of the oocyte in 100% and 94% of injected oocytes, respectively, before it becomes enriched at the DA by 40 minutes post-injection. This suggests that two purine stacks on the same side of a stem are not required for recognition of the GLS by Egl.

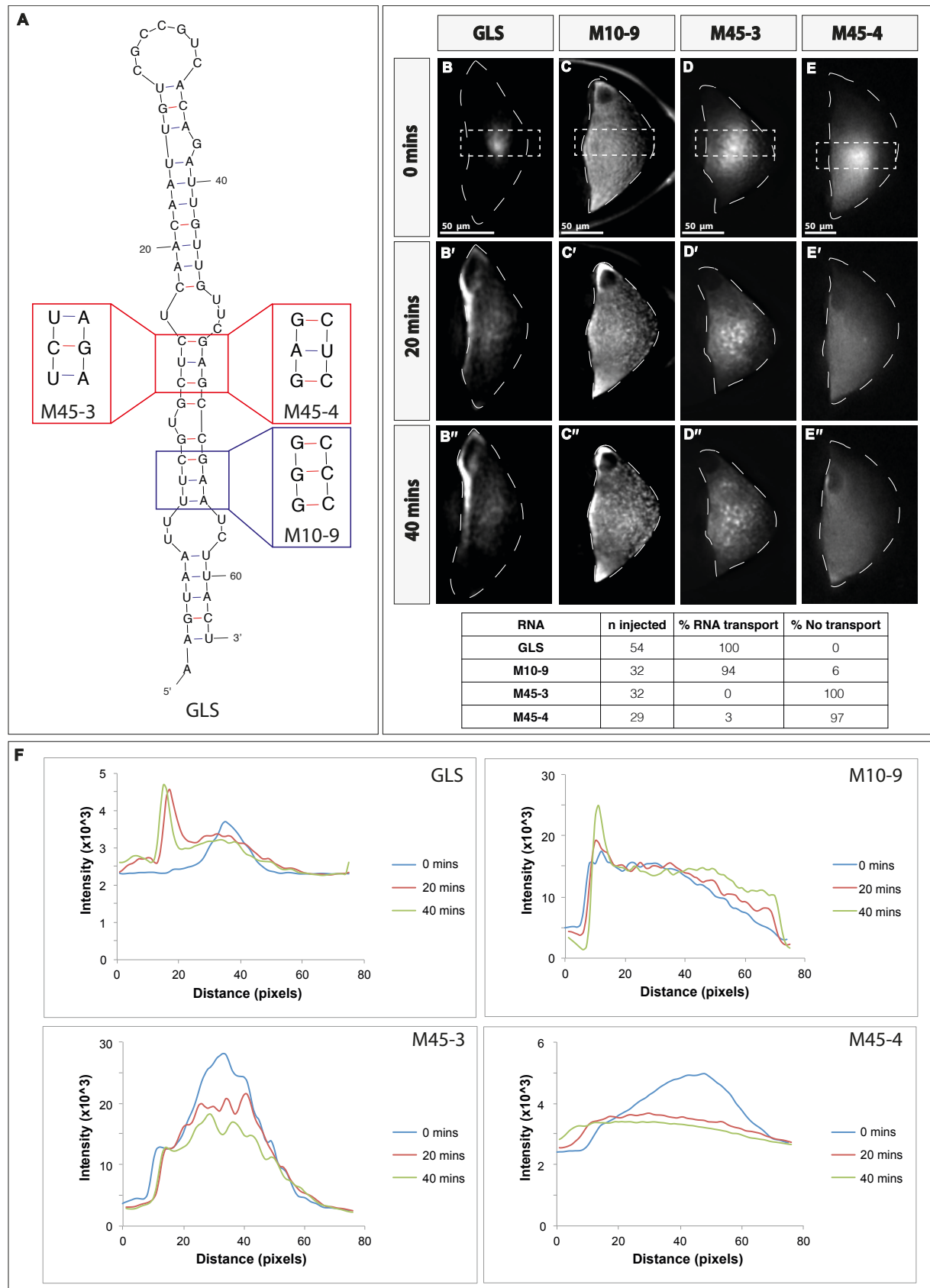


Figure 4.9. The sequence and orientation of the purine stack within stem 3 is essential for *grk* transport in the oocyte. (A) UNAFold predicted secondary structures of the GLS and mutants in stem 2 and stem 3. (B-B'') Wild type *grk* RNA is transported to the anterior of the oocyte after 20 minutes following injection and enriches at the DA within 40 minutes. (C-C'') Mutating stem 2 in M10-9 has no effect on the transport of the RNA. (D-D'') Changing the sequence of the purine stack in M45-3 within stem 3 disrupts *grk* transport. (E-E'') Switching the orientation of the purine stack in stem 3 disrupts *grk* transport resulting in a diffuse signal throughout the oocyte after 40 minutes. (F) Fluorescent intensities profiles through a cross-section of the oocyte (white boxes) show that the GLS and M10-9 enrich at the anterior within 20 minutes of injection, whereas M45-3 and M45-4 are diffuse throughout the cytoplasm. N-numbers and the percentage of oocytes showing RNA transport are represented in the table. Oocytes outlined using white dashed lines.

Altering the sequence of the second purine stack in stem 3 in a way that maintains a purine stack in M45-3 (15-17CUC>UCU, 49-51GAG>AGA; Figure 4.9A) causes the RNA to form a diffuse signal throughout the oocyte within 20 minutes of injection (Figure 4.9D-D'', 4.9F, n=32). Maintaining the sequence of stem 3 but switching the purine stack to the opposite side of the stem in M45-4 (Figure 4.9A), abolishing the potential for A'-form RNA, again results in a diffuse signal throughout the oocyte in 97% of cases within 20 minutes of injection (Figure 4.9E-E'', 4.9F, n=29). Taken together, these results suggest that the purine stacks in stem 2 and stem 3 are not required on the same side of the stem, and that both the sequence and orientation of stem 3 are essential for *grk* transport.

To test whether the failure of M45-3 and M45-4 to localise in the oocyte is due to an inability to interact with the transport machinery, recruitment of Egl to these mutants was assessed in the embryo. In contrast to the apical localisation and Egl-GFP enrichment seen for wild type *grk* RNA (Figure 4.10A-A', 4.10G), n=10), M45-3 shows poor transport to the apical cortex with the majority of the RNA remaining in the basal cytoplasm in 100% of embryos (Figure 4.10C-C'', 4.10G, n=9) in a similar manner to the non-localising control *hb* (Figure 4.10B, 4.10G, n=7). In addition, both *hb* (Figure 4.10B', 4.10G, n=7) and M45-3 (Figure 4.10D-D'', 4.10G, n=9) show no enrichment of Egl-GFP to the apical cortex in 100% of embryos. M45-4, on the other hand, drives apical accumulation of a slightly larger proportion of the injected RNA than M45-3, but much of the RNA still remains at the site of injection (Figure 4.10E-E'', 4.10G, n=7). This is reflected in a weak recruitment of Egl to the apical cortex of the embryo (Figure 4.10F-F'', 4.10G), suggesting that maintaining a purine stack but switching its

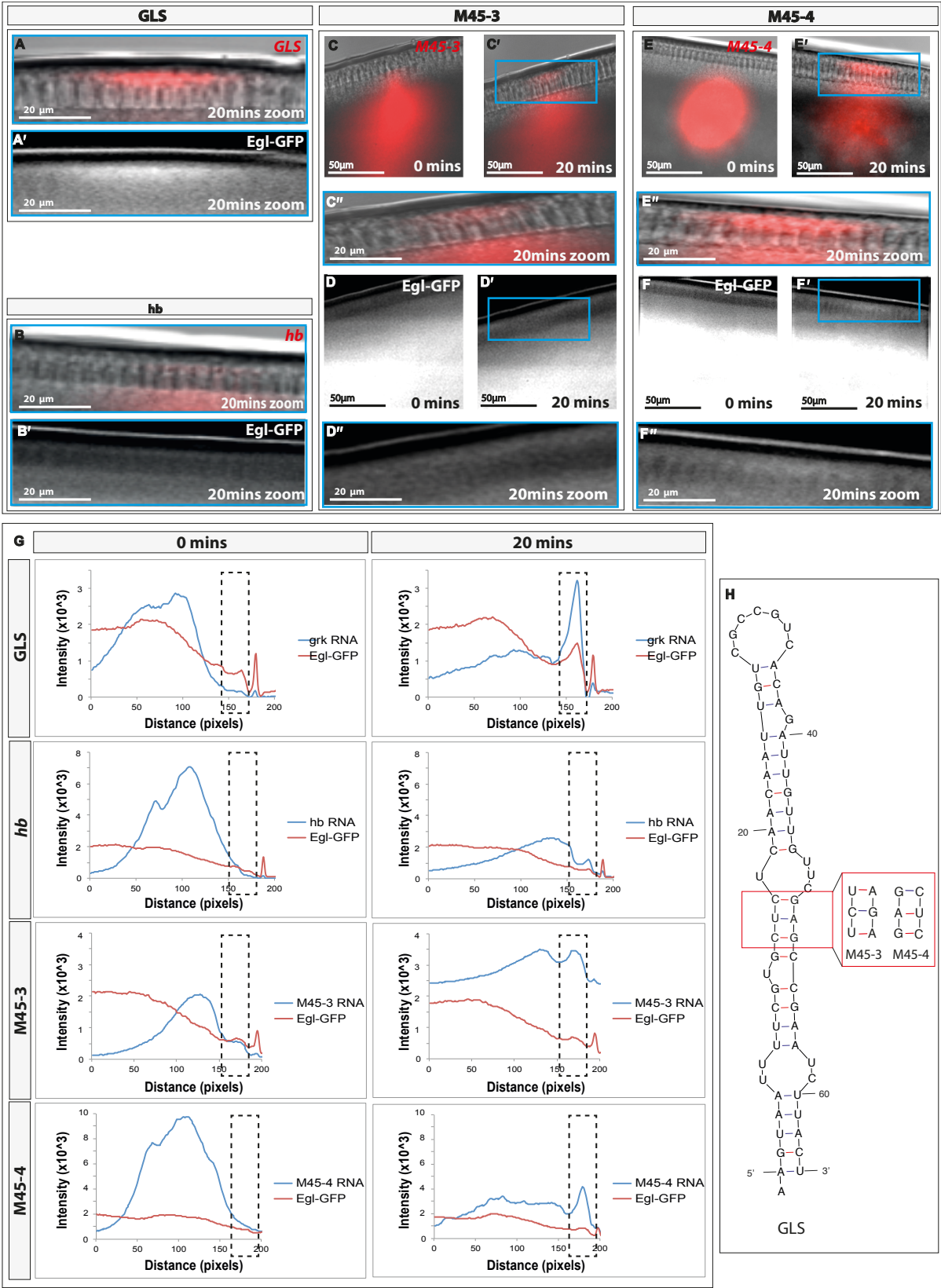


Figure 4.10. Changing the sequence of the purine stack in stem 3 disrupts transport of *grk* RNA due to an inability to recruit Egl. (A-A') Wild type *grk* RNA (red) localises apically in the embryo (DIC) and recruits Egl-GFP (greyscale) but a non-localising control *hb* does not (B-B'). (C-C'') M45-3 is not transported to the apical cortex of the embryo and does not recruit Egl-GFP (D-D'') whereas M45-4 shows low levels of transport (E-E'') and recruitment (F-F''). (G) Fluorescent intensity profiles through a cross-section of the embryo show that wild type *grk* strongly accumulates at the apical cortex and recruits Egl-GFP, whereas *hb* and M45-3 do not. M45-4 is transported weakly in the embryo and shows some Egl-GFP recruitment to the apical cortex. (H) UNAFold predicted secondary structure of the GLS and the mutants M45-3 and M45-4. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (G) refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

orientation has less of an effect on *grk* transport and recruitment of the transport machinery than changing the sequence of the stem in M45-3 (Figure 4.10H).

Taken together, the data from these mutants suggest that the presence of two purine stacks in the GLS is not essential for recognition by Egl. However, the sequence and orientation of the purine stack in stem 3 is essential for Egl recruitment and thus *grk* transport. As a result, it is possible that Egl doesn't recognise the GLS via A'-form RNA, although we can not rule out the presence of such structures within the mutant signals without structural analysis.

Conclusions A'-conformation predicted to form from two purine stacks is not required for recognition of the GLS by Egl. The purine stack within stem 2 is redundant for transport, but both the sequence and the orientation of the purine stack in stem 3 is essential.

The length of stem 4 is essential to maintain *grk* transport

The GLS-like signals identified in the *Drosophilids* and other Dipterans showed that the region of highest conservation is the AU-rich distal stem 4, indicating that this region could be important for the function of the GLS. In order to test this, I designed several mutants to truncate stem 4. Removing five base-pairs from the distal stem in M10-3 is predicted to have no effect on the secondary structure of the remaining stem-loop (Figure 4.11A). Injection of

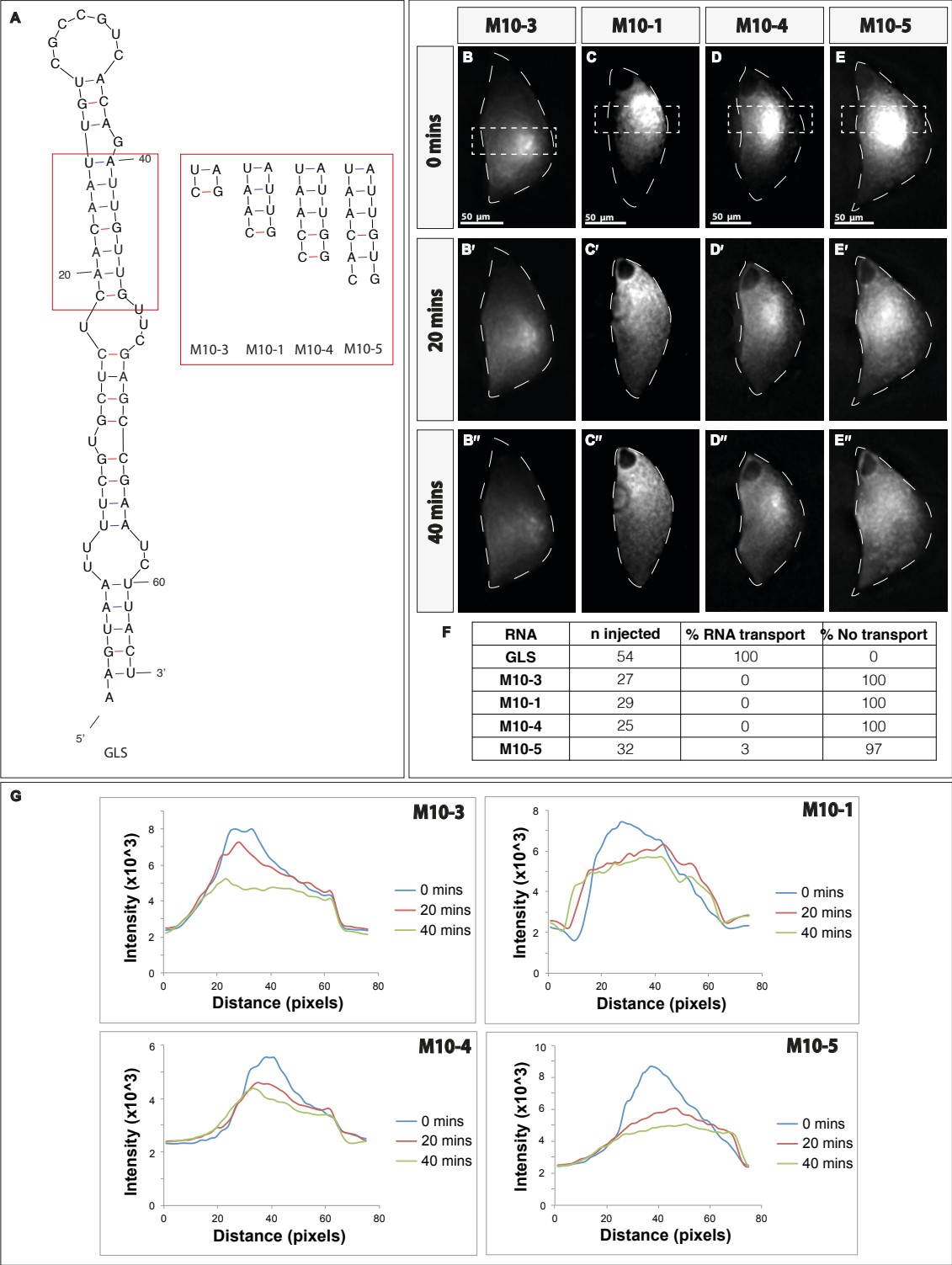


Figure 4.11. The length of distal stem 4 is essential for *grk* transport in the oocyte. (A) UNAFold secondary structure prediction of the GLS and the distal stem truncated mutants. M10-3 (B-B''), M10-1 (C-C''), M10-4 (D-D''), and M10-5 (E-E'') are unable to drive *grk* transport in the oocyte. (F) The percentage of oocytes injected with each mutant that show transport of the RNA to the anterior of the oocyte within 40 minutes of injection. (G) Fluorescent intensity profiles through a cross-section of the oocyte (white boxes) show that none of the truncated distal stem mutants can drive *grk* transport to the anterior of the oocyte. Oocytes outlined with dashed white lines.

M10-3 into the oocyte shows that the signal is unable to drive the transport of *grk* RNA to the anterior of 100% of the oocytes injected within 40 minutes (Figure 4.11B-B", 4.11F-G). The same can be seen for M10-1, M10-4 and M10-5 in which three, two and one base-pairs have been removed from the stem, respectively (Figure 4.11C-E", 4.11F-G). These results indicate that the full-length stem 4 is an absolute requirement for *grk* transport within the oocyte.

To test whether increasing the length of the stem has an effect on *grk* transport, I designed three mutants in which the length of the stem is either increased by three base-pairs in M10-6, or one base-pair in M10-7 and M10-8. In each case, increasing the length of the stem is predicted to have little effect on the structure of the rest of the stem-loop (Figure 4.12A). When compared to wild type injected *grk* RNA (Figure 4.12B-B", 4.12F, 4.12G), M10-6 shows poor *grk* transport with a diffuse distribution throughout the cytoplasm in 53% of oocytes injected (Figure 4.12C-C", 4.12F, 4.12G). However, 47% of oocytes injected with M10-6 show signs of RNA enrichment at the DA of the oocyte (white arrow in Figure 4.12C-C", 4.12F, 4.12G), with the remainder of the RNA diffuse throughout the cytoplasm. As a result, M10-6 either drives transport of a small proportion of RNA to the anterior, or the RNA diffuses there and is processed by the anchoring machinery. M10-7 and M10-8, on the other hand, are able to direct *grk* transport to the anterior of 96% and 86% of oocytes injected, respectively (Figure 4.12D-G), albeit less efficiently than wild type *grk* with more RNA remaining diffuse or in aggregates throughout the cytoplasm. As a result, lengthening the distal stem by only a single base-pair only slightly reduces *grk* localisation efficiency in the oocyte, whereas lengthening the stem by three base-pairs severely reduces *grk* transport.

To determine whether M10-4 and M10-6 (Figure 4.13A) can recruit the transport machinery, both signals were tested for Egl recruitment in the embryo and compared to wild type *grk* (Figure 4.13B-B', n=10) and *hb* (Figure 4.13C-C', n=7). M10-4 is unable to localise to the apical cortex of the embryo in 60% of cases, and only very weakly in the remaining 40% of embryos (Figure 4.13D-D", n=10). Furthermore, M10-4 does not recruit Egl-GFP in 90% of

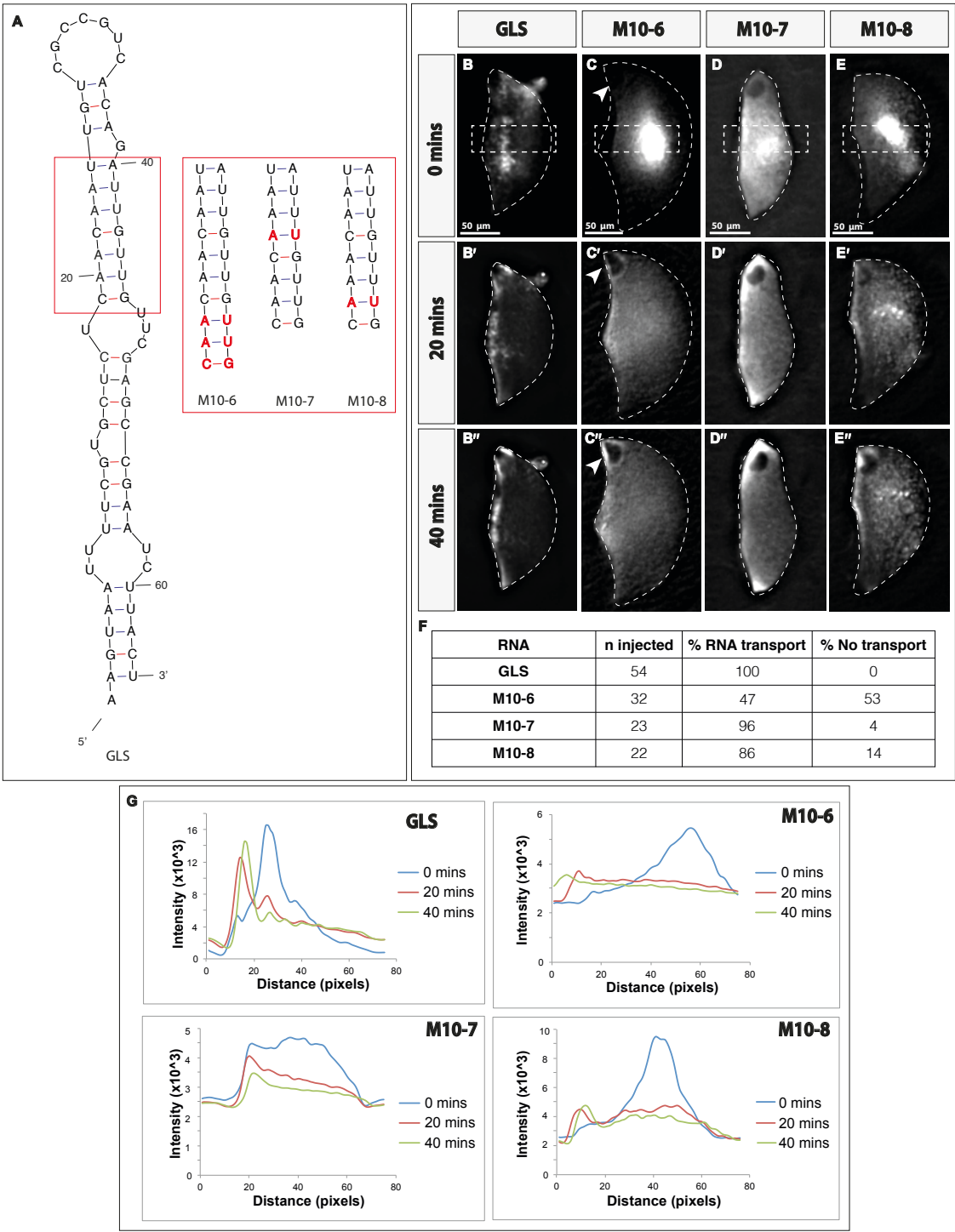


Figure 4.12. Increasing the length of distal stem 4 reduces *grk* transport in the oocyte. (A) UNAFold secondary structure prediction of the GLS and the distal-stem elongation mutants. (B-B'') Wild type *grk* RNA is transported to the anterior of the oocyte after 20 minutes following injection and enriches at the DA within 40 minutes. M10-6 (C-C''), M10-7 (D-D''), and M10-8 (E-E'') drive *grk* transport in the oocyte with reduced efficiency. (F) The percentage of oocytes injected with each mutant that show transport of the RNA to the anterior of the oocyte within 40 minutes of injection. (G) Fluorescent intensity profiles through a cross-section of the oocyte (white dashed rectangles) show that the distal stem 4 mutants can drive reduced levels of *grk* transport to the anterior of the oocyte. Oocyte outlined with white dashed lines.

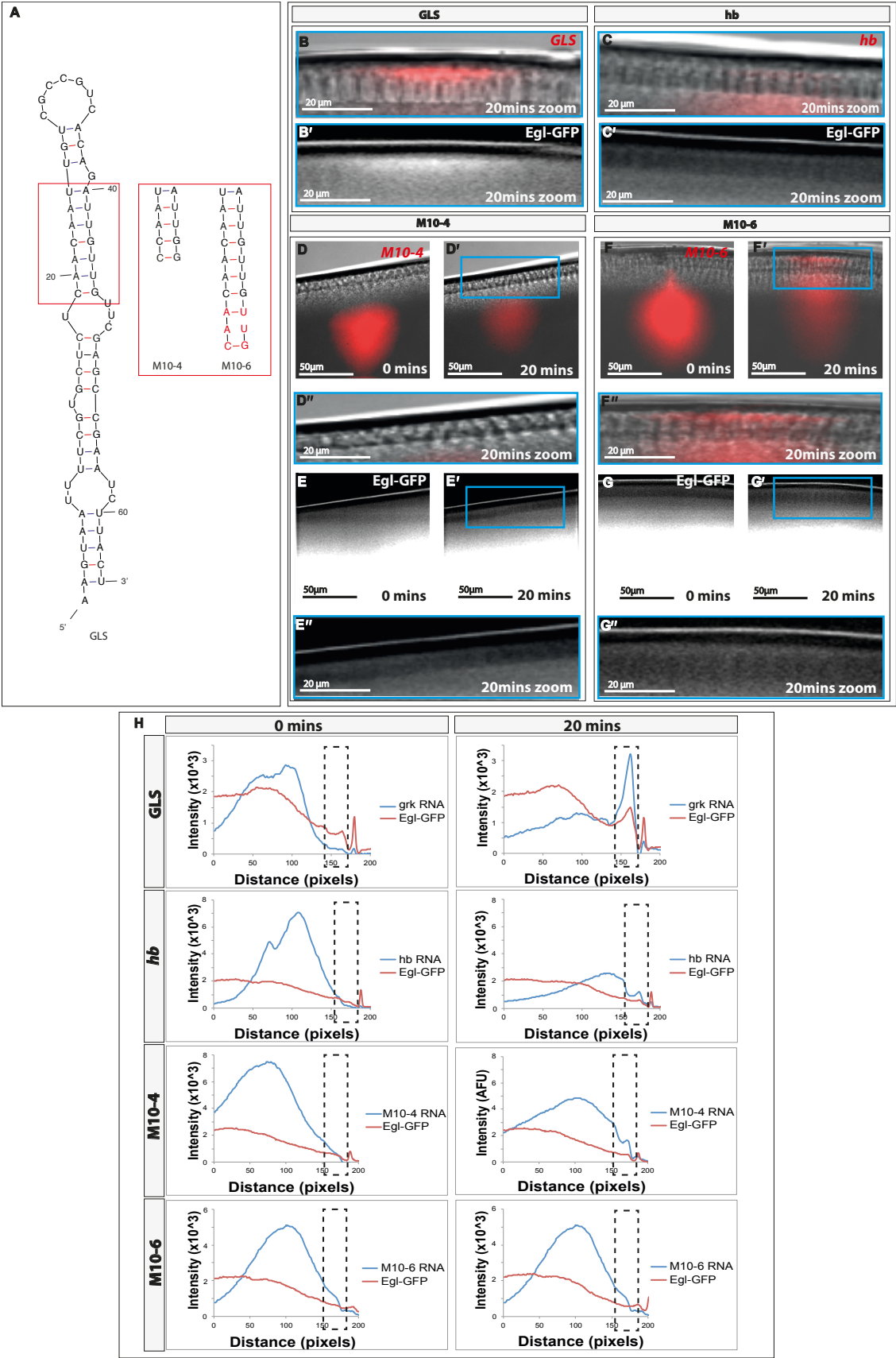


Figure 4.13. Changing the length of distal stem 4 greatly reduces *grk* transport in the embryo due to an inability to recruit the transport machinery component Egl. (A) UNAFold predicted secondary structure of the GLS and mutants M10-4 and M10-6. (B-B') The GLS drives *grk* RNA to the apical cortex of the embryo and recruits Egl-GFP. (C-C') The non-localising control *hb* does not localise in the embryo and cannot recruit Egl-GFP. M10-4 RNA is unable to localise in the embryo (D-D'') and does not recruit Egl-GFP (E-E''). M10-6 localises weakly to the apical cortex of the embryo (F-F'') and recruits low levels of Egl-GFP (G-G''). (H) Fluorescent line profiles through a cross-section of the embryo show that the GLS strongly localises and recruits Egl-GFP to the apical cortex (black box) whereas *hb*, M10-4 and M10-6 do not. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (H) refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

embryos injected (Figure 4.13E-E'', n=10). M10-6 shows weak localisation in 67% of embryos injected, with the majority of the RNA remaining in the basal cytoplasm, and the remaining 33% of injections show diffuse RNA at the site of injection (Figure 4.13F-F'', n=9). Weak levels of M10-6 localisation is associated with weak recruitment of Egl-GFP in 67% of cases, and no recruitment in the remaining 33% of embryos (Figure 4.13G-G'', n=9). This is reflected in the fluorescent intensity profiles taken through a cross-section of the injected embryos showing that the majority of M10-4 and M10-6 RNA remains at the site of injection throughout the assay in a similar way to *hb* (Figure 4.13H). The fact that M10-6 can drive low levels of *grk* transport could explain why a high percentage of oocyte injections resulted in enrichment and anchoring of a small proportion of the RNA at the DA.

Taken together, these data suggest that the length of the distal stem is essential for maintaining the function of the GLS in driving *grk* transport due to its ability to recruit the transport machinery. However, it is unclear whether it is the distances created by the stem, or sequences within the stem that are essential for maintaining RNA transport.

Conclusions Truncations within the distal stem abolish RNA localisation in the oocyte and greatly reduce transport in the embryo due to inefficient recruitment of Egl. Increasing the length of the distal stem by a single base-pair reduces transport efficiency of the RNA in the oocyte, and lengthening the stem by three base-pairs enhances this reduction in RNA transport efficiency in the oocyte and the embryo due to less efficient recruitment of the

transport machinery. Features within the distal stem such as the length, the sequence, or distances created by the stem, are therefore important for *grk* transport in the oocyte and the embryo.

The AU-richness of stem 4 is essential for *grk* transport

In addition to the length of the distal stem being essential for *grk* transport, the sequence or the structure may also be important. An outstanding feature of this stem is its AU-richness. In order to test whether the sequence or the AU-content of the stem is important for *grk* transport, I made transversion mutations in stem 4 to generate a GC-rich stem in M42-7 (20-21AA>GG, 23-24AA>GG, 41-42UU>CC, 44-45UU>CC), or reflected the stem to maintain the AU-content in M42-8 (20-24AACAA>UUGUU). In both cases prediction of the secondary structures suggests that the rest of the signal is unaffected by the mutations (Figure 4.14A). M42-7 is unable to localise in 100% of the oocytes injected and is diffuse throughout the cytoplasm (Figure 4.14C-C'', 4.14E, n=21) instead of being transported to the anterior of the oocyte as seen for wild type *grk* (Figure 4.14 (B-B'', E), n=54). M42-8, on the other hand, drives *grk* RNA to the anterior of the oocyte in 100% of cases (Figure 4.14D-D'', 4.14E, n=34). The anterior enrichment of the GLS and M42-8 can be seen in the fluorescent intensity profiles generated from a cross-section through the oocyte (Figure 4.14F). These data suggest that the sequence of the distal stem is not important for *grk* transport, but that the overall AU-richness of this region is essential.

To investigate whether the lack of M42-7 (Figure 4.15A) localisation in the oocyte is due to an inability to recruit the transport machinery, it was assessed for Egl-GFP recruitment in the embryo. Like *hb* RNA (Figure 4.15C-C', n=7), M42-7 is unable to localise in 100% of embryos injected (Figure 4.15D-D'', n=9) and does not recruit Egl-GFP to the apical cortex (Figure 4.15E-E'', n=9). M42-8, on the other hand, is able to drive efficient localisation of *grk* RNA and shows strong recruitment of Egl-GFP in 100% of embryos injected (Figure 4.15F-G'', n=7) in a similar manner to the wild type GLS (Figure 4.15B-B', n=10). These localisation and

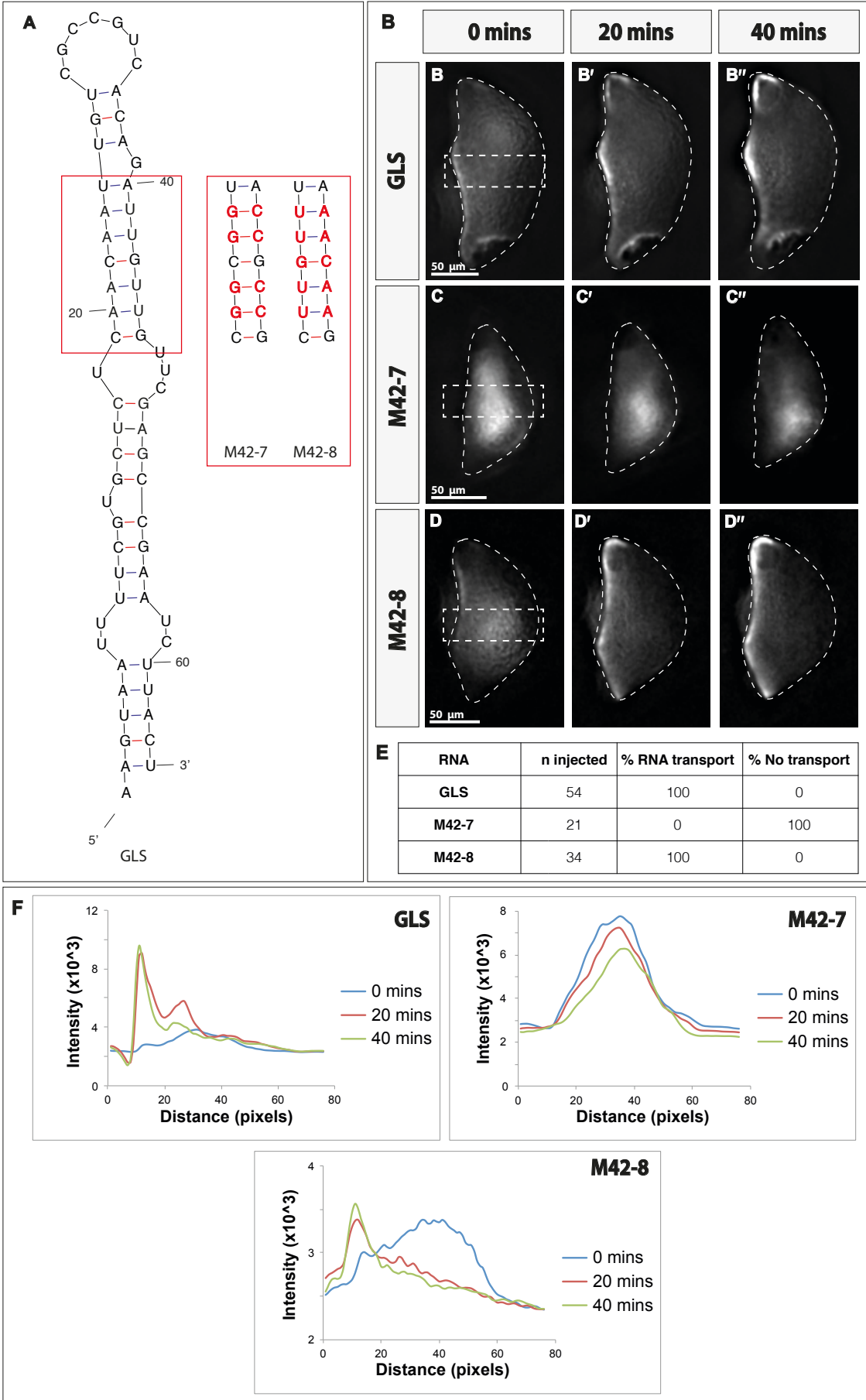


Figure 4.14. The AU-content of the distal stem is essential for driving *grk* transport in the oocyte. (A) UNAFold predicted secondary structures of the GLS and the mutants M42-7 and M42-8. (B-B'') The wild type GLS drives *grk* to the anterior of the oocyte. (C-C'') Injected M42-7 is unable to localise within the oocyte. (D-D'') M42-8 drives efficient *grk* transport to the anterior of the oocyte. (E) Percentages of oocytes showing efficient *grk* transport within 40 minutes of injection. (F) Fluorescent intensity profiles of the GLS and mutants M42-7 and M42-8 showing that M42-8 is not transported in the oocyte. Distance refers to the position through the cross-section of the oocyte (white boxes) from the posterior to the anterior.

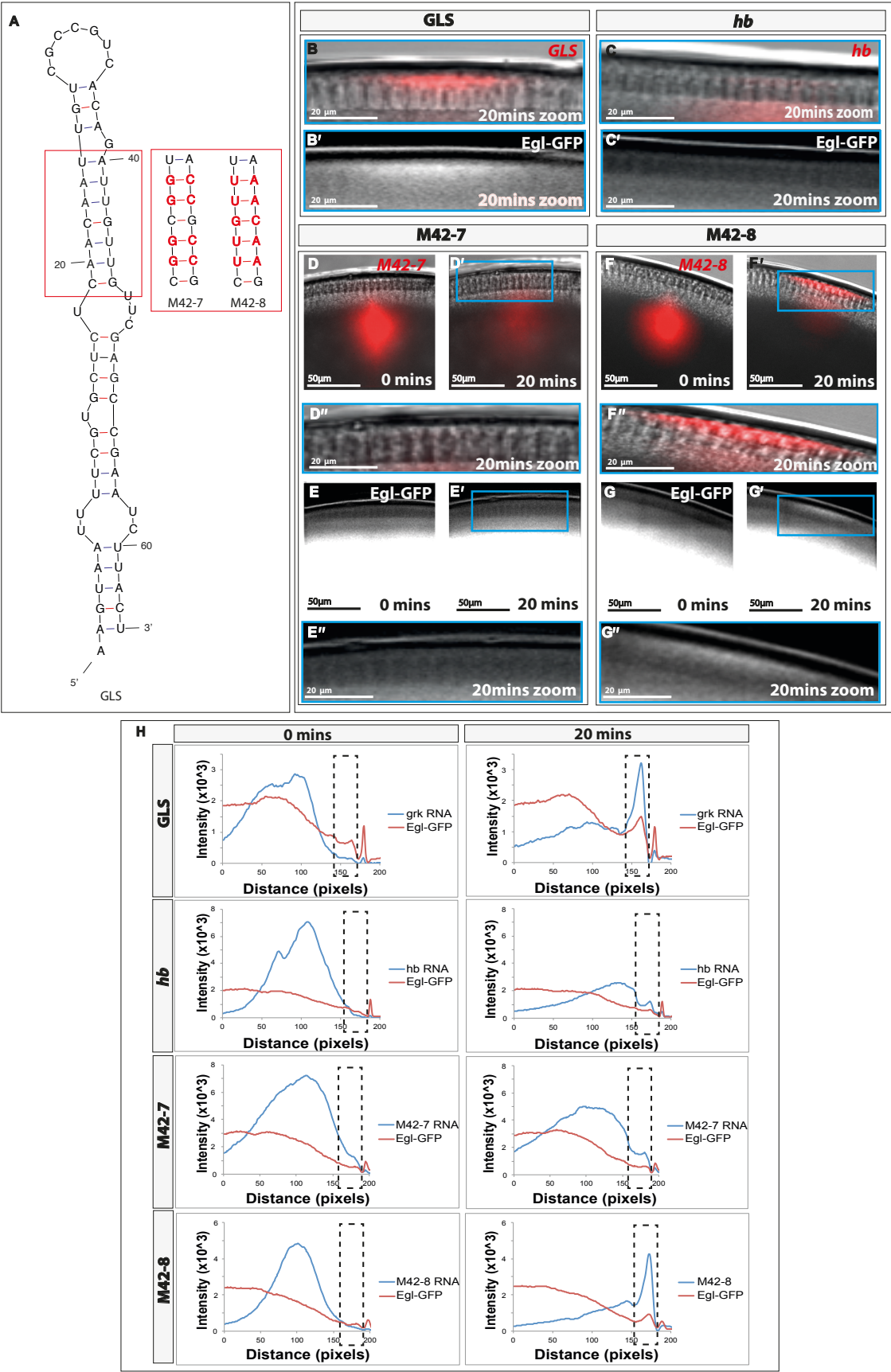


Figure 4.15. A GC-rich distal stem fails to drive *grk* localisation in the embryo due to an inability to recruit the transport machinery component Egl. (A) UNAFold predicted secondary structures of the GLS, M42-7 and M42-8. (B-B') The wild type GLS drives efficient transport of *grk* RNA (red) to the apical cortex of the embryo (DIC) and recruitment of Egl-GFP (greyscale). (C-C') *hb* cannot localise in the embryo and does not recruit Egl-GFP. (D-D'') The GC-rich stem in M42-7 is unable to transport *grk* RNA in the embryo and cannot recruit Egl-GFP (E-E''). The reflected AU-stem in M42-8 maintains *grk* localisation in the embryo (F-F'') and efficiently recruits Egl-GFP (G-G''). (H) Fluorescent intensity profiles taken through the centre of the embryo show that the GLS and M42-8 efficiently transport to the apical cortex (black box) and recruit Egl-GFP, whereas *hb* and M42-7 do not. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (H) refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

recruitment patterns are reflected in the fluorescent intensity profiles generated from a cross-section through the embryo (Figure 4.15H). Taken together, these results suggest that the AU-richness or the stability of the distal stem is essential for Egl recruitment and thus *grk* transport, whereas the sequence is not.

Conclusions The AU-content of distal stem 4 is essential for driving *grk* transport.

The sequence of stem 4 is not essential for driving *grk* transport but can increase efficiency

As reflecting the whole distal stem in M42-8 retains a large section of the sequence, it is still possible that sequences within the stem are required for *grk* transport. Alternatively, the stability of the stem may be important. To distinguish between these possibilities, a range of sequence modifications were designed within the distal stem and the mutants were tested for their ability to drive *grk* transport to the anterior of the oocyte. For full details of all the nucleotides mutated, their predicted secondary structures, and the overall effect on *grk* localisation see Figure 4.16 and Table 4.3.

Several of the mutants showed similar localisation to wild type *grk* RNA, being transported to the anterior of the oocyte within 20 minutes of injection, and enriching at the DA by 40 minutes (Figure 4.17A-A''; Table 4.3). For example, mutations to the central CG base-pair in

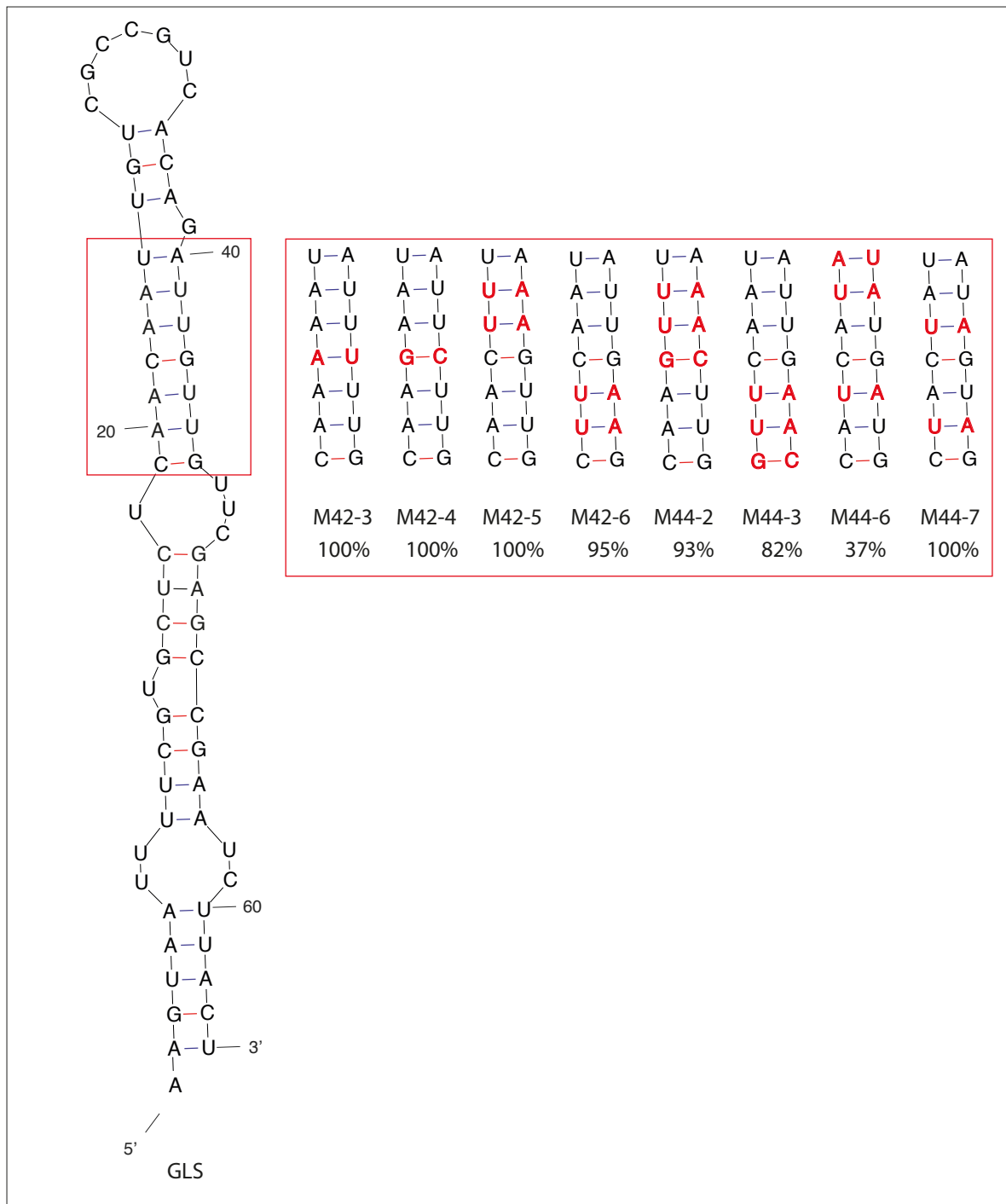


Figure 4.16. Sequence modifications designed to test whether the sequence of distal stem 4 is important for *grk* localisation. (A) UNAFold predicted secondary structures of the GLS and several mutants in the distal stem (red box). The mutations (highlighted in red) are predicted to have little effect on the secondary structure of the rest of the stem-loop. The ability of each mutant to drive *grk* transport is represented as the percentage of oocytes injected showing transport to the anterior within 40 minutes.

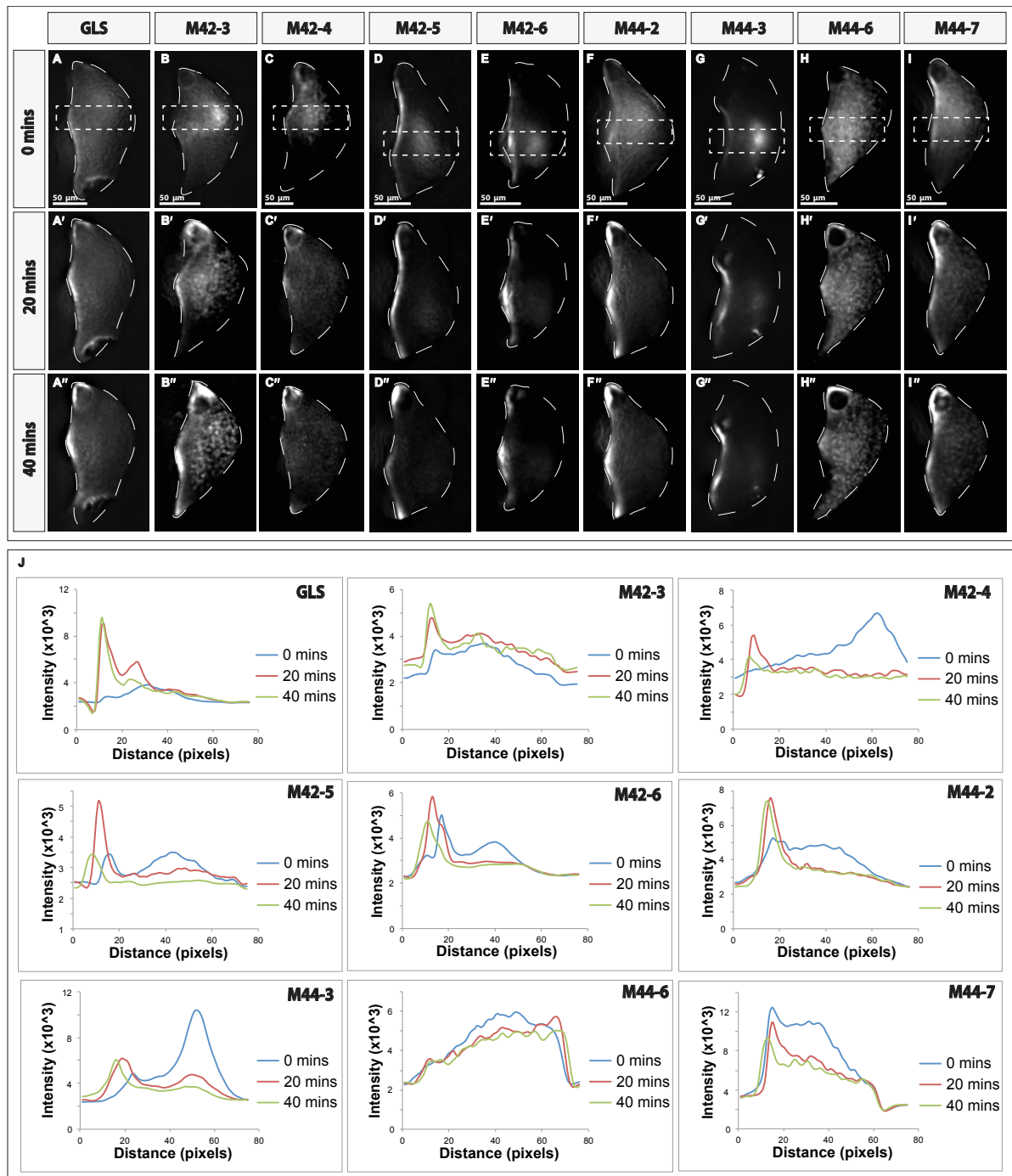


Figure 4.17. Sequence modifications within distal stem 4 have little effect on *grk* transport in the oocyte, but mutation of the distal U-A base pair greatly reduces *grk* transport. (A-A'') The wild type GLS drives efficient transport of *grk* RNA to the anterior of the oocyte within 20 minutes of injection, before it enriches at the DA by 40 minutes. M42-3 (B-B''), M42-4 (C-C''), M42-5 (D-D''), M42-6 (E-E''), M44-2 (F-F'') and M44-3 (G-G'') support efficient transport of *grk* RNA in the oocyte. M44-6 (H-H'') shows a vastly reduced function in driving *grk* localisation with much of the RNA remaining diffuse throughout the oocyte, likely due to disruption of the U-A base pair above distal stem 4. (I-I'') M44-7 efficiently localises *grk* in the oocyte. (J) Fluorescent intensity profiles of each mutant taken from a cross-section through the middle of the oocyte. Distance refers to the position through the cross-section of the oocyte (white dashed rectangles) from the anterior to the posterior.

RNA	Positions of sequence modifications within the GLS	n injected	% RNA transport	% No transport
GLS	N/A	54	100	0
M42-3	22C>A, 43G>U	27	100	0
M42-4	22C>G, 43G>C	19	100	0
M42-5	23-24AA>UU; 41-42UU>AA	18	100	0
M42-6	20-21AA>UU; 44-45UU>AA	20	95	5
M44-2	22-24CAA>GUU; 41-43UUG>AAC	28	93	7
M44-3	19-21CAA>GUU; 44-46UUG>AAC	28	82	18
M44-6	21A>U; 44U>A; 24A>U; 41U>A; 25U>A; 40A>U	30	37	63
M44-7	20A>U; 45A>U; 23A>U; 41U>A	25	100	0

Table 4.3. Specific sequence modifications in distal stem 4 reduce *grk* transport in the oocyte. Outline of the sequence modifications made to the GLS in each mutant designed to test whether sequences within distal stem 4 are required for *grk* transport in the oocyte. Changes to a nucleotide at the given position are marked by (>). The percentage of oocytes showing *grk* RNA transport to the anterior of the oocyte within 40 minutes of injecting each mutant are shown, along with the percentage that fail to show *grk* transport. N-numbers for each mutant are displayed.

M42-3 and M42-4 had no effect on *grk* localisation (Figure 4.16; Table 4.3; Figure 4.17B-B", 4.17C-C", 4.17J), indicating that it is not important for contributing to the sequence of the stem, nor to the GC-content.

