Regulation of mRNA stability by the RNA-Binding Protein Pumilio during Early *Drosophila* Embryogenesis

by

Alexander John Marsolais

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Biochemistry
University of Toronto

© Copyright by Alexander John Marsolais 2015
Regulation of mRNA stability by the RNA-Binding Protein Pumilio during Early Drosophila Embryogenesis

Alexander John Marsolais
Doctor of Philosophy
Department of Biochemistry
University of Toronto
2015

Abstract

The maternal-to-zygotic transition (MZT) is a characteristic phase of early metazoan development where control of embryogenesis transitions from products encoded by the mother to those encoded by the zygotic genome. Post-transcriptional regulation (PTR) plays a critical role in the MZT, particularly in the clearance of maternal mRNAs. Mechanisms of maternal transcript decay that rely exclusively on maternal protein factors and function early during the MZT, as well as mechanisms that require zygotic factors and function later during the MZT, have been characterized.

The Drosophila embryo has long-served as a model for the MZT. The RNA-binding protein (RBP) SMAUG (SMG) has been shown to function during the early (maternal) phase of degradation. In contrast, computational methods suggest the RBP PUMILIO (PUM) functions in the late (zygotic) phase of maternal mRNA degradation. Such a role is curious as PUM is maternally-contributed and functional during the maternal (early) phase of embryogenesis. I show here that: 1) PUM is required for the degradation of approximately 500 maternal mRNAs during the late (zygotic) wave of degradation; 2) degradation of PUM target mRNAs is likely delayed to the late (zygotic) phase due to the presence of sub-optimal PUM binding sites within
these target mRNAs, 3) degradation of PUM targets is dependent on additional factors such as the RBP BRAIN TUMOUR (BRAT) and a core component of the RNAi machinery, ARGONAUTE 1 (AGO1); and 4) a critical function of PUM appears to be clearance of smg mRNA, since in pum mutant embryos SMG protein persists post-MZT and is associated with an inappropriate down-regulation of SMG target transcripts. Taken together, these data support a multi-factorial view of RBP function, in which the activity of a given RBP is determined by other RBPs associated with a particular mRNA.
Acknowledgments

I would like to thank all the members of the Smibert and Lipshitz laboratories for reagents, advice and support over the years, in particular Mariana Kekis, Xiao Li, John Laver, Angelo Karaiskakis and Claudia Walser.

Howard Lipshitz has functioned as a second, unrecognized supervisor over the years of this project, and has provided a great deal of mentorship on many aspects of this project, large and small.

I would like to thank my committee, Angus McQuibban and Alex Palazzo, for their support and feedback over many committee meetings, departmental seminars and informal presentations.

Craig Smibert has been invaluable as a supervisor. I would like to thank him for countless hours of advice and enlightening discussions, and for fostering an atmosphere of collegiality and enthusiasm for the scientific process in the lab that was instrumental in overcoming the many challenges of a doctoral research project.

Finally, the support of my family, particularly my wife Erica, has been critical over the years. I would like to thank her for her unflagging encouragement over the years; every grad student needs someone like Erica in their corner.
# Table of Contents

Acknowledgments .......................................................................................................................... iv  
Table of Contents ............................................................................................................................ v  
List of Abbreviations Used ............................................................................................................ ix  
List of Tables ......................................................................................................................................... xi  
List of Figures ....................................................................................................................................... xii  

1 Introduction ........................................................................................................................................ 1  
   1.1 Post-transcriptional regulation .................................................................................................. 1  
      1.1.1 Transcript localization ........................................................................................................ 2  
      1.1.2 Translational regulation ................................................................................................... 5  
      1.1.3 Transcript degradation ...................................................................................................... 12  
      1.1.4 Multifactorial regulation of mRNA targets by RBPs and the pervasiveness of PTR .......... 18  
   1.2 The maternal-to-zygotic transition in *Drosophila* embryogenesis ............................................ 20  
      1.2.1 Oogenesis ....................................................................................................................... 20  
      1.2.2 Egg activation .................................................................................................................. 22  
      1.2.3 The maternal phase of embryogenesis .......................................................................... 22  
      1.2.4 Early or maternal mRNA degradation .......................................................................... 23  
      1.2.5 The zygotic phase of embryogenesis .............................................................................. 24  
      1.2.6 Late or zygotic mRNA degradation .................................................................................. 24  
      1.2.7 Unfertilized eggs .............................................................................................................. 27  
      1.2.8 miRNAs and late (zygotic) mRNA degradation ............................................................... 27  
      1.2.9 PUMILIO (PUM), BRAIN TUMOUR (BRAT) and other factors involved in late (zygotic) mRNA degradation .......................................................... 28  
   1.3 Pumilio-Fem3 binding factor (Puf) proteins .............................................................................. 29
1.3.1 The biological roles of Puf proteins ................................................................. 29
1.3.2 The binding specificity of Puf proteins ............................................................... 32
1.3.3 The structure of Puf proteins ............................................................................. 34
1.3.4 Mechanisms of Puf protein regulation ............................................................... 34
1.3.5 Puf proteins and the RNAi machinery ............................................................... 37
1.3.6 The complexity of PUM-mediated regulation during *Drosophila* embryogenesis ................................................................. 39

1.4 Thesis outline ...................................................................................................... 44

2 Materials and Methods .......................................................................................... 45

2.1 Fly stocks ............................................................................................................ 45
2.2 Transgenic reporters ............................................................................................ 45
2.3 Microarrays ......................................................................................................... 47
2.4 Motif enrichment analysis .................................................................................. 48
2.5 GO term analysis ................................................................................................ 49
2.6 RT-qPCR primers for *smg* and *RpL32*: ......................................................... 50
2.7 Northern and Western blots .............................................................................. 50
2.8 Purification of GST-PUM RBD .......................................................................... 52
2.9 Gel shifts ............................................................................................................ 52
2.10 Statistical analysis ............................................................................................. 53

3 Results .................................................................................................................... 54

3.1 Attributions ......................................................................................................... 54
3.2 PUM is a major regulator of mRNA stability in the early *Drosophila* embryo ........ 54
3.3 PUM negatively regulates mRNA stability through direct binding ................. 59
3.4 PUM downregulates mRNAs encoding transcriptional and post-transcriptional regulators, as well as the V-ATPase and proteasome complexes .................. 59
3.5 SMG target mRNAs are down-regulated in *pum* mutant embryos ............... 62
3.6 PUM regulates SMG expression ....................................................................... 65
3.7 PUM directly regulates smg mRNA through PBEs in the smg 3' UTR ................................. 66
3.8 PUM can induce transcript decay through an early pathway ........................................... 71
3.9 Mapping additional cis-elements required for smg mRNA decay ........................................ 76
3.10 smg mRNA decay requires BRAT .................................................................................. 76
3.11 smg mRNA decay requires AGO1 .................................................................................. 78
3.12 PUM, BRAT and AGO1 cooperate to regulate other targets in the Drosophila embryo. 80

4 Discussion and future directions .......................................................................................... 87
4.1 Discussion ......................................................................................................................... 87
  4.1.1 PUM is a critical regulator of the post-transcriptional landscape of the early
      Drosophila embryo .......................................................................................................... 87
  4.1.2 smg mRNA as a critical target of PUM ........................................................................ 87
  4.1.3 Functional PUM protein is contributed maternally to the embryo ......................... 89
  4.1.4 PUM-mediated zygotic transcript degradation depends on other co-factors .......... 90
  4.1.5 PUM, BRAT and AGO1 regulate similar transcripts in the embryo ..................... 92
  4.1.6 SMG target transcripts are not down-regulated in ago1 knockdown embryos .... 93
  4.1.7 A PUM-BRAT-AGO1 complex as a model for smg mRNA degradation .......... 94
  4.1.8 The CCR4-POP2-NOT deadenylase complex ....................................................... 97
  4.1.9 Multi-factorial post-transcriptional regulation in the early Drosophila embryo .. 98
4.2 Future directions ............................................................................................................... 98
  4.2.1 BRAT ..................................................................................................................... 98
  4.2.2 AGO1 ..................................................................................................................... 99
  4.2.3 A PUM-BRAT-AGO1 complex .......................................................................... 100
  4.2.4 Zygotic miRNA expression and timing of smg mRNA degradation .................. 102
  4.2.5 Constructing a reporter mRNA that is degraded with late kinetics ................. 105
  4.2.6 The downstream mechanisms of PUM-mediated degradation of smg mRNA... 106
  4.2.7 The biological significance of PUM-mediated mRNA degradation ................. 108
List of Abbreviations Used

4E-BP – eIF4E binding protein
AUBP – AU-rich binding protein
AGO1 – Argonaute 1
BCD – Bicoid
BRAT – Brain Tumour
GSC – germline stem cell
hb – hunchback
IRE – iron response element
miRNA – micro ribonucleic acid
mRNA – messenger ribonucleic acid
mRNP – messenger ribonucleoprotein particles
MBT – mid-blastula transition
MZT – maternal-to-zygotic transition
nt – nucleotide
ORF – open reading frame
PABP – poly(A) binding protein
PAR-CLIP - photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation
PBE – PUM binding element
PGC – primordial germ cell
PNG – PAN GU
PTR – post-transcriptional regulation
Puf - Pumilio-Fem3 mRNA binding factor
PUM - PUMILIO
Puf HD - Pumilio-Fem3 mRNA binding factor homology domain
RIP – RNA immunoprecipitation
RISC – RNA-induced silencing complex
RNA – ribonucleic acid
RNAi – RNA interference
RBD – RNA binding domain
RBP – RNA binding protein
SAM – significance of microarrays
shRNA – short hairpin RNA
siRNA – silencing ribonucleic acid
SMG – smaug
UE – unfertilized egg
ZGA – zygotic genome activation
List of Tables

Table 1 ........................................................................................................................................61
List of Figures

Figure 1. The ‘closed loop’ model of translation.................................................................7
Figure 2. Pathway of mRNA degradation..............................................................................13
Figure 3. Post-transcriptional regulation during early *Drosophila* embryogenesis..................25
Figure 4. PUM regulates hundreds of transcripts after the onset of zygotic transcription........55
Figure 5. PUM destabilizes hundreds of transcripts after the onset of zygotic transcription
through direct binding........................................................................................................57
Figure 6. In *pum* mutant embryos, SMG protein persists after the onset of zygotic transcription
and down-regulated transcripts are enriched for SMG target transcripts..............................63
Figure 7. *smg* mRNA decay requires the *smg* 3’ UTR.........................................................67
Figure 8. PUM directly regulates *smg* mRNA through PBEs located in its 3’ UTR..............69
Figure 9. High affinity PBEs mediate degradation through a maternal pathway....................72
Figure 10. mRNA destabilization mediated by the *smg* 3’ UTR requires PBEs and additional cis-
elements in nucleotides 1-438.........................................................................................77
Figure 11. *smg* degradation is dependent on AGO1 and BRAT...........................................79
Figure 12. Several hundred transcripts are up-regulated in 4-5 hour old *ago1* knockdown
embryos.................................................................................................................................82
Figure 13. PUM, BRAT and AGO1 cooperate to degrade transcripts during the late (zygotic
phase of decay ....................................................................................................................83
Figure 14. Proposed model of regulation by PUM, BRAT and AGO1 during the late (zygotic)
phase of mRNA decay.......................................................................................................96
1 Introduction

1.1 Post-transcriptional regulation

Post-transcriptional regulation (PTR) refers to various modes of gene regulation that occur at the level of the messenger RNA (mRNA) transcript, and includes processes such as cytoplasmic mRNA localization, translational regulation and mRNA degradation. Post-transcriptional regulatory mechanisms play a prominent role in controlling gene expression, and are critical to all cells. PTR has important roles in all major cell biological processes, including growth and development, signal transduction, cell division and responses to stress. Sometimes PTR is global in nature, with changes occurring to most mRNAs in the cell simultaneously in a sequence-independent manner (e.g., large-scale translational down-regulation during cellular stress (Shalgi R, et al. 2013)). In many cases, however, PTR is transcript-specific and mRNAs are recognized in a sequence-specific manner and targeted for a certain form of PTR. The sequence-specific regulation of transcripts is often accomplished through the action of RNA-binding proteins (RBPs) that recognize binding sites within target mRNAs. Sequence-specific regulation can also be achieved by non-coding RNAs. For example, microRNAs (miRNAs) and small interfering RNAs (siRNAs) can recognize mRNAs through base pairing with complementary sequences within the target. Collectively RBPs, miRNAs and siRNAs are referred to as ‘trans-acting factors’, and the sequence elements they recognize in target transcripts are referred to as ‘cis-acting elements’. While some trans-factors can directly regulate their targets, many achieve regulation through recruitment of other factors to the mRNA.

PTR refers to all forms of regulation that occur at the level of the mRNA and includes co-transcriptional and other nuclear mechanisms, such as splicing, nuclear polyadenylation, 5' capping, and nuclear export. While critical to the proper expression and regulation of eukaryotic
protein-coding genes, these nuclear forms of regulation will not be further discussed here. Instead, this survey will focus on the three main classes of cytoplasmic PTR: 1) transcript localization, 2) translational regulation, and 3) transcript degradation. In addition, several key concepts in PTR will be discussed, including the concept of the combinatorial regulation of mRNA target transcripts by multiple RBPs.

1.1.1 Transcript localization

After export from the nucleus, many transcripts are localized within the cytoplasm (Meignin C and Davis I 2010). Localization of transcripts is an important mechanism employed by the cell to restrict the gene product of an mRNA to a specific location within the cell, and is particularly important in large and/or highly polarized cell types, such as embryos and neurons. Most examples that have been described to date occur in eukaryotes, although some data exist which suggest that transcript localization occurs in prokaryotes as well (Nevo-Dinur K, et al. 2011). Well-known examples of mRNA localization include that of ASH1 mRNA which is transported to the daughter cell in Saccharomyces cerevisiae (Beach DL and Bloom K 2001), β-actin mRNA localization on the leading edge of fibroblasts (Latham VM Jr, et al. 1994; Lawrence JB and Singer RH 1986) and various transcripts (e.g., gurken, nanos, oskar and bicoid) localized along either the anterior-posterior or dorsal-ventral axes of the Drosophila oocyte and embryo (Martin KC and Ephrussi A 2009; Neuman-Silberberg FS and Schüpbach T 1993). In fact, a genome-wide survey of transcript localization in the Drosophila embryo found that as many as 70% of transcripts are localized (Lécuyer E, et al. 2007), and evidence suggests that the phenomenon is also widespread in vertebrate cells (Holt CE and Bullock SL 2009). Localization of transcripts can be achieved through one of three mechanisms; 1) active transport, 2) passive capture, or 3) selective degradation.
1.1.1.1 She2/3p and BICD/EGL: active transport of mRNA

In the active transport mechanism, mRNAs are specifically recognized by RBPs and are then associated with molecular motors that transport the mRNA along either microfilaments or microtubules. Examples of active transport include the transfer of \textit{ASH1} mRNA to the daughter cell in \textit{S. cerevisiae} (Jansen RP and Niessing D 2012) and \textit{CAMKII\textalpha} mRNA to the dendrites of mammalian neurons (Hirokawa N 2006; Mayford M, et al. 1996). The \textit{ASH1} mRNA is one of the better-known examples of an actively transported mRNA. The PTR of \textit{ASH1} begins in the nucleus with the binding of the RBP She2p to the so-called ‘zip code elements’ – \textit{cis}-acting elements located in the 3’ UTR of \textit{ASH1} mRNA required for its correct localization. In addition to She2p, several other factors associate with \textit{ASH1} mRNA in the nucleus – some of these factors only transiently bind to the transcript in the nucleus, while others remain associated with the mRNA after it is exported from the nucleus (such as Puf6p and Khd1p, which serve to translationally repress \textit{ASH1} mRNA during its transport). Once exported to the cytoplasm, She3p directly interacts with \textit{ASH1} mRNA, as well as She2p. She3p is in turn associated with the myosin motor, Myo4p. Once formed, this \textit{ASH1} messenger ribonucleoprotein particle (mRNP) is then actively transported to the daughter cell along actin filaments.

An additional example is the transcript localization carried out by EGL and BICD in \textit{Drosophila} oocytes and embryos. Recent data indicate that EGL is a non-canonical RBP which recognizes stem-loop structures in certain localized transcripts. EGL interacts with BICD, and both factors make contacts with the dynein/dynactin motor complex, allowing for minus-end transport of bound mRNAs along the microtubule network (Dienstbier M, et al. 2009).
1.1.1.2 nanos transcript: passive capture of mRNA

Passive capture occurs when an mRNA associates with a localized anchoring factor. Localization of the transcript occurs because once bound the molecule no longer diffuses or its diffusion is greatly reduced, leading to its concentration in the area in which it is anchored. One of the better known examples of this form of mRNA localization occurs in the *Drosophila* oocyte, where the nanos (nos) mRNA becomes anchored at the posterior of the developing oocyte (Forrest KM and Gavis ER 2003). The anchoring of nos mRNA to the posterior pole requires the presence of a complex of different cis-acting elements present in the nos 3' UTR (Gavis ER, et al. 1996). These sequence elements are in turn believed to be bound by several different trans-acting factors at the posterior pole, such as Rumpelstiltskin (RUMP) (Jain RA and Gavis ER 2008) and Aubergine (AUB) (Becalska AN, et al. 2011). This mRNP containing nos mRNA is anchored at the posterior through the action of additional factors, such as Hsp90 (Song Y, et al. 2007), although the exact mechanism remains to be determined.

1.1.1.3 Hsp83 and SMAUG: selective degradation of mRNA

The third form of transcript localization is selective degradation. In this mechanism, an mRNA is degraded in one part of a cell but not another, leading to its concentration in the area in which it is protected from decay. There are several examples of this phenomenon in early *Drosophila* development, such as the transcript Hsp83, which is degraded in the bulk of the embryo by the RBP SMAUG (SMG) through recognition of so-called SMG recognition elements (SREs) in the Hsp83 open reading frame (ORF) (Semotok JL, et al. 2008). Localization of Hsp83 occurs because, while the transcript is degraded in the bulk cytoplasm, it is protected at the posterior of the embryo in a specialized cytoplasmic region known as the germ plasm (Ding D, et al. 1993).
It is important to note that mechanisms of transcript localization are not mutually exclusive, and combinations (e.g., both active transport and passive capture) could be employed to regulate a single transcript.

1.1.2 Translational regulation

Another common set of PTR mechanisms control the rate of translation of mRNAs. Recent studies have measured the pervasiveness of this form of regulation in the cell by correlating both mRNA and protein expression levels using next-generation sequencing and quantitative mass spectrometry. In this analysis, a high correlation between mRNA and protein levels would indicate little reliance on translational regulation to set protein levels, i.e., if mRNA levels are similar to those of the proteins they encode, then gene expression is mostly determined by mRNA level (itself an equilibrium between the rate of transcription and the rate of mRNA degradation). In general, however, these studies have found a poor correlation between mRNA and protein levels, indicating that the cell makes extensive use of translational regulation to control gene expression (de Sousa Abreu R, et al. 2009; Vogel C and Marcotte EM 2012).

Although some forms of translational regulation are global in nature (e.g., the phosphorylation of eEF2A by various signaling pathways, which inhibits ribosome translocation and therefore decreases total protein synthesis (Celis JE, et al. 1990; Pyronnet S and Sonenberg N 2001)), many are transcript-specific.

1.1.2.1 The ‘closed loop’ model of translation

Examples of the translational regulation of mRNAs are numerous, and all target specific steps in the so-called ‘closed loop’ model of eukaryotic translation (Preiss T and Hentze MW
1999) (Fig. 1). In this model, the poly(A) tail of the mRNA, located at the 3' end of the transcript, is required along with the 5' cap structure for efficient initiation of translation. The cap structure is required because it is bound by the initiation factor eIF4E, a component of the eIF4F translation initiation complex, which is involved in recruitment of the 40S ribosomal subunit to the mRNA. The poly(A) tail is required because it is bound by poly(A) binding protein (PABP), which in turn associates with the initiation factor eIF4G (also a component of the initiation complex eIF4F) (Tarun SZ Jr and Sachs AB 1996). These interactions result in the formation of a eIF4E-eIF4G-PABP complex (Wells SE, et al. 1998), stabilizing the interaction of eIF4G with the mRNA. This enhances translation, as eIF4G recruits eIF3 to the mRNA, which in turn recruits the 40S ribosomal subunit to the 5' end of the transcript. The interaction of PABP with eIF4G implies that the 5' and 3' ends of an actively translated mRNA are in close proximity (hence the ‘closed loop’ model), and indeed electron micrographs of actively translated transcripts suggest this is the case (Madin K, et al. 2004). Once the 40S ribosome is recruited, it ‘scans’ the transcript in the 5'-to-3' direction until the start codon is recognized. At this stage, the 60S ribosome is recruited (and the 80S ribosome is formed), and translational elongation begins. Translation ends when the transiting ribosome recognizes a stop codon, leading to ribosome disassembly and polypeptide release.

It should be noted that this is a simplified model of eukaryotic translation, and that a large number of additional accessory factors are critical for the correct completion of translational initiation, elongation and termination. In addition, other modes of translation initiation have also been documented. One well-known example is that of internal ribosome entry sites (IRESs). While originally described in viral transcripts, many examples of IRESs in eukaryotic transcripts are now known (Hellen CU and Sarnow P 2001), and IRESs are now believed to play critical
Figure 1. The ‘closed loop’ model of translation. Simplified model of eukaryotic translation, showing the topography of the transcript (including 5’ cap structure and 3’ poly(A) tail), select initiation factors (blue), PABP (yellow) and ribosomal subunits (green). Actively translated transcripts are believed to adopt a loop conformation, wherein the 5’ and 3’ ends are in proximity to one another. eIF4E (a member of the eIF4F complex) directly binds to the 5’ cap, while PABP interacts with the poly(A) tail. Both make contact with another eIF4F member, eIF4G, stabilizing the association of the initiation complex with the transcript. The initiation complex in turn recruits the 40S ribosome, which then begins to scan the transcript for the start codon. Upon reaching the start codon, the 60S ribosome is recruited, forming the functional 80S ribosome. The 80S ribosome then translates the ORF and disassembles when it reaches the stop codon (not depicted). Mechanisms of translational regulation can occur at any stage of this process, although most occur during the early steps of initiation, in particular at the level of the eIF4E-eIF4G-PABP complex.
roles in initiating the translation of many cellular mRNAs, including key transcripts that are translated under stress conditions when cap-dependent translation is down-regulated (Komar AA and Hatzoglou M 2011).

