Developmental Analysis of DEAD Box 1

By

Matthew Robert Hildebrandt

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Experimental Oncology

Department of Oncology

University of Alberta

Abstract

DEAD box protein 1 (DDX1) is a member of the DEAD box family of helicases which bind and unwind double-strand (ds) RNA and are associated with all aspects of RNA metabolism. DDX1 was identified by differential screening of a subtracted cDNA library prepared from two retinoblastoma (RB) cell lines. *DDX1* is co-amplified with *MYCN* and overexpressed in a subset of *MYCN*-amplified neuroblastoma (NB) cell lines and tumours as well as in all *MYCN*-amplified RB cell lines examined to date. DDX1 has been shown to be associated with a number of different cellular processes including DNA repair and mRNA 3' processing, polyadenylation and transport. *In vitro* work has demonstrated a role for DDX1 in unwinding RNA-RNA and RNA-DNA duplexes (up to 29 base pairs) in an ADP-dependent manner. Finally, DDX1 can digest single-strand (ss) RNA in a magnesium-dependent and energy-independent manner.

The structures of a number of DEAD box proteins have been solved. These X-ray structures have been used to determine the mechanisms by which DEAD box proteins bind and modify RNA and ribonucleoprotein (RNP) complexes. I was able to generate a number of DDX1 protein crystals including some obtained in the presence of an RNA-DNA duplex; however, none of these crystals yielded diffraction patterns. An in-depth characterization of the ribonuclease activity of DDX1 revealed that DDX1 degrades single-strand (ss) RNA in a sequence-independent manner. My results suggest that DDX1 is a ribonuclease that cleaves ssRNA down to the 3-6 nucleotide range. Attempts to map the ribonuclease

activity of DDX1 identified the carboxyl-terminal region of DDX1 as being essential for ribonuclease activity.

I used a gene knockout model to examine the role of DDX1 during development. I found that *Ddx1* knockout embryos stall during development at the 2- or 4-cell stage. Intriguingly, I also noted that a subset of the wild-type embryos die between E3.5 and E6.5. Wild-type embryos that were dying arose from second generation intercrossed heterozygote mice. This phenomenon was observed in both the FVB and C57BL/6 mouse strains and represents a type of non-Mendelian inheritance not reported previously. Examination of the subcellular localization of DDX1 in fetal and adult mouse tissues revealed a primarily nuclear localization. However, when I examined the subcellular localization of DDX1 in oocytes and pre-implantation embryos, I found that the protein was localized primarily to the cytoplasm where it formed large aggregates. These aggregates were highly dynamic in nature, with size and numbers changing as a function of developmental stage. While I was not able to identify any organelles or proteins that co-localized with DDX1 in the aggregates, I was able to determine that the aggregates at the 2-cell stage were dependent on the presence of RNA. DDX1 aggregates in the cytoplasm of pre-implantation embryos may therefore represent a novel RNA containing granule that plays a role either in the protection or degradation of RNAs.

Preface

This thesis is an original work by Matthew R. Hildebrandt. No part of this thesis has been previously published.

The research project, of which this thesis is a part, received research ethics approval from the Cross Cancer Institute Animal Care Committee, #BC11185.

Acknowledgements

I would like to first thank my mentor and supervisor, Roseline Godbout. Her support and guidance through the years taught me the skills I needed to successfully complete this Ph.D. Thank you for continually pushing me to work harder and strive to take hold of my project. I would also like to thank Bart Hazes for agreeing to teach me crystallography and serving on my supervisory committee to help further my project in its diverse nature.

There are a few people in the lab I would like to thank specifically for their direct assistance with my project: Liz, for teaching me many of the techniques I used over the years and screening the multitude of mice over the years; Darryl, who taught me the complexity of fixation, sectioning, and antibody staining; and the fellow members of the DDX1 project, Lei and Devon. I would also like to thank other members of the Godbout lab for their support over the years: Miranda, Rong-Zong, Raja, Mike, Lucy, Tina, Saket, Indrani, David, Elizabeth, Jack, Kevin, and Stanley. Despite any difficulties during the week I always had the lunch break to unwind. Thanks Devon for always having a pack of cards and to Miranda and Mike for playing the numerous different games over the years.

My project would not have gone where it has without the support of two of our core facilities and those individuals that maintain them. I had substantial assistance from the vivarium staff, Gail, Dan and Daming, who maintained the mouse colony and all of the breeding requests for me over the years. The imaging

would not look as nice as it does without Dr. Sun and Gerry who taught me how to take a picture the correct way.

Finally, I would not be here without the support of my family. My parents (Robert and Karen) ingrained in me the drive to pursue higher education and without that drive I would not have made it as far as I did. I also could not have done this without the love of my life, Kerri. You have supported me through the tough times and been there to celebrate the successes. The wait is finally over and we can take that break and travel the world.

Table of Contents

Chapter 1

Introduction..... 1

Chapter 6

List of Tables

List of Figures

Abbreviations

-/-	Knockout
+/-	Heterozygote
+/+	Wild-type
Å	Angstrom
aa	Amino acid
ADP	Adenosine diphosphate
ADP-β-S	Adenosine 5'[β-thio]diphosphate
Ago	Argonaute
ANOVA	Analysis of variance
AR29	Polyadenoside$_{29}$
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATP-γ-S	Adenosine 5'[γ-thio]triphosphate
β-gal	β-galactosidase
BISMA	Bisulfite sequencing DNA methylation analysis
β-ME	β-mercaptoethanol
bp	Base pairs
BSA	Bovine serum albumin
C	Carboxyl
CCD	Cooled charged-coupled device
cDNA	Complementary DNA
Cdx	Caudal type homeobox
Ci	Curies
cm	Centimeter
COBRA	Combined bisulfite restriction analysis
CPEB	Cytoplasmic polyadenylation element binding protein
CPSF	Cleavage and polyadenylation specificity factor
CR29	Poly cytodine$_{29}$
CstF	Cleavage stimulation factor
CTP	Cytidine triphosphate
DAB	3,3'-diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole

DCT	Distal convoluted tubules
dd	Double distilled
DDX	DEAD box protein
DIC	Differential interference contrast
DNA	Deoxyribonucleic acid
Dnmt	DNA methyltransferase
ds	Double-strand
DTT	Dithiothreitol
E	Embryonic day
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EF	Eukaryotic elongation factor
eIF	Eukaryotic initiation factor
ES	Embryonic stem
EtOH	Ethanol
Exo1	Exonuclease 1
FBS	Fetal bovine serum
fmol	Femtomole
G	Gauge
GCL	Ganglion cell layer
GST	Glutathione s-transferase
GTP	Guanosine triphosphate
GV	Germinal vesicle
Gy	Gray
H2AX	Histone 2A isoform X
HIV	Human immunodeficiency virus
hnRNP	Heterogeneous ribonucleoprotein
HR	Homologous recombination
HRP	Horseradish peroxidase
hu	Human
ICR	Imprinting control region
IF	Immunofluorescent
IHC	Immunohistochemistry
INL	Inner nuclear layer

IP6	Inositol hexakisphosphate
IPL	Inner plexiform layer
IPTG	β-D-1-tiogalactopyranoside
IR	Ionizing radiation
JC virus	John Cunningham virus
kb	Kilobases
kDa	Kilo Daltons
LB	Lysogeny broth
LSM	Laser scanning microscope
M	Molar
MAFFT	Multiple Alignment using Fast Fourier Transform
Mb	Mega base
met	Methionine
μg	Microgram
mg	Milligram
Mg^{2+}	Magnesium ions
MI	Meiosis I
MII	Meiosis II
miRNA	MicroRNA
mjDEAD	*Methanococcus jannaschii* DEAD box protein
μl	Microlitre
ml	Millilitre
μm	Micrometers
μM	Micromolar
mm	Millimeter
mM	Millimolar
MRN	Mre11, Rad50 and Nbs1 complex
mRNA	Messenger RNA
MZT	Maternal to zygotic transition
N	Amino
NB	Neuroblastoma
neo	Neomycin
ng	Nanogram
NHEJ	Non-homologous end joining

nm	Nanometer
nM	Nanomolar
NTP	Nucleoside triphosphate
nts	Nucleotides
OD	Optical density
ONL	Outer nuclear layer
OPL	Outer plexiform layer
P	Postnatal
p	Probability
PA	Polyadenylation signal
PAGE	Polyacrylamide gel electrophoresis
PB	Processing bodies
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PCT	Proximal convoluted tubules
PEG	Polyethylene glycol
PK	Proteinase K
pmol	Picomol
PMSF	Phenylmethanesulfonylfluoride
RB	Retinoblastoma
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
Rnase	Ribonuclease
RNP	Ribonucleoprotein
RPE	Retina pigmented epithellium
rRNA	Ribosomal RNA
RT	Reverse transcription
SA	Splice acceptor
SDS	Sodium dodecyl sulphate
SF2	Super family 2
SG	Stress granule
siRNA	Small interfering RNA
SMN	Survival motor neuron
sn	Small nuclear

sno	Small nucleolar
Sox2	SRY-box containing gene 2
SPRY	Dual specific kinase sp1a of *Dictyostelium* and rabbit ryanodine receptors
ss	Single-strand
TBE	Tris borate EDTA
TBS	Tris-buffered saline
TE	Tris EDTA
TIF	Transcription intermediary factor
T_m	Melting temperature
TP	Transition nuclear protein
tRNA	Transfer RNA
U	Units
UTP	Uridine triphosphate
UV	Ultraviolet
V	Volt
Zar	Zygote arrest
ZGA	Zygotic gene activation

Chapter 1

Introduction

1.1 DEAD box protein family

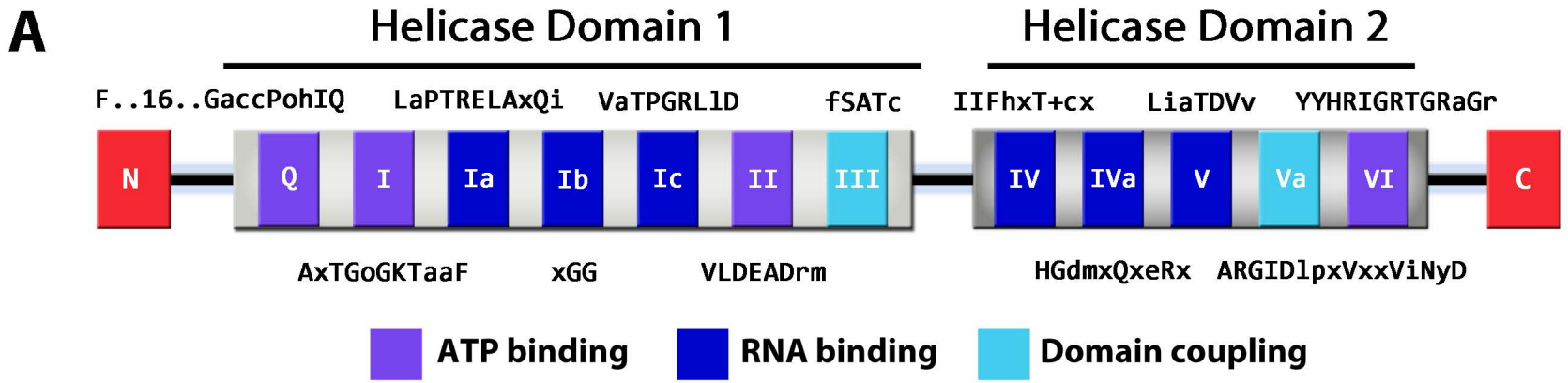
1.1.1 Identification and characterization of DEAD box proteins

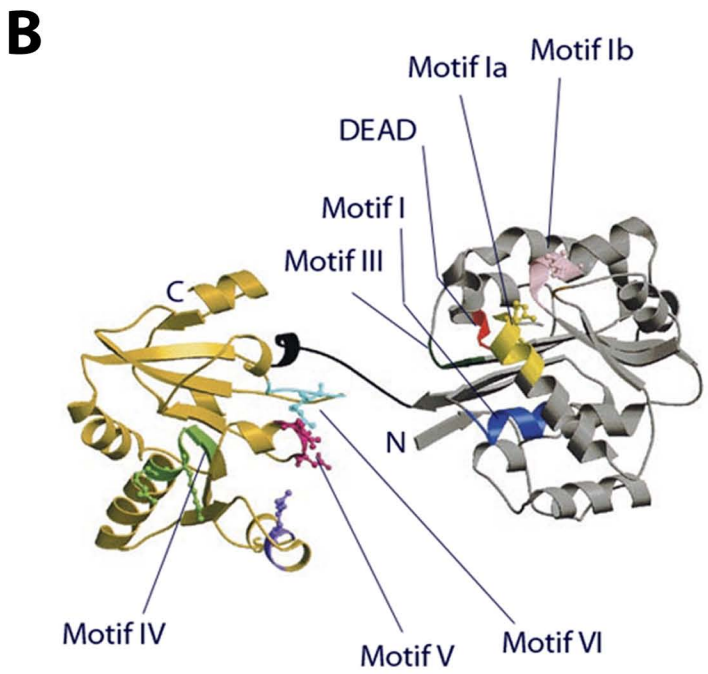
Helicases are a group of enzymes that are associated with all aspects of DNA and RNA metabolism (1-17). These proteins unwind DNA, RNA or hybrid duplexes in an ATP-dependent manner. The largest group of helicases is Super Family 2 (SF2) which primarily consists of the DEAD box protein family of RNA helicases (7). DEAD box proteins were first classified in 1988 and are characterized by 12 conserved motifs: Q, I, Ia, Ib, Ic, II, III, IV, IVa, V, Va and VI (15,18). The family name originates from the sequence of motif II: D(asp)-E(glu)-A(ala)-D(asp). DEAD box proteins are found throughout all three domains of life and even a few viruses encode DEAD box proteins (19). The human genome alone encodes 37 members of the DEAD box family (20-24).

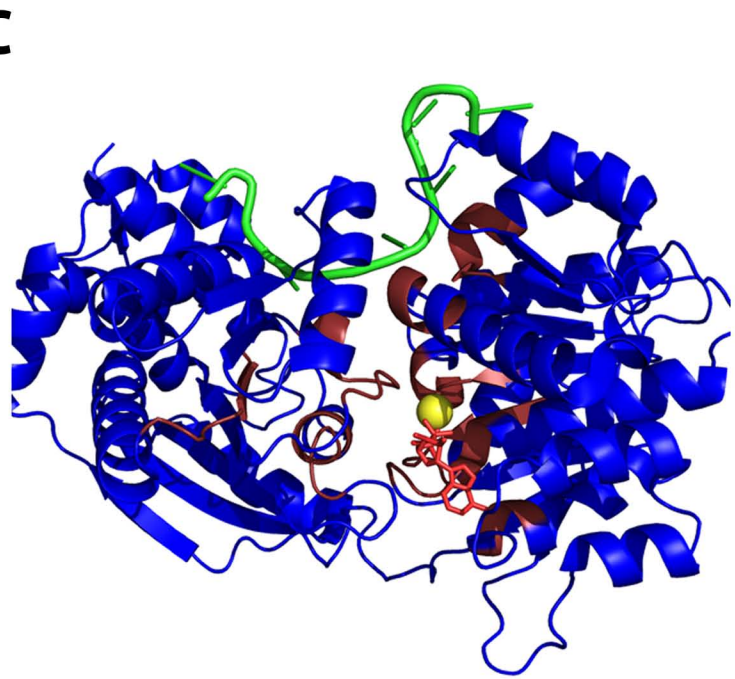
The twelve conserved motifs of DEAD box proteins make up a core region of approximately 350 to 400 amino acids (Figure 1.1). DEAD box proteins do not have other identifiable motifs within their core region, with the exception of DEAD box 1 (DDX1) and DDX24 (25,26). Although some DEAD box proteins only consist of the core region, most have amino (N) or carboxy (C) extensions which contribute to the diversity of functions associated with these proteins.

Figure 1.1 Structure of the DEAD-box helicase core.

(A) Schematic representation of the primary structure of the DEAD-box helicase core. Helicase domain 1 and 2 designate the two RecA-like helicase domains. Consensus sequences of the conserved motifs are shown. Motifs are colour labeled based on their functions: ATP binding (purple), RNA binding (blue) and domain coupling (cyan). (B) 3D representation of eIF4A indicating the location of the conserved DEAD box protein motifs. Nine motifs are highlighted in the N-terminal (grey) or C-terminal (yellow) domains. The dumbbell shape can be observed with the flexible linker (black) spanning the cleft between the two domains. The eIF4A structure was adapted from Schutz *et al.* 2008 (27). (C) Structure of DEAD box protein DDX19 (blue and maroon) bound to single-strand RNA (green), Mg^{2+} (yellow), and ADP (red). Conserved motifs are highlighted in dark red with the N-terminal domain on the right and the C-terminal domain on the left. The RNA binds across the cleft that separates the two domains. The structure was generated from [PDBID: 3G0H] using Pymol.


A
Helicase Domain 1
Helicase Domain 2
F..16..GaccPohIQ
LaPTRELAxQi
VaTPGRLlD
fSATc
IIFhxT+cx
LiaTDVv
YYHRIGRTGRaGr
N
Q
I
Ia
Ib
Ic
II
III
IV
IVa
V
Va
VI
C
AxTGoGKTaaF
xGG
VLDEADrm
HGdmxQxeRx
ARGIDlpxVxxViNyD
ATP binding
RNA binding
Domain coupling


B
Motif Ia
Motif Ib
DEAD
Motif I
Motif III
C
N
Motif IV
Motif V
Motif VI


C

1.1.2 DEAD box protein structure

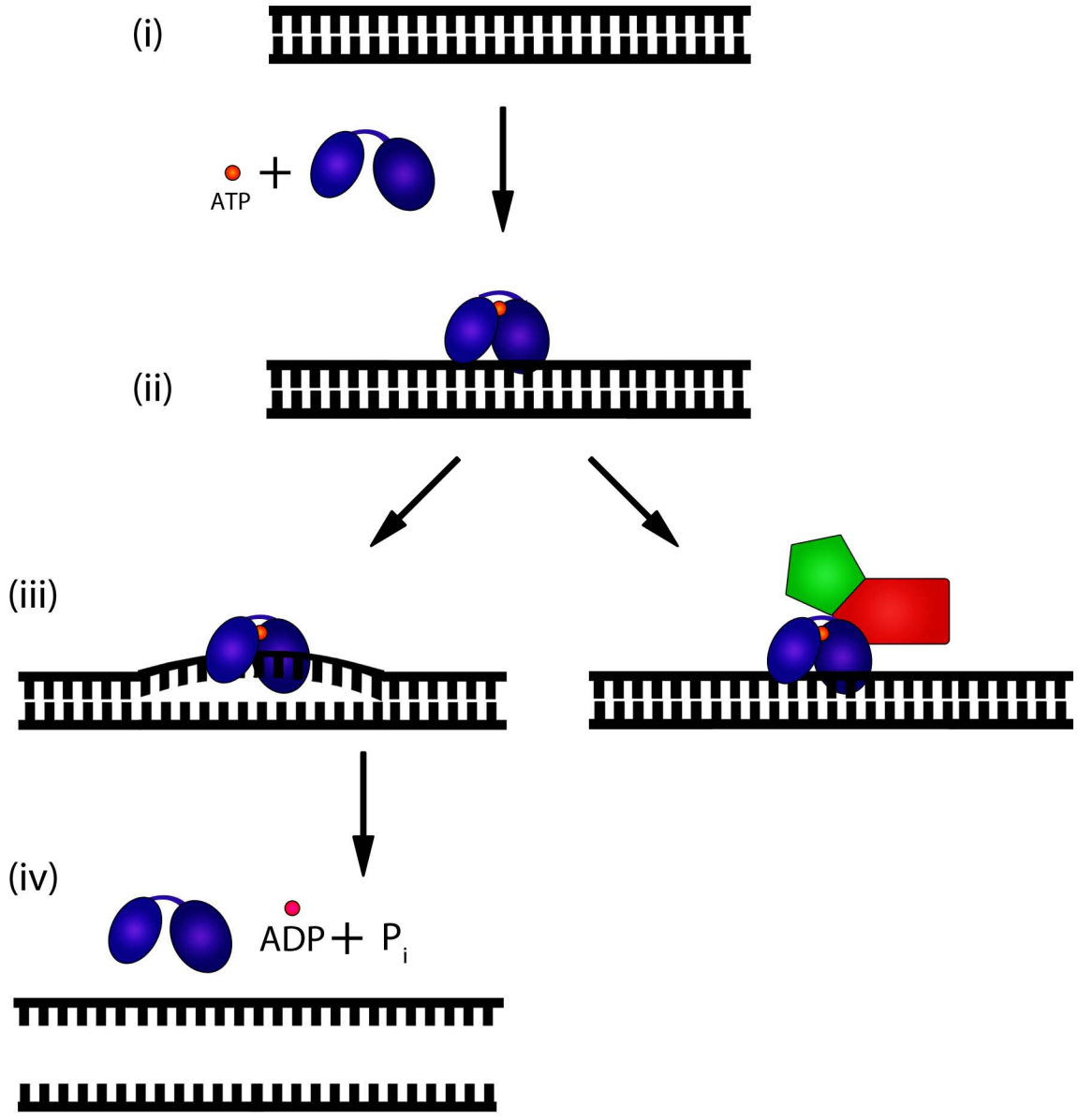
Structurally, DEAD box proteins resemble other non-ring helicases which fold into two tandem RecA-like domains (28,29). The first structures of DEAD box proteins (eIF4A and mjDEAD) were solved in 2000 and 2001, respectively (29,30). eIF4A and mjDEAD are two of the shortest DEAD box proteins and consist mainly of the DEAD box core. Analysis of these two structures revealed an N-terminal RecA-like domain which includes motifs Q – III and a C-terminal RecA-like domain which includes motifs IV – VI. Motifs III and Va are primarily responsible for domain coupling to stabilize a cleft formed between the two RecA-like domains (31-33). The ATP binding domain resides within this cleft. The Q motif is responsible for adenine base specificity while coordinating with motifs I, II and VI and a Mg^{2+} ion to stabilize the phosphate groups and complete the ATP binding pocket (34-36). Motifs Ia, Ib, Ic, IV, IVa and V are found on the opposite side of the cleft. Different residues in motifs IVa and V are exposed allowing interaction with the phosphate groups of the sugar-phosphate backbone of the RNA. This type of interaction suggests little to no RNA sequence specificity. Specific N- and C-terminal domains found in many DEAD box proteins likely play important roles in the recognition of specific substrates through protein-protein or protein-RNA interactions.

1.1.3 DEAD box protein functions

DEAD box proteins can modify RNA secondary structures or remodel ribonucleoprotein (RNP) complexes (Figure 1.2). In most cases, ATP first binds to the

Figure 1.2. A model of how DEAD box proteins modify RNP complexes and unwind RNA duplexes.

(i) Binding of ATP induces a high affinity RNA binding pocket in DEAD box proteins. (ii)-(iii) DEAD box proteins then bind to RNA and either induce localized strand separation (left) or act as a clamp to modify the RNP complex (right) by either recruiting other proteins (shown in red and green) or dissociating proteins already bound to the RNA (not shown). (iv) In the case of strand separation, ATP is hydrolysed following strand separation, which reduces the RNA binding affinity of the DEAD box protein and allows the release of the RNA.


(i)
ATP
(ii)
(iii)
(iv)
ADP + P i

DEAD box protein at the cleft located within the DEAD box protein core. ATP binding shifts the conformation of the other motifs resulting in the formation of an RNA binding domain (37-39). The protein then binds to double-strand (ds) RNA duplexes through a poorly understood loading mechanism. In most cases, this loading requires a single-strand (ss) RNA overhang which serves as a docking site for the DEAD box protein, with subsequent transfer of the protein to the dsRNA substrate. After docking to the RNA, DEAD box proteins can take two different paths depending on whether their role is to unwind RNA substrates or remodel RNP complexes.

Those family members that modify RNP complexes remain clamped to the RNA duplex and recruit other proteins to the complex and/or modify the binding of other proteins to the RNA substrate (39-41). DEAD box proteins that unwind RNA substrates do so by inducing localized strand destabilization (16). This is in contrast to other helicases which unwind their nucleic acid substrates in a processive manner by moving along the substrate (5,31,42-44). When bound to dsRNA, DEAD box proteins with RNA unwinding activity force the dsRNA into an unstable conformation which causes the two strands to dissociate. When the strands separate, bound ATP is hydrolysed to ADP which decreases the affinity of the DEAD box protein for the RNA substrate. At this point, both the RNA substrate and ADP are released from the protein and the cycle is repeated depending on availability of substrate. Several DEAD box proteins have been shown to have unwinding activity *in vitro* including eIF4a (DDX2), An3 (DDX3) and Vasa (DDX4) (45-47). In most cases, DEAD box proteins are able to unwind duplexes that are 9-15 base pairs in length with a single ATP molecule (47,48).

There are a few examples of protein-protein interactions that enhance the unwinding activity of DEAD box proteins. The prototypic DEAD box protein, translation initiation factor eIF4A (DDX2), is a key component in the translation initiation complex (49,50). eIF4A unwinds 5′ double strand mRNA structures during translation initiation (51). *In vitro* studies have demonstrated that eIF4A can unwind short RNA-RNA and RNA-DNA duplexes (52). eIF4A on its own has a low affinity for RNA when compared to other DEAD box proteins. However, when eIF4G binds to eIF4A, both RNA affinity and ATPase activity are enhanced (27). A second DEAD box protein, Dbp5 (DDX19), has also shown enhanced activity when bound to Gle1 (53). Of note, Gle1 and eIF4G both bind to the same region of their respective DEAD box proteins. Based on X-ray crystallography, Gle1 and eIF4G have similar structures despite their different sequences and functions (18). Other DEAD box proteins may also have co-factors that interact in this manner to either enhance or reduce their activity.

1.2 DDX1

1.2.1 Identification of DDX1

DEAD box 1 (DDX1) was first identified by screening a subtracted cDNA library prepared using human fetal tissue/LoVo cancer cell-depleted RNA from retinoblastoma cell lines (54). *DDX1* is located on chromosome 2p24, approximately 340 kb telomeric to *MYCN* (55,56). DDX1 orthologues are found in all animals, a few

species of algae and even in a few protists; however, it is absent in prokaryotes and fungi (Figure 1.3).

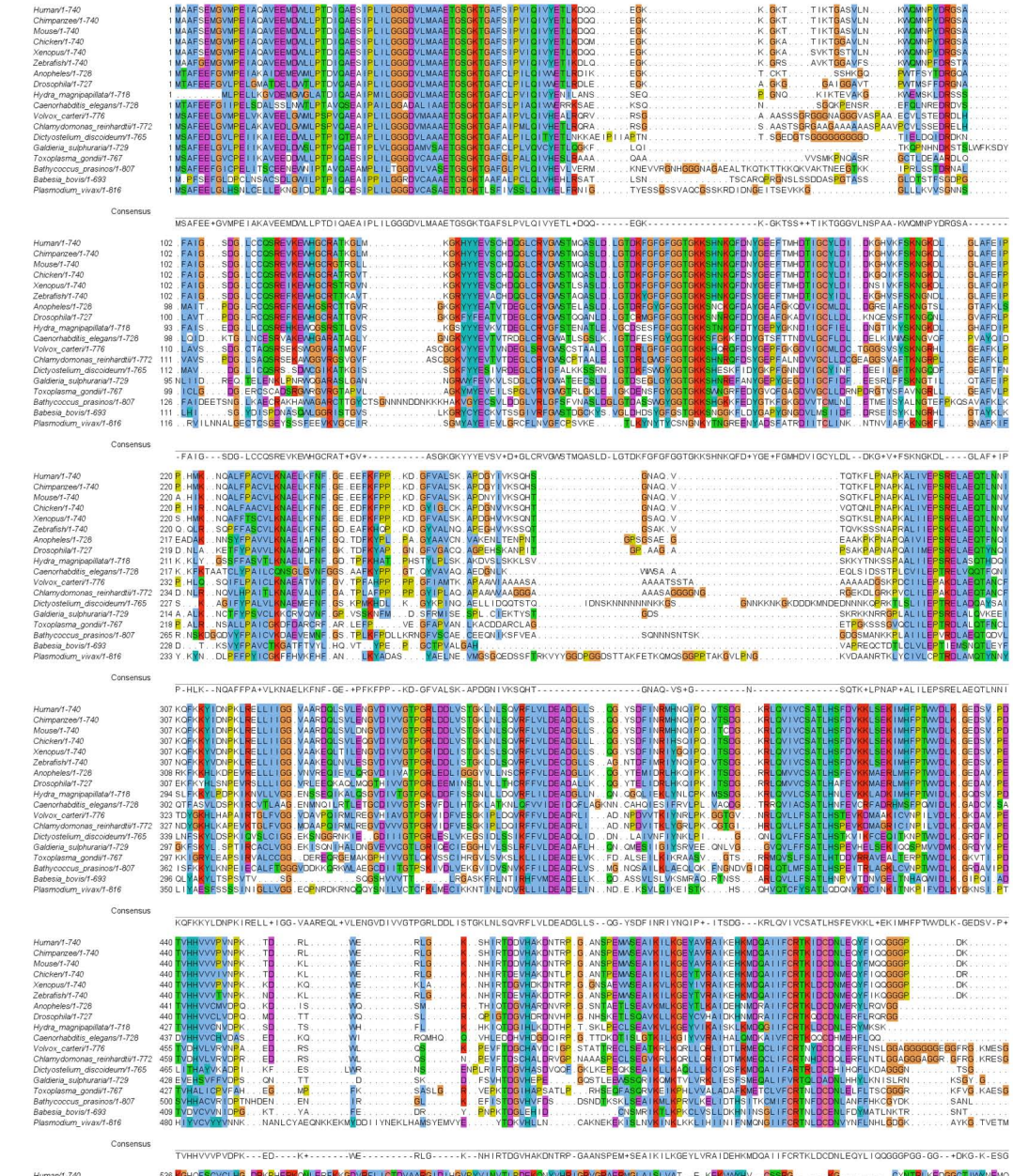
DDX1 shares high conservation with other DEAD box proteins over the twelve conserved motifs (Figure 1.4). Unlike other DEAD box proteins which have unique sequences both N- and C-terminal to the helicase core, the Q domain of DDX1 is found immediately next to the start codon. The C-terminus of DDX1 (downstream of the helicase core) has no homology to any other domains or proteins. Thus, the C-terminus of DDX1 may be associated with specific activities related to the function of DDX1. DDX1 is also one of only two DEAD box proteins (the other being DDX24) with an extra domain located within its core region. This internal domain was initially found to have high homology to hnRNPU, another RNA binding protein (26). This domain has been identified as a SPRY domain and is 130 amino acids long.

1.2.2 The SPRY domain

The SPRY domain is named for the combined dual specific kinase **sp**1a of *Dictyostelium* and rabbit **ry**anodine receptors (57). SPRY domain containing proteins have been shown to play roles in RNA metabolism, calcium signaling, cytokine signaling and immune response (58-66). The SPRY domain is found throughout the phylogenic tree, including all three domains of life, as opposed to the related B30.2 domain which has only been identified in human, mouse, chicken, and frog (58). Thus the SPRY domain appears to be ancestral to the B30.2 domain, with the latter being

Figure 1.3. Sequence comparison of DDX1 orthologues.

The amino acid sequences of several DDX1 orthologues were aligned using the MAFFT sequence alignment program with its global alignment mode. Sequences were displayed and coloured using Jalview v2.8.0b1.



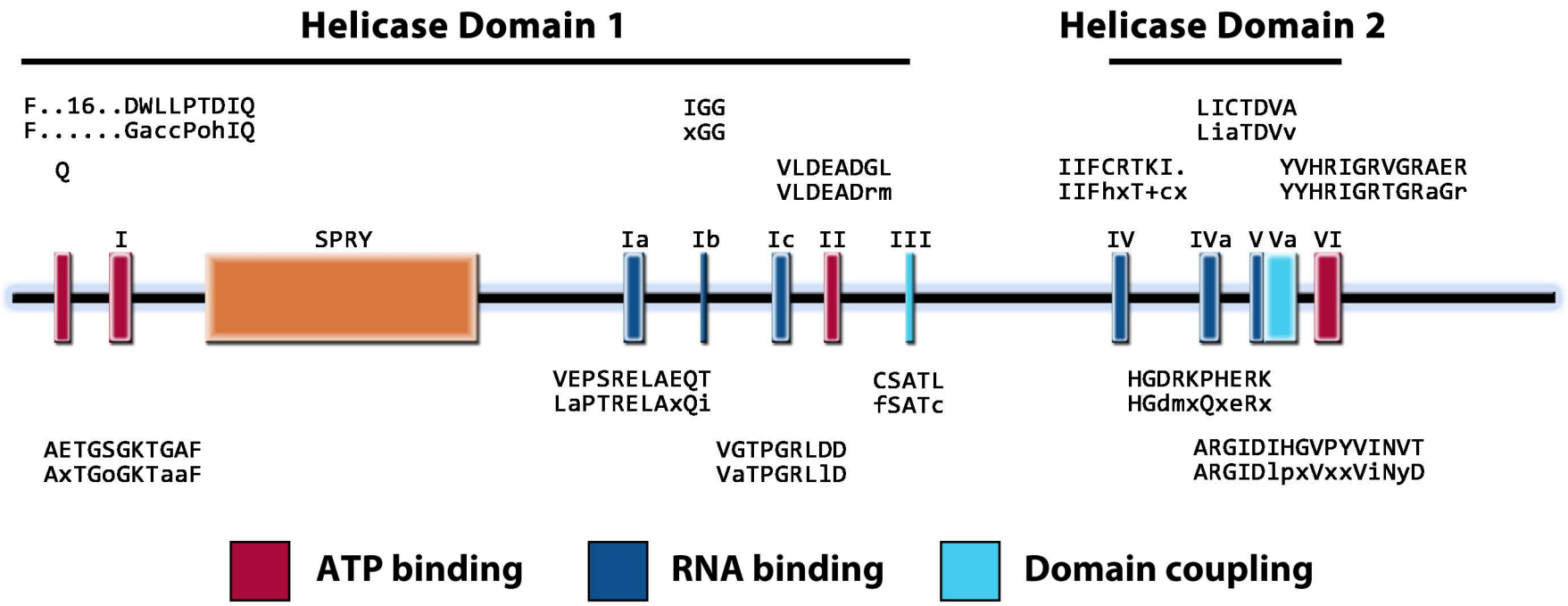


Figure 1.4 DDX1 motif structure.

DDX1 has the 12 conserved motifs found in all DEAD box proteins as well as a SPRY domain. Each motif is coloured to match its role in ATP binding (red), RNA binding (blue) or domain coupling (cyan). The amino acids spanning each conserved motif are indicated with DDX1 being the upper sequence and the consensus sequence the lower. In motif Q there is a 16 amino acid gap between the first and second conserved amino acids. Figure is drawn to scale.

larger in size with an extended N-terminal end (58). In humans, over 100 proteins containing a SPRY domain have been identified (65). The role of the SPRY domain in DDX1 is still unknown, but it has been proposed to be involved in protein-protein or protein-RNA interactions (67-69). The SPRY domain consists of two layers of β-sheets, with N- and C-terminal ends located structurally close to one another (70). Interactions Based on sequence similarity and multiple solved structures, it is believed that the SPRY domain is modular in nature. The SPRY domain consists of two layers of β-sheets, with N- and C-terminal ends located close to one another structurally (70). Compared to the SPRY domain, the B30.2 domain has additional α-loops at the N-terminal end. Interactions between the SPRY domain and its binding partners are thought to take place along the outer loops that connect the β-sheets. As an example, the SPRY domain of GUSTAVUS, a protein found in cytoplasmic ribonucleoprotein particles, interacts with a 20 amino acid sequence found in DDX4 (VASA) (69). The region of GUSTAVUS that associates with the 20 amino acid DDX4 sequence is conserved structurally amongst other SPRY domain containing proteins. It has been postulated that the overall core structure of SPRY domains fulfills a common role with sequence differences outside the core providing specificity of binding to short peptide sequences.

1.2.3 DDX1 in cancer

DDX1 was originally identified by screening a retinoblastoma cDNA library, and was subsequently found to be amplified and overexpressed in a subset of

retinoblastoma (RB) cell lines (71). RB is a tumour of the retina that predominantly arises in young children before age 5 with an incidence of 1/20,000 live births (72). If the tumour is relatively small when diagnosed, it can be treated with radiation, chemotherapy and/or cryotherapy in order to preserve the eye (73-77). However, when the tumours are large, the eye is removed by surgery. Although the cure rate for retinoblastoma is very high (>90%), the treatment of these young patients remains a challenge (78-80). Until recently, it was believed that the underlying cause of all retinoblastoma was loss of the tumour suppressor gene *RB1*. Recently, a subset of retinoblastoma tumours have been identified which have two normal copies of the *RB1* gene and apparent normal pRB function (81). These retinoblastoma tumours all showed amplification of the *MYCN* gene.

We have found that *DDX1* is co-amplified with the *MYCN* gene in all *MYCN*-amplified retinoblastoma cell lines tested to date (4/4) [(71,82); Godbout R., unpublished data]. *DDX1* maps to chromosome band 2p24, just 340 kb telomeric to *MYCN* (56). *MYCN* amplification is common in another childhood tumour, neuroblastoma (NB). NB is a tumour of neural crest cells which arises in a variety of different tissues (83), with its most common site being the adrenal gland (84). Two thirds of NB tumours and cell lines with *MYCN* amplification show *DDX1* co-amplification (85). As *DDX1* is located 340 kb from the *MYCN* gene and the *MYCN* amplicons range from 300 kb to 1 Mb in size, this suggests that *DDX1* is co-amplified with *MYCN* due to its close proximity. In addition to *DDX1,* two other genes are commonly found in *MYCN* amplicons: *NAG* and *NSEI (86-88)*. Although genes other

than *DDX1* have been found to be co-amplified with *MYCN,* Southern, northern, and western blot analyses have revealed a direct correlation between *DDX1* gene copy number, *DDX1* RNA, and DDX1 protein levels in retinoblastoma and neuroblastoma cells (71). As amplified genes that do not provide a growth advantage to the cancer cells are often mutated and not expressed, abundant expression of DDX1 in *DDXI*-amplified tumours suggest a role for DDX1 in tumour formation or progression.