The distal stem contains two repeats of the sequence motif CAA (or UUG on the opposite side of the stem). To investigate the importance of this motif, the mutants M42-5 and M42-6 were generated where one of the two CAA motifs are altered to CUU in each case (Figure 4.16). Both mutants are able to drive similar levels of *grk* transport to the wild type GLS (Table 4.3; Figure 4.17D-D", 4.17E-E", 4.17J), indicating that the orientation of the AU base-pairs, and therefore the second and third positions of the CAA motif, are not important for *grk* transport. Reflection of all three base-pairs in one of the two CAA motifs in M44-2 and M44-3 (Figure 4.16) has little effect on *grk* transport with M44-3 showing only a slight reduction in

the percentage of oocytes with RNA transport to the anterior of the oocyte (Table 4.3; Figure 4.17F-F", 4.17G-G", 4.17J). As M42-6 contains two of the same base-pair reflections as M44-3 but maintains wild type levels of *grk* transport, it is possible that the evolutionarily conserved CG base-pair formed from positions 19 and 46 is important for *grk* transport, whereas the CAA motifs are not. Alternatively, the repeated CAA sequence may provide two motifs that are recognised by the transport machinery, resulting in some redundancy.

To test whether the two CAA sequences act together to drive efficient *grk* transport in the oocyte, reflections of distant base-pairs within each of the CAA motifs were tested in mutants M44-6 and M44-7 (Figure 4.16) to disrupt both motifs. The sequence modifications in M44-6 cause a drastic reduction in transport, with *grk* being transported to the anterior of only 37% of the oocytes injected, and 43% showing signs of enrichment at the anterior or DA, but with a high signal of diffuse RNA throughout the cytoplasm (Table 4.3; Figure 4.17H-H", 4.17J). The remaining 20% of oocytes showed a diffuse signal in the cytoplasm with no enrichment at the anterior or DA within 40 minutes of injection. These data indicate that transport of M44-6 is greatly reduced, resulting in only a small proportion of the RNA being able to anchor at the DA. Sequence modifications in M44-7, on the other hand, have no effect on *grk* transport (Table 4.3; Figure 4.17I-I", 4.17J). The localisation pattern of all of the distal stem sequence mutations are summarised using fluorescent intensity profiles from a cross-section through the middle of the oocyte and clearly show that all of the mutants, except for M44-6, drive efficient localisation of *grk* RNA to the anterior of the oocyte (Figure 4.17J).

Taken together these results show that, in general, changes to the sequence of the distal stem do not abolish *grk* transport in the oocyte, suggesting that the AU-content is more important than the sequence. However, sequence changes can reduce the percentage of oocytes showing *grk* transport, suggesting that it may also play a role in the initial recognition of the RNA by the transport machinery. Combinatorial mutations to the third position in both of the repeated CAA motifs seem to have the most drastic effect on *grk* transport, as seen in M44-6. However, mutating the entire CAA motif has little effect suggesting that it may instead

be the UA base-pair formed from nucleotides 25 and 40, and mutated in M44-6, that is important for *grk* transport. In support of this hypothesis, the UA base-pair in question is entirely conserved among the GLS-like signals identified in other species (Figure 4.5A). However, further analysis of this base-pair will be required to distinguish between these two possibilities.

Conclusions The sequence of the entire distal stem 4 is not essential for driving *grk* transport. The UA base-pair directly above the distal stem is likely to be important for *grk* transport.

The location and accessibility of the distal stem is essential to maintain *grk* transport

Although I have shown that the sequence of the distal stem is not essential for driving *grk* transport in the oocyte, the stability or accessibility of the stem may be essential for binding of the transport machinery. In order to test this, I designed the mutant M50-2 that reduces the accessibility of the distal stem by introducing strong GC base-pairs on either side of bulge 2 (Figure 4.18A). The mutations in M50-2 abolish *grk* transport in 100% of oocytes with the RNA showing a diffuse distribution throughout the oocyte (Figure 4.18C-C'', n=20), as opposed to the anterior localisation seen within 20 minutes for the wild type GLS (Figure 4.18B-B'', n=54). Furthermore, injection of M50-2 into embryos expressing Egl-GFP shows that, unlike wild type *grk* RNA which localises apically (Figure 4.18D-D'', n=10) and recruits Egl-GFP (Figure 4.18E-E'', 4.18H n=10), M50-2 does not accumulate apically and largely diffuses from the site of injection (Figure 4.18F-F'', n=7), resulting in no recruitment of Egl-GFP (Figure 4.18G-G'', 4.18H n=7). These results indicate that increasing the stability of the distal stem, and thus reducing the accessibility, abolishes the association of Egl with the RNA. It is possible that a combination of the conserved UA base-pair implicated to be important for *grk* transport in M44-6, and its accessibility to the transport machinery, may be essential for *grk* recognition.

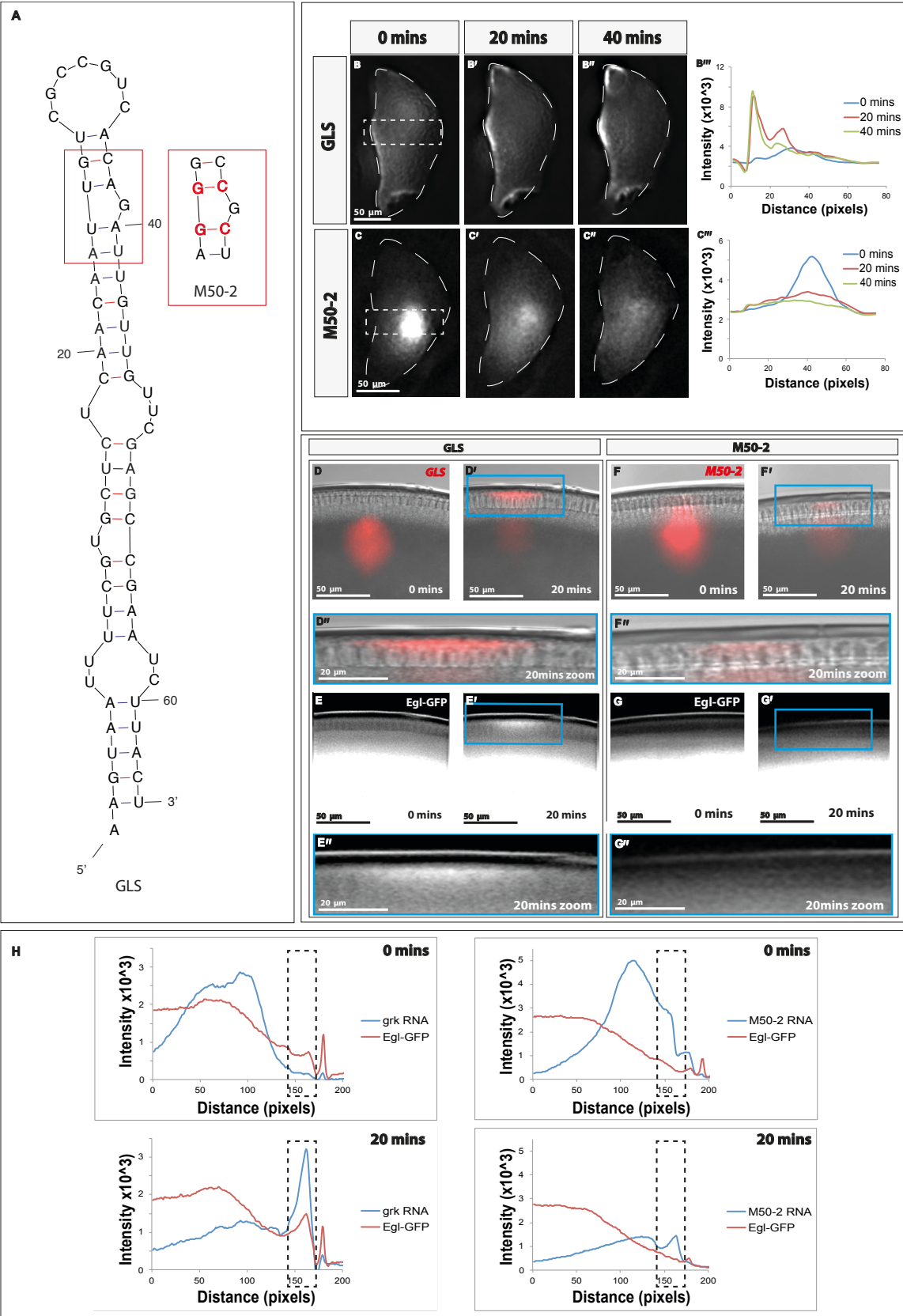


Figure 4.18. Increasing the stability of distal stem 4 abolishes *grk* transport in the oocyte and embryo due to an inability to recruit the transport machinery component Egl. (A) UNAFold predicted secondary structure of the GLS and the mutant M50-2 containing two strong GC base-pairs on either side of bulge 2, reducing accessibility of the distal stem. (B-B'') The wild type GLS is capable of driving *grk* to the anterior of the oocyte within 20 minutes of injection before it enriches at the dorsal-anterior, as shown in the fluorescent intensity profile (B''') taken through a cross-section of the oocyte (white box in B). (C-C'') M50-2 is not transported to the anterior of the oocyte within 40 minutes of injection, as shown in the fluorescent intensity profile (C''') taken through a cross-section of the oocyte (white box in C). (D-D'') The GLS drives *grk* RNA localisation to the apical cortex of the embryo resulting in Egl-GFP recruitment (E-E''). (F-F'') M50-2 is unable to drive *grk* localisation apically and cannot recruit Egl-GFP within 20 minutes of injection into the embryo (G-G''). (H) Fluorescent intensity profiles taken from a cross-section through the embryo at 0 minutes and 20 minutes following the injection of the GLS and M50-2. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in B''' and C''' refers to the position through the cross-section of the oocyte (white boxes in B and C) from the posterior to the anterior. Distance in H refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

To investigate whether the distal stem must be in an accessible region of the signal below bulge 2, the mutant M42-9 was generated where stem 4 is moved to a position within stem 1 (Figure 4.19A). The secondary structure prediction of M42-9 suggests that the surrounding regions of the stem-loop are unaffected by moving stem 4 (Figure 4.19A). In contrast to wild type *grk* RNA (Figure 4.19B-B''), M42-9 forms large aggregates throughout the cytoplasm in 70% of oocytes injected (Figure 4.19C-C'', n=30), and is diffuse throughout the cytoplasm in a further 20% of oocytes. These results are clearly displayed in the fluorescent intensity profiles taken through a cross-section of the oocyte (white box) at different time points throughout the experiment (Figure 4.19D). However, many of the oocytes also showed an enrichment of the RNA at the anterior or DA, suggesting that a low proportion of the RNA can be transported and anchored.

To investigate whether a proportion of the M42-9 mutant RNA can be recognised by the transport machinery, or whether it diffuses in the oocyte and is able to anchor, I tested its ability to recruit Egl-GFP in the embryo. Injected M42-9 localises apically in 100% of embryos injected (Figure 4.19G-G'', n=9) in a similar manner to wild type *grk* RNA (Figure 4.19E-E'', n=10), albeit at a slightly reduced level with more of the RNA remaining in the basal cytoplasm. However, M42-9 only shows a strong level of Egl-GFP recruitment to the apical

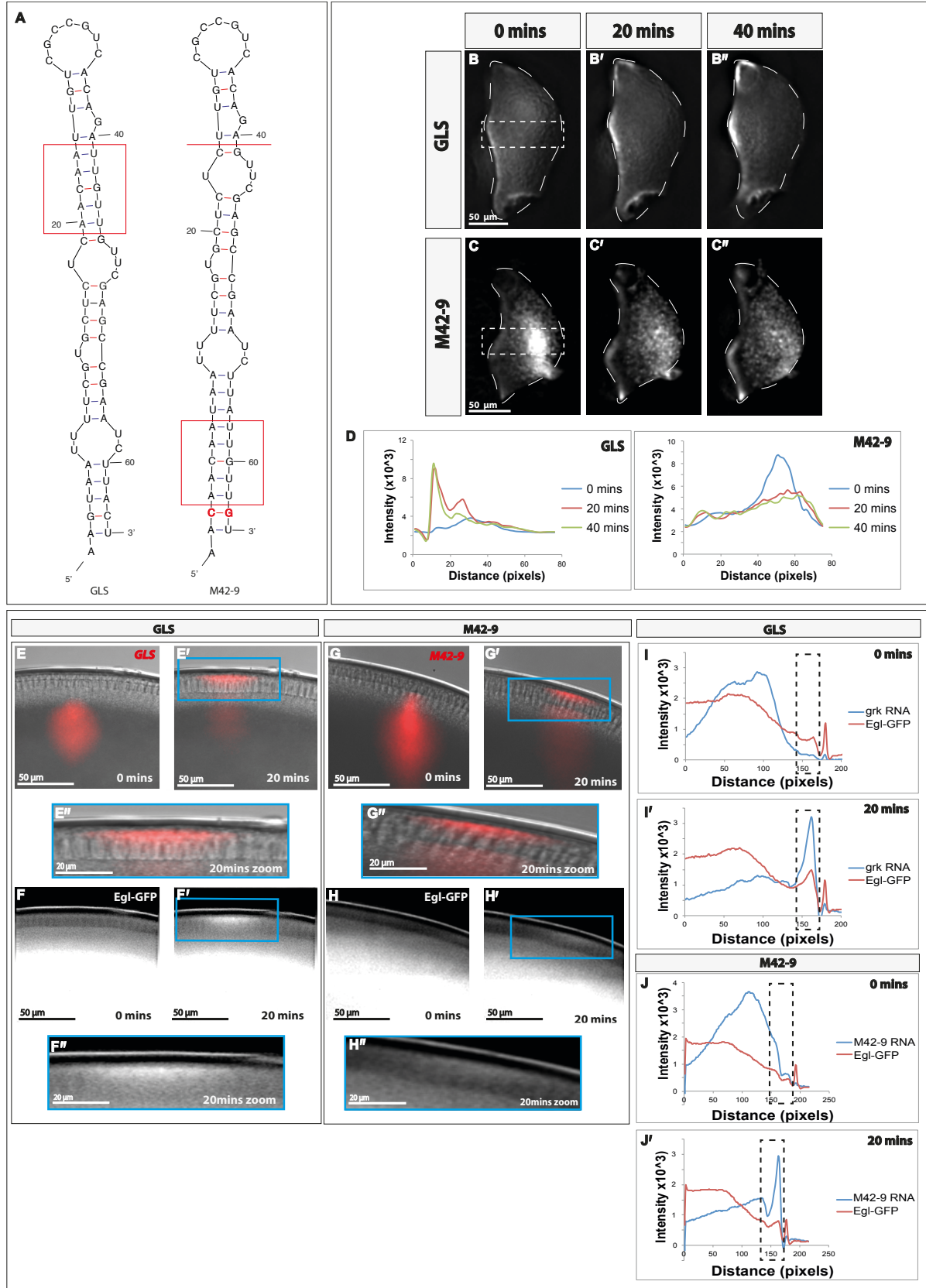


Figure 4.19. Moving the position of distal stem 4 reduces *grk* localisation in the oocyte and the embryo due to poor recruitment of the transport machinery component Egl.

(A) UNAFold secondary structure predictions of the GLS and the mutant M42-9 where the distal stem is moved to within the proximal stem. (B-B'') The GLS drives *grk* transport to the anterior of the oocyte within 20 minutes of injection. (C-C'') M42-9 greatly reduces *grk* transport in the oocyte and causes the RNA to form aggregates. (D) Fluorescent intensity profiles generated from a cross-section through the oocyte (white boxes in B and C) showing that the enrichment of *grk* RNA at the anterior of the oocyte is greatly reduced in M42-9. (E-E'') Wild type *grk* RNA (red) localises apically in the embryo (DIC) and recruits Egl-GFP (greyscale) (F-F''). (G-G'') M42-9 localises apically in the oocyte but recruits lower levels of Egl-GFP (H-H'') than wild type *grk*. Fluorescent intensity profiles taken through a cross-section of the embryo at 0 minutes and 20 minutes following injection of the GLS (I-I') and M42-9 (J-J'). (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (D) refers to the position through the cross-section of the oocyte (white boxes in B and C) from the posterior to the anterior. Distance in (I-J') refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

cortex in 56% of cases (Figure 4.19H-H'', J-J', n=9), with the remaining embryos showing much lower levels of Egl-GFP recruitment than wild type *grk* (Figure 4.19F-F'', I-I', n=10). Taken together, these results suggest that movement of distal stem 4 to stem 1 reduces transport of *grk* in the oocyte and embryo due to a reduction in the recruitment of Egl. However, in both the oocyte and the embryo a substantial proportion of the RNA is transported, indicating that binding of the transport machinery is not completely abolished.

Conclusions Increasing the stability of distal stem 4 abolishes *grk* localisation in the oocyte and the embryo due to a failure to recruit the transport machinery to the RNA. Moving the position of the distal stem greatly reduces the localisation efficiency but does not abolish association of the transport machinery entirely.

The orientation of internal loop 2 is essential for driving *grk* transport

Although the sequence of the two internal loops are variable among the GLS-like signals identified in other species (Figure 4.5C, E), both of the internal loops are always present and may provide ssRNA motifs for protein binding, or a way of further opening up the stem-loop to allow proteins access to dsRNA helices.³⁴⁸ The importance of internal loop 1 for *grk* localisation was previously tested in the lab through the mutant M13, where internal loop 1 is

removed (Van De Bor and Ball, unpublished), and in this thesis with the truncation M1 (Figure 4.6). Both mutants demonstrated that the loop is not essential for *grk* transport. However, previous studies have shown that internal loop 2 is absolutely essential for maintaining *grk* transport, as removing the loop in M12 results in an inability of the RNA to localise in the oocyte (Van De Bor and Ball, unpublished; Figure 4.20A). In addition, they showed that removing two uridines from internal loop 2 abolishes *grk* localisation (Figure 4.20A), indicating these nucleotides are required to provide a sequence or structural motif for the transport machinery.

To determine whether the sequence or the structure of internal loop 2 is required for *grk* transport, M40-2 and M40-4 were designed in which the internal loop is reduced by removing a single nucleotide (Figure 4.20A). M40-2 drives *grk* to the anterior of the oocyte within 20 minutes of injection (Figure 4.20C-C''; Figure 4.21B; Table 4.4), as seen for the wild type GLS (Figure 4.20B-B''; Figure 4.21A; Table 4.4). M40-4, on the other hand, is only able to drive low levels of *grk* transport in 50% of oocytes injected, with *grk* showing signs of enrichment at the anterior or DA, but with high levels of RNA diffuse throughout the cytoplasm. The remaining 50% of oocytes injected show signs of RNA aggregation and diffusion throughout the cytoplasm, with no enrichment of the RNA at the anterior (Figure 4.20D-D''; Figure 4.21C; Table 4.4). Taken together, these results indicate that a small proportion of M40-4 RNA is transported to the anterior, allowing it to become anchored.

To test whether M40-4 shows a reduction in localisation efficiency due to inefficient recruitment of the transport machinery, Egl-GFP recruitment was assessed in the embryo. Unlike the non-localising control *hb* (Figure 4.22C-C', n=7), M40-4 (Figure 4.22A) becomes enriched at the apical cortex of 100% of the embryos injected (Figure 4.22D-D'', n=7), although much more weakly than wild type *grk* (Figure 4.22B-B', n=10) with more RNA remaining in the basal cytoplasm. However, M40-4 only shows low levels of Egl-GFP recruitment in 86% of embryos injected (Figure 4.22E-E'', n=7), indicating that removing a

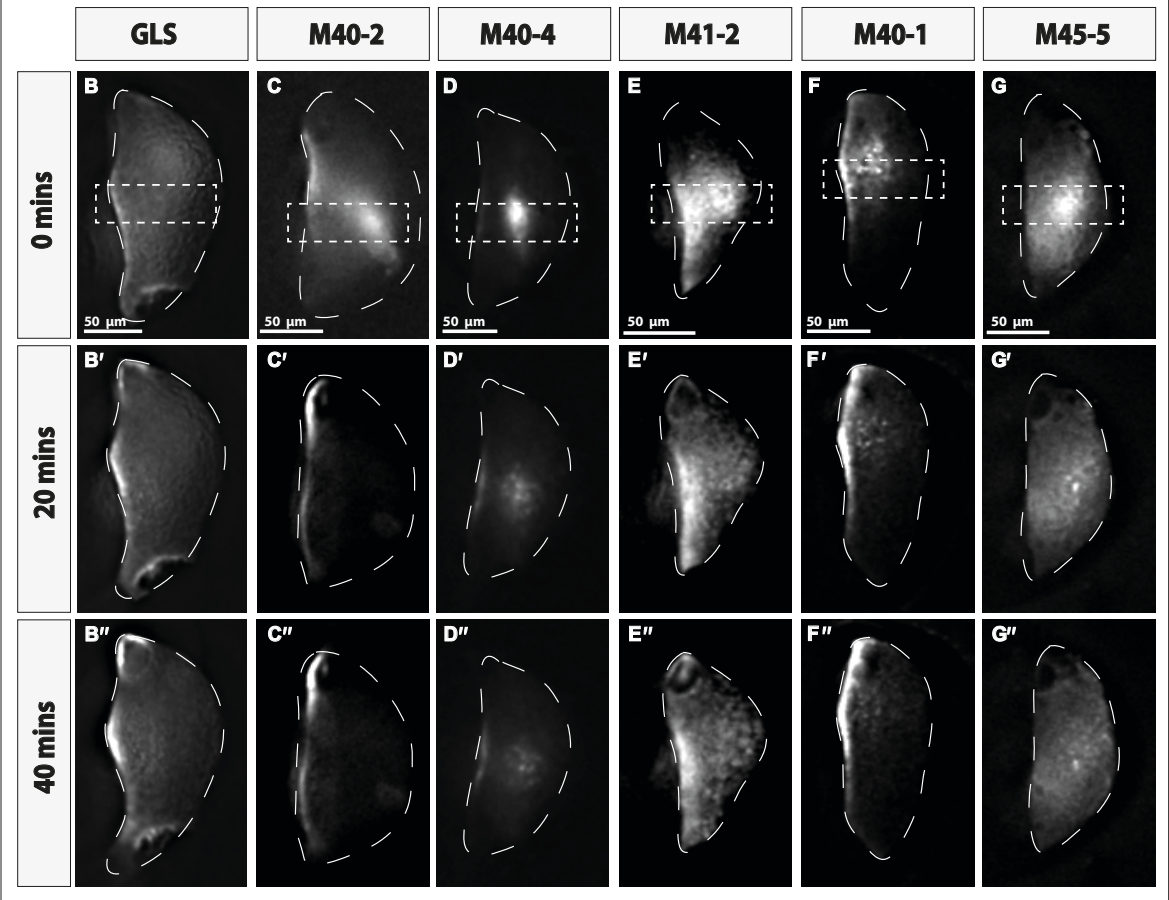
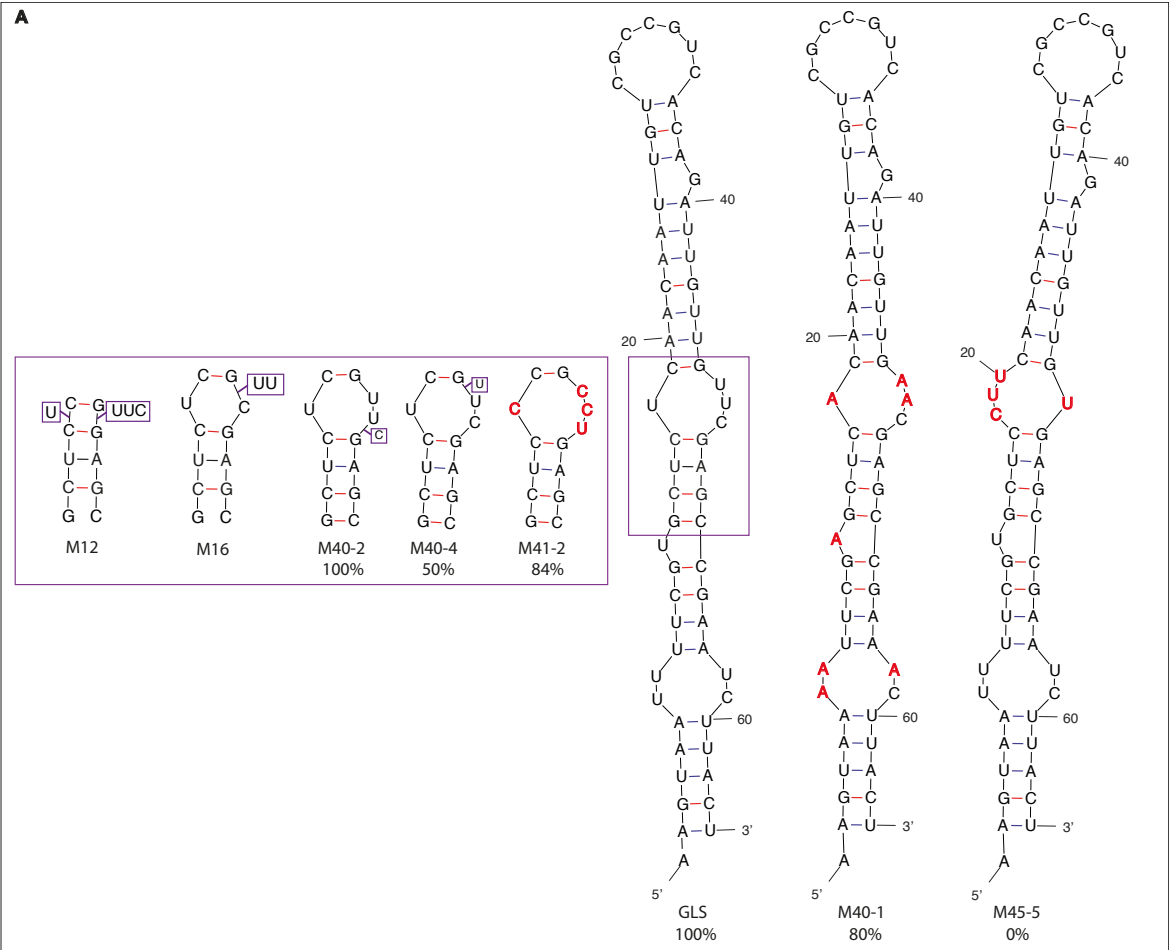


Figure 4.20. The sequence of internal loop 2 is not essential for *grk* transport but both the size and the orientation of the kink produced is. (A) UNAFold predicted secondary structures of the GLS and internal loop 2 mutants. Removing internal loop 2 (M12) and reducing the size by more than one nucleotide (M16) abolishes *grk* transport (Van De Bor and Ball, unpublished). (B-B'') The wild type GLS drives *grk* transport to the anterior of the oocyte within 20 minutes of injection. Reducing internal loop 2 by a single C nucleotide has no effect on *grk* transport (C-C''), whereas removing a U nucleotide greatly reduces *grk* transport (D-D''). Changing the sequence of the internal loops slightly reduces *grk* transport efficiency (E-F''). Changing the orientation of internal loop 2 abolishes *grk* transport (G-G''). White dashed rectangles show the regions of the oocyte where fluorescent intensity profiles were taken for Figure 4.21. Oocytes outlined by white dashed lines.

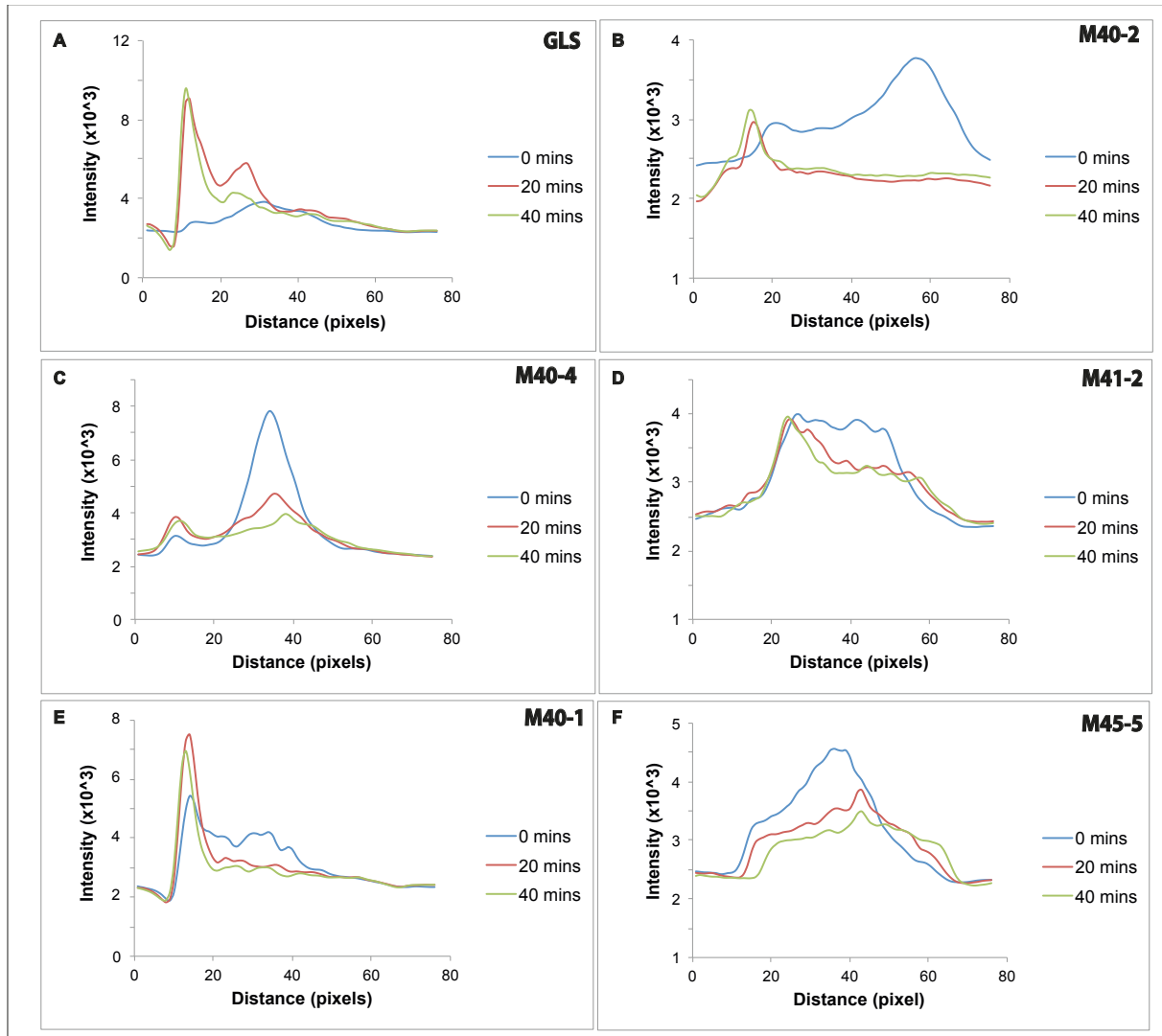


Figure 4.21. Changing the orientation of internal loop 2 abolishes *grk* enrichment at the anterior of the oocyte. (A-F) Fluorescent intensity profiles taken from a cross-section through the oocyte at different time points following injection of mutant RNAs. Distance refers to the position through the cross-section of the oocyte (white dashed rectangles in Figure 4.20 B-G) from the posterior to the anterior. M40-4 and M45-5 show poor *grk* transport to the anterior of the oocyte, corresponding to no clear peak at the left hand side of the graphs by 40 minutes.

RNA	Positions of sequence modifications within the GLS	n injected	% RNA transport	% No transport
GLS	N/A	54	100	0
M40-2	49ΔC	17	100	0
M40-4	47ΔU	28	50	50
M41-2	18U>C, 47-49UUC>CCU	31	84	16
M40-1	7-8UU>AA, 13U>A, 47-48UU>AA, 58U>A	50	80	20
M45-5	18U>CUU, 47-49UUC>U	35	0	100

Table 4.4. Modifications in internal loop 2 reduce *grk* transport in the oocyte. Outline of the sequence modifications made to the GLS in each mutant designed to test whether internal loop 2 is required for *grk* transport in the oocyte. Changes to a nucleotide at the given position are marked by (>). The percentage of oocytes showing *grk* RNA transport to the anterior within 40 minutes of injecting each mutant are shown, along with the percentage that fail to show *grk* transport. N-numbers for each mutant are displayed.

uridine from internal loop 2 greatly reduces recruitment of Egl. This is reflected in the fluorescent intensity profiles taken from a cross-section through the middle of the embryo at different time points throughout the assay (Figure 4.22F). It is possible that M40-4 is able to recruit low levels of Egl which are below the detection limit of the embryo recruitment assay, resulting in weak localisation of *grk* RNA in both the oocyte and embryo, with the localisation pattern being more profound in the embryo due to a shorter distance to be transported, or a more polarised microtubule network. Taken together, these results indicate that removing the cytidine nucleotide from internal loop 2 has no effect on *grk* transport, whereas removing a uridine reduces *grk* transport due to poor recruitment of Egl. Therefore, it is possible that the sequence of internal loop 2, rather than the size, may be important for *grk* transport.

To establish the importance of the sequence of internal loop 2, all of the nucleotides were mutated in M41-2. When injected into the oocyte, M41-2 transports *grk* to the anterior or DA within 40 minutes in 84% of cases (Figure 4.20E-E"; Figure 4.21D; Table 4.4), indicating that

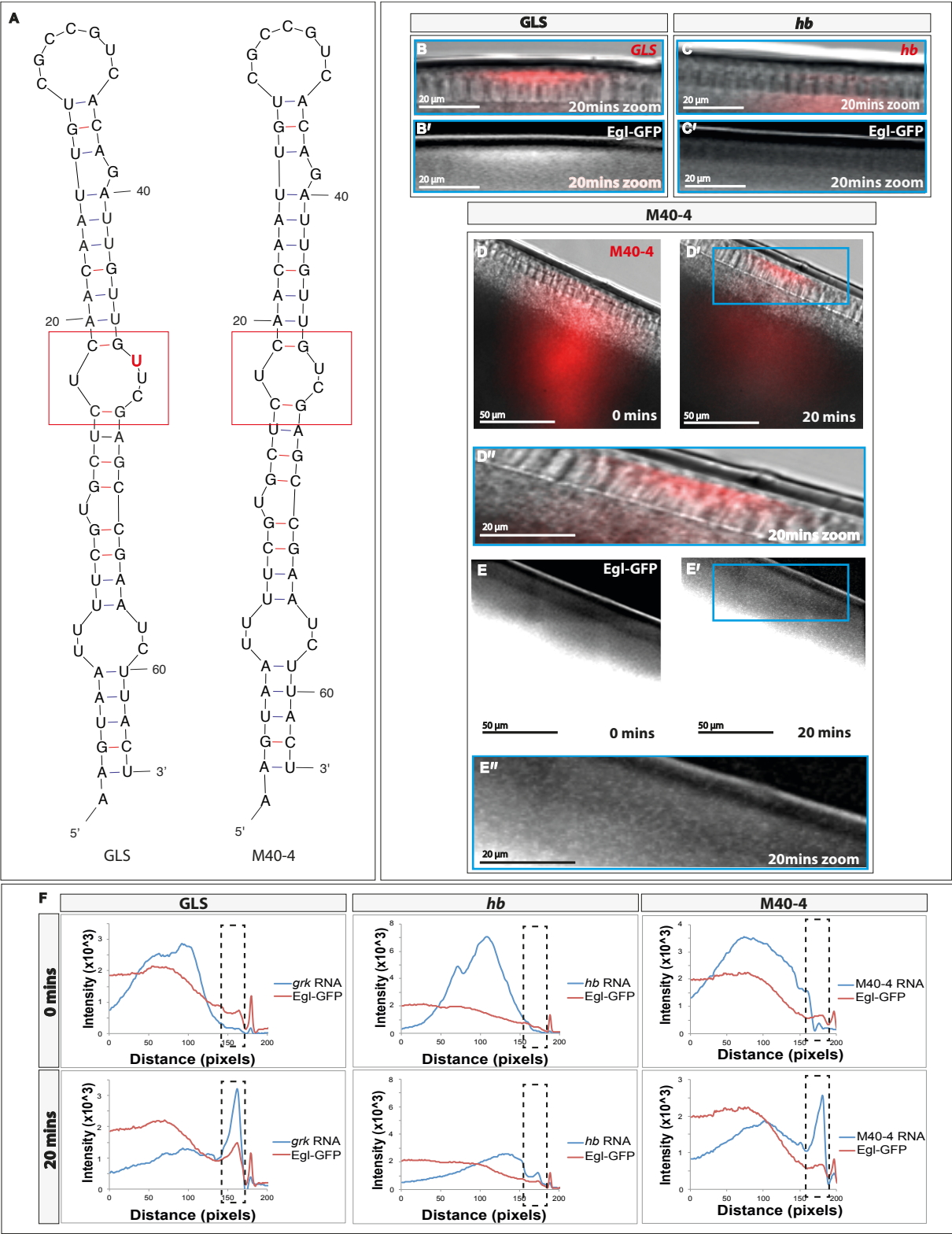


Figure 4.22. Reducing the size of internal loop 2 by removing a uridine reduces *grk* transport in the embryo and recruitment of the transport machinery. (A) UNAFold predicted secondary structures of the GLS and mutant M40-4. (B-B') The wild type GLS signal drives *grk* RNA (red) to the apical cortex of the embryo (DIC) and recruitment of Egl-GFP (greyscale), whereas the non-localising control *hb* does not (C-C'). (D-D'') Removing a uridine from internal loop 2 reduces *grk* transport in the embryo and recruitment of Egl-GFP (E-E''). (F) Fluorescent intensity profiles of the GLS, *hb* and M40-4 taken from a cross-section through the middle of the embryo. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

the sequence of internal loop 2 is not essential. However, internal loops 1 and 2 may work together for the binding of the transport machinery as they are similar in size and sequence, generating redundancy. To test this possibility, mutations were made to the uridines in internal loops 1 and 2, and bulge 1 to create M40-1 (Figure 4.20A). However, the mutations in M40-1 do not further reduce the percentage of oocytes showing *grk* transport following injection, with 80% still showing anterior or DA localisation (Figure 4.20F-F''; Figure 4.21E; Table 4.4). As a result, the sequences of internal loop 1 and internal loop 2 are not essential for *grk* transport, but they may be required for full efficiency.

The presence of internal loop 2 has been shown previously, using partial NMR structures, to generate a kink in the signal, forcing the distal half of the stem to fold towards the proximal half (Ball and McDermott, unpublished). The kink generated by the internal loop is represented to some extent in the predicted secondary structure of the signal (Figure 4.20A). In order to test whether this kink is important for *grk* transport, I generated sequence modifications in the GLS that reflects internal loop 2 to generate M45-5, shifting the kink in the opposite orientation (Figure 4.20A; Table 4.4). Injection of M45-5 into the oocyte shows a diffuse signal throughout the cytoplasm in 100% of cases (Figure 4.20G-G''; Figure 4.21F; Table 4.4), indicating that the orientation of the kink produced by internal loop 2 is essential for *grk* transport.

To investigate whether the failure of M45-5 (Figure 4.23A) to localise in the oocyte is due to an inability to bind the transport machinery, the RNA was injected into embryos expressing

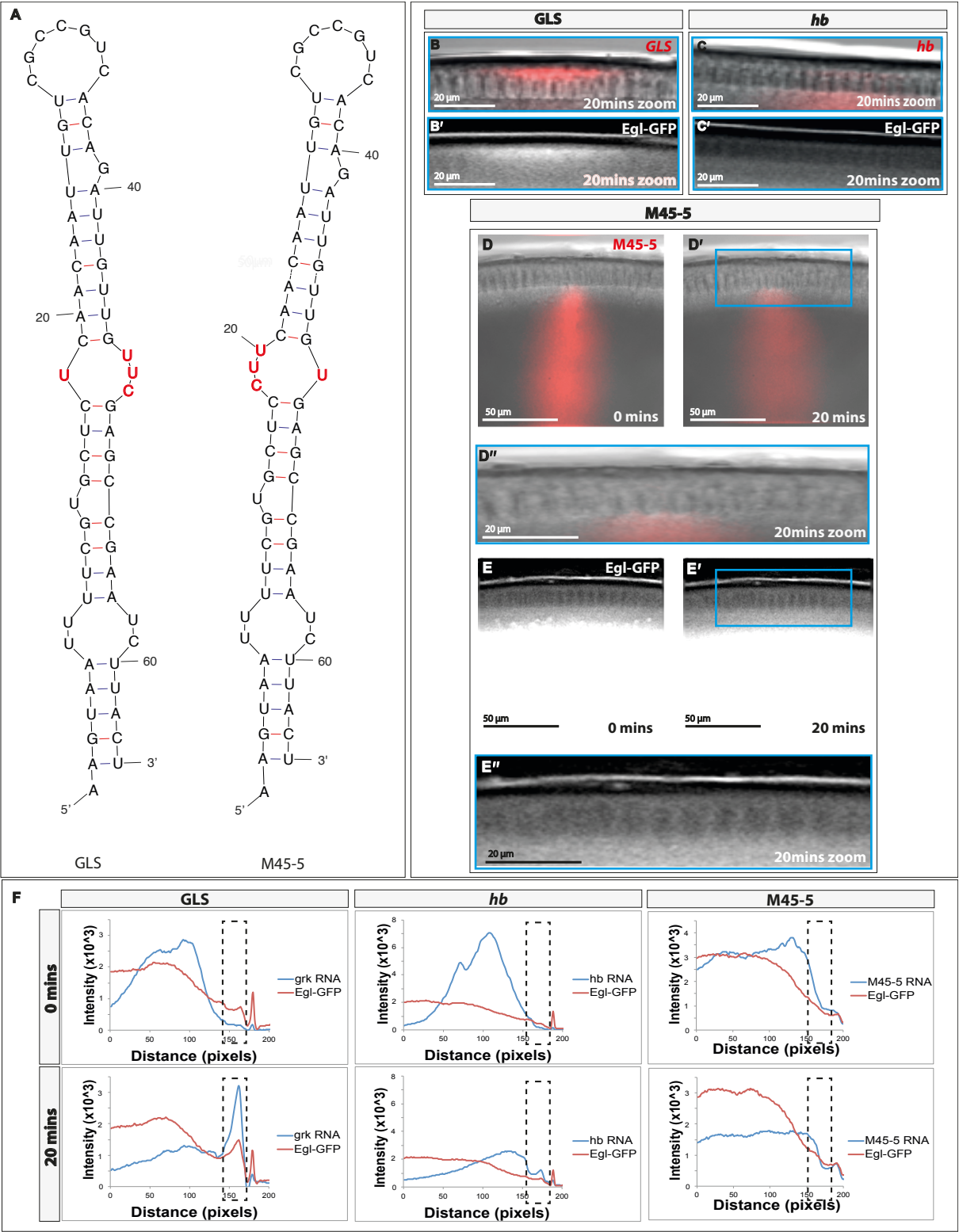


Figure 4.23. Switching the orientation of the kink produced by internal loop 2 abolishes *grk* transport in the embryo and recruitment of the transport machinery. (A) UNAFold predicted secondary structures of the GLS and mutant M45-5. (B-B') The wild type GLS signal drives *grk* RNA (red) to the apical cortex of the embryo (DIC) and recruitment of Egl-GFP (greyscale), whereas the non-localising control *hb* does not (C-C'). (D-D'') Switching the orientation of internal loop 2 abolishes *grk* transport in the embryo and recruitment of Egl-GFP (E-E''). (F) Fluorescent intensity profiles of the GLS, *hb* and M45-5 taken from a cross-section through the middle of the embryo. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 4.7) from the site of injection to the apical cortex (black box)).

Egl-GFP. Unlike the wild type GLS (Figure 4.23B-B', n=10), M45-5 is unable to localise apically (Figure 4.23D-D'', n=6) and cannot recruit Egl-GFP (Figure 4.23E-E'', 4.23F n=6), sharing the same characteristics as the non-localising control *hb* (Figure 4.23C-C', 4.23F, n=7). These results indicate that switching the orientation of the kink produced by internal loop 2 abolishes recruitment of the transport machinery to the RNA, and thus prevents the RNA from localising in both the oocyte and the embryo.

Conclusions The sequence of internal loops 1 and 2 are not essential for *grk* transport. The orientation of the kink in internal loop 2, and the degree of kink established by the size of the loop, are essential for *grk* transport and recruitment of the transport machinery.

The internal bulges are not required to drive *grk* transport

Bulges are often required in RNA to provide accessibility to sequences within dsRNA helices, ³⁴⁸ suggesting that bulges 1 and 2 may be important for recognition of the GLS by *trans*-acting factors, such as components of the transport machinery. To test whether bulges 1 and 2 are important for *grk* transport they were first removed in M3 and M38-5, respectively (Figure 4.24A), and their localisation patterns compared to the wild type GLS (Figure 4.24B-B''; Table 4.5). In both cases, removal of the bulge has no effect on *grk* transport, with M38-5 driving *grk* to the anterior of the oocyte within 20 minutes (Figure 4.24C-C''; Table 4.5), and

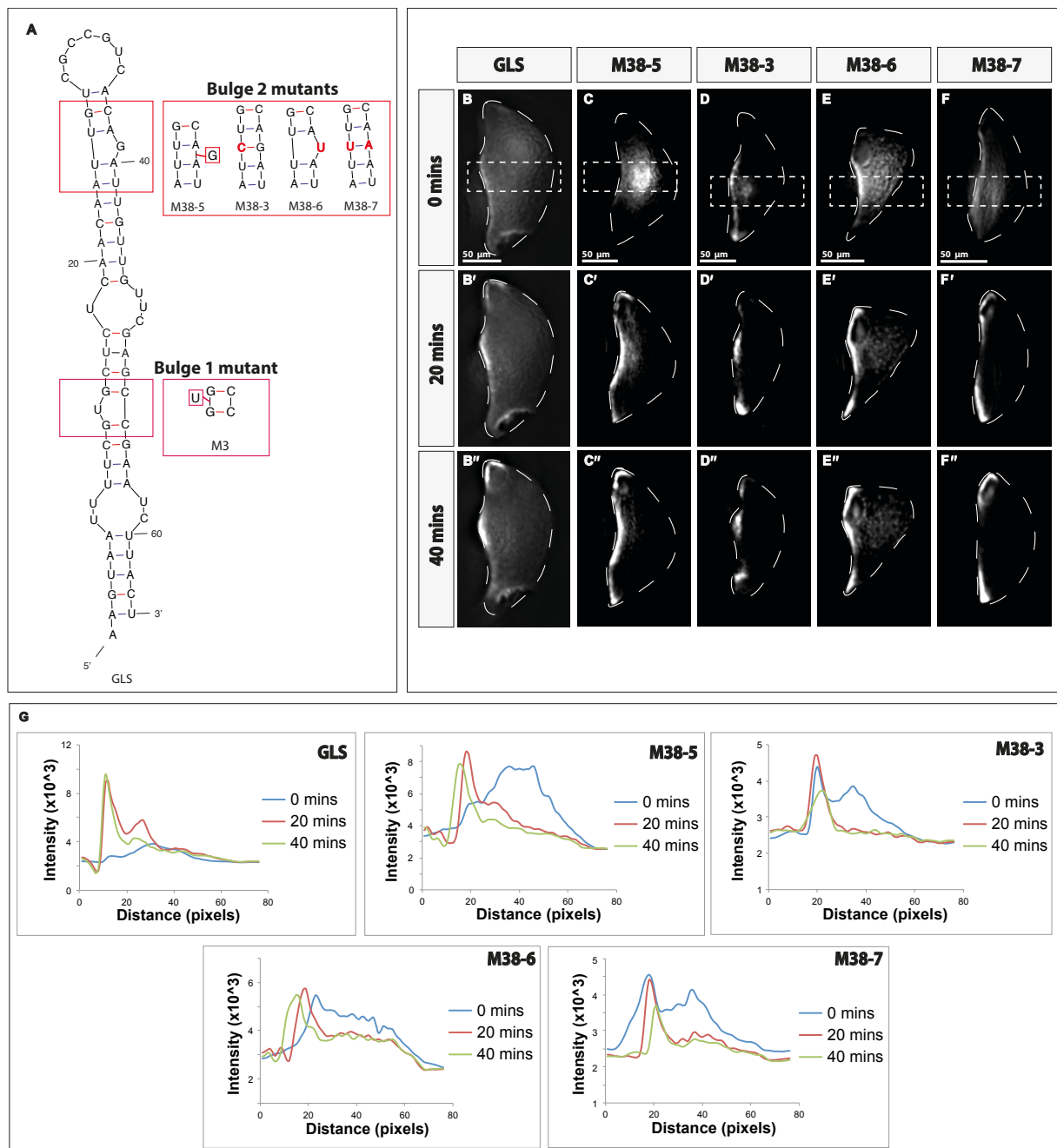


Figure 4.24. The presence of bulge 1 and bulge 2 is not required for efficient *grk* transport in the oocyte. (A) UNAFold predicted secondary structures of the GLS and mutants of bulge 1 and bulge 2. (B-B') The wild type GLS signal drives *grk* to the anterior of the oocyte within 20 minutes of injection. Removing bulge 2 (C-C'') and substituting the bulged nucleotide (E-E'') has no effect on *grk* transport. Forcing the bulge to create a strong CG base-pair only slightly reduces the efficiency of *grk* transport (D-D''), whereas changing the base-pair to a weaker AU has no effect (F-F''). (G) Fluorescent intensity profiles of the GLS and the bulge mutants taken from a cross-section through the middle of the oocyte (white dashed rectangles in B-F). Outline of oocytes shown by dashed white line.

RNA	Positions of sequence modifications within the GLS	n injected	% RNA transport	% No transport
GLS	N/A	54	100	0
M38-5	39ΔG	25	100	0
M38-3	25-26UU>UCU	66	91	9
M38-6	39G>U	17	100	0
M38-7	25-26UU>UUU; 38-40AGA>AAA	26	100	0

Table 4.5. Modifications in bulge 2 have no effect on *grk* transport in the oocyte.

Outline of the sequence modifications made to the GLS in each mutant designed to test whether bulge 2 is required for *grk* transport in the oocyte. Changes to a nucleotide at the given position are marked by (>). The percentage of oocytes showing *grk* RNA transport to the anterior within 40 minutes of injecting each mutant are shown, along with the percentage that fail to show *grk* transport. N-numbers for each mutant are displayed.

M3 previously described to localise normally (Figure 4.24A, Van de Bor and Ball, unpublished).

As bulge 1 is positioned within the proximal third of the GLS which was shown not to be required for *grk* transport, further mutations were made to bulge 2. Forcing bulge 2 to base-pair by adding in an extra nucleotide in M38-3 (Figure 4.24A) has very little effect on *grk* transport (Figure 4.24D-D''; Table 4.5). The loss of *grk* transport in 9% of oocytes may be due to a reduction in the flexibility of the essential stem 4 below the bulge, as replacing bulge 2 with a weaker U-A base-pair in M38-7 has no effect on *grk* transport (Figure 4.24F-F''; Table 4.5). Finally, changing the sequence of bulge 2 in M38-6 (Figure 4.24A) has no effect on *grk* transport in the oocyte (Figure 4.24E-E''; Table 4.5). Taken together, these data indicate that the presence of both bulge 1 and 2, and the sequence of bulge 2, is not essential for *grk* transport.

Conclusions The presence of bulge 1 and bulge 2 is not important for *grk* transport.

Discussion

Most of the novel GLS-like signals identified are able to drive *grk* RNA transport

My data reveal that divergent GLS-like signals from *grk*-like genes in both *Drosophilids* and other Dipterans are able to drive *grk* transport in the *D. melanogaster* oocyte, with the exception of *A. aegypti*. The lack of transport of the *A. aegypti* signal is likely due to primary sequence and predicted secondary structure divergence from the GLS, which may prevent its recognition by the *D. melanogaster* transport machinery. However, whether the signal drives transport of a similar axis-determining RNA in *A. aegypti* is currently unknown.

The case of *A. aegypti* aside, these data suggest that the GLS is highly conserved in sequence, structure and function throughout the *Drosophilidae* family and even in a subset of Dipterans. It would be interesting to investigate how far the GLS and transport machinery conservation extends outside of the Diptera to establish polarity in the oocyte. *grk* orthologs have not been identified outside of the Diptera due to its rapidly evolving nature. However, the EGF signalling network required in *D. melanogaster* for mediating the encapsulation of the oocyte by somatic follicle cells, establishing the polarity of the egg chamber, and for setting up the dorsal-ventral axis of the embryo, is broadly conserved in other insects.³⁸²

The conservation of the GLS can be used to identify regions important for function

Having demonstrated that many of the GLS-like signals identified are able to support *grk* transport in the *D. melanogaster* oocyte, I used both their primary sequences and predicted secondary structures to infer regions of the stem-loop that may be essential for transport. Although the primary sequence is likely to be highly constrained by both the requirement to

encode the protein and to drive mRNA localisation, it provides the most accessible way of testing the importance of different regions of the GLS in driving *grk* transport.

Regions of the GLS with the lowest primary sequence conservation predicted not to be essential for *grk* transport were stem 1, stem 2, internal loop 1, internal loop 2, bulge 2 and the hairpin loop. Generating mutations within these regions and testing their effects on *grk* transport confirmed that stem 1, internal loop 1 and bulge 2 are indeed redundant. The presence of stem 2, on the other hand, is required for high levels of *grk* transport, potentially by providing stability to the stem-loop secondary structure, as the sequence of the stem is not important.