1.1.2.2 Iron response elements and iron response element binding proteins

While in principle regulation of translation could occur at any of the above steps, many documented forms of translational regulation occur at the level of the eIF4F complex members eIF4E and eIF4G (Sonenberg N and Hinnebusch AG 2009). This may be because the 5'-methyl cap-dependent recruitment of the 40S ribosomal subunit by initiation factors eIF4E and eIF4G is believed to be the principle rate-limiting step in translation, and hence affords an efficient target for translational regulation. Note, however, that there are examples of translational regulation that occur at other phases of translation. A classic example of translational repression that occurs during translational initiation is dictated by the presence of so-called iron response elements (IREs) in the 5' UTRs of certain transcripts involved in iron metabolism (Muckenthaler MU, et al. 2008). These IREs are stem loop structures that in their unbound state permit 40S ribosome scanning to proceed. However, under conditions of low cellular levels of iron, IRE-binding proteins (IRPs) bind to these IREs and reduce the rate of translation of the transcript. One of the better characterized examples of a transcript containing such a 5' IRE is the ferritin transcript, whose gene product is not needed unless cellular concentrations of iron are high.

The regulation of transcripts such as the ferritin mRNA by 5' IREs and IRPs is quite complex, with multiple levels of regulation. Briefly, the presence of IREs in the 5' UTR of certain targets is required for both translational repression in the absence of iron, and translational up-regulation in the presence of iron. Under conditions of low iron, the IRE is
bound by IRP, which interferes with recruitment of eIF4F (the complex containing the eIF4G scaffolding protein) and prevents translation. However, under condition of high iron, Fe^{2+} interacts with both IRP (altering its binding specificity, and resulting in its dissociation from the transcript) as well as the IRE itself. The association of Fe^{2+} with the IRE alters its secondary structure, increasing the affinity of eIF4F for the 5' UTR of the transcript, in turn increasing the rate of translation (Ma J, et al. 2012).

1.1.2.3 eIF4E binding proteins and 4E-HP

Translational regulation at the level of initiation is often exerted through modulation of the eIF4E-eIF4G-PABP complex (e.g., by blocking the binding of eIF4E to the cap or interfering with the interactions of any of the members of the complex). For example, eIF4E binding proteins – 4E-BPs have well-characterized roles in repressing translation through binding to eIF4E and blocking its interaction with eIF4G (Igreja C, et al. 2014).

Cap-dependent translation is also negatively regulated by 4E-HP, a protein related to eIF4E, which interacts with the cap structure but cannot bind eIF4G. 4E-HP interacts with the RBP BICOID (BCD), and cooperates with BCD to repress the caudal (cad) mRNA in the Drosophila embryo. Recruitment of 4E-HP to the cad transcript by BCD results in competition between eIF4E and 4E-HP for cap binding, resulting in a loss of eIF4E recruitment to the transcript and translational repression (Cho PF, et al. 2005).

1.1.2.4 Deadenylation and cytoplasmic polyadenylation

Upstream of the eIF4E-eIF4G-PABP complex, the poly(A) tail is another common target. Many repressive RBPs recruit deadenylases to reduce or remove the poly(A) tail of a target
transcript, displacing PABP from the poly(A) tail and leading to a destabilization of the translation initiation complex. GW182/TNRC6 (which is recruited to target transcripts by Argonaute proteins) is also able to recruit deadenylases to transcripts in order to translationally repress them (Braun JE, et al. 2013).

While many forms of translational regulation are negative, transcripts can also be translationally up-regulated. For example, the poly(A) tail of an mRNA can be elongated in the cytoplasm, increasing the efficiency of PABP recruitment, thereby resulting in higher rates of translation. This mechanism is well known in the context of the so-called cytoplasmic ‘unmasking’ of transcripts during egg activation in many eukaryotes (Radford HE, et al. 2008), although it is now known to occur in other cell types as well (Charlesworth A, et al. 2013).

Interestingly, several of the factors involved in this ‘unmasking’ also play roles in translationally repressing the same target transcripts prior to egg activation. In the oocytes of several species, including *Xenopus*, transcripts containing so-called cytoplasmic polyadenylation elements (CPEs) in their 3' UTRs (such as *cyclin B*) are translationally repressed before oocyte maturation. This repression is accomplished through the trans-acting factors CPEB and Maskin (Mendez R and Richter JD 2001; Norbury CJ 2013). CPEB is an RBP that directly recognizes and binds to CPEs in a target transcript, while Maskin is a 4E-BP that interacts with eIF4E, blocking eIF4G recruitment and therefore down-regulating the translation of the target transcript. In addition, CPEB (along with the scaffolding protein symplekin) interacts with the PARN deadenylase. Recruitment of PARN serves to keep the repressed transcript in a deadenylated state, further contributing to its translational repression. Interestingly, the polyadenylase PAP/GLD-2 is also recruited to the transcript by CPEB. PAP/GLD-2 is constitutively active, however its ability to polyadenylate the transcript is counteracted by PARN, keeping the mRNA in a deadenylated state. Upon oocyte maturation, CPEB is phosphorylated, leading to the
rearrangement of the complex and the ejection of PARN. The departure of PARN results in a lengthening of the poly(A) tail by PAP/GLD-2 (which remains associated with the transcript), in turn leading to the recruitment of PABP. Maskin remains in the complex, but competition between Maskin and PABP for eIF4E binding is believed to shift the equilibrium towards eIF4G recruitment and translational initiation.

1.1.2.5 hnRNPK/E1 and downstream translational repression

Translational repression can also occur even further downstream during translation. One well-known example is the repression of lipoxygenase (LOX) mRNA by hnRNPK/E1 that occurs in undifferentiated reticulocytes. The transcript-specificity of LOX mRNA repression is governed by the presence of a cis-acting 'differentiation control element' in its 3' UTR (Ostareck-Lederer A, et al. 1994). This sequence element is recognized by the hnRNP proteins K and E1. Recruitment of hnRNPK/E1 to LOX mRNA does not affect 40S ribosomal subunit recruitment, but instead acts to repress 60S ribosomal subunit joining, which normally occurs after the 40S subunit has recognized the translation start codon. Since this mechanism lies downstream of the cap-binding activity of eIF4E, hnRNPK/E1 can repress both cap-dependent as well as IRES-driven transcripts in vitro (Ostareck DH, et al. 1997).

It is also important to note that mechanisms that regulate the translation of an mRNA are often coupled to mechanisms that control transcript stability. As will be discussed below, removal of the poly(A) tail from an mRNA often leads to its destabilization (Chen CY and Shyu AB 2011) as well as its translational repression. In addition, removal of the 5'-methyl cap from a transcript (a common downstream event in transcript degradation following deadenylation), would also lead to efficient translational repression due to the disruption of the eIF4E binding. These concepts will be further examined in the following section.
1.1.3 Transcript degradation

All mRNAs are ultimately turned over by the cell at the end of their life cycle. Not only will transcript degradation terminate protein production but, as discussed above, selective transcript degradation in specific subcellular locales can also function as a mechanism of transcript localization. Studies that have examined mRNA degradation at a global level have found a wide variation in transcript half-lives, from as little as a few minutes to many hours (Sharova LV, et al. 2009; Yang E, et al. 2003), suggesting that regulation of mRNA degradation is widespread in the cell.

1.1.3.1 Mechanisms of transcript decay

The mechanisms of mRNA turnover have been best studied in yeast and mammals, with several major pathways of transcript degradation described. The process of mRNA degradation is dependent on the recruitment of various nucleases to disassemble the transcript, and typically proceeds in a step-wise fashion with various protective elements of the mRNA being removed, before the eventual destruction of the transcript (Meyer S, et al. 2004; Parker R and Song H 2004). The two principle protective elements of an mRNA are the 5' cap structure and the 3' poly(A) tail. As discussed in section 1.1.2.1, the cap and poly(A) tail are also critical for efficient translation. Hence, mechanisms of translational repression and transcript decay are often intertwined, with the process of deadenylation a frequent component of both forms of regulation.

Generally speaking, transcript decay can begin with different rate-limiting steps: 1) deadenylation, 2) endonucleolytic cleavage, or 3) deadenylation-independent decapping (Fig. 2). All three pathways render the transcript vulnerable to rapid 5'-to-3' and/or 3'-to-5'
Figure 2. Pathways of mRNA degradation. Simplified model of a eukaryotic transcript and the major pathways of mRNA degradation. The transcript is protected at either end by the 5' cap structure and 3' poly(A) tail, structures which are also required for its efficient translation. Degradation of the transcript requires modification or removal of these protective elements. The most common mechanism begins with deadenylation, a reduction in the length of the poly(A) tail. This disrupts the ‘closed loop’ conformation of the transcript (see Figure 1), in turn rendering the 5' cap vulnerable to decapping enzymes. Removal of the 5' cap in turn leads to 5'-to-3' exonucleolytic decay. Alternatively, transcripts can be degraded through a direct decapping mechanism, wherein decapping enzymes act directly on the transcript without the need for deadenylation. This in turn leads to 5'-to-3' exonucleolytic decay. Finally, certain mechanisms of degradation rely on endonucleolytic cleavage of the transcript. This renders the mRNA vulnerable to both 3'-to-5' decay (towards the 5' cap), as well as 5'-to-3' decay (towards the 3' poly(A) tail).
exonucleolytic degradation, functions generally carried out by Xrn1/2 (5'-to-3') (Jones CI, et al. 2012) and/or the exosome (3'-to-5') (Houseley J, et al. 2006).

1.1.3.2 Deadenylation-dependent degradation

Deadenylation is believed to be the rate-limiting step in mRNA decay in most cases. The major deadenylase activity in eukaryotic cells is supplied by the Ccr4-Pop2-Not complex, in which the Ccr4 subunit represents the major deadenylase subunit, with deadenylase activity also attributable to Pop2 (Wahle E and Winkler GS 2013). In addition, in some species other deadenylases (such as PARN (Virtanen A, et al. 2013) and PAN2/PAN3 (Wahle E and Winkler GS 2013)) contribute to poly(A) tail removal. While critical for mRNA turnover, deadenylases such as the Ccr4-Pop2-Not complex do not recognize targets directly, but instead are recruited to transcripts through interactions with trans-factors which specifically recognize the mRNA.

Once deadenylated, transcripts can be degraded through a number of different pathways (Meyer S, et al. 2004; Parker R and Song H 2004). Without a poly(A) tail, PABP is displaced from the mRNA, leading to loss of the ‘closed-loop’ conformation as well as the cap-associated translational initiation factors such as eIF4E. This renders the 5' methyl cap of the mRNA vulnerable, and decapping factors such as Dcp1/2 can then act to remove the cap. Once the 5' cap is cleaved off, the mRNA is susceptible to 5'-to-3' exonucleolytic decay, through the activity of enzymes such as Xrn1/2. The deadenylation-dependent decapping followed by 5'-to-3' decay pathway is the principle mechanism by which mRNAs are turned over in most cell types. In addition, multiple other pathways play a role in transcript decay. For example, deadenylated transcripts can be degraded in a 3'-to-5' direction through the cytoplasmic action of enzyme complexes such as the RNA exosome, which also functions in nuclear rRNA trimming (Houseley J, et al. 2006).
1.1.3.3 UNR, AUBPs and the deadenylase complex

One of the first RBPs shown to recruit the Ccr4-Pop2-Not complex was UNR. UNR, which specifically recognizes cis-elements referred to as mCRDs, forms a complex with PABP and the Ccr4-Pop2-Not complex, promoting the deadenylation and degradation of its targets (Chang TC, et al. 2004). Another classic example of specific degradation being accomplished through promotion of deadenylation is directed by so-called AU-rich elements. AU-rich elements are well-characterized cis-elements typically found in the 3' UTR of transcripts that confer instability to the mRNA. They are bound by a class of proteins known as AU-rich binding proteins (AUBPs). The best-characterized AUBP is AUF1 (Gratacós FM and Brewer G 2010), which promotes deadenylation of transcripts through recruitment of deadenylases. The protein tristetraprolin (TPP) also regulates AU-rich element containing transcripts. TPP functions through numerous mechanisms, including recruitment of deadenylases, decapping enzymes and exonucleases, and stimulates both 5'-to-3' and 3'-to-5' exonucleolytic decay to eliminate its targets (Ciais D, et al. 2013). As will be discussed below, many other RBPs have subsequently been shown to recruit deadenylases to target transcripts in order to destabilize them.

1.1.3.4 Nonsense-mediated decay and direct decapping-dependent degradation

In certain cases, a transcript can be directly decapped, without first being deadenylated. Although there are several such mechanisms described in the literature, one of the better-known examples involves so-called nonsense-mediated decay (NMD). This form of transcript decay is typically triggered by the presence of nonsense mutations in the ORF of a given transcript. NMD bypasses the normal requirement for deadenylation and instead results in direct decapping of a target mRNA (Muhlrad D and Parker R 1994). The mechanism of NMD is quite complex and
still not fully understood. In most mRNAs, the 3' UTR is entirely contained within the final exon. Hence, there are no splice junctions downstream of the stop codon. In transcripts harboring nonsense mutations, the presence of splice sites downstream of the premature stop codon signals the presence of a nonsense mutation and marks the transcript for destruction. Splice junctions in mRNAs are initially marked by the exon-junction complex (EJC), deposited on the mRNA during nuclear splicing. The UPF1 protein interacts with terminating ribosomes and termination factors, and should there be an intact EJC downstream of the stop codon, UPF1 will recognize UPF2 and UPF3, which associate with the EJC. Following the recruitment of additional NMD factors, the mRNA is regulated through multiple mechanisms, including decapping and degradation in the 5'-to-3' direction (Chang YF, et al. 2007), deadenylation and 3'-to-5' degradation (Mitchell P and D 2003), and endonucleolytic cleavage followed by 5'-to-3' and 3'-to-5' degradation (Eberle AB, et al. 2009; Gatfield D and Izaurralde E 2004).

1.1.3.5 Argonaute/IRPs and endonucleolytic degradation

Another form of transcript degradation involves endonucleolytic cleavage of an mRNA, which renders it vulnerable to both 5'-to-3' and 3'-to-5' exonucleolytic degradation. The ‘slicer’ activity of some Argonaute proteins serve to degrade mRNAs in this manner (Song JJ, et al. 2004).

It should be noted that, as for regulation of translation, regulation of mRNA stability is frequently negative. Examples discussed above fall into the category of regulation through decreases in the stability of the mRNA. However, regulation of transcript stability can also be positive (i.e., can increase the stability of the mRNA). One well-known example of the positive regulation of transcript stability is that of the increase in stability of the transferrin receptor mRNA.
Whereas the presence of IREs in the 5' UTR of a transcript confers regulation on the mRNA at the level of translation, the recruitment of IRPs to IREs located in the 3' UTR of the mRNA such as the transferrin receptor transcript actually increases the stability of the transcript. The recruitment of an IRP to this IRE increases the half-life of the mRNA, leading to an increase in the steady state levels of the transferrin receptor (Posch M, et al. 1999). This is due to the fact that IRPs protect the transcript from endonucleolytic decay, presumably by interfering with the ability of an endonuclease to target the transcript (Binder R, et al. 1994).

1.1.3.6 GW182/TNRC6: translational repression and transcript degradation

As discussed above, many of the mechanisms that result in the degradation of an mRNA are similar to those employed during translational regulation, with deadenylation being a major and often rate-limiting step (Chen CY and Shyu AB 2011). A well-known example of this phenomenon is the regulation of mRNAs by Argonaute and GW182/TNRC6, in which GW182 has both a translational repression function and a degradation activity (Fabian MR, et al. 2009). Studies of the mechanism of both forms of GW182-mediated repression revealed that both are largely dependent upon the ability of GW182 to bind to PABP and to recruit deadenylases to remove the poly(A) tail of the target (Huntzinger E, et al. 2013; Zekri L, et al. 2009).

1.1.3.7 P bodies: sites of translational repression and transcript degradation

An important subcellular structure implicated in mRNA degradation (as well as translational repression) is the P body (Eulalio A, et al. 2007; Parker R and Sheth U 2007), which belongs to the larger family of so-called RNA granules, which includes compartments such as
stress granules and neuronal granules. P bodies are specialized, non-membrane bound structures composed largely of aggregates of protein, including mRNA decay enzymes, and mRNAs. Once selected for degradation, mRNAs can be recruited to P bodies, where they are translationally repressed and degraded. Most of the degradation pathway components described above (Ccr4, Dcp1/2, Xrn1) have been shown to be present in P bodies. In addition, in mammalian cells P bodies contain components of the RNA interference (RNAi) machinery, such as RNA-induced silencing complex (RISC) and GW182. Interestingly, mRNAs can also be mobilized out of P bodies, particularly during recovery from stress conditions when P bodies may be used for temporary repression of housekeeping mRNAs (Brengues M, et al. 2005). The role of P bodies in storing (but not degrading) transcripts is shared by a related class of RNA granule, the stress granule, which contain repressed mRNAs associated with stalled 48S preinitiation complexes and other components of the translational machinery (Anderson P and Kedersha N 2006).

1.1.4 Multifactorial regulation of mRNA targets by RBPs and the pervasiveness of PTR

In a simple model, the regulation of mRNAs by RBPs involves a single mRNA molecule bound by a single RBP, leading to regulation of the target. However, most endogenous mRNAs are bound and regulated by multiple RBPs. The evidence for this so-called multifactorial model of PTR comes from efforts to characterize the ‘RBP interactome’ – the interaction network between all of the RBPs in a cell with their target transcripts. Such studies have found that canonical RBPs are a very abundant and diverse group of proteins in the cell (often numbering in the hundreds in mammalian cells). Each RBP in turn typically binds many transcripts (frequently in the range of several hundred to thousands) (Baltz AG, et al. 2012; Hafner M, et al. 2010; Kwon SC, et al. 2013). In addition, studies that have measured RNA-binding activity proteome-
wide have detected a large class of so-called non-canonical RBPs. Non-canonical RBPs are typically proteins that have previously defined roles (e.g., as metabolic enzymes) but have been since shown to also have RNA-binding activity (Mitchell SF, et al. 2013). Given the prevalence RBPs and the number of targets that many have, it is likely that most mRNAs are bound by multiple RBPs. The association of RBPs with target transcripts can be considered both over the course of the lifecycle of the mRNA, as well as at any one time. Efforts to estimate the number of RBPs associated with an average mRNA supports the notion that an average yeast mRNAs is bound by ~30 RBPs over its lifetime (Hogan DJ, et al. 2008).

From a regulatory standpoint, the presence of multiple RBPs on a single transcript raises the possibility of cooperative or antagonistic relationships among RBPs in determining the ultimate fate of an mRNA. While the activity of a given RBP on a transcript may in certain cases be insensitive to the presence of other RBPs on the same molecule, many lines of evidence suggest that this is often not the case. Indeed, computational approaches suggest that different trans-factors often interact with and regulate the same transcripts. For example, the canonical AUBP AUF1 has been shown to bind to many of the same transcripts as Ago2, and statistical evidence suggests these two factors cooperate to degrade their common targets (Wu X, et al. 2013). As will be discussed below, similar observations have been made with regards to Argonaute and Puf proteins. Additional examples abound, such as the observation that binding of RBPs found in P bodies can in some cases be cooperative (i.e., the binding of one RBP increases the frequency that another RBP is bound to the same transcript) (Mitchell SF, et al. 2013).
1.2 The maternal-to-zygotic transition in *Drosophila* embryogenesis

As discussed above, PTR is a critical component of gene regulation, and plays indispensable roles in many biological processes. Early *Drosophila* development is a classic example, and has served as a model for PTR for many years. This thesis project has used the early *Drosophila* embryo as such a model. The general steps of early *Drosophila* embryogenesis will therefore be outlined below, with an emphasis on the role of PTR in this process.

*Drosophila* development can be divided into several major phases, such as oogenesis, embryogenesis and larval development. The processes of oogenesis and the earliest stages of embryogenesis are either partially or completely under maternal genetic control, while the zygotic genome is involved in development during later embryogenesis and larval development. The handover of genetic control from the maternal to the zygotic genome is referred to as the maternal-to-zygotic transition (MZT). The MZT occurs over an extended period of time, and includes events that take place during both oogenesis and embryogenesis. As both the oocyte proper as well as the early embryo are largely transcriptionally silent, the role of PTR is of particular importance during the MZT. In fact, all forms of PTR discussed in section 1.1 are represented during this period of development, including mRNA localization, translational regulation, and mRNA degradation.

1.2.1 Oogenesis

Oogenesis takes place in the female ovary, which has an architecture reflecting the various stages of oocyte development. Oocytes move from the anterior end of the ovary to the posterior in an assembly line fashion as they develop. The process begins at the anterior with a self-sustaining population of germline stem cells (GSCs). GSCs are located in the so-called
germarium, where asymmetric cell divisions maintain the population of progenitor cells as well as produce differentiating cystoblasts (Wong MD, et al. 2005). Once differentiated, cystoblast cells divide 4 times with incomplete cytokinesis at each division, resulting in a cyst of 16 cells connected by cytoplasmic bridges referred to as ring canals (Spradling A 1993). One of these 16 cells becomes the presumptive oocyte, while the other 15 cells become so-called nurse cells. The nurse cells enter the endoreplicative cell cycle. This results in polyploid cells whose role is to express large quantities of mRNAs and proteins which are then transported into the developing oocyte through the ring canals. As each cyst moves toward the posterior, it develops according to both cell autonomous signals (i.e., the maternal germline factors principally supplied by the nurse cells) as well as cell non-autonomous somatic factors chiefly originating from the ovarian follicle cells surrounding the cyst.

As oogenesis proceeds, the oocyte undergoes several profound transformations which are critical for the MZT to occur successfully. Firstly, it becomes polarized, with both the anterior-posterior and dorsal-ventral axes being determined. Axis specification involves such processes as localization of various maternal mRNAs within the developing oocyte. Those localized mRNAs are often regulated by translational controls that ensure only properly localized mRNA is translated. Polarization also requires somatic signals from the follicle cells. In addition, the oocyte must begin to negotiate the process of meiosis, in which the diploid maternal genome is recombined and reduced to haploid status in preparation for fusion with the male pronucleus after fertilization (Page SL and Hawley RS 2003). Maternal meiosis is therefore a prerequisite for the creation of the zygotic genome and ultimately for the MZT. This process also depends on both germline maternal factors and somatic factors.
1.2.2 Egg activation

Once oogenesis is complete, the oocyte passes from the ovary to the uterus, where it undergoes a process referred to as egg activation, after which it is fertilized. In many species, including mammals, egg activation is coupled to fertilization; in brief, sperm entry leads to a spike in cytoplasmic calcium concentration (resulting from release of intracellular calcium stores) which triggers egg activation (Horner VL and Wolfner MF 2008). In many insect species, however, egg activation and fertilization are uncoupled. In fruit flies (Heifetz Y, et al. 2001) and certain species of wasp (King PE and Rafai J 1970; Went DF and Krause G 1974), this has been shown to be because egg activation is at least partially triggered by the mechanical stress of the egg passing through the oviduct. In Drosophila, recent evidence has suggested that this may be due to the activation of ‘stretch-activated’ ion channels, allowing for the influx of extracellular calcium from the hypotonic liquid found in the fly uterus (Horner VL and Wolfner MF 2008). Because of this, a Drosophila egg can be activated while remaining unfertilized.