DDX1 was subsequently found to be co-amplified with *MYCN* in medulloblastoma, alveolar rhabdomyosarcoma, rat uterine endometrial carcinoma tumours and cell lines, and a subset of glioblastoma multiforme tumours maintained as xenografts in mice (54,56,82,89-92). Although the role of DDX1 in NB remains controversial, the largest study carried out to date suggests that *MYCN* and *DDX1* co-amplification correlates with a better prognosis than *MYCN* amplification alone (93,94). The significance of co-amplification of *DDX1* and *MYCN* in other tumours has not been investigated.

DDX1 overexpression is a prognostic indicator for early recurrence in primary breast cancer (95,96). Elevated levels of *DDX1* mRNA were correlated with early recurrence and death in primary breast cancer regardless of molecular subtype. Immunostaining analysis of a 120-patient breast cancer tissue microarray identified high cytoplasmic DDX1 as an independent marker of recurrence and death (96). DDX1 may thus be playing a role in the cytoplasm that promotes tumour progression. A conflicting report indicates that DDX1 is associated with increased survival in breast cancer (97). The latter study involved a different population of breast cancer patients,

early stage and node-negative, as opposed the less restricted populations of patients used in the former studies (97). Surprisingly, close to 50% of the breast cancer tissues analysed by Taunk, *et al*. were labeled as DDX1-negative. This is in contrast to our study where virtually every tumour expressed DDX1, albeit at different levels and with different subcellular distribution patterns (96). The discrepancy in our results may therefore reside in the quality of the breast cancer tissue analysed and/or the specificity and affinity of the anti-DDX1 antibody used for these experiments.

1.2.4 Expression of DDX1 in vivo

Previous *in vitro* analysis found ubiquitous expression of DDX1 in all cell lines tested (54,71). In chick embryos, DDX1 was detected in all tissues analyzed: retina, brain, heart, liver and kidney (98). The highest levels of DDX1 were observed in the retina and brain. By examining chick tissues at different stages of development, DDX1 was shown to be more highly expressed during early development. DDX1 protein levels decreased from embryonic day E4 to E16 in chick heart, liver and kidney. Knockout of the *Ddx1* gene in *Drosophila melanogaster* was serendipitously generated a number of years ago and reported to cause early embryonic lethality (99). Examination of *Ddx1* in *Drosophila* revealed highest levels of *Ddx1* RNA in early (0-2 hour) embryos with decreased expression observed in later stage embryos (100). These data suggest the need for elevated levels of maternal *Ddx1* in *Drosophila*. Together, these data suggest a role for DDX1 in highly proliferating cells both in early development and in tumours.

1.2.5 Localization of DDX1

DDX1 protein is detected in both the nucleus and cytoplasm of all cell lines tested to date (101). DDX1 is considerably more abundant in the nucleus of most cells examined, although absent from the nucleoli, and has a punctate distribution forming 3-4 bright foci (101). The DDX1 foci are present throughout the cell cycle with the exception of mitosis (101). The disappearance of DDX1 foci during mitosis is not accompanied by a decrease in DDX1 levels suggesting redistribution of DDX1 protein rather than degradation.

DDX1 foci are approximately 0.5 μm in diameter and have been shown to associate with other nuclear bodies such as cleavage bodies, Cajal bodies and gems (101,102). Cleavage bodies are nuclear foci of 0.5-1.5 μm in diameter that are dynamic in nature. These bodies are present during G1 and S phase and decrease or are absent in G2 and mitosis. These foci contain components of pre-mRNA processing, namely those associated with 3′ cleavage and polyadenylation. Antibodies to CPSF and CstF-64 proteins are commonly used to identify cleavage bodies (101-103). Both these proteins are essential for the coupling of transcription termination to the 3′ processing of pre-mRNA. Cleavage bodies co-localize with Cajal bodies which have been associated with various aspects of RNA metabolism including snRNP and snoRNP biogenesis, snRNA posttranscriptional modification, pre-rRNA processing and histone mRNA processing (104-106). Proteins used to identify Cajal bodies include Sm (107) and p80-coilin (108). Similar to cleavage bodies, Cajal bodies are found primarily during the G1 phase of the cell cycle and decrease as the cells progress to G2 (109). No

foci are present during mitosis. Gems are another type of nuclear body which associates with cleavage bodies and Cajal bodies. Gems are normally found adjacent to cleavage bodies and contain the SMN protein (110). Gems are believed to perform similar roles as Cajal bodies and are associated with snRNP biogenesis. The association between DDX1 and these bodies suggests that DDX1 may play a role in 3' processing or posttranscriptional modification of different RNAs.

Although the levels of nuclear DDX1 are generally higher than cytoplasmic DDX1 in many cell types, DDX1 has been found to associate with several RNA containing granules located in the cytoplasm (111-113). In neuronal cells, DDX1 is found in RNA containing granules called the RNA transport complex. These RNA granules shuttle RNA from the cell body along the axons and dendrites for localized translation in neurons (111,112). DDX1 also co-immunoprecipitates with components of RNA containing granules such as the purine rich element binding protein alpha (Purα), staufen, and EF-1α (111). Under conditions that induce stress, DDX1 is recruited to stress granules that contain stalled mRNA translation complexes (113-115).

High levels of DDX1 can also be found in the cytoplasm of *DDX1*-amplified and overexpressing retinoblastoma and neuroblastoma cancer cell lines and tumours (85). As previously mentioned, elevated levels of DDX1 in the cytoplasm of breast cancer tissue are associated with a poor prognosis (96). These data suggest a cytoplasmic role for DDX1 in cancer cells that is associated with tumour progression.

Infection with some viruses can result in dramatic redistribution of DDX1 within the infected cells. For example, infection of Vero cells with a coronavirus induces relocalization of DDX1 from the nucleus to the cytoplasm (116). Although DDX1 redistribution is only associated with specific viruses, DDX1 has been shown to play an important role in the efficient replication of several different types of viruses, including HIV-1, JC virus, and corona virus (116-121).

1.2.6 Biochemical activity of DDX1

A number of DEAD box proteins can unwind small stretches of double-strand RNA *in vitro* (43,45,46,48,122,123). Like other DEAD box proteins, DDX1 can unwind RNA-RNA duplexes *in vitro* (117,124). DDX1 can also unwind RNA-DNA duplexes *in vitro*, a property associated with a few DEAD box proteins such as eIF4A (125). DDX1's ability to unwind double-strand RNA-RNA and RNA-DNA duplexes does not require the presence of ssRNA overhangs indicating that DDX1 does not require ssRNA to initiate the loading process. As well, DDX1's unwinding activity is dependent on ADP rather than ATP, and DDX1 is able to unwind much longer duplexes (29 bp) than most other DEAD box proteins (9-15 bp) (124). While unwinding experiments were being conducted, it was discovered that DDX1 also has ribonuclease activity. When using substrates with ssRNA overhangs, rapid degradation of the ssRNA was observed leading to the generation of a blunt double strand substrate. This ribonuclease activity is energy independent but Mg^{2+}-dependent and heat sensitive (126).

1.2.7 Role of DDX1 in DNA repair

Cells are continuously exposed to agents that cause different types of DNA damage, including single-strand breaks and double strand breaks. The most detrimental DNA damage is the double-strand break. There are two major pathways to repair this type of damage: homologous recombination (HR) or non-homologous end joining (NHEJ) (127,128). Repair by HR requires proximity to the undamaged DNA template and occurs after the DNA has been replicated during the S phase of the cycle. Repair by NHEJ does not require a template as the two ends of the double-strand break are ligated together. NHEJ thus represents an error-prone form of repair whereas HR repair is usually error-free.

HR repair of double strand breaks is a complex process that requires a large number of proteins working in concert to carry out the repair process. Briefly, when double strand breaks occur, the MRN complex is recruited at the site of the break. The Ataxia Telangiectasia Mutated (ATM) protein is then recruited to the break site and phosphorylates several proteins to stimulate their activity or to act as scaffolds for the next stage in repair. One of the major targets of ATM is the histone 2A isoform H2AX. When phosphorylated, γ-H2AX acts as a scaffold to recruit downstream effectors such as CHK2 and p53 in order to delay cell cycle progression. CtIP is recruited to the site of damage to initiate 5′ to 3′ end resectioning, which is then carried out by exonuclease 1 (Exo1). The resulting ssDNA becomes bound by RPA and with the assistance of other factors, the ssDNA pairs to the sister chromatid sequence. This strand invasion results in the formation of a structure called a Holliday junction which is resolved when DNA

is synthesized and replaces the resected DNA thereby completing the repair of the break. Although the model of DNA double strand break repair is well established, the full spectrum of proteins involved in efficient DNA repair is not known.

There are 3-4 DDX1 foci per nucleus in a cell cultured under normal conditions (101). When cells are exposed to 5 Gy γ-radiation to induce DNA double strand breaks, the number of DDX1 foci increases to 30-40 per cell (124). The expression levels of DDX1 do not change following exposure to ionizing radiation (IR), indicating that these foci are formed by redistribution of DDX1. DDX1 IR-induced foci co-localize with a subset of γ-H2AX foci (30%), a marker of DNA double strand breaks. When cells are treated with RNase H to degrade RNA from RNA-DNA duplexes, these foci disappear (124). These data suggest that DDX1 localization to sites of double strand breaks requires the presence of RNA-DNA hybrids. We postulate that the double strand breaks to which DDX1 is recruited are in the vicinity of active transcription, and that the RNA-DNA duplex structures that naturally occur at these sites need to be cleared so that efficient DNA repair can occur. Unpublished data from the lab indicate that knockdown of DDX1 decreases HR repair efficiency, with DDX1 playing a role in end resection (Lei Li et al., unpublished data). Our data suggest that removal of RNA at DNA double strand breaks by DDX1 promotes repair by HR (Lei Li et al., unpublished data).

1.2.8 Role of DDX1 in RNA storage, turnover and transport

Cytoplasmic mRNA is associated with many different proteins during its lifespan. In some cases, mRNAs are packed together in discrete foci for delayed site-specific translation, translation stalling due to stress, or degradation (129). There are several different types of RNA containing foci including germ cell granules, chromatoid bodies, neuronal transport granules, stress granules, and processing bodies (111,112,129-134). Here, we focus on neuronal transport granules and two related mRNA containing granules found in cells: processing bodies (PBs) and stress granules (SGs).

Neuronal granules are RNP complexes found in neurons that contain stalled RNA transcripts required for translation at distant sites in the cell (111). Neurons have processes called dendrites and axons that need to form specific connections with other cells. Axons in particular have to cover long distances, over one meter in some cases, in order to reach their final destination. Rather than transporting proteins along the lengths of axons and dendrites, translationally stalled mRNA is packaged in RNA granules and translated at localized sites. These RNA granules contain mRNA, translation initiation factors, and other RNA binding proteins (e.g. Staufen, DDX3, DDX5 and CGI-99) (111,112,135,136). Some RNA granules also contain ribosomal subunits to further facilitate translation. Upon reaching their destination (e.g. synapses), the content of the RNA granules is released in response to external stimuli and the mRNAs undergo localized translation.

Different types of neuronal RNA granules have been identified based on their RNA and protein contents, with specific DEAD box proteins associated with different types of neuronal granules (112). DDX1- and DDX3-containing neuronal granules are associated with ribosomes. Granules that contain DDX1 and CGI-99, but not DDX3, are not associated with ribosomes. DDX6-positive granules also contain ribosomes, a surprising finding since DDX6 has been shown to be part of the processing bodies associating with mRNA decay (described below). It has been proposed that RNA transcripts containing ribosomes are more readily translated following transport as they have already been partially assembled onto the translation machinery.

Processing bodies (PBs) contain proteins associated with the microRNA (miRNA) pathway (Ago2, DDX6), 5′ decapping enzymes (DCP1/2), and 5′ to 3′ exonucleases (XRN1 and Lsm) (137-141). These proteins are part of the RNA interference and mRNA degradation pathways indicating that PBs play a role in mRNA silencing and degradation. PBs can be found in both stressed and unstressed cells. Different cell types have different numbers of PBs; for example, HeLa cells have several PBs in each cell, whereas U2OS cells only have PBs in 20-50% of cells (115). The abundance of PBs also changes during development. PBs disappear during oocyte maturation, but reappear during early stages of embryogenesis when maternal transcripts are undergoing degradation (142). PBs have also been shown to be dynamic structures that can move throughout the cytoplasm of embryos (143). Several components of PBs are also found in stress granules (discussed below), including Staufen (associated with RNA transport) and CPEB (associated with translational

control) (144,145). Drugs that induce SGs also induce the formation of PBs. In some cases, PBs have been found to interact with SGs (115,146). It has been proposed that transcripts are shuttled between PBs and SGs to target different transcripts for degradation; however, it's still unclear if mRNAs are actually transferred between PBs and SGs.

SGs are large (100-2000 nm) cytoplasmic bodies that form when cells are exposed to heat, UV, hypoxia, viral infection and some chemicals. SGs consist of stalled translation pre-initiation complexes bound to RNA binding proteins such as Staufen and TIA-1, as well as RNA degradation proteins such as Lsm1 (a 5′ to 3′ exonuclease also found in PBs) (113,114,132,146,147). The initial step in SG formation is phosphorylation of eIF2α which depletes the levels of eIF2α-GTP-$tRNA_i^{Met}$ preventing the large ribosomal subunit from binding to transcripts (147). As a result, the number of ribosomes that run off the mRNA transcript is increased compared to the number of ribosomes that loads onto the transcript. The shift in this ratio results in polysome disassembly and low levels of translation. Transcripts are only stored transiently in SGs, being shuttled back and forth between SGs and the cytosol (148). Together, SGs and PBs play an important role in the post-transcriptional regulation of mRNA. They contain similar protein constituents; however, SGs are primarily sites of temporary mRNA storage while PBs appear to play a more direct role in RNA silencing and degradation. DDX1 localizes to stress granules upon arsenite treatment or heat shock [(113), Lei Li, unpublished]. The role of DDX1 in these granules may be to stabilize the stalled transcripts.

In summary, DDX1 has been found to associate with neuronal RNA transport granules and stress granules. Although DDX1 has not been reported in PBs, similar protein components have been found in SGs and PBs, suggesting that DDX1 may also be found in PBs (111-113). In light of the known biochemical properties of DDX1 (nucleic acid unwinding and RNA degradation), it is possible that DDX1 associates with RNA containing granules and can assist in either the stabilization or degradation of transcripts depending on cellular context.

1.3 Mouse development

1.3.1 Gametogenesis

Early mouse development is a highly regulated event that begins in the gametes generated by the parents. The paternal contribution to the zygote is the sperm. This specialized cell carries the compacted paternal genome to the oocyte for fertilization. In mice, sperm development begins in the seminiferous tubules of the testes as a stem cell (spermatogonium) that divides via mitosis to produce two primary spermatocytes (149). The primary spermatocytes divide via meiosis I to yield secondary spermatocytes which in turn undergo meiosis II to yield spermatids. Cell division is incomplete during these stages leaving all the spermatids interconnected to transport RNA or proteins in a passive or active manner. The spermatids then undergo a final metamorphosis from round cells to individualized elongated spermatozoa. All of these steps occur while in contact with Sertoli cells which provide support and signalling. As the spermatid undergoes elongation, the Sertoli cells

phagocytose the spermatid cytoplasm that is shed. The final stages of maturation occur once the spermatids leave the seminiferous tubules and move along the epididymis where they are stored until mating.

Transcription is active during the initial stages of sperm development; however, during elongation the chromatin begins compacting and transcription stops. All mRNAs required for further development are stored in various RNP complexes (150-152). Protamines are small proteins that replace the histones in order to compact the genome into a size smaller than that of the somatic nucleus (153). Genome compaction is essential for proper maturation as mouse knockouts of genes involved in compaction such as the transition nuclear protein genes (*TP1* or *TP2*) or protamine genes (*Prm-1* or *Prm-2)* result in male infertility (154,155). The replacement of histones with testes-specific histone variants occurs before and during meiosis, ultimately generating a genome that is almost completely packed by protamines (156). In mice, 99% of histones are replaced with protamines with the few remaining histone-containing nucleosomes concentrated around the centromeres and telomeres (157,158). Human sperm retain a higher percentage of histones (10-15%) (159). It has been postulated that this higher retention of histone-containing nucleosomes is associated with specific genes related to embryo development (160-162).

In addition to mRNAs required for development of the embryo, various miRNAs have been identified in mature sperm although their roles in sperm remains poorly understood (163,164). There is evidence that some of these miRNAs may regulate gene expression in the developing embryo by causing altered epigenetic marks at

different genes as discussed below (165-168). The role of the sperm outside of genome transfer is poorly understood although there is increasing interest in the possibility that RNAs that are transferred to the oocyte upon fertilization play important roles in embryo development (163,166,169,170).

The oocyte is the maternal contribution to the zygote and contains the maternal genome as well as a large complement of RNAs and proteins. These maternal RNAs and proteins support the oocyte, and after fertilization the zygote, until the zygotic genome is activated. The first stage of oocyte development occurs during embryogenesis. Oogonia (primordial oocytes) develop into primary oocytes which arrest at meiosis I. Meiosis resumes upon sexual maturation and estrus cycle stimulation. In mouse, the number of oocytes which resume meiosis depends on the strain and ranges from 6-14 oocytes. As meiosis resumes, the primary oocytes undergo asymmetrical division where most of the cytoplasm is retained by the secondary oocyte and a small amount is ejected as a polar body. Soon after, the oocyte is ovulated and stalls at meiosis II until fertilization.

Oocyte development is accompanied by a large increase in size (from 25 to ~80 μm) as the result of accumulation of large amounts of mRNA, ribosomes, mitochondria, etc., required to support the newly formed zygote (171,172). While the oocyte itself generates many of these factors, it is supported by other cells in the follicle known as granulosa cells. There are two types of granulosa cells: mural granulosa cells which make up the bulk of the follicle and the cumulus cells which surround the oocyte. The cumulus cells are linked by gap-junctions to each other and

to the oocyte allowing the exchange of various metabolites. The cumulus cells are responsible for triggering further maturation of the oocyte. Most transcripts are directly translated, but some are stored for use later in development as discussed below (172).

1.3.2 Fertilization

The act of fertilization triggers a calcium cascade which results in the activation of a number of different pathways. Fertilization by additional sperm (polyspermy) is blocked by two mechanisms (173,174). The first block is a rapid change in membrane potential which prevents additional sperm from interacting with the oocyte surface. The second block, which is slower, occurs through the release of the cortical granule content found adjacent to the plasma membrane of the oocyte. The content of these granules is released into the perivitelline space and expands the glycoprotein matrix that surrounds the oocyte (zona pellucida) and generates an additional barrier that lifts sperm from the oocyte membrane and prevents them from penetrating the zona pellucida. Fertilization also triggers oocyte activation allowing completion of meiosis II which releases the second polar body (175). The fertilizing sperm is transported into the oocyte and the process of removing the protamines begins. The oocyte provides the initial histones that replace the protamines allowing reorganization of the chromatin into nucleosome structures (172). As the protamine-DNA structure is very stable, the removal of protamines is an energy dependent process and directed by maternally expressed proteins. In mice, over the next few hours, the two pronuclei

undergo replication prior to fusion of the gametes (syngamy). Following syngamy, the first mitotic cleavage event occurs.

1.3.3 Maternal to zygotic transition

Regulation of the initial developmental stages is completely controlled by the maternal complement of transcripts and proteins. However, in order for development to proceed, the zygote has to undergo genome activation. The maternal to zygotic transition (MZT) occurs just after fertilization and is mostly complete by the 2-cell stage in mice and by the 8-cell stage in humans and bovines (176-178).

The first component of MZT is the active degradation of oocyte-specific transcripts. The oocyte generates a large pool of transcripts that are necessary to produce the proteins required by the oocyte and the fertilized zygote. Many of these transcripts are stored for long periods of time, ranging from days to months in mice (179). The degradation of these transcripts occurs in a very short time frame which suggests a significant shift in transcript stability. In mice, >90% of the maternal mRNAs have been degraded by the 2-cell stage. This degradation is believed to be one of the developmental triggers that activates embryo transcription (178,180). Transcript degradation is associated with the RNA induced silencing complex (RISC) (181). The endoribonuclease Argonaute 2 (Ago2) is one of the key proteins involved in the degradation of maternal transcripts. Knockout of the *Ago2* gene results in embryonic lethality at the 2-cell stage (181). Several maternal transcripts (such as *Msog* and *Gbx2)* were found to be stabilized rather than degraded in the *Ago2* knockout embryos

which suggests that these maternal transcripts must be actively degraded for the embryo to progress past the 2-cell stage of development (181-183).

The second component of MZT is the reprogramming of embryonic genes through chromatin remodelling. The initial stages of chromatin remodelling take place during gamete formation. As previously mentioned, the sperm nucleus undergoes a unique type of chromatin compaction through the use of protamines rather than histones. The switch from histones to protamines means that paternal histone post-translational modifications cannot be passed on to their offspring. In mice, very few histones are retained; however, 15% of the genome retains paternal histone packing in humans, suggesting that histone post-translational modifications could be inherited from the paternal genome in humans. Following fertilization, protamines are actively removed from the paternal genome by glutathione produced by the oocyte during oocyte maturation (184). Transcription factors such as Sp1 and Oct-4 have been found to interact with the DNA resulting in transcriptional activity detected in the male pronucleus but not the female pronucleus, suggesting that these factors are involved in the chromatin remodelling of the paternal genome (185). The nucleosomes that are re-established in the male pronucleus contain elevated levels of acetylated histones compared to the female pronucleus suggesting transcriptional activation. An additional difference between the male and female pronuclei is their DNA methylation status. The oocyte DNA is highly methylated and becomes passively demethylated following each round of DNA replication as the embryo divides. The

sperm DNA is initially highly methylated but undergoes rapid demethylation prior to S-phase (within 4 hours of fertilization) (186,187).

The third component of MZT is embryo transcription activation to replace maternal transcripts. Several different models have been proposed to describe the mechanism and timing of zygotic gene activation (ZGA). In mice, ZGA appears to occur via a zygotic clock mechanism (179,183,188-193). There is a low level activation of transcription that occurs in the 1-cell zygote (early ZGA) shortly after fertilization. Most transcription that does occur at this developmental stage is specific to the paternal pronuclei. Early ZGA may in part be due to the major chromatin reorganization that occurs in the 1-cell zygote and occurs even if S-phase has not been completed. There appears to be a secondary method of regulation in 1-cell zygotes which results in a delay in translation of the newly expressed transcripts (194). The large burst of transcription associated with late ZGA does not occur until the 2-cell stage, and there is no delay in the translation of these transcripts (195).

One example of a protein associated with ZGA is Zygote arrest 1 (Zar1). Zar1 is an oocyte-specific cytoplasmic protein that is specifically expressed in oocytes and 1-cell embryos. *Zar1* knockout mice fail to develop to the 2-cell stage likely due to a complete block of ZGA (196,197). Another protein associated with ZGA is the Transcription Intermediary Factor 1 α protein (TIF1α) which is initially found in the cytoplasm of the germinal vesicle (GV) stage oocyte, the stalled primary oocyte (198). Following fertilization, TIF1α translocates to the nucleus where it is believed to play a role in ZGA by interacting with chromatin remodelers. Micro-injection of TIF1α siRNA

or antibodies in 2- or 4-cell stage embryos results in developmental arrest, indicating that TIF1α is essential for early gene regulation during ZGA (198).

1.3.4 Imprinting and paramutation

In addition to trans-acting transcription factors that bind to cis-acting regulatory elements, gene expression in the early embryo is also regulated by epigenetic modifications. The epigenetic marks found in the parents are mostly erased during gametogenesis and are reapplied during the early stages of development. Some epigenetic marks are maintained from the previous generation though different mechanisms including imprinting and paramutation.

Imprinting is the best studied mechanism of inherited epigenetic modification. Approximately 80 genes have been found to be regulated by imprinting and these genes tend to be in clusters (199). These clusters are regulated by DNA methylation of a regulatory region known as the imprinting control region (ICR). Non-coding RNA molecules have also been found to associate with these ICRs although their function is still poorly understood. Imprinted genes are generally expressed in a monoallelic manner and the stable transmission of a parental-specific epigenetic mark influences the expression of a cluster of genes. As previously mentioned, during early gametogenesis the genomes of both the oocyte and sperm are demethylated and acquire methylation as they mature. Upon fertilization, the female pronucleus undergoes passive demethylation, while the male pronucleus is actively demethylated. ICRs are differentially methylated between the maternal and paternal

genomes and they maintain their methylation patterns during early embryogenesis. *De novo* DNA methyltransferases Dnmt1, Dnmt3A and Dnmt3B are essential in maintaining ICR methylation patterns while the genome is undergoing global demethylation. *Dnmt1/3a/3b* knockout mice die early in development (200-202). Upon fertilization, either the maternal or paternal ICR influences the methylation pattern of the other so that monoallelic expression is achieved.

Another form of epigenetic inheritance is called paramutation. In contrast to imprinting which results in monoallelic expression, the term paramutation is used when there is an interaction between the two alleles of a single gene so that one allele affects the expression of the other. This type of interaction was discovered in *Zea mays* where plant or seed pigments were found in non-Mendelian ratios in offspring (203). One of the best described paramutations in plants is at the *b1* locus (204). The two phenotypes *B-I* and *B'* generate dark and light coloured plants, respectively. When intercrossed, the *B'* locus interacts with the *B-I* locus and changes it to *B'* resulting in a light coloured plant. The mechanism that changes *B-I* to *B'* is unknown, although there are some models that suggest that non-coding RNA molecules may play a role. Some paramutations in plants are very strong and have been shown to persist continually from generation to generation, while other paramutations are weak and persist over only a few generations (205).

The paramutation type of inheritance has recently been documented in animals (165,206,207). The first paramutation described in mice was associated with the *Kit* gene (165). Knockout of the *Kit* gene results in neonatal lethality. $Kit^{+/-}$ mice

bred in the C57BL/6 background were found to have white spots in their fur, white tail tips and white feet, in contrast to wild-type progeny which were uniformly brown. When *Kit*$^{+/-}$ mice were backcrossed to C57BL/6 wild-type mice, a fraction of wild-type offspring retained the spotting pattern observed in *Kit*$^{+/-}$ mice. When these affected wild-type mice (called paramutated) were crossed to other C57BL/6 wild-type mice, a small proportion of the progeny again showed the spotting pattern characteristic of heterozygous mice despite having two wild-type parents. This 'paramutant' state could be maintained for at least five generations. It was ultimately determined that miRNAs complementary to the *Kit* transcript (miR-221 and miR-222) were responsible for the paramutant state through downregulation of the *Kit* gene. Injection of these miRNAs into mice embryos induced the paramutant state in adult mouse.

Additional studies have shown that injection of other miRNAs in 1-cell embryos can induce other types of paramutations in mice. For example, when miR-1 was micro-injected into embryos, the resulting mice had cardiac hypertrophy (167). This phenotype could be propagated across several generations and was associated with the continued inheritance of miR-1. Of note, inheritance of miR-1 through the sperm was as effective as inheritance through the oocyte (167). The regulation of paramutations and their mechanism of inheritance are under investigation.

1.3.5 Morula and blastocyst development

In mice, blastomere (individual cell of the embryo) cleavage occurs with complete cytokinesis but the blastomeres do not divide in a synchronous manner. This

differs from developing *Drosophila* embryos which do not undergo cytokinesis for several divisions while the nuclei divide in a synchronous manner. At the 8-cell stage in mice, the blastomeres have formed a morula (ball of cells) which begins to form cell-to-cell interactions primarily through E-cadherin. *E-cadherin* knockout embryos develop normally up to the blastocyst stage but fail to progress beyond this stage (208).

Embryos normally develop apical-basal polarity following compaction at the 16-cell stage. The 8-cell morula lacks inner and outer cells as all cells are exposed to the extraembryonic environment. By the 16-cell stage, cells can be defined as "inner" or "outer" cells although the number of cells that make up either of these populations is not necessarily consistent in mice due to variability in the axis of cleavage. The outer cells become trophoblast cells which form the extraembryonic tissues. The inner cells form the inner cell mass which develops into the embryo. Trophoblast cells can be defined by the expression of the Tead4 transcription factor which upregulates a number of genes associated with trophoblast differentiation including the Caudal type homeobox 2 (Cdx2) (209). Pluripotency markers Oct-4, Nanog and SRY-box containing gene 2 (Sox2) are main regulators of inner cell mass development (210-212). The mechanisms underlying cell fate determination in mouse embryo is an active area of research.

The transition from morula to blastocyst occurs at the 32-cell stage. At this developmental stage, trophoblast cells have undergone differentiation and micro-injection of embryonic stem (ES) cells into 32-cell stage embryos results in their almost

complete association with the inner cell mass (213). The expression of water transport molecules increases in trophoblast cells and small fluid filled cavities are formed (214). Coinciding with cavity formation is the expression of tight junction components in the trophoblast cells to form a seal preventing water leakage. The small cavities fuse over time to form a large cavity known as the blastocoel. Once this cavity is formed the embryo is classified as a blastocyst. After an additional 24 hours (E4.5) of maturation, the blastocyst is ready to implant into the uterine wall.

1.3.6 Implantation and early gastrulation

The next stage of development for the embryo is the transition from a free floating mass of cells to a stationary mass implanted into the uterus. The embryo undergoes a number of changes in preparation for implantation. The first step is to "hatch" from the zona pellucida. Although the details are poorly understood, it is believed that several enzymes may play a role in zona pellucida digestion including strypsin and cathepsins which are released by the embryo allowing the embryo to hatch (215,216). Originally, it was believed that the blastocoel also provided pressure to assist in hatching; however, experiments where the blastocoel pressure was artificially reduced showed no decrease in hatching efficiency (217). After hatching, the trophoblast undergoes additional differentiation to form an inner layer (cytotrophoblast) and outer layer (syncytiotrophoblast). The cells in the outer layer fuse to form a multi-nucleated syncytium which is able to invade endometrial tissue, allowing the embryo to implant. The trophoblast layers also begin to develop into the

placenta by eroding the small maternal blood vessels to begin supplying the growing embryo with oxygen and other nutrients.

The inner cell mass undergoes its own differentiation program. The first germ layer to form is the endoderm which is determined by the late blastocyst stage and is based on the positioning of cells within the inner cell mass. The primitive endoderm forms a monolayer along the edge of the inner cell mass that faces the blastocoel. Cell position is key to cell fate determination in the early embryo as there is little cell-to-cell variation in the transcription factors that are expressed in the inner cell mass pre-E3.5 (213). It is only when the blastocyst matures that variations in the expression of transcription factors occur. Key transcription factors in promoting the endoderm cell fate while suppressing the epiblast fate are Gata4 and Gata6 (218,219). Nanog, on the other hand, is responsible for maintaining the epiblast fate (220). It should also be pointed out that the cells in the inner cell mass are very mobile with extensive cellular rearrangements occurring which position endoderm-fated cells at the outer edge of the inner cell mass (221). As the embryo continues to develop and increase in size, the epiblast expands and the embryo takes on an elongated shape to form an egg cylinder. The epiblast undergoes further differentiation as it progresses through gastrulation to form the three germ layers: endoderm, mesoderm and ectoderm. These layers in turn will ultimately differentiate into the various tissues that make up the fully developed mouse.

1.3.7 Knockout of DEAD box genes in mice

One approach to determining the role of DEAD box proteins is to generate mouse strains in which a DEAD box gene has been disrupted. Based on the published literature, seven DEAD box genes have been disrupted in mice to date. The phenotypes of the homozygous knockouts range from gender-specific infertility to preimplantation lethality. One of the earliest DEAD box gene knockouts in mice was that of *Mvh*/*Ddx4* or *Vasa*. Knockout of the *Ddx4* gene produces no phenotype in females; however, male *Ddx4* knockout mice are unable to produce mature spermatids (222). The developing spermatocytes do not differentiate, but rather undergo apoptosis. Similarly, *Ddx25*$^{-/-}$ mice are characterized by male infertility. Similar to *Ddx4* knockout mice, *Ddx25*$^{-/-}$ mice have a defect in spermatogenesis; however, the defect appears later in the differentiation pathway and appears when the round spermatids elongate to form mature spermatids (223).

In contrast to *Ddx4*$^{-/-}$ and *Ddx25*$^{-/-}$ mice which are viable, disruption of the *Ddx58* (*RIG-1*) gene results in an embryonic lethal phenotype with most of the mice dying at E12-14; although, death is also observed perinatally (224). The main defect observed in *Ddx58* knockout mice involves the liver which does not develop properly. These mice also have defects in the interferon response pathway. A second strain of *Ddx58*$^{-/-}$ knockout mice was found to be viable although these mice had a number of defects including the development of colitis after a few months and T-cell disruption (225). One similarity between the two strains of *DDX58*$^{-/-}$ mice was that the interferon response pathway was disrupted.

Ddx5 (p68) and *Ddx17* (p72) are the two most closely related members of the DEAD box protein family, with 78% similar amino acids (226). *Ddx17*$^{-/-}$ mice are fertile; however, their progeny have vasculature defects and die shortly after birth at postnatal day (P)2 (227). In contrast to *Ddx17*$^{-/-}$, *Ddx5*$^{-/-}$ mice die around E11.5. Double knockout of *Ddx5* and *Ddx17* results in even earlier lethality than *Ddx5* knockout, with embryos dying pre-E8.5 (227). Both DDX5 and DDX17 are associated with miRNA and rRNA processing by Drosha and their depletion affects cell survival and proliferation.

Mice with disrupted *ChlR1* (*Ddx11)* die earlier than single *Ddx5* knockout mice, with embryonic lethality observed at E10.5 (228). Gross defects are observed prior to this developmental stage and include an improperly formed placenta. In agreement with a role for DDX11 in maintaining sister chromatid and centromeric cohesion, *Ddx11*$^{-/-}$ embryos have significantly increased levels of aneuploidy.

The only DEAD box gene associated with a pre-implantation lethal phenotype is *Ddx20* (*DP103*), with *Ddx20*$^{-/-}$ embryos dying at the 2-cell stage (229). DDX20 is one of the proteins that is upregulated during MZT and has been postulated to be associated with the expression reprogramming that occurs at this stage of development (229,230). As DEAD box proteins have been linked with virtually every aspect of RNA metabolism, the wide variety of different phenotypes observed upon their disruption is not unexpected. Further characterization of the mechanisms underlying the phenotypic defects observed in DEAD box gene knockout mice will lead to a better understanding of the specific functions of these ubiquitous yet poorly understood proteins.

1.4 Thesis objectives

1.4.1 Chapter 3: DDX1 structure and enzymatic activity

Several DEAD box proteins have had their structures solved using X-ray crystallography. These structures have generated insight into how DEAD box proteins interact with and modify RNA structures. DDX1 is one of two proteins containing an additional domain (SPRY) within the conserved DEAD box protein core. The role of the SPRY domain in DDX1 is not known but may be involved in RNA interaction. The first component of Chapter 3 was to determine the structure of DDX1 both with and without bound RNA-DNA duplex substrate using X-ray crystallography and to compare these structures to other solved DEAD box proteins. The second component of Chapter 3 was to pursue the biochemical characterization of DDX1 by studying its ribonuclease activity and determining what region of the protein is responsible for this activity.

1.4.2 Chapter 4: Characterizing the Ddx1 knockout mouse

Knockout mouse models have been used to study the function of some of the DEAD box protein genes. Two mouse lines carrying a germ-line *Ddx1* mutation were previously generated in the lab and knockout of the *Ddx1* gene was found to be embryonic lethal. In Chapter 4, I identify the stage at which *Ddx1* knockout is embryonic lethality and address the mechanism underlying embryonic lethality. While carrying out these studies, I found that some of the wild-type mice were also dying

during development. The second component of Chapter 4 was to study this phenomenon and its non-Mendelian inheritance pattern.

1.4.3 Chapter 5: Subcellular localization of DDX1

DDX1 was previously shown to be highly expressed in proliferating cells and cells of neuroectodermal origin. DDX1 is primarily located in the nucleus of most cell lines; however, in *DDX1*-amplified RB and NB cell lines, DDX1 is primarily found in the cytoplasm. Furthermore, localization of DDX1 to the cytoplasm in breast cancer tumours is associated with a poor prognosis. In Chapter 5, I study the localization of DDX1 in a number of mouse tissues. The second component of Chapter 5 was to study the localization of DDX1 in developing embryos with a focus on pre-implantation development as the *Ddx1*$^{-/-}$ embryos were found to stall at the 2- or 4-cell stage. By studying the subcellular localization of DDX1 during early development, I found that DDX1 forms large aggregates in the cytoplasm that are highly dynamic in nature. While these DDX1 aggregates were dependent on the presence of RNA, I was not able to identify any proteins previously associated with RNA granules in the DDX1 aggregates.

Chapter 2

Materials and Methods

2.1 Molecular Techniques

2.1.1 PCR amplification and DNA insert purification

PCR amplification from plasmid DNA was carried out using a standard 30 cycle protocol with annealing temperatures based on the T_m of the oligonucleotide pairs (Table 2.1). The extension time was for 30 seconds to 1 minute depending on the size of the amplified product with 1000 bp per minute used as a guideline. Following PCR amplification and restriction enzyme digestion, DNA inserts were isolated by polyacrylamide gel electrophoresis (PAGE) using 4-8% gels, depending on fragment sizes, in 1X Tris-borate-EDTA (TBE) buffer. The DNA was visualized using ethidium bromide staining and the DNA fragments of the correct sizes were cut from the gel, placed in dialysis tubing with a molecular weight cut off of 12 kDa and subjected to electroelution in 1X TBE buffer at 80 V for a minimum of 3 hours. TE buffer was used to rinse the bag and bring the volume to 1.5 ml. The DNA was phenol/chlorophorm extracted and precipitated with 1/10 volume 5 M NaCl and 2 volumes 100% ethanol. After overnight precipitation at -20˚C, the DNA was pelletted, dried and resuspended in double distilled (dd) H_2O. Purified DNA was stored at -20˚C.