Although the sequence of internal loop 2 is not conserved, it is present in all of the GLS-like signals identified and previous structural studies have shown that the unpaired nucleotides are flipped out of the structure and generate a kink in the signal (Ball and McDermott, unpublished). As predicted, the sequence of internal loop 2 is not important for *grk* transport and thus recognition by the transport machinery. Instead, the orientation of the kink produced by internal loop 2 is essential, suggesting that the secondary structure of this region in some way contributes to providing a recognition motif for the transport machinery. In agreement with conservation studies of internal loop 2, the minimum size of the loop is two unpaired nucleotides on the right side of the stem. However, removing a single nucleotide has mixed effects on *grk* transport, with removal of the cytidine leaving *grk* transport unaffected, and removal of a uridine reducing *grk* transport and Egl recruitment. Although further investigations are required to establish the exact features of internal loop 2 required for *grk* transport, I favour a model where the internal loop plays a role in providing the correct distances required for binding of the transport machinery through generating a kink in the signal.

The sequence of the hairpin loop is not highly conserved and sequence variations are predicted to greatly affect the secondary structure of both the hairpin and bulge 2. As

predicted, sequence modifications within this region carried out previously were shown to have no effect on *grk* transport, further exemplifying the fact that bulge 2 is not required. It was instead shown that the size of the hairpin loop is essential for *grk* transport in the oocyte, with the minimum size of five nucleotides being required.

Although removal of bulge 2 has no effect on *grk* transport, mutation of the UA base-pairs on either side of the bulge to strong GC base-pairs abolishes *grk* transport in the oocyte and embryo and recruitment of Egl. It is possible that introducing the stronger GC base-pairs reduces accessibility to stem 4, or that the entirely conserved UA base-pair below bulge 2 is essential for providing some sort of recognition motif for the transport machinery. As strong GC base-pairs are present around the bulge in other GLS-like signals shown to be transported in the oocyte, I propose that *grk* transport requires the UA base-pair above stem 4, rather than weak base-pairing around bulge 2.

Stem 4 is one of the most highly conserved regions of the GLS, making it a likely candidate for playing a role in *grk* localisation. Although the sequence of stem 4 was shown not be essential for transport, extensive mutational analysis of this region confirms that both the AU-content and the length of the stem are essential for *grk* transport and recruitment of Egl, suggesting that it provides a recognition site for the transport machinery. In addition, mutations within this region support the requirement for the UA base-pair above stem 4 in *grk* transport. Modification of this base-pair alone would provide more insight into its importance in *grk* transport.

Bulge 1 is also highly conserved among the GLS-like signals identified and previous structural studies have shown that the unpaired nucleotide is indeed bulged out of the lower helix, making it a candidate for protein interactions. However, they also demonstrated that the presence of bulge 1 is not required for *grk* transport (Ball, Van De Bor and McDermott, unpublished). The highly conserved stem 3, on the other hand, is essential for *grk* transport with mutations to both the sequence and the orientation of this region abolishing *grk*

localisation in the oocyte and the embryo, due to an inability to recruit the transport machinery. As a result, stem 3 likely provides a recognition motif for Egl binding.

From the data obtained in this chapter I have identified several regions of the GLS that are essential for driving *grk* transport in the oocyte and the embryo (Figure 4.25A), allowing me to generate a model for the recognition of the GLS by the transport machinery.

A model for the recognition of the GLS by the transport machinery

In the oocyte, Egl is thought to link mRNA cargoes to the Dynein motor through recognition of A'-form structures in the RNA, as previously shown for *K10*.³⁸⁶ Although the purine stacks in stem 2 and stem 3 of the GLS were previously predicted to generate A'-form RNA,³⁸⁶ partial structural analysis suggests that the helical regions are likely to be in the A-form conformation (Ball and McDermott, unpublished). Furthermore, part of the purine stack in stem 2 is predicted to be located in an internal loop in some of the GLS-like signals. Finally, I have demonstrated that mutants predicted to disrupt A'-form RNA do not abolish *grk* transport suggesting that the A'-conformation may not be important for Egl binding to the GLS, at least not in the region of the signal where it was initially predicted to form.

Using the data presented in this chapter I propose an alternative model for the recognition of the GLS by Egl, and thus the transport machinery (Figure 4.25B). In this model, Egl recognises the secondary structure of the AU-rich stem 4. A UA base-pair positioned directly above the stem is essential for this interaction. Egl also interacts with the purine stack in stem 3, with the sequence of this stem being essential for recognition. The kink in the signal provided by internal loop 2 generates the distances required for formation of a processive molecular motor complex between stems 3 and 4. Finally, the size of the hairpin loop is in some way essential for *grk* transport, with the minimum number of single-stranded nucleotides required being five.

Limitations to this model include the fact that moving stem 4 to a position within stem 1 greatly reduces *grk* transport in the oocyte and embryo, but does not completely abolish *grk* localisation as one might expect from the model. A possible explanation for this is that Egl can bind the complete AU-rich stem at any position within the GLS and recruit the rest of the transport machinery. However, to generate high levels of RNA transport, Egl must bind on either side of internal loop 2, recruiting components of the transport machinery which form a much more stable and processive motor complex when positioned within the constraints of the kink in the GLS. Specific structural positioning of the motor complex and the rigidity of the RNA in this kinked conformation may present the motor complex to the microtubules in such a way that encourages processive movement.¹³⁰ It is also possible that the reduction in *grk* transport in this mutant is a result of disrupting stem 1 which has been shown to be required for full efficiency of *grk* localisation. Alternatively, if Egl interacts with both stem 3 and stem 4 there may be some redundancy allowing recognition of stem 3 in the mutant, driving low levels of *grk* transport in the oocyte. However, this is unlikely to be the case as mutations that disrupt the AU-content of stem 4 completely abolish *grk* transport and Egl recruitment, arguing against the role of redundancy between stem 3 and stem 4, unless allosteric binding takes place requiring the initial interaction of Egl with stem 4 before binding to stem 3.

Future Work

To confirm their ability to drive *grk* localisation, more of the GLS-like signals should be injected into the *D. melanogaster* oocyte. In addition, it would be interesting to investigate why the GLS-like signal from *A. aegypti* is unable to support *grk* transport by systematically mutating different regions of the signal back to the wild type GLS to try and rescue function, and testing whether the signal has co-evolved with the *A. aegypti* transport machinery through cross-species injections. Further analysis of the regions of the GLS shown here to be important for *grk* transport, such as biochemical and structural studies, would provide further insights into the exact features required for binding of the transport machinery. Searching for the features identified as being essential for the function of the GLS in other

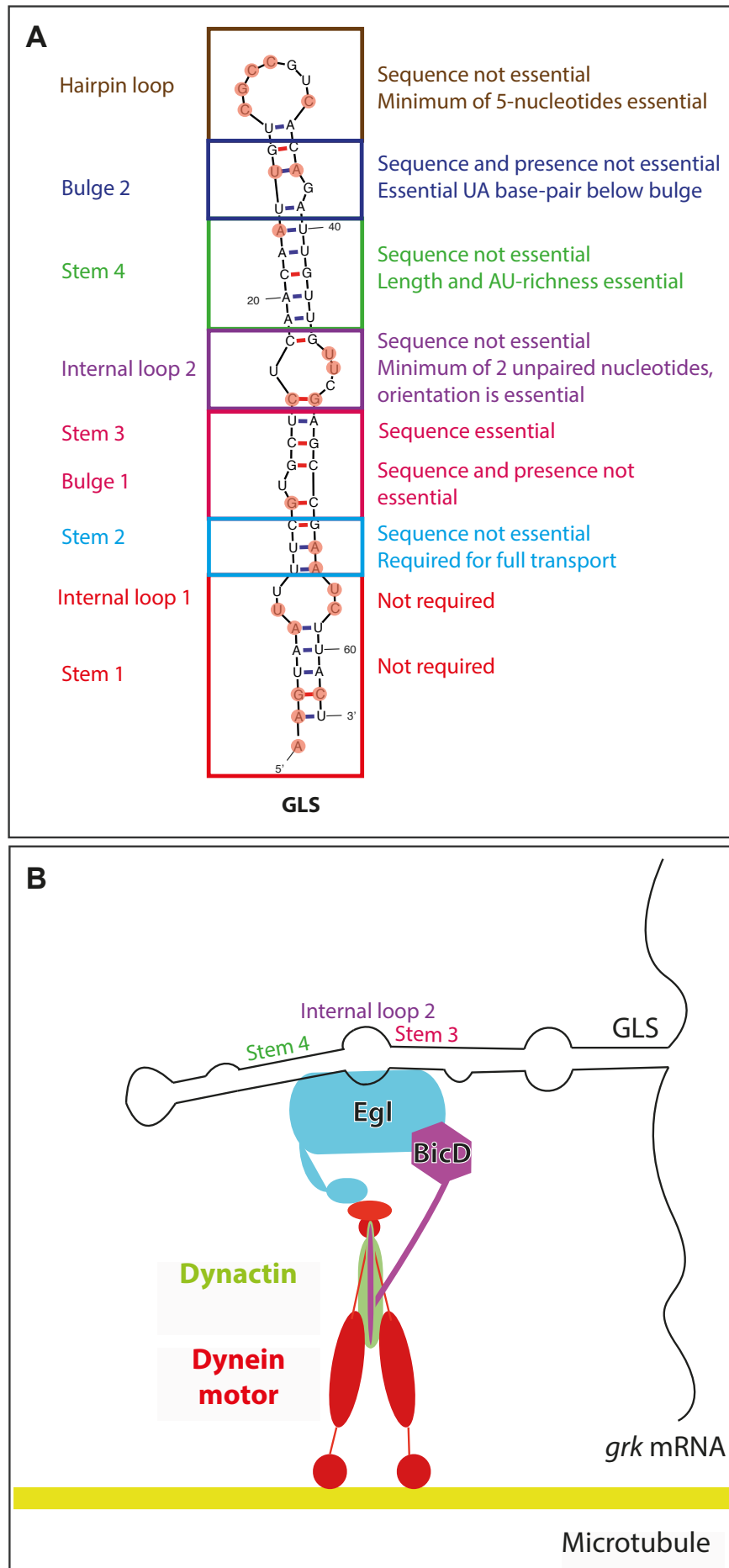


Figure 4.25. Schematic model for the recognition of the GLS by the transport machinery. (A) UNAFold predicted secondary structure of the GLS with non-conserved nucleotides in other GLS-like signals identified highlighted in orange. Several features of the signal have been identified to be important for *grk* transport and are summarised here. (B) Efficient recognition of the GLS by the transport machinery requires two helical stems positioned on either side of an internal loop. The most likely candidate for interacting with the RNA is Egalitarian which couples RNA cargoes to the dynein motor complex via direct interactions with the accessory protein BicD and the Dynein motor itself. The dynein motor complex then associates with microtubules to drive *grk* transport in the oocyte and the embryo.

dynein-transported transcripts would allow the identification of motifs that may be universally recognised by the transport machinery (see Chapter 7), thus providing the tools to predict localisation signals in other uncharacterised transcripts. A better understanding of the second anchoring step of *grk* localisation would allow us to determine whether some of the transport mutants showing enrichment at the anterior and the DA do so without a requirement for initial transport, or if all anchoring requires some form of inefficient transport (see Chapter 5). This would allow us to map the two steps of *grk* localisation to distinct regions of the GLS if the processes can be separated.

Final conclusions

Taken together, the data presented in this chapter lead me to conclude that both structural and sequence elements within the GLS are essential for recognition by the transport machinery, most likely through interactions with Egl. In agreement with conservation studies, the most important features that are absolutely essential for recognition by the transport machinery and for *grk* localisation include:

1. A hairpin loop of five or more nucleotides
2. An AU-rich stem 4 of the correct length with an essential UA base-pair positioned directly above
3. An internal loop between stem 3 and stem 4 that must be the correct size and orientation to produce a kink in the RNA
4. A purine stack in stem 3 whose sequence must be maintained
5. Stem 2 whose sequence is not essential but provides stability to the stem-loop

Chapter 5: Regions of the GLS required for the DA localisation and anchoring of *grk* RNA

Introduction

The localisation of *grk* RNA to the DA corner of the oocyte requires both active transport and anchoring.^{120, 133} Despite much research being directed towards understanding how *grk* RNA localisation is achieved and controlled, only Sqd and Dynein have emerged as factors required for anchoring, but exactly how they achieve anchoring is not clear. Moreover, although several other proteins have been identified using genetics and biochemistry,^{54, 133, 317} their precise roles in *grk* anchoring are unknown. Finally, the binding motifs in the GLS that confine *grk* to the DA corner of the oocyte are yet to be defined.

In this chapter I have used information gained from direct binding assays to predict putative Sqd binding sites in the GLS. Using the *in vivo* injection assay, I have tested the effect of a range of sequence modifications within the putative Sqd binding sites on the localisation pattern of *grk* RNA in the oocyte. Finally, several observations made during this process has led me to reinterpret the role of Sqd in *grk* anchoring, and how this is affected in a *sqd* mutant.

***grk* RNA is anchored at the DA of the oocyte**

grk localisation to the DA of the oocyte is a two-step process. First, *grk* is transported from the posterior to the anterior via microtubules using Dynein. Second, *grk* becomes restricted to the DA.^{120, 133} Studies using MS2-tagged transcripts have shown that *grk* particles at the anterior of the oocyte are highly dynamic, whereas those at the DA are less dynamic and are statically anchored.²⁸³ Several mutants have been identified which cause patterning defects due to the mislocalisation of *grk* RNA. For example, *sqd1*, *hrb27c*, *heph*, *uap56* and *fs(1)K10* mutants all lay eggs with dorsal-ventral defects as a result of inappropriate *grk* localisation

and/or translation along the anterior of the oocyte.^{15, 162, 299, 313, 394, 395} However, the exact molecular mechanisms directing *grk* anchoring at the DA of the oocyte are not well understood.

The role of Sqd during mid-oogenesis

Sqd is a conserved hnRNP, also known as Hrp40, and belongs to the HNRNPA family of RNA binding proteins.^{396, 397} This family of proteins is well conserved from *C. elegans* to humans and is thought to retain similar functions throughout evolution.³⁹⁸ Sqd contains an M9 nuclear import motif and is known to be required for regulating the nuclear export of *grk* RNA into the cytoplasm.³¹⁷ Alternative splicing of Sqd produces three isoforms: SqdA, SqdB and SqdS. Despite their similarities over the majority of the protein, each isoform has its individual role in the oocyte and distinct intracellular distributions. For example, SqdA and SqdS have overlapping but non-identical functions in *grk* localisation and translation. SqdB, on the other hand, is incapable of supporting wild type *grk* localisation and protein accumulation. Although all three isoforms are expressed in the germline, SqdB and SqdS are found in the nuclei, whereas SqdA is mainly found in the cytoplasm.³¹⁷ Sqd also plays a role in somatic tissues as null alleles of *sqd* are lethal. However, *sqd1* is a non-lethal allele containing a P-element insertion in the 5'UTR of *sqd*, which disrupts the germline promoter of all three isoforms and can be used to investigate the germline requirement for *sqd* during axis specification.²⁹⁹

Sqd is thought to be involved in both the correct localisation of *grk* mRNA and its translational silencing during transport,^{120, 133, 162, 193, 317} and its role in *grk* anchoring has been described previously.^{133, 283} In a *sqd1* mutant, *grk* mRNA is mislocalised such that it shows a stronger, more persistent distribution along the anterior of the oocyte than wild type,^{120, 317} suggesting a lack of anchoring in these mutants.^{283, 133} As a result, Grk is synthesised along the anterior of the oocyte where *grk* mRNA is mislocalised, resulting in a dorsalised egg phenotype.²⁹⁹ Sqd has been shown to switch Dynein from a processive motor into a static anchor at the

DA of the oocyte.^{120, 133} Here, *grk* is targeted to the edge of P bodies where it is actively translated due to its accessibility to ribosomes and translational activators such as Orb.¹⁹³ This suggests that anchoring of *grk* is tightly linked to its translational regulation in P bodies.¹⁹³ Both Sqd and Dynein have been implicated in maintaining *grk* in P bodies at the DA.¹³³ Egl and BicD, on the other hand, are only required for the active transport of *grk* RNA and not for the anchoring process.¹³³

Sqd has been shown to directly interact with the 3'UTR of *grk* RNA using UV cross-linking analysis, and is likely to interact with many other RNAs within the oocyte.³¹⁷ In addition, Sqd has been shown to immunoprecipitate with the GLS and potentially interacts with *grk* at various positions throughout the transcript.⁵⁴ Despite data showing that Sqd can interact with *grk* RNA, little is known about the sequence or structural motifs required within the RNA for Sqd recognition.

Aims of this chapter and experimental rationale

In this chapter I will use the GLS-like signals identified from other species in Chapter 3 to investigate whether sequences and structures within the GLS that confer *grk* anchoring are conserved. Furthermore, I will dissect the regions of the GLS required for *grk* anchoring. The specific questions addressed are:

1. Is the anchoring function of the GLS conserved in GLS-like signals from other species?
2. Which features of the GLS are essential for recognition by the anchoring machinery?

In order to address these questions I will use the *in vivo* injection assay developed in Chapter 3 to assess subtle differences in *grk* localisation. I will use a fly-line expressing MS2-tagged *grk* in a *sqd1* background to define the characteristics of *grk* in a *sqd* mutant, and will use these characteristics to assess whether GLS-like signals from other species maintain their function in driving *grk* anchoring. Finally, I will dissect the effect of sequence modifications within the GLS on *grk* anchoring.

Results

Transport of injected *grk* RNA to the oocyte anterior is unaffected in a *sqd1* mutant

grk has been shown to localise to the DA corner of the oocyte in a two-step process involving transport to the anterior, followed by anchoring at the DA (Figure 5.1F).^{15,120,133} Although a role for Sqd has been implicated in the second step of *grk* anchoring, it is not thought to be essential for *grk* transport to the anterior of the oocyte, as *grk* transport particles are unaffected in *sqd1* mutants and the RNA becomes localised to the anterior (Figure 5.1G).^{15,120, 283, 299, 317} To show that anterior transport of *grk* is indeed unaffected in *sqd1* mutant egg chambers, I measured the fluorescent intensities through a cross-section of the oocyte over time, following injection of *grk* RNA. At 0 minutes, *grk* RNA is enriched at the injection site in the middle of the oocyte (Figure 5.1A-A"). By 10 minutes, *grk* begins to enrich at the anterior of the oocyte (Figure 5.1B-B"). Between 20 and 30 minutes the signal at the anterior becomes stronger until all of the RNA is enriched at the anterior by 40 minutes (Figure 5.1C-E"). These results indicate that *grk* is transported to the anterior of the *sqd1/Df* oocyte in a similar manner to in wild type oocytes.

Conclusions The transport of injected *grk* to the anterior of the oocyte is unaffected in a *sqd1* mutant.

Injected and MS2-labelled *grk* RNA forms a strong anterior ring in a *sqd1* mutant

Despite the lack of requirement for Sqd in *grk* transport to the anterior of the oocyte, studies have shown that the second step of *grk* anchoring is disrupted for endogenous, MS2-labelled, and injected *grk* transcripts in the *sqd1* mutant, resulting in the mislocalisation of *grk*

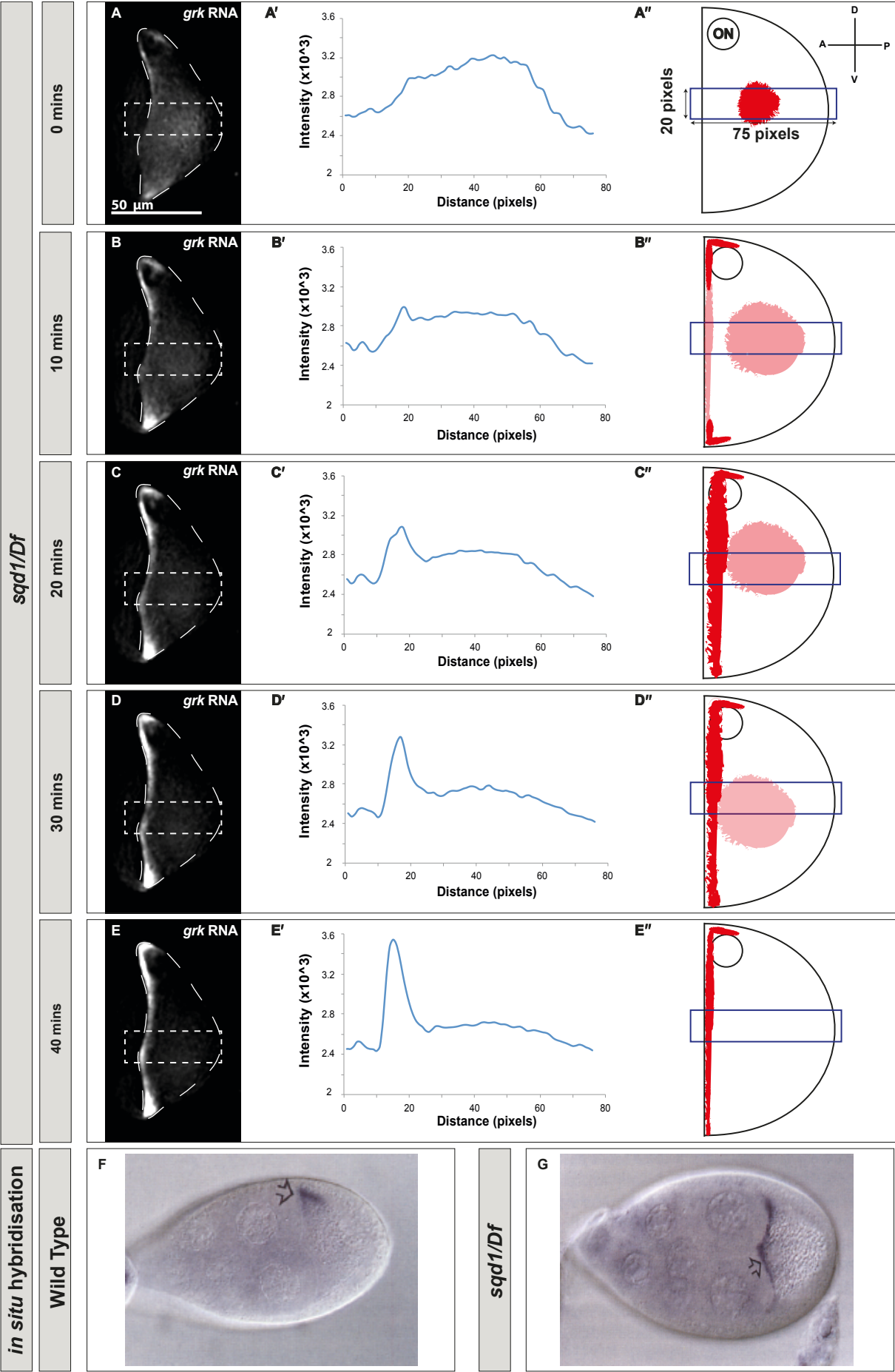


Figure 5.1. Injected *grk* RNA is rapidly transported to the anterior of the *sqd1/Df* oocyte. *grk* RNA moves from the site of injection in the centre of the oocyte (A-A'') towards the anterior between 10 and 40 minutes, enriching both at the DA and the ventral anterior, and forming a strong anterior ring (B-E'') (n=21). *in situ* hybridisation shows endogenous *grk* is restricted to the DA of wild type oocytes (F) and forms a strong anterior ring in *sqd1* mutants (G) (Neuman-Silberberg and Schupbach, 1993). Oocytes outlined with white dashed lines in A-E.

along the anterior of the oocyte.^{15, 133, 283} Therefore, Sqd is capable of anchoring both endogenous and injected *grk* RNA at the DA of the oocyte, possibly through its association with the GLS.^{54,133,193,283}

In order to dissect the regions of the GLS required for the association of Sqd, and thus *grk* anchoring, the localisation pattern of *grk* mRNA in a wild type and *sqd1* mutant background was characterised under the experimental conditions of the injection assay. Injection of *grk* RNA into wild type *OrR* flies results in localisation to the DA of the oocyte within 40 minutes, with only a faint anterior ring in 81% of injected oocytes (Figure 5.2A-A'', 5.2C). Injection of *grk* RNA into a *sqd1/Df* mutant, on the other hand, results in the formation of a stronger anterior ring and accumulation at the DA, forming a cap above the nucleus in 86% of the *sqd1/Df* oocytes (Figure 5.2B-B'', 5.2C). These results suggest differences in the distribution of *grk* RNA between wild type and *sqd1/Df* mutant oocytes, indicative of disrupted anchoring, as previously described in the literature.^{120,133,283} However, in contrast to the literature, they indicate that injected *grk* RNA in some way still associates with the DA corner of the oocyte in a *sqd1* mutant, despite the absence of Sqd.

To determine whether the DA association of *grk* seen in the *sqd1/Df* mutant oocyte is an artifact of the injection procedure, the pattern of MS2-tagged *grk* was assessed in a *sqd1/Df* mutant and wild type background. Imaging the pattern of *grk*GFP* localisation in an otherwise wild type background shows that *grk* localises to the DA of the oocyte forming a strong cap above the nucleus, and a weak anterior ring at stages early 8 to late 8 (Figure 5.3A-A''). By stage early 9, *grk* RNA is further restricted to the DA of the oocyte with much

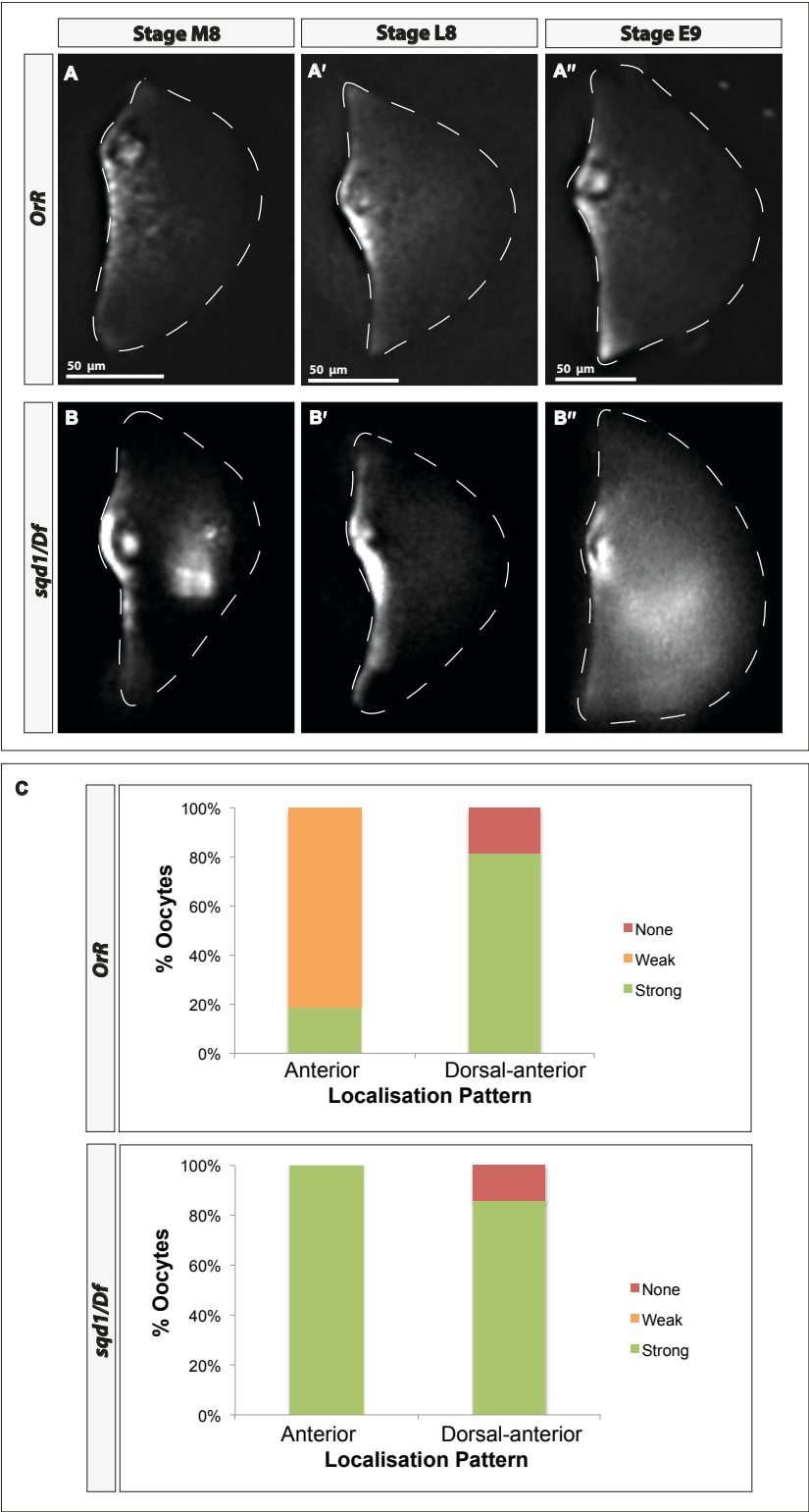


Figure 5.2. Injected *grk* RNA enriches at the DA of *OrR* and *sqd1/Df* egg chambers. (A-A'') *grk* RNA injected into the oocyte of wild type egg chambers enriches at the DA within 40 minutes (n=54). (B-B'') *grk* RNA injected into *sqd1/Df* oocytes also enriches at the DA (n=21). (C) Localisation pattern of *grk* RNA injected into wild type and *sqd1/Df* oocytes. Outline of oocytes shown with white dashed line. M8 = mid 8, L8 = late 8, E9 = early 9.

less signal along the anterior, as previously described in the literature (Figure 5.3A''').²⁸³ Flies expressing *grk*dendra* (*grk*-MS2 recombined with *MCP-dendra*; made by Rippei Hayashi and Richard Parton, unpublished), a photoswitchable chromophore similar to GFP that labels *grk* RNA in the expected distribution pattern throughout oogenesis in wild type egg chambers, crossed with *sqd1/Df* flies show that *grk* forms a much stronger anterior ring in stages early 8 to late 8 (Figure 5.3B-B'') than in wild type oocytes. Even at stage early 9, when *grk* is known to be anchored at the DA of the oocyte, a strong anterior ring can be seen in the *sqd1/Df* mutant (Figure 5.3B'''). However, *grk* also shows accumulation at the DA of 93% of the mutant oocytes (Figure 5.3B-B'''), compared with 100% in the wild type oocytes (Figure 5.3C). These data suggest that *grk* is able to associate with the DA corner in *sqd1* oocytes, but a large proportion of the RNA also enriches along the anterior, making it distinct from wild type *grk* localisation.

To confirm that the *sqd1/Df* flies are deficient in Sqd protein, I first assessed the eggs laid by the females for a *sqd1* phenotype.²⁹⁹ *grk*dendra* expressed in a wild type background fly produces eggs with normal dorsal-appendages showing that the MS2-MCP system does not disrupt the patterning of the embryo (Figure 5.4A). Eggs laid by females expressing *grk*dendra* in a *sqd1/Df* background, on the other hand, have fused dorsal appendages (Figure 5.4A') characteristic of the dorsal-ventral patterning defects previously described in *sqd1* mutants.²⁹⁹ I also carried out Western Blot analysis using a monoclonal anti-Sqd antibody (DHB 1B11)¹⁶² comparing Sqd protein levels in the wild type *OrR*, *sqd1/sqd1* and *sqd1/Df* genetic backgrounds (Figure 5.4B). Again, these data confirm that the flies are indeed deficient in Sqd protein, indicating that the accumulation of *grk* seen at the DA of the *sqd1* oocyte occurs in the absence of Sqd.

Although most studies investigating the role of Sqd in *grk* localisation have characterised *grk* as mislocalised along the entire anterior circumference of the *sqd1* mutant oocyte,^{15, 120, 133,}²⁸³ *grk* accumulation at the DA corner has also been documented in some cases.^{133, 15} In line with these studies, my results indicate that both MS2-labelled and injected *grk* RNA are able

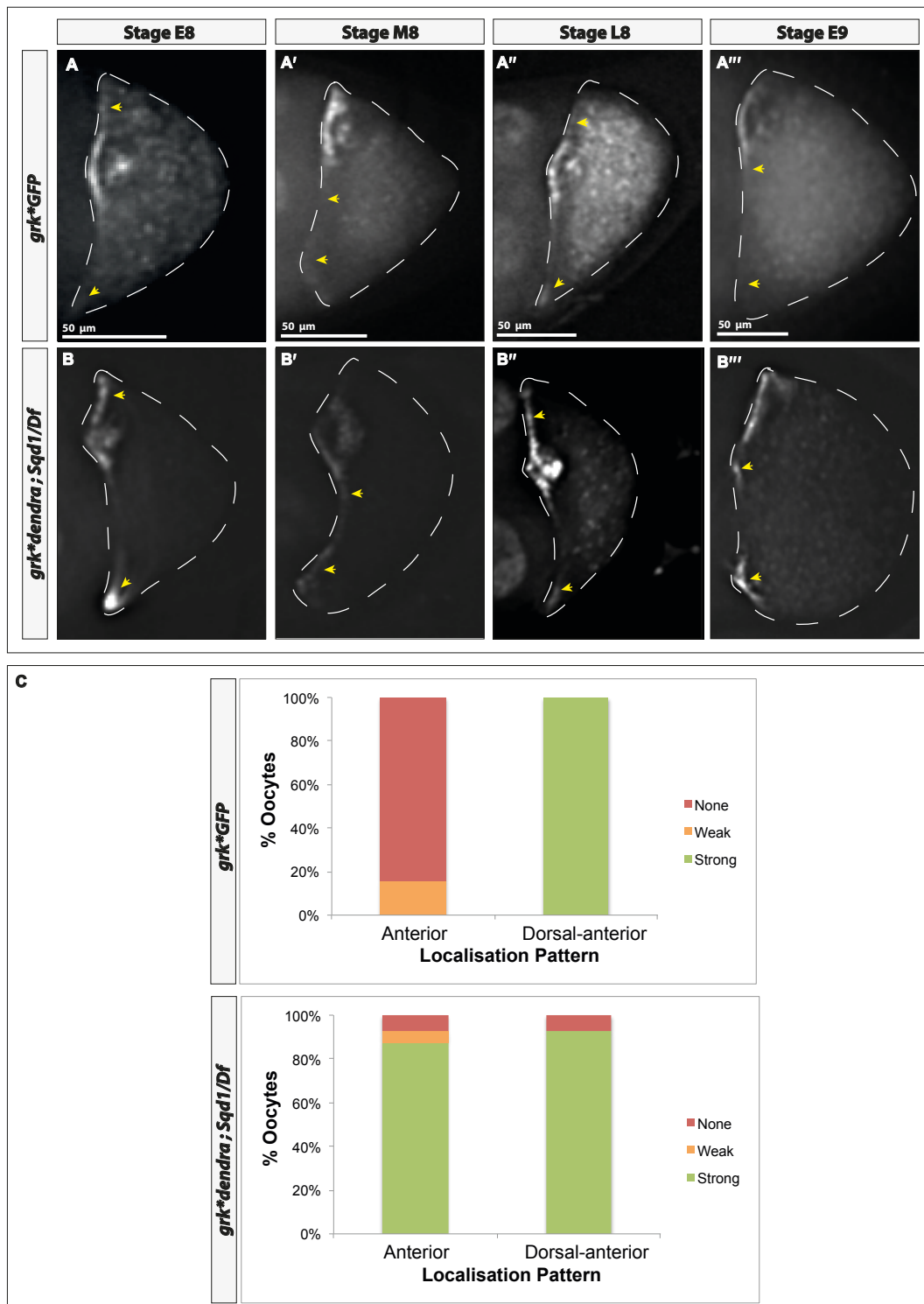


Figure 5.3. MS2-labelled *grk* RNA forms a strong anterior ring yet still associates with the DA in *sqd1* mutants. (A-A'') *grk*GFP* forms a strong DA cap and a weak anterior ring (yellow arrowheads) in a wild type background throughout stages early 8 to early 9 (n=19). (B-B'') *grk*dendra* forms a much stronger anterior ring (yellow arrowheads) during stages early 8 to early 9 when expressed in *sqd1/Df* egg chambers, but still shows association with the DA (n=40). (C) Localisation patterns of MS2-tagged *grk* RNA in a wild type genetic background and in *sqd1/Df* egg chambers. Outline of oocytes shown by white dashed lines. E8 = early 8, M8 = mid 8, L8 = late 8, E9 = early 9.

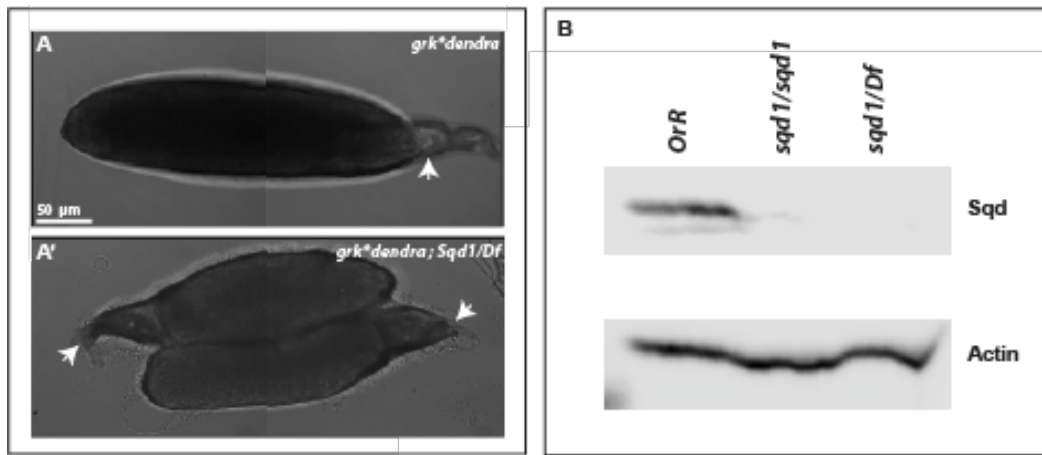


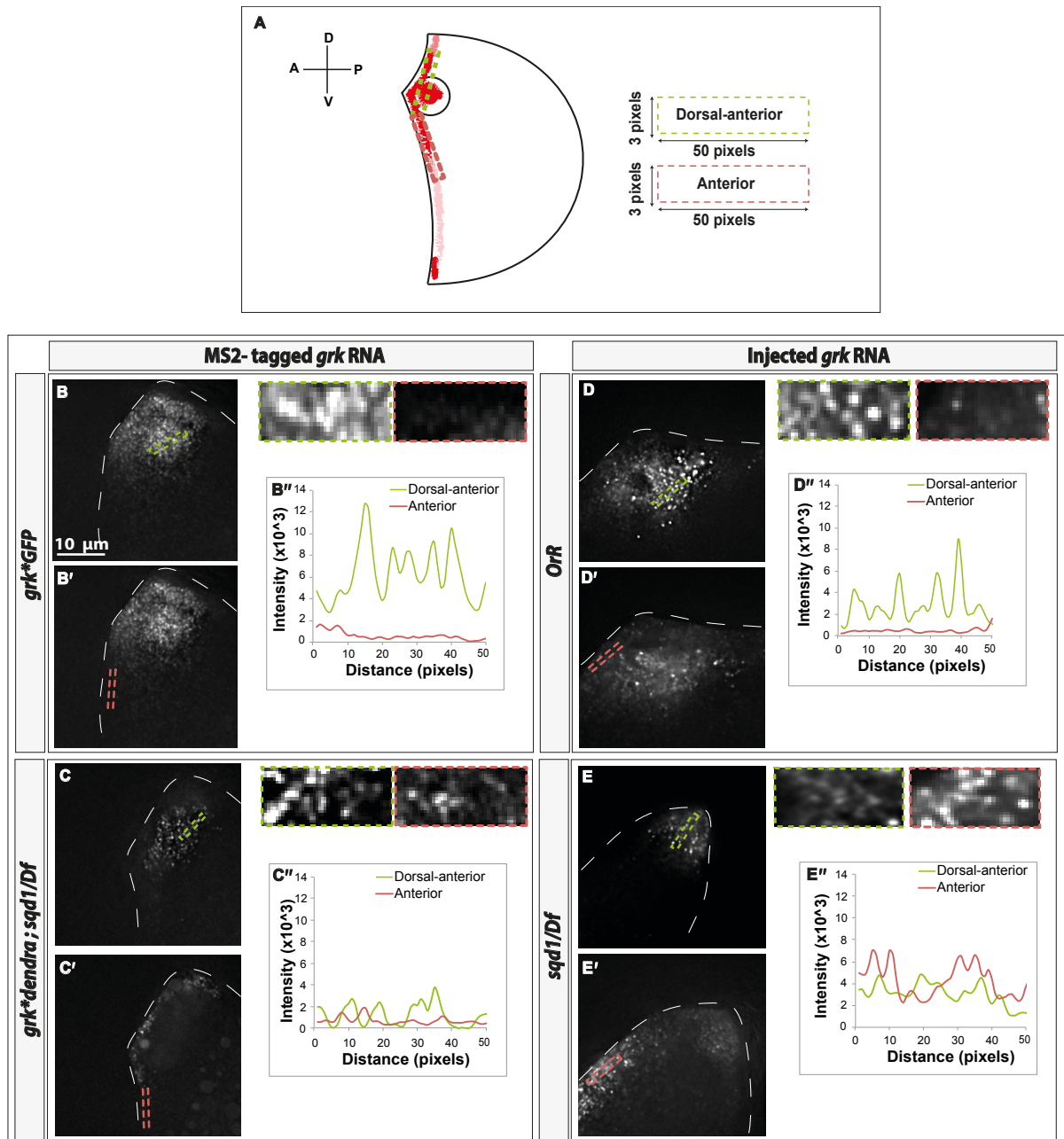
Figure 5.4. *grk*dendra;sqd1/Df* mutant flies lay dorsalisated eggs due to a depletion of Sqd protein. Eggs laid by females expressing *grk*dendra* have normal dorsal appendages (A, white arrows), whereas those laid by *grk*dendra;Sqd1/Df* females have fused dorsal appendages (A', white arrows). (B) *sqd1/sqd1* and *sqd1/Df* ovary extracts both show a depletion of Sqd protein when compared to wild type *OrR*.

to accumulate at the DA of the *sqd1* mutant oocyte, but *grk* at the DA corner does not appear to be as highly enriched over the anterior as in wild type. This difference is likely due to a decreased dosage of *grk* at the DA of *sqd1* mutants due to an increased concentration along the anterior.

Conclusions Both MS2-labelled and injected *grk* forms a strong anterior ring and accumulates at the DA of *sqd1* oocytes, despite the lack of Sqd protein. In both cases, *grk* at the DA corner of the mutant is still enriched over the anterior, although *grk* in a *sqd1* mutant is not as enriched as in wild type oocytes.

***grk* RNA forms large clusters at the anterior and DA of *sqd1* oocytes**

To further characterise the accumulation of *grk* at the anterior and DA of *sqd1/Df* oocytes, injected and MS2-tagged *grk* was imaged at high resolution and fluorescent intensities were measured through a cross-section of each region (Figure 5.5A). In wild type *OrR* oocytes *grk*GFP* is enriched at the DA of the oocyte resulting in large puncta with high fluorescent intensity peaks indicative of *grk* anchoring (Figure 5.5B, 5.5B''). In contrast, at the anterior of the oocyte only faint *grk* particles can be seen that do not cluster to form bright puncta



(Figure 5.5B', 5.5B''). In *sqd1/Df* oocytes, *grk***dendra* RNA forms bright puncta at both the DA and the anterior of the oocyte with similar fluorescent intensity peaks in both regions (Figure 5.5C-C''). Similar intensity patterns are seen for injected *grk* RNA at the DA of wild type oocytes (Figure 5.5D-D''), and both the DA and the anterior of *sqd1/Df* mutants (Figure 5.5E-E''), compared to the MS2-tagged RNAs. These results suggest that, in the absence of Sqd, *grk* forms large, bright puncta at the anterior and the DA of the oocyte that are similar to those found specifically at the DA in wild type oocytes.

To investigate whether the *grk* puncta in a *sqd1* mutant differ to those in wild type oocytes, the size of the clusters were measured at the anterior and the DA of the oocyte. In a wild type oocyte, MS2-tagged *grk* forms larger and brighter puncta at the DA than at the anterior (Figure 5.6A-A''). In *sqd1* oocytes, on the other hand, MS2-tagged *grk* forms large, bright clusters at both the DA and the anterior (Figure 5.6B-B''). The mean cluster size of *grk* RNA at the DA of wild type ($n=66$) and *sqd1/Df* ($n=77$) oocytes is comparable ($p=0.6817$) at $678 \pm 88 \text{ nm}^2$ and $637 \pm 52 \text{ nm}^2$, respectively (Figure 5.6C, Table 5.1). The mean size of *grk* clusters at the anterior of the wild type oocyte ($376 \pm 32 \text{ nm}^2$, $n=52$) is significantly smaller than at the DA of both the wild type ($p=0.004$) and the *sqd1/Df* oocytes ($p=0.0002$). Finally, the *grk* clusters formed at the anterior of the *sqd1/Df* oocytes are significantly larger ($795 \pm 73 \text{ nm}^2$, $n=65$) than those at the anterior in the wild type egg chambers ($p<0.0001$), and are comparable to those found at the DA in both wild type ($p=0.3046$) and *sqd1/Df* egg chambers ($p=0.0737$) (Figure 5.6C; Table 5.1). These data suggest that *grk* RNA forms clusters at both the anterior and the DA of *sqd1/Df* oocytes that are similar in size and brightness to those found at the DA of wild type oocytes, where the transcripts have been shown to be anchored at P bodies.

Conclusions In *sqd1* oocytes, *grk* mRNA forms large, bright clusters at both the anterior and the DA of the oocyte that are comparable to those found at the DA of wild type oocytes.

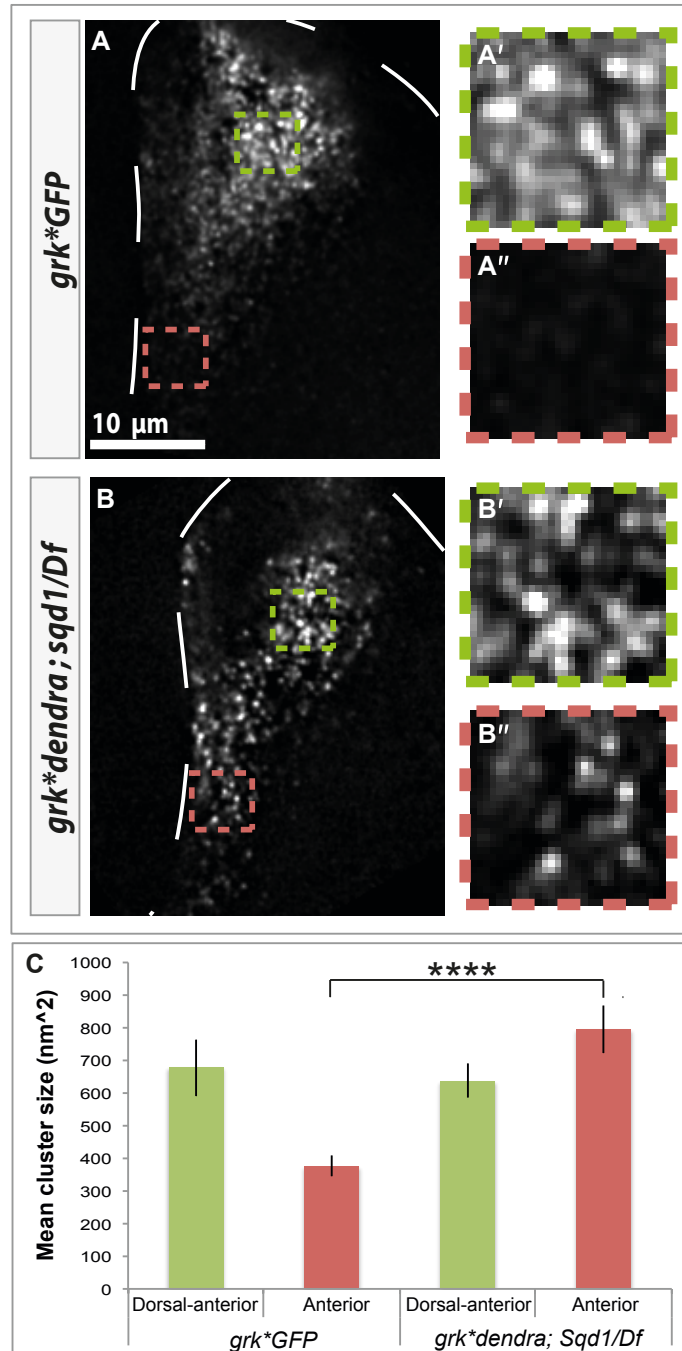


Figure 5.6. *grk* RNA clusters at the anterior of *sqd1* oocytes are similar in size to those at the DA of wild type oocytes. (A) MS2-labelled *grk* RNA forms a cap above the nucleus at the DA of the wild type oocyte (green box) and a weak anterior ring (red box). *grk* clusters at the DA (A') are brighter than those at the anterior (A''). (B) MS2-labelled *grk* RNA forms a cap above the nucleus at the DA of *sqd1/Df* mutants (green box) and a strong anterior ring (red box). *grk* puncta at the DA of the *sqd1/Df* oocyte (B') are bright and similar to those at the anterior (B''). (C) The mean size of *grk* clusters at the anterior of the *sqd1/Df* oocyte (n=65) is significantly larger (p<0.0001) than that at the anterior of wild type oocytes (n=52). *grk* clusters at the anterior of the *sqd1/Df* oocytes are similar in size to those at the DA of both *sqd1/Df* (n=77) and wild type (n=66) oocytes. Oocytes are outlined with white dashed line. Images are equally contrasted for each RNA. Error bars = S.E.M, Student's unpaired, two-tailed t-test.

Cluster type and size (n)	<i>Sqd1/Df</i> DA 637 ± 52 nm ² (77)	<i>Sqd1/Df</i> Anterior 795 ± 73 nm ² (65)	WT Anterior 376 ± 32 nm ² (52)
WT DA 678 ± 88 nm ² (66)	p=0.6817	p=0.3046	p=0.0039
WT Anterior 376 ± 32 nm ² (52)	p=0.0002	p=<0.0001	N/A
<i>Sqd1/Df</i> Anterior 795 ± 73 nm ² (65)	p=0.0737	N/A	N/A

Table 5.1. *grk* clusters at the anterior of *sqd1/Df* oocytes are significantly bigger than those at the anterior of wild type oocytes. Statistical analysis of the mean cluster size of *grk* RNA at the anterior compared to the DA of *sqd1/Df* and wild type oocytes using a Student's t-test. *grk* clusters at the DA of wild type oocytes are significantly larger than those at the anterior. Clusters at the DA and the anterior of *sqd1/Df* oocytes are comparable in size to those at the DA of wild type oocytes, and are significantly larger than those at the anterior of the wild type oocyte. Student's unpaired, two-tailed t-test. Significant values are highlighted in bold.

***grk* RNA clusters are statically anchored at the anterior of *sqd1* oocytes**

Studies investigating the dynamics of *grk* mRNA during mid-oogenesis using fluorescent recovery after photobleaching (FRAP) have shown that *grk* particles at the anterior of the oocyte are highly dynamic, whereas a high proportion of those at the DA are static.^{133, 283} In a *sqd1* mutant, *grk* particles along the anterior of the oocyte and the small amount found at the DA were shown to be highly dynamic, suggesting *Sqd* is required to statically anchor *grk* at the DA.^{133, 283} However, the large clusters of *grk* RNA identified at both the anterior and the DA of *sqd1* mutants are similar in size to those at the DA of wild type oocytes (Figure 5.6), suggesting that they may share similar characteristics.

To investigate whether the large *grk* clusters in *sqd1* oocytes show dynamics indicative of anchored *grk* RNA, fast time-lapse imaging of the *grk* clusters was carried out. At the DA of the wild type (Figure 5.7A; n=19) and *sqd1/Df* mutant (Figure 5.7B; n=37) oocytes, *grk**MS2

is heavily enriched forming large, bright clusters. A similar array of clusters is seen at the anterior of the *sqd1/Df* oocyte (Figure 5.7D; n=35). However, *grk* clusters at the anterior of the wild type oocyte form less distinct puncta (Figure 5.7C; n=19). Images taken at time zero and 10 seconds later show that the pattern of *grk* clusters does not drastically change over the time course, resulting in a clear distribution of *grk* RNA following projection over the time course at the DA of wild type oocytes (Figure 5.7A'-A'''; n=3), and both the DA (Figure 5.7B'-B'''; n=3) and anterior (Figure 5.7D'-D'''; n=3) of *sqd1/Df* oocytes. However, at the anterior of wild type oocytes, the distribution of the *grk* clusters appears to change over the time course (Figure 5.7C'-C'''; n=3), resulting in a hazy *grk* signal and only forming a few distinct puncta.

To distinguish between static and dynamic *grk* clusters over the time course, the time zero (red) and the projected time course (green) were merged. *grk* clusters that are statically anchored (yellow) should be present in all of the images captured over the time course, whereas those that are dynamic will disappear in the projected image. At the DA of both wild type (Figure 5.7A''''; n=3) and *sqd1/Df* mutant oocytes (Figure 5.7B''''; n=3), the majority of *grk* clusters are present at time zero and over the 10 second time course (yellow), indicating that they are static. Similarly, at the anterior of the *sqd1/Df* oocyte *grk* clusters also show high levels of static behaviour (Figure 5.7D''''; n=3). However, at the anterior of the wild type oocyte, most of the *grk* clusters are red or green, indicating that they were only present at time zero or appeared during the course of the time-lapse, respectively (Figure 5.7C''''; n=3). These data indicate that the *grk* clusters at the anterior of *sqd1/Df* oocytes show similar static behaviour to those anchored at the DA of the wild type oocyte.