1.2.3 The maternal phase of embryogenesis

Once fertilized and activated, the embryo is deposited by the female into the external environment, where embryogenesis proceeds. Egg activation triggers the completion of meiosis at this time, after which the male and female pronuclei fuse to form the zygotic genome. Once formed, the zygotic genome begins to undergo rapid rounds of duplication and mitosis. The zygotic nuclei undergo a truncated form of the cell cycle involving only DNA synthesis and mitosis, without intervening growth phases (so-called ‘S/M’ cycles). These mitoses occur without cell division, and produce in the space of 2-3 hours a large multi-nucleated syncytium containing approximately 6000 nuclei. One of the factors critical for these nuclear divisions is
the kinase PAN GU (PNG), which is activated during egg activation and is required to regulate CYCLIN B protein levels (Vardy L and Orr-Weaver TL 2007).

Little transcription occurs from the zygotic genome during these first 2-3 hours of embryogenesis. Thus, maternally-deposited factors control development during this period, and the embryo relies almost exclusively on post-transcriptional regulatory mechanisms to control gene expression and direct developmental events. This early period of embryogenesis is therefore referred to as the maternal phase. These maternal factors function in the replication of the zygotic genome described above, as well as further events in axis specification (typically involving localized translation of specific factors and the establishment of protein gradients along either the anterior-posterior or dorsal-ventral axes). In addition, the maternal phase is characterized by the large-scale degradation of many maternal mRNAs and proteins (Fig. 3).

1.2.4 Early or maternal mRNA degradation

Many post-transcriptional regulatory mechanisms of the early embryo act downstream of egg activation and PNG activity (Kronja I, et al. 2014). One prominent example of PNG-dependent PTR occurs through the PNG-dependent translation of maternally-deposited smg mRNA (Tadros W, et al. 2007). This leads to the appearance of the RBP SMG, which then proceeds to degrade (Tadros W, et al. 2007) and translationally repress (Chen L, et al. 2014) mRNAs present in the early embryo. The post-transcriptional changes brought about by SMG are profound, with approximately 2/3 of those transcripts that are degraded during the early phase of decay being targeted by SMG. Nevertheless, other factors must also be active in degrading mRNA during this period, since not all transcripts are targeted in a SMG-dependent manner. Recently, the RBP BRAIN TUMOUR (BRAT) has been shown to bind and degrade transcripts during the early phase (Laver JD, et al. 2015), and additional factors may also be involved.
1.2.5 The zygotic phase of embryogenesis

At approximately the same time, the rapid replication of the zygotic genome slows and ultimately ceases temporarily, an event which is permissive for the beginning of transcription from the zygotic genome (Fig. 3). This zygotic genome activation (ZGA) is actually detectable in the case of a handful of transcripts before the cessation of zygotic mitoses, but the main wave of ZGA begins with the arrest of the cell cycle in the third hour of embryogenesis. The majority of the arrested nuclei have by this time migrated to the periphery of the embryo, and begin to undergo a process termed cellularization, in which the plasma membrane of the embryo extends down between nuclei and envelops them, ultimately forming distinct cells. This process is the first morphological event in the development of the embryo that is dependent on ZGA, and is referred to as the mid-blastula transition (MBT).

1.2.6 Late or zygotic mRNA degradation

Once ZGA occurs, zygotic factors begin to replace maternal factors in the embryo, and the MZT is almost complete. Despite the increasing dominance of zygotic factors in directing development in the immediate aftermath of ZGA, there still remain large numbers of maternal mRNAs in the embryo at this time. Interestingly, one of the early functions of certain zygotic factors appears to be the further clearance of a subset of these maternal mRNAs (Fig. 3), often termed late or zygotic decay. This late decay is characterized by its kinetics, since transcripts targeted exclusively by late decay will remain stable during the first 2-3 hours of embryogenesis, and only begin to decay after the MBT.
A. Fertilized embryos

B. Unfertilized eggs

- Maternal stable
- Early (maternal) degraded
- Late (zygotic) degraded
- Zygotically expressed

C. Early (maternal) degradation

- Late (zygotic) degradation

- Zygotic transcription

- Egg activation
  - Endoreplication of zygotic nuclei
  - Cell cycle arrest
  - Cellularization

- 1h
- 2h
- 3h
Figure 3. Post-transcriptional regulation during early *Drosophila* embryogenesis. (A) Profile of different classes of mRNA during early *Drosophila* embryogenesis. At the point of egg activation (0 hours), the egg contains a large number of maternally-contributed transcripts. Many of these transcripts are stable over the first few hours of development, and persist in the embryo. However, many are also turned over as part of the maternal-to-zygotic transition (MZT – see text). Such transcripts are targeted during two broad phases of mRNA degradation. The first (yellow line) begins immediately after egg activation, is under the exclusive control of maternal factors and is referred to as early (or maternal) decay. After the onset of zygotic transcription and the appearance of zygotic transcripts (green line) and proteins, a second phase of mRNA degradation begins (red line). (B) In unfertilized eggs, only the early (maternal) phase of degradation occurs, and neither zygotic transcription nor late (zygotic) decay occur. (C) Simplified diagram of early *Drosophila* embryogenesis, illustrating the various phases of post-transcriptional regulation (PTR) occurring alongside major developmental and morphological milestones, such as egg activation, rapid endoreplication of zygotic nuclei, zygotic nuclei cell cycle arrest, and cellularization. Note that the actual number of zygotic nuclei (blue) approaches 6000 by cell cycle arrest and cellularization.
1.2.7 Unfertilized eggs

As discussed above, egg activation and fertilization are functionally uncoupled in fruit flies. In fact, activated but unfertilized eggs (UEs) can be collected in the laboratory from virgin female fruit flies. In UEs, the early, maternal form of transcript decay (including the principle, SMG-dependent early mRNA degradation) occurs normally. This is because it is triggered by egg activation and PNG activity, which proceed normally in UEs. In contrast, late or zygotic decay does not occur in UEs, since it is dependent on ZGA, which in turn cannot occur without the proper formation of the zygotic genome. Examination of gene expression in UEs is therefore a common approach to distinguishing early (maternal) from late (zygotic) forms of PTR (Tadros W, et al. 2007; Thomsen S, et al. 2010). This approach has already been used to show that zygotic degradation of mRNAs is widespread in *Drosophila* embryos, with several hundred mRNAs stabilized in UEs (Thomsen S, et al. 2010). Evidence also suggests that the late (zygotic) wave of mRNA degradation is conserved in vertebrates (Alizadeh Z, et al. 2005; Ferg M, et al. 2007; Giraldez AJ, et al. 2006; Hamatani T, et al. 2004; Mathavan S, et al. 2005), although it is not as well studied in other species.

1.2.8 miRNAs and late (zygotic) mRNA degradation

miRNAs appear to play an important role in zygotic mRNA turnover in both *Drosophila* and other species. A class of miRNAs (the *miR-309* cluster of miRNAs, which are produced as a single primary miRNA transcript) have been shown to target approximately 400 transcripts after the MBT in flies (Bushati N, et al. 2008; Thomsen S, et al. 2010). Additional miRNAs also play a role during zygotic degradation; miR-14 is required to degrade the *Hr78* transcript during this period and the seed sequences for other embryonic miRNAs are enriched in mRNAs degraded by
the zygotic degradation machinery (Thomsen S, et al. 2010). In zebrafish, the miRNA miR-430 is also expressed zygotically and functions to degrade transcripts after the MBT, suggesting the role of miRNAs in late (zygotic) mRNA degradation is conserved (Giraldez AJ, et al. 2006). Interestingly, miR-430 has also been shown to translationally repress target transcripts in 4 hour old zebrafish embryos (before gastrulation) in a deadenylation-independent manner, while degradation of target mRNAs only begins in 6 hour old embryos (post-gastrulation) (Bazzini AA, et al. 2012). This has been suggested to be a part of a general switch in the role of the poly(A) tail over the course of early development, with short poly(A) tails resulting in translationally repressed but stable transcripts before gastrulation, while the same signal results in transcript decay after gastrulation (Subtelny AO, et al. 2014).

1.2.9 PUMILIO (PUM), BRAIN TUMOUR (BRAT) and other factors involved in late (zygotic) mRNA degradation

Efforts have begun to identify other factors involved in late (zygotic) degradation. In addition to finding an enrichment of miRNA binding sites in late (zygotically) degraded transcripts, Thomsen et al. detected a significant overlap between transcripts degraded during this time period and transcripts bound by the RBP PUMILIO (PUM) (Gerber AP, et al. 2006; Thomsen S, et al. 2010; Zamore PD, et al. 1997; Zhang B, et al. 1997). A novel approach was employed by De Renzis et al. to examine late (zygotic) decay by measuring mRNA stabilities in embryos missing various fragments of all the major chromosomes of the Drosophila genome (De Renzis S, et al. 2007). In this analysis, an mRNA whose stability increases upon ablation of a given chromosome must depend on zygotic factor(s) originating from that chromosome. This study determined that the degradation of many transcripts after the MBT depends on factors encoded on all major chromosomes (i.e., the X, 2nd and 3rd chromosomes). Searching these
destabilized mRNAs for enriched *cis*-elements yielded several motifs. Firstly, several sequences containing ‘UGUU’ were identified. The authors speculate that these may represent binding sites for PUM, which are known to bind to similar sequences (see section 1.3.2). Interestingly, recent data suggests they may also represent binding sites for BRAT. This is because BRAT has been shown to have RNA binding activity (Loedige I, et al. 2014), and recognizes sites containing ‘UGUU’ sequences (Laver JD, et al. 2015), a property it may share with other TRIM-NHL domain-containing proteins (Loedige I, et al. 2013). In fact, while BRAT has been shown to exhibit a preference for ‘UGUU’ motifs, *in vitro* PUM also demonstrates some affinity for these motifs. As such, it is difficult at this time to determine whether the motifs identified by De Renzis et al. constitute PUM or BRAT binding sites.

In addition, De Renzis et al. discovered an enrichment for AU-rich motifs in their set of zygotically degraded mRNAs. AU-rich motifs have been shown to cause the destabilization of mRNAs that contain them through recruitment of AUBPs (Gingerich TJ, et al. 2004; Gratacós FM and Brewer G 2010) which recognize them as binding sites. Taken together, many factors such as PUM, BRAT, *miR-309* miRNAs, and potentially AUBPs (as well as others yet to be identified) are likely involved in degrading subsets of the mRNAs targeted by late (zygotic) degradation.

### 1.3 Pumilio-Fem3 binding factor (Puf) proteins

#### 1.3.1 The biological roles of Puf proteins

**1.3.1.1 Germline maintenance**

Pumilio-Fem3 binding factor (Puf) proteins are a conserved family of RBPs named for several founding members – *Drosophila* PUM and *C. elegans* Fem3 binding factor (Fbf-1/2)
(Puf – PUM and Fbf). All Puf proteins share a conserved RNA binding domain (RBD) (sometimes called the Puf homology domain – Puf HD – referred to hereafter as the Puf RBD), typically consisting of eight copies of the so-called Puf repeat.

*Drosophila* PUM was originally characterized in a screen of maternal effect lethal mutations, where it was found that PUM is required for the correct patterning of the posterior of the embryo (Lehmann R and Nüsslein-Volhard C 1987). It was subsequently determined that PUM is also required for maintenance of germline stem cells in the ovary of female flies (Forbes A and Lehmann R 1998; Lin H and Spradling AC 1997). As discussed in section 1.2.1, in the *Drosophila* germarium, GSCs give rise to differentiating egg cysts while maintaining their numbers through a process of asymmetrical division, in which a single GSC divides to give rise to another GSC as well as a differentiating cystoblast, which in turn gives rise to a mature egg. In ovaries lacking PUM activity, these asymmetric divisions do not occur correctly, and both daughter cells of a dividing GSC differentiate into cystoblasts. This ultimately results in the depopulation of GSCs from the germarium and eventual female sterility. In *C. elegans*, Fbf-1/2 have a similar role in maintaining the so-called ‘mitotic zone’ of the gonad (composed of pluripotent stem cells); in *fhf-1; fhf-2* double mutants, germline stem cells differentiate and the mitotic zone depopulates (Crittenden SL, et al. 2002). Fbf-1/2 additionally have a role in germline switching in *C. elegans* hermaphrodites, which initially produce sperm and then switch to produce oocytes. Fbf-1/2 directly recognize the *fem-3* transcript, whose translation is required for spermatogenesis and which is repressed in order to allow for the switch to oogenesis. Fbf-1/2 bind to and repress *fem-3* mRNA through recognition of cis-acting elements in the *fem-3* 3' UTR. This repression is necessary for the sperm to oocyte switch in wild-type animals, since in *fhf-1; fhf-2* double mutants, regulation of the *fem-3* mRNA is compromised, resulting in ‘masculinized’ germlines that produce only sperm (Zhang B, et al. 1997). A role for Puf proteins has also been
demonstrated in negatively regulating the differentiation of Dictyostelium cells upon starvation, suggesting that the general maintenance of pluripotency is an ancient function of Puf proteins (Souza GM, et al. 1999). In mouse, Pum1 has been shown to be required for maintenance of spermatogenic stem cells, although this is actually mediated through a downregulation of the p53 pathway and apoptosis and not through pluripotency maintenance (Chen D, et al. 2012).

1.3.1.2 Other roles

Puf proteins have subsequently been shown to have roles in other cell types, in particular in the nervous system. Mutations in the pum gene have been shown to result in long-term memory formation defects in Drosophila (Dubnau J, et al. 2003). Anatomically, the role of PUM is to modulate the number and size of synapses in developing neurons. It does so in part by binding to and regulating the eIF4E mRNA in the post-synaptic compartment, and thereby modulating the levels of the GluRIIa glutamate receptor (Menon KP, et al. 2004). Drosophila PUM also regulates the transcript encoding the sodium channel paralytic (para) and therefore indirectly regulates the electrophysiological properties of larval neurons (Muraro NI, et al. 2008). Mammalian Pum2 appears to be involved in the formation of RNA granules in neurons (Vessey JP, et al. 2010), and is required for the homeostasis of neuronal membrane excitability through the regulation of various genes, including (as in Drosophila) membrane ion channels (Driscoll HE, et al. 2013; Fiore R, et al. 2014; Mee CJ, et al. 2004). In addition, mutations to mammalian Pum2 result in learning and memory defects in mice (Siemen H, et al. 2011). Puf proteins likely have many other roles as well. For example, Puf proteins have been shown to regulate cell proliferation in mammalian primary fibroblasts through direct regulation of the 3′ UTR of the p27 mRNA, a tumour suppressor (Kedde M, et al. 2010).
1.3.2 The binding specificity of Puf proteins

All Puf proteins share the Puf RBD, which is composed of so-called Puf repeats (which are largely α-helical and composed of approximately 36 amino acids each) (Wang X, et al. 2001). Most Puf proteins contain 8 such repeats. Puf proteins recognize linear RNA sequences, usually between 8-10 nt in length (Gerber AP, et al. 2004; Gerber AP, et al. 2006). Interestingly, crystal structures of the Puf RBD bound to its cognate RNA show that each Puf repeat usually makes contact with a single nucleotide (Wang X, et al. 2002), providing a structural explanation for why Puf proteins recognize linear RNA sequences approximately 8 nt in length (referred to in the current work as PUM binding elements – PBEs). The fact that the binding specificity of Puf proteins is determined on a one repeat-one nucleotide basis raises the interesting possibility of ‘programming’ Puf proteins to recognize different targets by altering the order of Puf repeats in the Puf RBD. In fact, considerable progress has been made in this regard, with synthetic recombinant Puf proteins rationally designed to recognize different linear sequences of RNA and to target mRNAs for regulation (Campbell ZT, et al. 2014; Hall TM 2014).

Numerous studies have determined the PBE sequence recognized by Puf proteins in various species. Such studies typically begin by immunoprecipitating a Puf protein from a cellular extract, and then determining which mRNAs in the cell are enriched in the precipitate through detection by microarray or high-throughput sequencing. Once a list of bound mRNAs has been arrived at, de novo motif discovery approaches are used in which motifs that are enriched in the bound set versus the unbound set are determined. Such studies have been carried out in yeast, fruit fly and human cells, and the motif enriched in each case is highly similar (Gerber AP, et al. 2004; Gerber AP, et al. 2006; Morris AR, et al. 2008). It usually conforms to the 8-mer UGUA(N)AUA, where (N) is any nucleotide. In addition, the same motif was identified when the binding sites for human Pum2 were characterized in HEK 293 using
photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) (Hafner M, et al. 2010). The 5’ ‘UGUA’ core motif is the most highly conserved part of the binding sequence, with the 3’ AUA motif displaying greater variability.

The most comprehensive studies of the exact binding specificities of Puf proteins have been carried out in yeast. Yeast have the advantage of encoding multiple Puf proteins, which allows for comparison of the binding specificities of different Puf proteins in the same cell. Interestingly, Gerber et al. showed that for 5 of the *S. cerevisiae* Puf proteins (Puf1p-Puf5p) each bound to slightly different sequences (although each protein did recognize the UGUA core motif) (Gerber AP, et al. 2004). The binding specificities of yeast Puf proteins tended to vary most often in the number of N’s tolerated between the 5’ UGUA core motif and the 3’ AUA motif, with anywhere from 1 to 3 Ns tolerated depending on the Puf protein (i.e., UGUA(N1-3)AUA). Structural studies of the Puf4p RBD, which recognizes a 9 nt sequence with 2 Ns between the UGUA and AUA motifs (UGUA(NN)AUA), revealed that several key residues in the Puf4p RBD are responsible for this specificity, and that mutation of these residues to the equivalent residues in the Puf3p RBD (which recognizes only a single N between the UGUA and AUA motifs) changes the specificity of the Puf4p RBD to that of Puf3p (Miller MT, et al. 2008). Interestingly, Puf3p itself was shown to have a preference for a C 2nt upstream of the core motif, and therefore has a binding specificity of C(N)UGUA(N)AUA (Gerber AP, et al. 2004). The preference for a 5’ C is unusual for Puf proteins, and it was subsequently shown through solving the structure of the protein to be due to a non-conserved binding pocket outside of the canonical binding surface of the Puf3p RBD (Zhu D, et al. 2009). The authors furthermore show that the presence of the upstream C is critical for the recognition of the *COX17* 3’ UTR, a previously characterized Puf3p target (Olivas W and Parker R 2003). Taken together, this work indicates
that while Puf proteins largely recognize similar sequences that always contain the 5' UGUA core motif, slight changes in the Puf RBD can subtly shift the exact binding specificity.

1.3.3 The structure of Puf proteins

There are few defined domains shared by Puf proteins outside of the Puf RBD. There are, however, some conserved structural features found in many members of the Puf family. In particular, in many Puf proteins (including *Drosophila* PUM and the mammalian homologues Pum1/2) the N-terminus is thought to be largely disordered and contains repetitive asparagine- and glutamine-rich stretches characteristic of aggregation-prone or prion-like proteins. Interestingly, the presence of such aggregation-prone domains is a trait shared by many RBPs. *In vitro*, such domains have been shown to form ‘hydrogel’-like structures that may facilitate recruitment into subcellular structures such as RNA granules (Han TW, et al. 2012; Kato M, et al. 2012).

1.3.4 Mechanisms of Puf protein regulation

1.3.4.1 Direct regulation of target transcripts

Puf proteins regulate their target transcripts using several different mechanisms. As with many RBPs, Puf proteins do not generally directly regulate their target mRNAs, but rather recruit downstream factors to do so. There are, however, several exceptions to this rule. For example, *Xenopus* Pumilio 2 (XPum2) has been shown to directly bind to the 5′ cap structure, an activity required for its regulation of the *Xenopus* oocyte transcript RINGO/SPY (Cao Q, et al. 2010). Interestingly, the authors of this study identified a residue (W344) required for 5′ cap binding, and found that it is conserved in other vertebrate Puf homologues, suggesting this function may be conserved. Additionally, yeast Puf6p has been shown to directly interact with the translation
initiation factor eIF5B, blocking recruitment of the 60S ribosome to *ASH1* mRNA and therefore translationally repressing the target (Deng Y, et al. 2008).

1.3.4.2  **Deadenylation- and PABP-dependent regulation by Puf proteins**

With the exception described in section 1.3.4.1, however, Puf proteins are believed to mostly function through other co-factors, with the role of Puf proteins restricted to target recognition (i.e., Puf proteins are the specificity factor and not the regulator *per se*). The best-characterized mechanism by which Puf proteins regulate their target transcripts occurs through the interaction of the Puf RBD and the deadenylase component Pop2 (Goldstrohm AC, et al. 2006), a component of the Ccr4-Pop2-Not complex (Temme C, et al. 2014). The interaction of Puf homologues with Pop2 has been tested and appears to be conserved from yeast to human (yeast homologues Puf4p and Mpt5/Puf5p, *C. elegans* PUF-8, and human homologue Pum1 all bind to the species-specific Pop2 homologue as well as yeast Pop2p). The recruitment of the Ccr4-Pop2-Not complex to an mRNA bound by a Puf protein through the Puf-Pop2 interaction offers a simple model for how Puf proteins can negatively regulate their targets, since deadenylation of a transcript will typically lead to its translational repression and/or its degradation (see section 1.1.2.4).

It has also recently been shown that the *Drosophila* PUM RBD interacts directly with PABP in S2 cells (Weidmann CA, et al. 2014). Knockdown of PABP interferes with the ability of both the *Drosophila* PUM RBD and the human Pum2 RBD to repress a reporter RNA in S2 and HEK293 cells, respectively, suggesting there may be a conserved mechanisms by which PABP can cooperate with Puf proteins to repress target transcripts.
1.3.4.3 Additional mechanisms of regulation: \( p27 \) mRNA

While central to the current understanding of Puf protein-mediated regulation, the Puf-Pop2 interaction is not the only Puf protein-dependent mechanism to be described to date. For example, data exist which suggest that Puf proteins can also regulate their target transcripts through deadenylation-independent mechanisms (Chagnovich D and Lehmann R 2001; Van Etten J, et al. 2012). There are also individual cases where Puf proteins have been shown to regulate target transcripts through novel mechanisms. For example, one study examining the regulation of the transcript encoding the tumour suppressor \( p27 \) found that mammalian Pum1 cooperates with the miRNA machinery to down-regulate \( p27 \) mRNA in cycling fibroblasts (Kedde M, et al. 2010). The mechanism involves the modulation of the secondary structure of the \( p27 \) 3' UTR. The \( p27 \) mRNA is a target of the miRNAs miR-221 and -222, but the seed sequences recognized by the miRNAs are normally inaccessible due to local secondary structure. Phosphorylation of Pum1 in cycling fibroblasts leads to Pum1 binding to the \( p27 \) 3' UTR in the vicinity of the miRNA seed sequences, leading to local unfolding of the RNA and allowing for Argonaute recruitment to the transcript and mRNA degradation. Taken together, the data suggest that Puf proteins have the ability to invade local secondary structure, allowing for the binding of other trans-factors to target transcripts.