Table 2.1. Primer pairs for generation of DDX1 recombinant protein.		
aa 1	Sense	5′-AAAGGATCCGACGGGGTGAAGATGGCG
aa 620	Antisense	5′-TTGTACTCGAGTCATGCCACCAGGGAAATTGC
aa 633	Antisense	5′-CAGTACTCGAGTCAACGGCTGCTACATACATG
aa 645	Antisense	5′-CTGTTCTCGAGTCATTCCTTGAGTCTTGTGTTATA
aa 696	Antisense	5′-GTGTACTCGAGTCAAGCAGCCCTTTTCTGACC
aa 709	Antisense	5′-CAGTACTCGAGTCACAAAATATCCACATGGCCTT
aa 720	Antisense	5′-CTGTACTCGAGTCACTTTTCAAGGGCAGCCAA
Fly	Sense	5′-ATAGGATCCATGACTGCATTCGAAGAG
	Antisense	5′-TGATCTCGAGTTAGACCTTAAGACGTTTC
Chicken	Sense	5′-TGAGGATCCATGGCGGCGTTCTCGGAA
	Antisense	5′-TGTACTCGAGGCACAATCAGAATGTTCT
Mouse	Sense	5′-TGCGGATCCAAGATGGCGGCCTTCTCC
	Antisense	5′-GTTGCTCGAGTACATGACGTCAGAAGGT

2.1.2 DNA ligation

Each ligation was carried out using insert DNA and vector DNA at a molar ratio of 2:1, with total DNA concentration in the reaction in the range of 20 μg/ml. The reaction was performed in 20 μl volume containing 1X ligase buffer [50 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 20 mM dithiothreitol (DTT)], 0.1 μg/ul BSA, 1 mM ATP and 1 U T_4 DNA ligase. Reactions were incubated at 16°C overnight.

2.1.3 DNA sequencing

Automated sequencing was carried out according to the manufacturer's directions (Applied Biosystems) except that the reactions were halved. Approximately 200 ng of double-stranded DNA template, 4 μl 5X Big Dye Terminator Buffer v3.1 (Applied Biosystems), 0.5 μl Big Dye Terminator, 1.5 μM primer, and ddH_2O to a final volume of 10 μl were amplified in a thermal cycler (Perkin Elmer) under the following conditions: 25 cycles at 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes followed by a final extension for 7 minutes at 4°C. Amplified products were precipitated at room temperature for 15 minutes using 25 μl 100% ethanol and 1 μl 3 M NaoAc pH 5.2. The DNA was spun down and washed twice in 70% ethanol and dried. The pellets were resuspended in 12.5 μl formamide, boiled for 5 minutes and chilled on ice for 5 minutes. The samples were subsequently loaded into an automated sequencer (ABI PRISM 310 Genetic Analyser).

2.1.4 Recombinant protein expression in bacteria

Overnight cultures of 50 ml LB + 50 μg/ml ampicillin were inoculated with a colony of BL21 bacteria containing the plasmid of interest and grown at 37˚C. The following morning, 500 ml fresh LB was inoculated with 5 ml of the overnight culture and grown at 37˚C until the OD_{600} reached 0.6-1.0 (approximately 3.5 hours). Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a concentration of 0.2 mM. The culture was allowed to grow for an additional 4 hours at room temperature. The culture was then pelleted at 8000 x g for 12 minutes and frozen at -80˚C. The pellets were resuspended in 30 ml lysis buffer [25 mM Tris-HCl pH 7.5, 250 mM NaCl, 2 mM phenylmethanesulfonylfluoride (PMSF), 0.5X Complete protease inhibitors (Roche), 2 mM DTT] on ice. At this point the resuspended bacteria were lysed with either (1) the French press or (2) sonication.

1) The 30 ml French press cell was pre-chilled to 4˚C and the suspension was subjected to 16,000 PSI to lyse the bacteria.
2) Bacteria were incubated for 5 minutes in 20 mg/ml lysozyme. Bacteria were transferred to a 50 ml plastic beaker and placed on ice where the sonicator probe was placed approximately 1.5 cm from the bottom. The cells were subjected to 4-5 cycles of 1 minute on/ 2 minutes off sonication at 80% power.

The lysates were then incubated in Triton X-100 to a final concentration of 1% for 30 minutes on a rocker platform at 4˚C. During this time, a 1 ml 50% suspension of Glutathione Sepharose (Amersham Biosciences) was washed three times with 1.5 ml

lysis buffer minus the protease inhibitors. The lysates were centrifuged at 8000 x g for 30 minutes and the supernatants were collected and incubated with the washed beads for 4 hours on a rotator at 4˚C to allow binding of GST-fused recombinant protein to the Glutathione Sepharose. The mixture was then centrifuged at 500 x g. The supernatant was discarded and the beads were washed three times in lysis buffer minus protease inhibitors. The beads were finally suspended in 400 µl buffer to which 20 U thrombin (Sigma) was added to digest the recombinant protein from the beads. The digestion proceeded for 18-24 hours at 4˚C on a rotator. The following day, the beads were pelleted by centrifugation at 500 x g for 2 minutes and the supernatant removed. Two hundred µl of buffer were added to the beads and the beads were incubated for 5 minutes on a rotator. The suspension was centrifuged and the supernatant removed and collected separately from the first elution. A Bradford assay was performed to measure the protein concentrations of the eluate and additional washes were performed until the protein concentration was below 0.5 mg/ml. The eluates were aliquoted into 50 µl volumes and frozen at -80˚C. Purity was assessed by SDS PAGE followed by Coomassie Blue staining of the gel.

2.1.5 Automated crystallography

Through a collaboration with Dr. Bart Hazes (Department of Medical Microbiology and Immunology, University of Alberta), we had access to a Honeybee X32 sub-microlitre dispensing robot (Genomic Solutions). We utilized a variety of different screening suites offered by NeXtal (Qiagen) to screen for crystal formation

(Table 3.1). We used 96 well crystallography plates (Greiner Bio One) that contained 3 wells per reservoir. Two of the wells were filled with 0.1 μl protein, 0.1 μl mother liquor (the solution which contains the buffer, salt, and other precipitants), and 0.2 μl water. The third well was used as an internal control with the protein being replaced with 0.1 μl of the buffer used in the protein purification. Control chambers were used to determine whether the buffer itself could generate crystals. The plates were sealed and observed daily for the first five days followed by weekly observations for up to two months for the formation of crystals.

2.1.6 Hanging drop crystallography

Large scale droplets were set up in 24-well plates that were pretreated with vacuum grease. Each mother liquor was prepared and 500 μl were added to the well. Droplets were manually prepared by pipetting 1 μl protein onto a 20 x 20 mm glass coverslip and overlaying 1 μl mother liquor. A control drop was also generated which contained the protein purification buffer overlaid with 1 μl mother liquor. The coverslip was then inverted and placed over the filled reservoir and sealed by light pressure. The droplets were observed as described for the small scale plates. Crystals were harvested by capturing with a nylon loop attached to a mount and either immediately flash cooled in liquid nitrogen or first plunged into a 25% glycerol solution for cryoprotection.

2.2 Biochemical Techniques

2.2.1 End-Labeling RNA

Synthesized RNAs were purchased from Thermo Scientific (Table 2.2) and deprotected by adding 400 μl supplied deprotection buffer to the lyophilized RNA. The resuspended RNA was divided into 50 μl aliquot and incubated at 60˚C for 2 hours. The RNA was vacuum dried and each aliquot resuspended in water to a final concentration of 20 μM. The RNA was stored at -80˚C. Two μl of the RNA were added to 3 μl 10X polynucleotide kinase buffer (New England Biolabs), 2 μl T_4 polynucleotide kinase, 2 μl [γ-^{32}P]ATP (3,000 Ci/mmol, Perkin Elmer) and 21 μl ddH_2O. The reaction was incubated at 37˚C for 1.5 hours then precipitated overnight with the addition of 5 μl 3M NaoAc and 150 μl 100% EtOH. The precipitated RNA was spun at maximum speed (16,000 g) using a benchtop centrifuge for 15 minutes at 4˚C. The supernatant was removed by pipetting and the pellet was washed with 500 μl 70% EtOH, spun at room temperature at maximum speed in a benchtop centrifuge. The supernatant was removed by pipetting and the pellet was vacuum dried. The pellet was resuspended in 20 μl ddH_2O to a final concentration of 2 pmol/μl and stored at -20˚C. A working dilution of 1/30 was prepared and each reaction was carried out using 1 μl of RNA.

2.2.2 In Vitro Transcription

DNA templates (annealing pairs) were ordered from IDT and resuspended in 100 μl ddH_2O. Concentrations were measured by UV spectrophotometry (OD_{260}). The following oligonucleotides were used with the underlined portions representing the

Table 2.2. Synthesized RNA substrates used for end-labelled degradation assays.	
R29	5′-GAUCCUCUAGAGUCGACCUGCAGGCAUGC
AR29	5′-AAAAAAAAAAAAAAAAAAAAAAAAAAAAA
CR29	5′-CCCCCCCCCCCCCCCCCCCCCCCCCCCCC

transcribed sequence and the bolded T indicating the site of the labelled uridine: 2929 sense and antisense, 5′-ATCTCATAATACGACTCACTATAGGGGAAAAGAAAAGAAAAGAA AAGAAAAGAA**T** and 5′-ATTCTTTTCTTTTCTTTTCTTTTCTTTTCCCCTATAGTGAGTCGTATT ATGAGAT; 2905 sense and antisense, 5′-ATCTCATAATACGACCACTATAGGGGAAA**T**GAA AAGAAAAGAAAAGAAAAGAAA and 5′-TTTCTTTTCTTTTCTTTTCTTTTCATTTCCCCTATAGT GAGTCGTATTATGAGAT; 2915 sense and antisense, 5′-ATCTCATAATACGACTCACTATAG GGGAAAAGAAAAGAAA**T**GAAAAGAAAAGAAA and 5′-TTTCTTTTCTTTTCATTTCTTTTCTTT TCCCCTATAGTGAGTCGTATTATGAGAT. To generate each template, 33 μg of each oligonucleotide pair (sense and antisense) were added to 1X TE buffer to a final concentration of 20 pmol/μl duplex DNA. Oligonucleotide pairs were boiled for 5 minutes then allowed to slowly cool to room temperature. Duplex DNA (1.5 μl) was added to 5X *in vitro* transcription buffer (Invitrogen) consisting of 2.5 μl 10 mM ATP/GTP/CTP, 2.5 μl 1 mM UTP, 0.5 μl 150 mM DTT, 8 μl [α-^{32}P]UTP (3000 Ci/mmol, Perkin Elmer), 5 μl Ribolock (Thermo Scientific), 2 μl T7 RNA polymerase (Invitrogen) in a 40 μl reaction volume. The reaction was incubated at 37˚C for 3 hours followed by precipitation overnight with the addition of 11 μl 3 M NaoAc and 300 μl 100% EtOH. The DNA was spun down for 15 minutes at 4˚C and the supernatant was removed by pipetting. The pellet was washed with 500 μl 70% EtOH and spun for 5 minutes at room temperature. The supernatant was again removed by pipetting and the pellet was dried under vacuum and resuspended in 30 μl water. A working dilution of 1/30 was prepared and the labelled substrates were stored at -20˚C.

2.2.3 RNA degradation assay

Fifty fmol radiolabeled RNA were added to 0.6 μg (7 pmol) recombinant DDX1 protein in 20 mM Tris-HCl pH 7.5, 70 mM KCl, 2 mM MgoAc and 1.5 mM DTT in a final volume of 20 μl. Where indicated, ATP, ATP-γ-S, ADP, ADP-β-S, GTP, CTP or UTP (Sigma) was added to a final concentration of 1 mM. For reactions without Mg^{2+}, MgoAc was replaced with 10 mM EDTA pH 7.5. Reactions were incubated at 37˚C for 20 minutes unless otherwise indicated. The reactions were quenched on ice with the addition of 5 μl loading buffer (50 mM EDTA pH 8.0 and 40% glycerol). The total reaction was electrophoresed through a 20 cm 15% native polyacrylamide gel (29:1) in 1X TBE buffer for 2 hours at 80 V. Alternatively, half the reaction volume was electrophoresed through a 40 cm 15% native polyacrylamide gels. Gels were dried and exposed to X-ray film (124).

2.3 Mouse Manipulations

2.3.1 Collecting post-implantation embryos

Post-implantation embryos (E5.5 – E12.5) were collected by dissecting the uterus from pregnant females and placing it in a 35 mm dish containing PBS. Individual decidua were separated by cutting the uterus between each decidua and transferring to a clean dish of PBS. The washed decidua were kept in PBS on ice until dissected. The embryos were removed from the decidua following the method outlined in Manipulating the Mouse Embryo (231). Following dissection, embryos were washed in PBS. If the embryos were being genotyped they were processed as

described in Section 2.3.4. If the embryos were being sectioned and stained, they were processed as described in Section 2.3.5.

2.3.2 Collecting pre-implantation embryos

Naturally mated mice were checked for the presence of a vaginal plug indicating mating was successful. The plug date indicated the developmental stage of embryonic day (E) 0.5. Plugged mice were euthanized by CO_2 asphyxiation. The ovaries, oviducts and uterus were removed and placed in flush media [either warmed DMEM media supplemented with 5 mM HEPES pH 7.5 or warmed M2 media (Sigma)]. If embryos were being cultured, then drops of 25 μl KSOM (Millipore) or M16 media (Sigma) were prepared at least 30 minutes prior to collecting embryos. The medium was equilibrated by overlaying with mineral oil and incubating it at 37˚C in a 5% CO_2 incubator. The ovaries were cut from the oviducts and the oviduct was separated from the uterine horn by cutting the oviduct just before the uterotubal junction. To collect E0.5 to E2.5 embryos, the oviducts were flushed with flush medium. The oviducts were transferred to a fresh 35 mm dish containing flush media. Embryos were flushed from the oviducts from both ends by inserting a blunted 28G needle first into the cut end of the oviduct and flushing 0.25 ml media in pulses followed by flushing from the infundibulum with an additional 0.25 ml media. To collect E2.5 and E3.5 embryos, both the oviducts and uterine horns were flushed. An 18G needle was inserted through the vaginal end of the uterus into one horn and 0.5 ml media was pulsed to flush one side. This procedure was repeated to flush the other horn. E0.5 single cell embryos were

additionally treated with hyaluronidase (Sigma) for 30 seconds to 1 minute at 300 μg/ml to remove the cumulus cells.

2.3.3 Culturing pre-implantation embryos

At least 30 minutes before collecting embryos, KSOM media or M16 media was prepared and 25 μl drops were placed and covered with embryo tested mineral oil (Sigma) in a 24 well dish. Three larger (100 μl) drops of media covered with mineral oil were also prepared for washing the embryos following flushing. The dish was placed in a 5% CO_2 incubator to equilibrate the pH. After flushing the embryos into DMEM, embryos were rinsed 3 times in KSOM before they were transferred to the 25 μl droplets the embryos were then cultured at 37°C in 5% CO_2 (231).

2.3.4 Culturing Blastocysts

High glucose DMEM media with 15% FCS and 0.1 mM β– mercaptoethanol (β-ME) was supplemented with 1X sodium pyruvate (Gibco), 1X non-essential amino acids (Gibco) and penicillin and streptomycin. The wells of a 24 well dish were coated with 0.1% gelatin in PBS. For staining purposes, coverslips were washed with ethanol and allowed to dry before coating with gelatin. Collected embryos were washed once in supplemented DMEM media then added to gelatin-coated wells containing supplemented DMEM.

2.3.5 Collecting oocytes

To obtain germinal vesical (GV) stage oocytes, ovaries were dissected from adult female mice. Ovaries were placed in a 35 mm tissue culture dish containing prewarmed (37˚C) DMEM supplemented with 5 mM HEPES pH 7.5. The ovaries were poked extensively with a fine needle under a dissecting microscope to release oocytes. The oocytes were transferred to KSOM media to wash using a mouth-operated pipette connected to a glass capillary drawn to an internal diameter of ~350 μm. The cumulus cells were removed by gently pipetting the oocytes up and down the capillary. If oocytes were to be stained they were transferred to PBS and processed as described for preimplantation embryos (see below). Oocytes were allowed to mature by culturing in KSOM media under oil for 24-48 hours at 37˚C in 5% CO_2.

2.3.6 Preparation of mouse genomic DNA isolated from tail clips or embryos

Mouse tail clips were incubated with 500 μl TEN buffer (50 mM Tris-HCL pH 7.5, 100 mM EDTA, 100 mM NaCl) and 200 μg proteinase K (PK) for a minimum of 3 hours at 58˚C. After 3 hours, an additional 200 μg PK was added and the mixture was incubated overnight. Whole embryos (E6.5 to E9.5) were incubated with 100 μl TEN buffer and 40 μg PK overnight. DNA from embryos at E10.5 or later stages was collected by digesting the tail using the conditions described for whole embryos. The following day, genomic DNA was purified by phenol and phenol/chloroform extractions and EtOH precipitation. The purified DNA was air dried and resuspended in 100 μl TE buffer. The concentration of DNA was measured by UV spectrophotometry

(OD_{260}). The DNA was stored at 4˚C. E3.5 embryos and cultured blastocysts were collected in 20 μl PCR buffer supplemented with 40 μg/ml PK. The embryos were digested at 55˚C for 1 hour followed by 10 minutes at 90˚C to inactivate the PK. Embryos prior to E3.5 were collected as described in Section 2.3.2, washed 2X in PBS and 1X in water then transferred to individual PCR tubes in minimal volumes (approximately 1-2 μl). The embryos were processed using a modified Sigma REDExtract-N-Amp protocol. Briefly, the embryos were digested in 1.1 μl tissue preparation buffer and 4.4 μl extraction buffer for 30 minutes at 56˚C followed by heat inactivation for 5 minutes at 95˚C. A total volume of 4.4 μl neutralization buffer was then added and the DNA stored at 4˚C.

2.3.7 Analysis of Mouse Genomic DNA

Genomic DNA was initially analyzed by Southern blotting using DNA that had been digested with either *EcoR*I (gene-trap reporter), *BamH*I (wild-type *Ddx1*), or *Xho*I (conditional knockout) and separated in a 1% agarose gel. The DNA was transferred to nitrocellulose membranes and incubated with ^{32}P-labelled probes specific to *β-gal* or *Ddx1* (exons 10-17). Membranes were exposed to X-ray film. Routine genotyping was performed using multiplex PCR on DNA isolated from ear punches. Ear punch samples were processed with the E.Z.N.A Tissue DNA Kit (Omega) according to the manufacturer's instructions and stored at 4˚C. Ear punch and post-implantation embryo genotypes were analysed by PCR in a 20 μl reaction volume containing 1 μl DNA template, 2 μl 10X PCR buffer (GE Healthcare), 0.4 μM of each primer (RGo60:

5′-CTGGGGTTCGTGTCCTACAA, RGo63: 5′-ATTAGGAACTGGGCATGTATC, and RGo65: 5′-AGCACTAGTAAGTACCTACAC), 250 μM dNTP mix and 0.2 μl Taq polymerase. The reaction was PCR-amplified under the following conditions: initial heating to 94˚C for 5 minutes followed by 35 cycles at 94˚C for 1 minute, 60˚C for 1 minute and 72˚C for 1 minute followed by a final extension at 72˚C for 10 minutes and a hold at 4˚C.

Blastocysts were analysed by nested PCR. For the first round, we used 1 μl DNA template, 2 μl 10X PCR buffer, 0.8 μM of the following primers: RGo60 and either RGo62: 5′-GATGGAGACAGTCCTGGTT or RGo66: 5′-CCAAGCTCCACTATTATCCC), 250 μM dNTP mix and 0.2 μl Taq polymerase using the same amplification protocol described above. For the second round, we used 1 μl from the first round reaction, 2 μl 10X PCR buffer, 0.4 μM primers (RGo60 and either RGo63 or RGo65), 250 μM dNTP mix and 0.2 μl Taq polymerase using the same amplification protocol described above. DNA from stages E2.5 and earlier was amplified using a modified Sigma REDExtract-N-Amp protocol. For the first round of PCR amplification, the entire DNA preparation (10 μl) was added to 10 μl REDExtract-N-Amp PCR reaction mix (Sigma), and 0.8 μM of each primer RGo60, RGo63 and RGo65. The amplification conditions were as stated above except that 40 cycles were used instead of 35 cycles. In the second round of amplification, we used 5 μl of the previous reaction, 2 μl 10X PCR buffer, 0.4 μM of each primer (RGo60, RGo63 and RGo65), 250 μM dNTP mix and 0.2 μl Taq polymerase. PCR amplification conditions were as described for the first round of amplification. The reactions were electrophoresed in a 1% agarose gel and the DNA visualized using ethidium bromide.

2.3.8 RNase treatment of embryos

Embryos were collected and transferred to warm PBS in a 24 well dish. The embryos were then transferred to PBS + 0.1% TX-100 (PBST) for 10 minutes at room temperature then incubated in PBS containing 150 μg/ml RNase A for 30 minutes at 37˚C. Embryos were then washed in PBS twice before immunofluorescence staining. Method adapted from (124).

2.3.9 Preparation and analysis of protein lysates from mouse brain

Protein was isolated from P0 to P3 brain tissue that had been previously flash frozen and stored at -80˚C. All steps were performed on ice. Two hundred μl chilled lysis buffer (PBS containing 1% TX-100, 0.1% SDS, 1X Complete (Roche), 1 mM PMSF, and 1 mM DTT) were added to each sample. The samples were homogenized with a pestle and spun at 14,000 g for 10 minutes at 4˚C. The supernatants were checked for protein concentration by Bradford assay and the lysates frozen at -80˚C.

Cell lysates (50 μg per lane) were electrophoresed in an 8% SDS-polyacrylamide gel. The proteins were transferred to a nitrocellulose membrane. Membranes were blocked with 10% milk in TBS+T (0.01% Tween-20) for 1 hour, then sequentially immunostained with anti-DDX1 (batch 2910; 1:5,000 dilution) and anti-actin (Sigma; 1:100,000 dilution) or with both anti-DDX1 (2910) and anti-actin (Sigma) together in 5% milk in TBST at 4˚C overnight. Membranes were washed three times in TBST. Signal detection was using anti-rabbit (for DDX1) and anti-mouse (for actin) secondary antibodies conjugated to HRP (Molecular Probes; 1:50,000 dilution) in 5% milk in TBST

for 4 hours, followed by incubation with the ECL reagent (GE) and exposure to X-ray film.

2.3.10 Analysis of RNA isolates from pre-implantation embryos

β-gal expression was used to determine when *DDX1* is first expressed in pre-implantation embryos. Embryos at different pre-implantation stages were obtained from *Ddx1*$^{+/-}$ backcrosses where the female was wild-type and the male was heterozygous and pooled. Pools of oocytes were collected from *Ddx1*$^{+/-}$ female mice by puncturing the follicles and the cumulus cells were removed by pipetting. Embryos and oocytes were washed 2X in PBS followed by 1X in water. They were then placed in a PCR tube containing 0.5 μl RiboLock (Thermo Scientific), flash frozen in liquid nitrogen and stored at -80˚C. PCR amplification using gene-specific primers was performed using the OneStep RT-PCR kit (Qiagen). Primers specific to *β-gal* (sense, 5′-CCTGTCCGGTGCCCTGAATG, antisense, 5′-GAAGAACTCGTCAAGAAGGCG) were used in both first and second round amplifications. The first cycle was performed according to the manufacturer's protocol with the following amplification program: 30 minutes at 50˚C; 15 minutes at 95˚C; then 40 cycles of 1 minute at 94˚C, 1 minute at 50˚C, 1 minute at 72˚C; with a final 10 minute extension at 72˚C. Samples were then kept at 4˚C. The second round amplification was performed with 1 μl of sample from the first round, 2 μl PCR buffer, 0.4 μM of each primer (same as previous reaction), 250 μM dNTP mix and 0.2 μl Taq polymerase. The following conditions were used for the PCR amplification: an initial heating to 94˚C for 5 minutes followed by 35 cycles at 94˚C for

1 minute, 55˚C for 30 seconds and 72˚C for 1 minute followed by a final extension at 72˚C for 10 minutes and a hold at 4˚C. The reactions were electrophoresed in a 1% agarose gel to separate the amplified DNA.

2.3.11 Analysis of RNA isolated from mouse brain

RNA was isolated from P0 to P3 mouse brains by adding 1 ml Trizol (Life Technologies) and homogenizing the tissue with a pestle in a 1.5 ml tube. RNA was extracted as per the manufacturer's protocol. OD_{260} was used to measure RNA concentration. The purified RNA was stored at -80˚C. Complementary DNA (cDNA) was generated using Superscript II (Life Technologies) following the manufacturer's protocol using either oligo$(dT)_{12-18}$ or random hexamer primers and 5 μg RNA. The cDNA in a total volume of 50 μl was stored at -20˚C. Semi-quantitative (sq) RT-PCR was performed in a 20 μl reaction containing 1 μl cDNA, 2 μl 10X PCR buffer (GE Healthcare), 0.4 μM of each primer pair (5′ *DDX1*: sense, 5′-GGTGCATTTAGTATTCCCGTT and antisense, 5′-ACGTTGTTCAAGGTTTGCTC; 3′ *DDX1*: sense, 5′-AGAATTATGTGCACCG GATC and antisense, 5′-GCACCAGAGGGTTAGAGT; *β-gal*: sense, 5′-CCTGTCCGGTGCCCT GAATG and antisense, 5′-GAAGAACTCGTCAAGAAGGCG; *β-gal-DDX1* fusion: sense, 5′-CTGAAGAGCTTGGCGGCGAAT and antisense, 5′-TTTGGATCCATGTACATCATCAGTTCT AAT; *GAPDH*: sense, 5′-ACGGCAAATTCAACGGCAC and antisense, 5′-GAGAGCAATGCCA GCCCC), 250 μM dNTP mix and 0.2 μl Taq polymerase. The reaction was amplified using the following conditions: an initial heating to 94˚C for 5 minutes followed by 25 cycles (*GAPDH*), 29 cycles (*DDX1* or *β-gal*, or *β-gal-DDX1 fusion*) at 94˚C for 1 minute,

55˚C for 30 seconds and 72˚C for 1 minute followed by a final extension for 10 minutes at 72˚C and a hold at 4˚C. The reactions were electrophoresed in a 1% agarose gel to separate the amplified DNA.

2.3.12 Statistical analysis

Ratios of wild-type to heterozygote progeny were determined using Fischer exact tests. Expected values were determined based on Mendelian ratios of 1:2 for heterozygote intercrosses and 1:1 for backcrosses.

2.3.13 DNA methylation analysis

A combined bisulfite restriction enzyme analysis was performed in collaboration with Dr. Igor Kovalchuk (Department of Biology, University of Lethbridge) using a previously published protocol (232). In brief, genomic DNA prepared from *Ddx1*$^{+/+}$, *Ddx1*$^{+/-}$ and *Ddx1*$^{S/-}$ mice (in triplicate) was treated with bisulfite, amplified using primers sense, 5'-AATAGGGTTTGGTTATTAGGGAAAGTA and antisense, 5'-CCAAACAAAACAACATCATCTTTAC, and digested with either *HpyCH4*IV (B) or *Taq*I. The digested DNA was electrophoresed in a 1% agarose gel and the ratio between the digested DNA and the total amplified DNA yielded the approximate percent methylation. Bisulfite sequencing was performed on genomic DNA prepared from ear punches. One μg DNA was subjected to bisulfite conversion using the EpiTect Bisulfite kit (Qiagen) using the manufacturer's protocol. The promoter region was amplified using 1 μl template, 10X PCR buffer (GE), 0.4 μM of each primer (promoter:

sense, 5'-AAGTTTATAGGTTTTGAGTGAATTATT, antisense, 5'-CCAAACAAAACAACATCAT CTTTAC) 250 μM dNTP mix and 1 μl Taq polymerase in a 100 μl volume. The gene body region was amplified using 1 μl template, 10X PCR buffer (GE), 0.4 μM of each primer (promoter: sense, 5'-GTTGTTAGAGGGATTGATATTTATGG, antisense, 5'-ATAACCAAACC CAT CCTAAAAAATT), 250 μM dNTP mix and 0.5 μl Taq polymerase in a 50 μl volume. The products obtained for the gene body region could be directly purified using the GeneJet PCR cleanup kit (Thermo Scientfic) and cloned as described below. The products obtained for the promoter region generated multiple bands and had to be purified by acrylamide gel purification (6%) as described in Section 2.1.1. Purified DNA was ligated to the pGEM-T Easy (Promega) vector user the manufacturer's protocol with overnight ligation at 16˚C. *E. coli* DH5α competent cells were transformed with the ligated products and colonies selected using blue/white colour selection. White colonies were expanded and DNA was purified by boiling lysis of the bacteria (233). Purified DNA was sequenced as described in Section 2.1.2 using the T7 sequencing primer (5'-TAATACGACTCACTATAGGG). DNA sequences were then subjected to analysis by Bisulfite Sequencing DNA Methylation Analysis (BISMA) using default parameters (234).

2.4 Microscopy Techniques

2.4.1 Antibodies and Markers

The antibodies used are listed in Table 2.3. We also used the following chemicals or reagents for staining embryos: acridine orange (Sigma) at 2 μM to stain

RNA; Mitotracker Orange CMTMRos (Molecular probes) at 150 nM to stain mitochondria; phalloidin-Alexa 547 (Molecular probes) at 1:300 to stain filamentous actin; and 4',6-diamidino-2-phenylindole (DAPI) at 1 μg/ml in mounting media to stain DNA.

2.4.2 Fixing and embedding tissues and embryos for sectioning

Dissected embryos or tissues were fixed in 10% formalin prior to embedding in either paraffin (1) or OCT (2).

(1) Following fixation, tissues were dehydrated through an alcohol series. The alcohol was removed by washing the embryos in xylene. The tissue was then added to melted paraffin to infiltrate the tissue and the tissue was mounted in blocks in preparation for microtome sectioning. Four μm sections were immunostained (IHC) as described in Section 2.4.3.

(2) Following fixation, the embryos were cryoprotected in solutions of 12%, 16% and 18% sucrose in PBS for 2 hours each. The tissues were then embedded in OCT and flash frozen in a dry ice/methanol bath. OCT embedded tissues were stored at -80˚C until cryosectioned. Seven μm sections were immunostained (IHC) as described in Section 2.4.3.

Table 2.3. Antibodies used for western blotting, immunohistochemistry, and indirect immunofluorescence.			
Antibody	Host	Dilution	Source
DDX1 2910 (aa 1-186 denatured)	Rabbit	WB: 1:5000 IHC/IF: 1:1000	In house
DDX1 2290 (aa 187-740 native)	Rabbit	IF: 1:800	In house
DDX1 2923 (aa 1-186 native)	Rabbit	IF: 1:800	In house
Calnexin	Goat	IF: 1:200	Santa Cruz
S6	Mouse	IF: 1:50	Santa Cruz
Golgin 97	Mouse	IF: 1:50	Molecular Probes
RanBPM (RanBP9)	Goat	IF: 1:100	Abcam
CstF64	Mouse	IF: 1:1000	Dr. James Manley
Rif1	Goat	IF: 1:100	Santa Cruz
SMN	Mouse	IF: 1:1000	Santa Cruz
Tubulin	Mouse	IF: 1:200	Dr. Michael Klymkowsky
DDX3	Mouse	IF: 1:200	Santa Cruz
ExoSC5	Mouse	IF: 1:100	Abcam
GW182	Mouse	IF: 1:50	Dr. Marvin Fritzler
β-gal	Mouse	WB: 1:500; IF: 1:100	Santa Cruz
Anti-Rabbit Alexa-488	Donkey	IF: 1:300	Molecular Probes
Anti-Mouse Alexa-555	Donkey	IF: 1:300	Molecular Probes
Anti-Goat Alexa-555	Donkey	IF: 1:300	Molecular Probes

2.4.3 Immunohistochemistry

Sectioned tissues from both OCT- and paraffin-embedded tissue were treated as follows. Tissues embedded in OCT were washed with water to remove the OCT then subjected to antigen retrieval. Slides were placed in 10 mM citrate buffer containing 0.05% Tween-20 at pH 6.0 heated to 100˚C. Tissues embedded in paraffin were de-waxed in xylene (3X for 10 minutes each). They were then hydrated through a series of ethanol from 100% ethanol to ddH_2O and TBS. The citrate antigen retrieval buffer described above was heated to 100˚C in a pressure cooker. The slides were then immersed in the antigen retrieval buffer and microwaved at 750W for 10 minutes (7 minutes for neural tissue). The pressure cooker was cooled with slow running water and the slides were washed with TBST (0.05% Tween 20) for 5 minutes. From this point all tissues were treated the same. Slides were dried around the tissue and a hydrophobic barrier pen was used to allow retention of reagents on the tissue. The tissue was blocked for 30 minutes to 1 hour in 0.5% fish gelatin in TBST (0.1% Tween 20). The primary antibody was diluted in Dako Antibody Diluent for overnight incubation at 4˚C. The tissue was then washed 1X in TBS and 3X in TBS containing 0.05% Tween 20 for 10 minutes each. The tissue was blocked with 3% H_2O_2 in TBS for 15 minutes to remove background peroxidase activity. The slides were then washed 2X in TBST. Dakocytomation Envision+ System Labelled Polymer HRP secondary antibody was added for 1-2 hours. The slides were washed 1X in TBS and 3X in TBST for 10 minutes each. DAB chromagen in TBS was added to the slide and DAKO Liquid DAB+ Substrate Chromagen System was added for 1-5 minutes. Signal intensity was

measured by monitoring under a microscope. DAB was rinsed with tap water. The DAB signal was darkened by incubating the slides in a 1% copper (II) sulphate solution. Counterstain was applied by incubating the slide in hematoxylin for 1-5 minutes. The slides were washed with warm tap water for ~3 minutes. The slides were then incubated in saturated lithium carbonate for ~2 minutes (until the colour reached light blue). The slides were coverslipped with VectaMount AG (Vector Laboratories). Images were captured with an Axioskop 2 Plus microscope with Zeiss Plan-N eoFLUAR 10X/0.3, FLUAR 20X/0.75 and F FLUAR 40X/1.3 lenses.

2.4.4 Immunofluorescence labelling of embryos in suspension

Embryos in 24 well dishes were transferred from flush media to PBS. Embryos were then fixed in 4% paraformaldehyde (PFA) for 15 minutes, washed 3 times in PBST (5 minutes per wash) followed by permeabilization in PBS + 0.5% TX-100 for 10 minutes. The embryos were washed 3 times in PBST and incubated with primary antibody in PBST for a minimum of 1 hour at room temperature. Embryos were washed 3 times in PBST then incubated with secondary antibody in PBST for 1 hour at room temperature. During the incubations, the dish was wrapped with foil to prevent light exposure. After a final wash in PBST, the embryos were placed on a slide and mounted in Mowiol (Calbiochem) containing 1 mg/ml 4',6-diamidino-2-phenylindole (DAPI) to stain the DNA. The slide was coverslipped and the mounting medium allowed to polymerize overnight at room temperature in the dark. Method adapted from Rossant lab protocol: “Antibody staining early embryos”.

2.4.5 Imaris analysis

Embryos were imaged by confocal microscopy using a Zeiss Z1 laser scanning microscope (LSM). We used a Zeiss plan-apochromat 63X/1.4 NA lens to capture single sections. Z-stacks were taken in 0.35 μm intervals. Cytoplasmic DDX1 aggregates were analysed with Imaris software v.7.7.0 (Bitplane). All images were processed with a 3x3x1 median filter followed by surface analysis. Surfaces were not smoothed and an absolute threshold of 200 voxels (3D pixel) was used to define surfaces. This was followed by a minimum threshold of 200 voxels to define the minimum aggregate size. Six to ten embryos were analysed for each developmental stage and the volumes of the DDX1 aggregates were compared using a box and whisker plot (MedCalc v.12.4.0.0, MedCalc Software). Additional statistical analysis was performed by one-way ANOVA and pairwise comparisons were made using the Student-Newman-Keuls post-hoc test.

2.4.6 Differential Interference Contrast (DIC) live cell imaging

Twenty-four hours prior to embryo collection, 10 μl M16 medium was placed onto a glass bottom culture dish, covered with embryo tested mineral oil and placed in an incubator at 37˚C in a 5% CO_2 incubator. Embryos were prepared as described in Sections 2.3.2 and 2.3.3. The dish containing single embryos in droplets was placed in a heated humidified chamber supplemented with 5% CO_2. The lid of the culture dish was replaced with a glass topped lid. Images were collected with a CoolSNAP HQ (Roper Scientific) charged-coupled device (CCD) sensor attached to a Zeiss Axiovert

200M microscope. Embryos were imaged with a Zeiss plan-apochromat 10X/0.14 lens. Images were collected every 3 minutes for 1 second. Image processing was done using Metamorph v.7.8.3.0 (Molecular Devices).

Chapter 3

DDX1 structure and enzymatic activity

Like all DEAD box proteins, DDX1 has 12 conserved motifs involved in ATP binding, ATP hydrolysis and RNA unwinding (54). Previous work in the lab indicates that DDX1 can unwind RNA-RNA or RNA-DNA duplexes that are significantly longer than the duplexes unwound by other DEAD box proteins (29 base pairs as compared to 9-15 base pairs for most other DEAD box proteins), in an ADP-dependent manner (124). Furthermore, DDX1 has a ribonuclease activity that is distinct from that of ribonuclease A (RNase A) in that DDX1's ribonuclease activity is dependent on the presence of Mg^{+2}. Boiling destroys DDX1 ribonuclease activity in contrast to RNase A which remains active after boiling (124).

The structures of several DEAD box proteins have been determined using X-ray crystallography (27-29,33,39). From these structures it has been surmised that RNA helicases of the DEAD box protein family cause localized destabilization of the RNA-RNA duplexes which results in RNA-RNA unwinding. This is in contrast to DNA helicases which are processive enzymes. To further elucidate the mechanism of action of DEAD box proteins, structural analysis has been carried out in the presence of DEAD box interacting proteins to determine how co-factors can modulate the RNA binding or unwinding activity of DEAD box proteins. For example, Dbp5 (DDX19) is associated with mRNA export across the nuclear pore complex (235). The RNA remodeling activity of DDX19 was found to be greatly enhanced through interaction with Gle1 as well as the small molecule inositol hexakisphosphate (IP6) (53). Gle1 was found to interact with both helicase domains of DDX19 while IP6 bridged the N-terminal domain and Gle1. The addition of Gle1 and IP6 induced a conformation change in the structure of

DDX19 to a more open state which may promote RNA release. Biochemically, it was found that Gle1 and IP6 greatly increase the rate of ATP hydrolysis by DDX19. The interaction of Gle1 and IP6 with DDX19 facilitates the release of RNA which allows the protein to be recycled.