Conclusions In *sqd1* mutants, *grk* RNA anchored at the DA of the oocyte behaves like it does in wild type. However, *grk* RNA at the anterior of the *sqd1* mutant is less dynamic than it is in wild type oocytes. These observations suggest that *grk* RNA is prematurely anchored in a *sqd* mutant, rather than being unanchored.

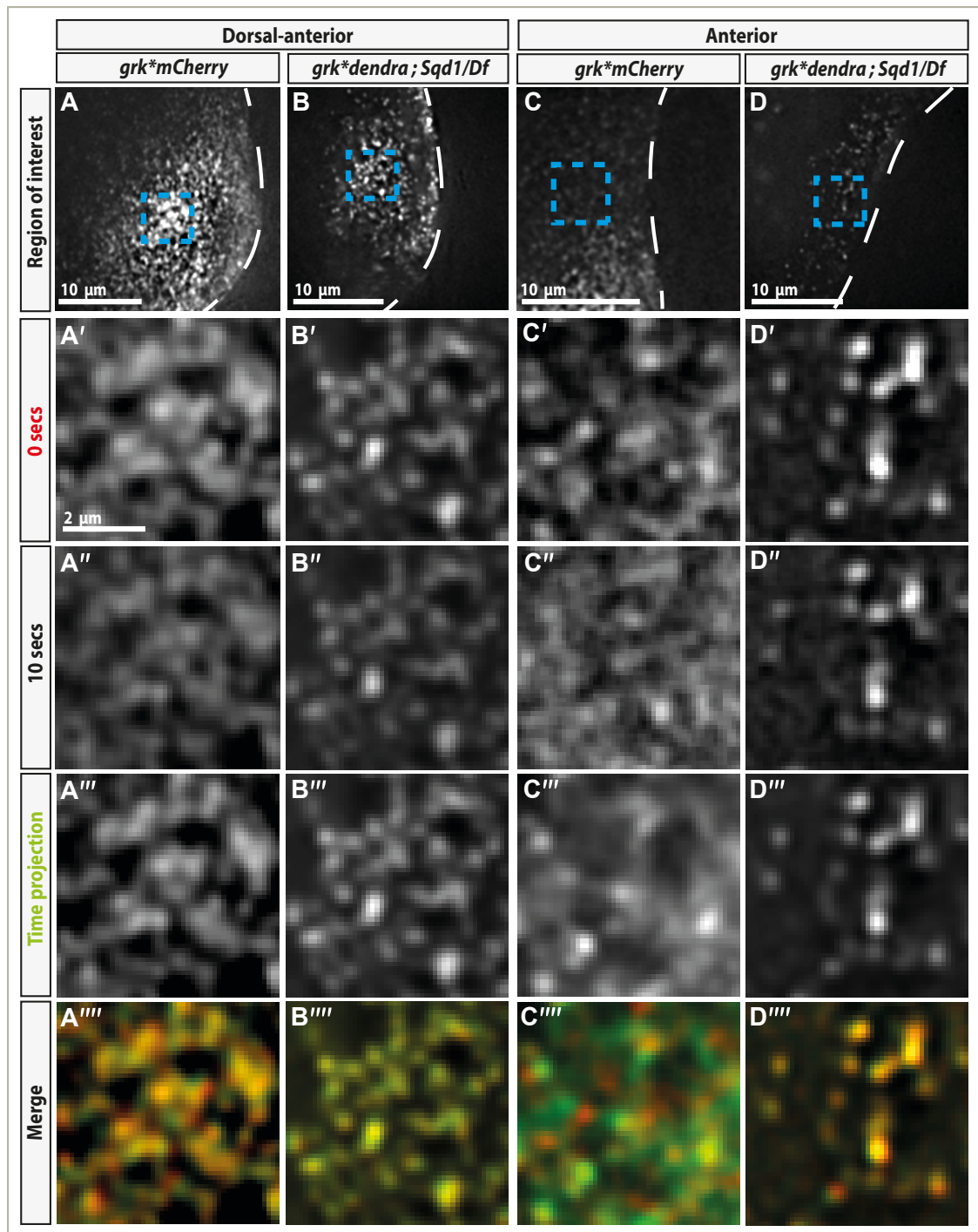


Figure 5.7. *grk* clusters at the anterior of *sqd1* oocytes show low levels of dynamics. *grk* forms large, bright clusters at the DA of wild type (A-A''') and *sqd1/Df* (B-B''') oocytes that maintain a similar distribution when projected over time. *grk* clusters at the anterior of wild type oocytes (C-C''') form less distinct puncta than those at the anterior of *sqd1/Df* oocytes (D-D''') and their distribution changes over time. (A''''-D'') Merge of the time 0 secs images (red) with the projection over the 10 secs time course (green). A large proportion of *grk* is static (yellow clusters) at the DA of wild type (A'') and *sqd1/Df* (B'') oocytes. *grk* is highly dynamic (discrete red and green clusters) at the anterior of the wild type oocyte (C'') and is much less dynamic at the anterior of *sqd1/Df* oocytes (D''). Images taken at 100x with auxiliary magnification. Oocytes are outlined with white dashed lines.

***grk* RNA interdigitates with Me31B at the anterior of *sqd1* oocytes**

In a *sqd1* mutant, *grk* RNA is localised along the anterior of the oocyte in transport particles and is ectopically translated there.^{15, 162, 299, 313} In wild type egg chambers, *grk* has been shown to anchor at the edge of P bodies at the DA of the oocyte where it is translationally activated.^{133, 193} As mislocalised *grk* at the anterior of the *sqd1* mutant oocyte shows similar dynamics to *grk* at the DA of the wild type oocyte in our assay, it is possible that the RNA is anchored and translationally activated at P bodies in a similar manner at both the anterior of mutant oocytes and at the DA of wild type oocytes.

To investigate whether *grk* RNA is associated with P bodies in a *sqd1* mutant, *grk* was injected into oocytes expressing the P body marker Me31B::GFP. At 40 minutes post-injection, *grk* RNA localises to the DA of wild type *OrR* oocytes expressing Me31B::GFP (Figure 5.8A-A'). Merging the two channels for *grk* RNA and Me31B::GFP shows that *grk* interdigitates with Me31B::GFP at the DA of the oocyte (Figure 5.8A'', yellow box), but that the *grk* particles that form the faint anterior ring do not associate with Me31B (Figure 5.8 A'', pink box). In a *sqd1/Df* background, on the other hand, *grk* enriches at the DA of the oocyte but also shows a strong anterior ring, distributing in a similar manner to Me31B::GFP (Figure 5.8B-B'). Merging the two channels shows that *grk* RNA at the DA of the *sqd1* oocyte interdigitates with Me31B (Figure 5.8B'', yellow box), as does *grk* RNA at the anterior of the oocyte (Figure 5.8B'', pink box). These results suggest that *grk* RNA is docked at the edge of P bodies at the DA and anterior of *sqd1* oocytes in a similar manner to at the DA of wild type oocytes. This is in agreement with the finding that *grk* is less dynamic at the anterior of *sqd1* mutant oocytes than in wild type (Figure 5.7), and the fact that *grk* is translated at the anterior of *sqd1* mutants.²⁹⁹ However, it is also possible that the inherently lower levels of *grk* RNA at the anterior of the wild type oocyte results in less interdigitation with Me31B due to differences in the concentration of the RNA. In order to distinguish between these possibilities it would be necessary to normalise the level of interdigitation with Me31B to the RNA intensity in each region of the oocyte.

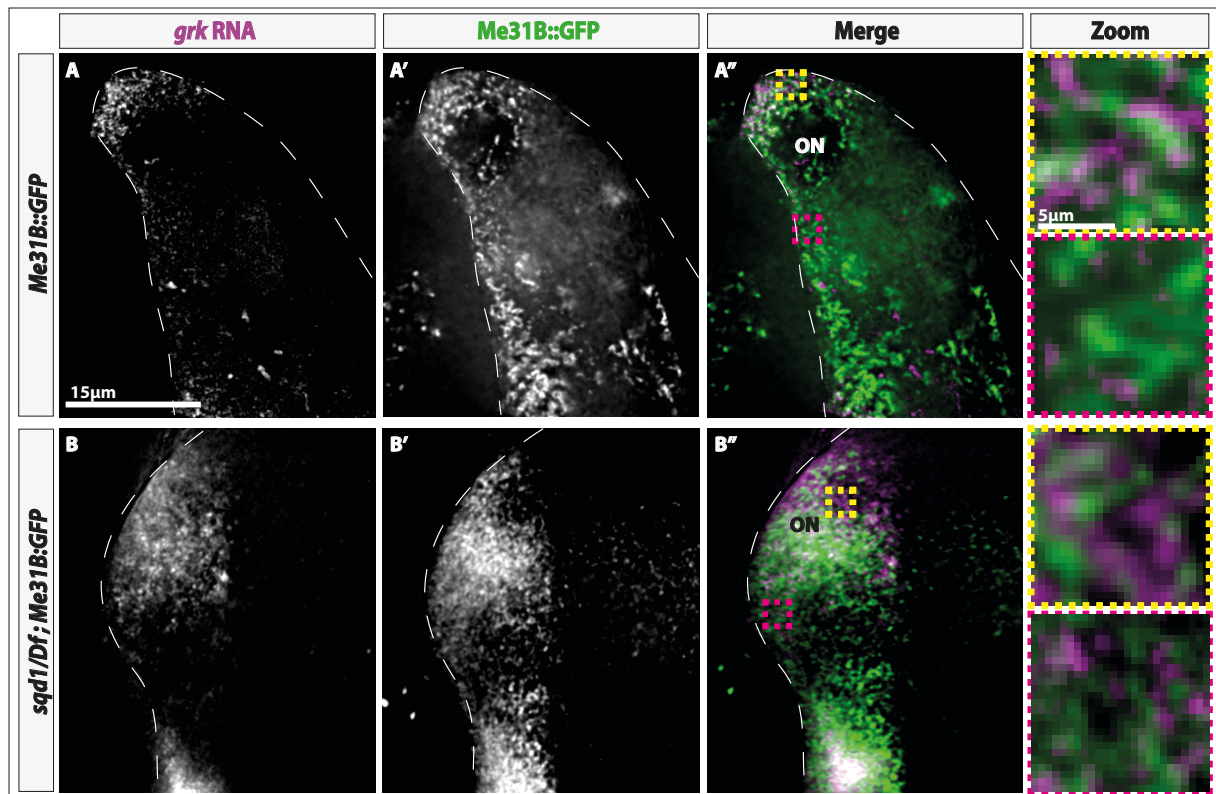


Figure 5.8. Injected *grk* RNA interdigitates with Me31B::GFP at the DA and the anterior of *sqd1/Df* oocytes. (A-A''') *grk* RNA injected into oocytes expressing Me31B::GFP localises within 40 minutes to the DA (yellow box, n=17) where it interdigitates with Me31B (n=3), whereas the low levels of *grk* RNA at the anterior of the oocyte do not associate with Me31B::GFP (pink box). (B-B''') *grk* RNA injected into *sqd1/Df* oocytes expressing Me31B::GFP localises within 40 minutes to both the DA (yellow box) and the anterior (pink box), and interdigitates with Me31B::GFP in both regions (n=3). Images contrasted for best display. Oocytes outlined with white dashed line.

Conclusions *grk* RNA is docked at the edge of P bodies at both the DA and the anterior of the *sqd1* oocyte.

Sqd is enriched at the oocyte anterior and is not recruited to injected *grk* RNA

The experiment showing that *grk* RNA is docked at the edge of P bodies in a *sqd1* mutant suggests that *grk* is still capable of anchoring in the absence of Sqd, but in the wrong place. In order to re-investigate the role of Sqd in *grk* anchoring, the distribution of Sqd in the oocyte was assessed. If Sqd plays a role in *grk* anchoring at the DA of the oocyte, one would expect the protein to be enriched there. However, in stage early 8 to late 9 egg chambers expressing Sqd::GFP, Sqd is highly enriched within the nuclei of the follicle cells, NCs and oocyte. Sqd::GFP is also weakly localised at the anterior of the oocyte, but shows no enrichment at the DA (Figure 5.9A-A’’).

Previous studies have shown that Sqd::GFP is recruited to injected *grk* RNA at the DA of the oocyte at 80 minutes post-injection.¹³³ To investigate Sqd::GFP recruitment to the DA in my shorter screening assay, *grk* was injected at high concentrations (600ng/ul) into the oocyte and recruitment was assessed after the reduced time of 40 minutes. Highly concentrated *grk* RNA localises to the DA of the oocyte and forms a strong anterior ring (Figure 5.9B-B’’), as previously described in the literature for both injected and over expressed endogenous *grk* RNA.¹⁹³ Sqd::GFP is not recruited to *grk* RNA at the DA cap (Figure 5.9C-C’, 5.9D-D’) after 40 minutes of injection, but shows low levels of recruitment to *grk* RNA at the anterior of the oocyte (Figure 5.9C’’-C’’’, 5.9D-D’’’). These results suggest that Sqd may function at the anterior of the oocyte to control *grk* localisation, rather than at the DA, possibly inhibiting the ectopic anchoring of *grk* RNA at the anterior of the oocyte.

Conclusions Injecting high concentrations of *grk* RNA into the oocyte results in DA enrichment and a strong anterior ring. Sqd is not recruited to the DA of the oocyte but enriches at the anterior following the injection of highly concentrated *grk* RNA.

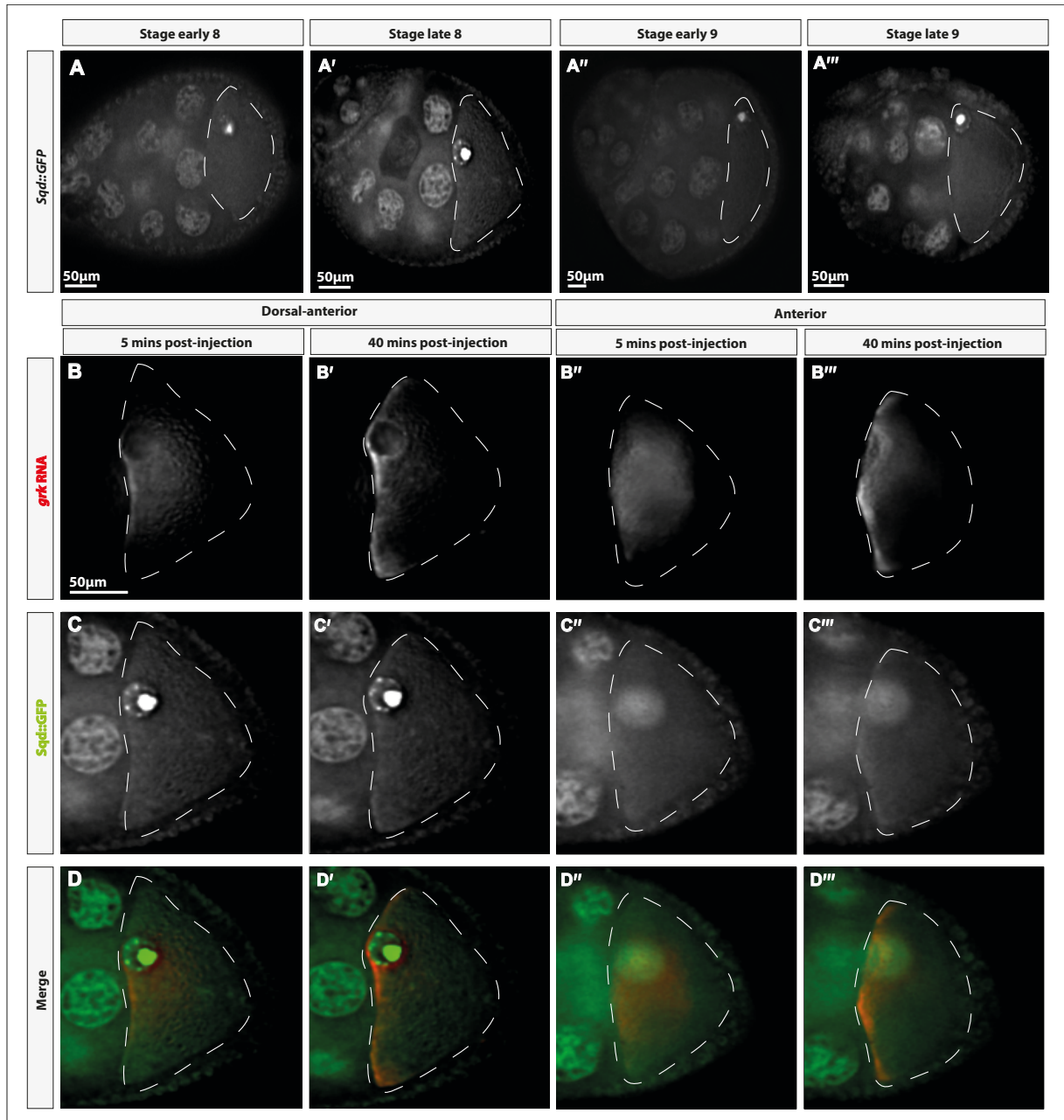


Figure 5.9. Sqd is weakly recruited to the anterior of the oocyte following *grk* RNA injection. (A-A''') Sqd::GFP is enriched in the nuclei of the nurse cells, follicle cells and the oocyte and is weakly enriched at the anterior of the oocyte (n=8). (B-B''') *grk* RNA injected at high concentrations localises to the DA of the oocyte and forms a strong anterior ring (n=4). (C-C''') Sqd::GFP is not recruited to the DA of the oocyte following *grk* injection (n=4) but becomes enriched at the anterior (n=3). (D-D''') Merged images of *grk* RNA and Sqd::GFP distribution at the DA and anterior of the oocyte at 5 minutes and 40 minutes post-injection of *grk*. Images taken in different focal planes to capture the DA and the anterior of the oocyte. Oocytes are outlined with white dashed lines in B-B'''.

***grk* RNA injected near the oocyte nucleus rescues the *sqd1* phenotype**

Taken together, the data presented so far suggest that Sqd may prevent inappropriate anchoring and association of *grk* with P bodies away from the DA of the oocyte. Therefore, I would predict that *grk* mRNA is anchored both at the DA and the anterior of *sqd1/Df* oocytes. To investigate whether the ectopic localisation of *grk* RNA at the anterior of the oocyte can be rescued in a *sqd1/Df* mutant by restricting the localisation of *grk*, RNA was injected near the nucleus of the oocyte and compared with injection in the centre (Figure 5.10A). *grk* RNA injected in the centre of wild type *OrR* oocytes localises to the DA and forms a weak anterior ring within 40 minutes (Figure 5.10B). In a *sqd1/Df* mutant, *grk* injected in the centre of the oocyte enriches at the DA of the oocyte but also forms a strong anterior ring (Figure 5.10C). However, when *grk* is injected near the nucleus of *sqd1/Df* oocytes it shows a strong enrichment at the DA within 40 minutes with only a weak anterior ring in 38% of oocytes (Figure 5.10D-D'). In 63% of cases, *grk* was highly restricted in a cap at the DA of the oocyte and showed no anterior ring (Figure 5.10E-F). These results indicate that *grk* may be anchored at the DA of the oocyte in a *sqd1/Df* mutant in a similar manner to in wild type oocytes, and that the anterior mislocalisation of *grk* can be rescued by DA anchoring in these mutants. Moreover, this result further supports the hypothesis that Sqd plays a role in inhibiting inappropriate *grk* anchoring at the anterior of the oocyte, rather than driving *grk* anchoring at the DA. Either way, regulation of *grk* anchoring is likely to require specific molecular interactions between Sqd and the GLS.

Conclusions Injecting *grk* RNA near the DA of the oocyte rescues the *sqd1* anchoring phenotype.

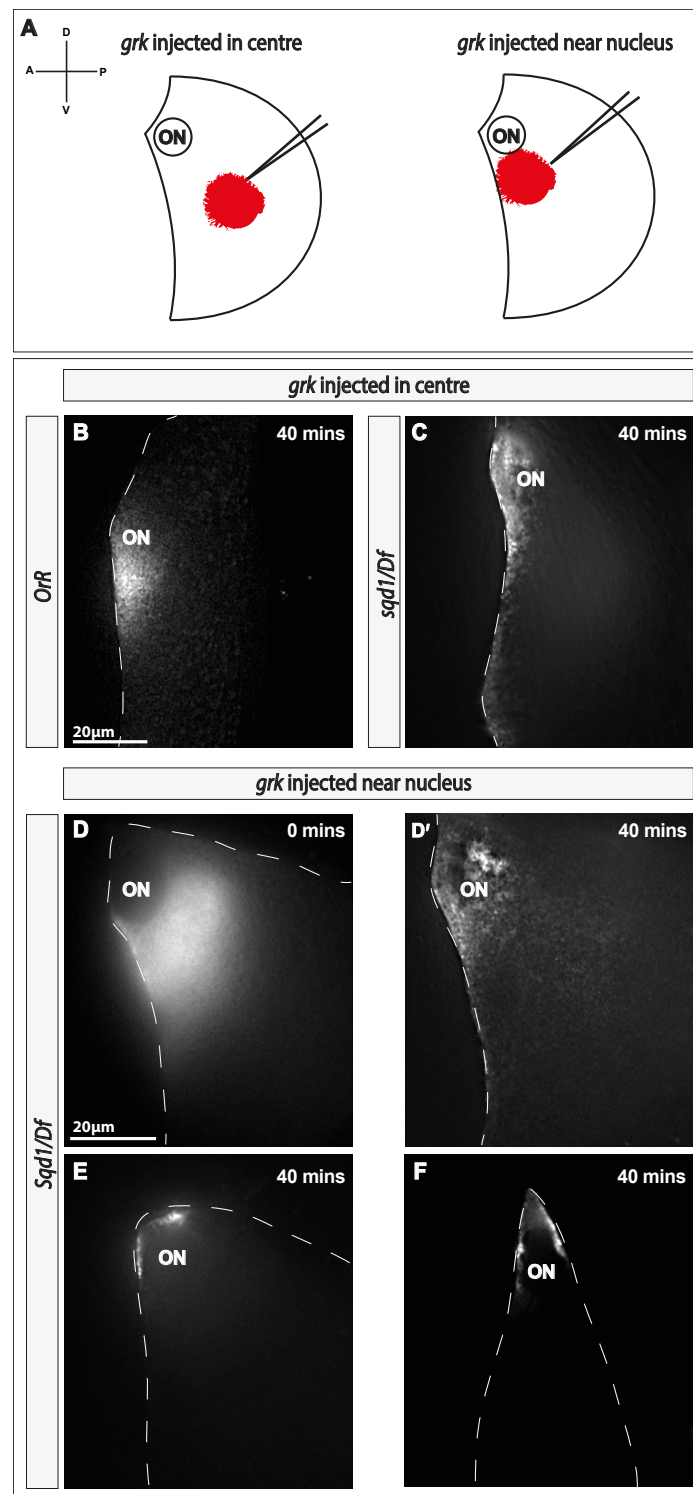


Figure 5.10. Injecting *grk* RNA near the nucleus rescues the *sqd1* mislocalisation phenotype. (A) Schematic representation of the position of *grk* injections into the oocyte. (B) *grk* RNA localises to the DA of the oocyte and forms a weak anterior ring when injected in the centre of a wild type *OrR* oocyte ($n=54$). (C) *grk* RNA enriches at the DA and forms a strong anterior ring when injected in the centre of a *sqd1/Df* oocyte ($n=21$). (D-D'') *grk* RNA injected near the nucleus of a *sqd1/Df* oocyte enriches at the DA and forms only a weak anterior ring, or no anterior ring (E-F) after 40 minutes ($n=8$). Images taken at 100x magnification and are contrasted to best show the RNA.

SELEX data can infer which regions of the GLS may be important for Sqd binding

In order to inform intelligent design of putative GLS anchoring mutants, the binding preferences of the Sqd (Hrp40) protein obtained from SELEX experiments was used to predict Sqd binding sites within the GLS.³⁹⁸ The binding motif generated from SELEX shows that Sqd preferentially binds RNA sequences with an adenine followed by a thymine in positions one and two, following by two nucleotides with less specificity, and ending with guanines and cytosines (Figure 5.11A). Searching for similar sequence motifs within the GLS reveals four potential binding sites for Sqd throughout the signal (Figure 5.11B, highlighted in pink).

Conclusions RNA-binding studies predict four putative Sqd binding sites within the GLS.

Previous GLS transport mutants can be used to inform anchoring mutant design

Based upon knowledge gained from Chapter 4 of this thesis, two of the putative Sqd binding sites (PSB1 and PSB4) fall within a region of the GLS that I have shown to not be essential for *grk* transport in the oocyte (Figure 4.6; Table 4.2). Truncation of the proximal third of the GLS in mutant M1 (Figure 5.12A) results in a slight reduction in the efficiency of *grk* transport, but enrichment at the DA of the oocyte is still seen in 77% of oocytes injected, with very little signal along the anterior of the oocyte (Figure 5.12E-E'', 5.12F). In addition, despite poor transport of the further truncated mutant M1-7, the RNA is able to enrich at the DA of the oocyte in 45% of oocytes injected (Figure 5.12D-D'', 5.12F), with only a weak anterior ring similar to that seen for wild type *grk* RNA (Figure 5.12B-B'', 5.12F). Further truncations to the proximal end of the signal in M1-5 result in only 23% of oocytes injected showing DA enrichment (Figure 5.12C-C'', 5.12F), although this is likely due to an inability of the RNA to be efficiently transported.

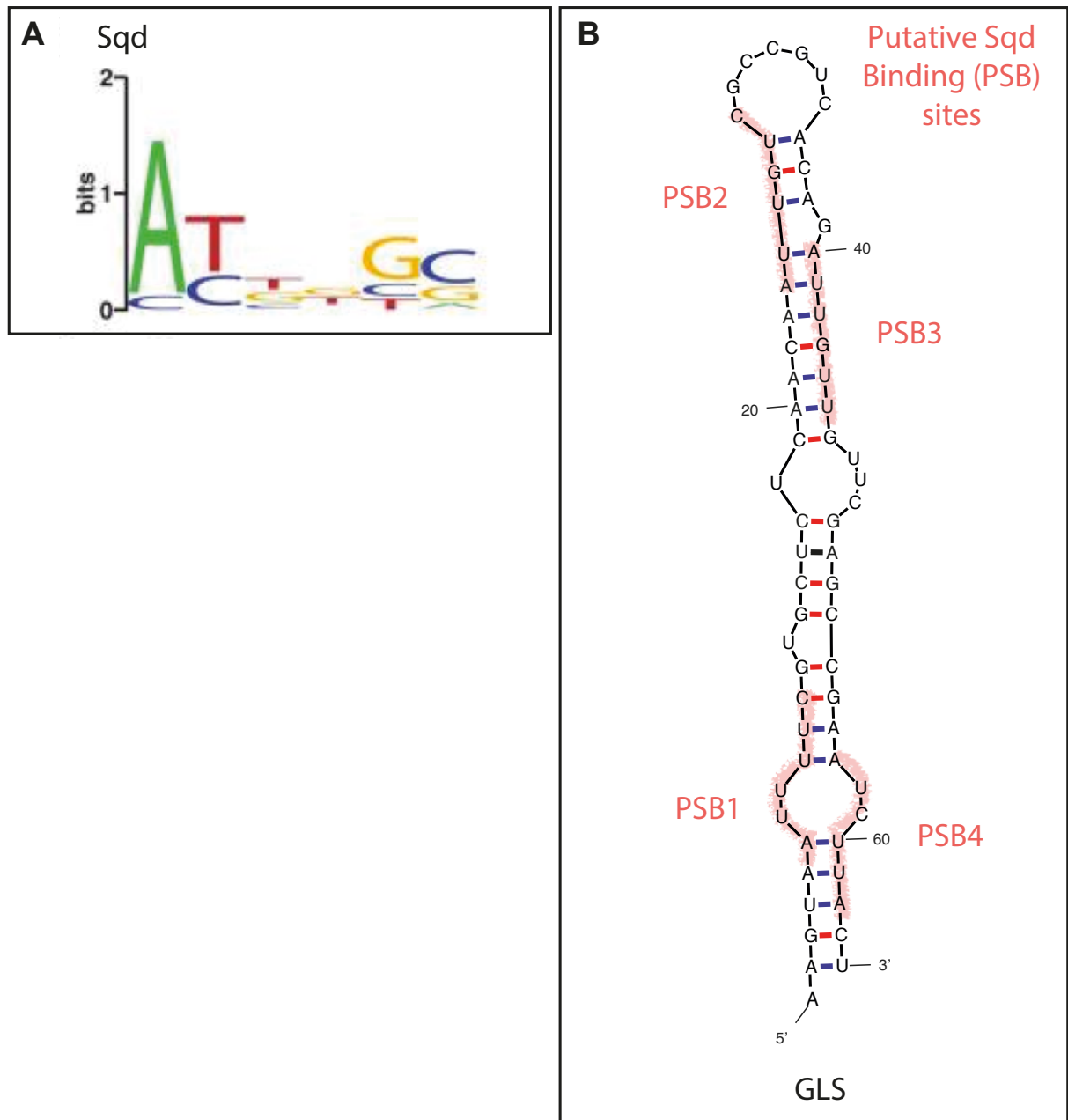


Figure 5.11. Predicted Sqd binding sites within the GLS. (A) Sequence logo generated from SELEX binding studies using hrp40 (Sqd) (Blanchette et al. 2009). The relative size of the letters represents their frequency in the sequences. (B) UNAFold predicted secondary structure of the GLS with the four putative Sqd binding sites predicted from SELEX data mapped in pink.

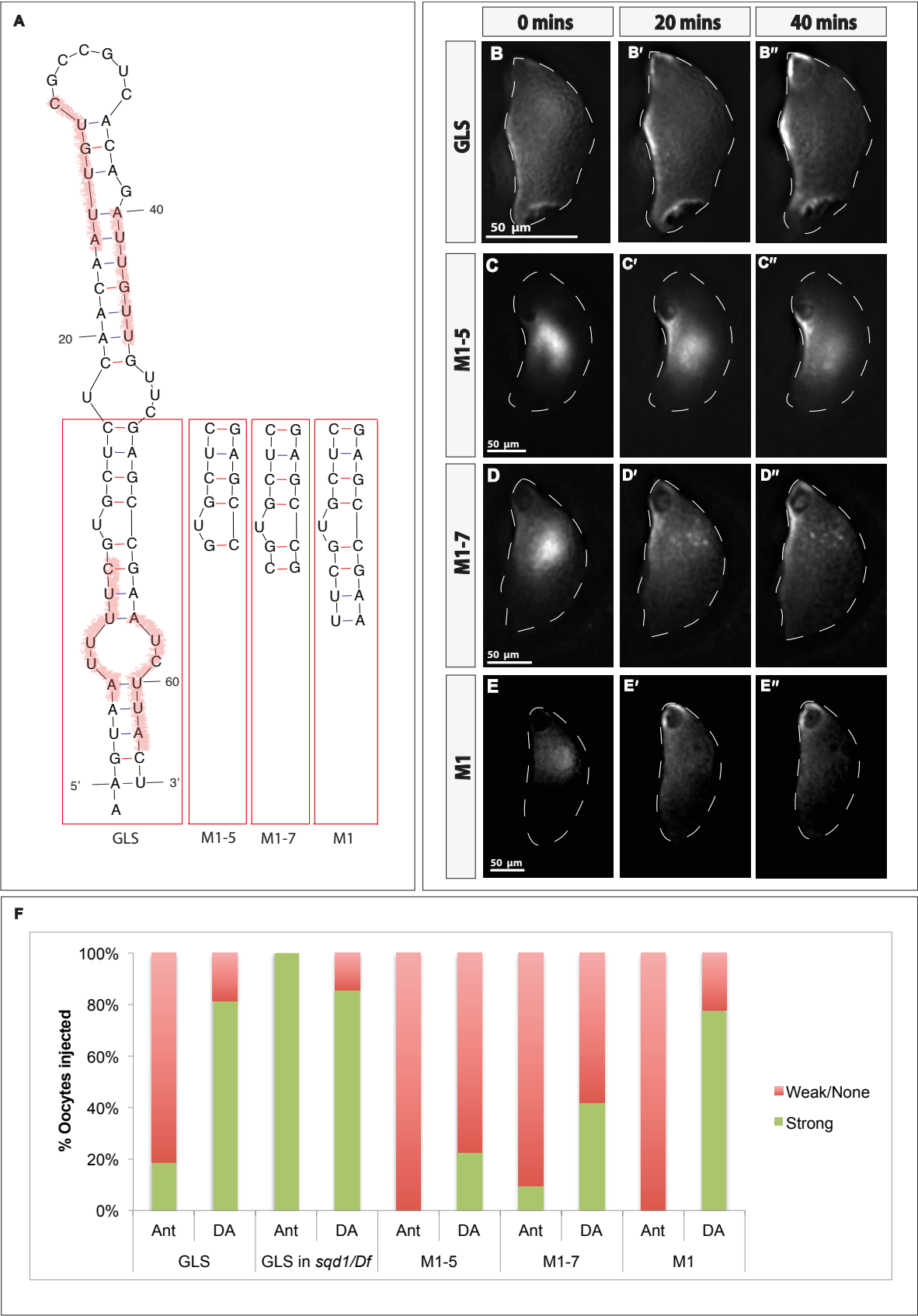


Figure 5.12. The two putative Sqd binding sites in the proximal third of the GLS are not essential for *grk* anchoring. (A) UNAFold predicted secondary structures of the GLS and mutants containing truncations within the proximal third of the signal. Putative Sqd binding sites are highlighted in pink. (B-B'') The GLS drives *grk* RNA to the DA of the oocyte within 40 minutes of injection (n=54). M1-5 (C-C'', n=31), M1-7 (D-D'', n=31) and M1 (E-E'', n=18) also drive *grk* to the DA of the oocyte within only a weak ring of *grk* at the anterior. (F) Distribution of *grk* RNA at the anterior and DA of the oocyte at 40 minutes post-injection showing that truncations to the proximal third of the GLS have little effect on the enrichment of *grk* RNA at the anterior of the oocyte, despite their reduction in transport efficiency (see Chapter 4). Oocytes outlined with white dashed line.

As a result, it is likely that the two putative Sqd binding sites PSB1 and PSB4 in the proximal third of the signal are redundant for restricting *grk* anchoring to the DA of the oocyte.

Conclusions Two putative Sqd-binding sites in the proximal third of the GLS are not required for restricting *grk* anchoring to the DA of the oocyte.

A selection of GLS-like signals from other species show aberrant anchoring in the oocyte

In Chapter 4 I showed that several of the GLS-like signals identified in other *Drosophilids* and Dipterans were able to drive the transport of *grk* RNA in the *D. melanogaster* oocyte. However, it is possible that sequence or structural variations in these signals (Figure 5.13A) have an affect on the association of the RNA with the anchoring machinery. To investigate whether the GLS-like signals from other species are capable of restricting *grk* anchoring to the DA of the oocyte, I assessed the degree of anterior localisation of the RNA in the *D. melanogaster* oocyte at 40 minutes post-injection. In a *sqd1* mutant, *grk* RNA forms a strong anterior ring in 100% of oocytes injected (Figure 5.13B, 5.13I), whereas in a wild type background, *grk* is restricted to the DA within 40 minutes in 84% of cases, leaving a strong anterior ring in only 19% of oocytes injected (Figure 5.13C, 5.13I). *D. takahashii*, containing an enlarged hairpin loop and a constricted internal loop 2 (Figure 5.13A), localises to the DA in 81% of cases, with 46% showing a strong anterior ring, indicating that this signal may have

a defect in *grk* anchoring (Figure 5.13D, 5.13I). The *D. biarmipes* GLS-like signal is able to drive *grk* transport to the anterior of the oocyte but the RNA remains highly enriched at the anterior in 86% of oocytes injected, with 72% also showing DA enrichment (Figure 5.13E, 5.13I). As a result, both the *D. takahashii* and the *D. biarmipes* signals show increased anterior RNA localisation suggesting they may have disrupted *grk* anchoring. Testing the localisation pattern of some of the GLS-like signals identified in species outside of the *Drosophilidae* family revealed that the *M. domestica* (Figure 5.13F), *B. cucurbitae* (Figure 5.13G) and *C. capitata* (Figure 5.13H) GLS-like signals also all drive higher levels of anterior *grk* RNA enrichment at 60%, 41% and 44%, respectively (Figure 5.13I).

Conclusions GLS-like signals from other species show increased *grk* enrichment at the anterior of *D. melanogaster* oocytes following injection.

The GLS-like signal from *C. capitata* forms large clusters that interdigitate with Me31B at the anterior of the oocyte

The GLS-like signal with the lowest conservation phylogenetically is that from *C. capitata* (Figure 3.4B). As a result, this GLS-like signal was chosen for further characterisation to investigate sequences or structures that may be required for restricting *grk* anchoring to the DA of the oocyte, and therefore Sqd binding. Unlike Egl, Sqd is not recruited to the apical cortex of the embryo following *grk* injection (data not shown). As a result, alternative *grk* anchoring assays were required to test the association of *grk* mutants with Sqd more indirectly.

To investigate whether the GLS-like signal from *C. capitata* drives similar anchoring characteristics to *grk* in a *sqd1* mutant, the RNA was injected into the oocyte and assessed for clustering along the anterior. *grk* RNA injected into *sqd1/Df* egg chambers forms large, bright puncta at the DA and anterior of the oocyte (Figure 5.14B-B''), similar to those anchored at the DA of wild type oocytes (Figure 5.14A-A''). Injected *C. capitata* RNA also

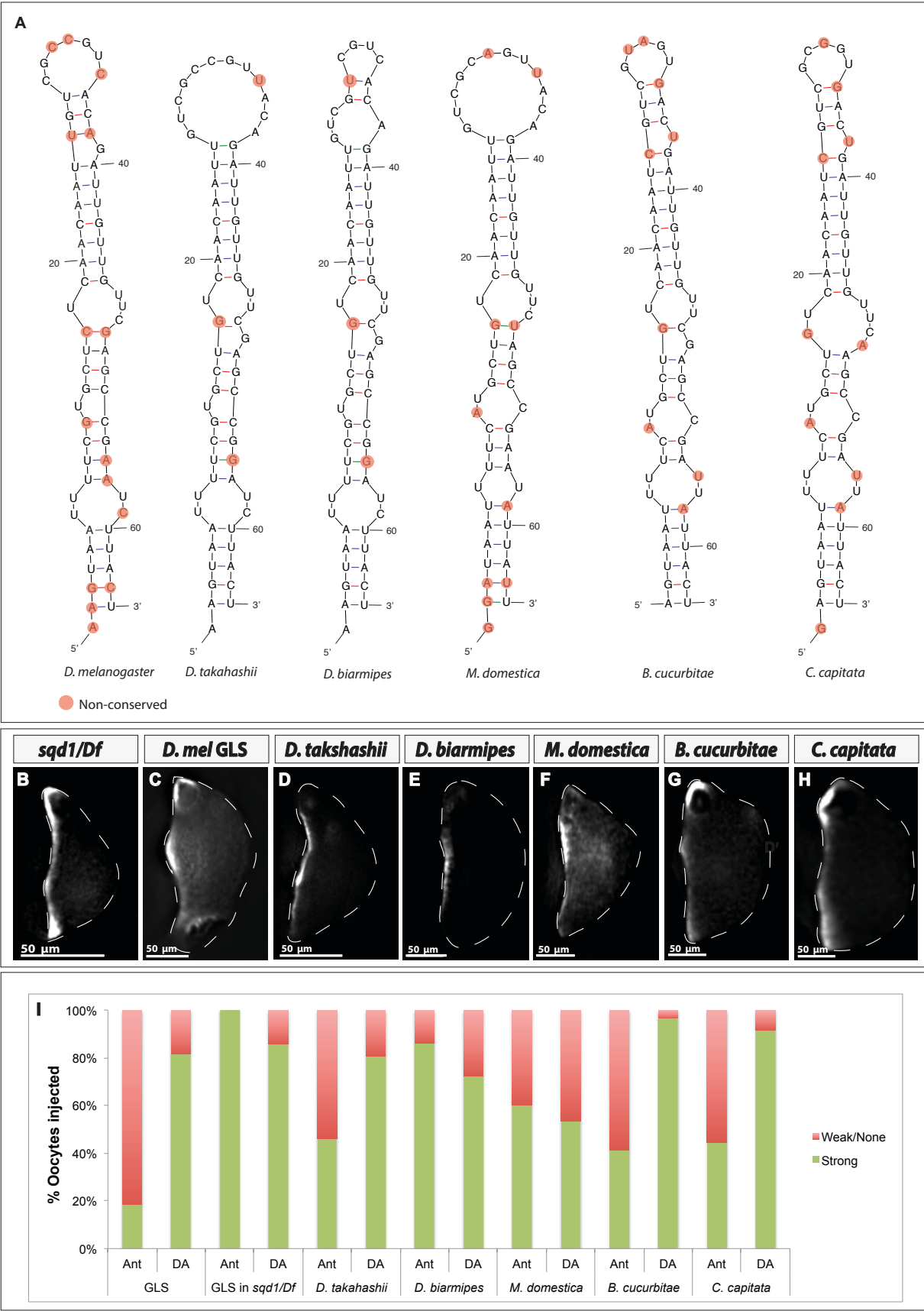


Figure 5.13. GLS-like signals from other species show increased levels of *grk* at the anterior of the *D. melanogaster* oocyte following injection. (A) UNAFold predicted secondary structures of the *D. melanogaster* GLS and GLS-like signals from other species. Nucleotides that diverge from the *D. melanogaster* GLS are highlighted in orange and mapped onto the *D. melanogaster* signal. (B) *grk* RNA injected into *sqd1/Df* oocytes enriches at the DA and forms a strong anterior ring within 40 minutes (n=21). (C) The *D. melanogaster* GLS drives *grk* localisation to the DA of the oocyte with only a weak anterior ring (n=54). The GLS-like signals from *D. takahashii* (D, n=26), *D. biarmipes* (E, n=29), *M. domestica* (F, n=30), *B. cucurbitae* (G, n=29) and *C. capitata* (H, n=36) all cause *grk* to form a strong anterior ring after 40 minutes following injection into the *D. melanogaster* oocyte. (I) Distribution of *grk* RNA at the anterior and DA of the oocyte at 40 minutes post-injection showing that the GLS-like signals from other species drive high levels of *grk* enrichment at the anterior of the oocyte; similar to that found in a *sqd1* mutant. Oocytes are outlined with dashed white lines.

forms these large, bright puncta at both the DA and anterior of the oocyte (Figure 5.14C-C''), suggesting the GLS-like signal from *C. capitata* may cause similar mislocalisation and clustering of *grk* RNA to that seen in *sqd1/Df* mutants.

To determine whether the clusters identified in *grk* RNA containing the *C. capitata* signal show similar associations with P bodies to *grk* RNA in a *sqd1/Df* mutant, the RNA was injected into oocytes expressing the P body marker Me31B::GFP. At the DA of the oocyte, *grk* RNA containing the *C. capitata* GLS-like signal interdigitates with Me31B (yellow box, Figure 5.14D-D''), as seen for wild type *grk* RNA (Figure 5.8A-A''). At the anterior (pink box) of the oocyte *grk* containing the *C. capitata* GLS-like signal shows signs of interdigitation with Me31B (Figure 5.14D-D''), as previously seen for *grk* RNA injected into the *sqd1/Df* oocyte (Figure 5.8B-B''). However, these preliminary data should be confirmed taking into account differences in the levels of RNA in each region.

Taken together, these results indicate that the GLS-like signal from *C. capitata* drives aberrant anchoring of *grk* RNA that phenocopies the *sqd1/Df* mutant, forming large clusters that are docked at the edge of P bodies. As a result, it is likely that the *C. capitata* GLS-like signal has impaired Sqd binding, resulting in anchoring of the RNA at both the DA and the anterior of the oocyte.

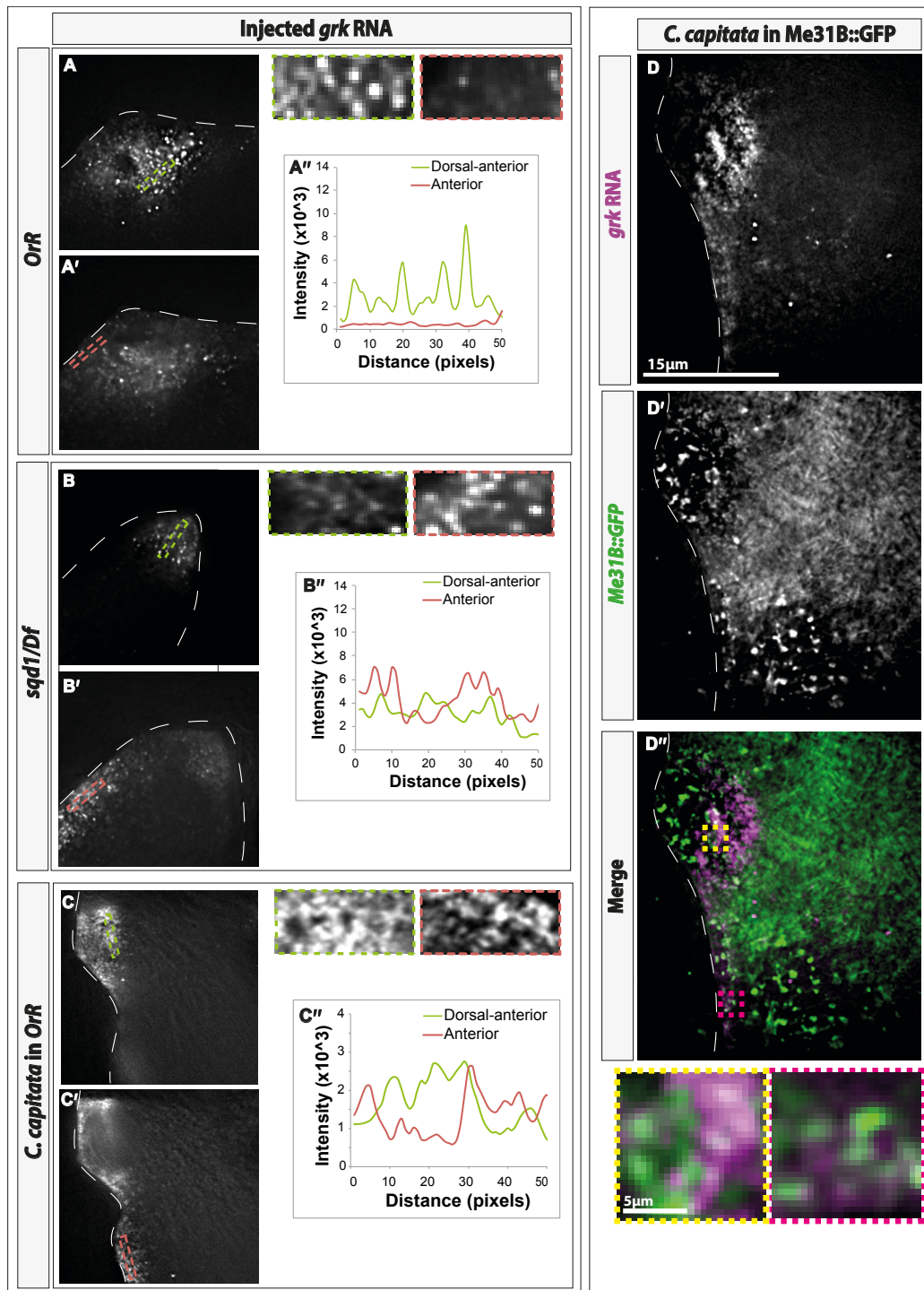


Figure 5.14. The GLS-like signal from *C. capitata* causes *grk* RNA to dock at the edge of P bodies in large clusters at the anterior and DA of the oocyte. (A-A'') Wild type injected *grk* RNA localises specifically to the DA of the oocyte where it forms large clusters upon anchoring (n=4). (B-B'') In a *sqd1* mutant, injected *grk* RNA forms large clusters at the DA of the oocyte and along the anterior, indicative of ectopic anchoring (n=4). (C-C'') The GLS-like signal from *C. capitata* causes *grk* to form large clusters at the DA and along the anterior of *D. melanogaster* oocytes in a similar manner to a *sqd1* mutant (n=5). (D-D'') In oocytes expressing Me31B::GFP, the GLS-like signal from *C. capitata* drives *grk* localisation to the DA (yellow box) and the anterior (pink box) where it interdigitates with Me31B::GFP, suggesting ectopic anchoring of the RNA (n=4). All images taken at 100x magnification. Images of the DA and anterior in A-C' taken in the same oocyte but at different focal planes. Oocytes outlined with white dashed line.

Conclusions The GLS-like signal from *C. capitata* is inappropriately anchored at the anterior of the *D. melanogaster* oocyte.

The distal third of the *C. capitata* GLS-like signal is responsible for aberrant *grk* localisation and anchoring

As *grk* RNA containing the GLS-like signal from *C. capitata* displays similar characteristics to wild type *grk* RNA in a *sqd1* mutant, it is likely that divergence in the sequence or structure of the signal has reduced the propensity of Sqd to bind the RNA. In order to determine which regions of the *C. capitata* GLS-like signal result in aberrant *grk* localisation and anchoring, mutagenesis of the RNA was carried out, replacing sections of the *C. capitata* signal with that of the wild type *D. melanogaster* GLS. The resulting UNAFold secondary structures of the mutants show that, following switching of different regions of the signal, they are still predicted to fold in the same way as their respective wild type signals (Figure 5.15A). As previously shown, the wild type GLS drives *grk* transport to the anterior of the oocyte with a weak ring at the anterior (Figure 5.15B-B", n=54), whereas the *C. capitata* signal strongly enriches *grk* RNA at both the DA and the anterior of the *D. melanogaster* oocyte (Figure 5.15C-C", n=36).

Replacing the central third and proximal third of the *C. capitata* signal with the same regions from the GLS to form *C. capitata-5* (Figure 5.15A), has no effect on the localisation of the *C. capitata* signal as it continues to direct *grk* RNA to both the DA and the anterior of the oocyte (Figure 5.15D-D", n=52), forming a strong anterior ring in 63% of oocytes injected (Figure 5.15F). However, replacing the distal third of the *C. capitata* GLS-like signal with the same region from the GLS to form *C. capitata-4*, results in *grk* enrichment at the DA of the oocyte in 60% of cases, with only a weak anterior ring in 93% of oocytes injected (Figure 5.15E-E", 5.15F, n=30). Taken together, these results indicate that the proximal and central thirds of the *C. capitata* GLS-like signal are not responsible for conferring aberrant *grk* localisation and anchoring. However, placing the distal third of the wild type GLS onto the *C. capitata* signal

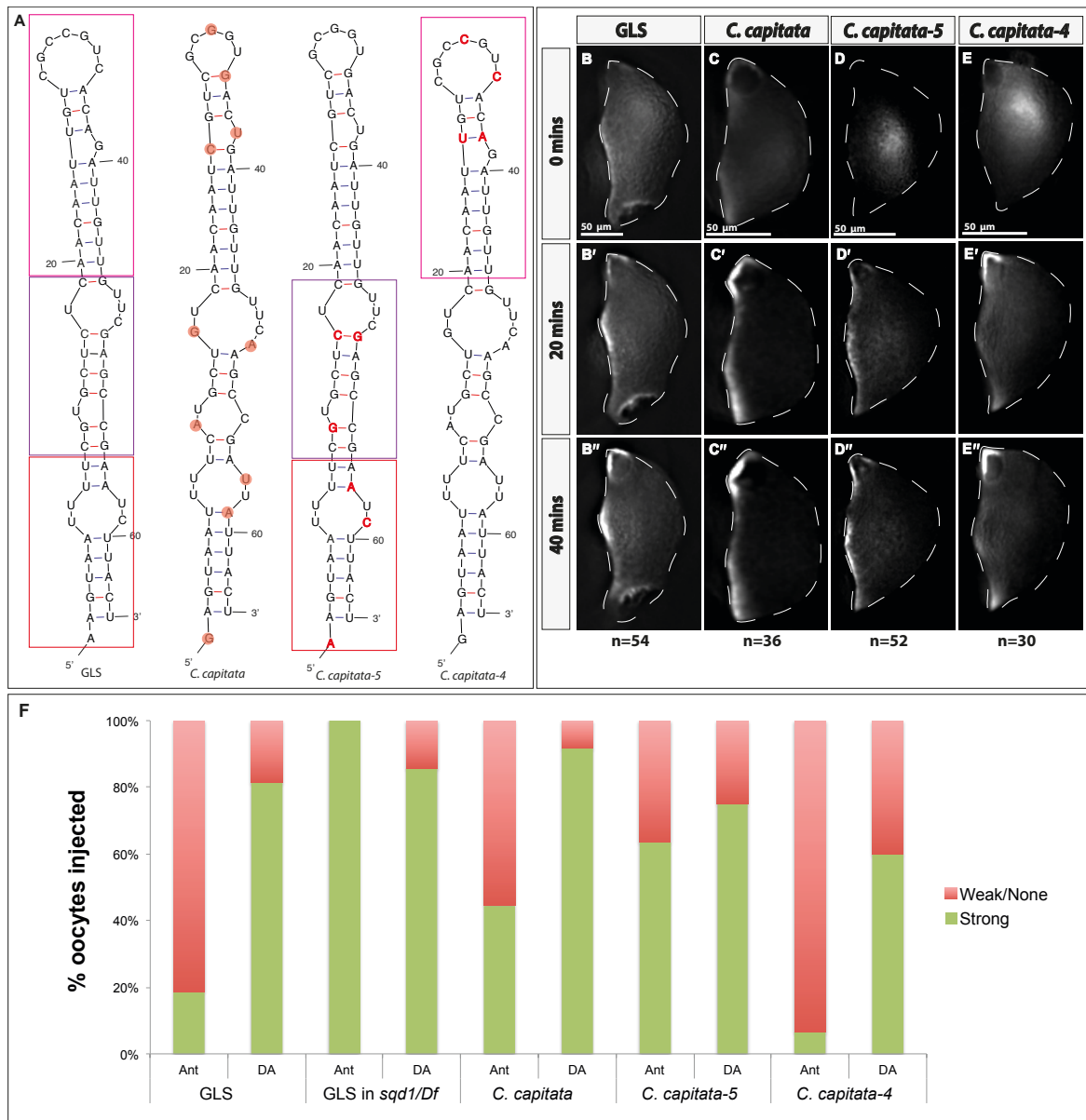


Figure 5.15. The distal third of the GLS-like signal from *C. capitata* is responsible for conferring aberrant *grk* anchoring at the anterior of the *D. melanogaster* oocyte. (A) UNAFold predicted secondary structures of the GLS, the GLS-like signal from *C. capitata*, and mutant *C. capitata* signals containing sections of the wild type GLS (pink box = distal third, purple box = central third, red box = proximal third of the GLS). (B-B'') Injected *grk* RNA containing the GLS localises to the DA of the oocyte within 40 minutes (n=54). (C-C'') The GLS-like signal from *C. capitata* drives injected *grk* RNA to the DA and the anterior of the *D. melanogaster* oocyte (n=36). (D-D'') *C. capitata-5*, containing the proximal two thirds of the GLS, localises to the DA of the oocyte and forms a strong anterior ring (n=52). (E-E'') A mutant form of the *C. capitata* GLS-like signal containing the distal third of the GLS (*C. capitata-4*) localises to the DA of the oocyte with only a weak anterior ring (n=30), similar to wild type *grk*. (F) Distribution of *grk* RNA at the anterior and DA of the oocyte at 40 minutes post-injection showing that replacing the distal third of the *C. capitata* signal with the distal third of the *D. melanogaster* signals rescues the *sqd1* anchoring phenotype. Oocytes outlined with white dashed line. (Ant = anterior, DA = dorsal-anterior).

to make *C. capitata*-4 (Figure 5.15A) rescues the *sqd1*-like phenotype of the *C. capitata* signal, and the distal third of the *C. capitata* signal can force the GLS to drive aberrant *grk* localisation (*C. capitata*-5), indicating that the two putative Sqd-binding sites in this region (PSB3 and PSB4, Figure 5.11) may indeed provide recognition motifs for Sqd binding.