1.3.4.4 Additional mechanisms of regulation: the PUM N-terminus

The mechanisms described above largely depend on the Puf RBD, with little role for domains outside of the RBD. This fact, along with the early observation that expression of the PUM RBD was able to partially rescue a weak \( pum \) mutant phenotype in embryos (Wharton RP, et al. 1998), has led to a general consensus within the Puf protein literature that the principle component of all Puf proteins is the RBD, with regions of the protein outside of this domain
having little or no function. However, recent data has challenged this and supported a model in which regions outside of the RBD such as the N-terminal aggregation-prone region also play a role in repressing target transcripts. For example, full-length PUM mediates greater repression of target transcripts than does the PUM RBD in *Drosophila* tissue culture cells and in vitro tethering experiments have demonstrated that the N-terminus of *Drosophila* PUM can repress targets independently of the RBD (Weidmann CA and Goldstrohm AC 2012). Interestingly, Weidmann and Goldstrohm demonstrated that while the aggregation-prone regions of the PUM N-terminus will repress a target mRNA, there are additional domains within the N-terminus that display autoregulatory functions, both positively and negatively modulating the repressive activity of the aggregation-prone domains. The *Drosophila* PUM N-terminus also appears to play a role *in vitro*, since overexpression of the aggregation-prone domains of the PUM N-terminus in fly muscles produces phenotypes similar to *pum* loss of function mutations (Salazar AM, et al. 2010).

### 1.3.5 Puf proteins and the RNAi machinery

There are several lines of evidence suggesting that Puf proteins directly interact with the RNAi machinery to regulate target mRNAs. For example, the *C. elegans* Puf protein Fbf-1 directly interacts with the non-canonical Argonaute protein CSR-1 in worm extracts, and together they form a complex with EFT-3 (the *C. elegans* translation elongation factor eEF1A homolog). This interaction attenuates the GTPase activity of EFT-3, suggesting a mechanism by which Puf and Argonaute proteins might translationally repress target transcripts during translational elongation (Friend K, et al. 2012). Furthermore, this mechanism appears to be conserved, since an equivalent mammalian complex was detected and mutation of homologous residues shown to be required for formation of the Fbf-1-CSR-1-EFT-3 complex in the
mammalian homologues were shown to abrogate regulation of a reporter mRNA carrying PUM binding sites in vitro. Interestingly, the mechanism by which Pum2, Argonaute 1 (Ago1) and eEF1A regulate reporter mRNAs was shown to be deadenylation-independent and in fact occurs during translational elongation, suggesting it is truly a different mechanism from the Puf-Pop2 interaction.

Interestingly, several bioinformatic studies have detected an enrichment of Puf protein binding sites within miRNA target transcripts, and vice versa (Galgano A, et al. 2008; Incarnato D, et al. 2013). While these studies do not directly support any particular mechanism, they lend further support to the notion that Puf proteins and the RNAi machinery cooperate to regulate mRNAs. In addition, individual mRNAs have been shown to be regulated by both Argonaute and Puf proteins, although whether Argonaute and Puf proteins are truly cooperating or are acting independently of each other in these cases remains unclear (Nolde MJ, et al. 2007).

It should be noted that some data exist which challenges the view that Puf protein and Argonaute proteins function together, or at least suggest that they do not always do so. In particular, the repression of Puf protein binding sites in various reporters in both HEK293 (human) and S2 (fly) cells was not dependent on Argonaute proteins, since neither 1) deletions in transfected Puf proteins that interfered with Puf-Argonaute binding, nor 2) knockdown of Argonaute proteins interfered with the ability of Puf proteins to repress reporter RNAs (Weidmann CA, et al. 2014). It therefore appears as if Puf proteins and Argonaute proteins can cooperate, but that they can also act independently of each other.
1.3.6 The complexity of PUM-mediated regulation during *Drosophila* embryogenesis

As discussed above, Puf-dependent regulation of target transcripts appears to be quite complex, with different mechanisms employed by Puf proteins in different cellular contexts. Studies of regulation mediated by PUM in the *Drosophila* embryo further underscore this complexity, and suggest that the mechanisms that underlie PUM-mediated regulation are highly variable. PUM-dependent regulation furthermore appears to be an excellent example of the concept of multifactorial PTR, as the form of regulation imposed on different PUM targets appears to vary according to the presence or absence of other *trans*-factors making direct contact with a given mRNA.

1.3.6.1 PUM-mediated repression of *hunchback (hb)* mRNA

Several mRNAs have been characterized as critical targets of PUM during early *Drosophila* development. The best-known example is the maternally-contributed mRNA *hunchback (hb)*, which PUM negatively regulates. HB is a transcription factor that functions in segmentation during early embryogenesis by directly and indirectly regulating the expression of genes such as *caudal*, *Kruppel*, *knirps*, and *giant* (Hülskamp M, et al. 1990; Schulz C and Tautz D 1995). It is critical for the function of HB that its expression be restricted to the anterior of the embryo.

Regulation of the *hb* transcript is dependent on two copies of a *cis*-acting element in the *hb* 3' UTR called the NANOS response elements (NRE), which contains within it a PBE. Regulation of *hb* mRNA by PUM requires at least two other key factors that both make direct contact with the *hb* 3' UTR. One is BRAT, which is a member of the TRIM-NHL family of proteins. BRAT was originally characterized for its role in regulating cell proliferation in the fly
nervous system, where mutations to the *brat* locus result in neoplasms (Arama E, et al. 2000; Kurzik-Dumke U, et al. 1992). BRAT was subsequently shown to interact with PUM in a complex with NANOS (NOS) (see below), an interaction that is required for the proper regulation of the *hb* mRNA (Sonoda J and Wharton RP 2001). The original model of PUM-BRAT regulation of *hb* mRNA posited that PUM and NOS recruited BRAT to *hb* mRNA without BRAT making direct contact with RNA. However, recent data shows that BRAT (and some related TRIM-NHL proteins (Loedige I, et al. 2013)) are in fact RBPs, and that BRAT makes direct contact with the *hb* mRNA through its TRIM-NHL domain (Loedige I, et al. 2014).

The other critical factor required for PUM-mediate regulation of *hb* mRNA is NOS. *nos* was originally characterized as one of the so-called ‘posterior group’ genes that are required for the correct specification of the posterior of the embryo, a group that includes *pum* (Nüsslein-Volhard C, et al. 1987). It was subsequently demonstrated that the NOS gene product is one of the critical determinants of posterior specification, with other posterior group genes functioning to localize NOS to the posterior (Lehmann R and Nüsslein-Volhard C 1991). Unlike PUM and BRAT, NOS protein is spatially restricted to the posterior of the embryo. This explains why *hb* mRNA is only repressed at the posterior, since it has been demonstrated that mislocalization of NOS protein to the anterior of the embryo results in *hb* mRNA repression in that compartment (Gavis ER and Lehmann R 1992). Interestingly, the only indispensable activity of NOS during embryogenesis is in *hb* mRNA repression, since *nos:**hb* double mutant embryos are viable (Irish V, et al. 1989).

NOS makes direct contact with nucleotides in the NRE, and this interaction is stabilized in the presence of PUM and BRAT (Sonoda J and Wharton RP 1999; Sonoda J and Wharton RP 2001).

The mechanism of PUM-mediated regulation of *hb* mRNA in the posterior of the embryo has been studied in some detail, and has served as the canonical model of PUM-mediated regulation more generally. Recruitment of the PUM/NOS/BRAT complex to the *hb* 3' UTR
results in deadenylation of the mRNA and down-regulation of both the *hb* mRNA and HB protein (Gamberi C, et al. 2002). As discussed above, there is a conserved interaction between Puf proteins and the Ccr4-Pop2-Not complex (Goldstrohm AC, et al. 2006). Although there is no direct evidence that the fly Ccr4 homolog (known as ‘Twin’) is required for PUM-mediated *hb* mRNA deadenylation, this remains the most likely model. There is also evidence that PUM-mediated regulation of *hb* mRNA involves deadenylation-independent mechanisms, since a deadenylated *hb* reporter RNA injected into the posterior of *Drosophila* embryos are regulated in a PUM-dependent manner (Chagnovich D and Lehmann R 2001). This deadenylation-independent mechanism may be related to the cap-binding protein 4E-HP, which has been shown to be involved in *hb* mRNA repression. 4E-HP is recruited by BRAT following formation of the PUM-NOS-BRAT complex on the *hb* NRE (Cho PF, et al. 2006). 4E-HP binds to the 5' cap of the mRNA, competing with eIF4E. Since 4E-HP cannot interact with eIF4G, its binding to the cap leads to the translational repression of the message.

1.3.6.2 PUM-mediated repression of *cyclin B (cycB)* mRNA

The requirement of the PUM-NOS-BRAT complex in *hb* mRNA regulation has served as a dominant paradigm of our understanding of PUM-mediated regulation, and it has been assumed that PUM is entirely dependent on NOS and BRAT in order to regulate its targets. However, more recent data examining other PUM targets in *Drosophila* embryos has challenged this view. In many cases, the regulation of these mRNAs does not depend on NOS or BRAT. Taken together, these data suggest the canonical model of PUM-mediated repression might be a simplistic view of PUM activity in the embryo.

One example is the PUM-mediated regulation of *cyclin B (cycB)* mRNA in the *Drosophila* embryo. The regulation of *cycB* mRNA by PUM appears to be a highly conserved
feature of early development, as there is a well-characterized role for a PUM homolog (XPum1) in \textit{cycB} mRNA regulation during \textit{Xenopus} oogenesis (Nakahata S, et al. 2001; Nakahata S, et al. 2003). XPum1-mediated repression of \textit{cycB} mRNA in \textit{Xenopus} occurs during oogenesis, and is dependent on the CPEB and the presence of both CPEs and PBEs in the transcript. This repression is relieved during egg activation, allowing CycB protein to be expressed. During \textit{Drosophila} egg activation and early embryogenesis, the PNG kinase is required for CycB expression in the bulk cytoplasm of the embryo, which in turn is required for the mitosis of somatic nuclei. Data indicate that PUM represses \textit{cycB} mRNA before egg activation, and that this repression is relieved by PNG activity at egg activation (Vardy L and Orr-Weaver TL 2007).

Interestingly, despite the fact that \textit{cycB} mRNA repression is relieved in the bulk cytoplasm at the beginning of embryogenesis, \textit{cycB} mRNA must continue to be translationally repressed in the so-called pole cells or primordial germ cells (PGCs), which form early in embryogenesis and ultimately go on to form the germline of the organism. Between the formation of the PGCs and their eventual arrival in the presumptive gonad, they must maintain cell cycle arrest. This cell cycle arrest is accomplished through repression of \textit{cycB} mRNA (Dalby B and Glover DM 1993).

This regulation is dependent on both PUM and NOS, as \textit{cycB} regulation is disrupted in both \textit{pum} mutant and \textit{nos} mutant PGCs (Asaoka-Taguchi M, et al. 1999; Kadyrova LY, et al. 2007). Although the PGCs are located at the posterior of the embryo, regulation of \textit{cycB} mRNA is not observed in the broader posterior compartment (as is the case with \textit{hb} mRNA) and is completely restricted to the PGCs (Asaoka-Taguchi M, et al. 1999). Interestingly, unlike for regulation of \textit{hb} mRNA, BRAT activity is dispensable for regulation of \textit{cycB} mRNA in the PGCs. This result is consistent with \textit{in vitro} data showing the \textit{hb} 3' UTR is bound by PUM, NOS and BRAT (Sonoda J and Wharton RP 2001), while the \textit{cycB} 3' UTR only interacts with PUM
and NOS (Kadyrova LY, et al. 2007). This suggests that the regulation of the \( hb \) and \( cycB \) transcripts is mechanistically different from each other. In support of this hypothesis, Kadyrova et al. demonstrated that swapping of the regulatory elements located in the \( hb \) and \( cycB \) 3' UTRs resulted in a corresponding reversal in regulation. In particular, inserting the portion of the \( cycB \) mRNA recognized by PUM in EMSA assays into the \( hb \) transcript resulted in ‘\( cycB \)-like’ regulation of \( hb \) mRNA (i.e., repression restricted to only the PGCs and not observed in the broader posterior). In addition, direct tethering of NOS to \( cycB \) mRNA is sufficient for its repression in both WT and \( pum \) mutant PGCs, suggesting that the role of PUM in \( cycB \) repression is restricted to NOS recruitment to the transcript. However, tethering of NOS to \( hb \) mRNA does not result in its repression in the posterior, suggesting the PUM (and BRAT) need to directly interact with the \( hb \) mRNA in order for it to be regulated.

### 1.3.6.3 PUM-mediated repression of \( bicoid \) (\( bcd \)) mRNA

PUM is also known to regulate \( bcd \) mRNA in the anterior of the embryo (Gamberi C, et al. 2002). \( bcd \) mRNA has a critical role in patterning the developing head. PUM regulation of \( bcd \) mRNA in the anterior is interesting because NOS protein is believed to be either absent or present at very low levels in this compartment. Loss of \( bcd \) mRNA regulation by PUM in the anterior of the embryo results in defects in the developing head skeleton. Interestingly, mutations to NOS also cause such defects, but the effects are much subtler, suggesting that PUM regulates \( bcd \) mRNA in a largely NOS-independent manner. While the dependence of \( bcd \) mRNA regulation on BRAT was not tested in this study, the canonical model of PUM-mediated regulation involving the formation of the PUM-NOS-BRAT complex is nevertheless challenged by this data.
1.4 Thesis outline

Transcripts bound by PUM (Gerber AP, et al. 2006) have been shown to significantly overlap with mRNAs degraded during the late (zygotic) phase of transcript decay in the fly embryo (Thomsen S, et al. 2010), despite the fact that a previous genome-wide survey of PUM’s role in transcript decay using weak pum alleles failed to demonstrate such a role. In this thesis I describe a transcriptomic analysis that uses strong pum alleles and that demonstrates a direct role for PUM in a late (zygotic) pathway of mRNA decay in the embryo. Analysis of these data and subsequent experiments demonstrate that an important target of PUM is the smg mRNA. PUM-induced degradation of smg mRNA is required for timely clearance of SMG protein from the embryo and disruption of this process results in downregulation of zygotic mRNAs that carry SMG-binding sites (SREs).

PUM is a maternally-contributed protein and is therefore present during the early (maternal) phase of decay (Macdonald PM 1992). Thus, PUM’s role in zygotic decay is curious. The data I present here indicate that the PUM-binding sites within the smg mRNA have reduced affinity for PUM and as such co-factors function with PUM to induce smg mRNA degradation. I have identified BRAT and AGO1 as two such co-factors and conclude that PUM, BRAT and AGO1 function together in a late mRNA decay pathway. Finally, I propose that one or more zygotically expressed miRNAs controls the timing of smg mRNA decay.
2 Materials and Methods

2.1 Fly stocks

All fly stocks were maintained in house using standard *Drosophila* husbandry techniques. Wild-type flies were *w*^1118^. The *pum^ET7^* and *pum^MSC^* alleles were maintained as balanced stocks (*pum^ET7^/TM3;Sb and *pum^MSC^/TM3;Sb*). *pum* mutant (*pum^ET7^/pum^MSC^*) flies were obtained by crossing *pum^ET7^/TM3;Sb* females with *pum^MSC^/TM3;Sb* males and collecting non-Sb flies. The *ago1* shRNA expressing line is *P{TRiP.HMS00610}attP2* (maintained as a homozygous stock - Bloomington stock ID # 33727) and the *mCherry* shRNA expressing line is *P{VALIUM20-mCherry}attP2* (maintained as a homozygous stock - Bloomington stock ID # 35785). The Gal4 driver line used to drive expression of shRNA constructs is *P(mat-tub-Gal4)mat67; P(mat-tub-Gal4)mat15* (originally generated by Daniel St. Johnston). *brat* mutant stocks were *brat^fs1^/CyO* and *Df(2L)TE37C- 7/CyO*; *brat* mutant flies were obtained by crossing the two stocks and collecting non-Cy progeny.

2.2 Transgenic reporters

Reporters were constructed in modified pCaSpeR vectors containing an attB site for site-directed transgenesis using the attP-attB-φC31 integrase method (Bischof J, et al. 2007). All constructs were inserted at the attP40 site on the *Drosophila* second chromosome. Constructs were either driven by the *aTub84B* or *smg* promoters cloned as Ascl (5') to NotI (3') fragments. All *aTub84B* promoter-driven reporters retained the endogenous *aTub84B* intron which has been shown to contain promoter/enhancer elements (O'Donnell KH 1994), and which lies immediately downstream of the *aTub84B* 5' UTR and ATG start codon. The fragment containing the *aTub84B* promoter, 5' UTR, start codon and intron was originally amplified from genomic DNA...
isolated from fly strain BL-2057 (Bloomington) as a ~ 2.7 kb fragment using primers (forward primer – 5'-CTTACCGATGTCGACGAAGAGG-3' and reverse primer 5'-CTGTGGATGAGGAGGAAGGGA-3'). This fragment has previously been shown to drive the maternal expression of various reporter genes in the fruit fly embryo (Jennifer Semotok, unpublished data). The endogenous \(\alpha\)Tub84B start codon was eliminated through QuickChange PCR using primers (sense primer – 5'-TTCCAATAAAAAACTCAATGTGGTGAGTACTTTTTTTTTTATTGGAA-3' and anti-sense primer – 5'-TTTTTTTAAAGTACTCACCACATTGAGTTTTTATTGGAA-3'). All smg promoter-driven constructs retained the endogenous intron located in the smg 5' UTR (isoforms RA and RD in Flybase). The smg promoter and 5' UTR fragment is approximately 6.5 kb and was originally amplified from genomic DNA.

Reporters contained either the eGFP or smg A647H ORF cloned as NotI (5') to SwaI (3') fragments downstream of the \(\alpha\)Tub84B or smg promoter fragments, with their own Kozak-like consensus sequences (‘CGAG’) and start codon. Downstream of the ORF, either the \(\alpha\)Tub84B or smg 3' UTRs and downstream genomic sequence were cloned as SwaI (5') to SbfI (3') fragments. The \(\alpha\)Tub84B 3' UTR fragment is 815 nucleotides long, and includes the 287 nucleotide \(\alpha\)Tub84B 3' UTR and 561 nucleotides of downstream genomic sequence. The smg fragment is 2217 nucleotides long, and includes the smg 3' UTR (of which there are several isoforms, the longest of which is 959 nucleotides) and approximately 1 kilobase of downstream genomic sequence.

For the TGT reporters bearing PBEs, a BamHI site was engineered at an endogenous BstEII site in the \(\alpha\)Tub84B 3' UTR, and PBEs were cloned into this site. PBEs were cloned as oligonucleotide pairs with BamHI cohesive ends. To clone fragments of the smg 3' UTR into TGT reporters, the \(\alpha\)Tub84B 3' UTR was further modified to include a Xhol site upstream of the
BamHI site, and fragments of the smg 3’ UTR were amplified with XhoI and BamHI sites and cloned. PBE- versions of smg 3’ UTR-bearing reporters were generated through overlap extension PCR.

2.3 Microarrays

Gene expression profiling in pum mutant embryos was performed by first collecting embryos (from pum mutant and w\textsuperscript{1118} females). Embryos were collected on apple juice agar for 1 h and then plates were further incubated as appropriate to obtain 0-1h, 1-2h, 2-3h, 3-4h, 4-5h old embryos. Three replicates were performed for both wild-type and pum mutant embryos for the 0-1h, 1-2h, 2-3h and 3-4h time points and two replicates were performed for the 4-5h time point. Embryos were dechorionated in bleach, collected and washed in an aspirator unit with 0.1% Triton X-100, and stored at -80°C in TRIzol (Invitrogen) or TRI reagent (Sigma-Aldrich) until they were processed for total RNA.

Total RNA was extracted using TRIzol or TRI reagent according to the manufacturer’s instructions, with minor modifications. Briefly, embryos were disrupted with a pestle in a microfuge tube in 400 µl of TRIzol. Homogenates were then spun at 13 K for 10 min to pellet cellular debris. The aqueous phase was then extracted twice with chloroform, followed by RNA precipitation with glycogen and ethanol. Samples were resuspended in DEPC-treated H\textsubscript{2}O and quantitated with a NanoDrop spectrophotometer. cDNA synthesis was performed with random primers as described (Kapranov P, et al. 2007). Custom Agilent microarrays (AMADID number 019871) were used containing ~44,000 probes, representing 12396 genes from D. melanogaster release 3.41. The array was designed using OligoPicker software (Wang X and B. 2001). Most
genes were represented by 3 probes on the array. Microarray slides were scanned with an Agilent High-Resolution C Scanner and quantified using ImaGene software.

Raw data was then RMA normalized and log2-transformed using the ArrayStar program. Normalized data were then subjected to one-class significance of microarray (SAM) (Tusher VG, et al. 2001) analysis using the Multi-experiment Viewer (MeV) software (Saeed AI, et al. 2006; Saeed AI, et al. 2003). Any gene having an FDR <5% in this analysis in at least one time point for at least one genotype was defined as expressed. The normalized data for each of these genes was compared in wild-type and pum mutant embryos using a two-class SAM analysis in the MeV software package. Differentially expressed genes were defined as those with an FDR<5% and fold difference in wild-type versus pum mutant of >1.5 fold.

Gene expression profiling in ago1 knockdown embryos was performed essentially as above, with minor modifications. ago1 and control knockdown embryos were collected from females expressing the Gal4 driver mated to males carrying either the ago1 or control shRNA constructs. Total RNA samples were collected and labelled cDNAs were generated following the Nimblegen user’s manual. Samples were hybridized to custom-designed Drosophila 12 x 135K NimbleGen arrays (GEO platform number: GPL8593). Data was processed, normalized and analyzed as above.

2.4 Motif enrichment analysis

The Drosophila melanogaster (BDGP5.4) transcript sequences were downloaded from Ensembl using BioMart (http://www.biomart.org/). When there were multiple isoforms for a gene, the longest isoform was analyzed. The motif enrichment test was performed using the RNA-READ pipeline (RNA Regulatory Elements Analysis and Discovery) (Xiao Li and Quaid...
Morris, unpublished). Specifically, RNA-READ searches a collection of previously defined motifs to find the motifs whose involvement in the regression significantly improves the fitting to the data, compared to ones based on the control features alone. The comparisons between the motif and the control features were restricted to the 5' UTR, the ORF or the 3' UTR region of transcripts.

We scored each regulatory region using a given motif by summing the accessibility of all the target sites, where a target site was defined as a perfect match to the IUPAC representation of the motif and the accessibility of a target site was defined as the average single-base accessibility of the bases in the site. A score of zero was assigned to those transcripts whose 5' UTR, ORF or 3' UTR regions did not contain a motif match. The single-base accessibility was assessed using RNAplfold with parameter settings $W=80$, $L=40$ and $U=1$. Although the analysis was applied to the specific regulatory region, the entire transcript was input into RNAplfold to ensure correct folding close to the start and stop codons.