The primary amino acid sequence of DDX1 makes it an interesting candidate for structural determination as it has a number of features that differentiate it from other DEAD box proteins. The major differences between DDX1 and other DEAD box proteins include: (i) a very short (3 aa) N-terminal region upstream of the first DEAD box motif, (ii) a SPRY domain found within helicase domain I, and (iii) a non-conserved carboxy (C)-terminal domain downstream from the DEAD box motifs. Structural analysis may provide insight into how DDX1 unwinds longer duplexes compared to other DEAD box proteins, and may help to identify the location of the ribonuclease domain.

3.1 Generating constructs for recombinant DDX1 protein production

To determine the structure of DDX1, we prepared both wild-type and truncated DDX1 expression constructs. Wild-type DDX1 expression constructs were prepared using cDNAs encoding the entire open reading frames of human, mouse, chicken and *Drosophila* DDX1 (Figure 3.1). The 5'-and 3'-primers used for PCR amplification were designed to include *BamH*I and *Xho*I restriction enzyme sites, respectively, to facilitate directional cloning (Table 2.1). The open reading frame of DDX1 was placed in-frame

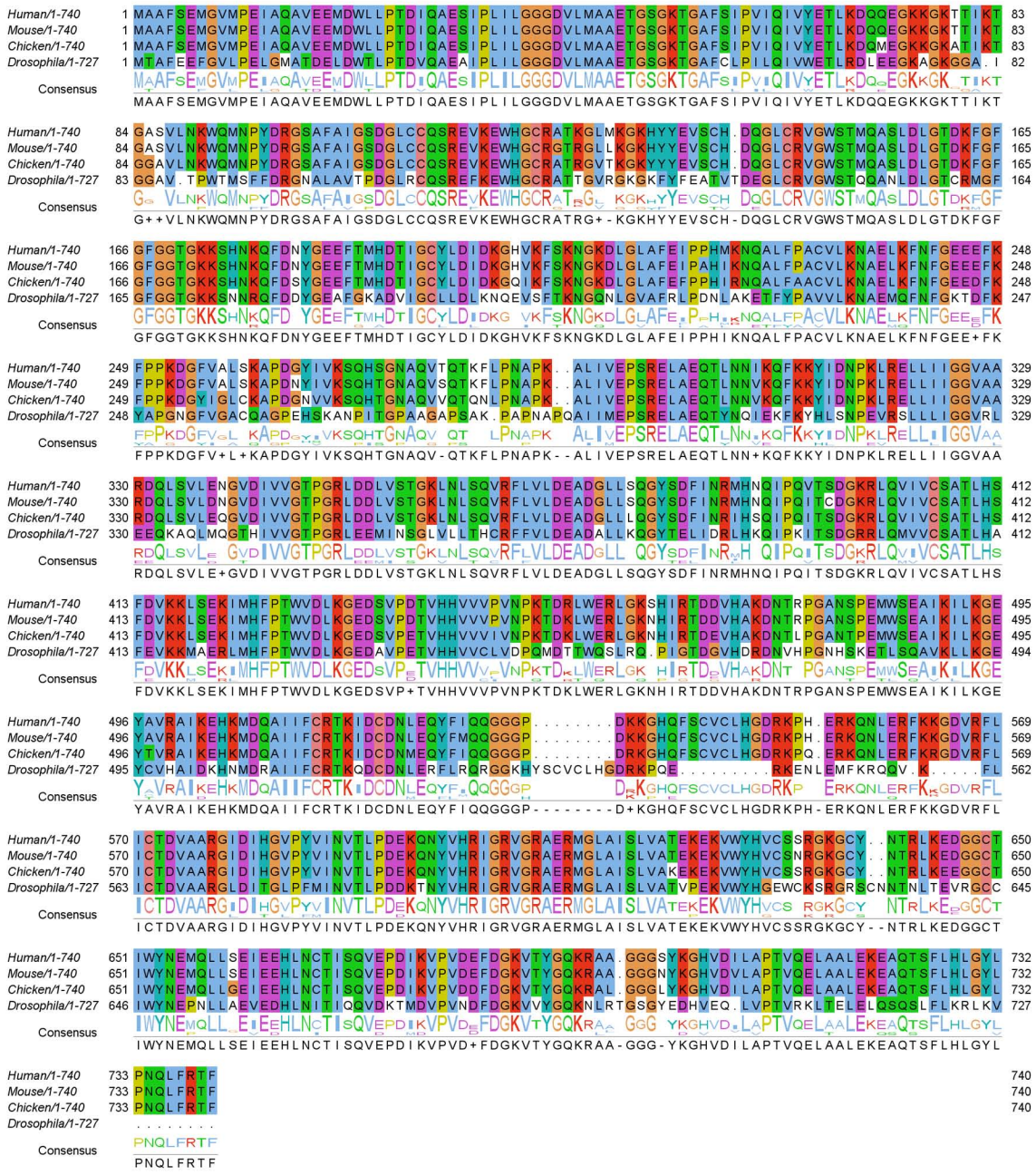


Figure 3.1 Alignment of human, chicken, mouse and Drosophila DDX1 proteins.

The amino acid sequences of human, chicken, mouse and Drosophila DDX1 orthologues were aligned using the MAFFT sequence alignment algorithm with its global alignment mode. Sequences were displayed and coloured using Jalview v2.8.0b1. The underlined sequences indicate the conserved DEAD box motifs.

with the upstream GST to generate GST-DDX1 fusion proteins. The amplified PCR products were digested with *BamH*I and *Xho*I, purified by polyacrylamide gel electrophoresis and phenol extraction, and ligated to the pGEX-4T3 vector. All constructs generated were tested by both restriction enzyme digestion and sequencing to confirm that the inserts were correct and that there were no mutations in the inserted DNA.

As DDX1 is a large protein (740 amino acids), so two sets of DDX1 truncated constructs were generated to improve the likelihood of generating crystals (Figure 3.1). The first set of truncated constructs was designed to split the N-terminal helicase domain and the C-terminal helicase domain. These mutants were predicted to be more likely to crystallize as they are much shorter in length than wild-type DDX1 and have less overall flexibility than the full-length protein. Solving the structure of just the N-terminal mutant would yield insight into how the SPRY domain interacts with the other DEAD box motifs as well as provide insight into how DEAD box motifs interact with each other. The C-terminal region of DDX1 downstream of motif VI shares no homology with any other proteins, domains, or motifs. Any structural information obtained with the C-terminal domain of DDX1 would thus provide insight into its function should structural homology with other proteins be identified. The drawback to solving the N-terminal and C-terminal structures independently is that we will be unable to determine how the domains interact with each other and how they compare to other solved DEAD box proteins.

The second set of truncated constructs was designed to shorten the full-length DDX1 protein at either the C-terminal end in order to reduce flexibility or the N-terminal end by deleting the SPRY domain. The C-terminal truncations were based on secondary structure prediction of DDX1 and involved amino acids located downstream of motif VI (Figure 3.2). The C-terminal region of DDX1 was predicted to have several helices (Figure 3.3) and truncation mutants were designed around these regions to reduce the amount of potential structural perturbations.

The DDX1 truncated constructs were generated by PCR amplification using pGEX-4T2-huDDX1 DNA as the template. DDX1 with deletion of the SPRY domain was generated by quickchange PCR whereby primers were designed to loop out the SPRY domain while amplifying the rest of the vector. Following PCR amplification, the parent vector was degraded by the addition of *Dpn*I. Some of the DDX1 orthologues and truncated DDX1 were subsequently subcloned into the pET vector as a means of generating recombinant proteins ($DDX1_{1-645}$, N-DDX1, and C-DDX1) for biochemical analysis.

3.2 Recombinant DDX1 protein production and purification

Recombinant GST-DDX1 proteins were purified using Glutathione Sepharose beads. Typical GST-DDX1 protein concentrations varied from 3 mg/ml to 8 mg/ml. As highly concentrated DDX1 is required in order to carry out X-ray crystallography, we attempted to concentrate the recombinant proteins using filter membranes with a 30 kDa (for >70 kDa recombinant proteins) or 10 kDa (for <70 kDa recombinant proteins)

Figure 3.2 Schematic representations of DDX1 truncation mutants.

The helicase domains and the conserved DEAD box motifs are shown to scale. The numbers listed represent the amino acid residues of interest. Two truncation mutants containing Helicase Domain 1 or Helicase Domain 2 were designated as N-terminal and C-terminal, respectively. $DDX1_{1\text{-}620}$ and $DDX1_{1\text{-}645}$ represent two different truncations at the C-terminus. In the SPRYless mutant, the SPRY domain consisting of amino acids 131-269 was deleted.

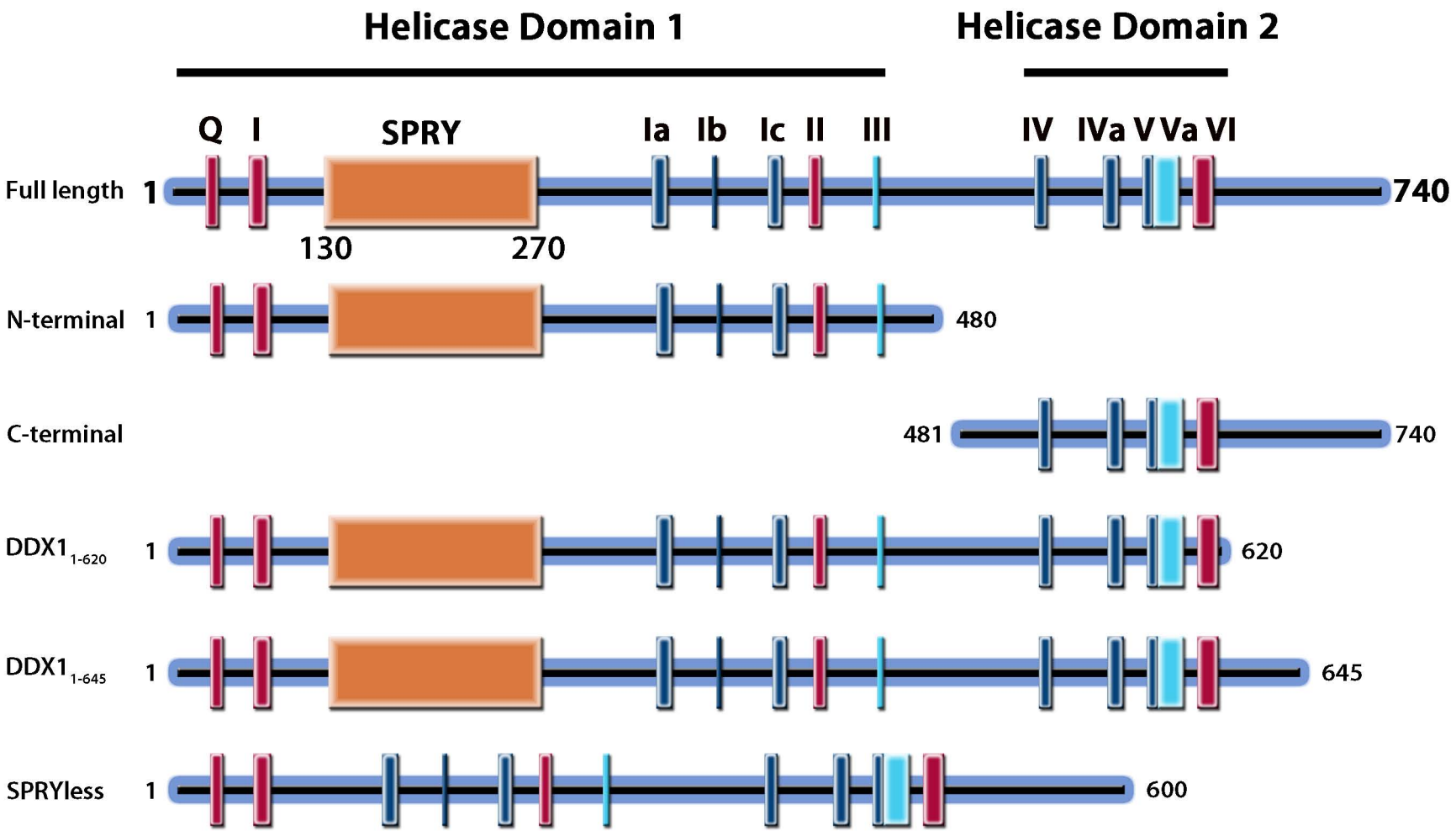
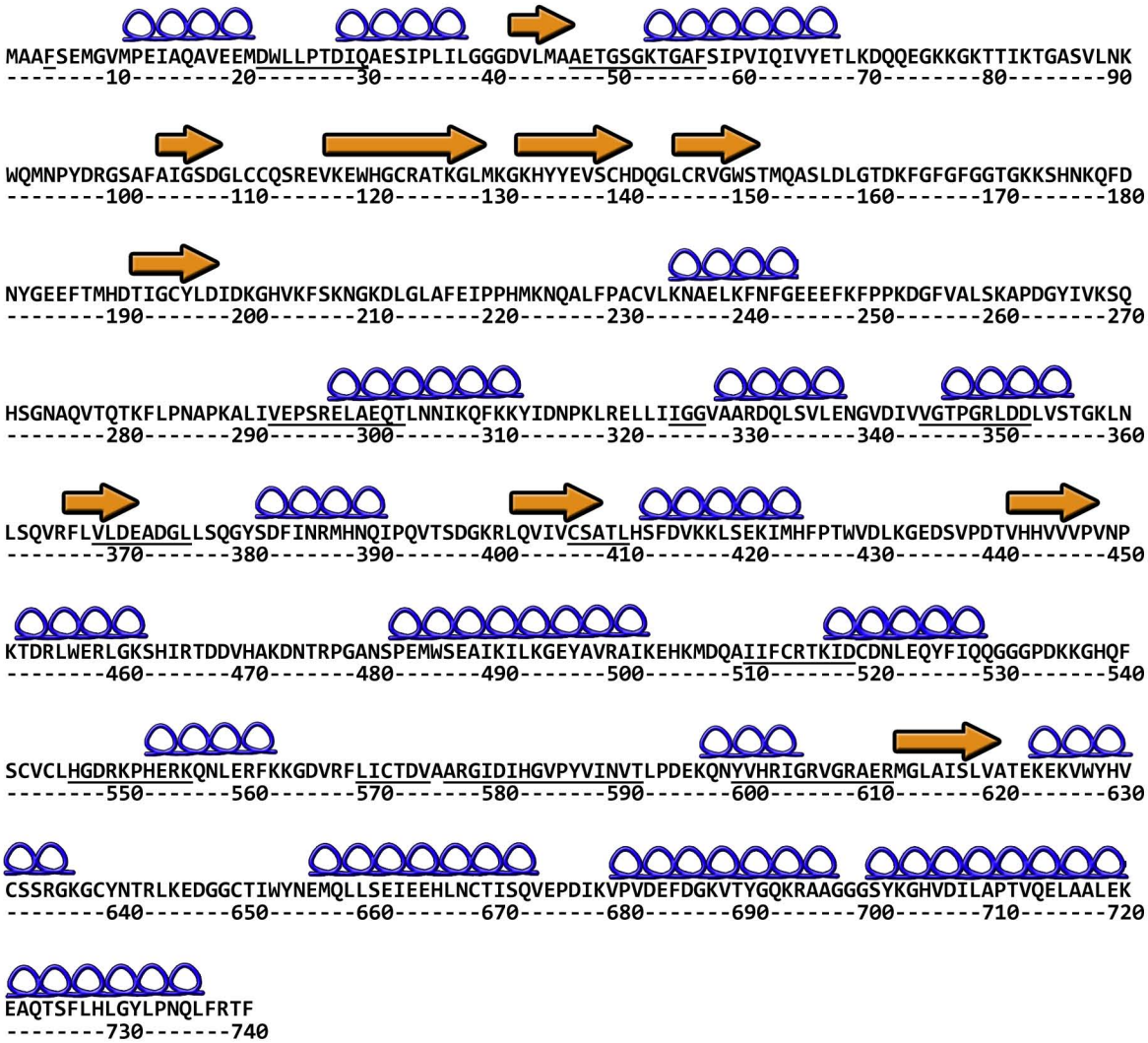

Helicase Domain 1
Helicase Domain 2
Q
I
SPRY
Ia
Ib
Ic
II
III
IV
IVa
V
Va
VI
Full length
1
740
130
270
N-terminal
1
480
C-terminal
481
740
DDX1 1-620
1
620
DDX1 1-645
1
645
SPRYless
1
600

Figure 3.3 Schematic representation of DDX1 secondary structure prediction.

The secondary structure of human DDX1 was predicted using Advanced Protein Secondary Structure Prediction Server (236). The predicted regions of α-helices or β-sheets were overlaid and represented with loops and arrows, respectively. Amino acids in conserved DEAD box motifs are underlined.


MAAFSEMGVMPEIAQAVEEMDWLLPTDIQAESIPLILGGGDVLMAAETGSGKTGAFSIPVIQIVYETLKDQQEGKKGKTTIKTGASVLNK
--------10--------20--------30--------40--------50--------60--------70--------80--------90
WQMNPYDRGSAFAIGSDGLCCQSREVKEWHGCRATKGLMKGKHYYEVSCHDQGLCRVGWSTMQASLDLGTDKFGFGFGGTGKKSHNKQFD
--------100-------110-------120-------130-------140-------150-------160-------170-------180
NYGEEFTMHDTIGCYLDIDKGHVKFSKNGKDLGLAFEIPPHMKNQALFPACVLKNAELKFNFGEEEFKFPPKDGFVALSKAPDGYIVKSQ
--------190-------200-------210-------220-------230-------240-------250-------260-------270
HSGNAQVTQTKFLPNAPKALIVEPSRELAEQTLNNIKQFKKYIDNPKLRELLIIGGVAARDQLSVLENGVDIVVGTPGRLDDLVSTGKLN
--------280-------290-------300-------310-------320-------330-------340-------350-------360
LSQVRFLVLDEADGLLSQGYSDFINRMHNQIPQVTSDGKRLQVIVCSATLHSFDVKKLSEKIMHFPTWVDLKGEDSVPDTVHHVVVPVNP
--------370-------380-------390-------400-------410-------420-------430-------440-------450
KTDRLWERLGKSHIRTDDVHAKDNTRPGANSPEMWSEAIKILKGEYAVRAIKEHKMDQAIIFCRTKIDCDNLEQYFIQQGGGPDKKGHQF
--------460-------470-------480-------490-------500-------510-------520-------530-------540
SCVCLHGDRKPHERKQNLERFKKGDVRFLICTDVAARGIDIHGVPYVINVTLPDEKQNYVHRIGRVGRAERMGLAISLVATEKEKVWYHV
--------550-------560-------570-------580-------590-------600-------610-------620-------630
CSSRGKGCYNTRLKEDGGCTIWYNEMQLLSEIEEHLNCTISQVEPDIKVPVDEFDGKVTYGQKRAAGGGSYKGHVDILAPTVQELAALEK
--------640-------650-------660-------670-------680-------690-------700-------710-------720
EAQTSFLHLGYLPNQLFRTF
--------730-------740

cutoff. Concentrating the DDX1 recombinant proteins was generally unsuccessful as they tended to adhere to the membrane except in high salt buffer (>500 mM NaCl) which would have interfered with the crystallography experiments. We were only able to concentrate N-terminal DDX1 to 9 mg/ml, with the concentration of the other recombinant GST-DDX1 proteins varying from 4 mg/ml to 7.5 mg/ml. The purity of the recombinant proteins was examined by SDS-PAGE separation followed by Coomassie Blue staining (Figure 3.4).

3.3 Growing and diffracting recombinant DDX1 protein crystals

To generate protein crystals, we screened a wide range of buffers, salts and precipitants against our recombinant proteins. Several conditions were found to generate crystals using small scale (0.1 μl) screening plates (Table 3.1). Some of the common features of these mother liquors were the presence of negatively charged ions including citrate and acetate. As DEAD box proteins are known to require Mg^{2+} for activity, we added 10 mM $MgCl_2$ directly to the recombinant protein just before screening. We found that some protein crystals also formed in the presence of Mg^{2+} that was present in the mother liquor.

Droplets that had fewer large crystals were selected for large scale experiments rather than those with smaller more numerous crystals because larger crystals tend to yield better results during data collection (Table 3.1). 24 well dishes were used for hanging drop vapor diffusion experiments to generate larger crystals. Drops of 1 μl protein and 1 μl mother liquor mixes were manually dispensed onto glass coverslips

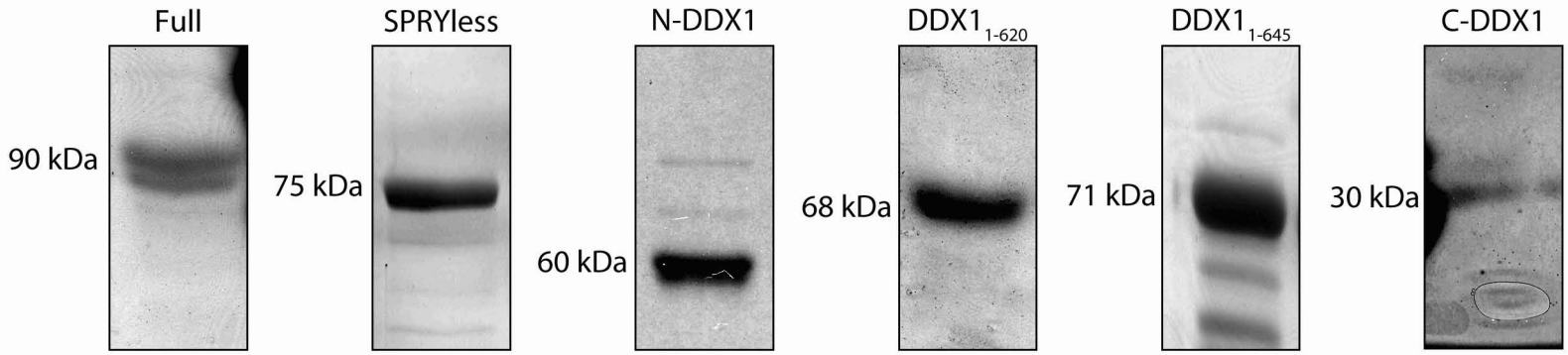


Figure 3.4 Commassie staining of purified recombinant DDX1 mutant proteins.

GST-fused DDX1 mutant proteins were expressed in BL21 (DE3) bacteria and purified with Glutathione Sepharose beads. The recombinant proteins were removed from the beads by cleaving the GST tag with thrombin. Four μg of each purified DDX1 mutant protein was loaded onto SDS-PAGE gels to examine purity. Gels were stained with Coomassie Blue overnight and destained for 2-4 hours.

inverted over 500 μl mother liquor and sealed with vacuum grease. Again there were control droplets placed on the same coverslip which consisted of 1 μl protein purification buffer and 1 μl protein mother liquor to ensure that the crystals formed were protein based. These hanging drops were observed using the same schedule described for small scale screening plates. Only a few of these large scale tests generated crystals and these were collected in a mounting loop and flash cooled in liquid nitrogen for storage. Those conditions with <50% PEG were subjected to a final cryopreservation step whereby the isolated crystals were plunged in a 25% glycerol solution prior to flash cooling in liquid nitrogen. Different forms of protein crystals were observed and some were imaged (Figure 3.5).

To collect diffraction data, the protein crystal was mounted in a nylon loop and placed in a stream of nitrogen to maintain a temperature of -160°C which preserves the crystal during data collection. We selected 5 crystals for preliminary tests based on their size, quality, and reproducible growth (Table 3.1). Preliminary tests to detect diffraction were performed with the detector set to 250 mm but all failed to generate any diffraction patterns. Further analysis of other potential crystals and crystallography conditions was not pursued.

3.4 DDX1 degrades single-strand RNA

DDX1 has previously been shown to degrade ssRNA *in vitro* (124). In comparison to RNAse A, different size products are obtained when ssRNA is incubated with DDX1, suggesting different mechanisms of RNA degradation for these two

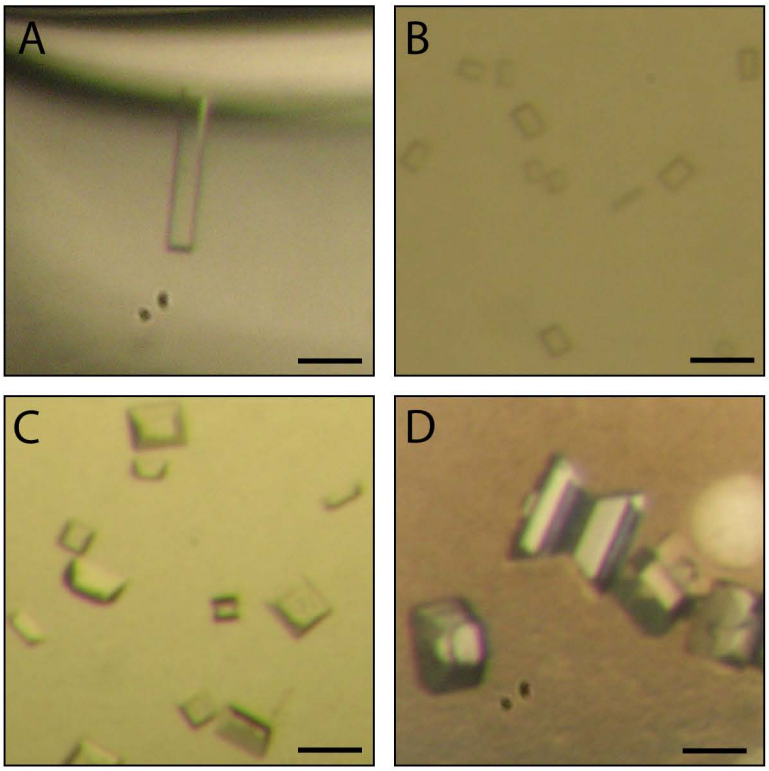


Figure 3.5 Images of DDX1 protein crystals that were further optimized for diffraction.

(A) Full length recombinant huDDX1 protein crystals formed in 0.2 M tri-lithium citrate and 20% (w/v) PEG 3350. (B) Full length recombinant huDDX1 protein crystals formed in 0.2 M sodium acetate, 0.1 M Tris-HCl pH 8.5 and 30% PEG 4000. (C) N-terminal recombinant DDX1 protein crystals formed in 0.1 M MES pH 6.5 and 40% (v/v) PEG 200. (D) N-terminal recombinant DDX1 protein crystals formed in 0.1 M MES pH 6.5 and 25% (v/v) PEG 600. Scale bar represents 50 mm.

Table 3.1 Summary of crystallography experiments including small and large scale and crystals diffracted.

Protein	Suite	Well	Additional additives	Large scale	Diffracted
1-480	Classics	20			
1-480	Classics	21			
1-480	Classics	23			
1-480	JCSG I	6			
1-480	JCSG I	11			
1-480	JCSG I	28			
1-480	JCSG I	52			
1-480	JCSG I	61			
1-480	JCSG III	12			
1-480	JCSG III	23			
1-480	JCSG III	63			
1-480	JCSG IV	32			
1-480	JCSG IV	34			
1-480	JCSG IV	39			
C-terminal	Classics	5			
C-terminal	JCSG I	16			
C-terminal	JCSG I	28			
C-terminal	JCSG I	40			
C-terminal	JCSG II	16			
C-terminal	JCSG III	47			
C-terminal	JCSG IV	89			
flyDdx1	JCSG I	95	Mg^{2+}/RNA-DNA Duplex		
flyDdx1	JCSG III	6			
flyDdx1	JCSG III	6			
flyDdx1	JCSG III	12			
flyDdx1	JCSG III	12			
flyDdx1	JCSG III	23			
flyDdx1	JCSG IV	22			
flyDdx1	JCSG IV	32			
flyDdx1	JCSG IV	89			
flyDdx1	Nucleix	46		X	
huDDX1	Nucleix	5	ATP		
huDDX1	Nucleix	56	ATP		
huDDX1	PACT	91	ATP		

Table 3.1 Continued					
Protein	Suite	Well	Additional additives	Large scale	Diffracted
huDDX1	JCSG IV	31	GTP		
huDDX1	JCSG IV	32	GTP		
huDDX1	PACT	23	GTP		
huDDX1	JCSG IV	32	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	46	Mg^{2+}/ATP/ Duplex	X	
huDDX1	Nucleix	56	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	13	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	27	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	31	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	39	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	42	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	64	Mg^{2+}/ATP/ Duplex		
huDDX1	Nucleix	89	Mg^{2+}/ATP/ Duplex		
huDDX1	PACT	17	Mg^{2+}/ATP/ Duplex		
huDDX1	PACT	20	Mg^{2+}/ATP/ Duplex		
huDDX1	PACT	34	Mg^{2+}/ATP/ Duplex		
huDDX1	PACT	35	Mg^{2+}/ATP/ Duplex		
huDDX1	PACT	96	Mg^{2+}/ATP/ Duplex		
huDDX1	JCSG I	4	Duplex		
huDDX1	JCSG I	30	Duplex		
huDDX1	JCSG IV	1	Duplex		
huDDX1	JCSG IV	12	Duplex		
huDDX1	JCSG IV	29	Duplex		
huDDX1	JCSG IV	32	Duplex		
huDDX1	JCSG IV	57	Duplex		
huDDX1	Nucleix	36	Duplex		
huDDX1	Nucleix	62	Duplex		
huDDX1	Nucleix	78	Duplex	X	
huDDX1	Nucleix	90	Duplex	X	
huDDX1	Nucleix	93	Duplex		
huDDX1	Cations	58			
huDDX1	Cations	59			
huDDX1	Classics Lite	2			
huDDX1	Classics Lite	65			

Table 3.1 Continued					
Protein	Suite	Well	Additional additives	Large scale	Diffracted
huDDX1	JCSG +	32		X	
huDDX1	JCSG III	26		X	
huDDX1	JCSG IV	18		X	
huDDX1	JCSG IV	26		X	
huDDX1	JCSG IV	28			
huDDX1	JCSG IV	32		X	
huDDX1	JCSG IV	89		X	
huDDX1	JCSG+	22			
huDDX1	JCSG+	88		X	X
huDDX1	JCSG+	89			
huDDX1	JCSG+	95			
huDDX1	Mb Class II	12			
huDDX1	Mb Class II	15			
huDDX1	PEGs	7		X	X
huDDX1	PEGs	95		X	X
huDDX1	PEGs	96		X	X
huDDX1	pH Clear II	17		X	
N-terminal	JCSG II	14	Mg^{2+}/ATP		
N-terminal	JCSG II	35	Mg^{2+}/ATP		
N-terminal	JCSG III	5	Mg^{2+}/ATP		
N-terminal	JCSG III	6	Mg^{2+}/ATP		
N-terminal	JCSG III	17	Mg^{2+}/ATP		
N-terminal	JCSG III	21	Mg^{2+}/ATP		
N-terminal	JCSG III	23	Mg^{2+}/ATP		
N-terminal	JCSG III	28	Mg^{2+}/ATP		
N-terminal	JCSG III	31	Mg^{2+}/ATP		
N-terminal	JCSG III	46	Mg^{2+}/ATP		
N-terminal	JCSG III	47	Mg^{2+}/ATP		
N-terminal	JCSG III	48	Mg^{2+}/ATP		
N-terminal	JCSG IV	9	Mg^{2+}/ATP		
N-terminal	Nucleix	27	Mg^{2+}/ATP		
N-terminal	JCSG II	14			
N-terminal	JCSG II	23			

Table 3.1 Continued					
Protein	Suite	Well	Additional additives	Large scale	Diffracted
N-terminal	JCSG III	4			
N-terminal	JCSG III	5			
N-terminal	JCSG III	6			
N-terminal	JCSG III	7			
N-terminal	JCSG III	9			
N-terminal	JCSG III	12			
N-terminal	JCSG III	17			
N-terminal	JCSG III	17			
N-terminal	JCSG III	21			
N-terminal	JCSG III	23			
N-terminal	JCSG III	28			
N-terminal	JCSG III	31		X	
N-terminal	JCSG III	46		X	X
N-terminal	JCSG III	47			
N-terminal	JCSG III	63			
N-terminal	JCSG IV	4			
N-terminal	JCSG IV	9			
N-terminal	JCSG IV	10			
N-terminal	JCSG IV	11			
N-terminal	JCSG IV	15			
N-terminal	JCSG IV	16			
N-terminal	JCSG IV	21			
N-terminal	JCSG IV	23			
N-terminal	JCSG IV	31			
N-terminal	JCSG IV	32			
N-terminal	JCSG IV	35			
N-terminal	JCSG IV	39			
N-terminal	JCSG IV	89			
N-terminal	Nucleix	27			
N-terminal	Nucleix	38			
N-terminal	Nucleix	39			
N-terminal	Nucleix	43			
N-terminal	Nucleix	46		X	
N-terminal	Nucleix	47			

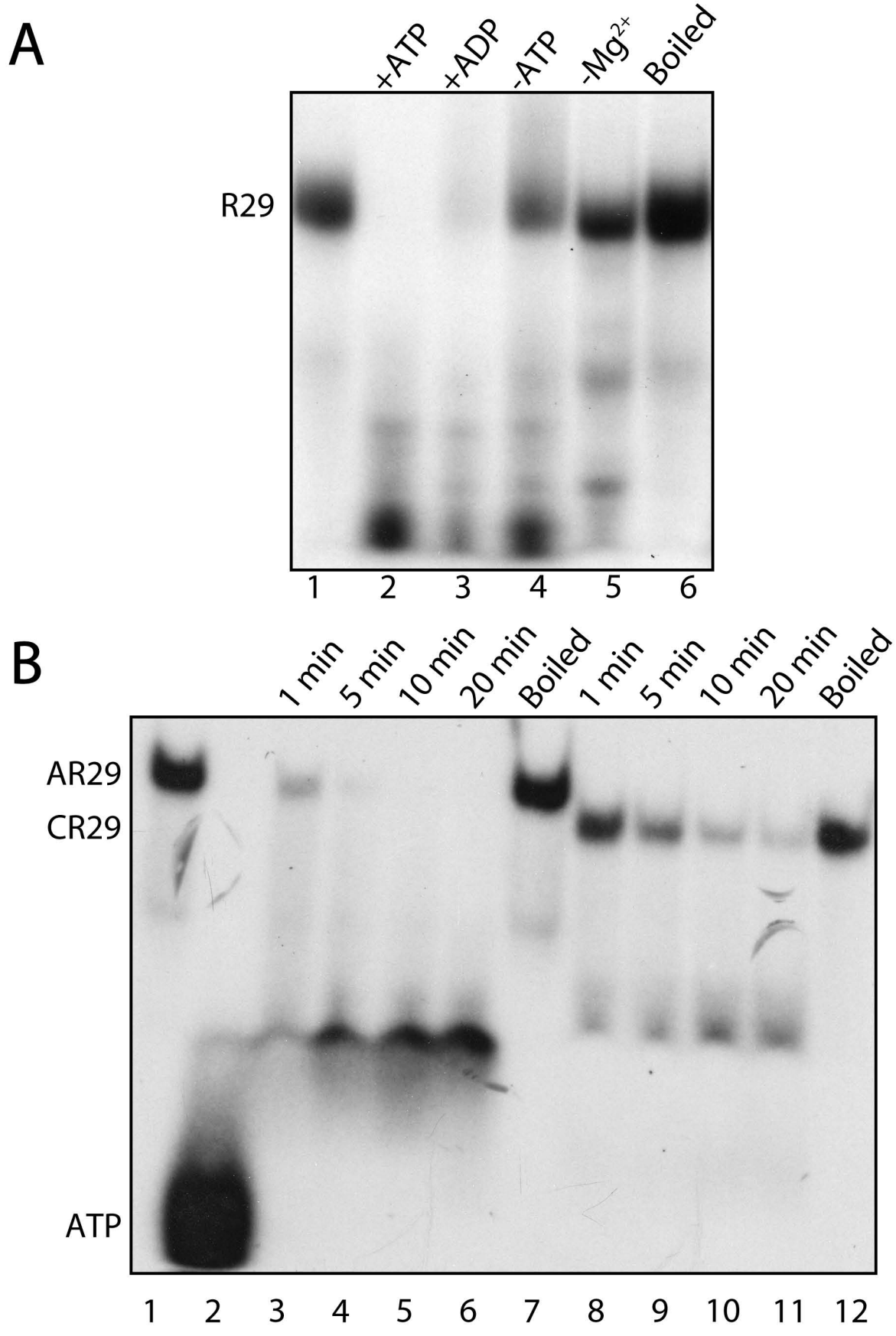
enzymes. To investigate whether DDX1 has sequence specificity for RNA degradation, we obtained commercial poly-A (AR29) and poly-C (CR29) RNA substrates that were each 29 nt in length (Table 2.2). These two new substrates along with the previously used 29-nt ssRNA substrate (R29) were radioactively end-labeled at the 5'-end with $^{32}P\gamma$-ATP using T4 polynucleotide kinase.

We first repeated our previous experiments using 0.6 μg recombinant DDX1 protein incubated with 50 fmol R29 in either Mg^{2+} containing buffer with ATP, ADP, or absence of nucleotide; or in Mg^{2+}-minus buffer that also contained EDTA to chelate any Mg^{2+} bound to the purified recombinant DDX1 protein. The reaction was incubated at 37°C for 20 minutes, then quenched and loaded on a 20 cm 10% polyacrylamide gel to separate the fragments. As previously reported, recombinant DDX1 protein degraded the R29 substrate in a nucleotide independent manner (lanes 2-4) and degradation was Mg^{2+}-dependent (lane 5) (Figure 3.6). We were also able to show that boiling DDX1 for 2 minutes was sufficient to abolish its ribonuclease activity (lane 6). These results demonstrate that the freshly prepared recombinant DDX1 protein shares the same properties as previously prepared batches of recombinant protein.

Next, we repeated the experiments with AR29 and CR29. Similar results were observed, with DDX1 degrading both substrates in the absence of nucleotides. As previously noted for R29, boiling DDX1 inactivated its ribonuclease activity towards AR29 and CR29 (Figure 3.6). We found that DDX1 was able to degrade both these

Figure 3.6 DDX1 degrades ssRNA substrates in the presence or absence of nucleotides *in vitro*.