Conclusions The distal third of *C. capitata* is responsible for conferring the ectopic localisation of *grk* at the anterior of the *D. melanogaster* oocyte.

The accessibility produced by bulge 2 is important for restricting *grk* to the DA of the oocyte

Based upon the SELEX data obtained for Sqd binding, two putative Sqd binding sites are predicted to form in the distal third of the GLS, PSB2 and PSB3 (Figure 5.11C; Figure 5.16A). Many RBPs have been shown to bind ssRNA in at least part of their binding site in order to provide accessibility for recognition of the bases.³⁶¹ As the putative binding sites PSB2 and PSB3 are found next to bulge 2, they are good candidates for Sqd binding. Based on the fact that the distal third of the GLS can restore the *sqd* phenotype of *C. capitata*, and that PSB3 is completely conserved in the *C. capitata* GLS-like signal, the role of PSB2 in *grk* anchoring was investigated.

To investigate whether bulge 2 of the GLS (Figure 5.16A) is required for Sqd binding, and therefore restricting *grk* to the DA of the oocyte, mutations were made in this region and their affect on *grk* localisation was assessed in the oocyte. The wild type GLS drives *grk* to the DA of the oocyte in 81% of cases, with a strong anterior ring in only 19% of injected oocytes (Figure 5.16B-B'', 5.16G). The mutant M38-5, in which bulge 2 is completely removed from the GLS (Figure 5.16A), drives *grk* to the DA of the oocyte in 88% of cases but also shows a strong anterior enrichment in 56% of oocytes injected (Figure 5.16C-C'', 5.16G), suggesting that bulge 2 may provide accessibility to Sqd binding sites within this region, or may form part of a sequence recognition motif.

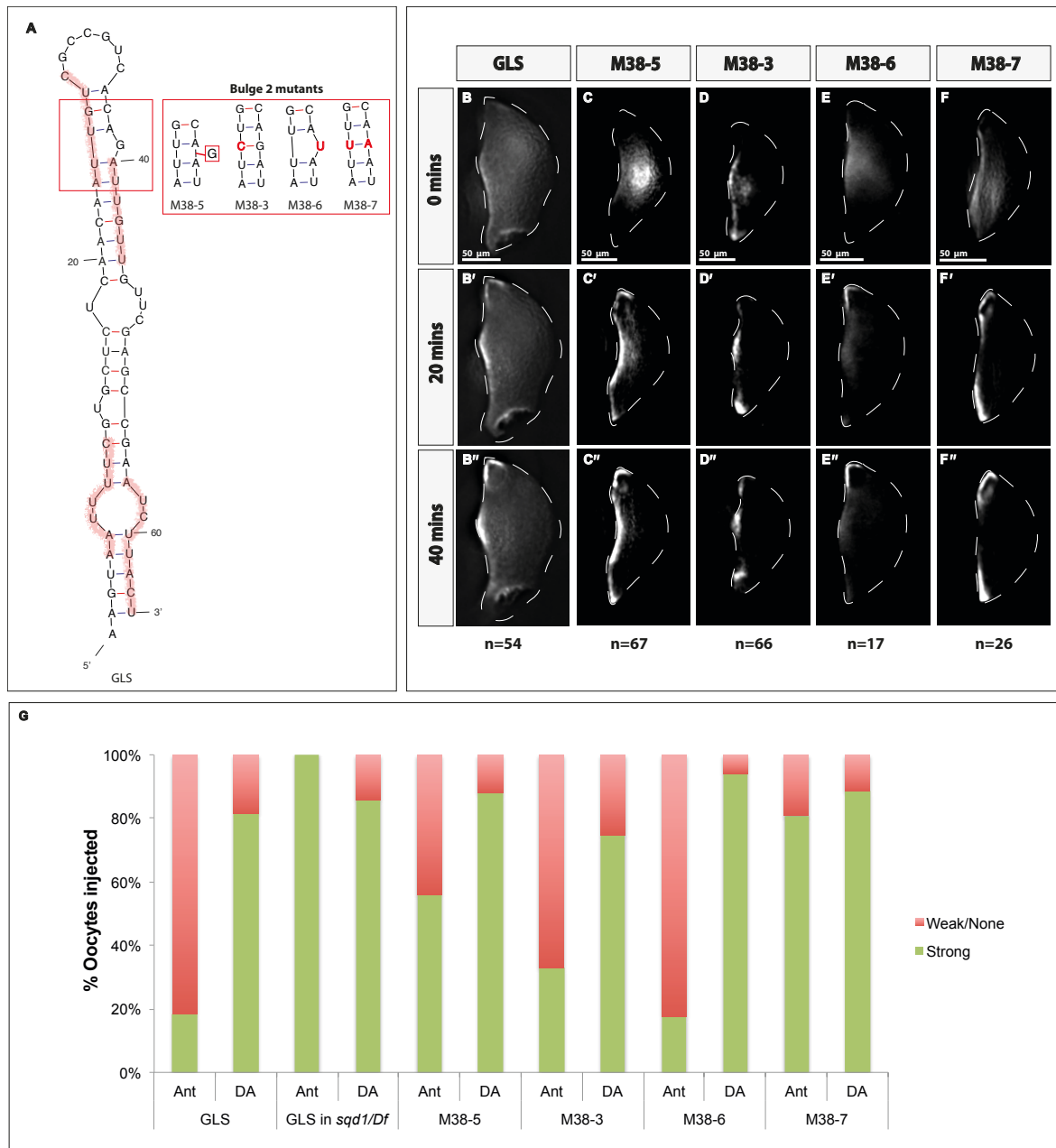


Figure 5.16. The accessibility provided by bulge 2 is important for efficiently restricting *grk* to the DA of the oocyte. (A) UNAFold secondary structures of the GLS and signals containing mutations in bulge 2. Putative Sqd binding sites are highlighted in pink. (B-B'') The GLS drives *grk* RNA to the DA of the oocyte within 40 minutes of injection (n=54). (C-C'') Removing bulge 2 in M38-5 increases the enrichment of *grk* RNA at the anterior of the oocyte. (D-D'') Replacing bulge 2 with a strong C-G base-pair increases the enrichment of *grk* RNA at the anterior of the oocyte (n=66). (E-E'') Replacing bulge 2 with a uracil in M38-6 has no effect on *grk* localisation to the DA of the oocyte (n=17). (F-F'') Replacing bulge 2 with a U-A base-pair in M38-7 greatly affects *grk* localisation, increasing its enrichment along the anterior of the oocyte (n=26). (G) Distribution of *grk* RNA in the oocyte at 40 minutes post-injection. *grk* RNA is driven to the DA of the oocyte in all of the mutants tested. Injecting *grk* into a *sqd1* mutant causes enrichment of the RNA at the anterior of the oocyte (data not shown). M38-5, M38-3 and M38-7 all cause an increase in *grk* enrichment at the anterior of the oocyte. Oocytes outlined with white dashed line. (Ant = anterior, DA = dorsal-anterior).

To investigate whether bulge 2 is required to provide accessibility for protein binding or to form part of a recognition sequence, the bulge was forced to form a strong base-pair by inserting a cytidine opposite the bulge in mutant M38-3 (Figure 5.16A). Injection of M38-3 into the oocyte shows that it is capable of driving *grk* RNA to the DA of the oocyte in 75% of cases, but forms a strong anterior ring in 33% of oocytes injected (Figure 5.16D-D'', 5.16G). Maintaining bulge 2 but replacing the bulged guanine with a uracil in M38-6 (Figure 5.16A), on the other hand, has little effect on *grk* localisation, with only 18% of oocytes injected showing a strong anterior ring, and 94% showing enrichment at the DA (Figure 5.16E-E'', 5.16F). Finally, replacing the bulge with a much weaker U-A base-pair in M38-7 (Figure 5.16A) greatly affects *grk* localisation with 81% showing a strong anterior enrichment and 88% showing DA enrichment (Figure 5.16F-F'', 5.16G).

To further investigate whether the bulge 2 mutations tested under low magnification in the oocyte show characteristics of *sqd* mutants, they were imaged at higher resolution and assessed for clustering along the anterior of the oocyte. In addition, interdigitation of the RNA with Me31B along the anterior of the oocyte was tested as a proxy for inappropriate anchoring of the RNA. The wild type GLS causes *grk* to form large clusters specifically at the DA of the oocyte which interdigitate with Me31B within 40 minutes of injection (Figure 5.17A-A''). In a *sqd1* mutant, on the other hand, these large *grk* clusters are also found at the anterior of the oocyte where they interdigitate with Me31B (Figure 5.17B-B''). The mutants M38-5, M38-3 and M38-7, which all showed increased enrichment at the anterior of the oocyte under low magnification, also produce *grk* clusters at both the DA and the anterior of the oocyte that interdigitate with Me31B (Figure 5.17C-C'', 5.17D-D'', 5.17F-F''). However, the mutant M38-6 that directs normal *grk* localisation under low magnification also shows wild type-like *grk* clustering that is restricted to the DA of the oocyte, and is not interdigitated with Me31B at the anterior (Figure 5.17E-E''). Taken together, these results suggest that the accessibility created by bulge 2 is important for efficiently restricting *grk* to the DA of the oocyte, rather than the sequence of the bulge, and therefore may provide a structural recognition motif for the association of Sqd with the putative binding site PSB2. In addition,

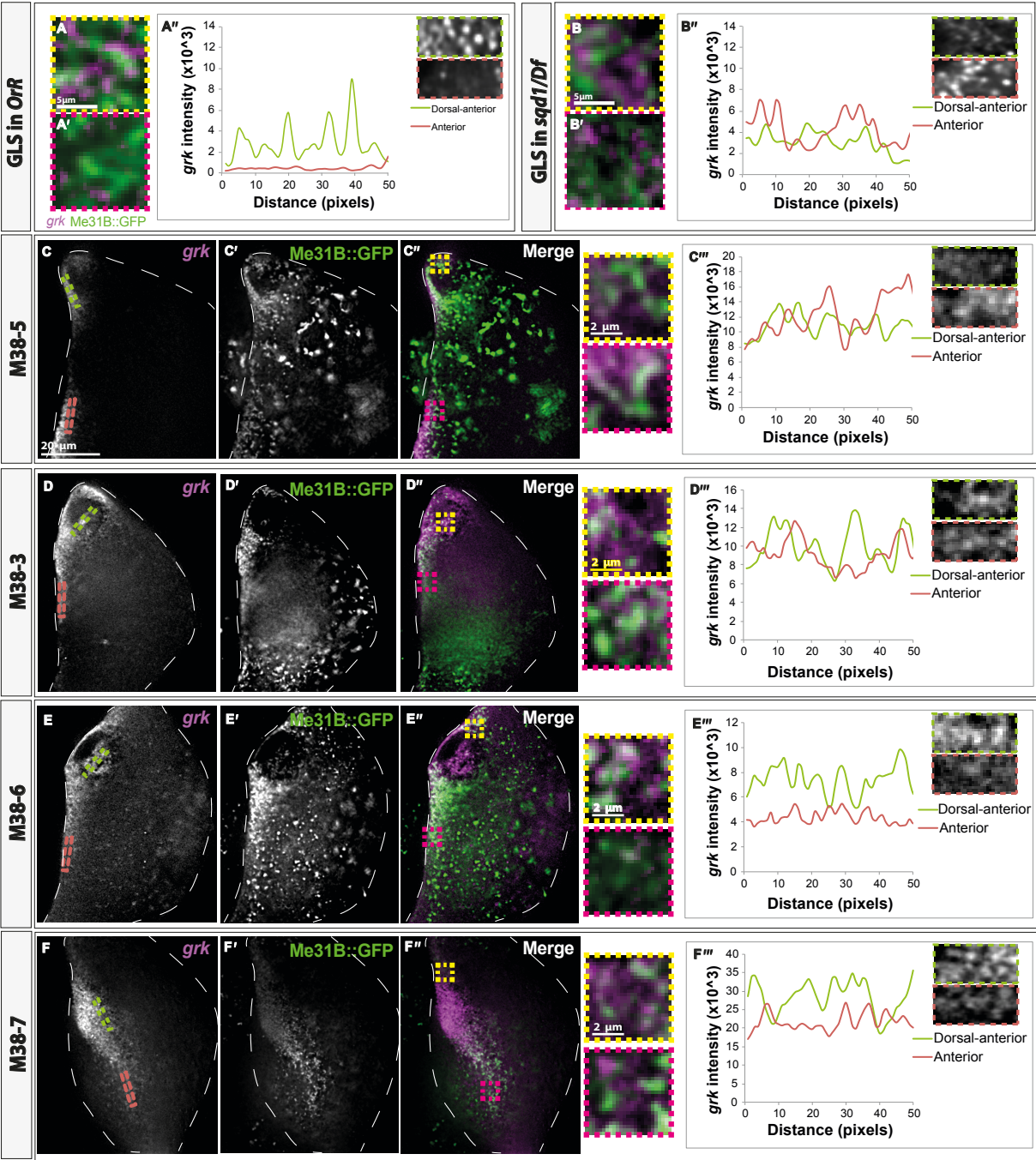


Figure 5.17. Bulge 2 is required to restrict *grk* anchoring to the DA of the oocyte. (A-A'') The GLS drives *grk* RNA (magenta) to the DA of the wild type oocyte where it interdigitates with Me31B (green) and forms large clusters indicative of anchoring at the edge of P bodies. At the anterior of the wild type oocyte only low levels of *grk* are seen that do not interdigitate with Me31B::GFP and do not form large clusters (n=3). (B-B'') In a *sqd1* mutant, *grk* RNA localises to both the DA and the anterior where it interdigitates with Me31B and forms large clusters (n=3). (C-C'') Removing bulge 2 in M38-5 causes inappropriate *grk* localisation and clustering at the anterior of the oocyte which interdigitates with Me31B::GFP, and normal clustering and anchoring at the DA. Replacing bulge 2 with a strong C-G base-pair in M38-3 (D-D'', n=3), or a weak U-A base-pair in M38-7 (F-F'', n=3) also causes *grk* enrichment in large clusters at the edge of P bodies at the anterior of the oocyte. (E-E'') Maintaining bulge 2 but changing the identity of the bulged nucleotide in M38-6 has no effect on *grk* anchoring with *grk* interdigitation with Me31B occurring only at the DA of the oocyte, and only a weak *grk* signal at the anterior of the oocyte (n=3). Fluorescent intensities were taken from a section through the DA (green box) or anterior (red box) of the oocyte and clustering was assessed. Interdigitation with Me31B was assessed at the DA (yellow box) and the anterior (pink box) of the oocyte. All Images taken at 100x magnification. Oocytes outlined with white dashed line.

the sequence of PSB2 may be essential for restricting *grk* anchoring but further studies are required to confirm this.

Conclusions Removing bulge 2 from the GLS disrupts *grk* localisation resulting in inappropriate anchoring along the anterior of the oocyte.

Discussion

***grk* RNA is localised to both the DA and the anterior of *sqd1* oocytes**

grk RNA is localised to the DA of the oocyte during mid-oogenesis in a two-step process that requires Dynein and the microtubule network. The first step of *grk* transport to the anterior of the oocyte is not thought to require *Sqd*, as the integrity of *grk* transport particles is maintained in a *sqd* mutant,¹³³ and I show that *grk* transport to the anterior of the *sqd1* oocyte following injection is comparable to wild type. However, the second step of *grk* localisation, restriction to the DA corner, is disrupted in *sqd1* mutants, resulting in *grk* being

mislocalised around the entire circumference of the oocyte.^{15, 133, 193, 283, 317} In addition to the strong anterior ring, however, I report an accumulation of both injected and MS2-labelled *grk* at the DA corner of *sqd1* oocytes, forming a similar distribution to *grk* in some *cup* mutants.³¹⁴ Similar DA corner *grk* localised has previously been reported in a *sqd1* mutant,^{15, 133} but the reasons behind why it forms in a mutant thought to show no *grk* anchoring has not been explored.

By comparing *grk* RNA at the DA and anterior of wild type and *sqd1* mutant oocytes, I showed that *grk* forms large, bright clusters at the DA in both cases, and at the anterior of the mutant. Measuring the size of clusters for MS2-tagged *grk* showed that they are significantly larger than the small population of *grk* particles found at the anterior of the wild type oocyte, suggesting that they may have further distinct characteristics.

The dynamics of *grk* RNA in both wild type and *sqd1* mutant egg chambers have been previously characterised.^{133, 283} In wild type oocytes, *grk* RNA forms highly dynamic transport particles that become less dynamic throughout oogenesis.²⁸³ As *grk* becomes restricted to the DA of the oocyte it is thought to become statically anchored at the edge of P bodies in a process that requires Dynein and Sqd.^{133, 400} In a *sqd1* mutant, *grk* has been shown to remain in transport particles, rather than forming P bodies, and is highly dynamic along the entire anterior of the oocyte. FRAP experiments showed that the anterior particles in wild type oocytes share similar dynamics to *grk* at both the anterior and DA of the *sqd1* mutant.^{133, 283} As a result, Sqd has been implicated in switching Dynein from a processive motor to a static anchor, restricting *grk* to the DA of the oocyte.¹³³

However, using fast time-lapse imaging, I observe low levels of dynamic behaviour for *grk* particles at the DA and the anterior of *sqd1* mutants that is comparable to that of *grk* at the DA of wild type oocytes. It is possible that the differences seen between the work presented here and in the literature are due to differences in the experimental set-up, or variations within the mutant fly lines as a result of picking up modifiers over time.^{133, 283} For example, in

these studies I have expressed *sqd1* with a deficiency line covering the region where *sqd* is encoded, with the aim of removing any effects of possible off-target mutations that could be present in a *sqd1/sqd1* mutant line. To test the possibility of differences in the fly line affecting the distribution and dynamics of *grk*, the experiments presented here should be repeated in a *sqd1/sqd1* mutant background, as used in previous publications.^{133, 283}

FRAP measures the average mobility of many molecules in a small area and is often carried out using confocal microscopy which is limited to a single z-plane and slow speeds. As a result of advances in imaging, both the resolution and acquisition speeds have increased, allowing us to resolve and track *grk* particles in multiple z-planes and at high speeds. FRAP experiments of MS2-tagged *grk* in a wild type background carried out on the OMX-V2 have shown that RNA at the DA is rapidly turned over but the overall distribution of the RNA is retained (Richard Parton, unpublished). This suggests that even the anchored RNA is highly dynamic, and could be explained by movements within the anchored assembly, or continued incorporation of new RNA particles. Nevertheless, additional studies will be required to further characterise the dynamics of *grk* in a wild type and *sqd1* mutant background.

***grk* RNA is associated with P bodies at the DA and anterior of *sqd1* oocytes**

Specifically at the DA corner of the oocyte, *grk* RNA docks and is statically anchored at the edge of P bodies where translational activators, such as Orb, are highly enriched, enabling translation of the RNA.^{133, 193} In a *sqd1* mutant, *grk* is translated along the anterior margin,^{15, 299, 300, 317} but how it is translationally activated there is currently unknown. Following injection of *grk* RNA into a *sqd1* mutant, most of the RNA was found in transport particles at the anterior. However, a small proportion was found in P bodies at the DA using electron microscopy (EM),¹³³ suggesting that *grk* may still associate to some extent with P bodies in the absence of Sqd. In support of this, I show that injected *grk* RNA interdigitates with Me31B::GFP at both the anterior and the DA of the *sqd1* oocyte after 40 minutes in a similar manner to at the DA of wild type oocytes. This suggests a possible mechanism for *grk*

translation at the anterior of a *sqd1* mutant. However, it is difficult to distinguish between differences in interdigitation with Me31B in GLS or *sqd1* mutants due to anchoring of the RNA, or simply due to a lower level of localised RNA at the anterior of wild type oocytes. As a result, levels of interdigitation should be quantified by normalising to the amount of RNA in each region, and statistical analysis should be carried out.

If Sqd is involved in the anchoring of *grk* specifically at the DA of the oocyte, one would expect Sqd to be enriched here with the RNA. Previous studies have shown that Sqd is recruited to the DA of the oocyte following the injection of *grk* RNA.¹³³ However, I was only able to show recruitment of Sqd to the anterior of the oocyte following *grk* injection. As Sqd is highly expressed in the oocyte nucleus, it is possible that the strong nuclear signal masks any weak enrichment at the DA corner. The Sqd::GFP signal at the anterior and DA of the oocyte could be enhanced by confocal microscopy which improves the imaging of weak, diffuse signal.

Taken together, these results have led me to re-interpret the function of Sqd in *grk* anchoring, where Sqd prevents inappropriate *grk* association and anchoring at P bodies at the anterior of the oocyte, potentially preventing ectopic translation of Grk. If this were the case, in the absence of Sqd, *grk* would be anchored at the edge of P bodies at both the anterior and the DA of the oocyte where it is translationally activated. In support of this, I have shown that the *sqd1 grk* mislocalisation phenotype can be rescued by injecting *grk* RNA in close proximity to the nucleus where it becomes highly enriched, and does not move towards the anterior of the oocyte. Furthermore, studies have shown that overexpressing *grk* RNA results in its localisation along the anterior of the oocyte and incorporation into the core of P bodies where it is presumably translationally repressed.¹⁹³ It is possible that overexpression of *grk* saturates the ability of Sqd to inhibit anchoring at the anterior of the oocyte or, using the current interpretation, Sqd is unable to anchor the excess *grk* RNA at the DA corner. Both of these interpretations result in the same localisation pattern of *grk* in the wild type oocyte, but through subtly distinct mechanisms.

Bulge 2 is required to efficiently restrict *grk* anchoring to the DA of the oocyte

Four putative Sqd binding sites are predicted within the GLS from SELEX binding studies.³⁹⁸ Truncations of the GLS show that the proximal two sites, PSB1 and PSB4, are not required to restrict *grk* to the DA of the oocyte. The putative site PSB2 is predicted to have diverged in sequence or structure in some of the GLS-like signals from other species, and injection of these signals into the *D. melanogaster* oocyte showed varying degrees of *grk* mislocalisation. The GLS-like signal from the most divergent species *C. capitata* was unable to efficiently restrict *grk* to the DA of the oocyte and caused the RNA to form large clusters that interdigitated with Me31B::GFP at both the anterior and DA of the oocyte. Mutation of the GLS-like signal from *C. capitata* revealed that the distal third of the signal is responsible for conferring *grk* mislocalisation, suggesting that divergences within this region result in inefficient recruitment of Sqd.

RBP's often require at least some degree of ssRNA in their binding sites to allow accessibility to the bases within the RNA.³⁶¹ As the putative Sqd binding sites PSB2 and PSB3 are adjacent to bulge 2, they are likely candidates for Sqd binding. Mutations to bulge 2 show that the accessibility of this region is indeed necessary for the efficient restriction of *grk* to the DA of the oocyte, suggesting that it may be required for Sqd binding. Removal of the bulge, or replacing it with a base-pair, greatly increases the proportion of oocytes showing a strong anterior ring of *grk* that interdigitates with Me31B, suggesting impaired Sqd binding. However, changing the identity of the bulge does not reduce *grk* localisation, and may actually improve restriction of injected *grk* to the DA of the oocyte. As a result, it is likely that bulge 2 provides the accessibility required for binding to a sequence within this region, such as PSB2 or PSB3. Due to the fact that PSB3 is completely conserved in the GLS-like signal from *C. capitata*, whereas PSB2 contains a divergent nucleotide, I favour the hypothesis that PSB2 provides a sequence recognition motif for Sqd binding. However, SELEX binding data represents an average of the binding preferences for all of the RBDs within a protein of

interest, making it difficult to predict sequence motifs for a protein containing more than one RBD.

Future work

To further investigate the role of Sqd in *grk* anchoring, particle dynamics are required for *grk* RNA at the anterior and the DA of the wild type and *sqd1* mutants, along with single molecule FISH to assess the endogenous pattern of *grk* localisation. First, FRAP experiments could be carried out with the better imaging conditions now available through use of the spinning disk. Second, the photoswitchable *grk*dendra* fly line used in this chapter could act as a tool for the development of an anchoring assay. Alternatively, fluorescence correlation spectroscopy (FCS) could be used to assess *grk* dynamics by measuring fluctuations in fluorescence as a result of molecules crossing an illuminated region. To characterise the regions of the GLS required for Sqd binding, the putative Sqd mutants could be tested *in vivo* as MS2-tagged RNAs to assess for aberrant *grk* localisation, translation along the anterior of the oocyte, or dorsalised egg defects. In addition, as many of the GLS-like signals identified in other species showed poor *grk* localisation to the DA of the *D. melanogaster* oocyte, it would be interesting to assess the egg shell appendages of these species to determine whether co-evolution of the anchoring machinery and localisation signal results in the GLS-like signals being capable of driving *grk* anchoring. The ability of putative Sqd mutants to interact with Sqd could be assessed using biochemical techniques. Finally, investigating the interdependence of *grk* translation and localisation by inhibiting translation would determine whether the two processes are separable.

Final conclusions

In this chapter, I have shown that the putative Sqd binding sites in the proximal third of the GLS are not important for correct *grk* anchoring in the oocyte. Instead, bulge 2 in the distal third of the GLS is important for providing accessibility to the anchoring machinery, and

therefore for ensuring that *grk* anchoring is restricted to the DA of the oocyte. In addition, the data presented in this chapter have led me to a revised interpretation of the role of Sqd in mediating *grk* anchoring and translation, preventing inappropriate association and anchoring with P bodies at the anterior of the oocyte. This is based upon my findings that:

1. *grk* forms a strong anterior ring and associates with the DA corner in *sqd1* mutant oocytes
2. Ectopic *grk* at the anterior of the *sqd1* mutant oocyte forms large clusters and interdigitates with Me31B in a similar manner to anchored *grk* at the DA of wild type oocytes
3. The *sqd1* mutant phenotype can be rescued by injecting *grk* RNA in close proximity to the nucleus, causing it to enrich specifically at the DA of the oocyte
4. Ectopic *grk* at the anterior of the *sqd1* mutant oocyte is highly static, suggesting that the RNA is anchored

Chapter 6: Regions of the GLS required for *grk* RNA association with P bodies

Introduction

***grk* mRNA is translationally silenced during its transport to the DA of the oocyte**

During its localisation to the DA of the oocyte, *grk* mRNA must be translationally silenced to prevent ectopic expression of Grk, which could be detrimental to the cell. The translational state of *grk* mRNA is thought to be controlled, in part, by the *trans*-acting factors Encore, Vasa, PABP, Me31B, Cup, Sqd and Bruno.^{162, 193, 314, 317, 322} The favoured model for translational regulation of *grk* was that Bruno binds to BRE-like sequences in the 3'UTR of *grk* to inhibit its translation,^{161, 317} as is the case for *osk* mRNA.¹⁶⁰ However, this model has not been conclusively tested and the Bruno-mediated silencing of Grk translation has been shown to be much weaker than for *osk*.⁴⁰¹ In addition, it has been suggested that translational silencing of *grk* mRNA is established by spatially restricting access to the translational activator Orb in the NCs.³²⁹ As a result of insufficient levels of Orb, *grk* is unable to associate with P bodies (Me31B-rich electron-dense regions in the cytoplasm involved in the translation and degradation of mRNAs) in the NCs.^{193, 329} Me31B has been implicated in the translational repression of several maternal mRNAs during their transport from the NCs into the oocyte; including *BicD*, *osk*, and *grk*, through packaging the mRNA into dense cytoplasmic RNP complexes and denying them access to ribosomes.^{193, 326} However, fast-live imaging suggests that Me31B moves independently from *grk* and *bcd* transport particles during active transport in the oocyte.¹⁹³

Translation of *grk* mRNA occurs at the edge of P bodies

Translational activation of *grk* mRNA only occurs upon its arrival at the DA of the oocyte. Here, *grk* is anchored in a Sqd-dependent manner at the edge of P bodies where ribosomes are enriched, along with the translational activator Orb which is required for both *grk* polyadenylation and Grk protein expression.^{133, 193, 314, 317, 402-404} Me31B is highly enriched at the core of P bodies, and is therefore a well-established marker for P bodies *in vivo*.³²⁶ At the DA of the oocyte, *grk* mRNA has been shown to interdigitate with Me31B at the edge of P bodies, whereas *bcd* mRNA colocalises with Me31B in the core of the body. As a result, differential association of mRNAs with the translationally active periphery of P bodies and the silent core is thought to control the activation and silencing of *grk* and *bcd* mRNA, respectively.¹⁹³

The role of P bodies in *grk* translation is further exemplified when *grk* RNA is injected into the oocyte: at moderate concentrations of RNA, *grk* localises to the edge of P bodies and interdigitates with Me31B.¹⁹³ However, at high concentrations injected *grk* RNA localises to the core of P bodies. In addition, overexpression of endogenous *grk* results in a similar distribution of *grk* at the P body core. *grk* has also been shown to accumulate in P bodies in the egg chamber in response to starvation.³³⁰ These results suggest that P bodies act as translational regulators by; repressing translation at the core which is devoid of ribosomes, sequestering excess *grk* mRNA, or silencing *grk* translation during stress conditions.

P bodies disassemble following egg activation

The formation and role of P bodies has long been debated. While it is still not proven whether P bodies are the cause or consequence of mRNA silencing,⁴⁰⁵ it is most widely accepted that P body formation is necessary for mRNA decay and translational repression.²⁰⁹ *bcd* mRNA colocalises with Me31B at the core of P bodies in the oocyte, rendering it translationally silent from stages 9-14 of oogenesis. However, translation of *bcd* mRNA must be activated in early

embryos to produce the Bcd gradient required for organising the anterior-posterior axis of the embryo.^{19, 280} During egg activation, *bcd* mRNA is released from its anchor at the anterior of the oocyte and loses its colocalisation with Me31B at the core of P bodies.^{193, 400} The dissociation of *bcd* mRNA from P bodies was shown to be a result of the dispersion of components of the P body, including Me31B and Trailerhitch (Tral).¹⁹³ Therefore, translational activation of *bcd* mRNA is likely to be a result of P body disassembly following egg activation. Despite these recent studies, little is known about how P bodies form and disassemble in the *Drosophila* oocyte and embryo to regulate translation of mRNAs, and how Me31B and Sqd act together in the regulation of *grk* translation.

Aims of this chapter and experimental rationale

In this chapter I will investigate the association between *grk* RNA and P bodies in the embryo, as Me31B is not seen to be recruited to injected *grk* RNA in the oocyte or the NCs (data not shown). I will address the following specific questions:

1. Can live embryo injection assays be used to study the association of *grk* with the P body marker Me31B?
2. Which regions of the GLS are required for *grk* association with the P body marker Me31B and therefore translational regulation of the transcript?

In order to address these questions I have used a well-established recruitment assay in the syncytial blastoderm embryo to test, for the first time, whether Me31B is recruited to ectopic *grk* RNA in the embryo. If *grk* and Me31B are found to associate in the embryo, I will compare their relationship to that found in the P bodies shown to regulate *grk* translation in the oocyte, and investigate whether mutations within the sequence or structure of the GLS can disrupt their interaction.

Results

Injected *grk* RNA recruits Me31B to the apical cortex in the embryo

Biochemical studies have shown that Me31B associates with *grk* mRNA in a GLS-specific manner.⁵⁴ In order to dissect the regions of the GLS required for this association, and thus translational regulation of *grk* in P bodies, an assay was required to study the recruitment of Me31B to mutant GLS signals. As Me31B is already enriched at the DA of uninjected oocytes, it is difficult to measure recruitment to *grk* RNA following injection into the oocyte. In addition, Me31B accumulates in the oocyte in response to stressors, such as long culturing times (Figure 3.7A-C'), which could affect the results of an Me31B recruitment assay in the oocyte. Therefore, an assay with short culturing times was required to keep the effects that stress may have on the results to a minimum.

In syncytial blastoderm embryos, ectopically injected *grk* RNA localises to the apical cortex and recruits components of the localisation machinery within minutes.¹¹⁶ To investigate whether this embryo recruitment assay can be used to study the recruitment of Me31B, *grk* was injected into embryos expressing Me31B::GFP, and localisation of the RNA and the protein was assessed after 20 minutes (Figure 6.1A). *grk* RNA localises to the apical cortex of 100% of embryos (Figure 6.1B-B'', 6.1D-D''; Figure 6.11B). When injected into wild type *OrR* embryos, no signal is seen at the apical cortex in the GFP channel (Figure 6.1C-C''), indicating there is no bleed through between the channels. Localised *grk* RNA at the apical cortex of embryos expressing Me31B::GFP, on the other hand, shows strong recruitment of Me31B in 83% of embryos within 20 minutes of injection (Figure 6.1E-E''; Figure 6.11B). The non-localising control *hb* is unable to localise in 100% of embryos injected (Figure 6.1F-F''), and Me31B::GFP is not enriched at the apical cortex (Figure 6.1G-G'). These patterns of localisation are further demonstrated by quantifying the fluorescent intensities of the RNA and GFP (Figure 6.1H-I').

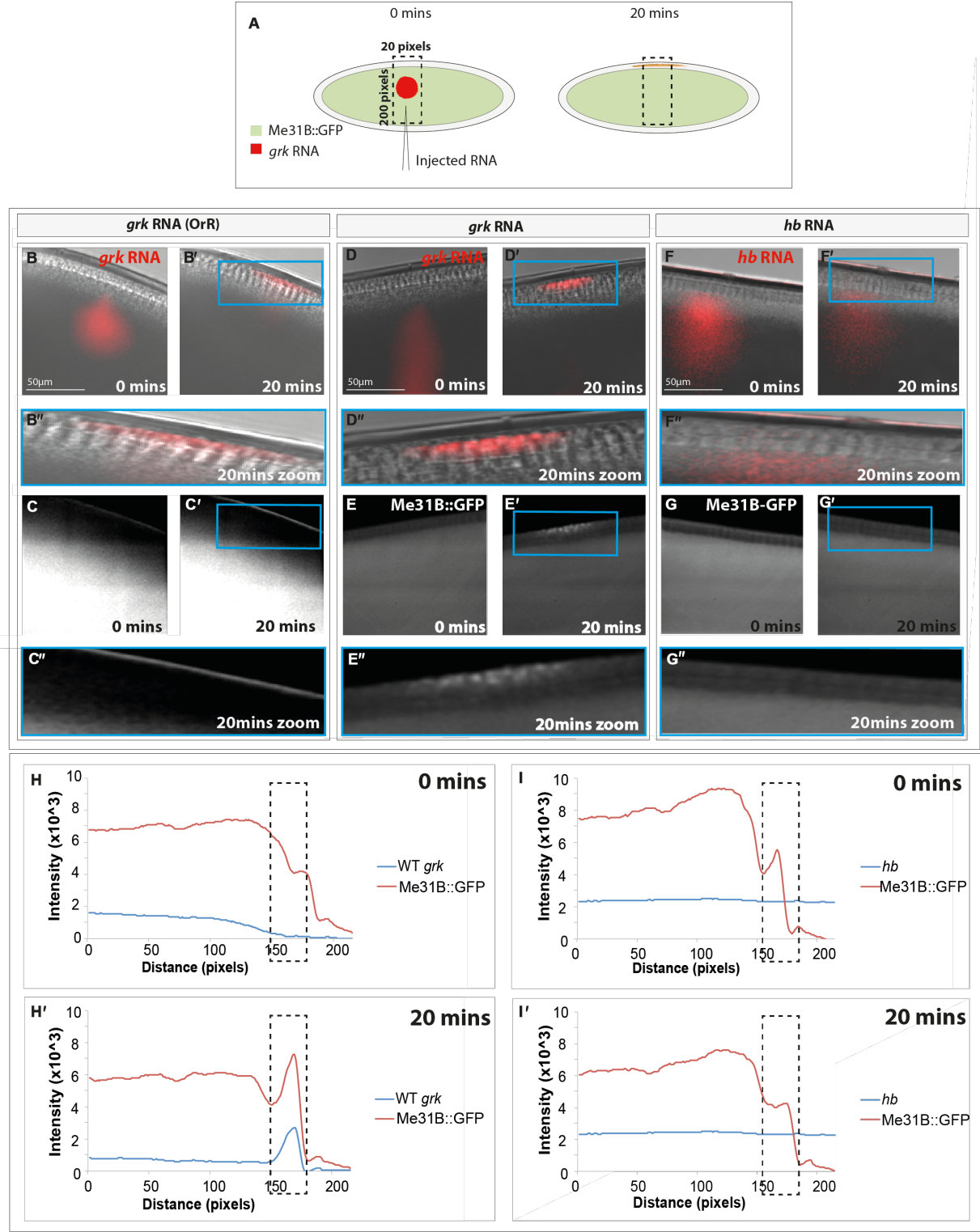


Figure 6.1. Me31B is recruited to apically localised *grk* RNA in the embryo. (A) Schematic representation of the recruitment injection assay in the syncytial blastoderm embryo. (B-B'') *grk* RNA localises to the apical cortex of the embryo within 20 minutes of injection into wild type *OrR* embryos, and the fluorescent signal does not bleed through in to the green channel (C-C'', n=3). (D-D'') *grk* injected into Me31B::GFP embryos localises apically and recruits Me31B within 20 minutes (E-E'', n=10). (F-F'') The non-localising control *hb* does not enrich apically and does not recruit Me31B (G-G'', n=3). Fluorescent intensity profiles taken from a cross-section through the centre of the embryo over time shows both *grk* and Me31B enrich at the apical cortex (H-H') whereas *hb* does not (I-I'). (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (H-I') refers to the position through the cross-section of the embryo from the site of injection to the apical cortex (black box)).

Although little is currently known about the role of Me31B in the translational regulation of mRNAs in the embryo, these results indicate that Me31B can accumulate in the embryo in response to injected *grk* RNA, and is recruited to the apical cortex following *grk* localisation. Therefore, it is possible that Me31B is involved in RNA localisation and/or translation in the embryo.

Conclusions *grk* RNA is able to recruit Me31B following its localisation to the apical cortex of the embryo.

Me31B is associated with *grk* RNA at the apical cortex but not in transport particles

In the oocyte, Me31B has been shown to associate with *grk* RNA specifically at the DA, but not in other regions of the oocyte, and is not found in *grk* transport particles.¹⁹³ To determine whether Me31B shows similar characteristics in the embryo, high resolution imaging was used to investigate the relationship between *grk* RNA and Me31B, both in transport particles and localised at the apical cortex. At 15 minutes post-injection *grk* RNA transport particles move through the cytoplasm towards the apical cortex of the embryo (Figure 6.2A-A'''). Me31B::GFP does not associate with 75% of *grk* transport particles in the cytoplasm (Figure 6.2B-C'''). Me31B also moves towards the apical cortex of the embryo following *grk* injection

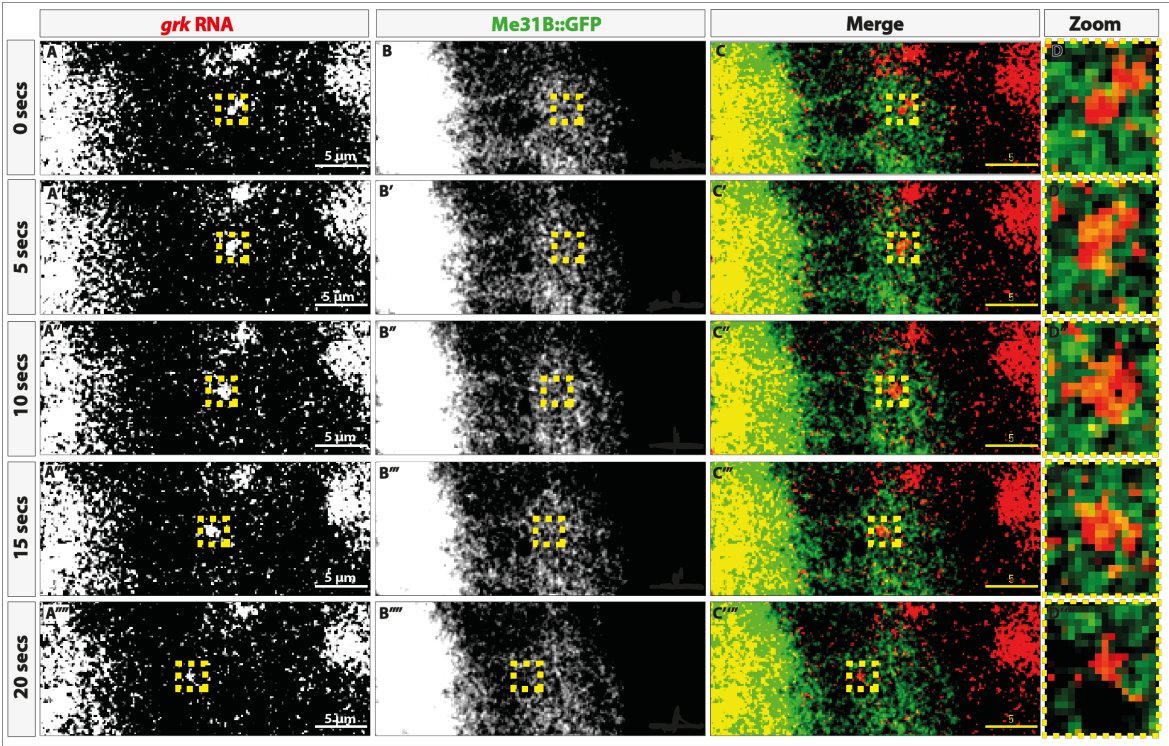


Figure 6.2. *grk* RNA moves independently of Me31B in transport particles. (A-A''') *grk* RNA transport particles move towards the apical cortex of the embryo following injection (yellow box, n=16). (B-B''') Me31B::GFP is not found associated with *grk* transport particles (yellow box, n=12). (C-C''') Merge of *grk* RNA and Me31B::GFP shows that *grk* and Me31B move independent during transport to the apical cortex of the embryo and do not colocalise (yellow box, zoom). Images were taken at 15 minutes post-injection and are equally contrasted for each channel over the time course. The apical cortex is positioned to the left of the images. Images taken at 100x with auxiliary magnification.

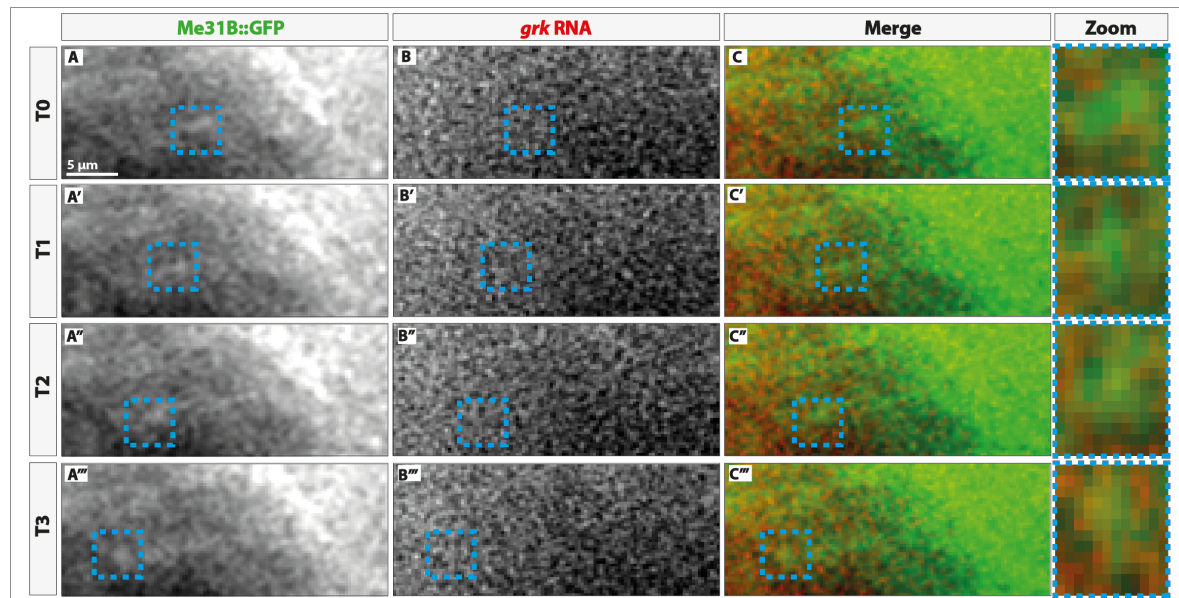


Figure 6.3. Me31B moves independently of *grk* RNA in transport particles. (A-A''') Me31B::GFP moves towards the apical cortex of the embryo following injection of *grk* RNA (blue box, n=30). (B-B''') *grk* RNA is not found associated with Me31B during transport (blue box, n=26). (C-C''') Merge of *grk* RNA and Me31B::GFP shows that Me31B moves independently of *grk* during transport to the apical cortex of the embryo, and they do not colocalise (blue box, zoom). Images were taken at 15 minutes post-injection and are equally contrasted for each channel over the time course. The apical cortex is positioned to the left of the images. Images taken at 100x with auxiliary magnification.

(Figure 6.3A-A'') and is not associated with *grk* RNA in 87% of cases (Figure 6.3B-C''), moving independently through the cytoplasm and between the apical nuclei. These results indicate that *grk* RNA and Me31B in the embryo behave in a similar manner to in the oocyte, moving independently of one another during transport.

Conclusions *grk* and Me31B travel independently to the apical cortex of the embryo.

***grk* RNA co-localises with Me31B at the apical cortex of the embryo**

In the oocyte, *grk* RNA is anchored and translated at the edge of P bodies at the DA corner, interdigitating with Me31B.^{133, 193} However, overexpression of *grk* RNA results in the colocalisation of *grk* and Me31B as *grk* RNA is targeted to the core of P bodies where it is translationally silenced. To investigate whether *grk* RNA is anchored at the edge of P bodies following injection into the embryo, Me31B and *grk* RNA were imaged at the apical cortex. Injection of *grk* RNA into the embryo results in its localisation to the apical cortex within 10 minutes in 100% of embryos (Figure 6.4A-A''; Figure 6.11B), and the formation of Me31B-rich P body-like structures in 92% of embryos (Figure 6.4B-B''; Figure 6.11B). Merged images of *grk* RNA and Me31B::GFP show that they colocalise as large clusters at the apical cortex in 100% of embryos injected (Figure 6.4C-C'', blue box), but not at the site of injection (Figure 6.4C, white asterisk). Higher magnification confirms that *grk* RNA (Figure 6.4A''') and Me31B (Figure 6.4B''') are colocalised (Figure 6.4C''') at the apical cortex. Injection of tetraspeck beads (beads stained with four different fluorescent dyes to assess co-registration) into the embryo confirm colocalisation with these imaging conditions (Figure 6.4D). These results show that, in contrast to at the DA of the oocyte where *grk* and Me31B interdigitate, *grk* RNA colocalises with Me31B at the apical cortex of the embryo, suggesting that *grk* may be targeted to the core of P bodies here, which is likely to result in translational repression of ectopic *grk* RNA.

Conclusions *grk* and Me31B colocalise at the apical cortex of the embryo

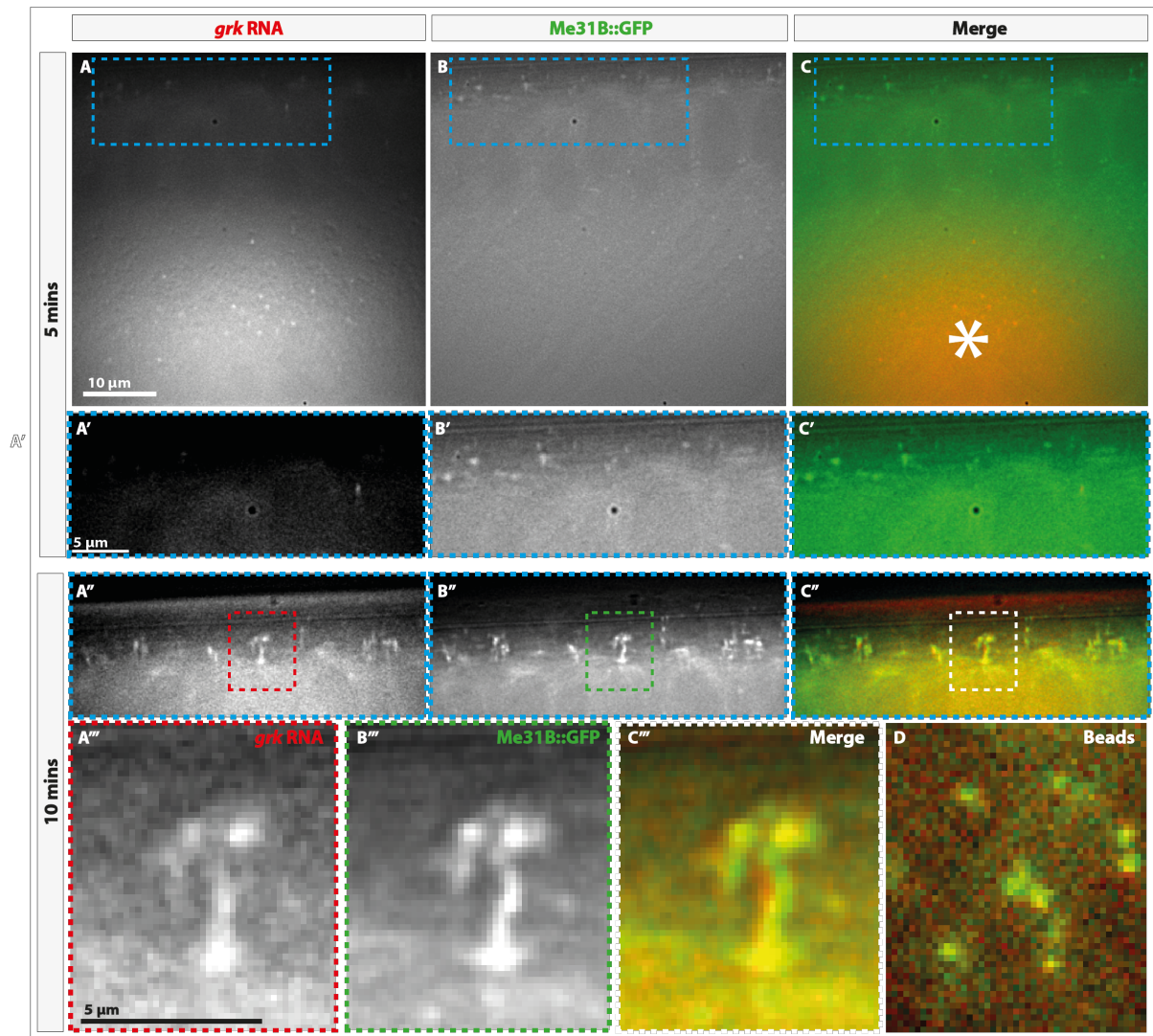


Figure 6.4. *grk* RNA and Me31B colocalise at the apical cortex of the embryo. (A-A'') Injected *grk* RNA localises to the apical cortex of the embryo within 10 minutes (n=12, blue box). (B-B'') Me31B::GFP accumulates at the apical cortex over time and is highly enriched by 10 minutes (n=12, blue box). (C-C'') *grk* RNA and Me31B::GFP colocalise at the apical cortex after 10 minutes (blue box, n=7), but not at the site of injection (white asterisk). Higher magnification at 10 minutes shows that *grk* RNA (red box, A'') and Me31B::GFP (green box, B'') colocalise at the apical cortex (white box, C'', n=7). (D) Injection of tetraspeck beads into the embryo show that the red and the green channel co-register (n=3), suggesting that *grk* RNA and Me31B::GFP are colocalised. Apical cortex positioned at the top of each image. A-C equally contrasted for both channels, A'-C'' equally contrasted within each channel. All images except A, B and C are taken at 100x with auxiliary magnification.