### 2.5 GO term analysis

Lists of differentially expressed genes were compared to expressed genes that were not differentially regulated using the DAVID algorithm (Huang da W, et al. 2009; Huang da W, et al. 2009). For components of the VTPase and proteasome complexes, microarray data was mined for expression data for components not originally detected as differentially expressed.
2.6 RT-qPCR primers for \textit{smg} and \textit{RpL32}:

Reverse transcriptions for RT-qPCR were performed with SuperScript II or III reverse transcriptase (Invitrogen) according to the manufacturer’s instructions, using either random primers (Fermentas) or gene specific primers (\textit{smg} – 5’-GGGTGTGGCAGGCAATTTAG-3’, \textit{RpL32} – 5’-CGTTGTGCACCAGGAACTTCT-3’) and 50 ng of total RNA extracted with TRIzol reagent (Invitrogen) or TRI reagent (Sigma-Aldrich) from homogenized embryos. qPCR was performed with Power SYBR Green PCR Master mix (Applied Biosystems), using primers for \textit{smg} (forward primer 5’-GGCCAAACAATGGGAACAT-3’ and reverse primer 5’-GTTGGAGTTTTTGCGGTGT-3’) and \textit{RpL32} (forward primer 5’-AGTCGGATCGATATGCTAAGCTG-3’ and reverse primer 5’-CGATGTTGGGATCATCAGATCAGATCTG-3’).

2.7 Northern and Western blots

For Northern blots, embryos were collected, dechorionated and washed as for microarray experiments. Total RNA was isolated using TRIzol or TRI reagent from homogenized embryos, and 4 μg of total RNA was loaded per lane onto 1 % agarose, 1 X MOPS, 1 M formaldehyde denaturing gels. RNA was blotted onto nitrocellulose membrane, UV cross-linked using a Stratalinker and the Autocrosslink setting and pre-hybridized in Church’s buffer (250 mM sodium phosphate pH 7.2, 7% SDS, 1% BSA, and 1mM EDTA) for 1h at 65 °C. Membranes were probed with $^{32}$P-labelled DNA probes in Church’s buffer overnight, washed 8 times with 1 X SSC + 0.2 % SDS and exposed on Phosphoimager Screens (GE). Screens were imaged using a Typhoon scanner (GE) and band intensities were quantitated using ImageJ (NIH).
To generate embryo extracts for Western blots, embryos were dechorionated and washed with 0.1% Triton X-100, and were then further washed with H2O followed by lysis buffer (100 mM KCl, 2 mM magnesium acetate, 30 mM HEPES-KOH pH 7.4). Embryos were then collected in microfuge tubes with lysis buffer supplemented with 1 mM AEBSF, 2 µg/mL leupeptide, 2 mM benzamidine, 2 µg/mL pepstatin A and 1 mM DTT. Crude embryo extracts were then obtained by disrupting embryos with a pestle, spinning at 13K for 15 min, and transferring aqueous phase to a clean microfuge tube. Extracts were stored at -80°C until use.

For Western blots, the protein concentration in crude embryo extracts were determined using the Bio-Rad protein assay reagent. To probe for SMG, 7.5 µg of extract was loaded per lane, and for beta-tubulin, 0.75 µg of extract was loaded per lane. In addition, standard curves were run for both SMG and beta-tubulin for each experiment, consisting of 3.75 and 15 µg of 0-1 and 2-3 hour extract for SMG blots and 0.375 and 1.5 µg of 0-1 hour extract for beta-tubulin blots. Samples were run on 8% SDS-PAGE gels and blotted onto nitrocellulose using a tank transfer system (Bio-Rad). Blots were then washed in PBS, blocked in PBS + 0.05 % Tween 20 + 1% dry milk powder for 1 h. Primary antibodies used were: guinea pig anti-SMG (Tadros W, et al. 2007) which was used at a dilution of 1:10,000; mouse anti-β-tubulin E7 (Developmental Studies Hybridoma Bank, Iowa City) which was used at a dilution of 1:10,000; mouse anti-AGO1 antibody 1B8 (Miyoshi K, et al. 2005) (1:1000, kind gift of Haruhiko Siomi). Incubations with primary antibody were carried out overnight at 4°C. The following day, blots were washed in PBS + 0.05 % Tween 20, exposed to the appropriate HRP-conjugated secondary antibody, washed again, and developed with ECL reagent (GE Healthcare). Blots were imaged on a VersaDoc (Bio-Rad), and bands were quantitated with ImageJ. Blots for AGO1 were performed in an identical manner (same amount of extract as for SMG and gel running and blotting conditions).
2.8 Purification of GST-PUM RBD

The GST-tagged Pum RBD has previously been described (Wharton RP, et al. 1998). It was expressed in BL21 *E. coli* bearing the pRARE plasmid (Novagen). Cells were grown in 2 X YT medium until an OD 0.6-0.8 was reached, and protein expression was then induced with 1 mM IPTG. Cells were incubated with shaking at 18°C overnight to allow for protein expression. Cells were then lysed with B-PER (Thermo Scientific) as per the manufacturer’s instructions. GST-PUM RBD was then bound to glutathione beads and washed first with wash buffer 1 (500 mM NaCl, 0.0007% β-mercaptoethanol) and then with wash buffer 2 (100 mM NaCl, 0.1 mM EDTA, 0.0007% β-mercaptoethanol, 20 mM HEPES pH 7.5). Protein was eluted in elution buffer (250 mM NaCl, 30 mM glutathione, 10% glycerol, 0.0007% β-mercaptoethanol, 50 mM Tris-HCl pH 8.8), quantitated using a Bradford assay, aliquoted, and stored at -80°C until use.

2.9 Gel shifts

Gel shift reactions were performed in a total volume of 10 µl. Protein (150 ng of GST-Pum-RBD in 5 µl elution buffer), competitor RNA (1 µl in H2O) and gel shift buffer (3 µl of 100 mM KCl, 25 % glycerol, 1 µg/µl heparin, 5 mM DTT, 100 mM HEPES-KOH, pH 7.4) were pre-incubated for 5 min, and then 5 ng of 32P-labelled RNA probe was added (1 µl in H2O). Reactions were incubated for a further 10 min, supplemented with 2 µl of 50% glycerol, and then run on 4 % native polyacrylamide gel (0.25 X TBE, 5 % glycerol). Gels were then dried and exposed on Phosphoimager Screens (GE). Screens were imaged using a Typhoon scanner and both shifted and unshifted band intensities were quantitated using ImageJ. “No competitor” and “no protein” controls were run on each gel. The molar excess of competitor giving 50% competition was determined by calculating the ratio of shifted to unshifted signal for each lane.
(setting the “no protein” ratio to 0% and the “no competitor” ratio to 100%), plotting the resulting values vs. the fold molar excess competitor, and then generating a line of best fit.

2.10 Statistical analysis

Statistical analysis of microarray experiments is described in section 2.3. Differences between either mRNA or protein expression levels (Northern or Western blot) were assessed in Excel using the Student’s t-test. The statistical significance of overlaps between various gene lists were assessed by Fisher’s exact test, using the fisher.test() function in the ‘stats’ package in R.
3 Results

3.1 Attributions

The work described in this chapter was performed by Alexander Marsolais, with the following exceptions:

1) Mariana Kekis (co-supervised by Howard Lipshitz and Timothy Hughes) collected embryos, isolated total RNA, labelled samples and performed hybridizations for microarray analysis of gene expression in *pum* mutant embryos.

2) Xiao Li (co-supervised by Howard Lipshitz and Quaid Morris) performed RNA-READ analysis of the list of genes up- and down-regulated in *pum* mutant embryos.

3.2 PUM is a major regulator of mRNA stability in the early *Drosophila* embryo

To test the role of PUM in mRNA degradation, two strong *pum* mutant alleles were employed: *pum*<sup>ET7</sup> (a nonsense allele that truncates PUM prior to its RBD) (Forbes A and Lehmann R 1998) and *pum*<sup>MSC</sup> (a chromosomal inversion that also removes the RBD of PUM) (Barker DD, et al. 1992). Total RNA was harvested from mutant *pum*<sup>ET7</sup>/ *pum*<sup>MSC</sup> mothers (hereafter referred to as *pum* mutant embryos) as well as to similarly staged wild-type embryos by Mariana Kekis, who subsequently labelled and hybridized these samples to microarrays. I then used Significance Analysis of Microarrays (SAM) (Tusher VG, et al. 2001) to identify mRNAs that were significantly up- or down-regulated with an FDR <5% and a fold-change of 1.5-fold at each time-point. Figure 4 shows the levels of each expressed mRNA in *pum* mutant versus wild-type embryos, with those significantly up- or down-regulated indicated as red and blue points, respectively. While only a small number of genes are mis-regulated in *pum* mutant embryos
Figure 4. PUM regulates hundreds of transcripts after the onset of zygotic transcription. Plots show RMA-normalized signal intensity of all transcripts in pum mutant versus wild-type embryos at the indicated time intervals of embryogenesis. Each indicated transcript was defined as expressed in at least one time point and in at least one genotype, as described in the Materials and Methods. mRNAs that were at least 1.5 fold up- or down-regulated with an FDR of <5% in pum mutant embryos are indicated in red or blue, respectively. Dashed lines indicate 1.5 fold increase or decrease in expression, while the solid diagonal line represents no change. Data from 0-1, 1-2, 2-3 and 3-4 hour old embryos represent three biological replicates, while data from 4-5 hour old embryos represent two biological replicates.
during the first three hours of development, several hundred were found to be both up- or down-regulated during the 3-4 and 4-5 time-points.

Puf proteins are generally thought to negatively regulate their targets. Since I wished to first examine putative direct targets of PUM, I began by examining those genes that were up-regulated in \textit{pum} mutant embryos, as genes negatively regulated by PUM would be predicted to be up-regulated in \textit{pum} mutant embryos. Our SAM analysis of the microarray data identified 21 transcripts that were up-regulated during 0-1, 1-2, and/or 2-3 hour old embryos. In contrast, in 3-4 hour old embryos I found 243 up-regulated genes, and in 4-5 hour old embryos I found 396 up-regulated genes (Fig. 5A). The list of genes up-regulated in 3-4 and 4-5 hour old embryos overlapped significantly (Fisher’s exact test, $P < 6 \times 10^{-102}$), and taking the union of both results in a list of 501 genes up-regulated at some point during the fourth and fifth hours in \textit{pum} mutant embryos (Fig. 5A). Taking the union of all five time-points results in a marginally longer list of 510 total up-regulated genes. These data indicate that PUM largely regulates the expression of transcripts after zygotic transcription initiates, consistent with the proposed role of PUM in a late mRNA degradation pathway that requires zygotic transcription (Thomsen S, et al. 2010).

We further analyzed the expression of the 510 mRNAs that are up-regulated in \textit{pum} mutant embryos by dividing their levels in \textit{pum} mutant by their levels in wild-type embryos and subjecting the resulting ratios to hierarchical clustering (Fig. 5B). In this analysis, a \textit{pum} mutant/wild-type value of 1 indicates no difference between \textit{pum} mutant and wild-type embryos, while values greater or smaller than 1 indicate up- or down-regulation in \textit{pum} mutant embryos, respectively. As with the SAM analysis, this approach highlights the fact that most genes that are up-regulated in \textit{pum} mutant embryos are stabilized after the onset of zygotic transcription, with the number of genes differentially regulated increasing between the 3-4 and 4-5 hour time-points.
A. **Genes up-regulated in *pum* mutant embryos**

- **Graph**
  - Time-point: 0.5h, 1h, 1.5h, 2h, 3h, 4h, 4.5h
  - # of genes: 0, 0, 0, 105, 138, 258

- **Venn diagrams**
  - Up-regulated 3-4h: 105 genes
  - Up-regulated 4.5h: 258 genes
  - Overlap: 138 genes
  - **Total genes detected = 5643**

B. **Heatmap**

C. **Table**

<table>
<thead>
<tr>
<th>Motif enriched</th>
<th>trans factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>GUAUUA</td>
<td>miR-983/PUM</td>
</tr>
<tr>
<td>UGUAAHK</td>
<td>PUM</td>
</tr>
<tr>
<td>UGUAMRK</td>
<td>PUM</td>
</tr>
<tr>
<td>UGUANAAUA</td>
<td>PUM</td>
</tr>
</tbody>
</table>

D. **Venn diagrams**

- **Up-regulated 3-4h + 4.5h**
  - Total genes detected = 5624
  - Gerber et al., PUM-bound: 61 genes
  - **P = 2.33 x 10^{-44}**

- **Up-regulated 3-4h + 4.5h**
  - Total genes detected = 4800
  - Laver et al., PUM-bound: 88 genes
  - **P = 2.81 x 10^{-8}**
Figure 5. PUM destabilizes hundreds of transcripts after the onset of zygotic transcription and destabilized transcripts are enriched for PBEs. (A) Histogram showing the number of genes significantly up-regulated in pum mutant embryos, as determined in Figure 4 (left). Genes up-regulated in 3-4 and 4-5 hour old pum mutant embryos were found to overlap significantly (Fisher’s exact test). (B) Hierarchical clustering of gene expression for transcripts up-regulated in pum mutant embryos. Differences in gene expression were visualized by dividing expression levels in pum mutant embryos by WT embryos. (C) Computational analysis demonstrated enrichment of potential PBEs within the 501 genes that are upregulated in pum mutant embryos. (D) Genes up-regulated in pum mutant embryos (this study) and PUM-bound mRNAs (Gerber AP, et al. 2006; Laver JD, et al. 2015) show significant overlap. Fisher’s exact test P values and the total number of expressed genes are indicated.
3.3 PUM negatively regulates mRNA stability through direct binding

We wished to assess whether our list of genes up-regulated in pum mutant embryos largely represents direct targets of PUM. Xiao Li performed a search for motifs enriched in our combined list of 501 genes up-regulated in 3-4 and 4-5 hour old pum mutant embryos. As predicted for direct PUM targets, the top enriched motifs are published consensus PBE sequences (Gerber AP, et al. 2006; Ray D, et al. 2013) which all include the core ‘UGUA’ motif (Fig. 5C).

I also compared our list with published lists of mRNAs bound by PUM in the Drosophila embryo (Fig. 5D). I detected significant overlaps between our list and those of Gerber et al. (Gerber AP, et al. 2006) and Laver et al. (Laver JD, et al. 2015) (Fisher’s exact test, $P < 3 \times 10^{-44}$ and $3 \times 10^{-8}$ respectively). Gerber et al. also identified transcripts bound by PUM in fly ovaries, and our list of up-regulated genes overlaps significantly with this list as well (Fisher’s exact test, $P < 4 \times 10^{-12}$). I therefore conclude that our list of genes up-regulated in pum mutant embryos is highly enriched for direct targets of PUM, and that PUM degrades hundreds of transcripts through a late (zygotic) mRNA decay pathway.

3.4 PUM downregulates mRNAs encoding transcriptional and post-transcriptional regulators, as well as the V-ATPase and proteasome complexes

I next examined the biological functions of PUM degradation of transcripts in the early embryo. I compared our list of genes up-regulated in pum mutant embryos at 3-4 and 4-5 hours with data from Gerber et al. which showed that PUM binds to mRNAs involved in anterior-posterior patterning (Gerber AP, et al. 2006). Of the original list of 10 such transcripts identified
as PUM targets by Gerber et al., I found 8 (cad, exu, bcd, aret, bicC, osk, smg, hh) were stabilized in *pum* mutant embryos, confirming that PUM is an important regulator of early body patterning in the embryo. I then searched the lists of genes up-regulated in *pum* mutant embryos at 3-4 and 4-5 hours for GO term enrichment using the DAVID tool (Huang da W, et al. 2009; Huang da W, et al. 2009) (Table 1). Only terms with an FDR <10% and a Benjamini *P* value of <0.1 were considered for downstream analysis. I obtained similar results for the 3-4 and 4-5 hour lists, as expected since the lists are highly similar. However, I found more terms at the later time point (likely because a larger number of transcripts were found to be up-regulated at 4-5 hours), and therefore focused on the 4-5 hour list for my analysis. One of the top terms detected was “negative regulation of translation”, and included the genes smg, pek, rpr, bcd, cup, and *aret*. Two of the genes in this category (smg and bcd) overlapped with the genes involved in body patterning discussed above. I found an enrichment for “vacuolar proton-transporting V-type ATPase complex” and related terms; interestingly, Gerber et al. (Gerber AP, et al. 2006) also found an interaction between PUM and transcripts encoding the V-ATPase. Of the 33 V-ATPase subunits and accessory subunits annotated in the *Drosophila* genome, I detect six as being stabilized in *pum* mutant embryos (*Vha26, VhaAC39-1, Vha55, Vha68-2, Vha100-2 and VhaM9.7-b*). In addition, I found an enrichment for “proteasome core complex” and related terms. Of the 14 subunits of the proteasome core complex, I detect 5 as being stabilized in *pum* mutant embryos (*Prosa4, Prosβ3, Prosα5, Prosα6 and Prosβ6*), while 2 of the remaining 9 genes were found not to be expressed in our microarray data set (*Prosa1* and *Prosa2*). Finally, various metabolic enzymes involved in both arginine and proline metabolism and beta-alanine metabolism (CG4300, Gdh, Gs1, Argk, CG7433, CG6543, CG12262) were also found to be enriched in genes destabilized by PUM in the embryo.
<table>
<thead>
<tr>
<th>Term</th>
<th>Representative genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative regulation of translation</td>
<td><em>smg, pek, rpr, bcd, cup, aret</em></td>
</tr>
<tr>
<td>Vacuolar proton-transporting V-type ATPase complex</td>
<td><em>Vha26, VhaAC-39-1, Vha55, Vha100-2, VhaM9.7-B</em></td>
</tr>
<tr>
<td>Proteasome core complex</td>
<td><em>Prosβ3, Prosα4, Prosα5, Prosα6, Prosβ6</em></td>
</tr>
<tr>
<td>Zinc finger, ZZ-type</td>
<td><em>dah, Ada2b, CG11984, ref(2)P</em></td>
</tr>
<tr>
<td>Arginine and proline metabolism</td>
<td><em>CG4300, Gdh, Gs1, Argk, CG7433, CG6543, CG12262</em></td>
</tr>
</tbody>
</table>

Table 1. GO term enrichments for genes up-regulated in *pum* mutant embryos. Genes up-regulated in *pum* mutant embryos during the 4-5 hour time-point were examined for GO term enrichment using the DAVID algorithm (Huang da W, et al. 2009; Huang da W, et al. 2009). Similar terms that share genes are grouped together, and the genes that populate each category are listed.
3.5 SMG target mRNAs are down-regulated in *pum* mutant embryos

We reasoned that some of the biological consequences of disrupting PUM function in the embryo might also be uncovered by examining the genes that are down-regulated in *pum* mutant embryos. Similar to what I observed for up-regulated transcripts, I found only 14 genes were down-regulated in *pum* mutant embryos during the first three hours of embryogenesis (one of which one was the *pum* transcript itself, as expected since the *pum*<sup>MSC</sup> allele likely fails to produce *pum* mRNA). However, in 3-4 hour and 4-5 hour old embryos I detect 514 and 1556 down-regulated genes respectively, using an FDR of <5% and a fold change of >1.5 (Fig. 6A); the union of these two lists yields 1668 genes down-regulated in 3-5 hour old *pum* mutant embryos. In order to gain an insight into why these transcripts might be down-regulated, Xiao Li searched these mRNAs using the RNA-READ pipeline (unpublished) for enriched motifs and found these transcripts are potentially enriched for SREs, the binding site for the RBP SMG (Fig. 6B). SREs are stem loop structures, where the loop includes the core sequence CNGG. Xiao Li detected the sequences CNGG and CNGGN, suggesting that this group of transcripts is enriched for SMG target mRNAs. Consistent with the hypothesis that genes down-regulated in *pum* mutant embryos are enriched for SMG target transcripts, I see significant overlaps between these mRNAs and lists of transcripts that: 1) are bound by SMG (Chen L, et al. 2014), 2) require SMG for their translational repression (Chen L, et al. 2014) and 3) require SMG for their degradation (Tadros W, et al. 2007) (Fisher’s exact test, \( P < 4 \times 10^{-5}, 3 \times 10^{-19} \) and \( 3 \times 10^{-25} \), respectively) (Fig. 6C).
A. Genes down-regulated in *pum* mutant embryos

- **Time-point**
  - 0h
  - 1.5h
  - 3h
  - 4h
  - 5h

- **No. of genes**
  - 0-500
  - 500-1000
  - 1000-1500
  - 1500-2000

B. Motif enriched trans factors

<table>
<thead>
<tr>
<th>Motif</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNGG</td>
<td>SMG</td>
</tr>
<tr>
<td>CNGGN</td>
<td>SMG</td>
</tr>
</tbody>
</table>

C. Down-regulated in *pum* mutant 3-4h + 4-5h

- **Total genes detected = 4022**
- **P-value = 3.4 x 10^-5**

D. Down-regulated in *pum* mutant 3-4h + 4-5h

- **Total genes detected = 4774**
- **P-value = 2.3 x 10^-29**

E. Down-regulated in *pum* mutant 3-4h + 4-5h

- **Total genes detected = 3491**
- **P-value = 3.0 x 10^-25**

D. smg mRNA levels

- **% RNA remaining (log10)**
  - 0-1h
  - 1-2h
  - 2-3h
  - 3-4h
  - 4-5h

E. smg protein levels

- **% protein of starting levels**
  - 0-1h
  - 1-2h
  - 2-3h
  - 3-4h
  - 4-5h

**wild-type**

**pum mutant**
Figure 6. In *pum* mutant embryos, SMG protein persists after the onset of zygotic transcription and down-regulated transcripts are enriched for SMG target transcripts. (A) Number of genes whose transcripts are down-regulated in *pum* mutant embryos per time-point as defined in Figure 4. (B) Computational analysis demonstrated enrichment of potential SREs within the 1668 genes that are down-regulated in *pum* mutant embryos. (C) Comparisons of list of genes down-regulated in *pum* mutant embryos (this study) to lists of SMG-targeted mRNAs (Chen L, et al. 2014; Tadros W, et al. 2007). Fisher’s exact test $P$ values and the total number of expressed genes are indicated. (D) Total RNA was harvested from wild-type and *pum* mutant embryos at the indicated time intervals and subjected to RT-qPCR (with reverse transcription performed with random primers) to measure levels of *smg* mRNA. After normalizing *smg* mRNA levels using *RpLP2* mRNA as a loading control, the amount of *smg* mRNA in 0-1 hour old embryos was set to 100% for each genotype. The *smg* transcript is significantly stabilized in the 3-4 and 4-5 hour time-points in *pum* mutant embryos (** = $P < 0.01$, student’s t-test). (E) SMG protein persists in *pum* mutant embryos after the onset of zygotic transcription. Western blots (left) on embryo extract from wild-type and *pum* mutant embryos were performed, and SMG protein levels were quantified and normalized to tubulin levels (right) (** = $P < 0.01$, student’s t-test). (D+E) Quantification represents the results of three biological replicates and error bars indicate standard error.
3.6 PUM regulates SMG expression

What mechanism underlies the down-regulation of SMG target mRNAs in 3-5 hour old *pum* mutant embryos? In wild-type embryos SMG protein is expressed at its highest levels during the first three hours of embryogenesis, while at subsequent time points its levels are substantially reduced (Benoit B, et al. 2009). The down-regulation of SMG target mRNAs in 3-5 hour old embryos suggests the possibility that PUM protein downregulates SMG protein expression. The fact that *smg* mRNA co-immunoprecipitates with PUM (Gerber AP, et al. 2006) suggests that PUM could regulate SMG protein expression through a direct interaction with *smg* mRNA. Consistent with this possibility, our microarray data show that in wild-type embryos, *smg* mRNA is degraded starting in the 3rd hour of embryogenesis, such that little is found in 3-4 and 4-5 hour embryos, while in *pum* mutant embryos *smg* mRNA is partially stabilized. Similar results were obtained using RT-qPCR (Fig. 6D).