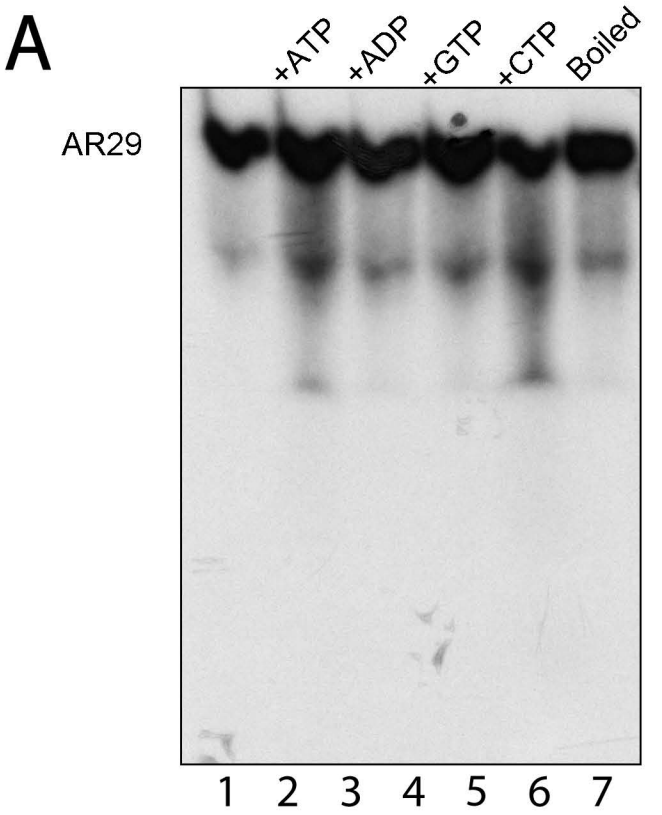
(A) Fifty fmol of 5′-labeled R29 substrate was incubated with 0.6 μg of recombinant DDX1 protein in the presence of 1 mM ATP (lane 2) or 1 mM ADP (lane 3). The R29 substrate was incubated in unwinding buffer without protein (lane 1) or with heat-inactivated DDX1 (lane 6). Mg^{2+} was omitted in lane 5. All reactions were incubated for 20 minutes at 37°C prior to electrophoresis in a 15% polyacrylamide gel followed by autoradiography. +, with; -, without. (B) Fifty fmol of either 5′-labeled AR29 (lanes 3-7) or CR29 (lanes 8-12) were incubated with 0.6 μg recombinant protein in the absence of ATP or ADP. Reactions were staggered to generate a time course to a maximum of 20 minutes. Substrates were mock-treated (lane 1) or incubated with heat-inactivated DDX1 for 20 minutes (lanes 7 and 12). Reaction mixtures were then electrophoresed in a 15% polyacrylamide gel followed by autoradiography.


A
+ATP
+ADP
-ATP
-Mg2+
Boiled
R29
1
2
3
4
5
6
B
1 min
5 min
10 min
20 min
Boiled
1 min
5 min
10 min
20 min
Boiled
AR29
CR29
ATP
1
2
3
4
5
6
7
8
9
10
11
12

substrates effectively, although AR29 was degraded faster than CR29, with complete degradation observed for AR29 by 10 minutes. In contrast, low amounts of CR29 could still be detected after 20 minutes of incubation. In an attempt to identify the types of degradation products produced by DDX1, we included free ^{32}P-ATP as a size marker for single nucleotides to determine whether the substrates were degraded to single nucleotides. We observed that the size of the degradation products was larger than a single nucleotide with a predicted size of approximately 3-8 nucleotides in length. The other difference noted was that there was also a larger product of approximately 15 nucleotides in length found in small amounts when R29 was degraded in the absence of nucleotides, as compared to when AR29 or CR29 were degraded. This larger product may be the result of R29 forming secondary structures that are refractory to DDX1 degradation. However, in the presence of ATP or ADP, DDX1 is able to unwind these secondary structures and degrade the ssRNA to fragment sizes of ~3-8 nucleotides. The AR29 and CR29 substrates, on the other hand, are simple chains of A or C which cannot form secondary structures, and thus are digested by DDX1 even in the absence of ATP or ADP to a size of 3-8 nucleotides.

When studying ribonuclease activity of recombinant proteins, there is always a concern that the protein preparation may be contaminated with RNase A. We have previously shown that DDX1's ribonuclease activity is different from that of RNase A when R29 is used as the ssRNA substrate. We therefore incubated our AR29 and CR29 substrates with RNase A under the same conditions as indicated for DDX1 (Figure 3.7). We found that RNase A was able to degrade CR29 under all conditions and was not

Figure 3.7. RNase A has sequence specific ribonuclease activity. Fifty fmol of 5′-labeled AR29 (A, a stretch of 29 As) or CR29 (B, a stretch of 29 Cs) substrates were incubated with 0.1 unit RNase A in the presence of 1 mM ATP, ADP, GTP or CTP (lanes 2-5). Boiled RNase A was used in lane 6. No RNase A was added in lane 1 in both (A) and (B). All reactions were incubated for 20 minutes at 37°C prior to electrophoresis in a 15% polyacrylamide gel followed by autoradiography.


A
+ATP
+ADP
+GTP
+CTP
Boiled
AR29
1 2 3 4 5 6 7
B
+ATP
+ADP
+GTP
+CTP
Boiled
CR29
1 2 3 4 5 6 7

affected by boiling. As expected, we also found that RNase A was unable to degrade AR29 as RNase A cleaves specifically after pyrimidines and is unable to degrade a purine substrate. In accordance with these data, we suggest that the ribonuclease activity observed upon incubation of CR29 and AR29 with DDX1 is solely due to DDX1 activity.

Next, we tested the consequence of adding nucleotides other than ATP or non-hydrolysable ATP and ADP analogues on DDX1 degradation activity. We used the AR29 substrate for these analyses to ensure that we were only observing degradation and not unwinding activity. We found that DDX1 was able to effectively degrade AR29 regardless of nucleotide added including ATP or ADP analogues (Figure 3.8), indicating that DDX1 can degrade ssRNA in a nucleotide-independent manner.

Our results indicate that DDX1 degrades R29 ssRNA to products of approximately 3-8 nucleotides; however, due to gel length we were not able to resolve the exact size of these small oligonucleotide products. In an attempt to accurately resolve the size of the degradation products, we utilized a 40 cm gel apparatus. A single nucleotide ladder was generated by acid hydrolysis. Recombinant DDX1 was incubated with AR29 at increasing 2 minutes intervals to identify intermediate as well as final products. We observed bands of virtually every possible size at the shorter incubation times. A significant enrichment of fragment sizes ranging from 3-6 nucleotides in length was observed in all lanes (Figure 3.9). One of the limitations with this gel is that the reaction seemed to be almost complete even after just 1 minute of incubation. Together, the degradation data may suggest that

DDX1 acts as an endoribonuclease as there are no single nucleotides generated following degradation. There is some laddering in the large gel and this may be due to the presence of contaminants or truncated products in our substrate preparation. An alternative explanation is that DDX1 may act as a 5′-3′ exonuclease. Additional experiments with less protein will be require to determine whether the laddering is due to DDX1 or inherent instability of the protein.

To characterize the ribonuclease activity of DDX1, we designed substrates that could be labeled at a single nucleotide located at different positions in the ssRNA (Table 2.2). To synthesize these substrates, we purchased custom-synthesized ssDNAs that consisted of a T7 promoter at the 5′-end followed by "AAAAG" repeats to generate 29-nucleotide DNA fragments. Three ssDNA substrates were designed so that the 5th, 15th or 29th nucleotide was substituted with uridine (2929: GAAAAGAAAAG AAAAGAAAAGAAAAGAA**U**, 2905: GAAA**U**GAAAAGAAAAGAAAAGAAAAGAAA, and 2915: GAAAAGAAAAGAAA**U**GAAAAGAAAAGAAA). We performed *in vitro* transcription in the presence of [α-^{32}P] UTP to generate ssRNA with the single uridine labeled with ^{32}P. These radioactive substrates were purified by acrylamide gel and phenol extraction. Recombinant DDX1 protein was incubated with the substrates in the absence of nucleotides for 20 minutes at 37°C. The resulting reactions were then electrophoresed in a 20 cm 10% acrylamide gel to separate the products (Figure 3.10).

The signals in the AR29 lanes were weak due to the use of older substrate that had undergone radioactive decay. Some differences in the size of the degradation

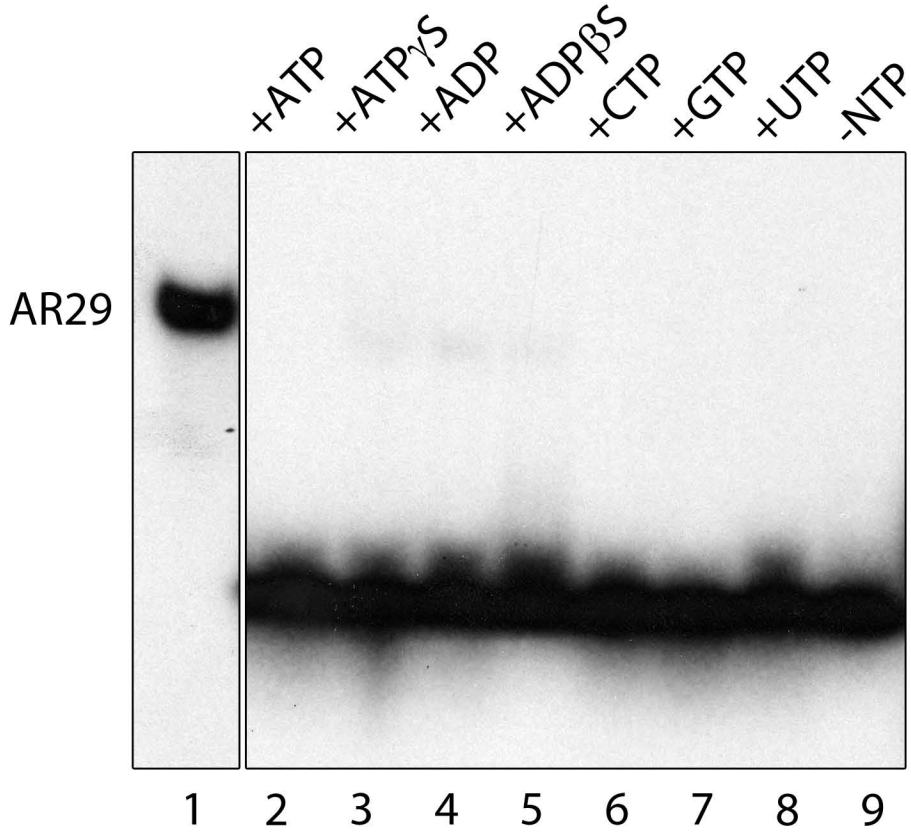
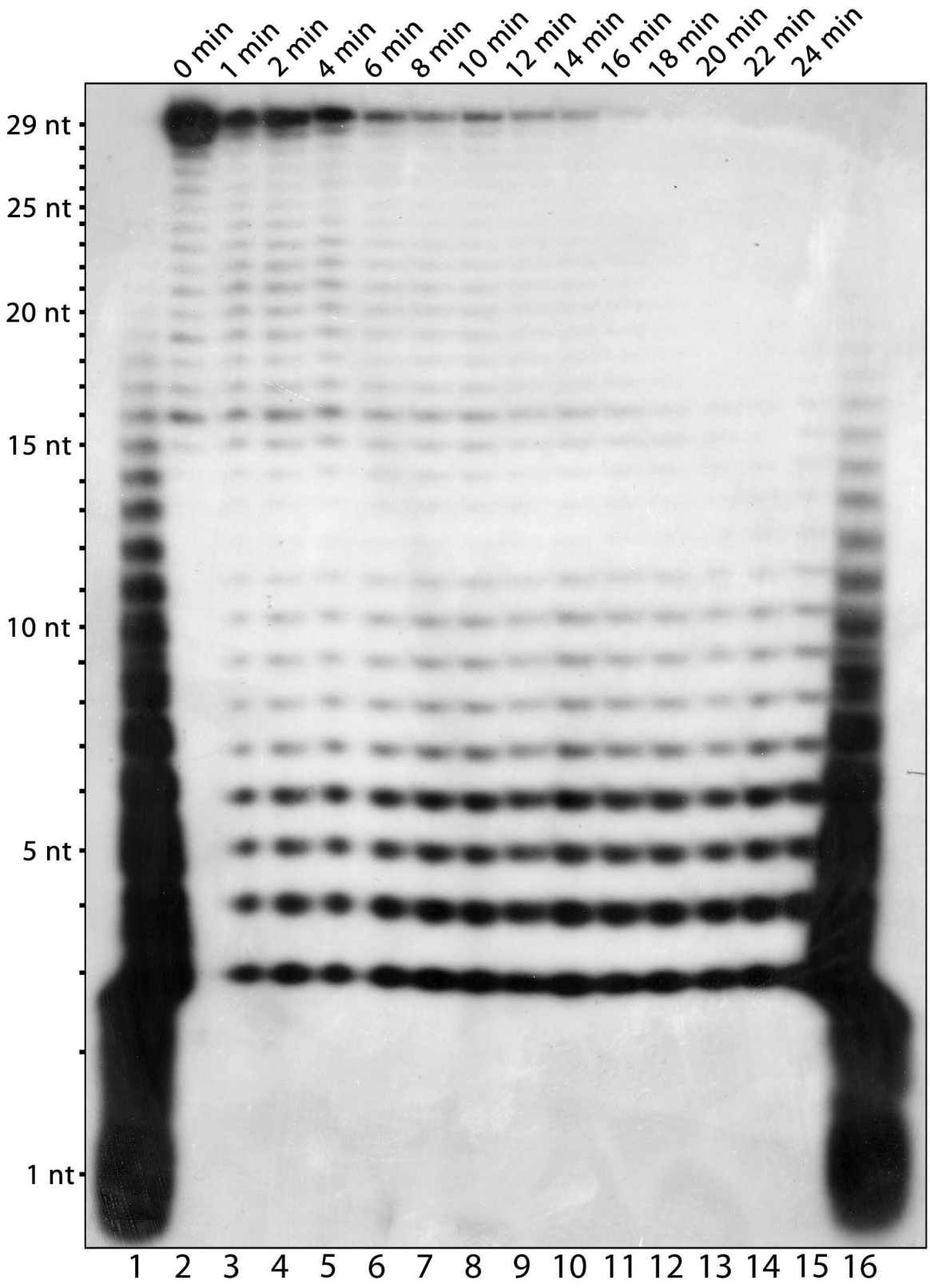


Figure 3.8 DDX1 degrades ssRNA in the presence or absence of nucleotides.

(A) Fifty fmol of 5′-labeled AR29 substrate (a stretch of 29 As) was incubated with 0.6 μg recombinant DDX1 protein (lanes 2-8) in the presence of 1 mM ATP, ATPγS, ADP, ADPβS, CTP, GTP, or UTP. No nucleotides were added to the reaction in lane 9. DDX1 was omitted in lane 1. All reactions were incubated for 20 minutes at 37°C prior to electrophoresis in a 15% polyacrylamide gel followed by autoradiography.

Figure 3.9 High resolution separation of ssRNA following degradation by recombinant DDX1.

(A) Fifty fmol of 5′-labeled AR29 substrate were incubated with 0.6 μg recombinant DDX1 protein (lanes 2-15). Reaction mixtures were incubated at 37°C for the indicated times. Mixtures were then loaded on a 40 cm 15% native polyacrylamide gel followed by autoradiography.


0 min
1 min
2 min
4 min
6 min
8 min
10 min
12 min
14 min
16 min
18 min
20 min
22 min
24 min
29 nt
25 nt
20 nt
15 nt
10 nt
5 nt
1 nt
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

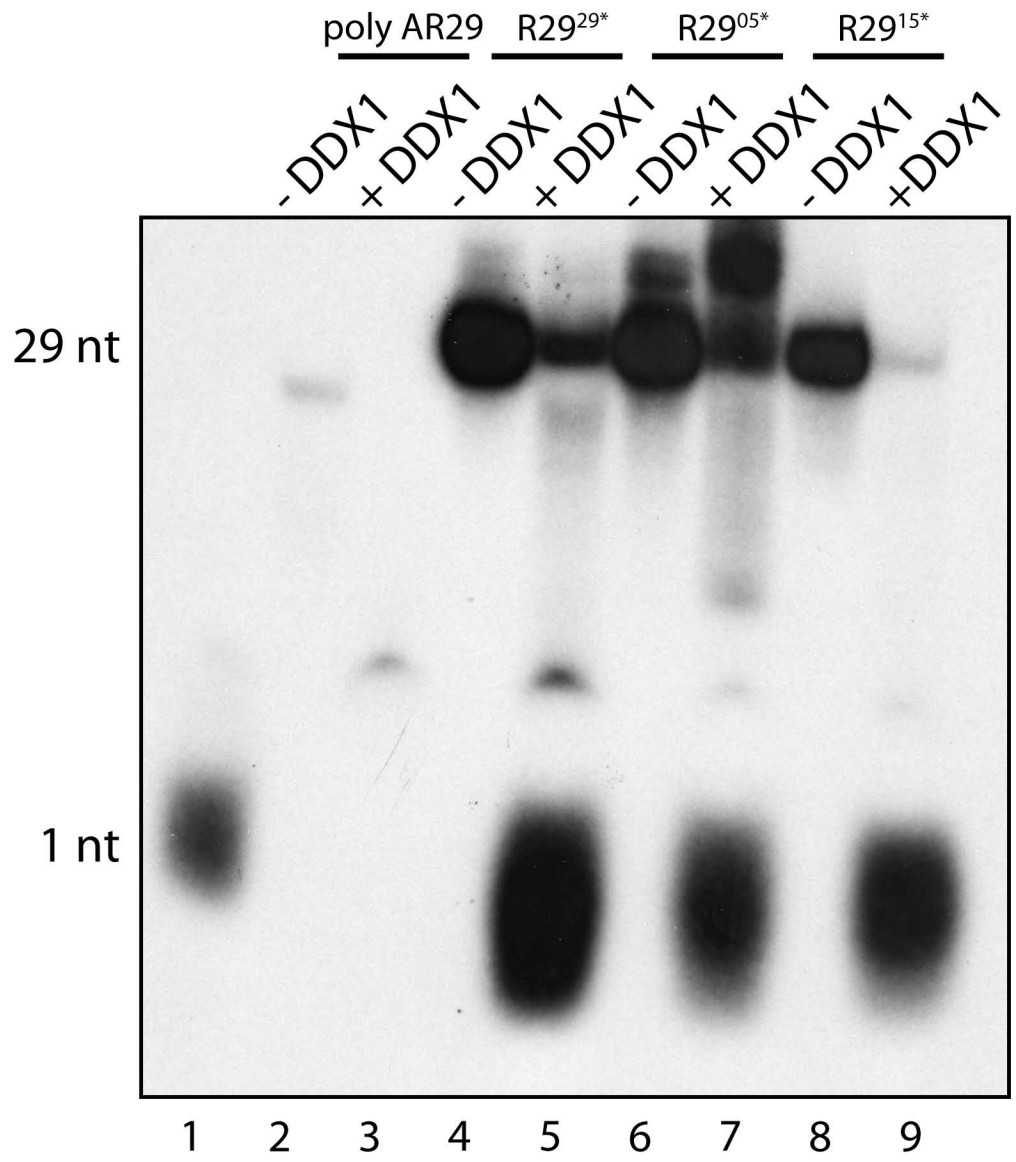


Figure 3.10 Degradation of 5′-labeled and internal labeled ssRNA substrates by recombinant DDX1.

(A) R29 substrate was radio-labeled either at 5′-end or internally (^{32}P-UTP). Fifty fmol of end labeled R29 substrate or 50 fmol of internal labeled R2929*, R2905*, R2915* ssRNA substrate was incubated with 0.6 μg recombinant DDX1 protein (lanes 3, 5, 7, and 9). The * represents the position of ^{32}P-labeled nucleotide. All reactions were incubated for 20 minutes at 37°C prior to electrophoresis in a 15% polyacrylamide gel followed by autoradiography.

products were observed between the two internally-labeled and end-labeled substrates. Both $R29^{29*}$ (*denoting the labeled nucleotide) and $R29^{05*}$ were poorly degraded by DDX1; however, $R29^{15*}$ was almost completely degraded after 20 minutes. In addition to a major band at 1 nucleotide, we also observed a minor band at around 3-6 nucleotides when the end-labeled substrate was incubated with DDX1. The single nucleotide band suggests that DDX1 may degrade ssRNA in an exonuclease manner in the 3'-5' direction. Additional experiments will be required to determine whether DDX1 is an endo- or exoribonuclease or both.

3.5 Mapping key regions of DDX1 required for ssRNA degradation

The amino acids or amino acid motifs responsible for DDX1's ribonuclease activity remain to be identified. Sequence comparisons reveal no regions of similarity between DDX1 and other ribonucleases. Furthermore, no other DEAD box proteins have been shown to have ribonuclease activity. One major region of interest is the C-terminal end of DDX1 as the sequence downstream of conserved DEAD box motif VI shares no homology with any other proteins. Previously, we had generated C-terminal truncations of DDX1 as part of our X-ray crystallography experiments. We selected two of these C-terminal truncation constructs, $DDX1_{1-620}$ and $DDX1_{1-645}$, to determine whether they retained ribonuclease activity. We also tested several other recombinant DDX1 proteins for ribonuclease activity including SPRYless DDX1, the N-terminal domain of DDX1, and a mutant DDX1 protein previously generated in the lab with motif II (D-E-A-D) mutated to AAAD.

R29 was incubated with the different recombinant DDX1 proteins under the same conditions used for full-length DDX1 protein analyses (Figure 3.11). We found that the SPRYless, N-terminal and $DDX1_{1-620}$ mutants were unable to degrade R29. The AAAD mutant showed very weak ribonuclease activity, whereas $DDX1_{1-645}$ had similar ribonuclease activity compared to full length. It is unlikely that the SPRY domain has an enzymatic role as the SPRY domain is found in several other proteins and is primarily responsible for protein-protein or protein-RNA interactions. The SPRYless mutant may not be functional due to misfolding of the N-terminal domain, or the SPRY domain may play a role in RNA binding. The lack of ribonuclease activity when the DEAD motif is mutated to AAAD is not surprising as the DEAD motif is responsible for ATP and Mg^{2+} binding, and wild-type DDX1 in the absence of Mg^{2+} has already been shown to have no ribonuclease activity. Thus, the most informative result from these experiments is the observation that $DDX1_{1-620}$ has no ribonuclease activity whereas $DDX1_{1-645}$ retains full ribonuclease activity, suggesting the presence of a ribonuclease motif in the C-terminus of DDX1.

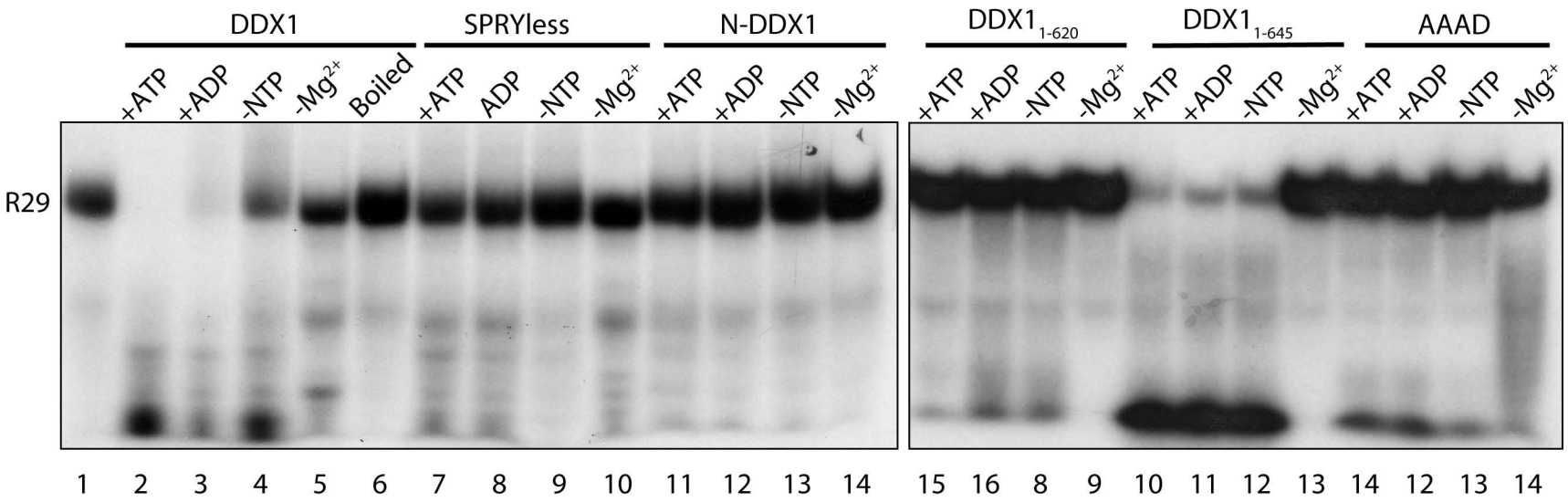


Figure 3.11 Ribonuclease activity of recombinant mutant DDX1 proteins.

Fifty fmol of 5'-labeled R29 substrate was incubated with 7.2 pmol of recombinant mutant DDX1 proteins in the presence of 1 mM ATP, ADP, in the absence of nucleotides or in the absence of Mg^{2+}. Heat-inactivated protein was used in lane 6. All reactions were incubated for 20 minutes at 37^{O}C prior to electrophoresis in a 15% polyacrylamide gel followed by autoradiography.

Chapter 4

Characterizing *Ddx1* knockout mouse

DDX1 is expressed in all tissues and cell lines tested to date, although at different levels. Proliferating cells and terminally differentiated cells derived from the neuroectoderm express the highest levels of DDX1 (98). DDX1 is also highly expressed in cancer cell lines (54,89,90,92,237), as well as in a subset of neuroblastoma (71,82,85,93,238) and breast cancer tumours (95,96). *Ddx1* knockout in *Drosophila* was reported to be embryonic lethal (99); however, work performed in our lab indicates that *Ddx1* knockout flies are viable but infertile (Devon Germain, unpublished data). DDX1 is highly expressed in *Drosophila* embryos (100). Together, these data suggest that DDX1 is important in a wide range of cell types, and that it may play a key role in proliferating cells during development.

Knockout mice are useful tools to study the role of disrupted genes. Mouse knockout models have been generated for a number of DEAD box family members (222-224,227-229). A variety of phenotypes have been observed in these DEAD box gene knockout models, in agreement with DEAD box proteins having a wide array of functions. Many of the phenotypes observed are developmental defects resulting in lethality as early as the 2-cell stage to 2 days postpartum.

DDX1 plays a role in the repair of DNA double strand breaks by homologous recombination [(124) and Lei Li, unpublished data]. Disruption of a number of genes involved in DNA repair has been shown to be developmentally lethal. For example, mutation of *Mre11*, whereby the nuclease activity was abolished, resulted in embryonic lethality at E8.5 (239). *ATR* knockout mice die by E7.5 with $ATR^{-/-}$ blastocysts

showing chromosomal instability and elevated levels of apoptosis after 5 days in culture (240).

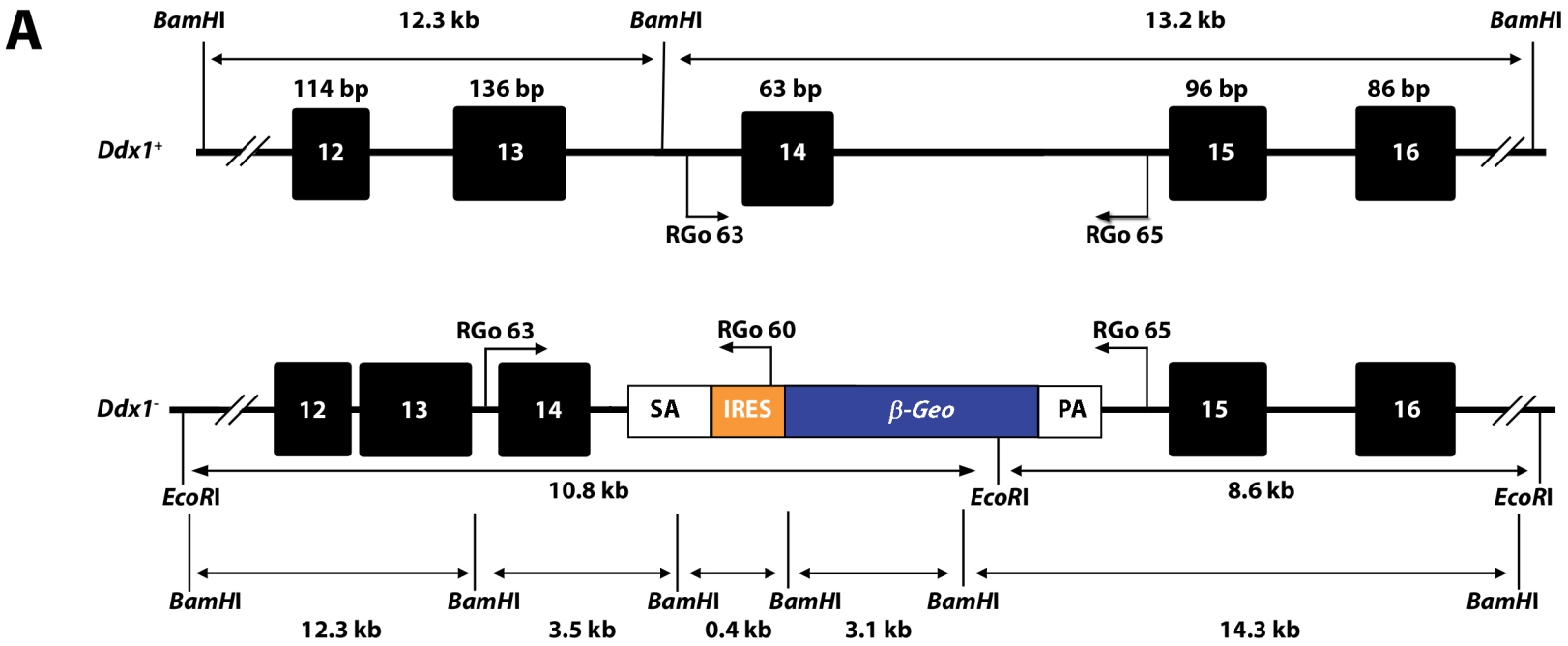
Generating a *Ddx1* knockout mouse model has a number of advantages including the possibility of studying the role of DDX1 in the context of the whole animal and generating *Ddx1*$^{-/-}$ lines that could then be used to carry out *in vitro* experiments to examine the consequence of loss of DDX1 on specific pathways. Here, we describe the generation and characterization of a *Ddx1* knockout mouse model. We also describe our attempts to generate a conditional *Ddx1* mouse knockout.

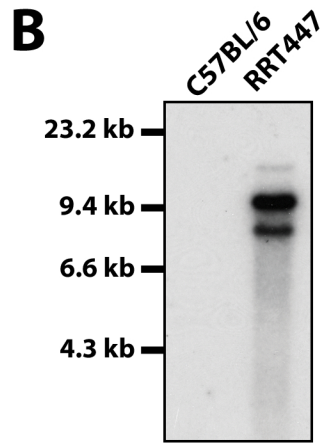
4.1 Generation of a constitutive *Ddx1* knockout mouse model

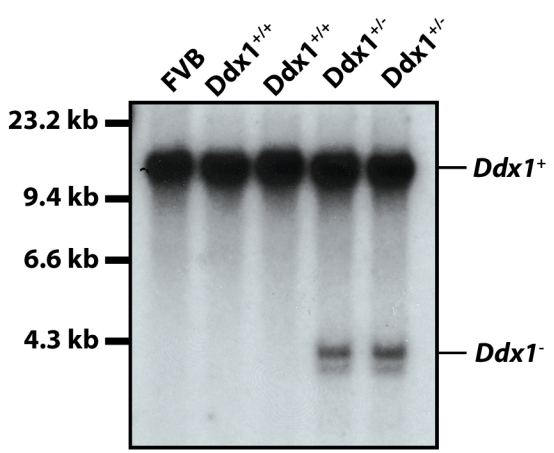
The mouse embryonic stem cell line (RRT447) containing an intronic gene trap within intron 14 of the *Ddx1* gene was purchased from BayGenomics. This gene trap results in the fusion of *Ddx1* exons 1-14 to the *β-geo* reporter gene (β-galactosidase/neomycin resistance). Chimeric *Ddx1* mice were generated by microinjecting this mouse embryonic stem cell line into C57BL/6 blastocysts. Chimeric mice were identified by coat colour. Male chimeric mice were mated to C57BL/6 females to obtain germ line transmission of the *Ddx1*$^{Gt(RRT447)RG}$ allele, which will be referred to as *Ddx1*$^{-}$ (Figure 4.1). Two lines of *Ddx1* gene knockout mice (abbreviated as C6 and C7) were generated using this strategy. To confirm *Ddx1* gene disruption at exon 14 and ensure that there was a single insertion site of the *β-geo* reporter gene in our two lines, we carried out Southern blot analysis. Using a ^{32}P-labeled cDNA probe to *β-geo*, we confirmed the presence of a single gene-trap insertion in the RRT447 cell

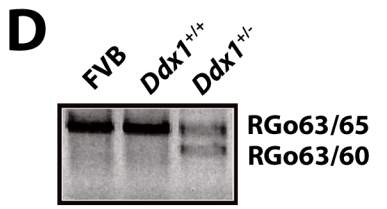
Figure 4.1 Genomic map of the gene-trap insertion site.

Embryonic stem cells that contain a single gene-trap insertion in *Ddx1* were purchased from BayGenomics. (A) The insertion containing a *β-geo* gene, splice acceptor (SA) and a polyadenylation signal (PA) is located between exons 14 and 15 of *Ddx1*. The insertion generates a truncated DDX1 protein fused to β-geo. Locations of primers (RGo) used for genotyping are also shown. (B) Southern blot analysis of the RRT447 cell line. Genomic DNA extracted from a RRT447 cell pellet and wild-type mouse tail was digested with *EcoR*I and separated in a 1% agarose gel. The DNA was transferred to a nitrocellulose membrane and hybridized with a ^{32}P-labeled cDNA probe specific to *β-geo*. The signal was detected by autoradiography. (C) Southern blot analysis of wild-type and DDX1$^{+/-}$ mice. Genomic DNA extracted from mice was digested with *BamH*I and separated in a 1% agarose gel. The DNA was transferred to a nitrocellulose membrane and hybridized with a ^{32}P-labeled cDNA probe specific to DDX1. The signal was detected by autoradiography. (D) PCR amplification of genomic DNA for routine genotyping using primers shown in (A).


A
BamHI
12.3 kb
BamHI
13.2 kb
BamHI
114 bp
136 bp
63 bp
96 bp
86 bp
Ddx1+
12
13
14
15
16
RGo 63
RGo 65
RGo 63
RGo 60
RGo 65
Ddx1-
12
13
14
SA
IRES
β-Geo
PA
15
16
EcoRI
10.8 kb
EcoRI
8.6 kb
EcoRI
BamHI
BamHI
BamHI
BamHI
BamHI
BamHI
12.3 kb
3.5 kb
0.4 kb
3.1 kb
14.3 kb


B
C57BL/6
RRT447
23.2 kb
9.4 kb
6.6 kb
4.3 kb


C
FVB
Ddx1+/+
Ddx1+/+
Ddx1+/-
Ddx1+/-
23.2 kb
9.4 kb
6.6 kb
4.3 kb
Ddx1+
Ddx1-


D
FVB
Ddx1+/+
Ddx1+/-
RGo63/65
RGo63/60

line (Figure 4.1B). Using a ^{32}P-labeled cDNA probe corresponding to exons 10-17 of *Ddx1*, we confirmed the presence of the mutant *Ddx1* allele, with a band of approximately 4 kb observed in heterozygous mice compared to wild-type mice (Figure 4.1C). Subsequent genotype analysis was performed by multiplex PCR with primers designed to specifically amplify the *Ddx1*$^{+}$ and *Ddx1*$^{-}$ alleles (Figure 4.1D).

4.2 *Ddx1*$^{-/-}$ embryos die pre-implantation

We initially genotyped 21 weaned pups obtained from heterozygote intercrosses by PCR analysis. No *Ddx1*$^{-/-}$ pups were identified, suggesting that the *Ddx1*$^{-/-}$ genotype was lethal. Ultimately, 408 weaned pups generated from heterozygote intercrosses from both the C6 and C7 lines were genotyped by PCR and not a single *Ddx1*$^{-/-}$ mouse was identified (Table 4.1). As no deaths were observed either perinatally or postnally from these heterozygote intercrosses, we surmised that *Ddx1*$^{-/-}$ lethality was prenatal. To determine when the *Ddx1*$^{-/-}$ mice were dying, embryos were collected between E6.5 and E10 and genotyped (Figure 4.2). No homozygous mutant embryos were identified at these stages, suggesting that *Ddx1*$^{-/-}$ embryos die pre-implantation.