C. capitata is able to recruit Me31B in the embryo

In Chapter 3 I identified several GLS-like signals from species outside of the *Drosophilidae* family. To investigate which regions of the GLS are responsible for the association between *grk* and Me31B, and therefore for the translational regulation of *grk* RNA, I tested the ability of the GLS-like signal from *C. capitata* (Figure 6.5A) to recruit Me31B in the embryo. When injected into syncytial blastoderm embryos expressing Me31B::GFP, both the wild type GLS (Figure 6.5B-B'') and the *C. capitata* GLS-like signal (Figure 6.5D-D'') drive *grk* RNA localisation to the apical cortex within 20 minutes of injection in 92% and 100% of embryos, respectively (Figure 6.11B). Both RNAs are able to recruit Me31B to the apical cortex but *C. capitata* shows a slight reduction in recruitment, with 67% showing strong recruitment and 33% weak recruitment (Figure 6.5C-C'', 6.5E-E''; Figure 6.11B). These localisation and recruitment patterns are reflected in the fluorescent intensity profiles taken from a cross-section through the middle of the oocyte and compared to the non-localising control *hb* (Figure 6.5F). These results indicate that the sequence and structure divergences in the GLS-like signal from *C. capitata* only slightly reduce its association with Me31B in the embryo. This observation is in agreement with previous data showing that *grk* RNA containing the *C. capitata* GLS-like signal interdigitates with Me31B at the DA and anterior of the oocyte (Figure 5.14D-D''). From these results we can predict two regions of the signal in particular that may be important for the association with Me31B due to their high degree of conservation between the wild type GLS and the GLS-like signal from *C. capitata*, namely stem 1 and stem 4 (Figure 6.5A).

Conclusions The GLS-like signal from *C. capitata* recruits Me31B in the *D. melanogaster* embryo.

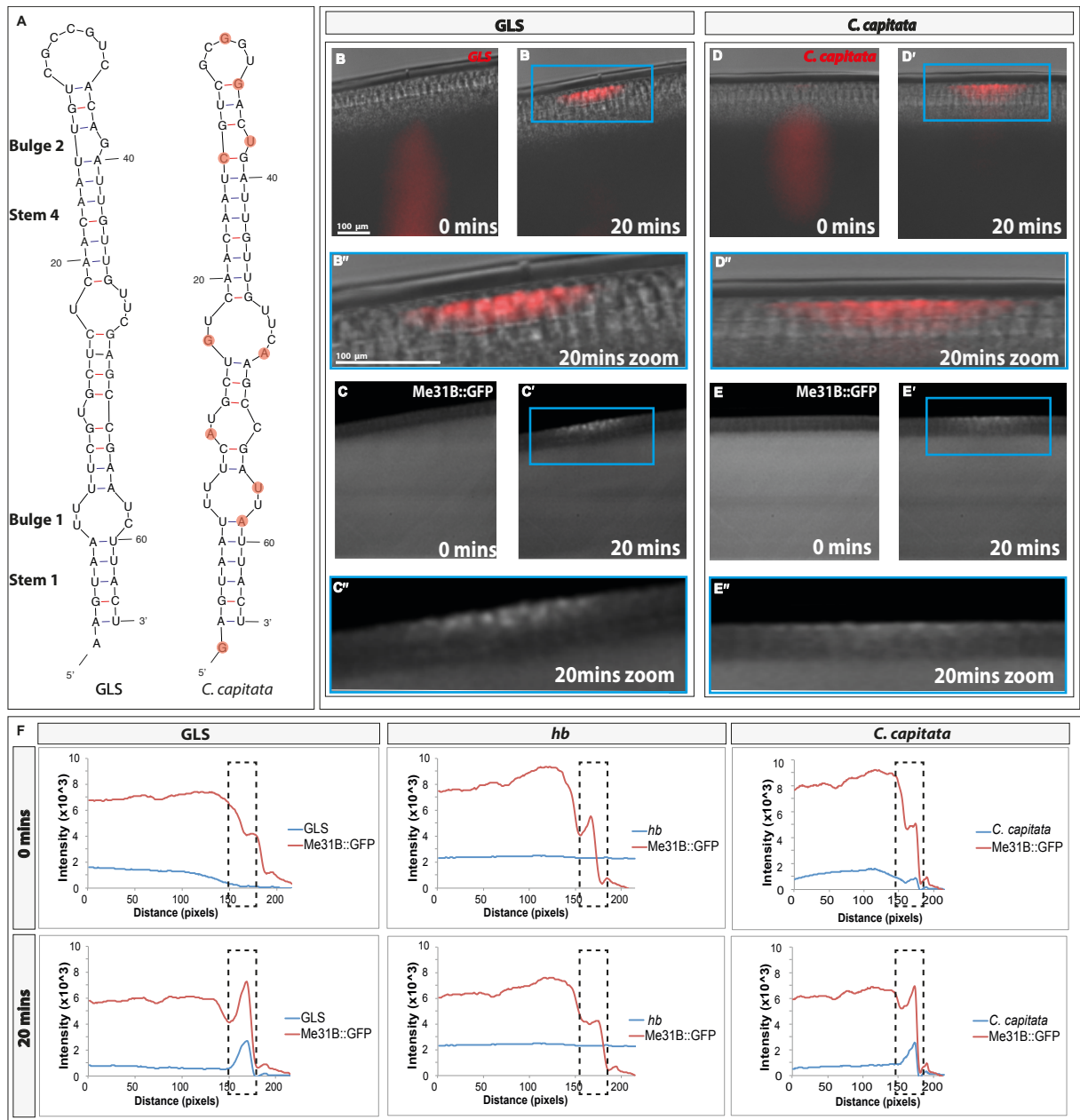


Figure 6.5. The GLS-like signal from *C. capitata* recruits Me31B to the apical cortex of the embryo. (A) UNAFold predicted secondary structures of the GLS and the GLS-like signal from *C. capitata*. The GLS (B-B'', n=12) and the GLS-like signal from *C. capitata* (D-D'', n=15) both drive *grk* RNA to the apical cortex of the embryo and are able to recruit Me31B within 20 minutes of injection (C-C'', E-E''), but the *C. capitata* signal is less efficient than the wild type. (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS and *C. capitata* conditions, but not in the non-localising control *hb*. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

The sequence of stem 4 is not required for Me31B recruitment

As the GLS-like signal from *C. capitata* is able to maintain interactions with Me31B, I carried out mutagenesis of the GLS to test whether conserved regions of the signal are required for its association with Me31B. Having shown that stem 4 is important for both *grk* transport and anchoring in Chapters 3 and 4, I tested whether this region is also important for the translational regulation of *grk* RNA through its interaction with Me31B. The mutant M42-8 contains several reflected base-pairs in stem 4 (Figure 6.6A) and was previously shown to transport efficiently in the oocyte (Figure 4.14D-D''). Both the wild type GLS (Figure 6.6B-B'') and M42-8 (Figure 6.6D-D'') localise to the apical cortex following injection into the embryo, with M42-8 showing strong localisation in 100% of embryos injected, and similar strong recruitment of Me31B to wild type in 100% of embryos (Figure 6.6C-C'', 6.6E-E''; Figure 6.11B). This is reflected in the fluorescent intensity profiles taken throughout the time course of the assay where the wild type GLS and M42-8 show substantial enrichment of Me31B at the same position as the RNA, whereas the non-localising control *hb* does not (Figure 6.6F). These results indicate that the sequence of the GLS stem 4 is not required for association of Me31B with *grk* RNA.

In order to investigate whether the position of stem 4 is important for Me31B association, the mutant M42-9, in which stem 4 is moved to the proximal third of the GLS (Figure 6.7A), was tested in the embryo recruitment assay. *grk* RNA containing the wild type GLS efficiently localises to the apical cortex of the embryo (Figure 6.7B-B'') and recruits Me31B::GFP (Figure 6.7C-C''), whereas the mutant M42-9 drives weaker localisation of *grk* RNA in 91% of embryos injected (Figure 6.7D-D''), due to inefficient recruitment of the transport component Egl (Figure 4.19H-H''), and shows a reduction in Me31B::GFP recruitment with only 45% of embryos showing strong levels of recruitment (Figure 6.7E-E''; Figure 6.11B). The ability of the wild type GLS and mutant M42-9 to drive *grk* localisation and Me31B recruitment can be compared to the non-localising control *hb* by measuring the fluorescent intensities of the RNA and protein through a cross-section of the embryo (Figure 6.7F). These data indicate

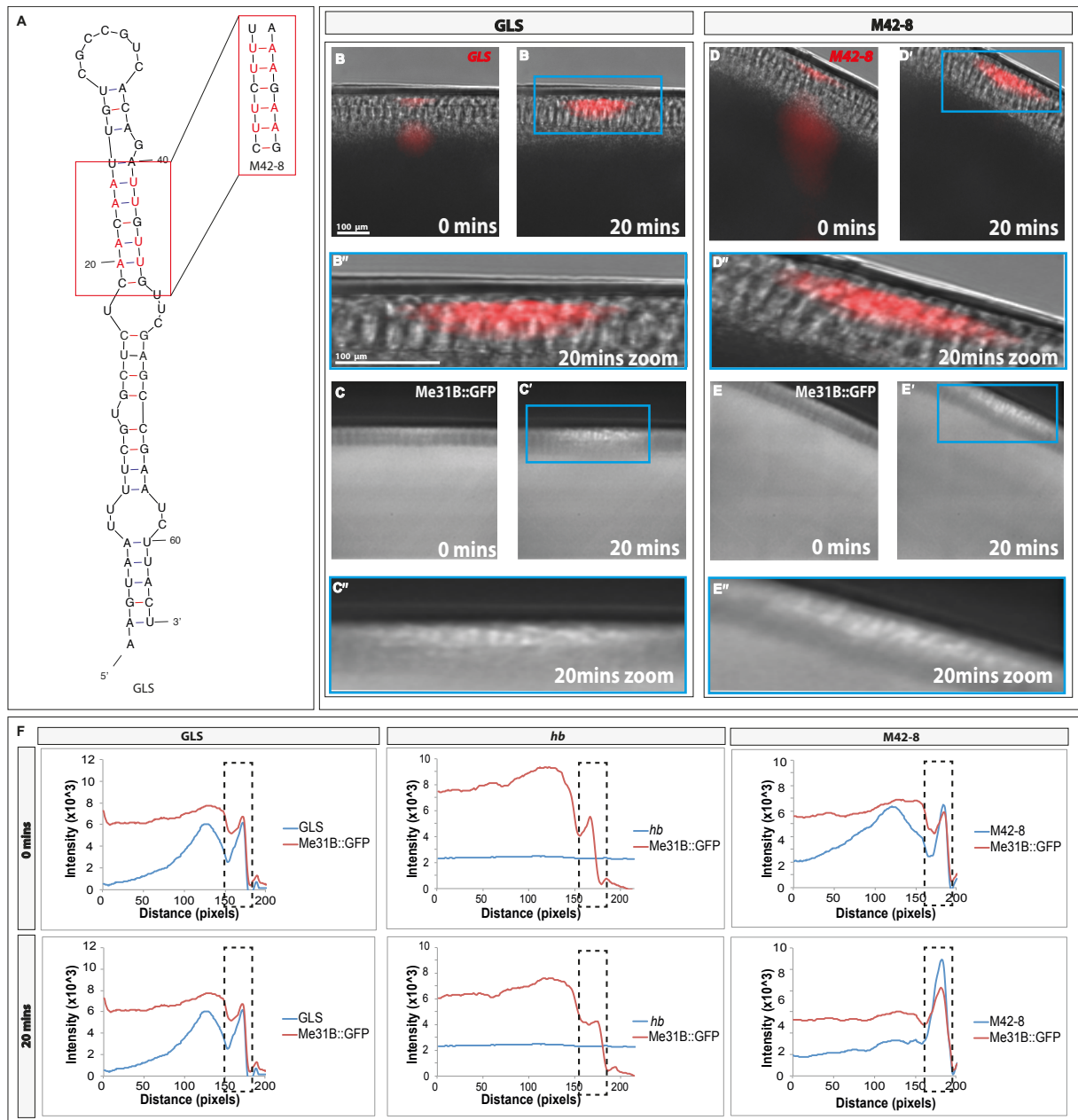


Figure 6.6. The sequence of stem 4 of the GLS is not required to recruit Me31B to the apical cortex of the embryo. (A) UNAFold predicted secondary structures of the GLS and the mutant M42-8 where several base-pairs in stem 4 are reflected. The GLS (B-B'', n=12) and M42-8 (D-D'', n=13) both drive *grk* RNA to the apical cortex of the embryo and are able to recruit Me31B within 20 minutes of injection (C-C'', E-E''). (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS and M42-8 conditions, but not in the non-localising control *hb*. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

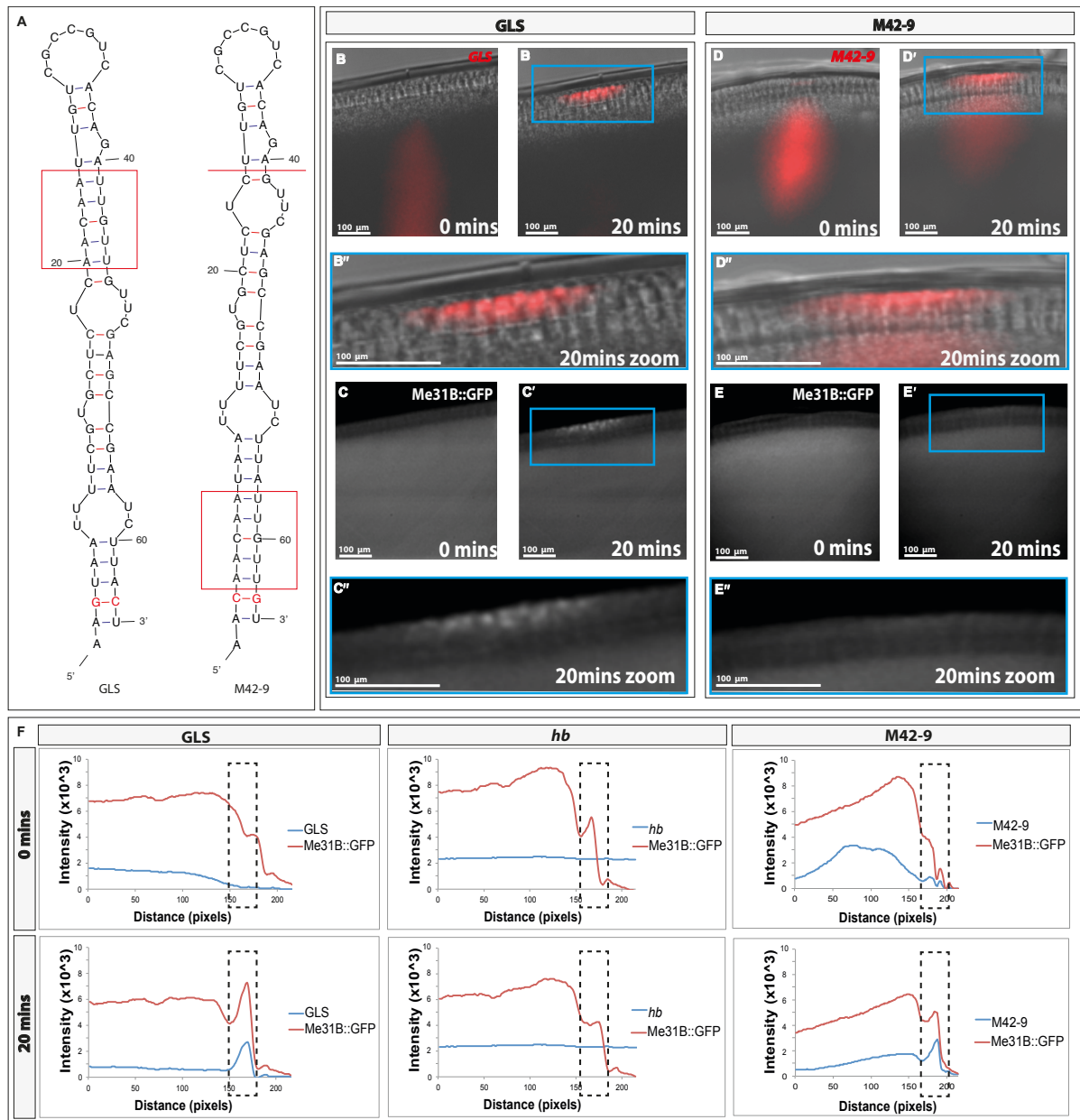


Figure 6.7. Moving stem 4 from the distal stem to the proximal stem of the GLS abolishes Me31B recruitment in the embryo. (A) UNAFold predicted secondary structures of the GLS and the mutant M42-9 where stem 4 is moved from the distal stem to within proximal stem 1. The GLS (B-B'') drives *grk* RNA to the apical cortex of the embryo and recruits Me31B within 20 minutes of injection (C-C'', $n=12$). M42-9 also drives *grk* RNA to the apical cortex, albeit less efficiently than the wild type GLS (D-D'', $n=11$), but only very weakly recruits Me31B (E-E'', $n=10$). (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS condition, but not in the non-localising control *hb*, and only the RNA intensity increases at the apical cortex for M42-9. (Blue box = region of interest at apical cortex used to compare RNA localisation and transacting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

that the position of stem 4 may be important for the association with Me31B, or that essential regions in the proximal stem are disrupted in this mutant.

Conclusions The sequence of stem 4 is not essential for the association between *grk* RNA and Me31B.

Stem 1 is essential for Me31B recruitment

To investigate whether M42-9 fails to recruit Me31B due to a change in the position of stem 4 or the disruption of stem 1, the truncated signal M1 was tested in the embryo recruitment assay. M1 is a 49-nucleotide signal truncated just above internal loop 1, removing both internal loop 1 and stem 1 (Figure 6.8A), and was shown to have slightly reduced transport efficiency in the oocyte (Figure 4.6F-F’). When injected into the embryo, the wild type GLS drives efficient transport of *grk* RNA to the apical cortex (Figure 6.8B-B’; Figure 6.11B). M1 shows strong apical localisation in 30% of embryos injected (Figure 6.8D-D’; Figure 6.11B), and weaker apical localisation in a further 70% with higher levels in the basal cytoplasm than in wild type. In contrast to the wild type GLS which recruits Me31B to the apical cortex in 92% of cases (Figure 6.8C-C’; Figure 6.1B), M1 does not recruit Me31B in 100% of embryos injected (Figure 6.8E-E’; Figure 6.11B). The failure of M1 to recruit Me31B despite its localisation to the apical cortex is reflected in the fluorescent intensity profiles taken through a cross-section of the embryo over time (Figure 6.8F). These results show that the proximal third of the GLS is essential for the association between *grk* RNA and Me31B, and is therefore likely to be required for correct translational regulation of the transcript. As disruption of stem 1 in mutant M42-9 only reduces the efficiency of Me31B recruitment but does not abolish it, it is likely that the lack of recruitment seen for M1 is related to a joint requirement for internal loop 1 and the sequence of stem 1.

Conclusions The proximal third of the GLS is essential for the association between *grk* RNA and Me31B.

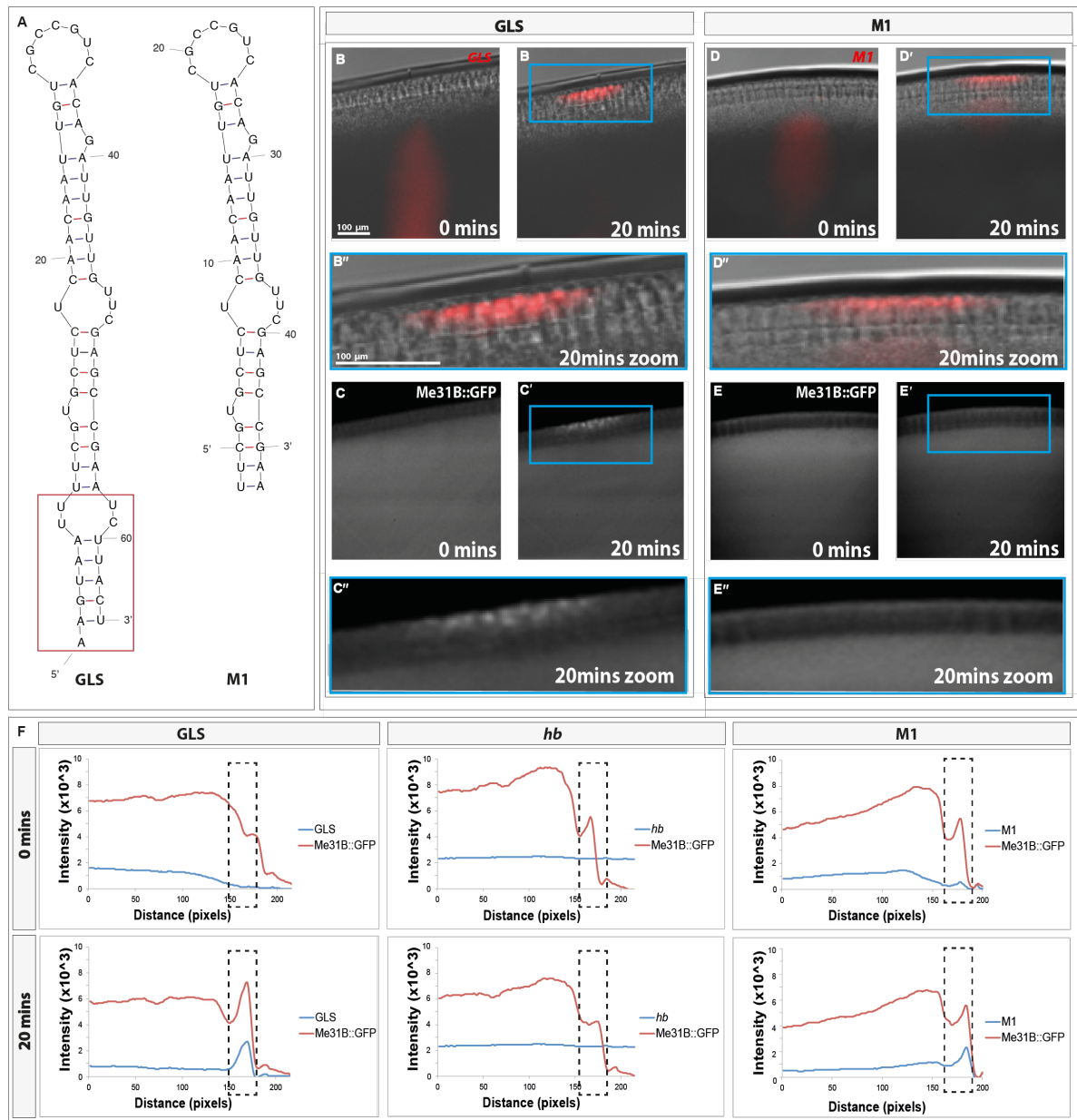


Figure 6.8. The proximal third of the GLS is essential for Me31B recruitment to the apical cortex of the embryo. (A) UNAFold predicted secondary structures of the GLS and the mutant M1 where stem 1 and internal loop 1 are removed from the proximal third of the signal. The GLS (B-B'', $n=12$) drives *grk* RNA to the apical cortex of the embryo and recruits Me31B within 20 minutes of injection (C-C'', $n=11$). M1 also drives *grk* RNA to the apical cortex (D-D'', $n=10$), but is unable to recruit Me31B (E-E'', $n=10$). (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS condition, but not in the non-localising control *hb*, and only the RNA intensity increases at the apical cortex for M1. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

The sequence of the internal loops is not essential for Me31B recruitment

To further investigate whether M1 fails to recruit Me31B due to its lack of stem 1 or internal loop 1, the sequence of the internal loops were mutated to create M40-1 (Figure 6.9A) which shows only a slight reduction in transport efficiency in the oocyte (Figure 4.20F-F"). Wild type *grk* RNA (Figure 6.9B-B") and M40-1 (Figure 6.9D-D") drive apical *grk* localisation in 92% and 100% of embryos, respectively (Figure 6.11B). However, M40-1 shows a reduction in the efficiency of Me31B recruitment, with only 50% of embryos showing strong levels of Me31B recruitment and 36% weak recruitment (Figure 6.9E-E"; Figure 6.11B) compared to 91% strong recruitment for wild type *grk* (Figure 6.9C-C"; Figure 6.11B). The enrichment of both *grk* and Me31B at the apical cortex of the embryo is displayed as two peaks at the same position in the fluorescent intensity profiles (Figure 6.9F). These results suggest that the sequence of the internal loops are required for maintaining the high efficiency of association between *grk* RNA and Me31B, suggesting that sequences required for Me31B binding may be disrupted in this mutant.

Conclusions The sequence of the GLS internal loops are important for the efficient recruitment of Me31B to *grk* RNA, but are not essential.

Bulge 2 is not required for Me31B recruitment

Having previously shown that bulge 2 is essential for *grk* RNA anchoring (Chapter 5, this thesis) I investigated whether it also played a role in translational regulation through the association with Me31B. In the GLS mutant M44-1, three base-pairs around bulge 2 are reflected, maintaining the predicted secondary structure of the RNA but altering the sequence (Figure 6.10A). Injection of M44-1 into the oocyte results in efficient transport in 66% of oocytes but poor anchoring at the DA (data not shown). Injection of the GLS (Figure 6.10B-B") and M44-1 (Figure 6.10D-D") into the embryo results in strong localisation to the apical cortex of 92% and 50% of embryos, respectively (Figure 6.11B).

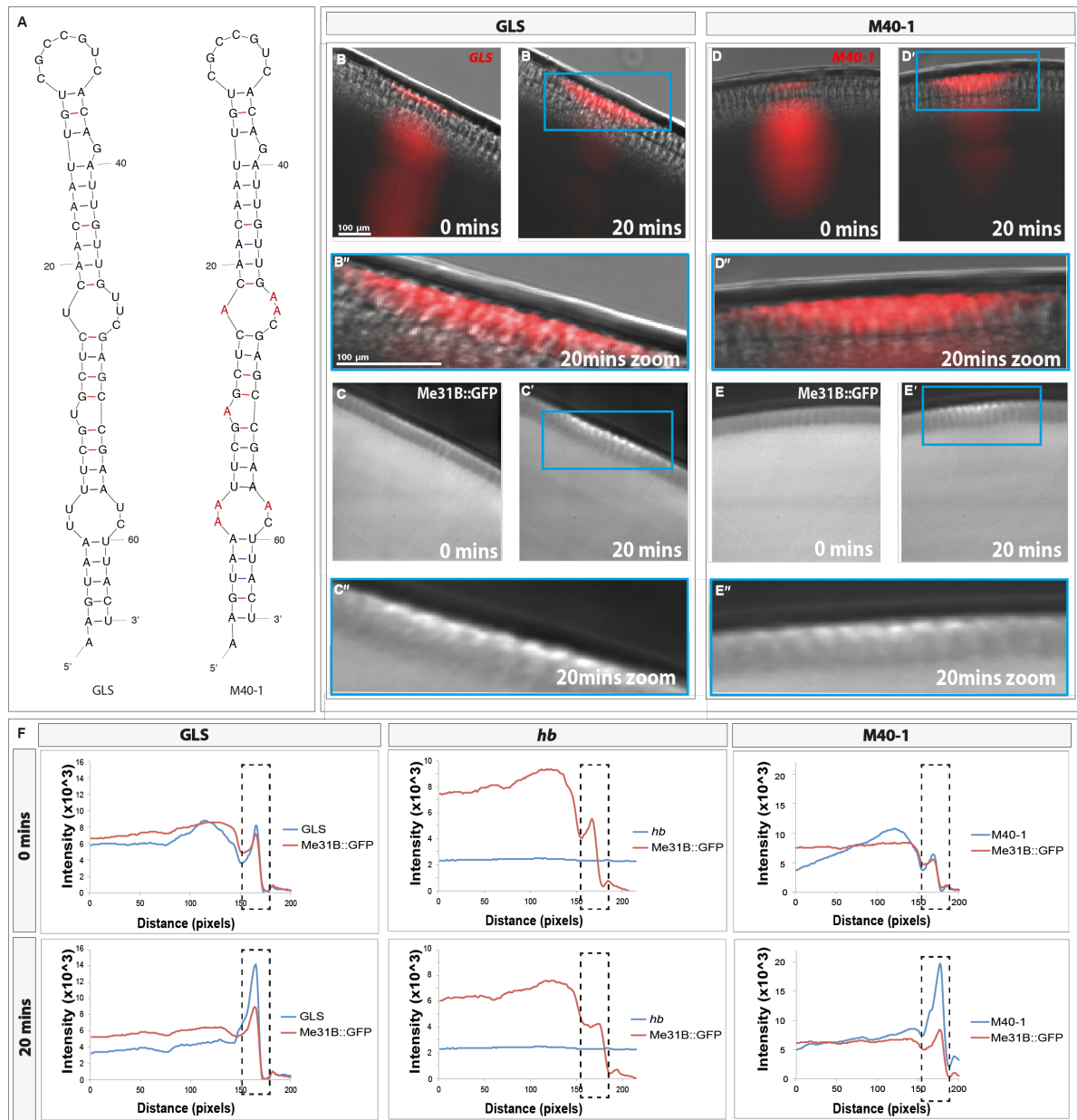


Figure 6.10. The sequence of the GLS internal loops is not essential for Me31B recruitment in the embryo. (A) UNAFold predicted secondary structures of the GLS and the mutant M40-1 where the uracil bases in the internal loops and bulge 1 are mutated to adenines. The GLS (B-B'', $n=12$) drives *grk* RNA to the apical cortex of the embryo and recruits Me31B within 20 minutes of injection (C-C'', $n=11$). M40-1 also drives *grk* RNA to the apical cortex (D-D'', $n=14$) and shows strong Me31B recruitment in half of the embryos injected (E-E'', $n=14$). (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS and M40-1 conditions, but not in the non-localising control *hb*. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

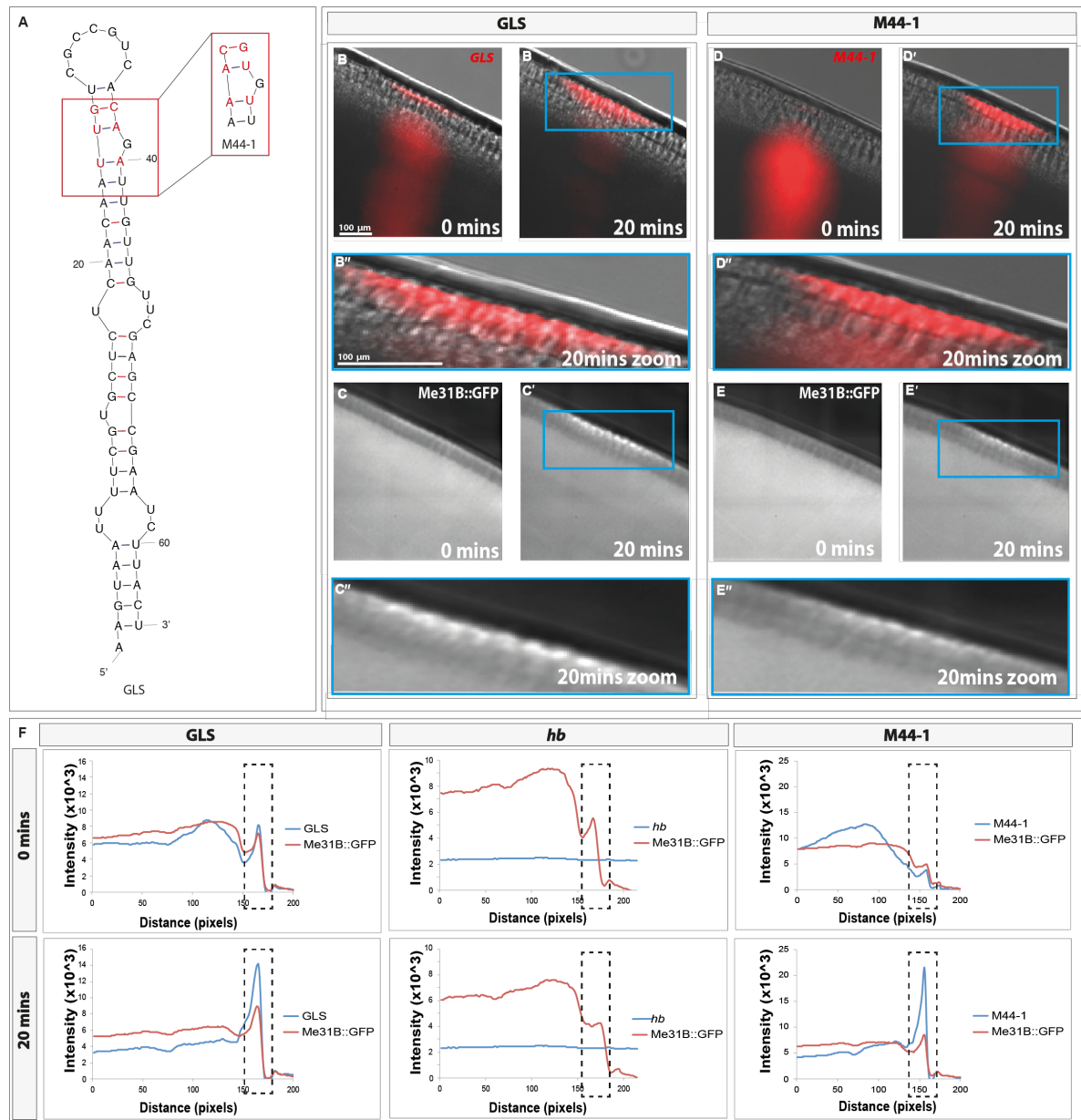


Figure 6.10. The sequence surrounding bulge 2 is important for efficient Me31B recruitment in the embryo. (A) UNAFold predicted secondary structures of the GLS and the mutant M44-1 where the base-pairs around bulge 2 are reflected. The GLS (B-B'', n=12) drives *grk* RNA to the apical cortex of the embryo. M44-1 (D-D'', n=6) localises more weakly than the GLS, with more RNA remaining in the basal cytoplasm, and recruits Me31B less efficiently than wild type *grk* RNA within 20 minutes of injection (C-C'', E-E''). (F) Fluorescent intensity profiles through a cross-section of the embryo show that both the RNA (blue line) and Me31B (red line) accumulate at the apical cortex (black box) over time in the GLS and M44-1 conditions, but not in the non-localising control *hb*. (Blue box = region of interest at apical cortex used to compare RNA localisation and *trans*-acting protein recruitment. Distance in (F) refers to the position through the cross-section of the embryo (see Figure 6.1) from the site of injection to the apical cortex (black box)).

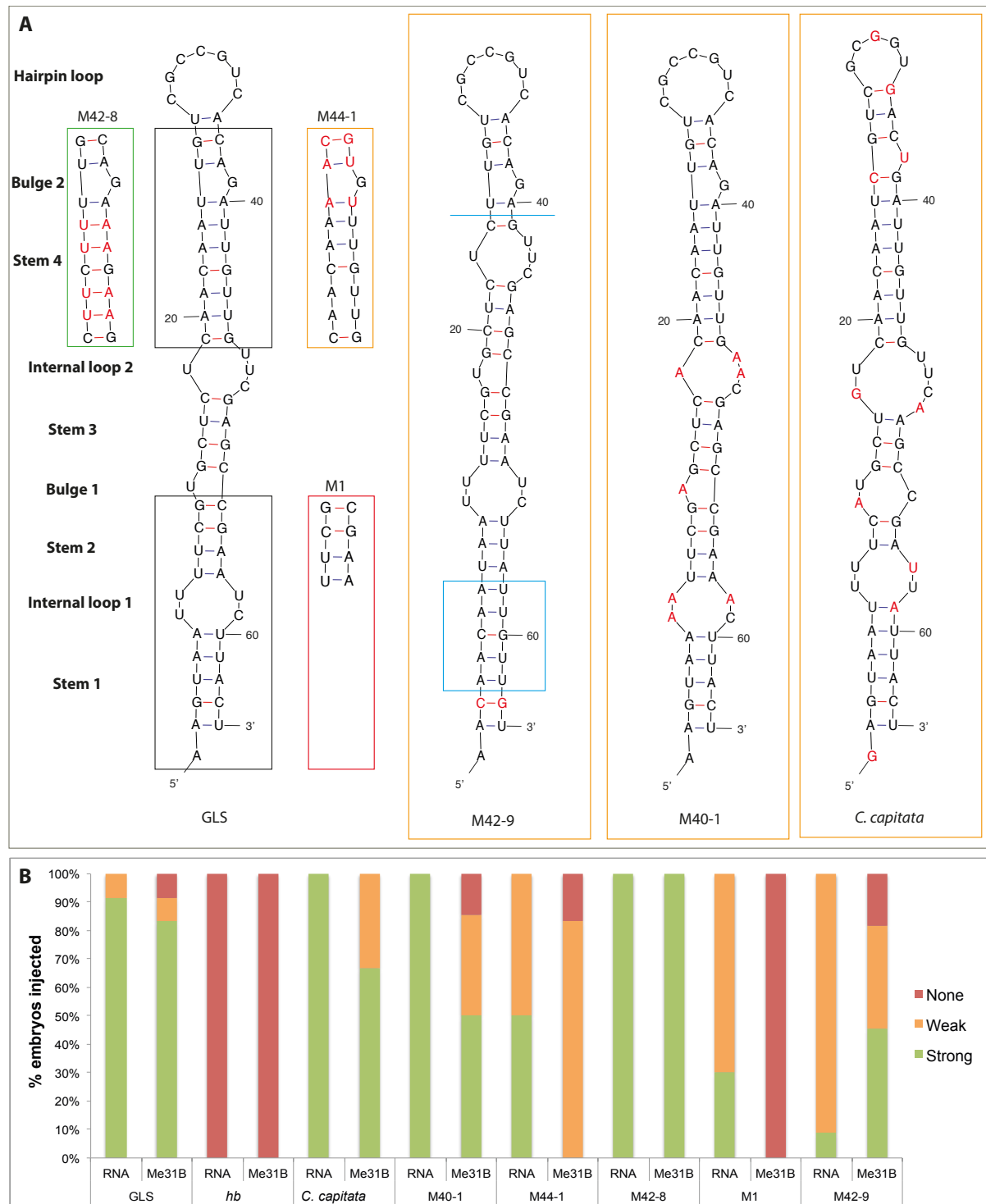


Figure 6.11. Summary of the affect of mutations within the GLS on recruitment of Me31B to the apical cortex of the embryo. (A) UNAFold predicted secondary structure of the GLS and the mutants investigated in this chapter (green box = strong recruitment of Me31B, orange box = weak recruitment, red box = no recruitment). (B) Graph summarising the percentage of embryos injected showing strong, weak and no apical localisation of *grk* RNA and Me31B::GFP. For definitions of strong, weak and none please refer to the materials and methods.

M44-1 is able to recruit Me31B::GFP to the apical cortex of 83% of embryos within 20 minutes (Figure 6.10E-E'') but more weakly than wild type GLS (Figure 6.10C-C''), as shown in the fluorescent intensity profiles generated for each RNA following injection (Figure 6.10F). The remaining 17% of embryos injected with M44-1 show no recruitment of Me31B (Figure 6.11B). These results indicate that the sequence of bulge 2 may be partly required for the efficient recruitment of Me31B, and therefore may be important for translational regulation of *grk* RNA.

Conclusions The sequence of bulge 2 is important for the efficiency of Me31B recruitment to *grk* RNA.

Discussion

***grk* RNA is actively transported to the apical cortex of the embryo where it recruits Me31B**

Although Me31B has previously been suggested to act as a repressor of *grk* translation, previous studies have shown that *grk* RNA and Me31B do not travel in the same transport particles in both the oocyte¹⁹³ and the NCs (Alex Davidson, unpublished). Moreover, in young Me31B mutant egg chambers Grk expression is restricted to the oocyte suggesting that Me31B is not required for translational repression of *grk* in the NCs.³²⁹ However, Me31B has been shown to associate with *grk* using biochemistry,⁵⁴ and *grk* interdigitates with Me31B at the edge of P bodies at the DA of the oocyte where it becomes translationally activated. In embryos, Me31B-containing P bodies disassemble upon egg activation allowing translationally repressed transcripts, such as *bcd*, to be released from the core of P bodies and to therefore be translated. Despite this research, the exact role of Me31B in regulating *grk* translation in the egg chamber, and regulating the translation of other mRNAs in other tissues such as the embryo, is currently unknown.

My data reveal that, upon injection into the embryo, *grk* RNA is transported to the apical cortex, as previously shown in the literature and Chapter 4 of this thesis.¹¹⁶ In addition, we show for the first time that localised *grk* RNA efficiently recruits Me31B to the apical cortex of the embryo. Immediately following injection into the embryo, *grk* RNA forms transport particles that are transported through the cytoplasm towards the apical cortex. *grk* RNA found at the site of injection, and being transported through the cytoplasm, does not colocalise with Me31B, suggesting that *grk* transport particles move independently of Me31B protein, although this would have to be tested in Me31B mutants. Upon injection of *grk* RNA, Me31B forms large puncta that move toward the apical cortex independently of *grk* RNA.

Me31B accumulation in the embryo may represent P body-like structures required for the translational regulation of *grk* RNA

At the DA corner of the oocyte, *grk* and Me31B interdigitate as a result of *grk* RNA docking at the edge of P bodies, where it is translationally activated through the accessibility to ribosomes and translational activators.^{133, 193} In contrast, at the apical cortex of embryos Me31B and *grk* colocalise in large puncta that resemble P bodies. It has previously been shown that injection of high concentrations of *grk* RNA, or overexpression of endogenous *grk*, results in targeting of the RNA to the core of P bodies causing *grk* and Me31B to colocalise at the DA of the oocyte.¹⁹³ It has been proposed that the core of P bodies lacks ribosomes and contains high levels of translational repressors, inhibiting the translation of RNAs.¹⁹³ Targeting of excess *grk* RNA to the core of P bodies is therefore likely to result in its translational inhibition and possibly degradation. The colocalisation of *grk* and Me31B at the apical cortex of the embryo suggests that, instead of docking at the edge of P bodies, *grk* is targeted to the core. As *grk* is not usually found in the embryo, targeting of the RNA to the P body core is likely to be a mechanism for translationally silencing, or degrading, the ectopic *grk* RNA. Alternatively, colocalisation of *grk* and Me31B could result from injection of high levels of *grk*, as seen in the oocyte.¹⁹³ However, we favour the former argument as *grk* RNA was injected at moderate concentrations (360ng/μl) compared to the concentrations used in

the embryo recruitment assays in the literature ($1\mu\text{g}/\mu\text{l}$).¹¹⁶ Moreover, the injected concentration used to represent overexpression of *grk* in the much smaller oocyte was $500\text{ng}/\mu\text{l}$,¹⁹³ suggesting that the concentration used here for the much larger embryo is unlikely to force *grk* to the core of P bodies. However, further characterisation of the Me31B clusters is required to determine if they truly represent P bodies.

The proximal third of the GLS is essential for the association between *grk* RNA and Me31B

In this chapter I have shown that the recruitment of Me31B to injected *grk* RNA in the embryo is somewhat maintained in the GLS-like signal from *C. capitata*, but can be abolished by mutations within the GLS. In line with biochemical studies showing that Me31B associates with the GLS,⁵⁴ these results suggest that the interaction between Me31B and *grk in vivo* may be orchestrated via the GLS. By testing a number of GLS mutants I have demonstrated that disrupting a number of regions of the GLS abolishes Me31B recruitment, suggesting that features of the GLS are necessary for the association between *grk* and Me31B. For example, the sequence of the region surrounding bulge 2, which I have shown in Chapter 5 to be required for restricting *grk* anchoring to the DA of the oocyte, is required for efficient Me31B recruitment. In addition, the sequence of stem 4 is not important for Me31B recruitment. Finally, the presence of the proximal third of the GLS, composed of stem 1 and internal loop 1, is absolutely essential for the association between *grk* and Me31B. Disrupting stem 1 by adding an additional sequence greatly reduces Me31B recruitment in the embryo, and changing the uracil bases of the internal loops also reduces Me31B recruitment. Taken together, these data suggest that sequences or structures within both stem 1 and internal loop 1 may be essential for the association between Me31B and *grk* RNA. It is possible that Me31B associates via a binding site within the proximal third of the GLS that overlaps between stem 1 and internal loop 1. As a result, changing the sequence of either of these regions would reduce the efficiency of Me31B binding, as seen in mutants M42-9 and M40-1,

and removing both regions completely would abolish Me31B recruitment, as seen in the mutant M1.

The data presented in this chapter indicate that the proximal third of the stem-loop is essential for the association between *grk* RNA and Me31B, and may therefore be important for the translational regulation of *grk* by controlling its association with P bodies. This is in agreement with data showing that the GLS-like signal from *C. capitata* is able to recruit Me31B in the embryo, albeit less efficiently than wild type *grk*, which may be a result of its high conservation within the proximal third of the signal. However, further mutational analysis of this region is required to dissect the sequence or structural motifs within the proximal stem that are essential for the association of Me31B.

Future work

To confirm whether the Me31B-rich structures that form in the embryo upon *grk* injection are P bodies, their further characterisation is required. For example, the presence of other components of P bodies could be investigated, such as an enrichment of Sqd and Orb at the periphery, and Cup and Tral at the core.¹⁹³ In addition, electron microscopy could be used to determine whether the Me31B clusters are indeed electron dense P bodies, and whether they play a role in the translational regulation of mRNAs endogenously expressed in the embryo, such as pair-rule transcripts. Moreover, the position of *grk* RNA within the Me31B-rich structure could be further studied using EM. This would determine whether *grk* RNA is located inside the P body core at the apical cortex of the embryo where it could be translationally silenced. The role of the Me31B-rich structures in the translational control of *grk* RNA could be investigated by staining for the ectopic expression of Grk protein in the embryo following injection of translationally-competent *grk* RNA. If the Me31B-rich structures are found to be P bodies, the recruitment assay in the embryo could provide a foundation upon which the temporal control of P body formation can be investigated in relation to translational regulation, addressing the question as to whether P bodies form as a cause or a

consequence of RNA silencing. Biochemical and structural studies could aid the identification of sequence or structural motifs within the proximal third of the GLS that are essential for the association between *grk* and Me31B, and are therefore likely to be important for the translational regulation of *grk* RNA. Expression of GLS mutants that fail to recruit and bind Me31B could then be expressed endogenously in the oocyte to determine whether the synthesis of Grk protein is misregulated. Finally, the exact role of Me31B in the translational regulation of *grk* in the egg chamber and at the DA P bodies needs to be established in order to dissect the relationship between *grk* anchoring and translation.

Final conclusions

Taken together, the data presented in this chapter demonstrate that injected *grk* RNA is able to recruit Me31B in the embryo, and stimulates the formation of novel P body-like Me31B accumulates at the apical cortex where *grk* becomes localised. *grk* RNA and Me31B do not travel together in transport particles but colocalise at the apical cortex, suggesting *grk* is closely associated with Me31B, rather than interdigitated as seen at the DA of the oocyte. Sequence or structural elements within the proximal third of the GLS are essential for the association between *grk* RNA and Me31B, suggesting it provides binding sites for the protein. Other regions of the GLS, including internal loop 1 and bulge 2 are not essential for the association between *grk* and Me31B, but are important for the efficiency of Me31B recruitment.

Chapter 7: Discussion and Future Perspectives

The information required for mRNA localisation, and localised translation at their target destinations, is encoded in *cis*-acting regions of the transcript which act through recruitment of a particular complement of *trans*-acting protein factors. However, the exact molecular nature of these interactions and how they allow the differential localisation of different species of transcripts is not known, except in a few limited cases, such as *ASH1* mRNA in yeast. Here I have carried out a detailed molecular dissection of the *grk* localisation signal (GLS, Figure 7.1A) in order to characterise its interaction with the key factors Egl, Sqd and Me31B at different stages throughout *grk* transport, anchoring and translation (Figure 7.1B).

GLS-like signals from other species can be used to predict regions of the GLS that are important for *grk* transport

The GLS was previously shown to be highly conserved in a subset of *Drosophilids* and able to localise following injection into the *D. virilis* oocyte.^{305, 307} Here, I have taken advantage of the recent sequencing of additional genomes to demonstrate that GLS-like signals also exist in a subset of Dipterans and are highly similar to the GLS in their sequence and predicted secondary structure. Most of the signals are able to drive *grk* transport, as assayed by injections into the *D. melanogaster* oocyte, suggesting that the transport machinery and adapter proteins are also conserved.

Using mutagenesis studies I have demonstrated that small regions of the GLS with low conservation, such as stem 1 and internal loop 1, often coincide with those regions that are redundant for *grk* transport. In contrast, regions of high conservation, such as stems 3 and 4, are essential for recruitment of the transport machinery and therefore *grk* transport.

An AU-rich stem, internal loop and purine stack motif may direct recruitment of the key adapter Egl

In the introduction I describe the interactions of motor proteins and the cytoskeleton and the critical role of cargo adapters, such as Egl, in transporting cargo.^{108, 116, 118, 270, 304} Studies investigating motifs within the RNA required for Egl binding have shown that Egl recognises two spatially registered A'-form RNA helices in *K10* RNA generated by two runs of three or more base-paired purines on the same side of a stem.³⁸⁶ This conformation of dsRNA widens the major groove increasing the accessibility in the major groove for protein interactions. However, it is currently unknown whether A'-form RNA is essential for interactions between Egl and all its RNA cargoes.

Here, I show that the two purine stacks in stems 2 and 3 (Figure 7.1A, pink) that are most likely to generate A'-form RNA in the GLS can be mutated in such a way that is predicted to disrupt this conformation, without affecting Egl recruitment. These results suggest that A'-form RNA in stems 2 and 3 of the signal is not essential for Egl to recognise the GLS. However, I have identified further regions of the GLS that affect *grk* transport which could potentially generate A'-form structures.

An AU-rich stem at the distal end of *K10* forms one of the two required A'-form RNA helices due to its high degree of purine stacking on the same side of the stem. The sequence of this region was shown not to be required for providing specificity for Egl binding.³⁸⁶ Although the GLS AU-rich distal stem 4 (Figure 7.1A, green) is not predicted to generate A'-form RNA due to the lack of purine stacking, in agreement with the *K10* studies, I show that the sequence of stem 4 is not required for transport, arguing against base-pair interactions providing specificity. However, in contrast to *K10*, transversion mutations of the AU-rich stem to a GC-rich stem abolishes *grk* transport, as does reducing the size of the GLS hairpin loop (Figure 7.1A, brown) to less than 5-nucleotides. This suggests that the accessibility of stem 4 for protein binding, and the AU-richness of this stem, may be essential for transport. This

accessibility could be achieved by the presence of a large hairpin loop, or unpredicted A'-form conformations in stem 4, but ultimately only structural analysis would test this definitively.

Regardless of whether the GLS AU-rich stem folds into an A'-form RNA, mutational analysis of the GLS led me to propose a possible mechanism for Egl binding. Namely, that Egl binding depends on both the AU-rich stem in the distal third of the GLS in a sequence-independent manner, and a purine stack in stem 3 in a sequence-dependent manner. I propose that a kink in the GLS between these two stems produced by internal loop 2 provides the correct distances required for the formation of a stable motor complex. Finally, a hairpin loop of more than five nucleotides is required to generate distortions in the helical AU-rich stem, widening the major groove and providing enough space and flexibility for *trans*-acting factors to gain access to nucleotides in the stem (Figure 7.1C, 7.1D). I propose that this could be achieved by either a single molecule of Egl binding to both stem regions, or through the binding of two Egl molecules, one to each stem on either side of the internal loop. Other regions of the GLS that are likely to contribute to the full efficiency of *grk* transport are stem 1 and internal loop 1 in the proximal third of the signal.