We next tested whether SMG protein persists in *pum* mutant embryos via Western blot analysis. Consistent with previous work (Benoit B, et al. 2009), I found that SMG protein levels are maximal in 0-3 hour old embryos, while levels are substantially reduced in 3-4 and 4-5 hour old embryos. In contrast, in *pum* mutant embryos, SMG protein levels are elevated in all five time-points, with statistically significant increases in 2-3h and 3-4 hour old embryos (Fig. 6E).

Taken together, these data are consistent with a model whereby *smg* mRNA is a direct target of PUM, and in which the persistence of SMG in a *pum* mutant embryo results in inappropriate degradation of SMG target mRNAs.
3.7  PUM directly regulates smg mRNA through PBEs in the smg 3' UTR

The persistence of SMG protein and the down-regulation of many SMG target transcripts in pum mutant embryos suggest that smg mRNA is an important target of PUM. To investigate this further, I set out to map the cis-acting sequences required for smg mRNA degradation, first testing the roles of the smg 5' UTR, ORF and 3' UTR in degradation of the transcript. This analysis made use of a transgenic reporter construct that expresses an mRNA that carries the αTub84B 5' UTR, the eGFP ORF and the αTub84B 3' UTR. I refer to this reporter mRNA as TGT, where the first letter indicates the origin of the 5' UTR, the second letter indicates the origin of the ORF, and the third letter indicates the origin of the 3' UTR. Related reporters are labeled using the same convention (Fig. 7A). TGT mRNA is stable over the first five hours of embryogenesis, as assayed by Northern blot analysis (Fig. 7B,7C). Replacement of the tubulin 5' and 3' UTR with the smg 5' and 3' UTR in TGT, to give SGS, resulted in an mRNA which is degraded with kinetics similar to endogenous smg mRNA. SGT mRNA, where the tubulin 5' UTR was replaced with the smg 5' UTR, is modestly degraded in the fifth hour of embryogenesis. Switching the 3' UTR sequences to give TGS resulted in an mRNA that degraded with similar kinetics to endogenous smg mRNA, albeit not to the same extent. Taken together, these data suggest that the smg 3' UTR makes a major contribution to the degradation of smg RNA, while the smg 5' UTR makes a more modest contribution.

I was unable to obtain transgenic lines expressing a TST mRNA, where the smg ORF replaced the GFP ORF, possibly because elimination of regulatory elements in the smg UTRs resulted in construct lethality. Thus, I made use of a mutant version of the smg ORF that encodes a protein where Ala 642 is changed to His and is defective for RNA binding
Figure 7. smg mRNA decay requires the smg 3' UTR. (A) Transgenic reporters constructed to test different fragments of the smg transcript for destabilizing activity. These reporters are driven by either the smg or αTub84B promoters. They incorporate fragments of smg, αTub84B and GFP sequence in order to test the smg transcript for destabilizing elements. (B) Total RNA was harvested from embryos expressing the transgenic mRNAs described in (A) at the indicated time intervals and subjected to Northern blot analysis using GFP (TGT, SGT, TGS, SGS) or αTub84B (TS\textsubscript{mut}T) probe to detect transgenic mRNAs. Probing for the stable RplP2 mRNA served as a loading control. Note that the αTub84B probe used to detect the TS\textsubscript{mut}T reporter (arrow) also detects a αTub84B isoform (*). (D) After normalizing transgenic mRNA levels using the loading control, the amount in 0-1 hour old embryos was set to 100%. Data represents four biological replicates, with error bars indicating standard error. The percentage of different transgenic mRNAs remaining at the last time point is indicated.
(Aviv T, et al. 2003), to generate a TS\textsubscript{mut}T reporter, and found that this TS\textsubscript{mut}T mRNA is stable, suggesting that the smg ORF does not play a role in degradation of smg mRNA.

I next wished to confirm that smg mRNA is a direct target of PUM. I therefore tested the ability of the PUM RNA binding domain (RBD), purified from \textit{E. coli} as a GST fusion, to interact with the 5' and 3' UTRs of smg \textit{in vitro} in gel mobility shift assays. I began by examining the smg 5' UTR and found a single putative PBE which contains the core ‘UGUA’ core motif and exactly matches the previously published consensus sequence of UGUA(C/A/U)AUA (Gerber AP, et al. 2006) (Fig. 8A). A 72 nucleotide probe centered around the putative PBE was bound by the PUM RBD, and this interaction was efficiently competed by an unlabeled version of the same RNA, but not by an mutant version of this RNA where the ‘UGUA’ core motif was mutated (Fig. 8B). Note that gel shifts were conducted with a standardized amount of PUM RBD which shifted all of the labeled probe, and that the ability of unlabeled RNAs to compete for PUM binding was quantitated by determining the molar excess of competitor (compared to probe) that reduces the amount of probe shifted by 50%.

I then proceeded to test the ability of other fragments of the smg UTRs to compete for binding to the PBE-containing smg 5' UTR probe (Fig. 8C). I tested four overlapping fragments derived from the smg 5' UTR. Three of these fragments lack the identified PBE and failed to compete for binding to the PUM RBD while the one fragment that contained the identified PBE did compete. Thus, the smg 5' UTR has a single PBE. I then tested fragments of the smg 3' UTR. Two of these fragments competed poorly for PUM binding, at levels similar to those seen with the mutant version of the PBE-containing smg 5' UTR fragment. In contrast, the 3' most fragment, which contains all 10 ‘UGUA’ motifs found in the smg 3' UTR, competed more efficiently. Furthermore, mutation of all 10 of these motifs decreased the ability of this fragment to compete, suggesting that some or all of these ‘UGUA’ motifs are functional PBEs.
A

\begin{align*}
& \text{smg 5'} \text{ UTR} \\
& \text{smg 5'} \text{ 1-245} \\
& \text{smg 5'} \text{ 246-491} \\
& \text{smg 5'} \text{ 150-306} \\
& \text{smg 5'} \text{ 315-491} \\
& \text{smg 3'} \text{ UTR} \\
& \text{smg 3'} \text{ 1-412} \\
& \text{smg 3'} \text{ 413-959} \\
& \text{smg 3'} \text{ 200-466}
\end{align*}

B

\begin{align*}
\text{Wild-type} & \quad 276-345 \\
& \text{US} \\
2.6 \pm 0.7 \\
\text{Mutant} & \quad 276-345 \\
& \text{US} \\
39.5 \pm 13.4
\end{align*}

C

\begin{tabular}{|c|c|c|c|}
\hline
Fragment & UGUA core motif? & Wild-type & Mutant (UGUA > ACA) \\
\hline
smg 5' 1-245 & No & >50 & N/A \\
smg 5' 246-491 & Yes & 1.5 \pm 0.055 & ND \\
smg 5' 276-345 & Yes & 2.6 \pm 0.7 & 39.5 \pm 13.4 \\
smg 5' 150-306 & No & >50 & N/A \\
smg 5' 315-491 & No & >50 & N/A \\
smg 3' 1-412 & No & 19.0 \pm 1.9 & N/A \\
smg 3' 413-959 & Yes & 5.3 \pm 3.4 & 9.2 \pm 2.8 \\
smg 3' 200-466 & No & 29.1 \pm 7.3 & N/A \\
\hline
\end{tabular}

D

\begin{align*}
tubulin 5' UTR & \quad eGFP ORF & tubulin 3' UTR \\
\text{mRNA levels} & \quad \text{TGT + 1-907} \\
\text{time-point} & \quad \text{TGT + 1-907 PBE-}
\end{align*}
Figure 8. PUM directly regulates smg mRNA through PBEs located in its 3' UTR. (A) Location of putative PBEs in the smg UTRs (‘UGUA’ core motifs). The only PBE which perfectly matches the UGUA(C/A/U)AUA consensus motif is located in the 5' UTR, and a fragment containing this site was used as a probe in a gel shift assay. (B) The mobility of the unbound probe fragment was determined by running it in the absence of PUM RBD. In the presence of the PUM RBD, the probe is shifted (US = unshifted, S = shifted, NP = no protein, P = protein). Binding of the PUM RBD to a given fragment was then tested by titrating increasing amounts of unlabeled fragment and measuring competition with the radiolabelled probe. The wild-type 5' nt 276-345 fragment competes with itself efficiently, but mutation of the PBE blocks competition. The molar excess of competitor that reduces the amount of shifted probe by 50% is indicated. (C) Summary of results of gel shift assays with fragments of the smg 5' and 3' UTRs, where the molar excess of competitor that reduces the amount of shifted probe by 50% is indicated for wild-type and PBE mutant versions of the relevant fragments of the smg 5' and 3' UTRs. Values represent 3 independent replicates and errors represent standard error. (D) The TGT reporter was modified to include either a wild-type version of nt 1-907 of the smg 3' UTR (TGT + nt 1-907) or a version with all potential PBEs mutated (TGT + nt 1-907 PBE-). Total RNA from transgenic flies expressing these reporters was harvested at the indicated time intervals and transgenic reporter mRNA levels were quantitated as described in Figure 7. The graph shows the results of four biological replicates, where error bars indicate standard error. The P values from Student’s t-test are indicated by * (P < 0.05) and ** (P < 0.01).
I next tested whether the ‘UGUA’ motifs within the \textit{smg} mRNA were necessary for regulation \textit{in vivo} (Fig. 8D). Given that the \textit{smg} 3’ UTR is largely sufficient for the decay kinetics observed for endogenous \textit{smg}, I restricted my analysis to the \textit{smg} 3’ UTR. I constructed new reporters utilizing the TGT construct as backbone, into which I cloned different fragments of the \textit{smg} 3’ UTR. Similar to the TGS reporter, the TGT + \textit{smg} 3' reporter recapitulated a decay pattern similar to endogenous \textit{smg} mRNA. Mutation of the ‘UGUA’ motifs within the \textit{smg} 3’ UTR resulted in significant stabilization of the reporter (TGT + PBE- \textit{smg} 3' UTR). In fact, the degradation of the TGT + PBE- \textit{smg} 3' UTR mRNA was very similar to endogenous \textit{smg} mRNA in \textit{pum} mutant embryos. I therefore conclude that PUM is necessary for regulation of \textit{smg} mRNA, and that this regulation occurs through direct binding to PBEs in the \textit{smg} 3’ UTR.

3.8 PUM can induce transcript decay through an early pathway

PUM protein is maternally contributed (Macdonald PM 1992), and is therefore present during both the early and late phases of transcript decay in the embryo. In light of this, our observation that most PUM-mediated transcript degradation (including that of \textit{smg} mRNA) occurs with late kinetics raises an important question; how is PUM-mediated degradation temporally regulated, such that most PUM targets are only degraded after the onset of zygotic transcription?

In order to better understand the mechanism of PUM-mediated regulation, I first wished to determine what affect PBEs have on an RNA in the absence of other \textit{cis}-acting elements. To this end, I compared the stability of transgenic TGT mRNA carrying 5 high-affinity PBEs (5xUGUAAUA) inserted into its 3’ UTR (TGT-5xPBE+) to TGT mRNA carrying the sequence complementary to the 5xPBE+ insert (5xACAUUUUAU – 5xcomp) (Fig. 9A). I first verified that
Figure 9. High affinity PBEs mediate degradation through a maternal pathway. (A) Artificial binding site fragments designed to test PUM regulation in the absence of other cis elements. The wild-type binding site (UGUAAAUA) corresponds to the smg 5' UTR PBE and matches the published consensus 8-mer motif for a PBE, UGUA(C/AU)AUA (Gerber AP, et al. 2006). The complement of this sequence (5xcomp) is used as a control. (B) Gel shift assays were performed using a radiolabelled probe corresponding to the 5xPBE+ fragment and included the indicated molar excesses, relative to the probe, of unlabeled 5xPBE+ and 5xcomp RNAs. (C) Transgenic flies were generated carrying the TGT reporter with either the 5xPBE+ and 5xcomp fragments inserted into the TGT 3' UTR. Total RNA was harvested at the indicated time intervals from wild-type, pum mutant or png mutant embryos and/or UEs, and transgenic reporter mRNA stability (along with that of endogenous smg mRNA) was assayed by Northern blot as described in Figure 7. (D) All possible 8-mer sequences beginning with a UGU core motif (UGUANNNN – 256 permutations) were analyzed for in vitro binding affinity to PUM using data from the RNAcompete project (Ray D, et al. 2009; Ray D, et al. 2013). The wild-type PBEs in the 5xPBE+ reporter are ranked 2 among the 256 permutations in the RNAcompete data, where the site with rank 1 has the highest affinity. The relative rank affinities for all putative PBEs in the smg 3' UTR are depicted where they occur in the UTR. (E) Additional transgenic TGT constructs were generated with either three rank 2 sites (i.e. the same site as in the 5xPBE+ reporter), or five rank 30 or five rank 195 sites, and the stability of the resulting mRNAs were assayed as in figure X by Northern blot. Quantification in (C) and (E) represents the results of four biological replicates, error bars indicate standard error and the percentage of mRNA remaining in the last time point for different mRNAs are indicated.
the 5xPBE+ sequence interacts with the PUM RBD in a gel shift assay and that an unlabeled version of this RNA competes for binding, while the 5xcomp RNA does not (Fig. 9B). Next I confirmed that TGT-5xcomp RNA was stable over the first five hours of embryogenesis (as assayed via Northern blot) confirming that insertion of sequences into the 3' UTR of TGT does not cause its destabilization. In contrast, TGT-5xPBE+ mRNA was unstable in wild-type embryos, and I confirmed that this degradation was mediated by PUM as the mRNA was stabilized in pum mutant embryos. Interestingly, TGT-5xPBE+ mRNA was degraded with early kinetics, suggesting its destabilization required only maternal gene products. I confirmed this by showing that the kinetics of TGT-5xPBE+ mRNA decay was very similar in fertilized and unfertilized eggs. This contrasts with the behavior of smg mRNA, which is degraded with late kinetics, and is therefore stabilized in unfertilized eggs (Fig. 9C). In addition, I examined the behavior of this reporter in png mutant embryos, in which the process of egg activation is compromised and which appear to have a general defect in maternal transcript destabilization (Tadros W, et al. 2003). Consistent with the notion that the TGT-5xPBE+ mRNA is degraded by a maternal decay mechanism, this transcript is stabilized in png mutant embryos.

The above data suggest that sequence differences between smg mRNA and the TGT-5xPBE+ transcript underlie their differential decay kinetics. One difference is that the TGT-5xPBE+ mRNA carries 5 high affinity PBEs while the smg 3' UTR, which harbors the bulk of the degradation activity of the smg mRNA, contains 10 PBEs which do not perfectly match the UGUAN(A)UA consensus motif. To assess the relative binding strengths of these various PBEs, I made use of data generated using the RNAcompete assay (Ray D, et al. 2009; Ray D, et al. 2013). Briefly, this assay involves mixing a complex pool of ~240,000 30-41mer RNAs that contains at least 16 copies of each 9mer sequence and at least 310 copies of each 8mer sequence with a GST-tagged RBP or RBD of interest. Protein/RNA complexes are then captured on
glutathione resin and bound RNAs are identified using microarrays. This assay is done in RNA excess, and thus the extent of capture of an RNA is a measure of its relative affinity for the protein of interest. Since the PUM RBD binds to linear 8-mer sequences (Wang X, et al. 2002), and because the 5' ‘UGUA’ core motif is the principle feature of published PBE consensus motifs (Gerber AP, et al. 2006), I used the RNAcompete data for PUM (Ray D, et al. 2013) to rank all 256 8-mers that begin with ‘UGUA’ (i.e., ‘UGUANNNN’ assigning a rank of 1 to the best binder) and then assessed the rank of all UGUA motifs within the smg 3' UTR. The PBE sequence I used in the TGT-5xPBE+ mRNA had a rank of 2 while those of the smg 3' UTR had lower ranks ranging from 26 to 239 (Fig. 9D) suggesting they have lower affinity for PUM, and by extension that this lower affinity might play a role in the timing of smg mRNA degradation.

To further explore the effect of PBE affinity on PUM function in the embryo, I generated a series of TGT reporters carrying reduced affinity PBEs, based on sites found in the smg 3' UTR (Fig. 9E). I found TGT mRNA carrying either five rank 30 PBEs or five rank 195 PBEs were both stable over the first five hours of embryogenesis as assayed by Northern blot. In addition, TGT mRNA carrying three rank 2 sites showed only very modest degradation over the same time course. Taken together these data suggest only several high affinity PBEs are, on their own, able to mediate mRNA decay in the early embryo, and then only during the early (maternal) phase. I therefore propose that low affinity PBEs in the smg 3’ UTR must function together with other cis-acting elements (and their corresponding trans-acting factors) to induce robust smg mRNA degradation during the late (zygotic) phase, since the requirement for zygotic transcription to induce the degradation of smg mRNA suggests that one or more of the trans-acting factors that targets smg mRNA for decay is zygotically expressed.
3.9 Mapping additional cis-elements required for smg mRNA decay

To map additional factors required for PUM-mediated degradation of smg mRNA, I set out to define a minimal fragment of the smg 3' UTR required for degradation. I divided the transcript into two halves, nt 1-438 (the 5' half of the UTR, that does not contain any PBEs) and nt 439-907 (which contains all ten potential UGUA PBEs within the smg 3' UTR) (Fig. 10A) and inserted these fragments into the TGT reporter construct (Fig. 10B). The reporter mRNA bearing nucleotides 1-438 of the smg 3' UTR was stable over the first 5 hours of embryogenesis, as assayed via Northern blot. Interestingly, the decay activity of the fragment containing nucleotides 439-907 of the smg 3' UTR was attenuated when compared to the full-length smg 3' UTR fragment (nucleotides 1-907). In addition, the degradation activity of nucleotides 438-907 was only slightly more robust than the PBE mutant version of the same fragment (nt 439-907, PBE-) or the PBE mutant version of the full-length smg 3' UTR (nt 1-907, PBE-). Taken together, these data argue that while PUM binding to the smg 3' UTR is necessary for robust degradation, other cis elements are required, and that some of these elements are located within nucleotides 1-438 of the 3' UTR. This region on its own has no degradation activity, suggesting that cis elements and trans factors interacting with both nucleotides 1-438 and 439-907 are required for full regulation of the transcript.

3.10 smg mRNA decay requires BRAT

Recent work from the Lipshitz and Smibert laboratories has shown that smg mRNA is bound by BRAT, and has determined the consensus sequence that BRAT recognizes (which consists of a core ‘UGUU’ motif) (Laver JD, et al. 2015). BRAT (previously thought to only function as a co-factor for PUM-mediated regulation) had recently been shown to directly bind
Figure 10. mRNA destabilization mediated by the smg 3' UTR requires PBEs and additional cis-elements in nucleotides 1-438. (A) Transgenic flies expressing TGT reporters containing fragments of the smg 3' UTR were generated. Where indicated the potential PBEs within the smg 3' UTR were mutated. Note that the TGT + nt 1-907 and TGT + nt 1-907, PBE- reporters are reproduced from Figure 8 for comparison. (B) RNA stabilities of the reporters in (A) were measured by Northern blot as in Figure 7. Asterisks refer to comparison between TGT + nt 1-907 and TGT + nt 439-907. Note that TGT + nt 1-907, PBE- and TGT + nt 439-907, PBE- are also significantly stabilized in the last two time-points. Quantification represents the results of four biological replicates, error bars indicate standard error and the percentage of mRNA remaining in the last time point for the different mRNAs is indicated. The P values from Student’s t-test are indicated by * ($P < 0.05$) and ** ($P < 0.01$).
RNA (Loedige I, et al. 2014), and make direct contact with the well-characterized PUM target, 
*hb* mRNA. Laver et al. detect *smg* mRNA in BRAT RNA immunoprecipitation (RIP) experiment (Laver JD, et al. 2015), suggesting BRAT binds to *smg* mRNA and that it may directly regulate it during *Drosophila* embryogenesis. Consistent with this, a search of the *smg* 3’ UTR indicates 13 ‘UGUU’ sequences (and therefore potential BRAT-binding sites) (Fig. 11A). In order to determine whether BRAT is required for *smg* mRNA degradation, I examined *smg* decay in embryos laid by *brat* FSl/Df(2L)TE37C-7 mothers (hereafter *brat* mutant embryos) via RT-qPCR. Consistent with a role for BRAT in *smg* mRNA decay, I found that *smg* mRNA was stabilized in *brat* mutant embryos (Fig. 11B). This confirms analysis of *brat* mutant embryos, which demonstrated a similar stabilization of *smg* mRNA (Laver JD, et al. 2015). Taken together, our data indicate that PUM and BRAT both contribute to the degradation of *smg* mRNA.

While important to the overall model of *smg* mRNA degradation, the identification of BRAT as a factor involved in the process fails to explain the zygotic nature of *smg* transcript decay. This is because BRAT is a maternally contributed factor, and is therefore present in the embryo hours before the transcript is targeted for turnover. I therefore wished to identify additional factors that might be expressed (or at least function) zygotically, in order to arrive at a more coherent model of *smg* mRNA regulation.

### 3.11 *smg* mRNA decay requires AGO1

Figure 11. smg mRNA degradation is dependent on AGO1 and BRAT. (A) Graphic of the smg 3' UTR showing the location of putative PBEs ('UGUA' sequences, lines) and BRAT binding sites ('UGUU' sequences, circles). (B) smg mRNA was measured in brat mutant embryos by RT-qPCR (with reverse transcription performed with gene specific primers) as described in Figure 6. (C) AGO1 protein levels were measured by Western blot in embryos expressing either a control or an ago1 siRNA. (D) AGO1 levels were normalized to tubulin and the amount of AGO1 protein in 0-1 hour old embryos was set to 100%. (E) smg mRNA stability was measured in control and ago1 KD embryos by RT-qPCR (with reverse transcription performed with random primers) as described in Figure 6. Quantification represents the results of three (C) or four (D and E) biological replicates, error bars indicated standard error and P values from Student’s t-test are indicated by * (P < 0.05) and ** (P < 0.01) or numerical values.
disrupted AGO1 expression in embryos using an AGO1 UAS-shRNA transgene. Expression of the short hairpin RNA (shRNA) was achieved by mating males carrying this transgene to females whose embryos express high levels of the Gal4 activator protein. Under these conditions, shRNA expression begins soon after zygotic transcription is permitted and AGO1 protein expression is inhibited around the time when the late mRNA decay pathways initiate (Fig. 11C and 11D). AGO1 protein levels increase as embryogenesis proceeds, and this newly synthesized AGO1 likely associates with newly synthesized zygotically expressed miRNAs. The processing and stability of miRNAs typically requires the presence of Argonaute protein that is not bound to a miRNA, since unloaded Argonaute proteins are a core component of the so-called RISC-loading complex (RLC), which processes pre-miRNAs into mature miRNAs (Winter J 2009). Therefore, the inhibition of AGO1 accumulation that is mediated by the AGO1 shRNA likely results in the down-regulation of zygotically-expressed miRNAs. Under these conditions, I found that smg mRNA is partially stabilized when AGO1 expression is knocked down (Fig. 11E). I therefore propose that zygotic miRNAs control the timing of smg mRNA decay and SMG protein clearance.