To further define when the *Ddx1*$^{-/-}$ embryos were dying, we collected fertilized embryos from heterozygote intercrosses at E0.5, removed the cumulus cells and placed them in culture. Embryos were observed daily. After 24 hours, most of the single cell embryos had divided into 2 cells (92%). By 48 hours, only 68% of the embryos were at the 4-cell stage and by 72 hours, 64% of the embryos were at the

Table 4.1. Genotyping weaned progeny of heterozygous matings				
Strain	Total	Genotype by PCR		
		+/+	+/-	-/-
C57BL/6/*Ddx1*$^{+/-}$	408	42 (10%)	366 (90%)	0

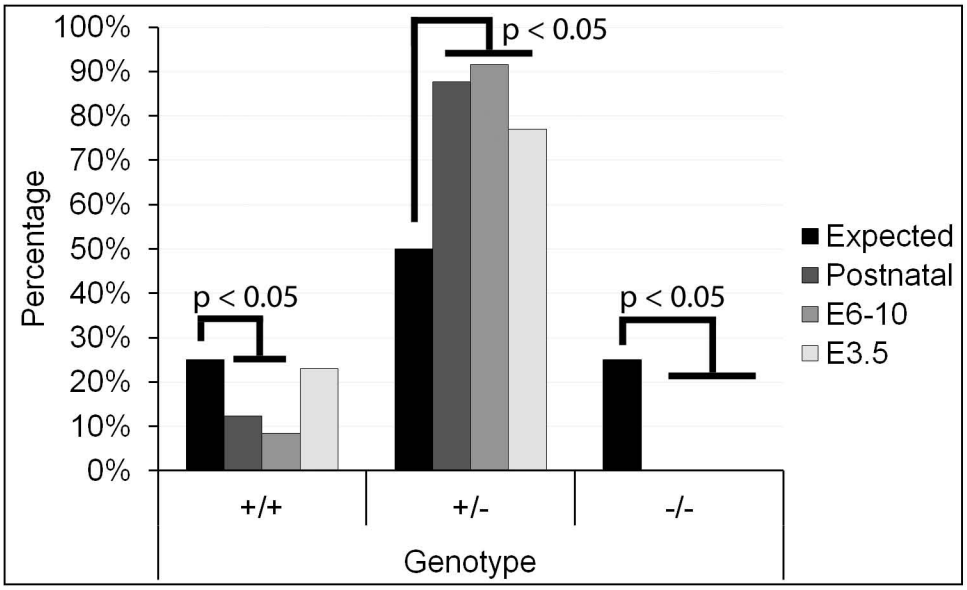


Figure 4.2 *Ddx1*$^{-/-}$ mice die pre-implantation.

Heterozygote mice were intercrossed and progeny genotyped at different stages. No *Ddx1*$^{-/-}$ mice were observed out of a total of 758 postnatal offspring, 225 E6-10 embryos and 91 E3.5 blastocysts genotyped. A significant decrease in the percentage of wild-type mice was observed post E3.5.

8-cell stage (Table 4.2). Our attempts to genotype the stalled embryos by PCR were unsuccessful, even when nested PCR was used. Stalled embryos were also fixed in paraformaldehyde and immunostained using anti-DDX1 (2910) antibody. Unless otherwise specified, we used the 2910 anti-DDX1 antibody for immunostaining or western blotting. Stalled embryos showed low DDX1 levels compared to wild-type embryos at the same stage (Figure 4.3). The residual staining in these embryos is likely due to maternally expressed protein and discussed further in chapter 5.

To determine when DDX1 is first expressed by the embryos, we crossed wild-type females to heterozygote males and collected 8 embryos at the 1-cell and 2-cell stages of development. Eight oocytes were also collected from heterozygote female ovaries. Using semi-quantitative RT-PCR with primers specific to *β-GAL*, we were able to detect the *Ddx1-β-GAL* fusion transcript in the oocyte (likely GV stage). At the 1-cell stage (E0.5), we were unable to detect the fusion transcript; however, the fusion transcript could again be detected at the 2-cell stage (E1.5) (Figure 4.4). Together, these data suggest that *Ddx1* is required by the 2-cell stage and that the *Ddx1*$^{-/-}$ embryos die between the 2- and 4-cell stages of development. Further characterization of DDX1 expression patterns during the development of normal embryos is included in Chapter 5.

Table 4.2. A portion of embryos generated from heterozygous intercrosses stall during preimplantation development				
		24 hours	48 hours	72 hours
Wild-type intercrosses (2 crosses)	8-cell	-	-	16 (94%)
	4-cell	-	16 (94%)	0
	2-cell	15 (88%)	0	0
	1-cell	2 (12%)	1 (5%)	1 (5%)
	Total embryos	17	17	17
Heterozygote intercrosses (17 crosses)	8-cell	-	-	94 (64%)
	4-cell	-	99 (68%)	16 (11%)
	2-cell	134 (92%)	35 (24%)	24 (16%)
	1-cell	12 (8%)	12 (8%)	12 (8%)
	Total embryos	146	146	146

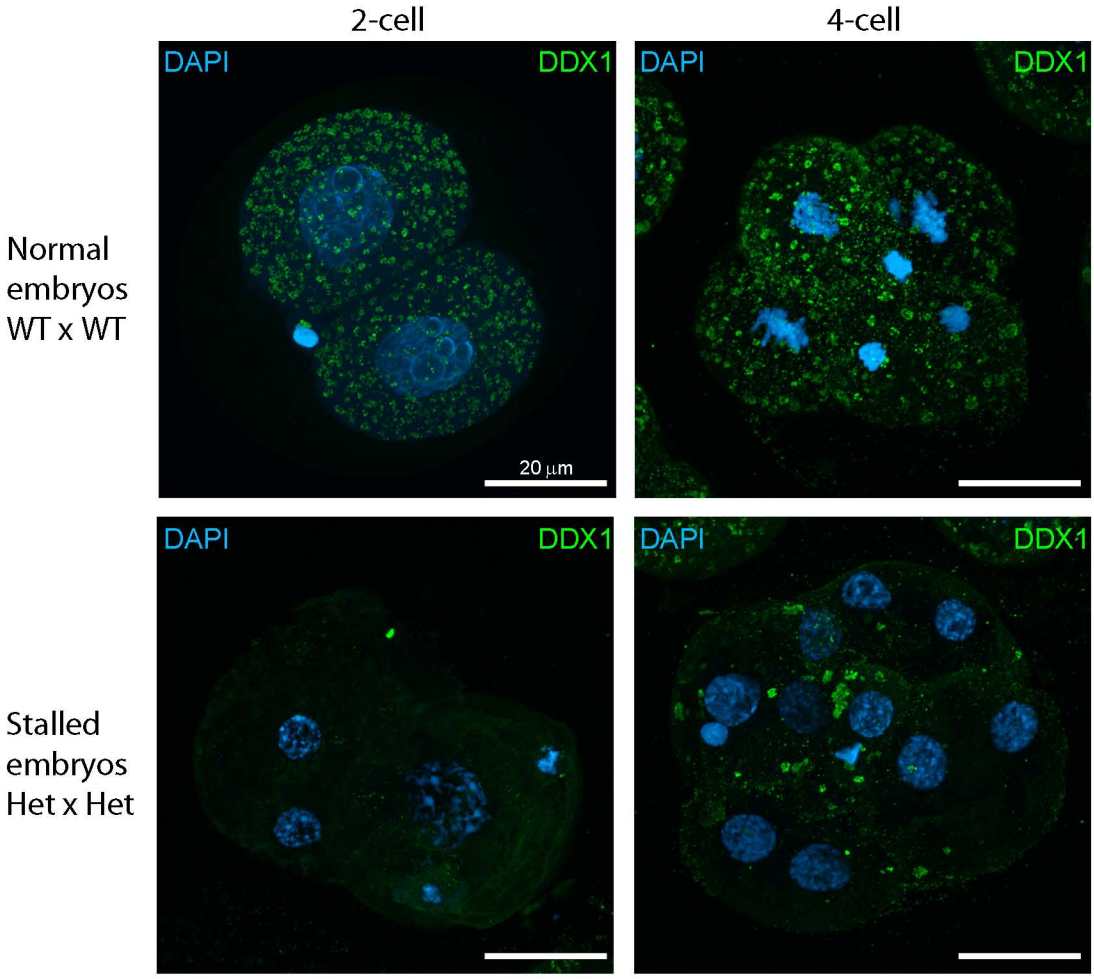


Figure 4.3 Immunofluorescent staining of developmentally stalled embryos.

Wild-type embryos were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were stained with anti-DDX1 antibody (batch 2910) and DAPI. Stalled embryos were collected from cultured embryos generated from *Ddx1*$^{+/-}$ intercrosses. Embryos (from wild-type and heterozygote crosses) were collected at E0.5 and the cumulus cells were removed by incubation in 0.3 mg/ml hyaluronidase for 30 seconds. The embryos were cultured in 25 μl droplets of KSOM media under mineral oil at 37^{O}C in 5% CO_2 incubator for 72 hours. After 72 hours, embryos should be at the 8-16 cell stage. Embryos were displayed as 2D projections of Z-stacks imaged by confocal microscopy. Scale bars = 20 μm.

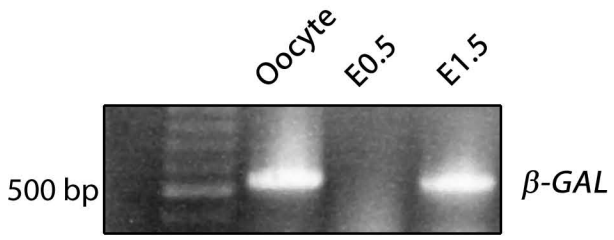


Figure 4.4 Expression of *DDX1* is initiated at the 2-cell stage.

Wild-type female mice were mated to *Ddx1* heterozygote males and the resulting offspring were collected at E0.5 and E1.5. Eight embryos were pooled and RNA was isolated. cDNA was generated using primers specific to *β-GAL*. Semi-quantitative RT-PCR was performed with primers specific to *β-GAL* and the products were run on a 1% agarose gel.

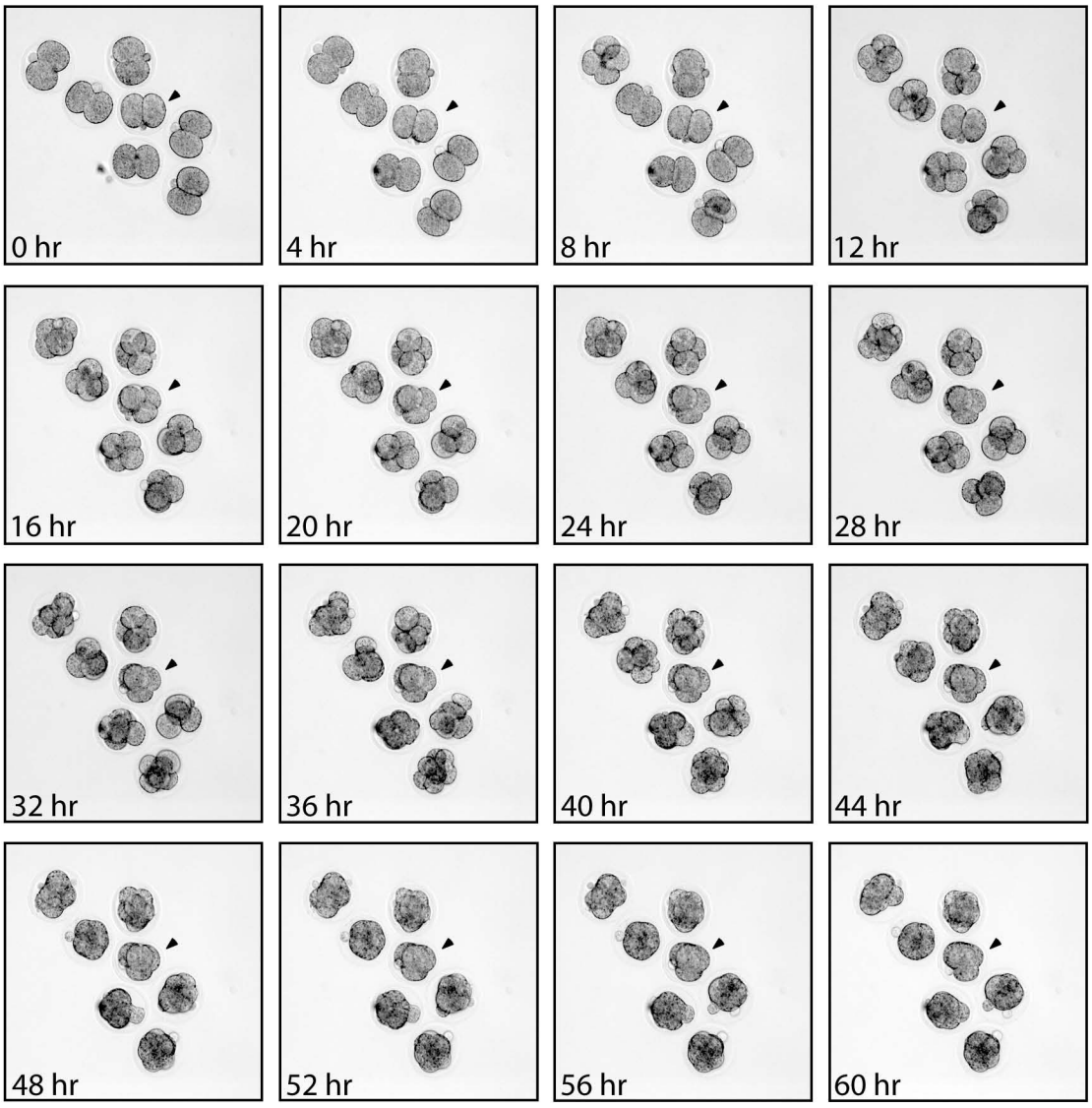


Figure 4.5 Live cell imaging of embryos derived from *Ddx1*$^{+/-}$ intercrosses.

Embryos were collected at E1.5 in M2 media. These embryos were transferred to glass bottom culture dishes containing10 μl droplets of M16 media under oil at 37^{O}C in 5% CO_2 in air. The lid of the culture dish was replaced with a glass topped lid. Images were collected every 3 minutes with a Zeiss Axiovert 200M microscope and a Zeiss plan-apochromat 10X/0.14 lens. Image processing was done using Metamorph v.7.8.3.0 (Molecular Devices). The arrowhead indicates the stalled embryo.

4.3 Live cell imaging of embryos collected from heterozygote intercrosses.

The 2- and 4-cell embryos that were stalled in their development also had multiple nuclei per cell and the nuclei looked abnormal. We used digital interference contrast (DIC) imaging in conjunction with time lapse photography to determine if there were any obvious problems with cell cycle progression (Figure 4.5). For these experiments, we collected embryos at the 2-cell stage, and recorded their development in culture over a period of 48 hours. While the embryos were able to divide normally, stalled embryos did not seem to be able to initiate cytokinesis. After the completion of the live cell imaging experiments, stalled embryos were immunostained with anti-DDX1 antibody. Reduced levels of DDX1 were observed as well as abnormal subcellular localization, with staining patterns that were virtually identical to those shown in Figure 4.3.

4.4 Non-Mendelian ratios are observed in progeny generated by intercrossed Ddx1$^{+/-}$ mice

The expected Mendelian ratio of progeny from intercrossed heterozygous mice is 1 wild-type to 2 heterozygotes to 1 knockout. The predicted ratio is 1 wild-type to 2 heterozygotes, as the *Ddx1*$^{-/-}$ genotype is lethal pre-blastocyst. Surprisingly, when we genotyped 408 weaned pups, the resulting ratio was 1 wild-type to 9 heterozygotes (Table 4.3). To rule out the possibility that reduced numbers of wild-type progeny

Table 4.3. Genotypes of weaned progeny from heterozygote matings				
Strain	Total	Genotype by PCR		
		+/+	+/-	-/-
C57BL/6/*Ddx1*$^{+/-}$	408	42 (10%)	366 (90%)	0
FVB/*Ddx1*$^{+/-}$	292	34 (12%)	258 (88%)	0

from heterozygote intercrosses was due to a recessive lethal allele linked to the wild-type *Ddx1* loci in the C57BL/6 background, *Ddx1*$^{+/-}$ mice were backcrossed to FVB mice for six generations. After 6 generations, >99% of the genome will be derived from the FVB background. At this point, we initiated heterozygote intercrosses with FVB/*Ddx1*$^{+/-}$ mice. We observed a genotype ratio of 1:9 from an initial count of 40 weaned pups. A total of 292 weaned pups from FVB intercrossed heterozygotes were ultimately genotyped, generating a ratio of 1:7.3 wild-type to heterozygote pups. As the apparent lethal effect of the *Ddx1*$^{+}$ allele was independent of background strain, we concluded that wild-type lethality in intercrossed heterozygote mice must be related to the *Ddx1*^{S} allele itself.

4.5 Identification of two distinct populations of *Ddx1*$^{+/-}$ mice

To examine the distribution of *Ddx1*$^{+/+}$ progeny per litter, we plotted the proportion of wild-type progeny per litter for all litters that contained at least 5 pups (Figure 4.6A). An unequal bimodal distribution was observed, suggesting that two distinct groups were present within the data set. The majority of litters generated no *Ddx1*$^{+/+}$ progeny; however, a second peak was observed with the expected proportion of wild-type to heterozygote mice (approximately one third wild-type). Given that *Ddx1*$^{+/-}$ mice are genetically indistinguishable, the only difference between the different crosses is parental genotype.

Distinct genotypic distributions were observed when heterozygote intercrosses were separated into the two following groups, one where the parental cross was a

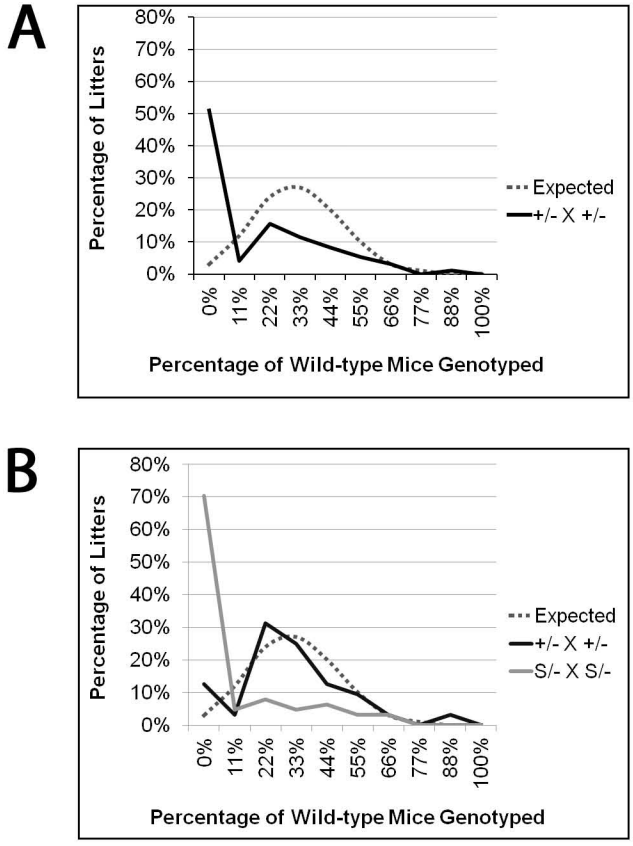


Figure 4.6 Bimodal distribution of wild-type progeny.

(A) All litters of heterozygote intercrossed mice that contained at least 5 pups were examined for the percentage of wild-type mice generated (n= 178). A normal distribution plotted around the expected value of 33% wild-type is included for comparison. (B) *Ddx1*$^{+/-}$ and *Ddx1*$^{S/-}$ intercrosses were separated (n= 32 and n= 146 respectively) and the percentage of wild-type mice generated was plotted.

heterozygous intercross (+/- X +/-) and the other where the parental cross was a backcross (WT x +/-). *Ddx1*$^{+/-}$ mice generated from the latter parental crosses (WT X +/-) generated the expected proportion of wild-type progeny (~33%) (Figure 4.6B). In contrast, *Ddx1*$^{+/-}$ mice generated from parental heterozygous intercrosses produced significantly reduced numbers of wild-type progeny (~5%). Thus, heterozygote mice which had inherited their "wild-type" allele from a wild-type parent produced the expected ratio of wild-type to heterozygote progeny. However, mice which had inherited their "wild-type" allele from heterozygote parents produced very few wild-type progeny. As the lethal effect was associated with "+" alleles inherited from heterozygous parents, we designated *Ddx1*$^{+}$ alleles inherited from heterozygous mice as *Ddx1*^{S} (S for super allele) in order to differentiate them from *Ddx1*+ alleles. While *Ddx1*$^{+}$ and *Ddx1*^{s} alleles are genetically identical, the lack of wild-type progeny associated with *Ddx1*^{s} suggests that this allele has been marked in some way, perhaps by epigenetic modification.

4.6 *Ddx1*^{S} associated lethality

Our data indicate that the homozygous *Ddx1*^{S} state results in embryonic lethality. Next, we examined whether the presence of a single *Ddx1*^{S} allele could also affect viability. To address this possibility, we analysed the genotype of all crosses between *Ddx1* heterozygotes and wild-type mice. Similar to heterozygote crosses, a reduction in wild-type progeny was observed in heterozygote X wild-type backcrosses. Genotyping of 562 progeny from *Ddx1*$^{+/-}$ backcrosses to wild-type FVB or

wild-type C57BL/6 mice resulted in the expected ratio of approximately 1 wild-type to 1 heterozygote (Figure 4.7). Notably, genotyping 295 progeny from *Ddx1*$^{S/-}$ backcrosses resulted in a ratio of approximately 1 wild-type to 3 heterozygotes, indicating reduced viability of *Ddx1*$^{S/+}$ mice. This effect is similar, albeit less severe, than that observed in *Ddx1*$^{S/S}$ mice (approximately 1 wild-type to 19 heterozygotes). These data lend further support to the idea that *Ddx1*$^{+}$ and *Ddx1*^{S} alleles can be distinguished from one another, and that the *Ddx1*^{S} allele is the cause of lethality in wild-type embryos. By comparing expected and actual numbers of wild-type progeny, we have determined that approximately 95% of *Ddx1*$^{S/S}$ and 50% of *Ddx1*$^{S/+}$ mice die.

Next, we analyzed the ratio of wild-type to heterozygote progeny from *Ddx1*$^{S/-}$ intercrosses at early developmental stages. At E3.5, a slight, but non-significant, reduction in the number of wild-type progeny was observed compared to *Ddx1*$^{+/-}$ intercrosses (Figure 4.8). After implantation (E6.5), a statistically significant reduction in wild-type progeny was observed, with *Ddx1*$^{S/-}$ intercrosses generating approximately 5% wild-type progeny. As this percentage is similar to the percentage of wild-type pups recovered, we propose that the lethal effect of *Ddx1*^{S} occurs after E3.5, and before E6.5. Thus, *Ddx1*$^{-/-}$ mice die at the 2- or 4-cell stage, whereas *Ddx1*$^{S/S}$ mice die in the post-blastocyst stage around the time of implantation suggesting two different mechanisms for *Ddx1*-mediated death of mouse embryos.

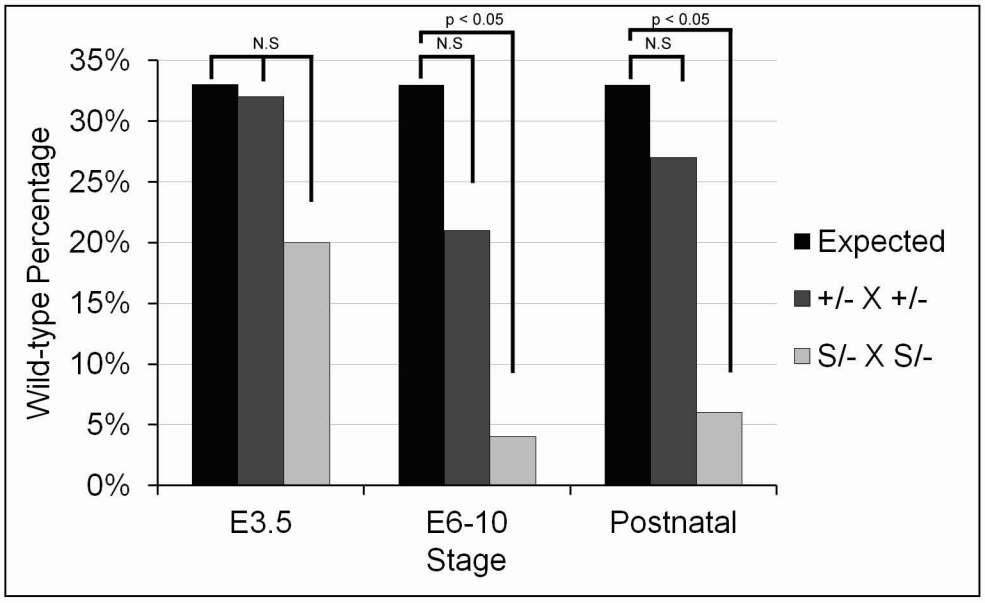


Figure 4.7 Percentages of wild-type progeny were decreased in *Ddx1*$^{S/-}$ intercrosses compared to *Ddx1*$^{+/-}$ intercrosses.

The percentage of wild-type mice at ages E3.5, E6-10 and P0 from *Ddx1*$^{+/-}$ (n= 22, 61, 229, respectively) or *Ddx1*$^{S/-}$ (n= 69, 164, 529, respectively) intercrosses were plotted against the expected percentage. Fischer exact tests were performed to determine significant differences between the expected and experimental groups.

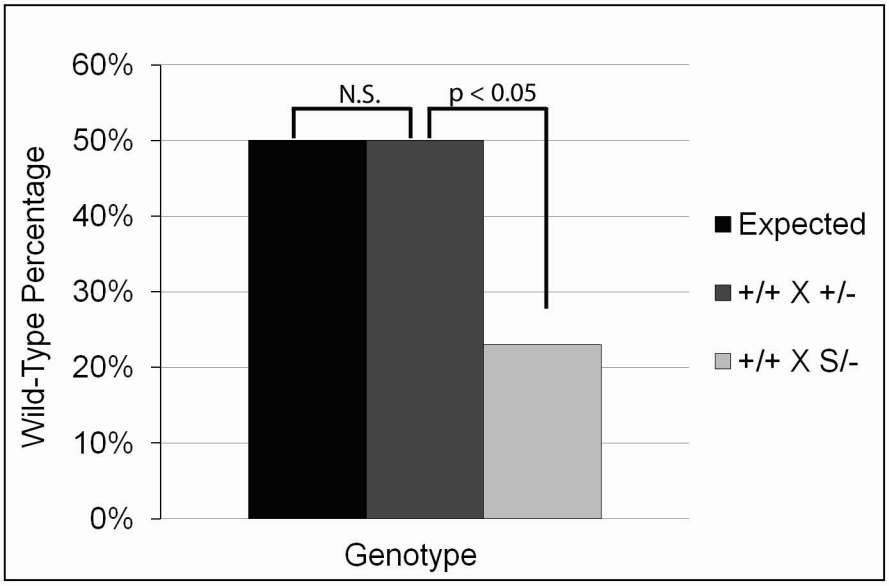


Figure 4.8 Wild-type lethality in progeny derived from backcrosses.

Percentages of wild-type progeny derived from either *Ddx1*$^{S/-}$ or *Ddx1*$^{+/-}$ backcrossed were plotted (n= 562 and 295, respectively). The expected percentage (50%) is shown. The Fischer exact test was performed to determine significant differences between the expected and experimental groups.

4.7 DDX1 dose compensation in heterozygote mice

Previous work has suggested that DDX1 levels may be tightly controlled in cells. For example, multiple attempts to generate *Ddx1* transgenic mice that overexpress DDX1 have been unsuccessful in spite of generating more than 12 lines of mice carrying multiple copies of the *Ddx1* gene (85). It is therefore possible that compensatory mechanisms come into play upon mating heterozygote mice to ensure that overall DDX1 protein levels are maintained in developing $Ddx1^{+/-}$ mouse embryos. Increased production of *Ddx1* transcripts from the "+" allele of heterozygote animals may therefore cause the conversion of this allele from "+" to "S".

To address the possibility of a compensatory mechanism underlying the observed wild-type lethality, we carried out western blot analysis of newborn mouse brain tissue. Similar levels of DDX1 protein were observed in all brain tissue tested whether the brain was obtained from $Ddx1^{+/+}$ or $Ddx1^{+/-}$ mice (Figure 4.9A). Semi-quantitative RT-PCR analysis of a region of *Ddx1* unique to the wild-type allele showed comparable levels of mRNA in wild-type and heterozygote mice (Figure 4.9B). β-gal expression from the mutant allele was also confirmed, thus establishing bi-allelic expression of the *Ddx1* locus. These results are in agreement with the existence of a compensatory mechanism to ensure that sufficient *Ddx1* transcripts are produced when one of the two *Ddx1* alleles is mutated.

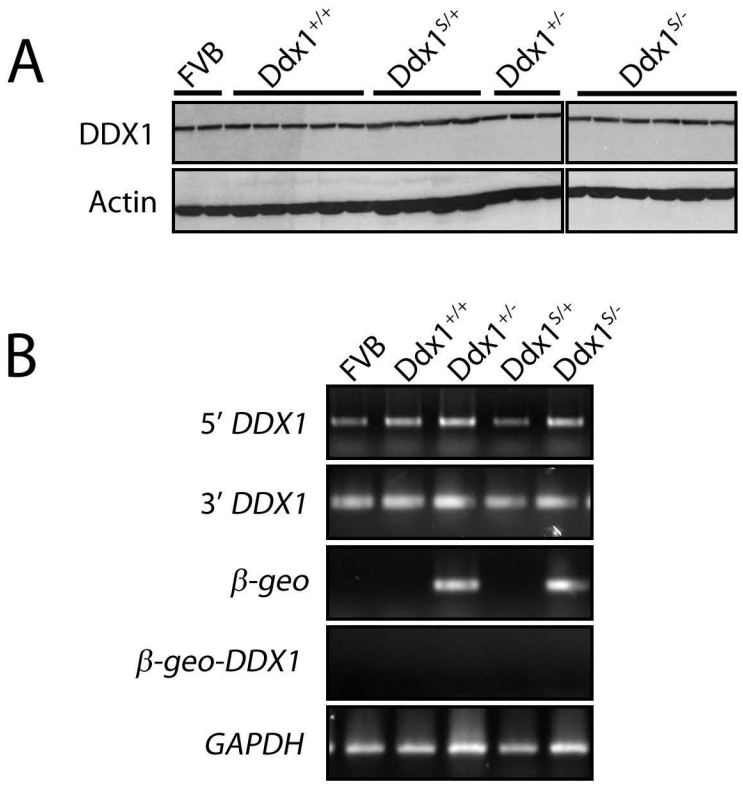


Figure 4.9 DDX1 mRNA and protein expression in mouse brains of different genotypes of *Ddx1* mice.

(A) Fifty μg of whole brain lysates from P0-3 mice electrophoresed in a 10% SDS-polyacrylamide gel. Proteins were transferred to a nitrocellulose membrane and immunostained with anti-DDX1 antibody (top) and anti-actin antibody (bottom). (B) Semi-quantitative RT-PCR was performed on cDNAs generated from P0-3 mouse brain RNA collected from mice of different *Ddx1* genotypes. Equal amounts of cDNA were amplified with primers 5′ to the gene-trap, 3′ to the gene-trap, specific to the insertion, spanning the 3′ end of the insertion into the 3′ region of *DDX1*, and to *GAPDH* as a control for the amound of cDNA. PCR products were electrophoresed in a 1% agarose gel.

4.8 Epigenetic inheritance of the *Ddx1*^{S} allele

Multi-generational matings of *Ddx1*^{S} allele-carrying mice suggest that *Ddx1* transcript levels can be modified from one generation to the next in order to compensate for the absence of *Ddx1* transcripts from one of the *Ddx1* alleles. Genomic DNA methylation is a common way of regulating gene expression epigenetically. With the exception of genomic imprinting, all forms of genomic DNA methylation are removed at the zygotic stage of development. Genomic imprinting allows propagation of genomic DNA methylation from one generation to the next. Genomic imprinting can be maternal or paternal, with the methylation pattern of one allele affecting the other allele. This process occurs early in development as the imprinted alleles maintain their methylation patterns despite the global demethylation that occurs during the pre-implantation stages of development. If genomic imprinting is responsible for generating the *Ddx1*^{S} allele, then the lethal effect should only be observed when the allele is either maternally or paternally inherited.

When backcrossed to wild-type mice, male and female *Ddx1*$^{+/-}$ mice generate normal proportions of wild-type progeny (Figure 4.10). In contrast, both male and female *Ddx1*$^{S/-}$ mice generate a significantly lower proportion of wild-type progeny when backcrossed to wild-type mice (~50% reduction from the total expected). As this effect is observed whether the "S" allele is maternally or paternally inherited, genomic imprinting is not a plausible explanation for the modification of the *Ddx1* allele. Thus, our results indicate that: (i) the *Ddx1*^{S} allele can be inherited (Figure 4.11A), (ii) the *Ddx1*^{S} allele is associated with increased *Ddx1* transcription or transcript stability, (iii)

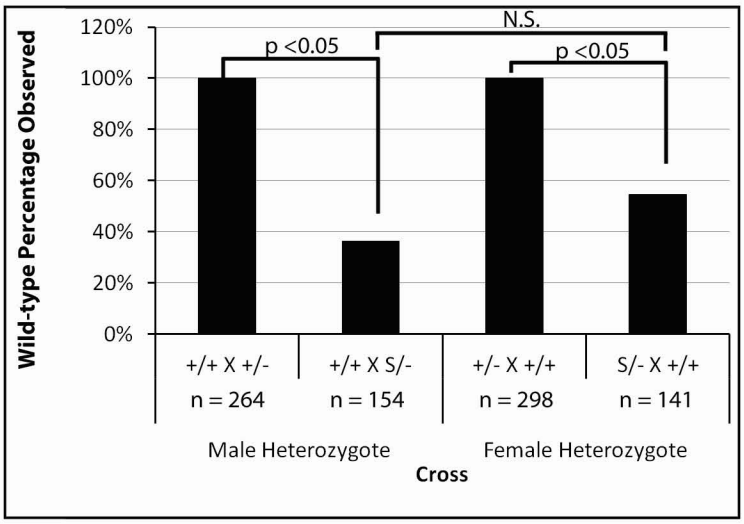


Figure 4.10 Decreased percentage of wild-type progeny generated from backcrosses of male and female *Ddx1*$^{S/-}$ mice.

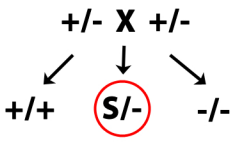
Backcrosses from both FVB and C57BL/6 mice were separated by genotype and gender of the heterozygote. The percentage of wild-type mice was normalized to the *Ddx1*$^{+/-}$ backcross. Fischer exact tests were used to determine significance.

Figure 4.11. Inheritance model of the *Ddx1*^{S} allele.

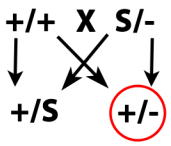
(A) There are two types of wild-type alleles determined by parental crosses. (B) When *Ddx1*$^{S/-}$ mice are intercrossed they produce an abnormal ratio of wild-type to heterozygote progeny, whereas the *Ddx1*$^{+/-}$ intercrosses produced the expected ratio. (C) This effect is also partially observed in backcrosses using *Ddx1*$^{S/-}$ mice, where the *Ddx1*$^{+/-}$ mice produce the expected ratio of wild-type to heterozygote. (D) We propose that in *Ddx1*$^{+/+}$ mice each allele contributes to the total level of DDX1. The wild-type allele (+) undergoes an unknown epigenetic change when both parents are heterozygotes which results in twice the normal levels of DDX1 (S = super allele). The increased dosage to ~3x from 2x results in partial lethality of the *Ddx1*$^{S/+}$ mice. This lethality is more apparent in the *Ddx1*$^{S/S}$ mice which are proposed to produce ~4x rather than 2x DDX1.

A

Distinguishing the two wild type alleles "+" vs "S"



"S" allele derived from + allele inherited from a +/- (or S/-) mouse



"+" allele is inherited from a +/+ mouse

B

Heterozygous intercrosses

Parents	S/- X S/-			+/- X +/-		
Progeny	S/S	S/-	-/-	+/+	S/-	-/-
Expected ratio	1	2	0*	1	2	0*
Observed ratio	1	9	0*	1	2	0*

* -/- lethality

C

Homozygous X heterozygous backcrosses

Parents	+/+ X S/-		+/+ X +/-	
Progeny	S/+	S/-	+/+	+/-
Expected ratio	1	1	1	1
Observed ratio	1	3	1	1

D

Proposed mechanism

	+/+	S/+	S/S
Expression level relative to normal conditons	1x + 1x	~2x + 1x	~2x + ~2x
Total DDX1	2x	~3x	~4x
Observed lethality	None	Partial	High

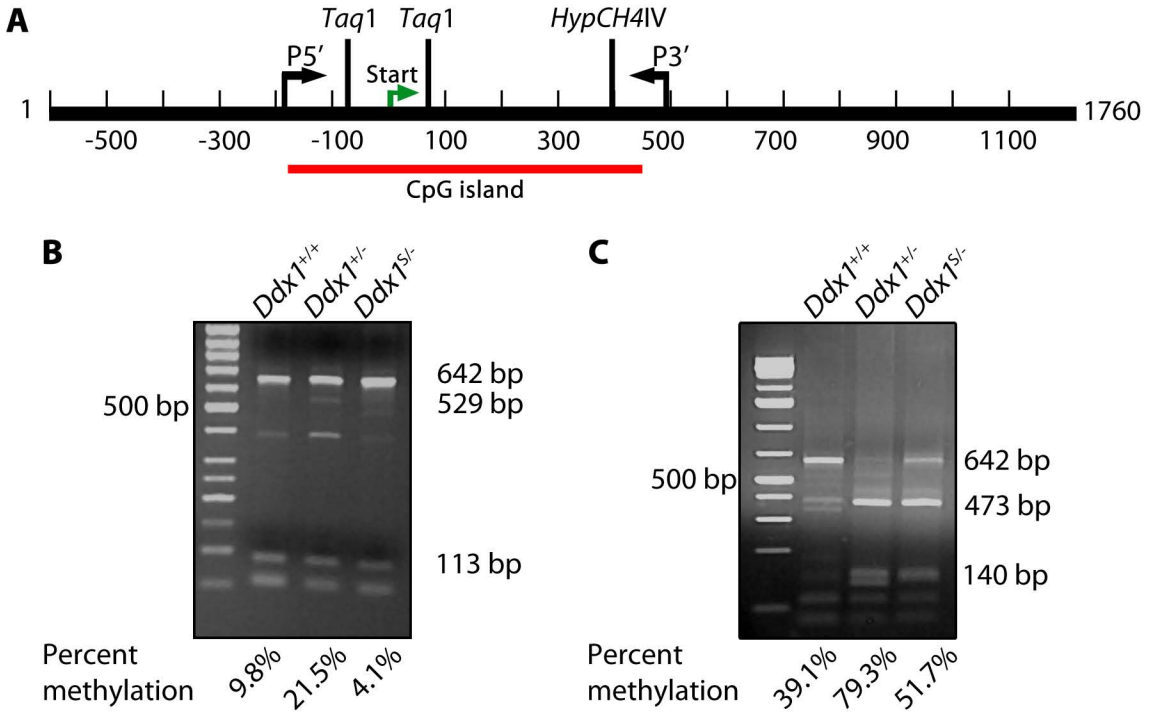
both the maternal and the paternal *Ddx1* loci can be modified, and (iv) *Ddx1* loci modification is not through genomic imprinting.