The finding that moving the AU-rich stem to the proximal end of the GLS within stem 1 reduces, but does not completely abolish, *grk* transport and Egl recruitment favours the notion that multiple copies of Egl bind the GLS. This is supported by studies showing that individual cargoes can simultaneously associate with multiple motors,⁴⁰⁶ and that increasing the copy number of dynein-dynactin per RNP can increase the probability of processive unidirectional minus-end directed movement.¹³⁰ For example, *ASH1* mRNA in yeast recruits multiple copies of Myo4 via the adaptor proteins She2 and She3 at four different *cis*-acting localisation signals, enhancing transport of the mRNA along the actin cytoskeleton. Moreover, the binding of Egl to mRNAs has been suggested to be somewhat promiscuous, with overexpression of the protein resulting in the targeting of a small amount of nonlocalising transcripts to the apical cortex of the embryo.¹⁰⁸ Therefore, I propose that Egl

can bind stems 3 and 4 independently to drive some limited *grk* transport. However, processive movement of *grk* requires Egl to bind to both the AU-rich stem 4 and stem 3 on either side of the internal loop, thus generating a processive motor complex. Structural studies of the GLS in complex with Egl could shed light on whether multiple copies of the protein bind the RNA.

Many transcripts are localised via the Egl-BicD-Dynein motor complex in several tissues and are likely to contain cryptic recognition motifs for Egl binding as they share little primary sequence conservation.³⁰⁴ As my results suggest that, for *grk*, A'-form conformation does not appear to be essential for transport, I searched for the motifs required for *grk* localisation, including an AU-rich stem and an internal loop followed by a purine stack, in other Egl-transported transcripts, and showed that these features are highly conserved (Table 7.1). Mutational analysis of these regions would be required to test their role in Egl recruitment and to explain why some Dynein-transported mRNAs have an internal kink in their localisation signal and some do not.

As many mRNAs that localise to distinct regions of the cell use the same transport machinery, cytoskeletal tracks, and often the same motor adaptor proteins, additional proteins must be required to generate specificity, including anchoring factors. For example, Sqd has been shown to be required previously for anchoring in the case of *grk*.^{133, 283}

The role of the GLS in directing *grk* anchoring and translation can be mapped to distinct regions of the signal

The anchoring and translation of mRNAs are predicted to be tightly linked, and translational silencing has been proposed to be required for mRNA localisation.⁴⁰⁷ For example, *ASH1* mRNA anchoring requires specific secondary structures within the ORF and active translation of the C-terminal domain,¹⁴³ and *grk* is anchored and translated at the edge of P

bodies where translational activators such as Orb are highly enriched.^{133, 193} It is worth pursuing

RNA	AU-rich stem	Internal kink	Purine stack
<i>grk</i> (GLS)	✓	✓	✓
<i>I-factor</i> (ILS)	✓	✓	✓
G2 (G2LS)	✓	✓	✓
<i>ftz</i> (FLE)	✓	✓	✓
<i>eve</i>	✓	✓	✓
<i>bicoid</i> (BLE)	✓	✓	✓
<i>hairy</i> (HLE)	✓	✓	✓
<i>K10</i> (TLS)	✓		✓
<i>orb</i>	✓		✓
<i>wingless</i> (WLE3)	✓		✓
<i>oskar</i> (OES)	✓	✓	✓
<i>nanos</i>	✓	✓	✓

Table 7.1. Features of the GLS shown to be essential for *grk* transport are conserved in several Egl-localised transcripts. Several transcripts localise to the minus-ends of microtubules in the oocyte and embryo using the Dynein motor and the accessory protein Egl. Most of the localisation signals responsible for directing transport of these transcripts contain similar features to regions of the GLS shown to be essential for *grk* transport, including an AU-rich stem, a purine stack with an internal kink between them. These features may represent universal recognition motifs for Egl.

whether the function of the GLS in controlling these two processes can be mapped to distinct regions of the signal. My data show that mutations affecting the accessibility of bulge 2 (Figure 7.1, brown) in the distal third of the GLS cause aberrant *grk* interdigitation with the P body component Me31B at the anterior of the oocyte in addition to the DA. These data are in line with two of the putative Sqd-binding sites predicted in the GLS from SELEX data, and suggest that Sqd binds to sequences within the helical regions surrounding bulge 2, with accessibility to the nucleotides being provided by the presence of the bulge (Figure 7.1C, 7.1D).

I show that the proximal third of the GLS (Figure 7.1A, red), on the other hand, is essential for recruitment of the P body component Me31B in the embryo, with sequences within the internal loops and bulge 2 being required for full efficiency. As Me31B is involved in the translational regulation of *grk*, these results suggest that the regions of the GLS responsible for controlling its interaction with Me31B may be required for translational regulation of the transcript (Figure 7.1E). Therefore, the multiple functions of the GLS in controlling *grk* localisation and translation are likely to map to distinct motifs within the signal, with overlapping regions being required for full efficiency. Future work would involve characterising the specific motifs required for Egl, Sqd and Me31B binding in these regions of the GLS using biochemical techniques or structural studies.

Some mRNAs form an anterior ring but no DA cap during mid-oogenesis. The most studied example is *bcd* which localises to the anterior of the oocyte where it is translationally repressed during mid-oogenesis. Although Sqd has been shown to bind to *bcd* mRNA in biochemical studies,^{395, 54} *bcd* localisation is unaffected in a *sqd*^{4B4} mutant and Sqd has not been genetically identified as being required for *bcd* anchoring in later stages.⁴⁰⁸ In later stages of oogenesis, *bcd* is instead maintained at the anterior of the oocyte by continued active transport along microtubules.¹²¹ Characterising the behaviour of *bcd* during mid- and

late-oogenesis could provide insights into the different mechanisms at play that mediate translational repression and promote or inhibit mRNA anchoring in this region of the oocyte. Does Sqd inhibit anchoring of both *bcd* and *grk* at the anterior of wild type oocytes? Furthermore, it would be interesting to test whether adding the GLS to *bcd* mRNA could promote its anchoring at the DA during mid-oogenesis.

Re-interpreting the mechanism of Sqd in *grk* anchoring and translation

Current studies suggest a dual role for Sqd in both the repression of *grk* translation during transport,^{314, 317, 383, 403} and its translational activation by switching Dynein from a processive motor to a static anchor, docking *grk* at the edge of DA P bodies.^{133, 193} However, in a *sqd1* mutant Grk is ectopically expressed around the anterior of the oocyte but not in the rest of the egg chamber,^{317, 329} raising several unanswered questions. If Sqd is required for repressing *grk* translation during transport, why is *grk* not ectopically translated in the rest of the *sqd1* mutant egg chamber? Moreover, is the role of Sqd in *grk* translational repression different between these cell types; and does Sqd play a more specific role in repressing *grk* translation at the anterior of the oocyte? Alternatively, is Grk expressed throughout the *sqd1* mutant egg chamber at low levels making it undetectable by antibody staining in regions where the RNA is not specifically enriched?

Sqd has been shown to be required in the NCs for translational repression of *grk* at the ventral-anterior (VA) of the oocyte.^{317, 383} In the NCs, however, *grk* translation is instead inhibited by the absence of the translational activator Orb. At the DA of the oocyte, hyperphosphorylated Orb is enriched and is responsible for recruiting the polyadenylase Wispy to anchored *grk*, resulting in lengthening of the *grk* poly(A) tail and activation of translation (Figure 7.1E).^{404, 411} Inappropriate translation of *grk* in NCs through over-expression of Orb involves association with P bodies, and it is not clear whether this association is also a requirement for the ectopic expression along the anterior margin of the oocyte in *sqd1* mutants.³²⁹ In the *grk* anchoring process, how does Sqd discriminate between

P bodies at the DA and elsewhere in the NCs and oocyte? How is Sqd switched from a repressor to an activator of translation specifically at the DA?

Whilst developing an assay to screen for *sqd*-like GLS mutants, I made several interesting observations in the *sqd1* mutant. First, injected and MS2-labelled *grk* is partly enriched at the DA of the oocyte in our *sqd1/Df* mutant flies, albeit at a lower proportion than in wild type oocytes. Although several studies have observed a similar DA enrichment in the *sqd* mutant,^{15, 133} this pool of *grk* at the DA has not been thoroughly investigated.

Second, *grk* RNA seems to form larger puncta at the anterior and DA of the *sqd1* mutant than at the anterior of wild type oocytes. *grk* RNA at the DA of the *sqd1* oocyte shows similar characteristics to that anchored at the DA of wild type oocytes in terms of dynamics and its association with P bodies. This is in agreement with studies showing that a small proportion of *grk* RNA is found in P bodies at the DA corner of *sqd1* mutants.¹³³ In addition, my data suggests that *grk* RNA at the anterior of the *sqd1* mutant is less dynamic than it is in wild type oocytes, and also appears to associate with P bodies.

Third, injecting *grk* near the DA of *sqd1* oocytes results in restriction of the RNA to the DA, reducing the anterior mislocalisation phenotype of the mutant. It is possible that a low level of *grk* anchoring at the DA is retained in the *sqd* mutant, or that the RNA is maintained at the DA by continual transport along the dense network of microtubules around the nucleus. Therefore, further characterisation of the dynamics of this DA enriched RNA would be required to determine whether it is truly anchored.

Finally, I show that Sqd is slightly enriched at the anterior of the oocyte, but not at the DA where it is expected to play a role in *grk* anchoring. This anterior enrichment is increased by recruitment of Sqd following the injection of highly concentrated *grk* RNA. Although I was unable to show Sqd recruitment to anchored *grk* at the DA, others have demonstrated this.¹³³ In addition, I was unable to demonstrate Sqd recruitment to injected *grk* in the embryo (data

not shown). It is possible that the accumulation of Sqd::GFP is too weak to see using widefield microscopy, and could be improved using confocal microscopy. Alternatively, my assessment of Sqd recruitment in the oocyte at 40 minutes may be premature, as the previous studies were carried out at 80 minutes post-injection of *grk*.¹³³

Taken together, these observations suggest that *grk* is prematurely anchored at P bodies at the anterior of a *sqd* mutant, rather than being un-anchored, and may provide insights into how *grk* is ectopically translated at the anterior of the mutant. I propose a revision to the current interpretation of the role of Sqd in promoting *grk* anchoring, where Sqd instead mediates anchoring and translation by preventing inappropriate association with P bodies away from the DA, facilitating anchoring specifically at the DA. It is possible that Sqd is exchanged with other factors at the DA, such as the Dynein anchor and the translational activator Orb, or another uncharacterised protein (Figure 7.1B, 7.1E). Both interpretations would result in the same outcome, with *grk* being restricted to the DA of the oocyte, but would generate subtle differences in *grk* mislocalisation phenotypes in a *sqd* mutant. Alternatively, *grk* localisation along the anterior of a *sqd1* mutant could be the result of inappropriate translation rather than inappropriate anchoring, possibly due to dysregulated translational inhibition.

Differences in the observations I have made here and those in previous publications could be a result of variations in the mutant background. It is possible that the *sqd* allele has picked up modifiers over time, or that the *sqd1/sqd1* line used in many of the previous publications has an additional mutation that is rescued with *Df(3R)urd* used in this study. Furthermore, despite showing that the *sqd1/Df* flies used in this study have the *sqd* dorsalised egg phenotype, it is possible this is a result of aberrant translation rather than an mRNA localisation defect as I find a less severe *grk* mislocalisation phenotype than in some previous publications. Therefore, differences between these fly lines should be tested for subtle variations in the *sqd1* phenotype, and the dorsalised egg phenotype should be quantified in my flies. Moreover, in order to truly determine whether *grk* forms a DA cap in a *sqd1* mutant, the

distribution of endogenous *grk* should be tested using single-molecule FISH, and the dynamics of *grk* RNA in wild type and mutant oocytes should be further addressed using particle tracking, FCS or *grk*-MS2 labelled with photoswitchable dyes.

There are several further questions that should be addressed. Why is Sqd enriched at the edge of P bodies at the DA of the oocyte? Are alternative *trans*-acting factors required for *grk* anchoring at the DA? Future work would aim to address these questions by testing the role of other *trans*-acting factors whose mutants have been shown to have *grk* mislocalisation at the anterior of the oocyte, such as K10, Cup, Hrb27C, Otu, Spn-F, Enc and Glo.^{116, 162, 314, 315, 319, 409} It has previously been demonstrated that mislocalised *grk* at the anterior of *enc* and *spn-F* mutants is not translated,^{320, 409} suggesting that the mechanisms to inhibit *grk* translation at the anterior are still intact. Investigating whether *grk* is anchored at the anterior of these mutants would provide insights into whether a lack of Grk expression is a result of an inability to anchor, and whether this is due to the presence of Sqd. Finally, ultrastructural analysis of the P body components at different regions of wild type and *sqd1* mutant oocytes could shed light on why *grk* is translationally activated specifically at the DA.

Translational control of *grk* in P bodies

As the localisation and translational regulation of *grk* are thought to be closely linked, it is interesting to investigate the mechanisms behind *grk* localisation in the embryo where it is not expressed. I show that ectopic injection of *grk* RNA into the embryo results in the accumulation of Me31B specifically at the apical cortex where it becomes closely associated with localised *grk*. An interesting interpretation of these results would be that ectopic *grk* causes the formation of P bodies in the embryo which could inhibit *grk* translation by targeting the transcript to the translationally inactive core. However, EM studies and immunostaining of flies expressing fluorescently tagged components of P bodies, such as Tral, would be required to determine whether these Me31B puncta are actually P bodies. Staining the embryos with an antibody against Grk could determine whether *grk* is translated at the apical

cortex of the embryo. However, if Grk is not expressed, distinguishing between whether lack of expression is due to *grk* association with P bodies, or whether *grk* is not competent to be ectopically translated in the embryo, would be a challenge. Although I have shown that the non-localising control *hb* does not recruit Me31B in the embryo, it would be interesting to see whether other RNAs that localise to the apical cortex and are expressed in the embryo also show Me31B accumulation.

Improving our understanding of how specificity is generated during RNA transport, anchoring and translation

Establishing specificity of mRNA localisation, anchoring and translation is difficult as many *trans*-acting factors function together on several mRNA cargoes destined for different destinations within the cell. For example, *sqd* has been shown to bind *grk*, *bcd* and *osk*; three mRNAs that localise to distinct regions of the *Drosophila* oocyte. My data, suggesting an alternative role of Sqd in inhibiting inappropriate mRNA anchoring at the anterior of the oocyte, could explain why Sqd also associates with *bcd* (localised to the anterior of the oocyte but is not anchored or translated until much later in oogenesis) and *osk* (transiently localised to the anterior of the oocyte but is not anchored or translated there).

Sqd is highly conserved and is most closely related to the hnRNPA family of proteins in humans. In both *Drosophila* and humans, Sqd/hnRNPA contains two RRM-domains and shuttle in and out of the nucleus using an M9 domain, allowing the protein to package pre-mRNAs into hnRNP particles and transport mRNAs from the nucleus into the cytoplasm. hnRNPA1 is also involved in hepatitis C virus (HCV) replication through binding to the 5'- and 3'-UTR of the viral RNA.⁴¹² It would therefore be interesting to investigate the degree of conservation between the RRMs in Sqd and hnRNPA1, and whether the putative Sqd binding sites are conserved within the HCV transcript UTRs.

mRNAs that utilise the same transport machinery do not appear to share common sequences within their localisation signals, making it difficult to determine how specificity is generated. My findings that Egl may not recognise A'-form RNA conformations previously shown to be required in the case of *fs(1)K10* indicate that the same adaptor proteins may recognise distinct motifs within different mRNA cargoes, adding further complexity to the system. Egl is conserved in *C. elegans* and functions together with the homologs of BicD and Dynein during nuclear migration in different tissues in both *Drosophila* and *C. elegans*. It is therefore interesting to investigate whether this transport complex also plays a role in mRNA localisation in *C. elegans*.

Specifically targeting mRNAs to distinct regions of the *Drosophila* oocyte is further complicated by the convoluted array of MTs observed during mid-oogenesis, making it difficult for the motor proteins to navigate the poorly polarised and unorganised cytoskeleton when compared to tissues with a highly polarised MT network, such as the *Drosophila* embryo. Such differences in the complexities of the MT network are likely to be the cause of the reduced transport efficiency I observe for some of the GLS mutants in the oocyte compared to in the embryo, and make the embryo injection assay important for reducing the complexity of the system when assessing the transport of mutant RNAs.

As many *trans*-acting factors act together in several steps of mRNA processing it is difficult to test the role of each protein and its specific interaction with cargo mRNAs. The most high-throughput way of assessing the association between transport and anchoring factors and *grk* is through injection of the RNA into the oocyte or embryo. However, these assays have their limitations. For example, injecting the RNA into the oocyte denies the RNA access to proteins in the NCs that may be important. In addition, *grk* mRNA may not travel from the centre of the oocyte as it is unclear if the same pool of mRNA moves from the posterior to the DA, or whether new mRNA arrives from the NCs. Moreover, the concentration of injected RNA is likely to be substantially higher than endogenous levels. Therefore, it may be more biologically relevant to test the localisation of RNAs by injecting them into the NCs as

opposed to the oocyte, or by endogenously expressing MS2-MCP labelled versions of the RNA. Finally, assessing the localisation pattern of the mRNA quantitatively can be challenging as the egg chamber is a large and thick three-dimensional structure.

We are currently investigating whether it is possible to quantitate the pattern of RNA localisation by generating ratios of the fluorescent intensities in each region of interest, such as the anterior and DA. This is achieved by choosing the z-section with the highest intensity for each region and measuring the mean fluorescent intensity within the region of interest in that section. However, due to the thickness of the sample, imaging and quantification of fluorescent intensities in the ring around the anterior of the oocyte is challenging when choosing different sections of the sample. It may therefore be necessary to assess localisation patterns using confocal microscopy where problems with out of focus light are minimised, although this would drastically reduce the throughput of the assay.

Final conclusions

The fine regulation of transport, anchoring and translation of many localised mRNAs is achieved through interactions between a combination of *trans*-acting factors and *cis*-acting signals in the transcript. In this thesis I aimed to dissect the molecular interactions between the GLS and a selection of *trans*-acting partners involved in *grk* localisation and translation. Identifying universal motifs within localisation signals that are responsible for the binding of *trans*-acting factors is limited by the fact that transport, anchoring and translation are interlinked, and proteins are likely to act together in these processes to provide specificity. Despite this, I have suggested detailed molecular mechanisms for critical steps of *grk* targeting by the GLS that should now be assessed for other transcripts. I also suggest a possible revision to the current interpretation for the role of Sqd in *grk* anchoring and translation at the anterior of the oocyte which would be interesting to pursue further. Dissecting the role of individual proteins in controlling mRNA localisation and translation would help us to understand how these processes are regulated in health and disease.

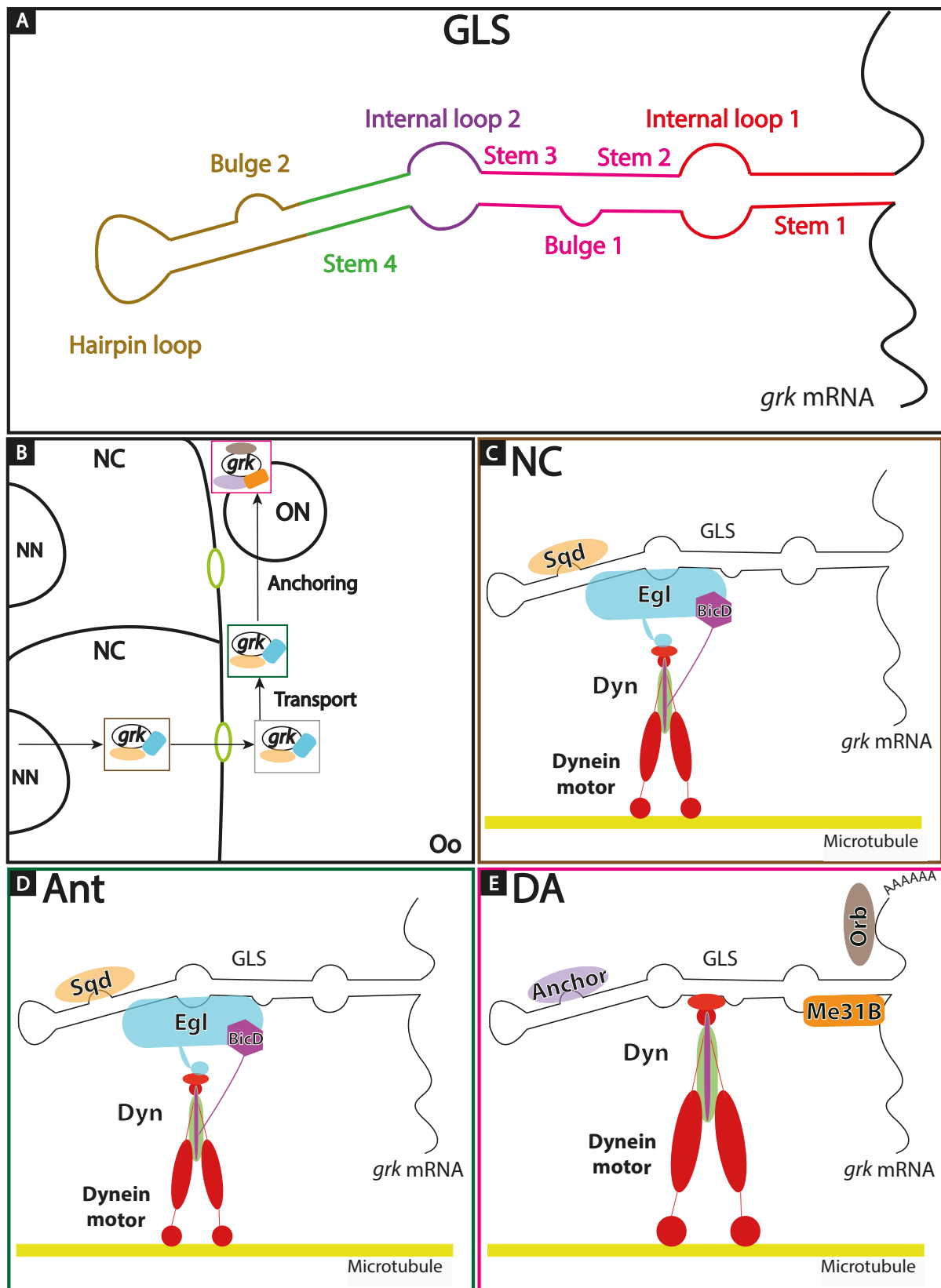


Figure 7.1. Model for the role of the GLS in *grk* localisation from the NCs to the DA of the oocyte. (A) Schematic of the predicted secondary structure of the GLS within the coding region of *grk* mRNA. The GLS is a stem-loop composed of two internal loops (red and purple), two bulges (pink and brown), four helical stem regions (red, pink and green) and a hairpin loop (brown). (B) The *Drosophila* egg chamber is composed of supporting nurse cells (NC) connected to a large oocyte (Oo) via ring canals (light green circles). The maternal mRNA *grk* is transcribed in the NC nuclei (NN) and undergoes remodelling throughout its journey to the dorsal-anterior (DA, pink box) of the oocyte. *grk* associates with Sqd (light orange) and Egl (blue) throughout its transport from the NC cytoplasm (brown box), into the oocyte cytoplasm (grey box), and to the anterior of the oocyte (dark green box). *grk* is then restricted to a region above the oocyte nucleus (ON) at the DA (pink box) via an anchoring mechanism that docks the RNA to the edge of P bodies. Anchoring requires remodelling of the complex. The transport adaptor protein Egl dissociates, Me31B associates with *grk*, and an unknown anchoring factor (light purple) replaces Sqd. (C) *grk* is transported into the oocyte through the ring canals along the MT network (yellow) using the Dynein motor (red), Dynamin (green), Egl (blue), BicD (purple), and the *grk* localisation signal (GLS). Egl interacts with an AU-rich stem 4 (green in A) whose sequence is not essential for transport, and a purine stack in stem 3 (pink in A) whose sequence is absolutely required for transport. The size and orientation of internal loop 2 (purple in A) is essential to produce a kink within the RNA of an appropriate size for the formation of a processive motor complex. It is currently unknown whether one or more motor complexes associate with the GLS in this manner to promote efficient *grk* transport. (D) *grk* is first transported to the anterior (Ant) of the oocyte (dark green box in B) where Sqd prevents inappropriate anchoring of the RNA, resulting in dynamic *grk* particles. Sqd is likely to associate with sequences around bulge 2 (brown in A), with the presence of the bulge being required for providing accessibility to such sequences in the surrounding helical regions. (E) *grk* becomes restricted to the dorsal-anterior (DA) of the oocyte (pink box in A), forming a cap above the ON, and becomes anchored at the edge of Me31B-rich P bodies. This association with P bodies is likely to be regulated by regions within stem 1 and internal loop 1 (red in A), and may be achieved by Me31B association with this region of the signal. Restriction of *grk* to the DA is likely to require relief from the inhibitory anchoring function of Sqd, possibly by replacement of Sqd with another anchoring factor (light purple), and requires Dynein, but not Egl or BicD. Ribosomes and the translational activator Orb are enriched at the edge of the DA P bodies. Hyperphosphorylated Orb (light brown) associates with *grk* and recruits the polyadenylase Wispy (not shown), resulting in extension of the *grk* poly(A) tail and stimulating localised Grk expression.

Chapter 8: References

1. Jambor, H. *et al.* Systematic imaging reveals features and changing localization of mRNAs in *Drosophila* development. *eLife Sciences* **4**, e05003 (2015).
2. Lécuyer, E. *et al.* Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174–187 (2007).
3. Jansen, R. P. mRNA localization: message on the move. *Nat Rev Mol Cell Biol* **2**, 247–256 (2001).
4. Blower, M. D. Molecular insights into intracellular RNA localization. *Int Rev Cell Mol Biol* **302**, 1–39 (2013).
5. Chartrand, P., Meng, X. H., Hüttelmaier, S., Donato, D. & Singer, R. H. Asymmetric Sorting of Ash1p in Yeast Results from Inhibition of Translation by Localization Elements in the mRNA. *Mol Cell* **10**, 1319–1330 (2002).
6. Long, R. M. *et al.* Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* **277**, 383–387 (1997).
7. Takizawa, P. A., Sil, A., Swedlow, J. R., Herskowitz, I. & Vale, R. D. Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* **389**, 90–93 (1997).
8. Medioni, C., Mowry, K. & Besse, F. Principles and roles of mRNA localization in animal development. **139**, 3263–3276 (2012).

9. Mingle, L. A. *et al.* 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 (2005).
10. Batada, N. N., Shepp, L. A. & Siegmund, D. O. Stochastic model of protein-protein interaction: why signaling proteins need to be colocalized. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6445–6449 (2004).
11. Broadus, J., Fuerstenberg, S. & Doe, C. Q. Staufer-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate. *Nature* **391**, 792–795 (1998).
12. Hughes, J. R., Bullock, S. L. & Ish-Horowicz, D. Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in *Drosophila* neuroblasts. **14**, 1950–1956 (2004).
13. Li, P., Yang, X., Wasser, M., Cai, Y. & Chia, W. Inscuteable and Staufer mediate asymmetric localization and segregation of prospero RNA during *Drosophila* neuroblast cell divisions. *Cell* **90**, 437–447 (1997).
14. Melton, D. A. Translocation of a localized maternal mRNA to the vegetal pole of *Xenopus* oocytes. *Nature* **328**, 80–82 (1987).
15. Neuman-Silberberg, F. S. & Schüpbach, T. The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF α -like protein. *Cell* **75**, 165–174 (1993).
16. Simmonds, A. J., dosSantos, G., Livne-Bar, I. & Krause, H. M. Apical localization of wingless transcripts is required for wingless signaling. *Cell* **105**, 197–207 (2001).

17. Zhang, J. *et al.* The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell* **94**, 515–524 (1998).
18. Liao, G., Mingle, L., Van De Water, L. & Liu, G. Control of cell migration through mRNA localization and local translation. *Wiley Interdiscip Rev RNA* **6**, 1–15 (2015).
19. Driever, W. & Nüsslein-Volhard, C. The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* **54**, 95–104 (1988).
20. Ephrussi, A., Dickinson, L. K. & Lehmann, R. Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* **66**, 37–50 (1991).
21. Gavis, E. R. & Lehmann, R. Localization of nanos RNA controls embryonic polarity. *Cell* **71**, 301–313 (1992).
22. Dahm, R., Kiebler, M. & Macchi, P. RNA localisation in the nervous system. *Semin Cell Dev Biol* **18**, 216–223 (2007).
23. Lawrence, J. B. & Singer, R. H. Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell* **45**, 407–415 (1986).
24. Liu-Yesucevitz, L. *et al.* Local RNA translation at the synapse and in disease. *J Neurosci* **31**, 16086–16093 (2011).
25. Paquin, N. & Chartrand, P. Local regulation of mRNA translation: new insights from the bud. *Trends Cell Biol* **18**, 105–111 (2008).

26. Cody, N. A. L., Iampietro, C. & Lécuyer, E. The many functions of mRNA localization during normal development and disease: from pillar to post. *WIREs Dev Biol* n/a–n/a (2013). doi:10.1002/wdev.113
27. Zhang, J. & King, M. L. *Xenopus* VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. **122**, 4119–4129 (1996).
28. Birsoy, B., Kofron, M., Schaible, K., Wylie, C. & Heasman, J. Vg 1 is an essential signaling molecule in *Xenopus* development. **133**, 15–20 (2006).
29. Weil, T. T. mRNA localization in the *Drosophila* germline. *RNA Biol* **11**, 1010–1018 (2014).
30. Huang, L. & Muthuswamy, S. K. Polarity protein alterations in carcinoma: a focus on emerging roles for polarity regulators. *Curr Opin Genet Dev* **20**, 41–50 (2010).
31. Martin, T. A., Watkins, G., Mansel, R. E. & Jiang, W. G. Loss of tight junction plaque molecules in breast cancer tissues is associated with a poor prognosis in patients with breast cancer. *European Journal of Cancer* **40**, 2717–2725 (2004).
32. Vikesaa, J. *et al.* RNA-binding IMPs promote cell adhesion and invadopodia formation. *EMBO J* **25**, 1456–1468 (2006).
33. Dictenberg, J. B., Swanger, S. A., Antar, L. N., Singer, R. H. & Bassell, G. J. A Direct Role for FMRP in Activity-Dependent Dendritic mRNA Transport Links Filopodial-Spine Morphogenesis to Fragile X Syndrome. *Dev Cell* **14**, 926–939 (2008).

34. Kao, D.-I., Aldridge, G. M., Weiler, I. J. & Greenough, W. T. Altered mRNA transport, docking, and protein translation in neurons lacking fragile X mental retardation protein. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15601–15606 (2010).
35. Bassell, G. J. & Warren, S. T. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* **60**, 201–214 (2008).
36. Ham, B. K. *et al.* A Polypyrimidine Tract Binding Protein, Pumpkin RBP50, Forms the Basis of a Phloem-Mobile Ribonucleoprotein Complex. *THE PLANT CELL ONLINE* **21**, 197–215 (2009).
37. Uniacke, J. & Zerges, W. Chloroplast protein targeting involves localized translation in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1439–1444 (2009).
38. Lough, T. J. & Lucas, W. J. Integrative plant biology: role of phloem long-distance macromolecular trafficking. *Annu Rev Plant Biol* **57**, 203–232 (2006).
39. Lehmann, M. *et al.* Intracellular transport of human immunodeficiency virus type 1 genomic RNA and viral production are dependent on dynein motor function and late endosome positioning. *J Biol Chem* **284**, 14572–14585 (2009).
40. Mouland, A. J. *et al.* RNA trafficking signals in human immunodeficiency virus type 1. *Mol Cell Biol* **21**, 2133–2143 (2001).
41. Braunschweig, U. *et al.* Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* **24**, 1774–1786 (2014).
42. Baurén, G. & Wieslander, L. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* **76**, 183–192 (1994).

43. Shandilya, J. & Roberts, S. G. E. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochim Biophys Acta* **1819**, 391–400 (2012).
44. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J* **19**, 6860–6869 (2000).
45. Natalizio, B. J. & Wente, S. R. Postage for the messenger: designating routes for nuclear mRNA export. *Trends Cell Biol* **23**, 365–373 (2013).
46. Saulière, J. *et al.* CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat Struct Mol Biol* **19**, 1124–1131 (2012).
47. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 750–764 (2012).
48. Montpetit, B. *et al.* A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **472**, 238–242 (2011).
49. Powrie, E. A., Zenklusen, D. & Singer, R. H. A nucleoporin, Nup60p, affects the nuclear and cytoplasmic localization of ASH1 mRNA in *S. cerevisiae*. *RNA* **17**, 134–144 (2011).
50. Rodríguez-Navarro, S. & Hurt, E. Linking gene regulation to mRNA production and export. *Curr Opin Cell Biol* **23**, 302–309 (2011).

51. Valkov, E., Dean, J. C., Jani, D., Kuhlmann, S. I. & Stewart, M. Structural basis for the assembly and disassembly of mRNA nuclear export complexes. *Biochim Biophys Acta* **1819**, 578–592 (2012).
52. Parton, R. M., Davidson, A., Davis, I. & Weil, T. T. Subcellular mRNA localisation at a glance. **127**, 2127–2133 (2014).
53. Amrute-Nayak, M. & Bullock, S. L. Single-molecule assays reveal that RNA localization signals regulate dynein-dynactin copy number on individual transcript cargoes. *Nat Cell Biol* **14**, 416–423 (2012).
54. McDermott, S. M., Meignin, C., Rappsilber, J. & Davis, I. Drosophila Syncrip binds the gurken mRNA localisation signal and regulates localised transcripts during axis specification. *Biology Open* **1**, 488–497 (2012).
55. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**, 1393–1406 (2012).
56. Castello, A. *et al.* System-wide identification of RNA-binding proteins by interactome capture. *Nat Protoc* **8**, 491–500 (2013).
57. Ghosh, S. *et al.* Acetylation of the SUN protein Mps3 by Eco1 regulates its function in nuclear organization. *Mol Biol Cell* **23**, 2546–2559 (2012).
58. St Johnston, D., Driever, W., Berleth, T., Richstein, S. & Nüsslein-Volhard, C. Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. **107 Suppl**, 13–19 (1989).

59. St Johnston, D., Beuchle, D. & Nüsslein-Volhard, C. Staufén, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51–63 (1991).
60. Horne-Badovinac, S. & Bilder, D. Dynein regulates epithelial polarity and the apical localization of stardust A mRNA. *PLoS Genet.* **4**, e8 (2008).
61. Trcek, T. & Singer, R. H. The cytoplasmic fate of an mRNP is determined cotranscriptionally: exception or rule? *Genes & Development* **24**, 1827–1831 (2010).
62. Munro, T. P. *et al.* Mutational analysis of a heterogeneous nuclear ribonucleoprotein A2 response element for RNA trafficking. *J Biol Chem* **274**, 34389–34395 (1999).
63. Macdonald, P. M. & Struhl, G. Cis- acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos. *Nature* **336**, 595–598 (1988).
64. Macdonald, P. M. & Kerr, K. Mutational analysis of an RNA recognition element that mediates localization of bicoid mRNA. *Mol Cell Biol* **18**, 3788–3795 (1998).
65. Macdonald, P. M. & Kerr, K. Redundant RNA recognition events in bicoid mRNA localization. *RNA* **3**, 1413–1420 (1997).
66. Macdonald, P. M., Kerr, K., Smith, J. L. & Leask, A. RNA regulatory element BLE1 directs the early steps of bicoid mRNA localization. **118**, 1233–1243 (1993).
67. Lewis, R. A. *et al.* Conserved and clustered RNA recognition sequences are a critical feature of signals directing RNA localization in *Xenopus* oocytes. *Mech Dev* **121**, 101–109 (2004).

68. Chartrand, P., Meng, X. H., Singer, R. H. & Long, R. M. Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo. *Current Biology* **9**, 333–338 (1999).
69. Hartswood, E., Brodie, J., Vendra, G., Davis, I. & Finnegan, D. J. RNA:RNA interaction can enhance RNA localization in *Drosophila* oocytes. *RNA* **18**, 729–737 (2012).
70. Besse, F., López de Quinto, S., Marchand, V., Trucco, A. & Ephrussi, A. *Drosophila* PTB promotes formation of high-order RNP particles and represses oskar translation. *Genes & Development* **23**, 195–207 (2009).
71. Chekulaeva, M., Hentze, M. W. & Ephrussi, A. Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. *Cell* **124**, 521–533 (2006).
72. Bashirullah, A. *et al.* Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in *Drosophila melanogaster*. *EMBO J* **18**, 2610–2620 (1999).
73. Forrest, K. M. & Gavis, E. R. Live Imaging of Endogenous RNA Reveals a Diffusion and Entrapment Mechanism for nanos mRNA Localization in *Drosophila*. *Current Biology* **13**, 1159–1168 (2003).
74. Dahanukar, A. & Wharton, R. P. The Nanos gradient in *Drosophila* embryos is generated by translational regulation. *Genes & Development* **10**, 2610–2621 (1996).
75. Gavis, E. R., Lunsford, L., Bergsten, S. E. & Lehmann, R. A conserved 90 nucleotide element mediates translational repression of nanos RNA. (1996).

76. Smibert, C. A., Wilson, J. E., Kerr, K. & Macdonald, P. M. smaug protein represses translation of unlocalized nanos mRNA in the *Drosophila* embryo. *Genes & Development* **10**, 2600–2609 (1996).
77. Reck-Peterson, S. L., Provance, D. W., Jr., Mooseker, M. S. & Mercer, J. A. Class V myosins. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1496**, 36–51 (2000).
78. Vale, R. D. Myosin V motor proteins. *J Cell Biol* **163**, 445–450 (2003).
79. Caviston, J. P. & Holzbaur, E. L. F. Microtubule motors at the intersection of trafficking and transport. *Trends Cell Biol* **16**, 530–537 (2006).
80. Tekotte, H. & Davis, I. Intracellular mRNA localization: motors move messages. *Trends Genet* **18**, 636–642 (2002).
81. Vale, R. D., Reese, T. S. & Sheetz, M. P. Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell* **42**, 39–50 (1985).
82. Kolomeisky, A. B. Motor proteins and molecular motors: how to operate machines at the nanoscale. *J. Phys.: Condens. Matter* **25**, 463101 (2013).
83. Gibbons, I. R. STUDIES ON THE PROTEIN COMPONENTS OF CILIA FROM TETRAHYMENA PYRIFORMIS. *PNAS* **50**, 1002–1010 (1963).
84. Paschal, B. M. & Vallee, R. B. Retrograde transport by the microtubule-associated protein MAP 1C. *Nature* **330**, 181–183 (1987).

85. Paschal, B. M., Shpetner, H. S. & Vallee, R. B. MAP 1C is a microtubule-activated ATPase which translocates microtubules in vitro and has dynein-like properties. *J Cell Biol* **105**, 1273–1282 (1987).
86. Schroer, T. A., Steuer, E. R. & Sheetz, M. P. Cytoplasmic dynein is a minus end-directed motor for membranous organelles. *Cell* **56**, 937–946 (1989).
87. Vallee, R. B., Shpetner, H. S. & Paschal, B. M. The role of dynein in retrograde axonal transport. *Trends Neurosci* **12**, 66–70 (1989).
88. Nishiura, M. *et al.* A single-headed recombinant fragment of Dictyostelium cytoplasmic dynein can drive the robust sliding of microtubules. *J Biol Chem* **279**, 22799–22802 (2004).
89. Reck-Peterson, S. L. *et al.* Single-molecule analysis of dynein processivity and stepping behavior. *Cell* **126**, 335–348 (2006).
90. Pfister, K. K. *et al.* Genetic analysis of the cytoplasmic dynein subunit families. *PLoS Genet.* **2**, 11–26 (2006).
91. King, S. M. The dynein microtubule motor. *Biochim Biophys Acta* **1496**, 60–75 (2000).
92. Burgess, S. A., Walker, M. L., Sakakibara, H., Oiwa, K. & Knight, P. J. The structure of dynein-c by negative stain electron microscopy. *J. Struct. Biol.* **146**, 205–216 (2004).
93. Gennerich, A. & Vale, R. D. Walking the walk: how kinesin and dynein coordinate their steps. *Curr Opin Cell Biol* **21**, 59–67 (2009).

94. Sakato, M. & King, S. M. Design and regulation of the AAA+ microtubule motor dynein. *J. Struct. Biol.* **146**, 58–71 (2004).
95. Samsó, M., Radermacher, M., Frank, J. & Koonce, M. P. Structural characterization of a dynein motor domain. *J. Mol. Biol.* **276**, 927–937 (1998).
96. Chowdhury, S., Ketcham, S. A., Schroer, T. A. & Lander, G. C. Structural organization of the dynein-dynactin complex bound to microtubules. *Nat Struct Mol Biol* **22**, 345–347 (2015).
97. King, S. J. & Schroer, T. A. Dynactin increases the processivity of the cytoplasmic dynein motor. *Nat Cell Biol* **2**, 20–24 (2000).
98. Schroer, T. A. Dynactin. *Annu Rev Cell Dev Biol* **20**, 759–779 (2004).
99. Waterman-Storer, C. M., Karki, S. & Holzbaur, E. L. The p150Glued component of the dynactin complex binds to both microtubules and the actin-related protein centractin (Arp-1). *PNAS* **92**, 1634–1638 (1995).
100. Holleran, E. A., Karki, S. & Holzbaur, E. L. The role of the dynactin complex in intracellular motility. *Int. Rev. Cytol.* **182**, 69–109 (1998).
101. King, S. M. AAA domains and organization of the dynein motor unit. *J. Cell. Sci.* **113**, 2521–2526 (2000).
102. Eckley, D. M. *et al.* Analysis of dynactin subcomplexes reveals a novel actin-related protein associated with the arp1 minifilament pointed end. *J Cell Biol* **147**, 307–320 (1999).

103. Bananis, E., Murray, J. W., Stockert, R. J., Satir, P. & Wolkoff, A. W. Regulation of early endocytic vesicle motility and fission in a reconstituted system. *J. Cell. Sci.* **116**, 2749–2761 (2003).
104. Nath, S. *et al.* Kif5B and Kifc1 interact and are required for motility and fission of early endocytic vesicles in mouse liver. *Mol Biol Cell* **18**, 1839–1849 (2007).
105. Ling, S.-C., Fahrner, P. S., Greenough, W. T. & Gelfand, V. I. Transport of Drosophila fragile X mental retardation protein-containing ribonucleoprotein granules by kinesin-1 and cytoplasmic dynein. *PNAS* **101**, 17428–17433 (2004).
106. Gross, S. P. Hither and yon: a review of bi-directional microtubule-based transport. *Physical biology* **1**, R1–11 (2004).
107. Ross, J. L., Wallace, K., Shuman, H., Goldman, Y. E. & Holzbaur, E. L. F. Processive bidirectional motion of dynein-dynactin complexes in vitro. *Nat Cell Biol* **8**, 562–570 (2006).
108. Bullock, S. L., Nicol, A., Gross, S. P. & Zicha, D. Guidance of bidirectional motor complexes by mRNA cargoes through control of dynein number and activity (Supplemental). **16**, 1447–1452 (2006).
109. Clark, I., Giniger, E., Ruohola-Baker, H., Jan, L. Y. & Jan, Y. N. Transient posterior localization of a kinesin fusion protein reflects anteroposterior polarity of the Drosophila oocyte. *Current Biology* **4**, 289–300 (1994).
110. Clark, I. E., Jan, L. Y. & Jan, Y. N. Reciprocal localization of Nod and kinesin fusion proteins indicates microtubule polarity in the Drosophila oocyte, epithelium, neuron and muscle. **124**, 461–470 (1997).

111. Theurkauf, W. E., Smiley, S., Wong, M. L. & Alberts, B. M. Reorganization of the cytoskeleton during *Drosophila* oogenesis: implications for axis specification and intercellular transport. **115**, 923–936 (1992).
112. Li, R. & Gundersen, G. G. Beyond polymer polarity: how the cytoskeleton builds a polarized cell. *Nat Rev Mol Cell Biol* **9**, 860–873 (2008).
113. Bartolini, F. & Gundersen, G. G. Formins and microtubules. *Biochim Biophys Acta* **1803**, 164–173 (2010).
114. Zimyanin, V. L. *et al.* In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell* **134**, 843–853 (2008).
115. Parton, R. M. *et al.* A PAR-1-dependent orientation gradient of dynamic microtubules directs posterior cargo transport in the *Drosophila* oocyte. **194**, 121–135 (2011).
116. Bullock, S. L. & Ish-Horowicz, D. Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* **414**, 611–616 (2001).
117. Liu, Y. *et al.* Bicaudal-D uses a parallel, homodimeric coiled coil with heterotypic registry to coordinate recruitment of cargos to dynein. *Genes & Development* **27**, 1233–1246 (2013).
118. Mach, J. M. & Lehmann, R. An Egalitarian-BicaudalD complex is essential for oocyte specification and axis determination in *Drosophila*. *Genes & Development* **11**, 423–435 (1997).

119. Navarro, C., Puthalakath, H., Adams, J. M., Strasser, A. & Lehmann, R. Egalitarian binds dynein light chain to establish oocyte polarity and maintain oocyte fate. *Nat Cell Biol* **6**, 427–435 (2004).
120. MacDougall, N., Clark, A., MacDougall, E. & Davis, I. *Drosophila* gurken (TGF α) mRNA localizes as particles that move within the oocyte in two dynein-dependent steps. *Dev Cell* **4**, 307–319 (2003).
121. Weil, T. T., Forrest, K. M. & Gavis, E. R. Localization of bicoid mRNA in late oocytes is maintained by continual active transport. *Dev Cell* **11**, 251–262 (2006).
122. Carson, J. H., Worboys, K., Ainger, K. & Barbarese, E. Translocation of myelin basic protein mRNA in oligodendrocytes requires microtubules and kinesin. *Cytoskeleton* **38**, 318–328 (1997).
123. Kanai, Y., Dohmae, N. & Hirokawa, N. Kinesin transports RNA: Isolation and characterization of an RNA-transporting granule. *Neuron* **43**, 513–525 (2004).
124. Katz, Z. B. *et al.* β -Actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development* **26**, 1885–1890 (2012).
125. Sundell, C. L. & Singer, R. H. Requirement of microfilaments in sorting of actin messenger RNA. *Science* (1991).
126. Barros, C. S., Phelps, C. B. & Brand, A. H. *Drosophila* Nonmuscle Myosin II Promotes the Asymmetric Segregation of Cell Fate Determinants by Cortical Exclusion Rather Than Active Transport. *Dev Cell* **5**, 829–840 (2003).

127. Schuh, M. An actin-dependent mechanism for long-range vesicle transport. *Nat Cell Biol* **13**, 1431–1436 (2011).
128. Dahlgaard, K., Raposo, A. A. S. F., Niccoli, T. & St Johnston, D. Capu and Spire assemble a cytoplasmic actin mesh that maintains microtubule organization in the *Drosophila* oocyte. *Dev Cell* **13**, 539–553 (2007).
129. Bullock, S. L., Zicha, D. & Ish-Horowicz, D. The *Drosophila* hairy RNA localization signal modulates the kinetics of cytoplasmic mRNA transport. *EMBO J* **22**, 2484–2494 (2003).
130. Soundararajan, H. C. & Bullock, S. L. The influence of dynein processivity control, MAPs, and microtubule ends on directional movement of a localising mRNA. *eLife Sciences* **3**, e01596 (2014).
131. Gross, S. P., Welte, M. A., Block, S. M. & Wieschaus, E. F. Dynein-mediated cargo transport in vivo. A switch controls travel distance. **148**, 945–956 (2000).
132. Fusco, D. *et al.* Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. **13**, 161–167 (2003).
133. Delanoue, R., Herpers, B., Soetaert, J., Davis, I. & Rabouille, C. *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 (2007).
134. Loiseau, P., Davies, T., Williams, L. S., Mishima, M. & Palacios, I. M. *Drosophila* PAT1 is required for Kinesin-1 to transport cargo and to maximize its motility. **137**, 2763–2772 (2010).

135. Sung, H.-H. *et al.* *Drosophila* ensconsin promotes productive recruitment of Kinesin-1 to microtubules. *Dev Cell* **15**, 866–876 (2008).
136. Kloc, M. & Etkin, L. D. RNA localization mechanisms in oocytes. **118**, 269–282 (2005).
137. Yisraeli, J. K., Sokol, S. & Melton, D. A. A two-step model for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of Vg1 mRNA. **108**, 289–298 (1990).
138. Babu, K., Cai, Y., Bahri, S., Yang, X. & Chia, W. Roles of Bifocal, Homer, and F-actin in anchoring Oskar to the posterior cortex of *Drosophila* oocytes. **18**, 138–143 (2004).
139. Baum, B. *Drosophila* oogenesis: generating an axis of polarity. **12**, R835–7 (2002).
140. Cha, B.-J., Serbus, L. R., Koppetsch, B. S. & Theurkauf, W. E. Kinesin I-dependent cortical exclusion restricts pole plasm to the oocyte posterior. *Nat Cell Biol* **4**, 592–598 (2002).
141. Delanoue, R. & Davis, I. Dynein anchors its mRNA cargo after apical transport in the *Drosophila* blastoderm embryo. *Cell* **122**, 97–106 (2005).
142. Bullock, S. L. *et al.* Differential cytoplasmic mRNA localisation adjusts pair-rule transcription factor activity to cytoarchitecture in dipteran evolution. **131**, 4251–4261 (2004).
143. Gonzalez, I., Buonomo, S. B. C., Nasmyth, K. & Ahsen, von, U. ASH1 mRNA localization in yeast involves multiple secondary structural elements and Ash1 protein translation. *Current Biology* **9**, 337–340 (1999).

144. Vanzo, N. F. & Ephrussi, A. Oskar anchoring restricts pole plasm formation to the posterior of the *Drosophila* oocyte. **129**, 3705–3714 (2002).
145. Vanzo, N., Oprins, A., Xanthakis, D., Ephrussi, A. & Rabouille, C. Stimulation of Endocytosis and Actin Dynamics by Oskar Polarizes the *Drosophila* Oocyte. *Dev Cell* **12**, 543–555 (2007).
146. Heasman, J., Wessely, O., Langland, R., Craig, E. J. & Kessler, D. S. Vegetal Localization of Maternal mRNAs Is Disrupted by VegT Depletion. *Dev Biol* **240**, 377–386 (2001).
147. Etkin, L. D. & Kloc, M. Delocalization of Vg1 mRNA from the vegetal cortex in *Xenopus* oocytes after destruction of Xlsirt. *Science* **265**, 1101–1103 (1994).
148. Kloc, M. *et al.* Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of *Xenopus* oocytes. **132**, 3445–3457 (2005).
149. Shi, Z. & Barna, M. Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. *Annu Rev Cell Dev Biol* **31**, 31–54 (2015).
150. Richter, J. D. & Sonenberg, N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433**, 477–480 (2005).
151. Jackson, R. J., Hellen, C. U. T. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**, 113–127 (2010).

152. Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* **6**, 318–327 (2005).
153. Braat, A. K., Yan, N., Arn, E., Harrison, D. & Macdonald, P. M. Localization-Dependent Oskar Protein Accumulation. *Dev Cell* **7**, 125–131 (2004).
154. Riechmann, V., Gutierrez, G. J., Filardo, P., Nebreda, A. R. & Ephrussi, A. Par-1 regulates stability of the posterior determinant Oskar by phosphorylation. *Nat Cell Biol* **4**, 337–342 (2002).
155. Markussen, F. H., Michon, A. M., Breitwieser, W. & Ephrussi, A. Translational control of oskar generates short OSK, the isoform that induces pole plasma assembly. **121**, 3723–3732 (1995).
156. Rongo, C., Gavis, E. R. & Lehmann, R. Localization of oskar RNA regulates oskar translation and requires Oskar protein. **121**, 2737–2746 (1995).
157. Benton, R. & Johnston, D. S. Cell Polarity: Posterior Par-1 Prevents Proteolysis. *Current Biology* **12**, R479–R481 (2002).
158. Giorgi, C. *et al.* The EJC Factor eIF4AIII Modulates Synaptic Strength and Neuronal Protein Expression. *Cell* **130**, 179–191 (2007).
159. Hachet, O. & Ephrussi, A. Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* **428**, 959–963 (2004).
160. Kim-Ha, J., Kerr, K. & Macdonald, P. M. Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* **81**, 403–412 (1995).

161. Filardo, P. & Ephrussi, A. Bruno regulates gurken during *Drosophila* oogenesis. *Mech Dev* **120**, 289–297 (2003).
162. Goodrich, J. S., Clouse, K. N. & Schüpbach, T. Hrb27C, Sqd and Otu cooperatively regulate gurken RNA localization and mediate nurse cell chromosome dispersion in *Drosophila* oogenesis. **131**, 1949–1958 (2004).
163. Gunkel, N., Yano, T., Markussen, F.-H., Olsen, L. C. & Ephrussi, A. Localization-dependent translation requires a functional interaction between the 5' and 3' ends of oskar mRNA. *Genes & Development* **12**, 1652–1664 (1998).
164. Yano, T. *et al.* Hrp48, a *Drosophila* hnRNPA/B Homolog, Binds and Regulates Translation of oskar mRNA. *Dev Cell* **6**, 637–648 (2004).
165. Gu, W., Pan, F., Zhang, H., Bassell, G. J. & Singer, R. H. A predominantly nuclear protein affecting cytoplasmic localization of β -actin mRNA in fibroblasts and neurons. *J Cell Biol* **156**, 41–52 (2002).
166. Nakamura, A., Sato, K. & Hanyu-Nakamura, K. *Drosophila* Cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA Translation in Oogenesis. *Dev Cell* **6**, 69–78 (2004).
167. Paquin, N. *et al.* Local Activation of Yeast ASH1 mRNA Translation through Phosphorylation of Khd1p by the Casein Kinase Yck1p. *Mol Cell* **26**, 795–809 (2007).
168. Ostareck, D. H., Ostareck-Lederer, A., Shatsky, I. N. & Hentze, M. W. Lipoxxygenase mRNA Silencing in Erythroid Differentiation. *Cell* **104**, 281–290 (2001).