3.12 PUM, BRAT and AGO1 cooperate to regulate other targets in the Drosophila embryo

Our stability measurements of smg mRNA indicate that it is targeted by three different factors – PUM, BRAT and AGO1. smg mRNA may represent a key target of all three factors, since stabilization of smg mRNA in pum mutant embryos is correlated with an aberrant increase in SMG protein and down-regulation of SMG target transcripts in the post-zygotic embryo. However, I wished to determine at a global level whether these factors cooperate to regulate other transcripts in the embryo, and to what extent they can function independently of each other.
I undertook microarray experiments to compare mRNA expression in \textit{ago1} and control knockdown embryos. Figure 12 compares gene expression levels in these embryos. Note that the microarrays used employed transcript-specific probe sets; hence, the scatterplots in Figure 12 represent expression levels for specific transcripts, with certain genes represented multiple times. During the first four hours of embryogenesis, I observe few transcripts that are up- or down-regulated in \textit{ago1} knockdown embryos. In the 4-5 hour time-point, in contrast, several hundred transcripts were up-regulated in \textit{ago1} knockdown embryos, with a much smaller number down-regulated.

I next subjected the data to SAM analysis to generate a list of high-confidence targets of AGO1 in the early embryo (Fig. 13A). I applied both an FDR 5\% as well as a 1.5-fold change cut-off and filtered the resulting lists for unique genes. I detected <10 genes up- or down-regulated in the first three hours of embryogenesis. In the 3-4 hour time-point, I observed <10 genes up-regulated and 37 genes down-regulated. Finally, in the 4-5 hour time-point I observe the largest number of regulated genes, with 211 unique genes up-regulated and 36 unique genes down-regulated in \textit{ago1} knockdown embryos. Taken together, our microarray data suggest that our genetic manipulation in this system results predominantly in up-regulation of genes in the 4-5 hour time-point.

I first subjected the list of genes stabilized in \textit{ago1} knockdown embryos to GO term enrichment using the same approach as for the genes stabilized in \textit{pum} mutant embryos. After applying the previously used filters of FDR <10\% and a Benjamini \textit{P} value of <0.1, no significant terms were detected.

The \textit{miR-309} cluster of miRNAs encodes several miRNAs that are zygotically expressed, and whose deletion results in the up-regulation of several hundred transcripts in the early embryo (Bushati N, et al. 2008). I found a modest overlap between my list of 211 genes up-regulated in
Figure 12. Several hundred transcripts are up-regulated in 4-5 hour old ago1 knockdown embryos. Plots show RMA-normalized signal intensity of all transcripts in ago1 knockdown embryos versus control knockdown embryos at the indicated time intervals of embryogenesis. Each indicated transcript was defined as expressed in at least one time point and in at least one genotype, as described in the Materials and Methods. mRNAs that were at least 1.5-fold up- or down-regulated with an FDR of <5% in ago1 knockdown embryos are indicated in red or blue, respectively. Dashed lines indicate 1.5 fold increase or decrease in expression, while the solid diagonal line represents no change. Data from all time intervals represent results of three biological replicates.
Figure 13. PUM, BRAT and AGO1 cooperate to degrade transcripts during the late (zygotic) phase of decay. (A) Histogram showing the number of unique genes significantly up-regulated in ago1 knockdown embryos, as determined in Figure 12. (B-E) The list of genes up-regulated inago1 knockdown embryos was compared with: those up-regulated in miR-309 mutant embryos (Bushati N, et al. 2008) (B); those bound by PUM (Gerber AP, et al. 2006) (C); those bound by BRAT (Laver JD, et al. 2015) (D); and those upregulated in pum mutant (my work) and brat mutant embryos (Laver JD, et al. 2015) (E). The list of genes down-regulated inbrat mutant embryos (Laver JD, et al. 2015) was compared with lists of genes 1) bound by SMG, 2) translationally repressed by SMG and 3) degraded by SMG (Chen L, et al. 2014; Tadros W, et al. 2007) (F). The Fisher’s exact test P values and the total number of expressed genes are indicated.
ago1 knockdown embryos with genes found to be up-regulated miR-309 cluster null mutant embryos (Fig. 13B, Fisher’s exact test, $P < 0.05$). The relatively small number of genes common to both lists may reflect the differences in developmental staging (the miR-309 mutant data was derived from 2-3 hour old embryos, whereas the ago1 knockdown-stabilized transcripts were detected in 4-5 hour old embryos). It may also reflect the incomplete knockdown of AGO1 expression in my experiments. Also noteworthy is the fact that genes up-regulated in miR-309 null mutant embryos are not enriched in predicted binding sites for miR-309 cluster miRNAs (Bushati N, et al. 2008), which could suggest that many of the up-regulated mRNAs are not direct targets of these miRNAs.

I then compared the list of genes up-regulated in ago1 knockdown embryos to transcripts bound by PUM (Gerber AP, et al. 2006) in the Drosophila embryo. I found a significant overlap between the two lists (Fisher’s exact test, $P < 2 \times 10^{-26}$) (Fig. 13C). I likewise compared the list of genes up-regulated in ago1 knockdown embryos with the list of genes bound by BRAT (Laver JD, et al. 2015), and found a more modest, albeit significant, overlap (Fisher’s exact test, $P < 0.05$) (Fig. 13D).

I next wished to compare the list of genes up-regulated in ago1 knockdown embryos with the lists of genes up-regulated in pum and brat mutant embryos, to determine at a global level the prevalence of co-regulation between PUM, BRAT and AGO1 in regulating common transcripts. All three lists were found to significantly overlap with each other (Fig. 13E, Fisher’s exact test, pum and brat stabilized lists $P < 4 \times 10^{-68}$, pum and ago1 stabilized lists $P < 2 \times 10^{-73}$, brat and ago1 stabilized lists $P < 7 \times 10^{-60}$). I found 73 transcripts in common between all three lists, including smg mRNA. Taken together my data strongly support a model in which these three factors frequently cooperate, in pairs or all together, in late mRNA decay.
I have shown that in addition to causing the degradation of several hundred transcripts during the late (zygotic) wave of degradation, loss of PUM also results in up-regulation of SMG protein and a corresponding down-regulation of SMG target transcripts. Considering smg mRNA is a common target of PUM, AGO1 and BRAT, one prediction of the model would be that similar phenomena would be observed both in ago1 knockdown and brat mutant embryos.

I first examined the list of transcripts down-regulated in ago1 knockdown embryos for evidence of overlap with the lists of SMG target transcripts previously compared with those genes down-regulated in pum mutant embryos (Chen L, et al. 2014; Tadros W, et al. 2007). I found no significant overlaps. It is worth noting that the list of genes down-regulated in ago1 knockdown embryos is small (47 genes in total between 3-4 and 4-5 old embryos), possibly explaining the lack of overlap.

I then compared the list of genes down-regulated in both 3-4.5 and 4.5-6 hour old brat mutant embryos with the lists of SMG target genes described above. As for the list of genes down-regulated in pum mutant embryos, I find a significant overlap between all three (Fig. 13F). This result confirms the role of BRAT in smg mRNA regulation, and suggests that this may be a central role of BRAT in the early embryo as it appears to be for PUM.
4 Discussion and future directions

4.1 Discussion

4.1.1 PUM is a critical regulator of the post-transcriptional landscape of the early *Drosophila* embryo

PUM is a founding member of the Puf family of RBPs, which are conserved from yeast to humans (Quenault T, et al. 2011). PUM has well-characterized roles in germline maintenance in the *Drosophila* ovary, patterning of the embryo, as well as roles in learning and memory. In the current work I have demonstrated that, in addition to regulating the handful of PUM target mRNAs previously characterized, PUM directly down-regulates approximately 500 mRNAs in the early *Drosophila* embryo. In agreement with previous computational analyses (Thomsen S, et al. 2010), I show here that PUM is a major regulator of the zygotic mRNA degradation pathway, degrading almost all of its targets after the onset of zygotic transcription.

I have also shown that PUM indirectly up-regulates an even larger number of mRNAs (~1500) during the same period of development. This is likely at least partly through the direct regulation of one of its targets, *smg* mRNA. *smg* mRNA encodes SMG protein, and I show that those genes down-regulated in *pum* mutant embryos are enriched for SREs. I also demonstrate that SMG protein persists in *pum* mutant embryos. Taken together, our data suggest that PUM-mediated down-regulation of *smg* mRNA is required to help clear SMG protein from the embryo during the onset of zygotic transcription, allowing for the expression of mRNAs that carry SREs during this period.

4.1.2 *smg* mRNA as a critical target of PUM

SMG is a major regulator in the early embryo through its ability to repress translation and induce the degradation of thousands of mRNAs through recognition of its binding site, the SRE,
within these targets (Chen L, et al. 2014; Tadros W, et al. 2007). SMG-mediated turnover of maternal transcripts is an integral part of the MZT, as removal of maternal transcripts is thought to allow for the handover of genetic control from the mother to the embryo. Consistent with an essential role for SMG in completion of the MZT is the fact that loss of SMG protein results in defects in blastoderm cellularization, cell cycle regulation and induction of zygotic transcription. SMG protein is dramatically down-regulated during the onset of zygotic transcription (Benoit B, et al. 2009), suggesting the possibility that once SMG has functioned in the maternal phase of embryogenesis, it needs to be cleared to allow further steps in development to occur. If so, this may represent a critical role for PUM, as this clearance is compromised in a pum mutant embryo.

SMG protein clearance could be important due to the relative lack of complexity of the SRE. This lack of complexity would suggest that SREs would be found frequently within transcripts. As a consequence, SMG clearance could be important to ensure that zygotic transcripts that contain potential SREs are not inappropriately down-regulated.

The clearance of SMG protein could also be required because certain mRNAs originally degraded by SMG need to be re-expressed from the zygotic genome. One model to explain why the embryo might degrade a transcript through the maternal decay machinery only to re-express it through zygotic transcription relies on the observation that many transcripts present in the embryo display spatially-restricted patterns of expression. In particular, it has been observed that the fraction of genes displaying spatially-restricted patterns of expression increases during the onset of zygotic transcription (Lécuyer E, et al. 2007). This may be because while it is possible for maternally-contributed transcripts to be localized within the embryo, zygotic transcription can produce extremely fine expression patterns through selective transcription of a gene in specific subsets of the ~6000 nuclei present in the embryo at ZGA. The requirement for more spatially-restricted forms of gene expression may also increase as development proceeds, since
each successive stage of early development involves more complex degrees of body organization and tissue specification. The degradation and re-expression of key developmental factors could therefore represent an opportunity to ‘reuse’ such factors to specify more restricted compartments of the organism. In agreement with this model, previous analyses have demonstrated that re-expressed genes are statistically enriched for spatially restricted expression patterns (De Renzis S, et al. 2007). Thus, it is possible that PUM-mediated turnover of SMG protein through direct regulation of smg mRNA facilitates the re-expression of many such transcripts and plays a critical role in body patterning.

One question that remains is whether PUM-mediated degradation of smg mRNA is solely responsible for turnover of SMG protein, or whether additional mechanisms are involved. For example, it is possible that loss of smg mRNA is sufficient to down-regulate SMG protein, provided the protein has an inherently short half-life such that without continuous translation its levels drop rapidly. However, it is also possible that the mechanisms of SMG protein turnover are more complex, and that in addition to smg mRNA degradation, SMG protein is also targeted directly for decay during the onset of zygotic transcription. Future work examining the issue of SMG protein stability in the early embryo will hopefully address this issue, and determine the exact role of SMG protein degradation during the MZT.

4.1.3 Functional PUM protein is contributed maternally to the embryo

The requirement for zygotic factors in PUM-mediated degradation of smg mRNA is surprising, since PUM protein is maternally contributed (Macdonald PM 1992) and is therefore present several hours before smg mRNA decay begins. In fact, the best characterized PUM target transcript, hb mRNA, is translationally repressed by PUM before the onset of zygotic transcription (Gamberi C, et al. 2002; Murata Y and Wharton RP 1995).
In addition, I have shown that placement of high affinity PBEs into a reporter RNA produces PUM-dependent decay during the maternal phase of degradation, independent of zygotic transcription, confirming that PUM can under certain circumstances target mRNAs before the onset of zygotic transcription. In contrast, fewer high affinity PBEs or reduced affinity PBEs have little or no ability to induce destabilization of a reporter mRNA. Taken together, these data suggest that the late decay kinetics of smg mRNA likely relates to the fact that the smg 3' UTR contains only reduced affinity PBEs.

As discussed below, PUM appears to generally function with other co-factors to regulate its target mRNAs. It is important to stress that, as discussed above, the vast majority of endogenous mRNAs targeted by PUM are regulated after the onset of zygotic transcription. Therefore, the PUM-dependent, maternal mechanism of degradation of the TGT reporter carrying five high affinity PBEs would appear to be little utilized for the endogenous targets of PUM in the early embryo. The experiment does serve to confirm, however, that PUM protein is present and competent to degrade transcripts during the maternal phase of mRNA decay, raising the question of how PUM-mediated degradation is temporally regulated (i.e., how the majority of PUM activity is restricted to the zygotic phase of degradation). I propose that the vast majority of PUM target mRNAs in the early embryo are similar to smg, in that they also have reduced affinity PBEs.

4.1.4 PUM-mediated zygotic transcript degradation depends on other co-factors

My data suggest that reduced affinity PBEs on their own are not able to induce early or late mRNA decay, indicating that other cis-acting elements within the smg 3' UTR cooperate with the PBEs in smg mRNA decay. In this model, one of the additional co-factors would be
zygotically-contributed, explaining the requirement for zygotic transcription for \textit{smg} mRNA degradation and, by extension, for the bulk of PUM-mediated degradation.

The stabilization of \textit{smg} mRNA in \textit{brat} mutant embryos suggest that one additional co-factor is the RBP BRAT. However, BRAT is also maternally-contributed, hence while it may be required for PUM-mediated degradation of \textit{smg} mRNA, it cannot be responsible for the timing of its decay.

We also present evidence here that AGO1 is an additional factor required for \textit{smg} mRNA degradation. In order to test the role of AGO1 in \textit{smg} mRNA degradation, I knocked down AGO1 expression in early embryos, and found that this significantly stabilizes the \textit{smg} transcript. Interestingly, our Western blot analysis of AGO1 protein levels indicates that the protein begins to accumulate early (during the first two hours of embryogenesis), before the onset of zygotic transcription (which occurs in the third hour). This suggests AGO1 is translated from maternally-contributed transcript. As further evidence for the maternal translation of \textit{ago1} transcript, our knockdown results in significant reductions in AGO1 protein levels as early as the second hour of embryogenesis, which could only occur if the protein was being actively translated at this time. Since our data argue that AGO1 accumulates before the main onset of zygotic transcription and as much as one hour before \textit{smg} mRNA is degraded, it is unlikely that AGO1 protein itself is the missing zygotically-contributed factor for \textit{smg} mRNA degradation.

However, Argonaute proteins do not directly recognize target transcripts, and instead require a guide RNA molecule such as a miRNA. Hence, it is possible that the zygotically-contributed factor required for the proper timing of \textit{smg} mRNA degradation is a miRNA. This hypothesis is consistent with the fact that miRNAs appear to play a conserved role in zygotic mRNA degradation. For example, the \textit{miR-309} cluster of miRNAs is expressed zygotically and is required for the zygotic degradation of a subset of transcripts in the \textit{Drosophila} embryo
(Bushati N, et al. 2008), and the binding sites for other miRNA families are also enriched in
zygotically degraded transcripts (Thomsen S, et al. 2010). In addition, miR-430 has been shown
to degrade hundreds of transcripts during the zygotic wave of decay in zebrafish embryos,
suggesting the role of miRNAs in this process is conserved (Giraldez AJ, et al. 2006). I would
propose a model in which zygotic transcription of miRNAs is a major and conserved feature of
the zygotic decay pathway, with a corresponding requirement for de novo Argonaute protein
synthesis. This model might explain why, despite only modest knockdown of AGO1 protein
levels, I observe a robust stabilization of smg mRNA. This is because if the miRNAs which
target smg mRNA are only transiently expressed, then a brief reduction in the rate of AGO1
protein production occurring at the same time might have a large impact on the amount of a
given miRNA loaded into RISC, even if the total amount of AGO1 protein is only modestly
affected.

4.1.5 PUM, BRAT and AGO1 regulate similar transcripts in the embryo

I have measured changes in global mRNA stability upon AGO1 knockdown by
microarray. In embryos where AGO1 expression is knocked down, 211 genes are up-regulated in
4-5 hour old embryos. Comparison of genes regulated by PUM and AGO1 shows that while
these two factors do regulate transcripts independently of each other, the mRNAs targeted by
both overlap significantly.

In addition, comparison of PUM target mRNAs with the list of BRAT target mRNAs
generated by Laver et al. (Laver JD, et al. 2015) shows that PUM and BRAT also regulate
similar sets of transcripts at the level of transcript stability. Similarly, comparison of AGO1- and
BRAT-regulated transcripts again demonstrates that both factors regulate similar lists of genes.
Taken together, these data argue that PUM, BRAT and AGO1 frequently regulate the same
transcripts in the embryo, and suggests they may be cooperating together during the late (zygotic) wave of mRNA degradation.

It should be noted that Laver et al. has demonstrated through RIP experiments that, counterintuitively, PUM and BRAT bind to largely different subsets of transcripts, suggesting the two factors may also function independently of each other to regulate transcripts in a manner distinct from transcript destabilization (for example, through translational repression). Similarly, while I have shown that the list of genes stabilized in ago1 knockdown embryos overlaps significantly with the lists of genes bound by both PUM and BRAT in the embryo, the degree of overlap is not as great as for those genes which are stabilized in pum and brat mutant embryos. This suggests that while PUM, BRAT and AGO1 may cooperate to degrade target mRNAs, they nonetheless can also bind and regulate transcripts independently of each other as well.

4.1.6 SMG target transcripts are not down-regulated in ago1 knockdown embryos

The overlap between genes up-regulated in pum and brat mutant embryos as well as ago1 knockdown embryos suggests they regulate common targets, including smg mRNA. If so, it is reasonable to hypothesize that the persistence of SMG protein and down-regulation of SMG target transcripts I have documented in pum mutant embryos would also be observed in both brat mutant and ago1 knockdown embryos. In order to test this hypothesis, I began by comparing the list of genes down-regulated in ago1 knockdown embryos with lists of genes 1) bound by SMG, 2) translationally repressed by SMG, and 3) destabilized by SMG. I did not, however, observe any significant overlaps with ago1 knockdown down-regulated genes.

It is important to note that the knockdown of AGO1 protein in these embryos is far from complete; the knockdown is in fact quite modest. It is therefore possible that the residual AGO1
protein is sufficient to regulate \textit{smg} mRNA properly, resulting in the proper clearance of the protein and allowing for the proper expression of SMG target genes after the MZT. This observation has to be reconciled, however, with the fact that a significant stabilization of \textit{smg} mRNA is observed in \textit{ago1} knockdown embryos. In fact, the degree of stabilization is similar to that observed in both \textit{pum} and brat mutant embryos. I would therefore propose a more complex model, in which the residual presence of both PUM and BRAT in the embryo is sufficient to translationally repress \textit{smg} mRNA despite its stabilization. This model is dependent on the notion of SMG protein instability discussed above, and represents another rationale for investigating that issue.

Interestingly, I performed a similar analysis comparing genes down-regulated in brat mutant embryos and the lists of SMG target transcripts described above, and found significant overlaps between them all. This supports a model in which PUM and BRAT (and possibly AGO1) are required for the proper clearance of SMG protein during the MZT, in order to allow for the (re)-expression of SMG target mRNAs. In support of this model, it will be important to confirm that SMG protein is indeed up-regulated in brat mutant embryos, an experiment that has not yet been performed.

4.1.7 A PUM-BRAT-AGO1 complex as a model for \textit{smg} mRNA degradation

Our data argue that recognition of the \textit{smg} 3’ UTR by PUM is not sufficient for transcript decay, and that recruitment of other factors (such as BRAT and AGO1) is also required. In addition, global analysis of gene expression patterns in \textit{brat} mutant embryos (Laver JD, et al. 2015) and in embryos where AGO1 expression has been knocked down (the current work) suggest that PUM, BRAT and AGO1 regulate similar lists of transcripts during the zygotic phase
of transcript degradation. One interesting possibility is that these three factors cooperate in regulation of target transcripts as part of a protein complex (Fig. 14).

This model is consistent with previous data concerning the dependence of PUM on other factors to regulate its targets. For example, one of the earliest described functions of PUM is to regulate \(hb\) mRNA, in order to allow for posterior differentiation during early embryogenesis. PUM-mediated regulation of \(hb\) mRNA requires BRAT as well as the posterior-localized co-factor NOS. As discussed previously, PUM-mediated regulation of \(hb\) mRNA is believed to involve the formation of a PUM-BRAT-NOS complex on \(cis\) elements in the \(hb\) 3′ UTR (Sonoda J and Wharton RP 1999; Sonoda J and Wharton RP 2001). It was originally thought that BRAT did not make direct contact with the target RNA, but rather was recruited through the presence of both PUM and NOS. Recent data shows that BRAT has RNA binding activity and makes direct contact with the \(hb\) 3′ UTR. Importantly, the binding of BRAT to \(hb\) mRNA is believed to be enhanced by the presence of PUM, and vice versa (Loedige I, et al. 2014), suggesting a model of direct interaction and cooperative binding between these two factors.

The regulation of the \(hb\) transcript through the formation of PUM-BRAT-NOS complex has become the canonical model for how PUM regulates its targets. However, since NOS protein is restricted to the posterior, the PUM-BRAT-NOS complex cannot underlie the majority of PUM-mediated degradation I have observed by microarray. In fact, the only essential role of NOS has been shown to be regulation of \(hb\) mRNA (Irish V, et al. 1989), suggesting the PUM-BRAT-NOS complex has only limited function in the embryo.