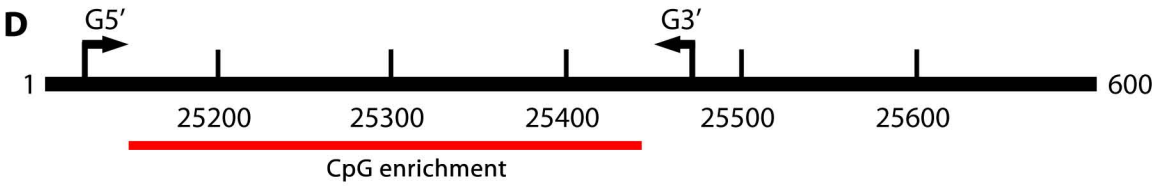
Importantly, our results indicate that the observed $Ddx1^{S}$ effect is cis-acting. The observed cis inheritance is most likely epigenetic in nature since $Ddx1^{S/-}$ mice generate $Ddx1^{+/-}$ progeny when backcrossed to wild-type mice, as demonstrated by normal ratios of wild-type mice observed from first generation heterozygous (Figure 4.11C) mice. If the effect was trans in nature, we would predict a similar effect to that observed in the *c-kit* mouse line where a miRNA induces a paramutation which is propagated by only one type of heterozygote mice (see Discussion) (165). In the case of *Ddx1*, we clearly have two different heterozygotes as determined by their ability to generate viable wild-type mice. These combined observations all support cis-acting epigenetic modification of the *Ddx1* gene.

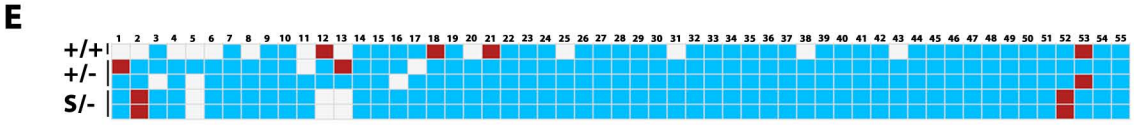
Although genomic imprinting as an explanation for the inheritance of the $Ddx1^{S}$ allele has been ruled out, and all other DNA methylation marks are believed to be removed at the zygotic stage, we still pursued the investigation of DNA methylation pattern changes between the different *Ddx1* genotypes. We examined two regions of the *Ddx1* loci predicted to contain CpG islands: the promoter region (Figure 4.12A) and a region in the gene body towards the 3' end of the gene (Figure 4.12D). In collaboration with Dr. Igor Kovalchuk (Department of Biology, University of Lethbridge), we used a digestion based method (COBRA) to detect differences in the methylation status of the *Ddx1* promoter region in $Ddx1^{+/+}$, $Ddx1^{+/-}$, and $Ddx1^{S/-}$ mice (232). Our initial results suggested that there was some difference between the three

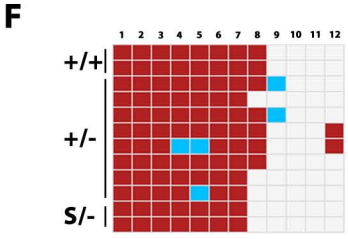
Figure 4.12. Methylation analysis of *Ddx1*.

(A) *Ddx1* has a CpG island containing 55 potential sites of modification from ~-200 to 450 bp spanning the transcription start site (241). Genomic DNA prepared from *Ddx1*$^{+/+}$, *Ddx1*$^{+/-}$ and *Ddx1*$^{S/-}$ mice (in triplicate) was treated with bisulfite, amplified using the indicated primers, and digested with either *HpyCH4*IV (B) or *Taq*I (C). The digested DNA was electrophoresed in a 1% agarose gel and the ratio between the digested DNA and the total amplified DNA yielded the approximate percent methylation. (D) *Ddx1* is predicted to have a CpG-enriched region containing 12 potential sites of modification from ~25150 to 25430 bp. The CpG island spanning the transcription start site (E) and the CpG-enriched region in the gene body (F) were bisulfite sequenced. The white squares indicate no data, the blue squares indicate no methylation and the red squares indicate sites of methylation.


A
Taq1
Taq1
HypCH4IV
P5′
Start
P3′
1
1760
-500
-300
-100
100
300
500
700
900
1100
CpG island
B
Ddx1+/+
Ddx1+/-
Ddx1S/-
500 bp
642 bp
529 bp
113 bp
Percent methylation
9.8%
21.5%
4.1%
C
Ddx1+/+
Ddx1+/-
Ddx1S/-
500 bp
642 bp
473 bp
140 bp
Percent methylation
39.1%
79.3%
51.7%


D
G5′
G3′
1
600
25200
25300
25400
25500
25600
CpG enrichment


E
+/+
+/-
S/-


F
+/+
+/-
S/-

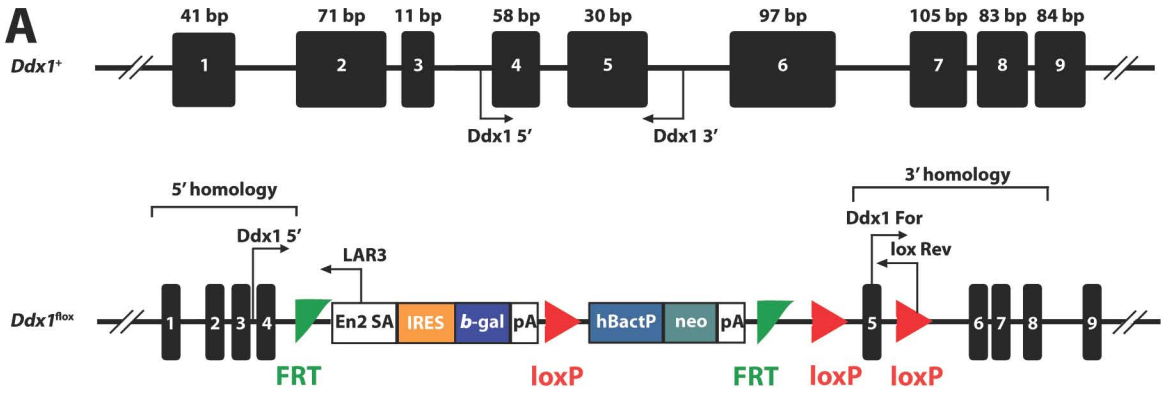
different genotypes (Figure 4.12B,C). The methylation of the *Ddx1* promoter was found to be similar for *Ddx1*$^{+/+}$ and *Ddx1*$^{S/-}$ mice; however, increased methylation was observed in the promoter of *Ddx1*$^{+/-}$ mice compared to *Ddx1*$^{+/+}$ and *Ddx1*$^{S/-}$ mice. We also used a direct bisulphite sequencing method to determine the methylation status of the *Ddx1* promoter region as well as an identified CpG rich region in the gene body. Genomic DNA from all three genotypes (*Ddx1*$^{+/+}$, *Ddx1*$^{+/-}$ and *Ddx1*$^{S/-}$) were analysed. Unlike the previous digestion-based analysis, the sequenced bisulphite converted DNA showed no difference between any of the samples (Figure 4.12E, F). The promoter region was generally unmethylated and the gene body was completely methylated. Although preliminary, these data indicate that it's unlikely that 5-methylcytosine is the epigenetic mark responsible for the conversion of *Ddx1*$^{+}$ to *Ddx1*^{S}. Other modifications may be responsible for the observed effect on *Ddx1*, such as 5-hydroxymethylcytosine. Alternatively, the epigenetic mark may be through a different mechanism altogether.

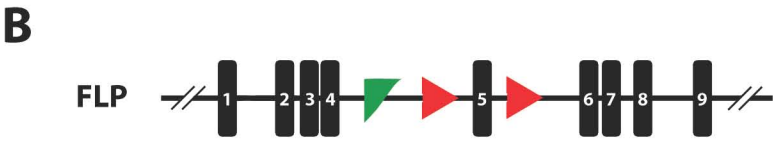
4.9 Attempts to generate a conditional *Ddx1* knockout mouse

As *Ddx1* is essential for early development thereby precluding analysis of DDX1 function, we decided to generate a conditional *Ddx1* knockout mouse to study its role at later stages of development. Three embryonic stem cell lines (B06, B07 and F05) that had been tested with gene-specific primers to ensure proper targeting of *Ddx1* were obtained from EUCOMM (Figure 4.13A). Exon 5 was the targeted exon in all three

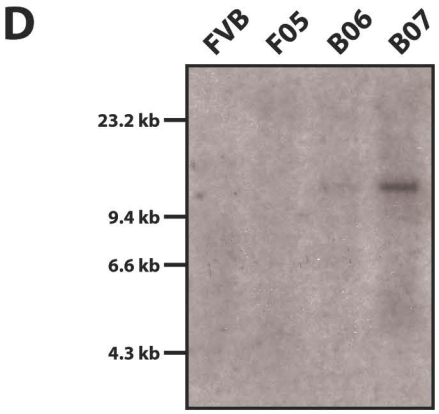
Figure 4.13. Strategy for generating conditional *Ddx1* knock-out mice.

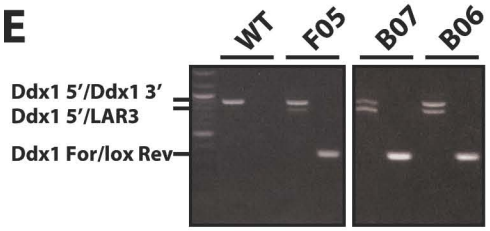
(A) Embryonic stem cells purchased from EUCOMM contain a *Ddx1* allele with *loxP* sites flanking exon 5. The targeted allele also contained *β-gal* and *neo* as reporters. In this state, the targeted allele would generate a protein consisting of DDX1 (exons 1-4) fused to *β-GAL* due to the presence of the *En2* splice acceptor. (B) The constitutive knockout could be made conditional by crossing $Ddx1^{+/flox}$ mice to FLP recombinase-expressing mice to remove the two reporters. (C) Crossing $Ddx1^{flox/fox}$ to mice expressing Cre-recombinase will delete exon 5 and as a result, cause a frame shift in *Ddx1*. (D) Southern blot analysis of a wild-type mouse and the three targeted embryonic stem cell lines. Genomic DNA extracted from FVB mice, B06, B07 and F05 cell pellets was digested with *Xho*I and separated in a 1% agarose gel. The blot was hybridized with a ^{32}P-labeled cDNA probe specific to β-gal. (E) PCR amplification protocol used for routine genotyping of genomic DNA. Primer locations are indicated in (A). (F) PCR amplification of genomic DNA extracted from a chimeric mouse and a second collection of cell pellets from the three targeted cell lines.

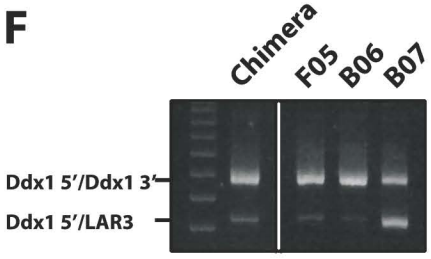

A
Ddx1+
41 bp
71 bp
11 bp
58 bp
30 bp
97 bp
105 bp
83 bp
84 bp
1
2
3
4
5
6
7
8
9
Ddx1 5′
Ddx1 3′
5′ homology
3′ homology
Ddx1 5′
LAR3
Ddx1 For
lox Rev
Ddx1flox
En2 SA
IRES
b-gal
pA
hBactP
neo
pA
FRT
loxP
FRT
loxP
loxP


B
FLP
1
2
3
4
5
6
7
8
9


C
Cre
1
2
3
4
6
7
8
9


D
FVB
F05
B06
B07
23.2 kb
9.4 kb
6.6 kb
4.3 kb


E
WT
F05
B07
B06
Ddx1 5′/Ddx1 3′
Ddx1 5′/LAR3
Ddx1 For/lox Rev


F
Chimera
F05
B06
B07
Ddx1 5′/Ddx1 3′
Ddx1 5′/LAR3

lines, with loxP sites flanking this exon and FRP sites flanking the reporter located immediately upstream of exon 5. By mating *Ddx1*$^{+/flox}$ mice with FLP recombinase-expressing mice, we would remove the reporter and generate mice ready for conditional knockout (Figure 4.13B). The final *Ddx1*$^{flox/flox}$ mice could then be crossed to Cre recombinase-expressing mice for tissue-specific removal of exon 5 resulting in a frameshift mutation which should effectively knock out *Ddx1* in the tissue of interest (Figure 4.13C).

B06, B07 and F05 cells were sent to the University of Calgary's Centre for Mouse Genomics for generation of *Ddx1* knockout mice. Here, the three clones were expanded and karyotyped prior to microinjection. Cell lines B06, B07, and F05 were 86%, 95%, and 81% diploid, respectively. Lysates of each clone were collected from 12-well dishes and genomic DNA was prepared for PCR and Southern blot analysis (Figure 4.13D). A strong band was observed in the B07 line using a probe specific to *β-gal*, indicating that the cells contained the floxed allele. Weaker bands were observed in F05 and B06, suggesting that not all cells had the targeted allele. PCR analysis confirmed the Southern blot results with the presence of the reporter and *loxP* sequence detected using either a multiplex reaction that contained primers to both *Ddx1* and the *β-gal* reporter, or a single reaction where the *Ddx1-loxP* sequence was amplified (Figure 4.13E).

Two methods were used to generate germ line transmission of the targeted *Ddx1* allele: (i) all three embryonic stem cell lines were microinjected into albino C57BL/6 hosts; and (ii) CD1 morula were combined with ES cells to generate

aggregates. From the microinjection experiments, we obtained 8 chimeric males and 1 chimeric female from lines B06 and F05, with varying degrees of coat chimerism. No chimeric animals were obtained from the B07 line. From the aggregated embryos, we obtained 9 chimera males and no females (5 from F05, 1 from B06 and 3 from B07). Chimeric mice were mated to C57BL/6 mice to generate first generation heterozygotes. Tail samples were taken from these mice and genomic DNA was isolated and screened by PCR for the presence of the targeted allele. Unfortunately, none of the 62 genotyped mice produced heterozygote offspring indicating that none of the chimeras had germline transmission of the targeted allele.

To address the reason for the absence of germline transmission, fresh lysates were prepared from the F05, B06 and B07 cell lines and again screened by PCR. Surprisingly, there was a significant decrease in the targeted allele band intensity in cell lines B06 and F05 compared to the previous PCR results (Figures 4.13F). The intensity of the band representing the targeted allele remained the same in B07.

Although both the B06 and F05 cell lines generated chimeras, none of the chimeras produced offspring with the targeted allele. One possible explanation for the lack of germ line transmission in chimeras produced from the B06 and F05 lines is that these lines consisted of a mixture of cells, only a small proportion of which had the targeted allele. The B07 cell line generated three chimeras which were obtained using the aggregation technique; however, these chimeras also failed to produce offspring.

Chapter 5

Subcellular localization of DDX1

DDX1 has been shown to play various roles in somatic cells, including DNA repair and RNA transport (112,124). Previous analysis of DDX1 expression patterns in tissues has been primarily conducted by *in situ* hybridization (98). In chicken neuroectodermal tissues such as the retina and brain, *DDX1* RNA was found to be highly expressed throughout these tissues early in development (98). As the tissues matured, the expression pattern became more restricted to specific cell types, some of which were still undergoing proliferation at E16, the latest stage analysed.

Although *in situ* hybridization is very useful for determining which cells express *DDX1* RNA, it provides no information regarding subcellular localization of the protein. In most cell lines, DDX1 localizes primarily to the nucleus (101,102). Its localization to the cytoplasm has been observed in breast cancer as well as in *DDX1*-amplified and overexpressing neuroblastoma and retinoblastoma cell lines (85,96). In breast cancer, the presence of elevated levels of DDX1 in the cytoplasm has been shown to be a negative prognostic factor (96). As shown in the previous chapter, DDX1 is essential for development pre-implantation as *Ddx1*$^{-/-}$ embryos die at the 2- or 4-cell stage. The role that DDX1 is playing at this stage of development is unknown and may be associated with its RNA transport function. The localization of DDX1 in adult and embryonic tissues may shed light on the different roles that DDX1 plays in the cell.

5.1 DDX1 localization in mouse tissues

We collected heart, liver, small intestine, kidney, ovary, testis, brain and retina tissues from mice. Tissues were fixed with 4% formalin and paraffin-embedded. The

sectioned tissues were labeled with anti-DDX1 antibody. DDX1 was expressed in all tissues tested although not in all the cells of those tissues. Approximately half of the nuclei of the cardiac cells were positive for DDX1 (Figure 5.1 A, B). DDX1 was also found in the cytoplasm of a subset of cardiac cells where it had a filamentous staining pattern (arrowheads). DDX1 was expressed in the majority of kidney cells (Figure 5.1 C/D) but was found in the nucleus of only a subset of cells in the glomerulus (arrows). The tubules of the kidney had both nuclear and cytoplasmic DDX1. The proximal convoluted tubules (PCT) had higher levels of DDX1, especially in the cytoplasm, compared to the distal convoluted tubules (DCT) which showed a primarily nuclear staining pattern (arrowheads). The staining of DDX1 in the liver was consistent throughout the tissue with high nuclear localization in the hepatocytes (Figure 5.1 E/F). The cytoplasmic staining of DDX1 in the liver was granular in appearance. In the small intestine, DDX1 was expressed in most cells with highest levels in the epithelium cells of the villi (arrows) (Figure 5.1 G/H). The staining of DDX1 was also stronger in the nucleus of cells found in the crypts (arrowheads) as compared to the outer columnar epithelial cells (arrows). Similar to the cardiac muscle, DDX1 was found at low levels in the smooth muscle lining the villi (asterisk) and only present in some of the muscle cell nuclei.

DDX1 was differentially expressed in the various cell types of the retina. At P1, DDX1 was most highly expressed in ganglion cells, with lower levels in the amacrine cell layer (Figure 5.2). A similar expression pattern was observed at P15, although fewer amacrine cells were positive for DDX1. In adult retina, the ganglion cells were

Figure 5.1 DDX1 expression in adult mouse tissues.

Formalin fixed paraffin-embedded sections of adult C57BL/6 mouse heart, kidney, liver and small intestine were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. Images were taken using an Axioskop 2 Plus microscope with 20X/0.75 and 40X/1.3 lenses. The final images were processed using Adobe CS6. Arrowheads in B indicate the DDX1 filamentous staining. Arrows in D indicate the cells with nuclear staining of DDX1 in the glomerulus. The arrowheads identify two different tubules in the kidney, the PCT and the DCT. The arrowheads in F indicate the granular staining that was present in the liver hepatocytes. The arrows in H indicate the predominantly nuclear staining of DDX1 in the columnar epithelia of the villi. The arrowheads indicate the stronger nuclear and cytoplasmic staining of DDX1 present in the crypt cells. The asterisks indicate cells of the smooth muscle with DDX1 staining.

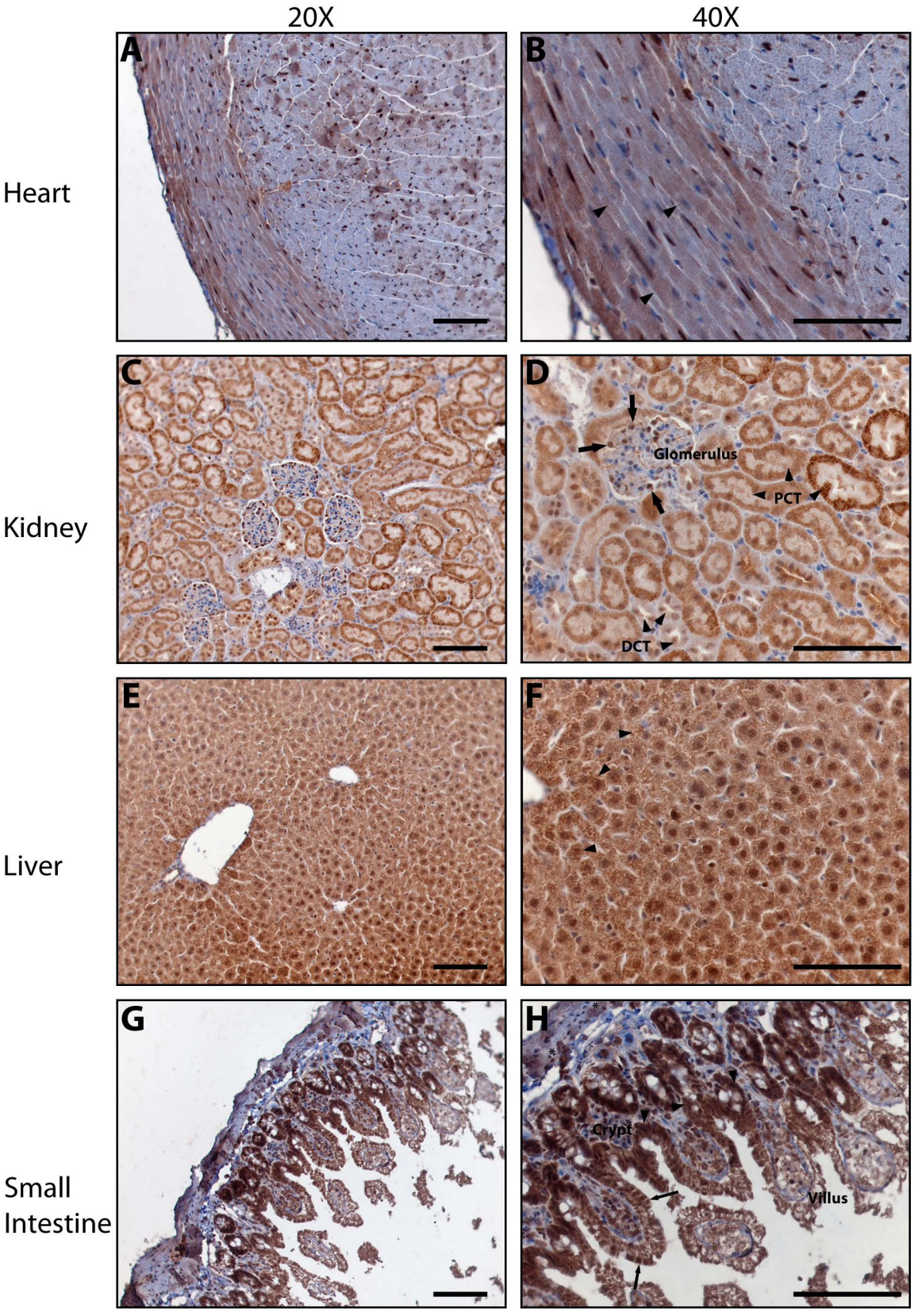
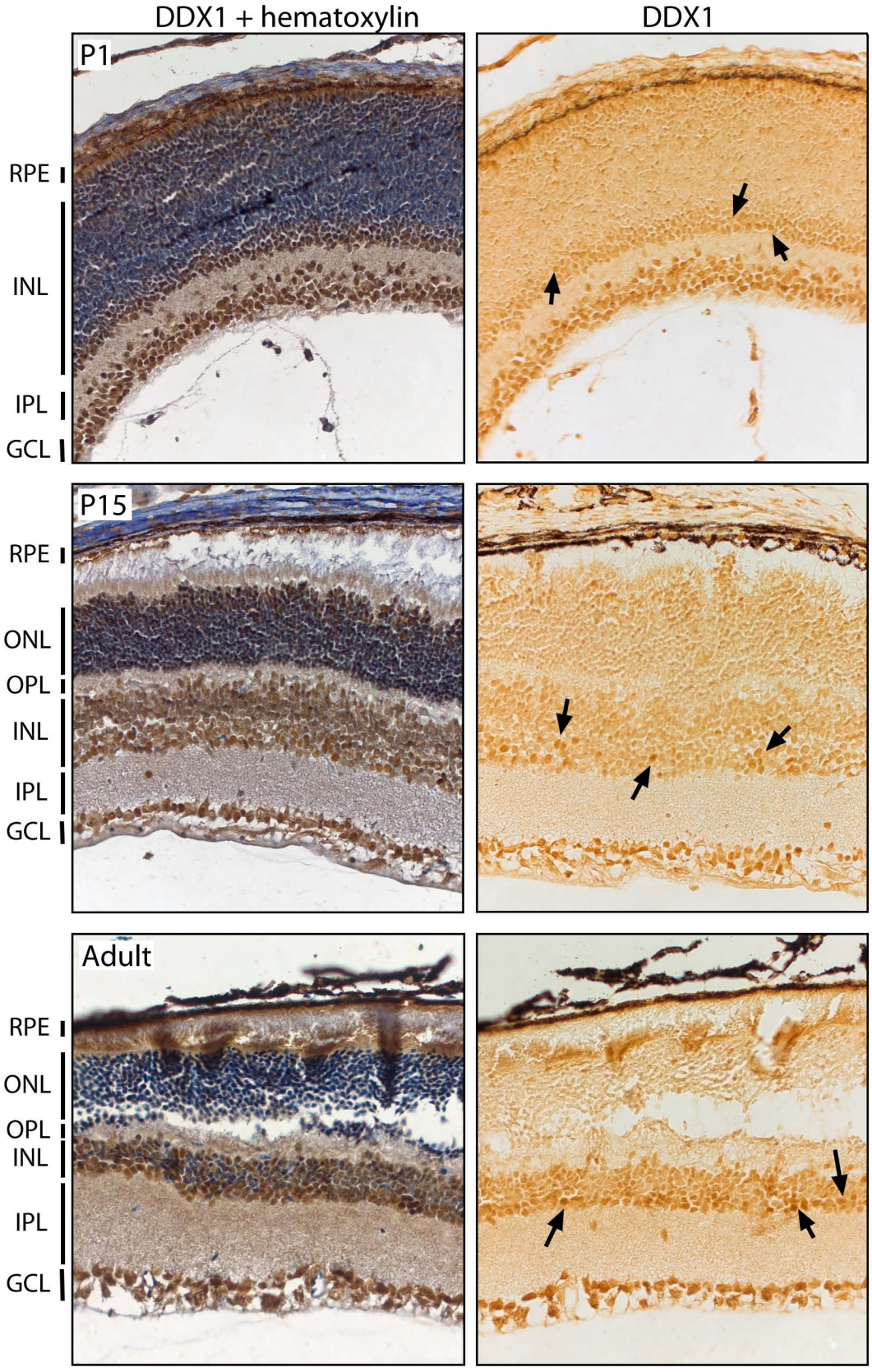

20X
40X
Heart
Kidney
Liver
Small Intestine
A
B
C
D
E
F
G
H
Glomerulus
PCT
DCT
Crypt
Villus

Figure 5.2 DDX1 expression patterns in mouse retina.

Formalin-fixed paraffin-embedded sections of P0, P15 and adult C57BL/6 mouse eyes were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. One slide from each stage of development was counterstained with hematoxylin (left panels). Images were taken using an Axioskop 2 Plus microscope with a 20X/0.75 lens. The final images were processed using Adobe CS6. The positions of the different layers of the retina are indicated. The arrows point to a subset of amacrine cells with higher expression of DDX1 in the inner nuclear layer.


DDX1 + hematoxylin
DDX1
P1
RPE
INL
IPL
GCL
P15
RPE
ONL
OPL
INL
IPL
GCL
Adult
RPE
ONL
OPL
INL
IPL
GCL

still strongly positive for DDX1, along with one layer of amacrine cells located closest to the inner plexiform layer. DDX1 was primarily found in the nucleus of both ganglion and amacrine cells.

We also carried out a more detailed analysis of DDX1 expression in developing mouse brain. At P0, DDX1 was found in most cells of the brain and localized primarily to the nucleus (Figure 5.3). DDX1 levels were higher in the olfactory tubercle and at the dorsal surface of the midbrain compared to most other regions of the brain. Considerable variation in DDX1 expression was observed in the P0 cerebellum: DDX1 was barely detectable in the granule cell progenitors (arrowheads), with higher levels of DDX1 observed in the adjacent cell layer which contains Purkinje cells and Bergmann glial cells (arrows). In the adult brain, DDX1 was again found in most cells and localized to the nucleus (Figure 5.4). DDX1 was highly expressed in the thalamus (Figure 5.4 C, D), olfactory bulb (Figure 5.4 E, F), and cerebellum (Figure 5.4 G, H). In the adult cerebellum, DDX1 was expressed in the granule cell layer in addition to the cell types previously shown to express DDX1 in P0 cerebellum.

The last adult tissues examined were the gonads. In the testis, DDX1 was found in both the developing spermatids and the supporting Sertoli cells (Figure 5.5 A, B). In the ovary, DDX1 was predominantly found in the developing oocytes and the surrounding granulosa cells (Figure 5.5 C, D). The stromal cells between the oocytes generally had low to no expression of DDX1. The cells of the corpus luteum were positive for nuclear DDX1 with weak diffuse cytoplasmic staining. All oocytes,

Figure 5.3 DDX1 expression in neonatal mouse brain.

Formalin-fixed paraffin-embedded sections of P0 mouse brain were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. Multiple images spanning the entire brain were taken using an Axioskop 2 Plus microscope with a 10X/0.3 lens. The multiple images were merged using Adobe CS6. (A) Tissue section immunostained with anti-DDX1 antibody. (B) Tissue section immunostained with anti-DDX1 antibody and counterstained with hematoxylin to detect the nuclei. (C, D) The olfactory tubercle has elevated levels of DDX1 next to the ventral edge (arrows). Some cells in this region have low levels of DDX1 (arrowheads). (E, F) Most of the cells in the dorsal mid-brain also show elevated expression of DDX1. (G, H) The Purkinje cells of the developing cerebellum are positive for DDX1 (arrows), with little or no DDX1 detected in the granule cell progenitors (arrowheads).

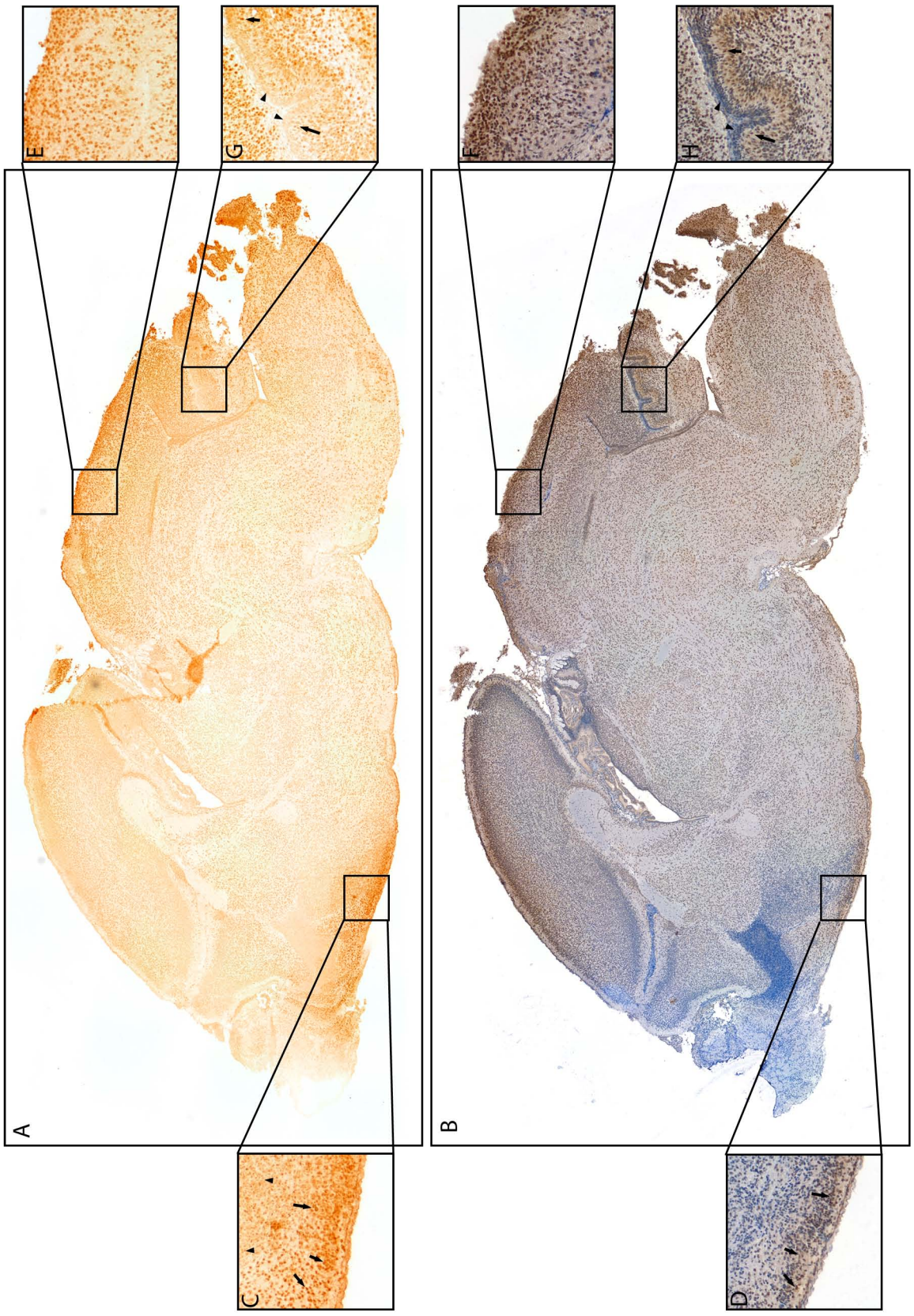
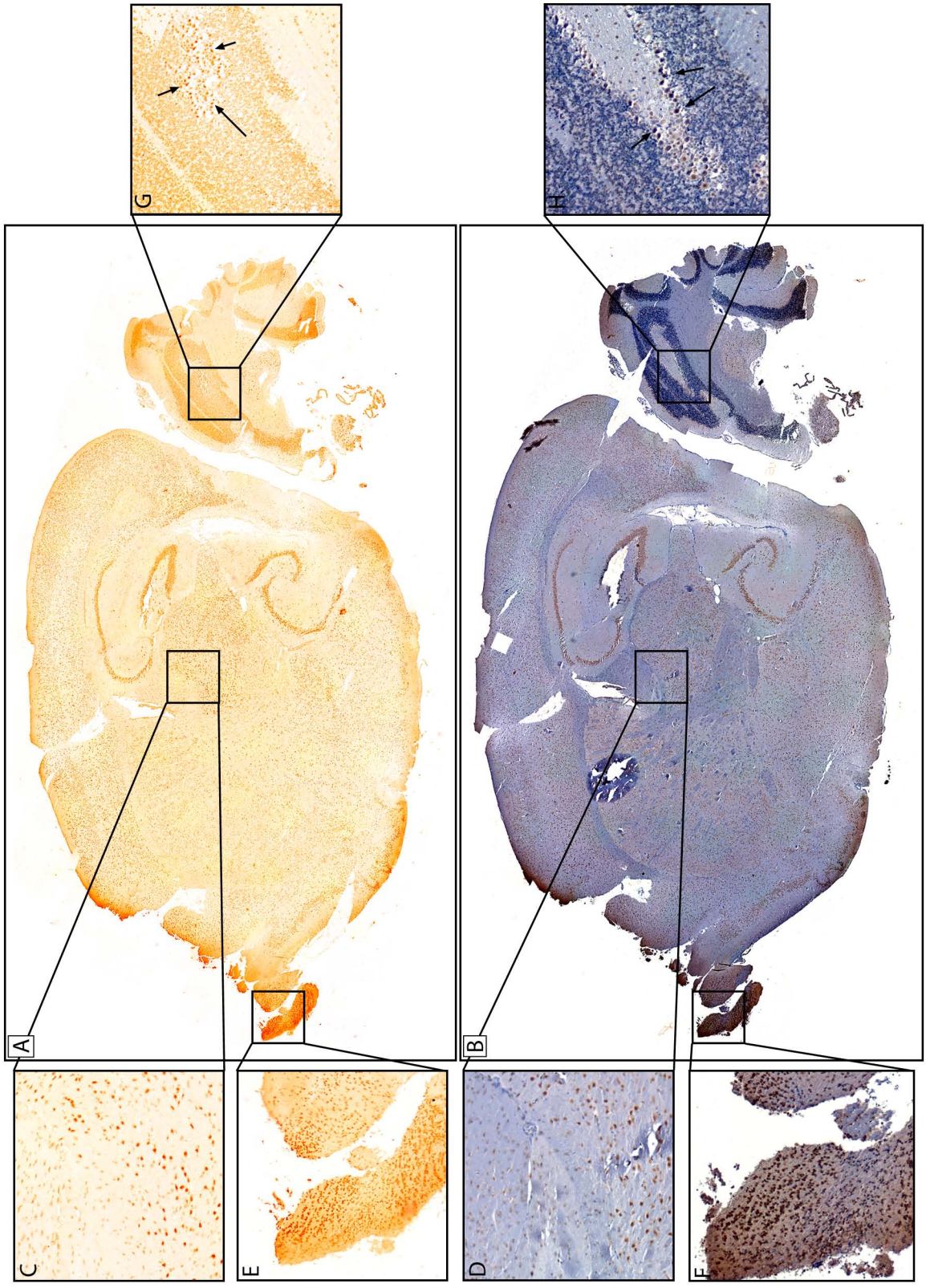

A
B
C
D
E
F
G
H

Figure 5.4 DDX1 expression in adult mouse brain.

Formalin-fixed paraffin-embedded sections of adult mouse brain were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. One slide from each stage of development was counterstained with hematoxylin (bottom panel). Multiple images spanning the entire brain were taken using an Axioskop 2 Plus microscope with a 10X/0.3 lens. The multiple images were merged using Adobe CS6. (A) Tissue sections were immunostained with anti-DDX1 antibody. (B) Tissue sections were immunostained with anti-DDX1 antibody and counterstained with hematoxylin to detect the nuclei. High levels of DDX1 were found in the thalamus (C, D) and olfactory bulb (E, F). The cerebellum was also strongly positive for DDX1, particularly the Purkinje cells (arrows) (G, H).