169. Frank-Vaillant, M., Jesus, C. & Ozon, R. *Two distinct mechanisms control the accumulation of cyclin B1 and Mos in Xenopus oocytes in response to progesterone.* (Molecular Biology of ..., 1999). doi:10.1016/S0248-4900(01)01126-1/abstract
170. Ivshina, M., Lasko, P. & Richter, J. D. Cytoplasmic Polyadenylation Element Binding Proteins in Development, Health, and Disease. <http://dx.doi.org/10.1146/annurev-cellbio-101011-155831> **30**, 393–415 (2014).
171. Mendez, R. & Richter, J. D. Translational control by CPEB: a means to the end. *Nat Rev Mol Cell Biol* **2**, 521–529 (2001).
172. de Moor, C. H. & Richter, J. D. The Mos pathway regulates cytoplasmic polyadenylation in *Xenopus* oocytes. *Mol Cell Biol* **17**, 6419–6426 (1997).
173. Richter, J. D. & Lasko, P. Translational control in oocyte development. *Cold Spring Harb Perspect Biol* **3**, a002758–a002758 (2011).
174. Fox, C. A., Sheets, M. D. & Wickens, M. P. Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUUAU. *Genes & Development* **3**, 2151–2162 (1989).
175. Hake, L. E. & Richter, J. D. CPEB is a specificity factor that mediates cytoplasmic polyadenylation during *Xenopus* oocyte maturation. *Cell* **79**, 617–627 (1994).
176. Bilger, A., Fox, C. A., Wahle, E. & Wickens, M. Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. *Genes & Development* **8**, 1106–1116 (1994).

177. Dickson, K. S., Bilger, A., Ballantyne, S. & Wickens, M. P. The cleavage and polyadenylation specificity factor in *Xenopus laevis* oocytes is a cytoplasmic factor involved in regulated polyadenylation. *Mol Cell Biol* **19**, 5707–5717 (1999).
178. McGrew, L. L., Dworkin-Rastl, E., Dworkin, M. B. & Richter, J. D. Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes & Development* **3**, 803–815 (1989).
179. Paris, J. & Richter, J. D. Maturation-specific polyadenylation and translational control: diversity of cytoplasmic polyadenylation elements, influence of poly(A) tail size, and formation of stable polyadenylation complexes. *Mol Cell Biol* **10**, 5634–5645 (1990).
180. Mandel, C. R. *et al.* Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**, 953–956 (2006).
181. Huarte, J. *et al.* Transient translational silencing by reversible mRNA deadenylation. *Cell* **69**, 1021–1030 (1992).
182. Kim, J. H. & Richter, J. D. Opposing Polymerase-Deadenylase Activities Regulate Cytoplasmic Polyadenylation. *Mol Cell* **24**, 173–183 (2006).
183. Sachs, A. & Wahle, E. Poly (A) tail metabolism and function in eucaryotes. *Journal of Biological Chemistry* (1993).
184. Barnard, D. C., Ryan, K., Manley, J. L. & Richter, J. D. Symplekin and xGLD-2 Are Required for CPEB-Mediated Cytoplasmic Polyadenylation. *Cell* **119**, 641–651 (2004).

185. Cao, Q. & Richter, J. D. Dissolution of the maskin-eIF4E complex by cytoplasmic polyadenylation and poly(A)-binding protein controls cyclin B1 mRNA translation and oocyte maturation. *EMBO J* **21**, 3852–3862 (2002).
186. Stebbins-Boaz, B., Cao, Q., de Moor, C. H., Mendez, R. & Richter, J. D. Maskin Is a CPEB-Associated Factor that Transiently Interacts with eIF-4E. *Mol Cell* **4**, 1017–1027 (1999).
187. Wilhelm, J. E., Hilton, M., Amos, Q. & Henzel, W. J. Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. **163**, 1197–1204 (2003).
188. Lasko, P. Cup-ling oskar RNA localization and translational control. *J Cell Biol* **163**, 1189–1191 (2003).
189. Pinder, B. D. & Smibert, C. A. microRNA-independent recruitment of Argonaute 1 to nanos mRNA through the Smaug RNA-binding protein. *EMBO reports* **14**, 80–86 (2013).
190. Bicker, S. *et al.* The DEAH-box helicase DHX36 mediates dendritic localization of the neuronal precursor-microRNA-134. *Genes & Development* **27**, 991–996 (2013).
191. Muddashetty, R. S. *et al.* Reversible Inhibition of PSD-95 mRNA Translation by miR-125a, FMRP Phosphorylation, and mGluR Signaling. *Mol Cell* **42**, 673–688 (2011).
192. Kamenska, A. *et al.* Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. *Nucl. Acids Res.* **42**, 3298–3313 (2014).

193. Weil, T. T. *et al.* Drosophila patterning is established by differential association of mRNAs with P bodies. *Nat Cell Biol* **14**, 1305–1313 (2012).
194. Salles, F. J., Lieberfarb, M. E., Wreden, C., Gergen, J. P. & Strickland, S. Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. *Science* **266**, 1996–1999 (1994).
195. Surdej, P. & Jacobs-Lorena, M. Developmental regulation of bicoid mRNA stability is mediated by the first 43 nucleotides of the 3' untranslated region. *Mol Cell Biol* **18**, 2892–2900 (1998).
196. Wilsch-Bräuninger, M., Schwarz, H. & Nüsslein-Volhard, C. A sponge-like structure involved in the association and transport of maternal products during Drosophila oogenesis. *J Cell Biol* **139**, 817–829 (1997).
197. Aizer, A. *et al.* The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol Biol Cell* **19**, 4154–4166 (2008).
198. Beckham, C. J. & Parker, R. P bodies, stress granules, and viral life cycles. *Cell Host Microbe* **3**, 206–212 (2008).
199. Minshall, N., Kress, M., Weil, D. & Standart, N. Role of p54 RNA helicase activity and its C-terminal domain in translational repression, P-body localization and assembly. *Mol Biol Cell* **20**, 2464–2472 (2009).
200. Barbee, S. A. *et al.* Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. *Neuron* **52**, 997–1009 (2006).

201. Zaessinger, S., Busseau, I. & Simonelig, M. Oskar allows nanos mRNA translation in *Drosophila* embryos by preventing its deadenylation by Smaug/CCR4. **133**, 4573–4583 (2006).
202. Hüttelmaier, S. *et al.* Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* **438**, 512–515 (2005).
203. Rosenthal, E. T., Tansey, T. R. & Ruderman, J. V. Sequence-specific adenylations and deadenylations accompany changes in the translation of maternal messenger RNA after fertilization of *Spisula* oocytes. *J. Mol. Biol.* **166**, 309–327 (1983).
204. Craig, A. W., Haghighat, A., Yu, A. T. & Sonenberg, N. Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392**, 520–523 (1998).
205. Wells, S. E., Hillner, P. E., Vale, R. D. & Sachs, A. B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **2**, 135–140 (1998).
206. Bergsten, S. E. & Gavis, E. R. Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. **126**, 659–669 (1999).
207. Nelson, M. R., Leidal, A. M. & Smibert, C. A. *Drosophila* Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. *EMBO J* **23**, 150–159 (2004).
208. Decker, C. J. & Parker, R. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol* **4**, a012286–a012286 (2012).

209. Decker, C. J. & Parker, R. A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes & Development* **7**, 1632–1643 (1993).
210. Dunkley, T. & Parker, R. The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. *EMBO J* **18**, 5411–5422 (1999).
211. Hsu, C. L. & Stevens, A. Yeast cells lacking 5'→3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol Cell Biol* **13**, 4826–4835 (1993).
212. Muhlrads, D., Decker, C. J. & Parker, R. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'→3' digestion of the transcript. *Genes & Development* **8**, 855–866 (1994).
213. Muhlrads, D., Decker, C. J. & Parker, R. Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol* **15**, 2145–2156 (1995).
214. Steiger, M., Carr-Schmid, A., Schwartz, D. C., Kiledjian, M. & Parker, R. Analysis of recombinant yeast decapping enzyme. *RNA* **9**, 231–238 (2003).
215. van Dijk, E. *et al.* Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J* **21**, 6915–6924 (2002).
216. Hu, W., Sweet, T. J., Chamnongpol, S., Baker, K. E. & Coller, J. Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature* **461**, 225–229 (2009).

217. Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M. & Parker, R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371–382 (2005).
218. Isken, O. & Maquat, L. E. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes & Development* **21**, 1833–1856 (2007).
219. Muhlrads, D. & Parker, R. Premature translational termination triggers mRNA decapping. *Nature* **370**, 578–581 (1994).
220. Trcek, T., Sato, H., Singer, R. H. & Maquat, L. E. Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes & Development* **27**, 541–551 (2013).
221. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006).
222. Harigaya, Y. & Parker, R. No-go decay: a quality control mechanism for RNA in translation. *Wiley Interdiscip Rev RNA* **1**, 132–141 (2010).
223. Frischmeyer, P. A. *et al.* An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* **295**, 2258–2261 (2002).
224. van Hoof, A., Frischmeyer, P. A., Dietz, H. C. & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262–2264 (2002).

225. Bushati, N., Stark, A., Brennecke, J. & Cohen, S. M. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in *Drosophila*. **18**, 501–506 (2008).
226. Giraldez, A. J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–79 (2006).
227. Baugh, L. R., Hill, A. A., Slonim, D. K., Brown, E. L. & Hunter, C. P. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. **130**, 889–900 (2003).
228. De Renzis, S., Elemento, O., Tavazoie, S. & Wieschaus, E. F. Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. *PLoS Biol* **5**, e117 (2007).
229. Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. H. Dynamics of Global Gene Expression Changes during Mouse Preimplantation Development. *Dev Cell* **6**, 117–131 (2004).
230. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. **136**, 3033–3042 (2009).
231. Bregman, A. *et al.* Promoter Elements Regulate Cytoplasmic mRNA Decay. *Cell* **147**, 1473–1483 (2011).
232. Reiter, L. T., Potocki, L., Chien, S., Gribskov, M. & Bier, E. A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res* **11**, 1114–1125 (2001).

233. Lloyd, T. E. & Taylor, J. P. Flightless flies: *Drosophila* models of neuromuscular disease. *Annals of the New York Academy of Sciences* **1184**, E1–E20 (2010).
234. Spradling, A. C. *Developmental genetics of oogenesis*. (The development of *Drosophila melanogaster*, 1993).
235. Bolvar, J. *et al.* Centrosome migration into the *Drosophila* oocyte is independent of BicD and egl, and of the organisation of the microtubule cytoskeleton. **128**, 1889–1897 (2001).
236. Cox, R. T. & Spradling, A. C. A Balbiani body and the fusome mediate mitochondrial inheritance during *Drosophila* oogenesis. **130**, 1579–1590 (2003).
237. de Cuevas, M. & Spradling, A. C. Morphogenesis of the *Drosophila* fusome and its implications for oocyte specification. **125**, 2781–2789 (1998).
238. Grieder, N. C., de Cuevas, M. & Spradling, A. C. The fusome organizes the microtubule network during oocyte differentiation in *Drosophila*. **127**, 4253–4264 (2000).
239. Huynh, J. & Johnston, D. S. The role of BicD, egl, orb and the microtubules in the restriction of meiosis to the *Drosophila* oocyte. **127**, 2785–2794 (2000).
240. Huynh, J. R., Shulman, J. M., Benton, R. & Johnston, D. S. PAR-1 is required for the maintenance of oocyte fate in *Drosophila*. **128**, 1201–1209 (2001).
241. McGrail, M. & Hays, T. S. The microtubule motor cytoplasmic dynein is required for spindle orientation during germline cell divisions and oocyte differentiation in *Drosophila*. **124**, 2409–2419 (1997).

242. Cox, D. N., Seyfried, S. A., Jan, L. Y. & Jan, Y. N. Bazooka and atypical protein kinase C are required to regulate oocyte differentiation in the *Drosophila* ovary. *PNAS* **98**, 14475–14480 (2001).
243. Huynh, J.-R. & St Johnston, D. The origin of asymmetry: early polarisation of the *Drosophila* germline cyst and oocyte. **14**, R438–49 (2004).
244. KING, R. C. & Burnett, R. G. Autoradiographic Study of Uptake of Tritiated Glycine, Thymidine, and Uridine by Fruit Fly Ovaries. *Science* **129**, 1674–1675 (1959).
245. Spradling, A. C. Germline cysts: communes that work. *Cell* **72**, 649–651 (1993).
246. Godt, D. & Tepass, U. *Drosophila* oocyte localization is mediated by differential cadherin-based adhesion. *Nature* **395**, 387–391 (1998).
247. González-Reyes, A. & Johnston, D. S. The *Drosophila* AP axis is polarised by the cadherin-mediated positioning of the oocyte. **125**, 3635–3644 (1998).
248. Horne-Badovinac, S. & Bilder, D. Mass transit: Epithelial morphogenesis in the *Drosophila* egg chamber. *Dev Dyn* **232**, 559–574 (2005).
249. Grammont, M. & Irvine, K. D. fringe and Notch specify polar cell fate during *Drosophila* oogenesis. **128**, 2243–2253 (2001).
250. McGregor, J. R., Xi, R. & Harrison, D. A. JAK signaling is somatically required for follicle cell differentiation in *Drosophila*. **129**, 705–717 (2002).
251. Roth, S. *Drosophila* oogenesis: Coordinating germ line and soma. *Current Biology* **11**, R779–R781 (2001).

252. Koch, E. A. & Spitzer, R. H. 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 (1983).
253. Pokrywka, N. J. & Stephenson, E. C. Microtubules Are a General Component of mRNA Localization Systems in *Drosophila* Oocytes. *Dev Biol* **167**, 363–370 (1995).
254. Theurkauf, W. E., Alberts, B. M., Jan, Y. N. & Jongens, T. A. A central role for microtubules in the differentiation of *Drosophila* oocytes. **118**, 1169–1180 (1993).
255. Ran, B., Bopp, R. & Suter, B. Null alleles reveal novel requirements for Bic-D during *Drosophila* oogenesis and zygotic development. **120**, 1233–1242 (1994).
256. Swan, A., Nguyen, T. & Suter, B. *Drosophila* Lissencephaly-1 functions with Bic-D and dynein in oocyte determination and nuclear positioning. *Nat Cell Biol* **1**, 444–449 (1999).
257. King, R. C. *Ovarian development in Drosophila melanogaster*. (1970).
258. Botchan, M. & Levine, M. A Genome Analysis of Endoreplication in the *Drosophila* Ovary. *Dev Cell* **6**, 4–5 (2004).
259. Calvi, B. R. & Spradling, A. C. Chorion Gene Amplification in *Drosophila*: A Model for Metazoan Origins of DNA Replication and S-Phase Control. *Methods* **18**, 407–417 (1999).
260. González-Reyes, A. & Johnston, D. S. Patterning of the follicle cell epithelium along the anterior-posterior axis during *Drosophila* oogenesis. **125**, 2837–2846 (1998).

261. Wu, X., Tanwar, P. S. & Raftery, L. A. *Drosophila* follicle cells: Morphogenesis in an eggshell. *Semin Cell Dev Biol* **19**, 271–282 (2008).
262. Januschke, J. *et al.* Polar transport in the *Drosophila* oocyte requires Dynein and Kinesin I cooperation. **12**, 1971–1981 (2002).
263. Zhao, T., Graham, O. S., Raposo, A. & Johnston, D. S. Growing Microtubules Push the Oocyte Nucleus to Polarize the *Drosophila* Dorsal-Ventral Axis. *Science* **336**, 999–1003 (2012).
264. Montell, D. J. Border-cell migration: The race is on. *Nat Rev Mol Cell Biol* **4**, 13–24 (2003).
265. Serbus, L. R., Cha, B.-J., Theurkauf, W. E. & Saxton, W. M. Dynein and the actin cytoskeleton control kinesin-driven cytoplasmic streaming in *Drosophila* oocytes. **132**, 3743–3752 (2005).
266. Gutzeit, H. O. The role of microfilaments in cytoplasmic streaming in *Drosophila* follicles. *J. Cell. Sci.* **80**, 159–169 (1986).
267. Kronja, I. *et al.* Quantitative proteomics reveals the dynamics of protein changes during *Drosophila* oocyte maturation and the oocyte-to-embryo transition. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 16023–16028 (2014).
268. Saunders, C. & Cohen, R. S. 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 (1999).

269. Cáceres, L. & Nilson, L. A. Production of gurken in the nurse cells is sufficient for axis determination in the *Drosophila* oocyte. **132**, 2345–2353 (2005).
270. Clark, A., Meignin, C. & Davis, I. A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte. **134**, 1955–1965 (2007).
271. Van Buskirk, C. & Schüpbach, T. Versatility in signalling: multiple responses to EGF receptor activation during *Drosophila* oogenesis. *Trends Cell Biol* **9**, 1–4 (1999).
272. Schüpbach, T. Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in *Drosophila melanogaster*. *Cell* **49**, 699–707 (1987).
273. González-Reyes, A., Elliott, H. & St Johnston, D. Polarization of both major body axes in *Drosophila* by gurken-torpedo signalling. *Nature* **375**, 654–658 (1995).
274. Roth, S., Neuman-Silberberg, F. S., Barcelo, G. & Schüpbach, T. cornichon and the EGF receptor signaling process are necessary for both anterior-posterior and dorsal-ventral pattern formation in *Drosophila*. *Cell* **81**, 967–978 (1995).
275. Meignin, C., Alvarez-Garcia, I., Davis, I. & Palacios, I. M. The salvador-warts-hippo pathway is required for epithelial proliferation and axis specification in *Drosophila*. **17**, 1871–1878 (2007).
276. Polesello, C. & Tapon, N. Salvador-warts-hippo signaling promotes *Drosophila* posterior follicle cell maturation downstream of notch. **17**, 1864–1870 (2007).

277. Arn, E. A., Cha, B. J., Theurkauf, W. E. & Macdonald, P. M. Recognition of a bicoid mRNA localization signal by a protein complex containing Swallow, Nod, and RNA binding proteins. *Dev Cell* **4**, 41–51 (2003).
278. Schnorrer, F., Bohmann, K. & Nüsslein-Volhard, C. The molecular motor dynein is involved in targeting swallow and bicoid RNA to the anterior pole of *Drosophila* oocytes. *Nat Cell Biol* **2**, 185–190 (2000).
279. Berleth, T. *et al.* The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO Journal* **7**, 1749–1756 (1988).
280. Driever, W. & Nüsslein-Volhard, C. A gradient of bicoid protein in *Drosophila* embryos. *Cell* **54**, 83–93 (1988).
281. Brendza, R. P., Serbus, L. R., Duffy, J. B. & Saxton, W. M. A function for kinesin I in the posterior transport of oskar mRNA and Staufen protein. *Science* **289**, 2120–2122 (2000).
282. Kim-Ha, J., Smith, J. L. & Macdonald, P. M. oskar mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* **66**, 23–35 (1991).
283. Jaramillo, A. M., Weil, T. T., Goodhouse, J., Gavis, E. R. & Schüpbach, T. The dynamics of fluorescently labeled endogenous gurken mRNA in *Drosophila*. *J. Cell. Sci.* **121**, 887–894 (2008).
284. Neuman-Silberberg, F. S. & Schüpbach, T. The *Drosophila* TGF- α -like protein Gurken: expression and cellular localization during *Drosophila* oogenesis. *Mech Dev* **59**, 105–113 (1996).

285. Sapir, A., Schweitzer, R. & Shilo, B. Z. Sequential activation of the EGF receptor pathway during *Drosophila* oogenesis establishes the dorsoventral axis. **125**, 191–200 (1998).
286. Wasserman, J. D. & Freeman, M. An Autoregulatory Cascade of EGF Receptor Signaling Patterns the *Drosophila* Egg. *Cell* **95**, 355–364 (1998).
287. Queenan, A. M., Barcelo, G., Van Buskirk, C. & Schüpbach, T. The transmembrane region of Gurken is not required for biological activity, but is necessary for transport to the oocyte membrane in *Drosophila*. *Mech Dev* **89**, 35–42 (1999).
288. Peri, F., Bökel, C. & Roth, S. Local Gurken signaling and dynamic MAPK activation during *Drosophila* oogenesis. *Mech Dev* **81**, 75–88 (1999).
289. Ghigliione, C. *et al.* Mechanism of activation of the *Drosophila* EGF Receptor by the TGF α ligand Gurken during oogenesis. **129**, 175–186 (2002).
290. Shmueli, A., Cohen-Gazala, O. & Neuman-Silberberg, F. S. Gurken, a TGF- α -like Protein Involved in Axis Determination in *Drosophila*, Directly Binds to the EGF-Receptor Homolog Egfr. *Biochem Biophys Res Commun* **291**, 732–737 (2002).
291. Goentoro, L. A. *et al.* Quantifying the Gurken Morphogen Gradient in *Drosophila* Oogenesis. *Dev Cell* **11**, 263–272 (2006).
292. Yakoby, N. *et al.* A Combinatorial Code for Pattern Formation in *Drosophila* Oogenesis. *Dev Cell* **15**, 725–737 (2008).

293. James, K. E., Dorman, J. B. & Berg, C. A. Mosaic analyses reveal the function of *Drosophila* Ras in embryonic dorsoventral patterning and dorsal follicle cell morphogenesis. **129**, 2209–2222 (2002).
294. Peri, F., Technau, M. & Roth, S. Mechanisms of Gurken-dependent pipe regulation and the robustness of dorsoventral patterning in *Drosophila*. **129**, 2965–2975 (2002).
295. Technau, M., Knispel, M. & Roth, S. Molecular mechanisms of EGF signaling-dependent regulation of pipe, a gene crucial for dorsoventral axis formation in *Drosophila*. *Dev. Genes Evol.* **222**, 1–17 (2012).
296. Sen, J., Goltz, J. S., Stevens, L. & Stein, D. Spatially Restricted Expression of pipe in the *Drosophila* Egg Chamber Defines Embryonic Dorsal–Ventral Polarity. *Cell* **95**, 471–481 (1998).
297. Nilson, L. A. & Schüpbach, T. Localized Requirements for windbeutel and pipe Reveal a Dorsoventral Prepattern within the Follicular Epithelium of the *Drosophila* Ovary. *Cell* **93**, 253–262 (1998).
298. Dorman, J. B., James, K. E., Fraser, S. E., Kiehart, D. P. & Berg, C. A. bullwinkle is required for epithelial morphogenesis during *Drosophila* oogenesis. *Dev Biol* **267**, 320–341 (2004).
299. Kelley, R. L. Initial organization of the *Drosophila* dorsoventral axis depends on an RNA-binding protein encoded by the squid gene. **7**, 948–960 (1993).
300. Neuman-Silberberg, F. S. & Schupbach, T. Dorsoventral axis formation in *Drosophila* depends on the correct dosage of the gene gurken. **120**, 2457–2463 (1994).

301. Pokrywka, N. J. & Stephenson, E. C. Microtubules mediate the localization of bicoid RNA during *Drosophila* oogenesis. **113**, 55–66 (1991).
302. Brendza, R. P., Serbus, L. R., Saxton, W. M. & Duffy, J. B. Posterior localization of dynein and dorsal-ventral axis formation depend on kinesin in *Drosophila* oocytes. **12**, 1541–1545 (2002).
303. Rom, I., Faicevici, A., Almog, O. & Neuman-Silberberg, F. S. *Drosophila* Dynein light chain (DDL1) binds to gurken mRNA and is required for its localization. *Biochim Biophys Acta* **1773**, 1526–1533 (2007).
304. Dienstbier, M., Boehl, F., Li, X. & Bullock, S. L. Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor. *Genes & Development* **23**, 1546–1558 (2009).
305. Lan, L., Lin, S., Zhang, S. & Cohen, R. S. Evidence for a transport-trap mode of *Drosophila melanogaster* gurken mRNA localization. *PLoS ONE* **5**, e15448 (2010).
306. Thio, G. L., Ray, R. P., Barcelo, G. & Schüpbach, T. Localization of gurken RNA in *Drosophila* oogenesis requires elements in the 5' and 3' regions of the transcript. *Dev Biol* **221**, 435–446 (2000).
307. Van De Bor, V., Hartswood, E., Jones, C., Finnegan, D. & Davis, I. gurken and the I factor retrotransposon RNAs share common localization signals and machinery. *Dev Cell* **9**, 51–62 (2005).
308. Lehmann, R. & Nüsslein-Volhard, C. hunchback, a gene required for segmentation of an anterior and posterior region of the *Drosophila* embryo. *Dev Biol* **119**, 402–417 (1987).

309. Manseau, L. J. & Schüpbach, T. cappuccino and spire: two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the *Drosophila* embryo. *Genes & Development* **3**, 1437–1452 (1989).
310. Clegg, N. J. *et al.* maelstrom is required for an early step in the establishment of *Drosophila* oocyte polarity: posterior localization of grk mRNA. **124**, 4661–4671 (1997).
311. Emmons, S. *et al.* Cappuccino, a *Drosophila* maternal effect gene required for polarity of the egg and embryo, is related to the vertebrate limb deformity locus. *Genes & Development* **9**, 2482–2494 (1995).
312. Serano, T. L. & Cohen, R. S. A small predicted stem-loop structure mediates oocyte localization of *Drosophila* K10 mRNA. **121**, 3809–3818 (1995).
313. Wieschaus, E., Marsh, J. L. & Gehring, W. fs(1)K10, a germline-dependent female sterile mutation causing abnormal chorion morphology in *Drosophila melanogaster*. *Wilhelm Roux' Archiv* **184**, 75–82 (1978).
314. Clouse, K. N., Ferguson, S. B. & Schüpbach, T. Squid, Cup, and PABP55B function together to regulate gurken translation in *Drosophila*. *Dev Biol* **313**, 713–724 (2008).
315. Geng, C. & Macdonald, P. M. Imp associates with squid and Hrp48 and contributes to localized expression of gurken in the oocyte. *Mol Cell Biol* **26**, 9508–9516 (2006).
316. Kalifa, Y., Armenti, S. T. & Gavis, E. R. Glorund interactions in the regulation of gurken and oskar mRNAs. *Dev Biol* **326**, 68–74 (2009).

317. Norvell, A., Kelley, R. L., Wehr, K. & Schüpbach, T. Specific isoforms of squid, a *Drosophila* hnRNP, perform distinct roles in Gurken localization during oogenesis. **13**, 864–876 (1999).
318. Chang, J. S., Tan, L., Wolf, M. R. & Schedl, P. Functioning of the *Drosophila* orb gene in gurken mRNA localization and translation. **128**, 3169–3177 (2001).
319. Hawkins, N. C., Thorpe, J. & Schüpbach, T. Encore, a gene required for the regulation of germ line mitosis and oocyte differentiation during *Drosophila* oogenesis. **122**, 281–290 (1996).
320. Hawkins, N. C., Van Buskirk, C., Grossniklaus, U. & Schüpbach, T. Post-transcriptional regulation of gurken by encore is required for axis determination in *Drosophila*. **124**, 4801–4810 (1997).
321. Styhler, S., Nakamura, A., Swan, A., Suter, B. & Lasko, P. vasa is required for GURKEN accumulation in the oocyte, and is involved in oocyte differentiation and germline cyst development. **125**, 1569–1578 (1998).
322. Tomancak, P., Guichet, A., Zavorszky, P. & Ephrussi, A. Oocyte polarity depends on regulation of gurken by Vasa. **125**, 1723–1732 (1998).
323. Van Buskirk, C., Hawkins, N. C. & Schüpbach, T. Encore is a member of a novel family of proteins and affects multiple processes in *Drosophila* oogenesis. **127**, 4753–4762 (2000).
324. Czaplinski, K. *et al.* Identification of 40LoVe, a *Xenopus* hnRNP D family protein involved in localizing a TGF-beta-related mRNA during oogenesis. *Dev Cell* **8**, 505–515 (2005).

325. Duncan, K. *et al.* Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2 mRNA 3' UTR: translational repression for dosage compensation. *Genes & Development* **20**, 368–379 (2006).
326. Nakamura, A., Amikura, R., Hanyu, K. & Kobayashi, S. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis. *Development* **128**, 3233–3242 (2001).
327. Christerson, L. B. & McKearin, D. M. orb is required for anteroposterior and dorsoventral patterning during *Drosophila* oogenesis. *Genes & Development* **8**, 614–628 (1994).
328. Roth, S. & Schüpbach, T. The relationship between ovarian and embryonic dorsoventral patterning in *Drosophila*. **120**, 2245–2257 (1994).
329. Davidson, A., Parton, R. M., Rabouille, C., Weil, T. T. & Davis, I. Localized Translation of gurken/TGF- α mRNA during Axis Specification Is Controlled by Access to Orb/CPEB on Processing Bodies. *CellReports* **14**, 2451–2462 (2016).
330. Burn, K. M. *et al.* Somatic insulin signaling regulates a germline starvation response in *Drosophila* egg chambers. *Dev Biol* **398**, 206–217 (2015).
331. Dawson, A., Hartswood, E., Paterson, T. & Finnegan, D. J. A LINE-like transposable element in *Drosophila*, the I factor, encodes a protein with properties similar to those of retroviral nucleocapsids. *EMBO J* **16**, 4448–4455 (1997).
332. Fawcett, D. H., Lister, C. K., Kellett, E. & Finnegan, D. J. Transposable elements controlling I-R hybrid dysgenesis in *D. melanogaster* are similar to mammalian LINEs. *Cell* **47**, 1007–1015 (1986).

333. Seleme, M. D. C. *et al.* In vivo RNA localization of I factor, a non-LTR retrotransposon, requires a cis-acting signal in ORF2 and ORF1 protein. *Nucl. Acids Res.* **33**, 776–785 (2005).
334. Schüpbach, T. Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in *Drosophila melanogaster*. *Cell* **49**, 699–707 (1987).
335. Hamilton, R. S. *et al.* A bioinformatics search pipeline, RNA2DSearch, identifies RNA localization elements in *Drosophila* retrotransposons. *RNA* **15**, 200–207 (2009).
336. Orjalo, A., Jr, Johansson, H. E. & Ruth, J. L. Stellaris™ fluorescence in situ hybridization (FISH) probes: a powerful tool for mRNA detection. *Nat Methods* (2011).
337. Wilkie, G. S. & Davis, I. *Drosophila* wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles. *Cell* **105**, 209–219 (2001).
338. Glotzer, J. B., Saffrich, R., Glotzer, M. & Ephrussi, A. Cytoplasmic flows localize injected oskar RNA in *Drosophila* oocytes. *Current Biology* **7**, 326–337 (1997).
339. Lall, S. *et al.* Squid hnRNP protein promotes apical cytoplasmic transport and localization of *Drosophila* pair-rule transcripts. *Cell* **98**, 171–180 (1999).
340. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **2**, 437–445 (1998).
341. Weil, T. T., Parton, R. M. & Davis, I. Making the message clear: visualizing mRNA localization. *Trends Cell Biol* **20**, 380–390 (2010).

342. Kislauskis, E. H., Zhu, X.-C. & Singer, R. H. β -Actin Messenger RNA Localization and Protein Synthesis Augment Cell Motility. *J Cell Biol* **136**, 1263–1270 (1997).
343. Chao, J. A. *et al.* ZBP1 recognition of beta-actin zipcode induces RNA looping. *Genes & Development* **24**, 148–158 (2010).
344. Ross, A. F., Oleynikov, Y., Kislauskis, E. H., Taneja, K. L. & Singer, R. H. Characterization of a beta-actin mRNA zipcode-binding protein. *Mol Cell Biol* **17**, 2158–2165 (1997).
345. Patel, V. L. *et al.* Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control. **26**, 43–53 (2012).
346. Lemieux, S. & Major, F. RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucl. Acids Res.* **30**, 4250–4263 (2002).
347. Chastain, M. & Tinoco, I., Jr. in **41**, 131–177 (Elsevier, 1991).
348. Hermann, T. & Patel, D. J. RNA bulges as architectural and recognition motifs. *Structure* **8**, R47–R54 (2000).
349. Leontis, N. B., Lescoute, A. & Westhof, E. The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* **16**, 279–287 (2006).
350. Hamilton, R. S. & Davis, I. Identifying and Searching for Conserved RNA Localisation Signals. *Methods in molecular biology (Clifton, NJ)* **714**, 447–466 (2011).
351. Markham, N. R. & Zuker, M. UNAFold: software for nucleic acid folding and hybridization. *Methods Mol. Biol.* **453**, 3–31 (2008).

352. Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R. & Stadler, P. F. RNAalifold: improved consensus structure prediction for RNA alignments. *BMC Bioinformatics* **9**, 474 (2008).
353. Li, X., Quon, G., Lipshitz, H. D. & Morris, Q. Predicting in vivo binding sites of RNA-binding proteins using mRNA secondary structure. **16**, 1096–1107 (2010).
354. Gorodkin, J. & Ruzzo, W. L. *RNA Sequence, Structure, and Function: Computational and Bioinformatic Methods*. **1097**, (Humana Press, 2014).
355. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**, 479–490 (2007).
356. Overall, C. M. *et al.* Protease degradomics: mass spectrometry discovery of protease substrates and the CLIP-CHIP, a dedicated DNA microarray of all human proteases and inhibitors. *Biol. Chem.* **385**, 493–504 (2004).
357. Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* **1**, 302–307 (2006).
358. Gerber, A. P., Herschlag, D. & Brown, P. O. Extensive association of functionally and cytotoxically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* **2**, E79 (2004).
359. Hogan, D. J., Riordan, D. P., Gerber, A. P., Herschlag, D. & Brown, P. O. Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol* **6**, e255 (2008).

360. Ray, D. *et al.* Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat Biotechnol* **27**, 667–670 (2009).
361. Ellis, J. J., Broom, M. & Jones, S. Protein-RNA interactions: structural analysis and functional classes. *Proteins* **66**, 903–911 (2007).
362. Rajkowitsch, L. & Schroeder, R. Dissecting RNA chaperone activity. *RNA* **13**, 2053–2060 (2007).
363. Fried, M. & Crothers, D. M. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucl. Acids Res.* **9**, 6505–6525 (1981).
364. Garner, M. M. & Revzin, A. The use of gel electrophoresis to detect and study nucleic acid- protein interactions. *Trends Biochem Sci* **11**, 395–396 (1986).
365. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
366. Ule, J., Jensen, K., Mele, A. & Darnell, R. B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376–386 (2005).
367. Spitzer, J. *et al.* PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. *Meth Enzymol* **539**, 113–161 (2014).
368. Huppertz, I. *et al.* iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* **65**, 274–287 (2014).

369. Sambrook, J. & Russell, D. W. *Molecular Cloning*. (CSHL Press, 2001).
370. Weil, T. T., Parton, R. M. & Davis, I. Preparing individual *Drosophila* egg chambers for live imaging. *J Vis Exp* e3679–e3679 (2012). doi:10.3791/3679
371. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* **32**, 1792–1797 (2004).
372. Boc, A., Diallo, A. B. & Makarenkov, V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucl. Acids Res.* **40**, W573–9 (2012).
373. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190 (2004).
374. Brunel, C. & Ehresmann, C. Secondary structure of the 3' UTR of bicoid mRNA. *Biochimie* **86**, 91–104 (2004).
375. Macdonald, P. M. bicoid mRNA localization signal: phylogenetic conservation of function and RNA secondary structure. *Development* **110**, 161–171 (1990).
376. Simon, B., Masiewicz, P., Ephrussi, A. & Carlomagno, T. The structure of the SOLE element of oskar mRNA. *RNA* **21**, 1444–1453 (2015).
377. Cha, B. J., Koppetsch, B. S. & Theurkauf, W. E. In vivo analysis of *Drosophila* bicoid mRNA localization reveals a novel microtubule-dependent axis specification pathway. *Cell* **106**, 35–46 (2001).

378. Bullock, S. L. & Ish-Horowicz, D. Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* **414**, 611–616 (2001).
379. Vendra, G., Hamilton, R. S. & Davis, I. Dynactin suppresses the retrograde movement of apically localized mRNA in *Drosophila* blastoderm embryos. **13**, 1860–1867 (2007).
380. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucl. Acids Res.* **41**, W597–W600 (2013).
381. Gupta, A. & Gribskov, M. The role of RNA sequence and structure in RNA-protein interactions. *J. Mol. Biol.* **409**, 574–587 (2011).
382. Lynch, J. A., Peel, A. D., Drechsler, A., Averof, M. & Roth, S. EGF signaling and the origin of axial polarity among the insects. **20**, 1042–1047 (2010).
383. Cáceres, L. & Nilson, L. A. Translational repression of *gurken* mRNA in the *Drosophila* oocyte requires the hnRNP Squid in the . *Dev Biol* **326**, 327–334 (2009).
384. Bannai, H. *et al.* An RNA-interacting protein, SYNCRIP (heterogeneous nuclear ribonuclear protein Q1/NSAP1) is a component of mRNA granule transported with inositol 1,4,5-trisphosphate receptor type 1 mRNA in neuronal dendrites. *J Biol Chem* **279**, 53427–53434 (2004).
385. Elvira, G. *et al.* Characterization of an RNA granule from developing brain. *Mol. Cell Proteomics* **5**, 635–651 (2006).

386. Bullock, S. L., Ringel, I., Ish-Horowicz, D. & Lukavsky, P. J. A'-form RNA helices are required for cytoplasmic mRNA transport in *Drosophila*. *Nat Struct Mol Biol* **17**, 703–709 (2010).
387. Nilson, L. A. & Schüpbach, T. EGF receptor signaling in *Drosophila* oogenesis. *Curr Top Dev Biol* **44**, 203–243 (1999).
388. Hoogenraad, C. C. *et al.* Mammalian Golgi-associated Bicaudal-D2 functions in the dynein–dynactin pathway by interacting with these complexes. *EMBO J* **20**, 4041–4054 (2001).
389. Kugler, J.-M. & Lasko, P. Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during *drosophila* oogenesis. *Fly* **3**, 15–28 (2009).
390. Nashchekin, D. & St Johnston, D. Egalitarian recruitment of localized mRNAs. *Genes & Development* **23**, 1475–1480 (2009).
391. Dubowy, J. & Macdonald, P. M. Localization of mRNAs to the oocyte is common in *Drosophila* ovaries. *Mech Dev* **70**, 193–195 (1998).
392. Mische, S., Li, M., Serr, M. & Hays, T. S. Direct observation of regulated ribonucleoprotein transport across the nurse cell/oocyte boundary. *Mol Biol Cell* **18**, 2254–2263 (2007).
393. Davis, I. & Ish-Horowicz, D. Apical localization of pair-rule transcripts requires 3' sequences and limits protein diffusion in the *Drosophila* blastoderm embryo. *Cell* **67**, 927–940 (1991).

394. Meignin, C. & Davis, I. UAP56 RNA helicase is required for axis specification and cytoplasmic mRNA localization in *Drosophila*. *Dev Biol* **315**, 89–98 (2008).
395. McDermott, S. M. & Davis, I. *Drosophila* Hephaestus/polypyrimidine tract binding protein is required for dorso-ventral patterning and regulation of signalling between the germline and soma. *PLoS ONE* **8**, e69978 (2013).
396. Matunis, E. L., Matunis, M. J. & Dreyfuss, G. Characterization of the major hnRNP proteins from *Drosophila melanogaster*. *J Cell Biol* **116**, 257–269 (1992).
397. Matunis, M. J., Matunis, E. L. & Dreyfuss, G. Isolation of hnRNP complexes from *Drosophila melanogaster*. *J Cell Biol* **116**, 245–255 (1992).
398. Blanchette, M. *et al.* Genome-wide Analysis of Alternative Pre-mRNA Splicing and RNA-Binding Specificities of the *Drosophila* hnRNP A/B Family Members. *Mol Cell* **33**, 438–449 (2009).
399. Weil, T. T. *et al.* Distinguishing direct from indirect roles for bicoid mRNA localization factors. *Development* **137**, 169–176 (2010).
400. Weil, T. T., Parton, R., Davis, I. & Gavis, E. R. Changes in bicoid mRNA anchoring highlight conserved mechanisms during the oocyte-to-embryo transition. **18**, 1055–1061 (2008).
401. Reveal, B. *et al.* BREs mediate both repression and activation of oskar mRNA translation and act in trans. *Dev Cell* **18**, 496–502 (2010).

402. Lantz, V., Ambrosio, L. & Schedl, P. The *Drosophila orb* gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. **115**, 75–88 (1992).
403. Li, W., Klovstad, M. & Schüpbach, T. Repression of Gurken translation by a meiotic checkpoint in *Drosophila* oogenesis is suppressed by a reduction in the dose of eIF1A. **141**, 3910–3921 (2014).
404. Norvell, A., Wong, J., Randolph, K. & Thompson, L. Wispy and Orb cooperate in the cytoplasmic polyadenylation of localized gurken mRNA. *Dev Dyn* **244**, 1276–1285 (2015).
405. Eulalio, A., Behm-Ansmant, I., Schweizer, D. & Izaurralde, E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* **27**, 3970–3981 (2007).
406. Vershinin, M., Carter, B. C., Razafsky, D. S., King, S. J. & Gross, S. P. Multiple-motor based transport and its regulation by Tau. *PNAS* **104**, 87–92 (2007).
407. Giorgi, C. & Moore, M. J. The nuclear nurture and cytoplasmic nature of localized mRNPs. *Semin Cell Dev Biol* **18**, 186–193 (2007).
408. Steinhauer, J. & Kalderon, D. The RNA-binding protein Squid is required for the establishment of anteroposterior polarity in the *Drosophila* oocyte. *Development* **132**, 5515–5525 (2005).
409. Abdu, U., Bar, D. & Schüpbach, T. spn-F encodes a novel protein that affects oocyte patterning and bristle morphology in *Drosophila*. **133**, 1477–1484 (2006).

410. Davis, I. & Parton, R. M. Selection of appropriate imaging equipment and methodology for live cell imaging in *Drosophila*. *CSH Protoc*, pdb.ip21–pdb.ip21 (2006).
411. Derrick, C. & Weil, T. T. Translational control of *gurken* mRNA in *Drosophila* development. *Cell Cycle* **16**, 23-32 (2017).
412. Kim, C.S., Seol, S.K., Song, O.K., Park, J.H. & Jang, S.K. An RNA-binding protein hnRNP A1, and a scaffold protein, Septin 6, facilitate hepatitis C virus replication. *Journal of Virology* **81**, 3852-3865 (2007).

Appendices

Appendix A: GLS protein interactors

Protein	Function	Associated with localised RNAs
Me31B	Translational Regulation	<i>oskar, bicoid, gurken</i> , Neuronal RNA granules
IGF-II mRNA - binding protein (Imp)	mRNA localisation	<i>gurken, oskar, Vg1, β-actin</i>
Squid (Sqd) A	mRNA localisation, localised translation	<i>gurken, oskar</i>
Dodecca satellite binding protein 1 (Dp1)	Translational regulation, heterochromatin formation	<i>Hsp83</i>
Bicaudal C (BicC)	Translational Regulation	<i>gurken, oskar</i>
Thioredoxin-like protein Deadhead (Dhd)	Thioredoxin-like protein, cell redox homeostasis	
Exuperentia (Exu)	mRNA localisation	<i>bicoid, oskar</i>
CG5205	Unknown	
Rigor mortis (Rig)	Regulation of transcription, response to ecdysone	
Polyadenylate-binding protein (PABP)	Translational Regulation	<i>gurken, bicoid, Hsp83, vasopressin, BC200</i>
Ribosomal protein L14 (RpL14)	Structural constituent of the ribosome	
Syncrin (Syp)	mRNA localisation, localised translation	<i>CamKIIa, IP₃R, Arc, BC200 (Syp)</i>
Fibrillarin (Fib)	rRNA processing	
CG6745	rRNA, tRNA, snRNA processing (inferred)	
Polypyrimidine tract binding protein (PTB)	mRNA localisation, localised translation, mRNA processing	<i>oskar, Vg1, β-actin</i>
Ribosomal protein S19 (RpS19)	Structural constituent of the ribosome	

Proteins that specifically interact with the *grk* 5'ORF containing the GLS. (McDermott et al. 2012) A list of proteins shown to specifically interact with the 5'ORF region of *grk* RNA containing the GLS. RNAs highlighted in bold have been shown to be directly bound by the protein.

Appendix B: Fly strains

Name in thesis	Genotype	Description	Source
<i>OrR</i>	<i>OrR</i>	Wild type stock	
<i>grk*GFP</i>	<i>grk-MS2¹²/Cyo</i>	<i>grk</i> mRNA with 12 MS2 stem-loops, MCP-GFP	Jaramillo <i>et al.</i> 2008
<i>grk*dendra</i>	<i>grk-MS2¹²/Cyo</i>	<i>grk</i> mRNA with 12 MS2 stem-loops, MCP-dendra	Rippe Hayashi and Richard Parton, unpublished
<i>Me31B::GFP</i>	<i>Me31B::GFP</i>	Me31B protein trap GFP line	Buszczak et al 2007, CG4916
<i>Sqd::GFP</i>	<i>Sqd::GFP</i>	Sqd protein trap GFP line	Norvell et al. 2005 CG16901
<i>sqd1</i>	<i>sqd1/TM3</i>	Sqd mutant for all isoforms	Kelley 1993
<i>Df</i>	<i>Df(3R) urd</i>	Third chromosome deficiency covering Sqd	Lindsley and Zimm (1992)

Appendix C: Imaging parameters and assessing RNA localisation in the oocyte

Low magnification oocyte images were taken with a 30µm z-stack at 1µm intervals to scan through the entire egg chamber. Images were taken every 10 minutes for a total of 40 minutes. High magnification oocyte images to assess *grk* clustering and dynamics were taken with a 0.75µm z-stack at 0.25µm intervals. Egg chambers are oriented with the posterior to the right, and the dorsal to the top. All images were deconvolved using the SoftWoRx Resolve 3D constrained iterative deconvolution algorithm from Sedat/Agard (Applied Precision) in order to reassign out-of-focus light to the point of origin and improve the signal-to-noise ratio.⁴¹⁰ Images were contrasted for best display unless otherwise stated. All image analysis was carried out using SoftWoRx or FIJI (Image J).

The distribution of *grk* was assessed by measuring fluorescent intensities from a cross-section (20 pixels x 75 pixels) through the oocyte at 0 minutes, 20 minutes and 40 minutes,

generating an intensity profile with the anterior of the oocyte at the left hand side of the plot and the injection site in the middle (see Chapter 3). Transport was characterised according to the extent of RNA relocation from the injection site at the oocyte centre to the anterior at 20 minutes. No transport was characterised as a diffuse signal or aggregation throughout the oocyte at 40 minutes post-injection. DA localisation was assessed by defining several 25 pixels x 25 pixels regions (ROI): background, DA, anterior, centre (injection site), and measuring the fluorescent intensities in each region every 10 minutes for a total of 40 minutes. Localisation at the DA or anterior was characterised at 40 minutes as strong (majority of the signal from injection relocates to the DA or anterior ROI), weak (intensity level in the DA/anterior ROI is higher than background but accounts for only a fraction of the initial injected amount), none (intensity level in the DA/anterior ROI is consistently less than or equal to the centre of the oocyte over time).

grk clustering (accumulation of RNA to form large, bright puncta) was assessed at high resolution (100x 1.3 NA SI lens) with a z-stack of 5.25µm at 0.25µm intervals to facilitate proper comparison of the anterior and DA of the same oocyte which are generally at different focal planes. Fluorescent intensity line profiles were generated from ROIs at the DA and anterior of 3 pixels x 50 pixels to assess the degree of clustering and the brightness of the clusters. The size of the clusters was estimated by creating a manual outline of each cluster in a 30 pixels x 30 pixels ROI at the DA and anterior of the oocyte using Image J, and calculating the mean size of each cluster. Cluster size was compared between the DA and anterior of wild type and mutant oocytes using SEM and a Student's unpaired, two-tailed t-test.

grk particle dynamics were also assessed at high resolution (100x 1.3 NA SI lens) with time-lapse imaging. Images were acquired every 300ms for 10 seconds and were processed by equalising the average intensity across the time points to reduce the effects of bleaching before, removing the background signal to increase the contrast, and projecting the images over time. The projected time-lapse image was then merged with the time zero image as red

and green colour channels, respectively, to distinguish moving *grk* clusters (red and green) from static clusters (yellow). A ROI (50 pixels x 50 pixels) was chosen at the DA and anterior of the oocyte to compare *grk* dynamics in these regions.

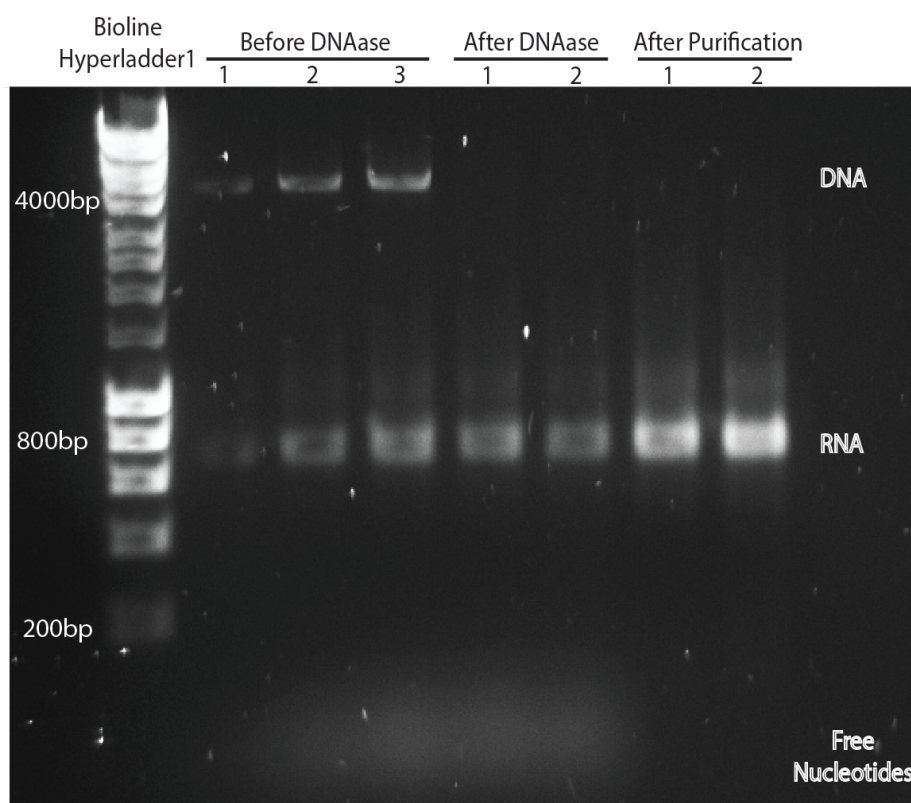
Appendix D: Imaging parameters and assessing RNA localisation in the embryo

All images were deconvolved (see Appendix C) and average intensity projected to improve the signal. Low resolution embryo images looking for RNA localisation and recruitment of *trans*-acting factors were taken with a 20x 0.75NA dry objective and a 1.05µm z-stack at 0.35µm intervals. Images were taken every 5 minutes following injection for a total of 20 minutes. High resolution embryo images investigating colocalisation of RNA and *trans*-acting factors at the apical cortex were taken with a 100x 1.3 NA SI lens and a 0.75µm z-stack at 0.25µm intervals. High resolution embryo images tracking *grk* particles and Me31B::GFP were taken with a 100x 1.3 NA SI lens with a 0.75µm z-stack at 0.25µm intervals and a time-lapse of 500ms for a total of 1 minute. Time points were equalised to reduce the effects of bleaching. All embryos were oriented with the apical cortex to the top unless otherwise stated.

Localisation of RNA and recruitment of GFP-tagged *trans*-acting factors was assessed after 20 minutes unless otherwise stated. Fluorescent intensities were taken from a cross-section (20 pixels x 200 pixels) through the embryo, generating an intensity profile with the apical cortex at the right hand side of the plot and the injection site in the middle (see Chapter 4). RNA localisation was categorised as strong (most of the RNA enriched at the apical cortex), weak (apical enrichment but a large proportion remains basally) or unlocalised (no apical enrichment of the RNA). Protein recruitment was categorised as strong (highly enriched above basal signal), weak (appears enriched at the apical cortex but the signal is not above basal signal) or none (does not appear enriched at the apical cortex and signal is much lower than the basal signal).

Appendix E: Assessing RNA quality following *in vitro* transcription

The quality and integrity of *in vitro* transcribed, fluorescently labelled RNA was assessed using agarose gel electrophoresis at several stages throughout the protocol including; immediately following transcription, after removal of the DNA template using DNase, after removal of unincorporated nucleotides using a G-50 Sephadex column (Roche). An example 1% agarose gel can be seen below for wild type *grk* transcripts.



Assessing the quality and integrity of *in vitro* transcribed *grk* RNA. Agarose gel electrophoresis of samples taken from different stages of the *in vitro* transcription reaction. Immediately following transcription the DNA template is removed using DNase, leaving the RNA and small unincorporated nucleotides. These nucleotides are then removed, leaving the RNA only. Precipitation of the RNA is later carried out and the RNA is further cleaned to remove salts.