We have also shown that AGO1 is required for \(smg\) mRNA decay, in agreement with previous work showing extensive cooperation between Puf and Argonaute proteins (Galgano A, et al. 2008; Incarnato D, et al. 2013; Kedde M, et al. 2010). Data also exist which indicate there may be a conserved PUM-AGO1 physical interaction (Friend K, et al. 2012), although more
Figure 14. Proposed model of regulation by PUM, BRAT and AGO1 during the late (zygotic) phase of mRNA decay. Before the onset of late (zygotic) decay, PUM and BRAT are present in the embryo, but AGO1 levels are low. In addition, it is possible that miRNAs targeting the smg 3' UTR have not been expressed yet, such that what AGO1 is present cannot be recruited to the transcript (A). Over the first few hours of embryogenesis, AGO1 protein accumulates, and it is also possible that smg mRNA-targeting miRNAs begin to be expressed. I propose that the regulation of smg mRNA during the late (zygotic) phase of mRNA decay involves the formation of a PUM-BRAT-AGO1 complex on the 3' UTR of target transcripts, leading to deadenylation and turnover through the ability of all three factors to interact with the CCR4-NOT-POP2 deadenylase complex (in the case of AGO1, this occurs through the GW182/TNRC6 family of co-factors - GW) (B).
recent data suggest this interaction is not always strictly required for PUM-mediated repression (Weidmann CA, et al. 2014). Nevertheless, our data support a model in which PUM, BRAT and AGO1 directly interact on the *smg* 3' UTR in the form of a PUM-BRAT-AGO1 complex, a mechanism that may involve cooperative binding to the target transcript.

It should be noted that PUM-mediated degradation is likely to occur in other modes as well. For example, regulation of *cycB* mRNA in the pole cells of the embryo is dependent on PUM and NOS and at least one other posterior-localized factor, but is independent of BRAT activity (Kadyrova LY, et al. 2007). This suggests that other configurations of PUM containing complexes may exist in the embryo. Future work on other transcripts (such as *cycB* mRNA) may more clearly define such associations.

### 4.1.8 The CCR4-POP2-NOT deadenylase complex

Upon binding of PUM, BRAT and AGO1 to *smg* mRNA, transcript decay is likely induced by the CCR4-POP2-NOT deadenylase complex. This is because Puf proteins (Goldstrohm AC, et al. 2006; Goldstrohm AC, et al. 2007), BRAT (Laver JD, et al. 2015; Temme C, et al. 2010) and Argonaute proteins (through their interacting partners, GW182/TNRC6) (Pfaff J and Meister G 2013) have been shown to interact with this complex. Deadenylation by the CCR4-POP2-NOT complex, in turn, is often the rate limiting step in eukaryotic transcript decay, and its recruitment by sequence-specific RBPs to an mRNA is a common mechanism of transcript turnover (Miller JE and Reese JC 2012; Zaessinger S, et al. 2006). Determining whether PUM, BRAT and AGO1 form a complex, whether this complex enhances the recruitment of each individual component and whether such a complex functions through recruitment of the CCR4-POP2-NOT complex all represent promising future avenues of research into the mechanisms of *smg* mRNA degradation.
4.1.9 Multi-factorial post-transcriptional regulation in the early Drosophila embryo

Based on my data I propose that instead of evolving a large number of RBPs to individually direct very specific forms of PTR for a small handful of transcripts, the embryo instead utilizes a small number of RBPs and regulates mRNAs by having multiple RBPs bind to a given mRNA. In this model, the complexity of PTR observed in the embryo is achieved through the unique complement of RBPs recruited to a given transcript, and the cooperation or antagonism of each RBP of the function of the others.

4.2 Future directions

4.2.1 BRAT

*smg* mRNA is stabilized in brat mutant embryos and Laver et al. (Laver JD, et al. 2015) have demonstrated that *smg* mRNA is enriched in BRAT RIPs, suggesting that BRAT directly regulates *smg* mRNA. However, it will be important to demonstrate conclusively that BRAT does indeed regulate *smg* mRNA through direct binding. Doing so would involve determining where in the *smg* 3' UTR BRAT makes direct contact. Loedige et al. recently demonstrated that BRAT directly binds to the *hb* mRNA (Loedige I, et al. 2014) and described a gel shift protocol for measuring such interactions *in vitro*. This protocol could easily be adapted to the *smg* 3' UTR. Such an experiment would be aided by the fact that Laver et al. have defined the core motif of a Brat binding site, which consists of the 4-mer ‘UGUU’. The *smg* 3' UTR has 14 such motifs, and gel shifts could be used to determine whether these sites are functional *in vitro*. If the gel shift data suggests that BRAT does make direct contact with the *smg* 3' UTR through UGUU
containing cis elements, the role of such BRAT binding sites could be assayed as for the smg 3' UTR PBEs described in the current work (by mutating the sites in the context of a reporter RNA). In addition, such reporters could be used in further RIP experiments; if immunoprecipitation of BRAT from transgenic embryos carrying a BRAT binding site mutant reporter failed to pull down the reporter, this would provide further evidence for BRAT binding to the smg 3' UTR. Together, such data would support the hypothesis that BRAT regulates the smg transcript through direct binding.

4.2.2 AGO1

In the same vein, additional data are required to demonstrate that smg mRNA represents a direct target of AGO1. As for BRAT, the first step would be to generate data consistent with a direct interaction between AGO1 and smg mRNA. The most straightforward approach to address this issue would be to perform RIP experiments, which would involve immunoprecipitating AGO1 from embryo extracts and performing RT-qPCR on associated RNA to detect an enrichment of the smg transcript.

Once a direct interaction between AGO1 and smg mRNA is confirmed, it will be critical to determine which miRNA binding sites are responsible for recruiting AGO1 to the transcript. This effort is complicated by the fact that there are many putative miRNA binding sites in the smg 3' UTR. As discussed below, I propose various experiments to narrow the number of potential miRNA binding sites that could be responsible for the zygotic degradation of smg mRNA. Once a small number of potential miRNA binding sites are identified, similar cis element mapping experiments, similar to those proposed for BRAT, could be performed. Specifically, such miRNA binding sites could be mutated in the context of reporter RNAs,
followed by stability measurements and AGO1 R IPs to determine whether the sites are functional \textit{in vivo}.

4.2.3 A PUM-BRAT-AGO1 complex

If it is established that BRAT and AGO1 both make direct contact with \textit{smg} mRNA, work could then begin to better understand how PUM, BRAT and AGO1 cooperate to regulate the \textit{smg} transcript. It will be important to examine potential protein-protein interactions between these various factors to begin to understand the mechanisms of \textit{smg} mRNA regulation in greater detail. Ample evidence indicates that both PUM and BRAT make direct contact with one another, and the PUM$^{G1330D}$ mutation has been shown to disrupt this interaction (Sonoda J and Wharton RP 2001). One informative experiment would be to generate PUM$^{G1330D}$ rescue constructs, in order to determine whether a direct PUM-BRAT interaction exists in embryos (through co-immunoprecipitation experiments). It should be noted that the use of RNAse during such co-immunoprecipitation experiments would be important. This is because the observation that two RBPs continue to interact when the extract is incubated with RNAse offers stronger evidence that protein-protein interactions underlie their association, and helps to exclude the possibility that co-immunoprecipitation occurs simply because both RBPs are simultaneously bound to the same transcript(s).

In addition, rescue with such a transgene could be used to determine whether the PUM-BRAT interaction is required for \textit{smg} mRNA regulation; loss of \textit{smg} mRNA regulation upon expression of the mutant (but not wild-type) rescue construct in \textit{pum} mutant embryos would indicate the interaction was required.

As well, evidence suggests Puf proteins and Argonaute proteins may make direct contact with one another (Friend K, et al. 2012; Weidmann CA, et al. 2014). Therefore, determining
whether a direct PUM and AGO1 interaction is required for smg mRNA degradation is another promising avenue of investigation. Such experiments, however, will have to contend with some controversy in the literature. In particular, Friend et al. (Friend K, et al. 2012) determined that a point mutation of human Pum2 (T752E) abrogated binding to human Argonaute proteins and that this binding was required for repression of a target mRNA in vitro. In contrast, Weidmann et al. (Weidmann CA, et al. 2014) found that this residue was not required for the interaction between human Pum2 and AGO2, nor was the corresponding residue required for the interaction between Drosophila PUM and AGO2. These authors furthermore show that the interaction between Drosophila PUM and AGO2 is dispensable for repression in S2 cells, and importantly are also unable to detect an interaction between PUM and AGO1. Hence, it is difficult to say at this time whether an interaction between PUM and AGO1 will be detectable in the Drosophila embryo, and whether such an interaction would prove to be required for smg mRNA degradation. Nevertheless, considering the strong overlap between PUM and AGO1 targets during zygotic degradation, investigating the possibility of a physical interaction between these two factors by co-immunoprecipitation is an important first step.

I propose that BRAT and AGO1 are required for regulation of the smg transcript because PUM does not stably associate with the mRNA on its own. The same is perhaps true of BRAT and AGO1 as well, consistent with a model in which the formation of larger complexes strengthens the binding of any one factor. Several RIP experiments could be performed to investigate this possibility. For example, one could perform BRAT or AGO1 RIPS on the PBE-reporters generated during this project. If loss of PBEs results in loss of binding of BRAT and/or AGO1 to the transcript, this would support a model of cooperative binding. Similar experiments could be performed with BRAT binding site mutant (using PUM or AGO1 RIPS) or miRNA binding site mutant (using PUM or BRAT RIPS) reporters. It would also be informative to
perform all such experiments both before and after the onset of zygotic transcription, in order to
determine whether the formation of a PUM-BRAT-AGO1 complex underlies the temporal
regulation of smg mRNA degradation.

Similarly, the protein-protein interaction mutant rescue constructs described above could
be used here. For example, one could express a BRAT-binding deficient PUM transgene in *pum*
mutant embryos, and determine whether BRAT (or AGO1) RIPs still detect smg mRNA.

### 4.2.4 Zygotic miRNA expression and timing of smg mRNA degradation

A critical outstanding question raised by this project centers on the timing of smg mRNA
degradation in the embryo. As discussed above, I propose that one or more zygotically expressed
miRNA(s) is required to control the timing of smg mRNA decay.

A first step to better understand the role of miRNAs in the zygotic wave of mRNA
degradation in general and for smg mRNA regulation in particular would be to profile miRNA
expression throughout this period of development. Expression profiling for miRNAs can be
accomplished through deep sequencing (Creighton CJ, et al. 2009). The aim of this effort would
be to generate lists of miRNAs expressed during embryogenesis, with a particular focus on
distinguishing maternally- vs. zygotically-contributed miRNAs. As for previous efforts to
characterize maternal vs. zygotic mRNA degradation (Tadros W, et al. 2007; Thomsen S, et al.
2010), these experiments could begin by using wild-type embryos and UEs. The comparison of
embryos to UEs is particularly powerful in this context, since only maternally-contributed
miRNAs will be expressed in UEs, allowing for their identification. In addition, examining
miRNA expression over a time-course covering the period of embryogenesis spanning the onset
of zygotic transcription would be informative. This is because only maternally-contributed
miRNAs would detectable before zygotic transcription, and only zygotic miRNAs would see
their levels increase after zygotic transcription begins. In addition to miRNA expression profiling in wild-type embryos and UEs, it would be informative to perform similar experiments in ago1 knockdown embryos. It should be noted that, since knockdown of ago1 involves mating females bearing the Gal4 driver to males bearing the shRNA construct that targets ago1 mRNA, it will not be possible to examine UEs where ago1 expression has been knocked down. However, comparing the miRNA complement in ago1 knockdown embryos to control knockdown embryos would allow me to determine which miRNAs are significantly reduced when ago1 is knocked down. Such miRNAs would be potential candidates for targeting smg mRNA, since ago1 knockdown disrupts smg mRNA decay.

It would also be important to determine which pools (maternal vs. zygotic) of miRNAs are associated with AGO1 over this period of development. One approach would be to immune-precipitate AGO1 protein and sequence associated miRNAs over the course of early embryogenesis, in order to determine which mature miRNAs are present in RISC over the relevant period of development. Current work in our laboratory to generate synthetic antibodies against RBPs in the embryo has yielded antibodies against AGO1 which would be useful in this regard (Laver JD, et al. 2012).

There are also computational approaches available to generate a list of zygotically-expressed miRNAs potentially involved in smg mRNA decay. In particular, the list of genes up-regulated in ago1 knockdown embryos could be examined for enriched motifs. One would predict that this gene list should be enriched for miRNA binding sites, since they represent putative AGO1 target transcripts. Comparison of this data with data generated from either the miRNA expression profiling or AGO1-associated miRNA analyses described above should in principle result in a list of high-confidence zygotically-expressed miRNAs that could be important for smg mRNA decay.
Having determined all zygotically-expressed miRNA, and having perhaps identified particular miRNAs whose levels are down-regulated in ago1 knockdown embryos, work could then begin to determine whether smg mRNA degradation is dependent on such zygotically-transcribed miRNAs. As discussed above, this question is currently complicated by the fact that there are numerous predicted miRNA binding sites in the UTR. Intriguingly, many of these predicted miRNA binding sites reside in nucleotides 1-438 of the smg 3' UTR, which I have demonstrated to be required for robust degradation of the mRNA. During the course of this project, I attempted to more finely map nt 1-438 in order to identify cis elements residing in this fragment required for smg mRNA turnover. Briefly, I divided nucleotides 1-438 into four sections of approximately 110 nt each, and constructed reporter RNAs containing the entire length of the smg 3' UTR (nucleotides 1-907) but with either single, double or triple deletions of each of these sections. I then measured RNA stability as for our other transgenic reporters. The objective was to identify a single section that was required for full regulation of the reporter, and to eliminate the rest of nucleotides 1-438 from further study. However, these experiments produced ambiguous results in which modest stabilization was observed upon deletion of all the sections tested. While difficult to interpret, the data was consistent with a model in which, far from containing a single cis element, nucleotides 1-438 actually contains several cis elements spread throughout the sequence. While additional experiments would be required to support such a model, this scenario is consistent with the notion that the large number of miRNA binding sites predicted in nucleotides 1-438 of the smg 3' UTR are functional and required for degradation of the transcript.

A potential way forward would be to compare the list of zygotically-transcribed miRNAs detected through expression profiling with the list of putative smg mRNA-targeting miRNAs. If such a comparison yields a relatively short list of potential candidate miRNAs responsible for the
timing of *smg* mRNA, additional reporters modelled on our TGT + *smg* 3' UTR reporters could be constructed to test the requirement for these cis elements. Specifically, identified candidate miRNA binding sites could be mutated in the context of the full-length *smg* 3' UTR, and the stability of the resulting reporter could be tested in transgenic embryos.

### 4.2.5 Constructing a reporter mRNA that is degraded with late kinetics

Our data argue that PUM, BRAT and AGO1 cooperate to regulate many transcripts after the onset of zygotic transcription in the embryo. In addition, the experiments proposed above may provide evidence for a PUM-BRAT-AGO1 complex. I therefore propose a ‘ground-up’ approach to determine the minimal *cis*-elements required to target an mRNA for late decay. This approach would allow one to determine whether 1) PUM, BRAT and/or AGO1 are the only necessary factors, and 2) the number of required binding sites for each. Such reporters could be modelled on the TGT reporters described in the current work.

The approach here would be modelled on my experiments for PUM and PBEs. Initially I would test the effect that BRAT-binding sites of various affinities and numbers would have on TGT mRNA stability. Similar experiments would test the effect of various numbers and affinities of binding sites for a zygotically expressed miRNA. Relative affinities for different BRAT sites would be based on RNAcompete data (Ray D, et al. 2009; Ray D, et al. 2013), while relative affinities for miRNA binding sites would be based on the extent of complementarity with the miRNA. I would then combine BRAT and/or miRNA binding sites that have little or no effect on TGT mRNA stability with 5 R30 PBEs (which are based on endogenous PBEs found in the *smg* 3' UTR and which I showed are not sufficient for decay on their own). As a controls, reporters with mutant and wild-type binding sites would all be tested. If a combination of sites induces
TGT destabilization while the sites on their own do not, this would indicate that the corresponding trans-acting factors are able to cooperate with one another other.

Although less biologically relevant, another interesting experiment that could be performed would involve using binding sites for PUM and BRAT as above, paired with maternally-expressed miRNA binding sites. If maternally-expressed miRNA binding sites could cooperate with PUM and/or BRAT binding sites to induce maternal decay, this would provide strong evidence for the hypothesis that miRNAs play an important role in the timing of degradation mediated by these RBPs.

### 4.2.6 The downstream mechanisms of PUM-mediated degradation of smg mRNA

It will also be important to determine the mechanisms that act downstream of PUM, BRAT and AGO1 to induce turnover of smg mRNA. As discussed previously, Puf proteins are known to associate with and recruit the CCR4-POP2-NOT deadenylase complex to target transcripts, through direct binding to the POP2 subunit of the complex. Argonaute proteins (including AGO1) are also known to repress their targets through recruitment of the deadenylase complex (although this occurs indirectly through the bridging factor GW182/TNRC6), as is BRAT. Therefore, an important first experiment in the effort to better understand the mechanism of smg mRNA degradation would be to examine the role of deadenylation in this process.

There are several approaches that can be employed to measure deadenylation of a given mRNA. The most straightforward is to perform Northern blots on an mRNA of interest, using changes in mobility as a measure of poly(A) tail length. In order to achieve the resolution required, the 3’ end of the transcript is typically separated from the rest of the mRNA through annealing of a complementary oligonucleotide to a region just a few hundred nucleotides
upstream of the end of the transcript followed by RNase H digestion. The digested sample is then run on an acrylamide gel, blotted and probed for the 3’ end of the RNA of interest. As a control, a sample is RNase-treated in the presence of the above oligonucleotide and oligo(dT). In this sample, the poly(A) tail of the RNA of interest will be removed, providing a convenient size marker for a fully deadenylated RNA. Samples in which the RNA of interest runs slower than the deadenylated control are interpreted as being poly-adenylated, and differences in poly(A) status are inferred from changes in the mobility of the RNA between samples. A correlation between the deadenylation and degradation of smg mRNA and defects in smg mRNA deadenylation in pum and brat mutant embryos as well as ago1 knockdown embryos would provide strong evidence that PUM, BRAT and AGO1 trigger smg mRNA degradation through deadenylation.

It is also possible to take a genetic approach to this problem. For example, mutants exist in components of the CCR4-POP2-NOT complex in Drosophila. In particular, mutants in the Drosophila CCR4 homolog, TWIN, have been shown to stabilize other transcripts whose degradation is deadenylation-dependent (Zaessinger S, et al. 2006), including the cyclin B mRNA, which is targeted by PUM in Drosophila PGCs (Kadyrova LY, et al. 2007). It should be noted, however, that residual deadenylase activity mediated by the POP2 complex member, which is also a deadenylase (Daugeron MC, et al. 2001) could confound this experiment. However, examining smg mRNA degradation in twin mutant embryos would be an important first step in order to determine whether smg mRNA is dependent on the deadenylase complex.
4.2.7 The biological significance of PUM-mediated mRNA degradation

The experiments described in the previous section are focused on the mechanisms of PUM-mediated degradation. Such research will shed greater light on the functioning of Puf proteins, in particular their documented requirement for co-factors. Another important avenue of research concerns the biological significance of PUM-mediated mRNA degradation in the early *Drosophila* embryo. This avenue of research is complicated by the sheer number of transcripts both directly and indirectly targeted by PUM in the early embryo.

One of the most important targets of PUM during *Drosophila* embryogenesis is the *hb* transcript. Regulation of *hb* mRNA also requires BRAT and NOS. Interestingly, regulation of *hb* mRNA appears to be the only absolutely required function of NOS protein, since *hb nos* double mutant embryos are viable whereas single mutant *nos* embryos die due to ectopic expression of HB protein (Irish V, et al. 1989). Building on this observation, it would be interesting to examine the phenotype of *hb pum* double mutant embryos. In such embryos, the early lethality caused by ectopic HB expression observed in *pum* mutant embryos would in theory be suppressed, allowing one to examine other phenotypes caused by loss of PUM-mediated mRNA degradation. It is possible that, as for *hb nos* double mutant embryos, *hb pum* double mutant embryos might prove to be viable, if *hb* mRNA is the only indispensable target of PUM (as it is for NOS). However, it is also possible that *hb pum* double mutant embryos would prove to be non-viable or display other phenotypes, if there are other critical targets of PUM in the early embryo. Such phenotype(s) of *hb pum* double mutant embryos might provide important preliminary indications of what such targets might be.

In the current work I have proposed a model in which *smg* mRNA is a critical target of PUM, since loss of PUM results in stabilization of *smg* mRNA, persistence of SMG protein, and
a concomitant down-regulation of SMG target transcripts. If our model is correct and clearance of smg mRNA is a core function of PUM during the zygotic wave of transcript degradation, it is possible that the phenotypes of hb pum double mutant embryos might actually be due to the inappropriate down-regulation of SMG targets after the onset of zygotic transcription. One approach to more directly determine to what extent PUM regulation of smg mRNA impacts Drosophila embryogenesis would be to devise an experimental system in which SMG protein persists after the onset of zygotic transcription in the absence of other effects caused by loss of PUM protein. This could be accomplished by constructing a SMG rescue construct under the control of the PBE- smg 3' UTR generated during the course of the current project. As per the data described in this work, such a construct should continue to express SMG protein after the onset of zygotic transcription, since it would escape regulation by PUM. Importantly, the regulation of other PUM target transcripts would not be perturbed, so the effect of persistent SMG expression could be isolated from the other effects of pum mutations in transgenic embryos expressing this construct. I would also construct such a transgene with mutations in BRAT and/or miRNA binding sites in the smg 3' UTR, once they are identified in the experiments described above. It should be noted that this approach might be complicated by the potential ectopic expression of SMG protein during oogenesis, since smg mRNA is translationally repressed during this period through factors that remain to be determined but which may include PUM and BRAT (Lipshitz laboratory, unpublished data). Ectopic SMG expression during oogenesis represents an obstacle because it interrupts egg development and results in females who produce no eggs (Semotok JL, et al. 2005). However, the fact that pum mutant and brat mutant females both lay eggs suggests the possibility that PBE or BRAT binding site mutant smg transgenes may not interfere with this process.
While the current work indicates that stabilization of smg mRNA does result in elevated levels of SMG protein, it is also clear from our Western blots of SMG protein in pum mutant embryos that SMG protein is unstable and is rapidly turned over during this period of embryogenesis. Current work in the laboratory of Howard Lipshitz is aimed at understanding the mechanisms that underlie this instability (Wendy Cao, unpublished data). One of the long-term aims of this project is to generate SMG transgenic constructs in which the protein is stabilized after the onset of zygotic transcription through mutation of either specific residues (e.g., sites of ubiquitination) that are required for the turnover of the protein. Such a construct would represent another opportunity to examine the effects on early embryogenesis resulting from persistence of SMG protein.
References


Cao Q, Padmanabhan K, JD R (2010). Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition. RNA 16:221-227


Jones CI, Zabolotskaya MV, SF. N (2012). The 5' → 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development. Wiley Interdiscip Rev RNA 3:455-468


King PE, Rafai J (1970). A possible mechanism for initiating the parthenogenetic development of eggs in a parasitoid Hymenopteran, Nasonia vitripennis (Walker) (Pteromalidae). Entomologist 106:118-120


Sharova LV, Sharov AA, Nedorezov T, Piao Y, Shaik N, MS K (2009). Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. DNA Res 16:45-58


Wharton RP, Sonoda J, Lee T, Patterson M, Y. M (1998). The Pumilio RNA-binding domain is also a translational regulator. Mol Cell 1:863-872