G
H
A
B
C
E
D
F

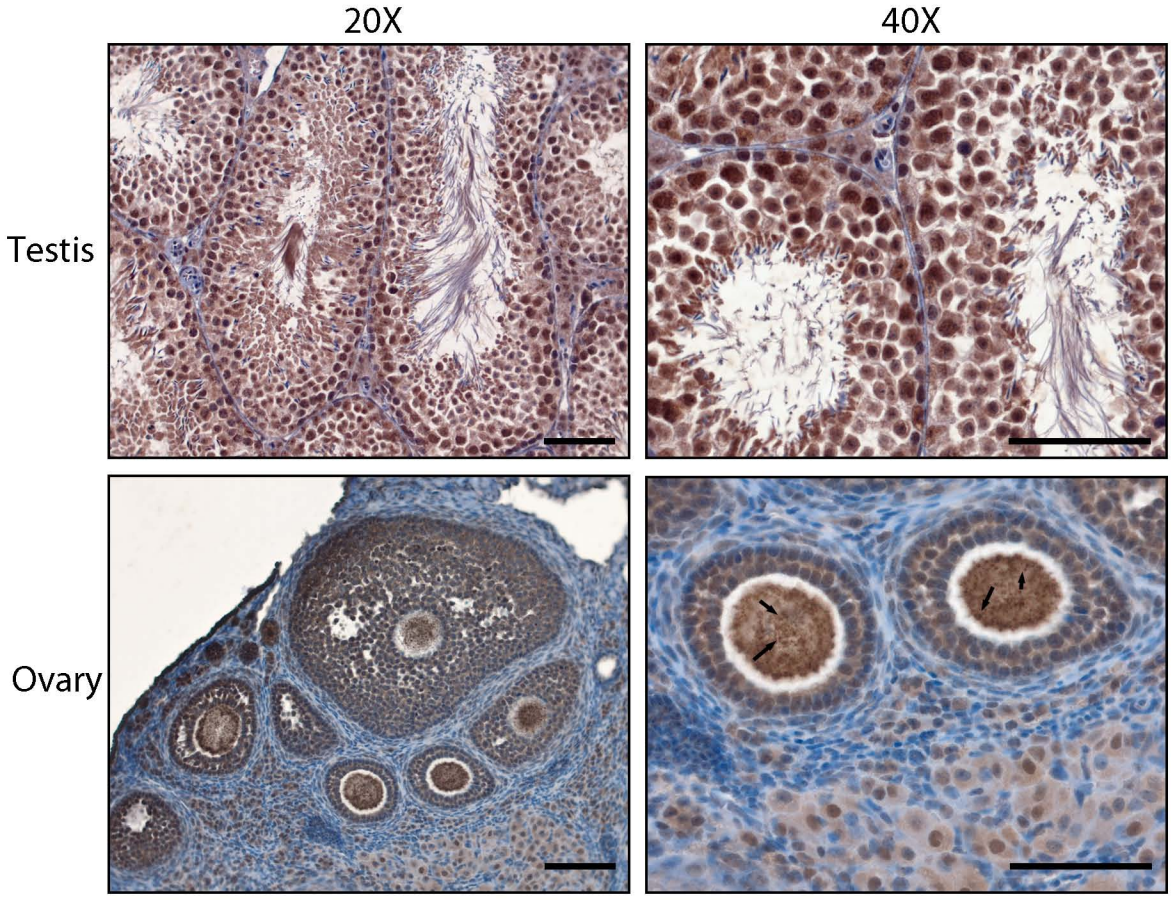


Figure 5.5 DDX1 expression in mouse ovary and testis.

Formalin-fixed paraffin-embedded sections of adult mouse ovaries and testes were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. Tissues were counterstained with haematoxylin. Images were captured using an Axioskop 2 Plus microscope with 20X/0.75 and 40X/1.3 lenses. Arrows point to DDX1 granules present in the developing oocyte. Scale bars represent 200 μm.

regardless of maturation stage, showed strong DDX1 cytoplasmic staining. Intriguingly, DDX1 formed large granules in the cytoplasm, suggesting significant aggregation of DDX1 in specific regions of the cytoplasm. These granules appeared quite different from those found in the other adult tissues such as the liver.

Our data indicate that DDX1 is highly expressed during development, with some differential staining in the post-natal tissue compared to adult. Next, we examined DDX1 expression in whole embryos at E6.5. Embryos were fixed in 4% formalin and mounted in paraffin or OCT. Sectioning was along the sagittal plane. DDX1 was present at low levels in the cells of the maternal decidua (Figure 5.6, see arrowheads). Red blood cells (indicated by asterisks) present around the embryo (caused by invasion of the implanting embryo) were negative for DDX1 expression. In comparison to the maternal decidua, the embryo proper had much higher levels of DDX1. However, unlike adult tissues, DDX1 was localized primarily to the cytoplasm as opposed to the nucleus. Nuclear DDX1 was observed in a few cells located in the extraembryonic ectoderm which will develop into the amniotic sac (marked by circles). Some of the trophoblast cells of the ectoplacental cone had large DDX1 granules in their cytoplasm (marked by arrows). These cells are highly invasive and are required for embryonic implantation into the decidua.

In conclusion, DDX1 was highly expressed in cells undergoing proliferation, such as those found in early development or in adult small intestine. DDX1 localized to the nucleus of most cells, although some cells also showed cytoplasmic staining. Some

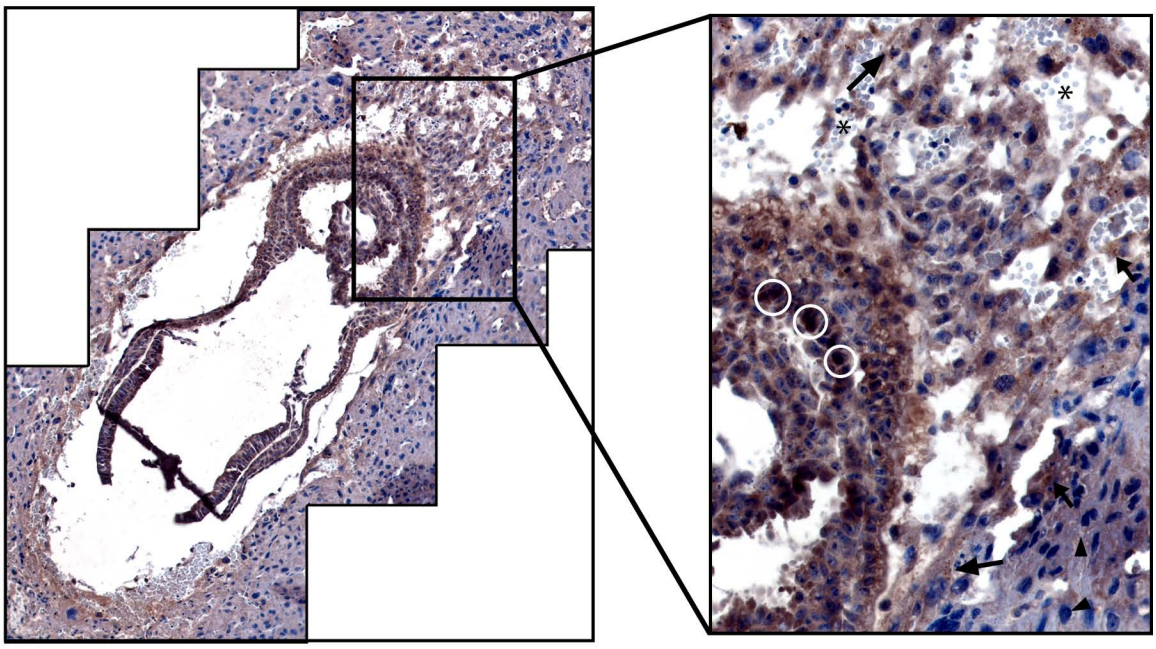


Figure 5.6 DDX1 expression in E6.5 embryos.

OCT-embedded sections of E6.5 decidua were subjected to immunohistochemical staining with anti-DDX1 (2910) antibody. Multiple images spanning the entire embryo were taken using an Axioskop 2 Plus microsope and a 20X/0.75 lens. The multiple images were merged using Adobe CS6. Cells of the maternal decidua had low levels of DDX1 expression (arrowheads). Trophoblast cells had large DDX1-positive granules in the cytoplasm (arrows). Some of the ectoplacental cone cells showed strong nuclear localization of DDX1 (circles).

tissues had no or low levels of DDX1 including the stromal cells of the ovary, the decidua, some muscle cell nuclei, endothelial cells, red blood cells, and some of the glomerulus cells. The early embryo was unique in expressing DDX1 primarily in the cytoplasm with little to no nuclear DDX1.

5.2 DDX1 in pre-implantation embryos

Immunostaining of E6.5 embryos with anti-DDX1 antibody demonstrated a predominantly cytoplasmic pattern, suggesting different roles for DDX1 at early developmental stages compared to later stages of development and mature cells. As indicated in Chapter 4, *Ddx1*$^{-/-}$ and *Ddx1*$^{S/S}$ embryos die pre-implantation. To address the DDX1 staining patterns in pre-implantation embryos, we first carried out immunofluorescence analysis of E3.5 blastocysts. These embryos were immunostained with three different anti-DDX1 antibodies (batch 2910, 2923 and 2290) in order to compare consistency of staining patterns using different antibodies. Similar to the staining patterns observed in oocytes and E6.5 embryos, all three antibodies produced a predominantly cytoplasmic signal (Figure 5.7). Anti-DDX1 antibodies from batches 2910 and 2923 both generated similar staining patterns although the signal obtained with the 2910 antibody was much cleaner with low background staining. DDX1 was found in large granules or aggregates throughout the cytoplasm of the E3.5 blastocysts (Figure 5.7).

In contrast to the 2910 and 2923 antibodies, the 2290 antibody (generated using a 3' region of DDX1 that contains most of the conserved motifs found in DEAD

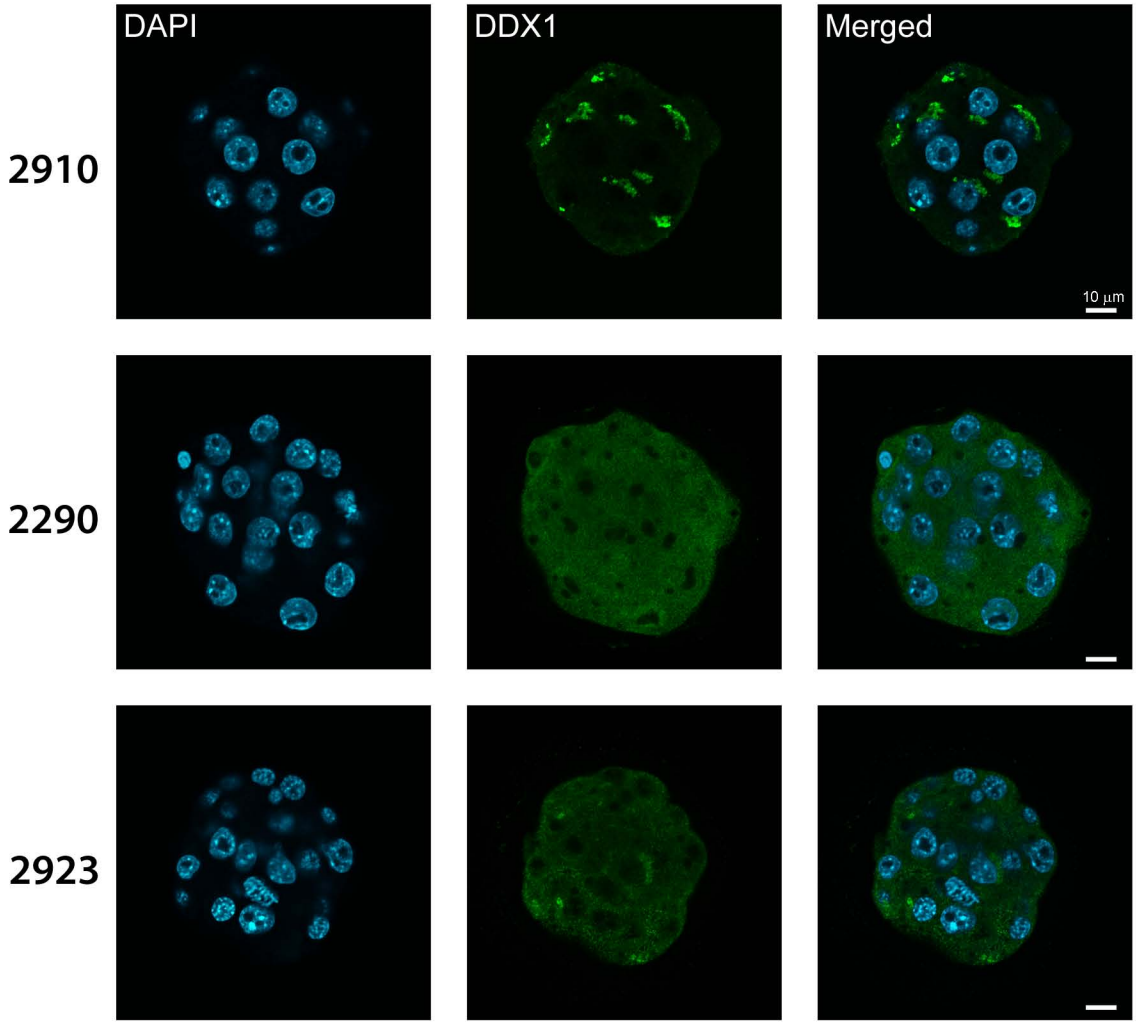


Figure 5.7 DDX1 staining in blastocysts with different DDX1 antibodies.

E3.5 stage embryos were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were then immunostained with different DDX1 antibodies: 2910 (generated using denatured N-terminal DDX1 (aa 1-186) as antigen), 2290 (generated using native C-terminal DDX1 (aa 187-740) as antigen), and 2923 (generated using native N-terminal DDX1 (aa 1-186) as antigen). Nuclei were visualized with DAPI. Embryos were imaged as single optical sections by confocal microscopy. Scale bars = 10 μm.

box proteins) produced a diffuse staining pattern throughout the blastomeres with some patches of stronger cytoplasmic or nuclear staining. As both 2910 and 2923 anti-DDX1 antibodies have been shown to specifically recognize DDX1 (101), all subsequent immunostaining was done using the 2910 antibody as it produced a strong signal with low background.

The granular appearance of DDX1 in the cytoplasm of blastocyts was different from all previously documented DDX1 immunostaining. In particular, the staining pattern of cytoplasmic DDX1 in other cell types shown to have elevated levels of DDX1 in the cytoplasm (e.g. neuronal cells, breast cancer cells and *DDX1*-amplifed retinoblastoma and neuroblastoma cells (85,112)), was much more diffuse than that observed in blastocysts. To determine whether the large DDX1 aggregates were specific to E3.5 blastocysts, we collected 1-, 2-, 4-, 8-cell, and blastocyst stage embryos and labeled them with 2910 anti-DDX1 antibody. Dramatic differences in staining patterns were observed at these different stages of development (Figure 5.8). Multiple images of each embryo were collected using laser scanning confocal microscopy to generate a Z-stack. These 3D images were then flattened using a maximum intensity projection algorithm in Zen (Zeiss). The resulting 2D image has all the layers stacked in order to show the staining throughout the embryo as opposed to a single optical slice. Using Imaris software the size of the DDX1 foci/aggregates was calculated using surface rendering (Figure 5.9). 1-cell embryos underwent an additional step of deconvolution prior to the calculations to reduce background noise.

Figure 5.8 DDX1 localization in pre-implantation embryos.

Embryos at from 1-cell to blastocyst stages were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were then immunostained with anti-DDX1 antibody (2910). Nuclei were stained with DAPI. Embryos were displayed as 2D projections of Z-stacks imaged by confocal microscopy. Scale bars = 20 μm.

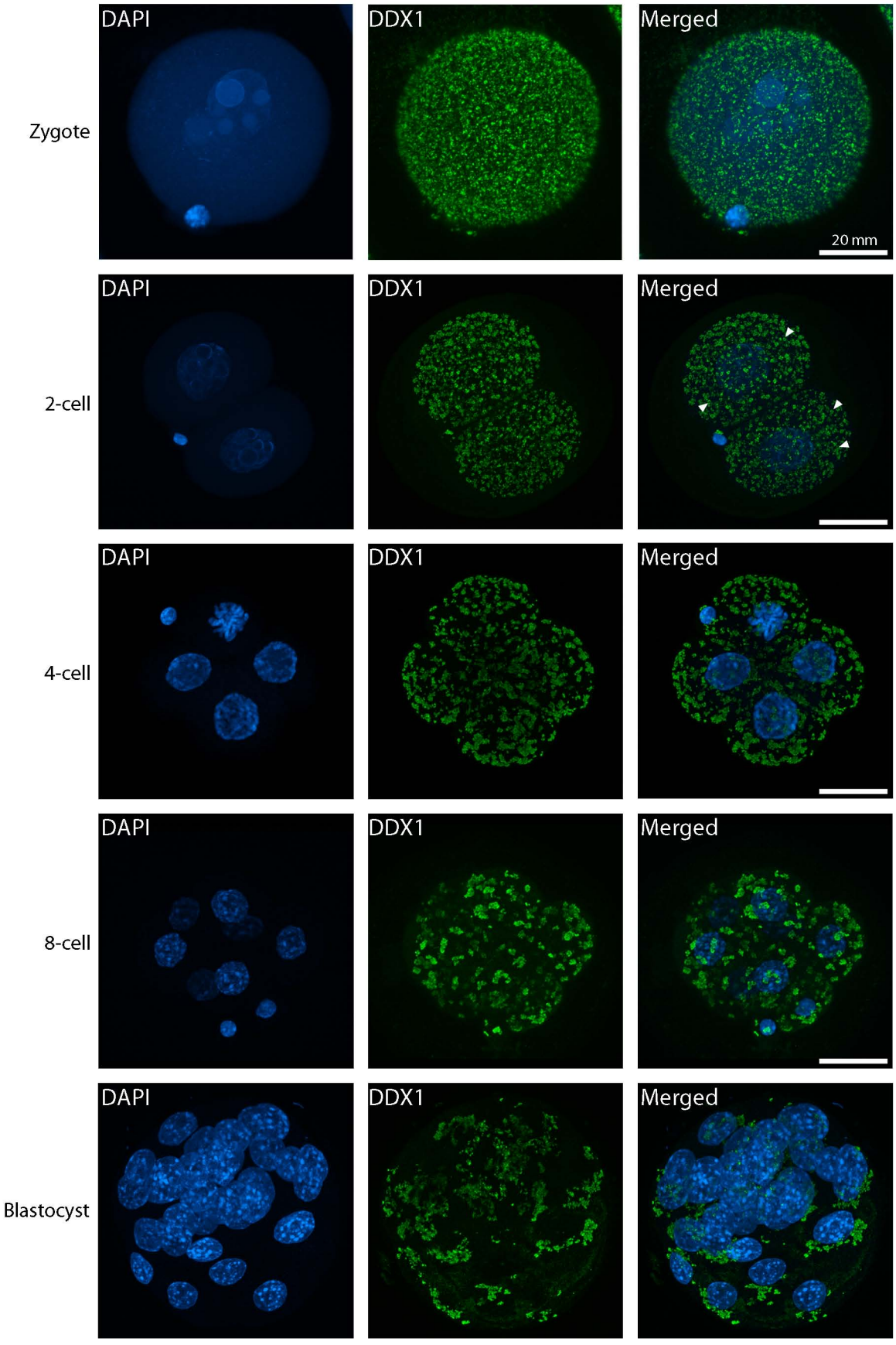
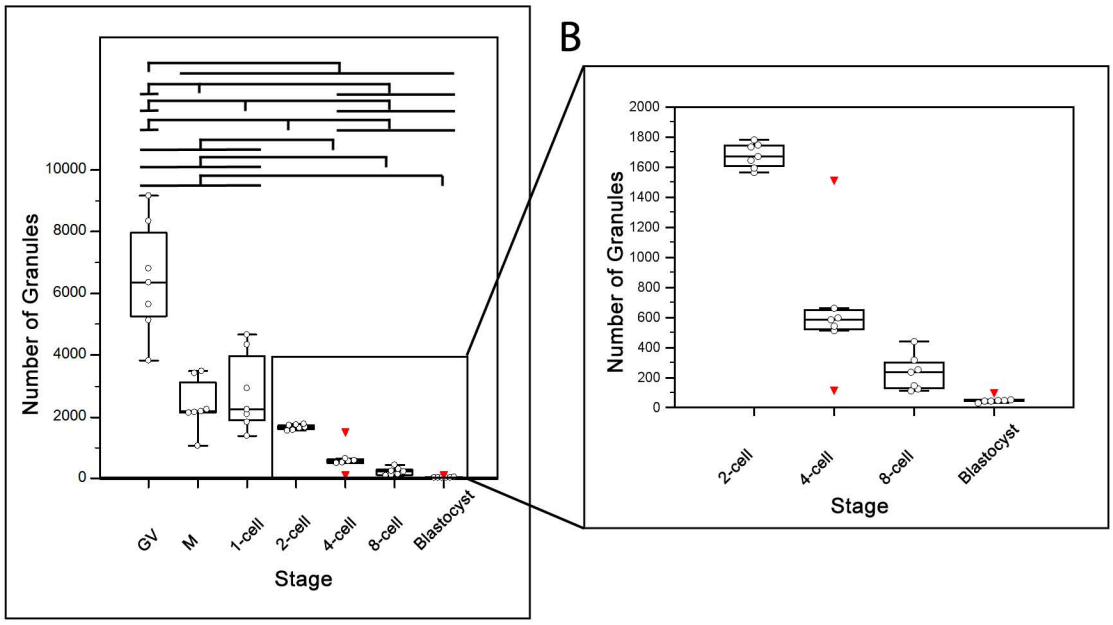

DAPI
DDX1
Merged
Zygote
20 mm
DAPI
DDX1
Merged
2-cell
DAPI
DDX1
Merged
4-cell
DAPI
DDX1
Merged
8-cell
DAPI
DDX1
Merged
Blastocyst

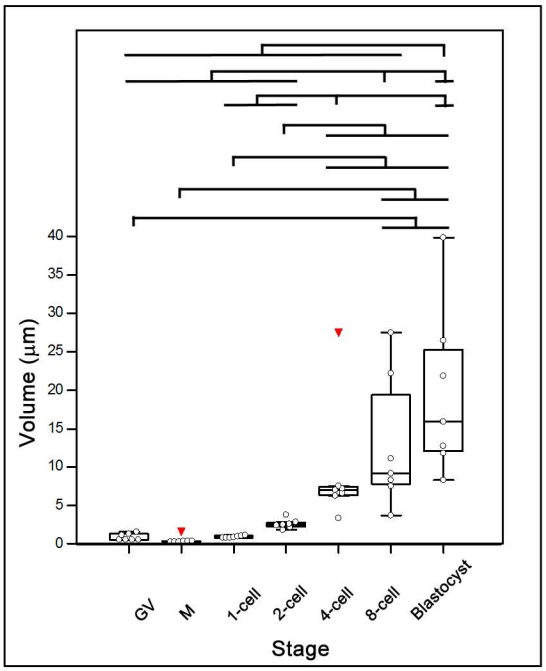
Figure 5.9 Analysis of DDX1 aggregates in pre-implantation and oocyte development.

Surface rendering by Imaris software was used to calculate the number (A, B) and volume of DDX1 aggregates (C) at the different stages of oocyte and preimplantation development. A total of 8 embryos or oocytes were analyzed at each developmental stage. Statistical differences were calculated by one-way ANOVA and Student-Newman-Keuls was used as a post-hoc test to determine significant differences between the groups. Significant differences ($p < 0.05$) between the different groups are shown by the overlying lines. Statistical outliers are indicated by the red triangles.

A


Number of Granules
10000
8000
6000
4000
2000
0
GV
M
1-cell
2-cell
4-cell
8-cell
Blastocyst
Stage
B
Number of Granules
2000
1800
1600
1400
1200
1000
800
600
400
200
0
2-cell
4-cell
8-cell
Blastocyst
Stage

C


Volume (µm)
40
35
30
25
20
15
10
5
0
GV
M
1-cell
2-cell
4-cell
8-cell
Blastocyst
Stage

In the 1-cell embryo, DDX1 was found in numerous (~3200) foci (0.9 μm^3) throughout the cytoplasm of the embryo. Single optical slices showed the complete absence of DDX1 from the nucleus. By the 2-cell stage, these small foci had transitioned into larger rounded aggregates which were evenly distributed throughout the cytoplasm (Figure 5.8). The average 2-cell embryo had ~1700 aggregates which were ~2.5 μm^3 in size (Figure 5.9). Although the aggregates were found throughout the embryo, there were more prevalent near the cell membrane. The 4-cell stage embryos had on average larger aggregates than 2-cell embryos (~7 μm^3 compared to ~2.5) but there were far fewer aggregates (~600 compared to ~1700) per embryo. The aggregates had a more clumped appearance at the 4-cell stage. There was one outlier in the 4-cell embryo sample group, but this may have been an early 4-cell embryo in which the small DDX1 aggregates were still transitioning to larger aggregates.

By the 8-cell stage, the aggregates had again increased in size to ~10 μm^3 although there was significant variation within the group. By this stage of development, there were only ~200 aggregates per embryo. The largest aggregates which were also the most variable in size were observed at the blastocyst stage. At this developmental stage, DDX1 aggregates averaged 15 μm^3 in size but some were as large as 60 μm^3. The number of aggregates was very low (~15) in E3.5 embryos. While DDX1 was not detected in every blastomere, in those blastomeres that did have DDX1 aggregates, the aggregates were found next to the nucleus. Their localization next to the nucleus appeared to be random and not associated with any cell polarity. There

was no pattern to the location of the blastomeres that lacked DDX1 staining, with no preferential association with either trophoblast cells or the inner cell mass.

5.3 Subcellular localization of DDX1 during oogenesis

DDX1 is expressed in developing oocytes as well as single cell embryos indicating that DDX1 is part of the maternally derived complement of proteins as embryonic translation does not occur until the 2-cell stage in mouse. To further investigate DDX1 protein localization during oocyte development, we collected germinal vesicle (GV) to meiosis I (MI) stage oocytes from ovaries obtained from mature C57BL/6 and FVB females. To examine MI oocytes (i.e. embryos undergoing germinal vesicle breakdown), GV oocytes were denuded of the surrounding granulosa cells and cultured for 24-48 hours. Embryos were staged as MI if they lacked a polar body but the chromosomes were aligned. MII oocytes were obtained from superovulated females in collaboration with Dr. Heather McDermid (Department of Biological Sciences, University of Alberta).

Oocytes were labeled with anti-DDX1 (2910) antibody. Z-stacks were collected using laser scanning confocal microscopy to generate 3D images of the oocytes (Figure 5.10). DDX1 was dispersed throughout the GV stage oocytes and formed small round foci averaging 1.2 μm^3 in size (Figure 5.9). There were a few instances of larger aggregates throughout the cytoplasm but these appeared different from the aggregates observed at the 2-cell stage. Approximately 7000 foci were observed in GV oocytes. Based on the examination of single optical sections, DDX1 is absent from the

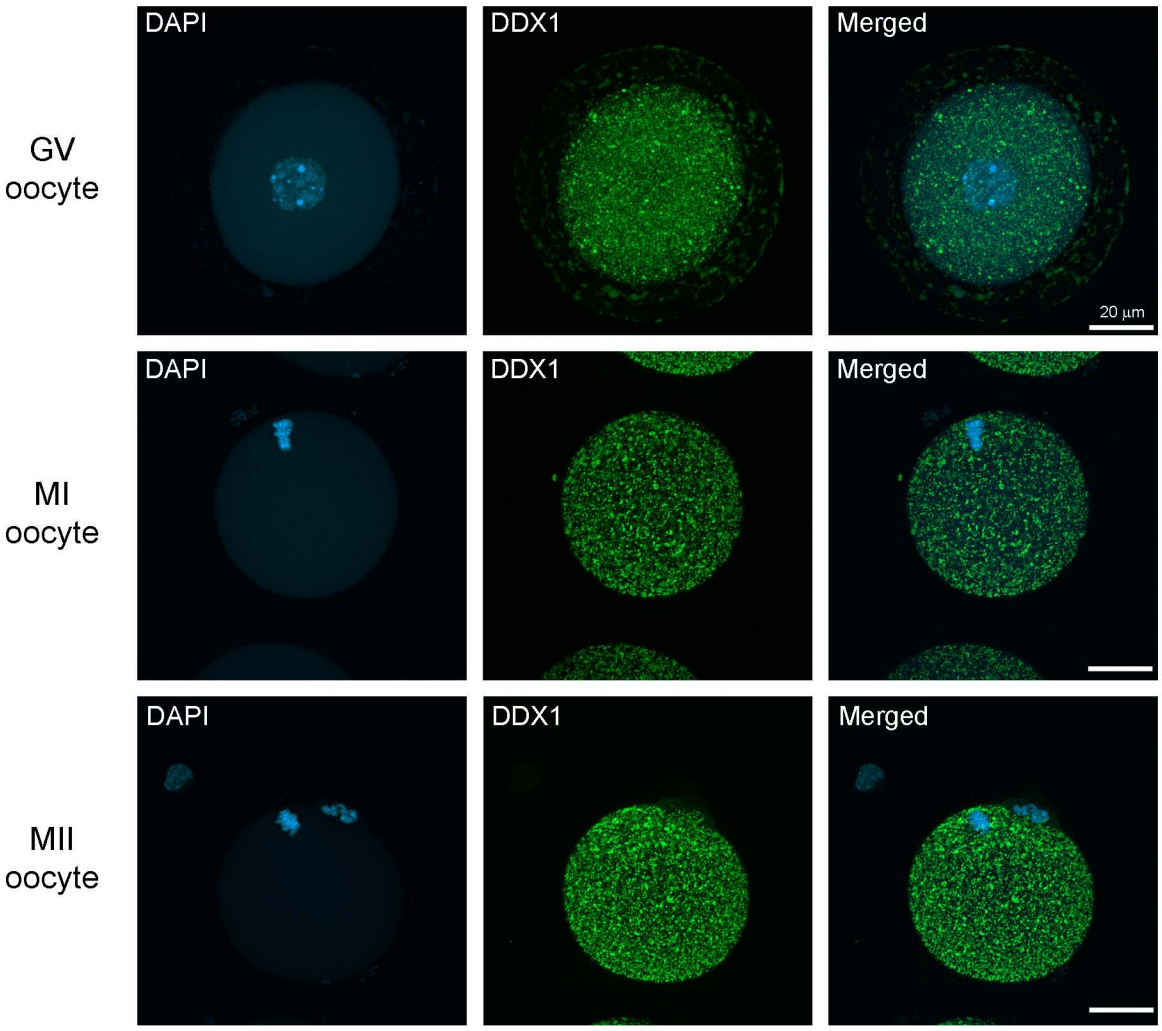


Figure 5.10 Staining for DDX1 in germinal vesicle stage oocytes and maturing oocytes.

GV oocytes were collected by puncturing ovaries collected from adult female mice. MI oocytes were obtained by culturing GV stage oocytes denuded of granulosa cells for 24-48 hours in KSOM media under oil. MII stage oocytes were collected from superovulated female mice. All oocytes were fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. DDX1 was labeled with anti-DDX1 antibody 2910. Nuclei were visualized with DAPI. Embryos were displayed as 2D projections of Z-stacks imaged by confocal microscopy. Scale bar = 20 μm.

nucleus of GV oocytes. By MI, DDX1 foci were more evenly distributed throughout the cytoplasm. At MII, larger aggregates similar to those found in the fertilized oocyte were observed.

5.4 Co-immunostaining of DDX1 and other cellular components

The subcellular localization of DDX1 and the size of the DDX1 foci/aggregates are highly dynamic during development. During oogenesis, DDX1 is found in small foci distributed throughout the cytoplasm. As the embryo develops into a blastocyst, DDX1 is found in increasingly larger aggregates. As DDX1 aggregates are relatively large and numerous, we first tested whether these aggregates might be associated with cell organelles such as ribosomes, mitochondria, endoplasmic reticulum and the Golgi complex by co-labeling cells with anti-DDX1 antibody and markers specific to each type of organelle (Table 2.3). Embryos were collected at different pre-implantation stages of development and were treated as previously described for immunofluorescence labeling. The embryos were visualized by laser scanning confocal microscopy. DDX1 did not co-localize with any of the organelles mentioned above (Figure 5.11).

To determine whether DDX1 might be co-localizing with previously identified DDX1-interacting partners in early stage embryos, we carried out co-immunostaining analysis of embryos with anti-DDX1 antibodies and antibodies to the following DDX1-interacting proteins: RanBPM, CstF64, RIF1 and SMN. RanBPM is a scaffold protein that is associated with nucleocytoplasmic transport and signal transduction (68,242-244). It

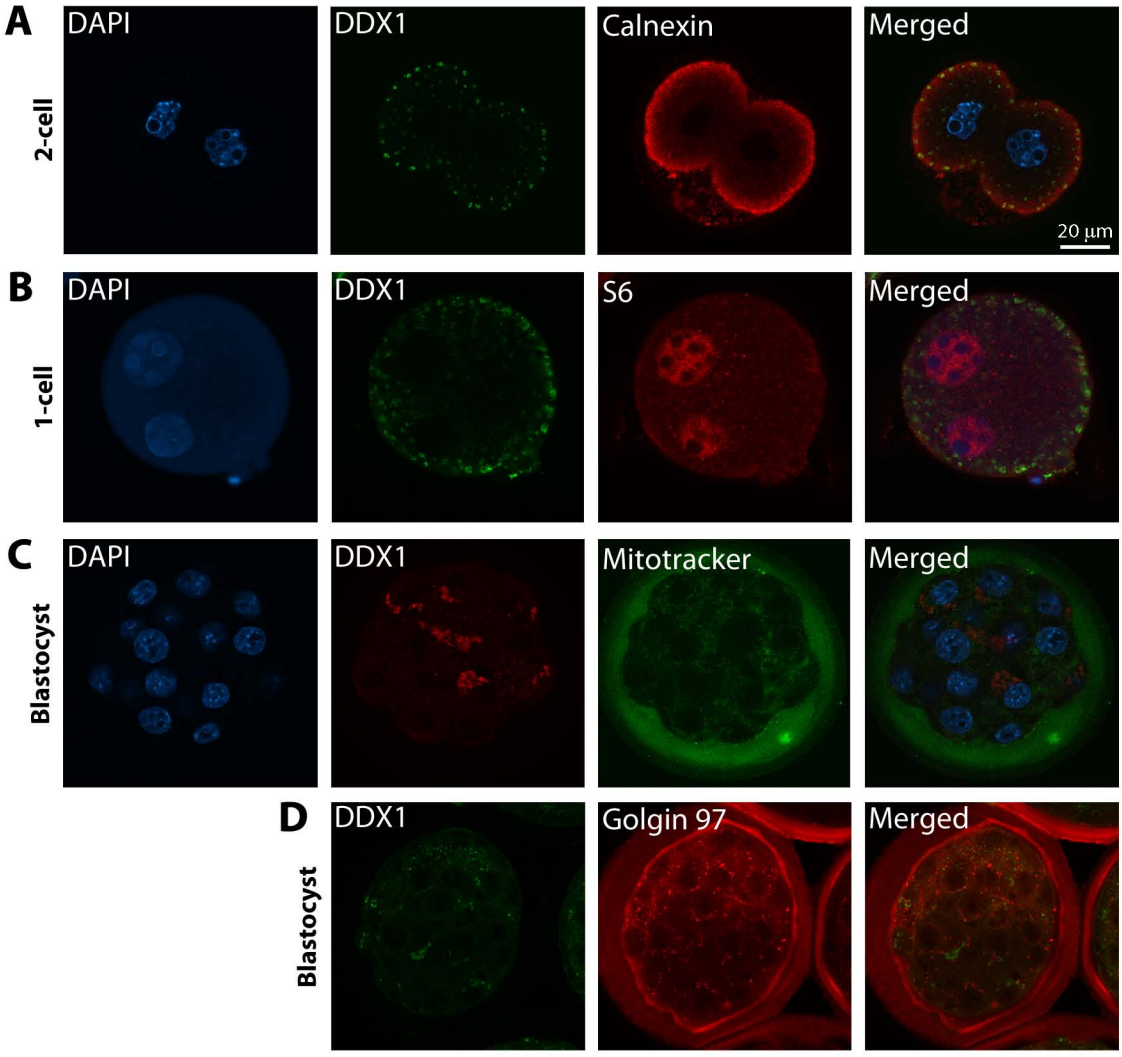


Figure 5.11 Localization of DDX1 and cytoplasmic organelles at various stages of pre-implantation development.

Embryos were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were then double-stained with anti-DDX1 antibody and either anti-calnexin (A) for endoplasmic reticulum, anti-S6 (B) for ribosomes, mitotracker (C) for mitochondria, or anti-golgin97 (D) for the Golgi complex. Nuclei were stained with DAPI. Note that images shown represent a single optical slice through the cell to facilitate co-localization analysis. Embryos were imaged as single optical sections with confocal microscopy. Scale bar = 20 μm.

has been previously found to interact with DDX1 although the reason for this interaction is not understood. CstF64 is a component of cleavage bodies which are associated with 3′ processing of pre-mRNAs. CstF64 was shown to co-localize with DDX1 foci in the nuclei of mammalian cell lines (102). RIF1 is a large protein that is associated with telomere maintenance and NHEJ-directed repair of DNA double-strand breaks (245-251). RIF1 co-localizes with DDX1 foci in mammalian cells and co-immunoprecipitates with DDX1 (Lei Li, unpublished data). SMN is associated with small nuclear RNP complex assembly and co-localizes with DDX1 in nuclear foci called gems as well as in stress granules (113,252,253). DDX1 did not co-localize with any of these proteins (Figure. 5.12) suggesting that the interaction of DDX1 with these proteins is not pertinent to the developing embryo.

Next, we examined whether DDX1 foci or aggregates might be associated with cytoskeletal elements such as actin or microtubules. Embryos were stained with either phalloidin conjugated to Alexa 546 to label actin filaments or anti-β-tubulin to label the microtubules (Figure 5.13). There was no co-localization of DDX1 and microtubules, with microtubules primarily found just below the cell surface membrane whereas DDX1 foci were located throughout the cytoplasm. Phalloidin staining revealed actin filaments throughout the embryo. There was no clear co-localization pattern for DDX1 foci and aggregates and filamentous actin.

DDX1 is a member of a protein family with well-recognized roles in RNA metabolism. We therefore focused on the possibility that RNA or other RNA binding

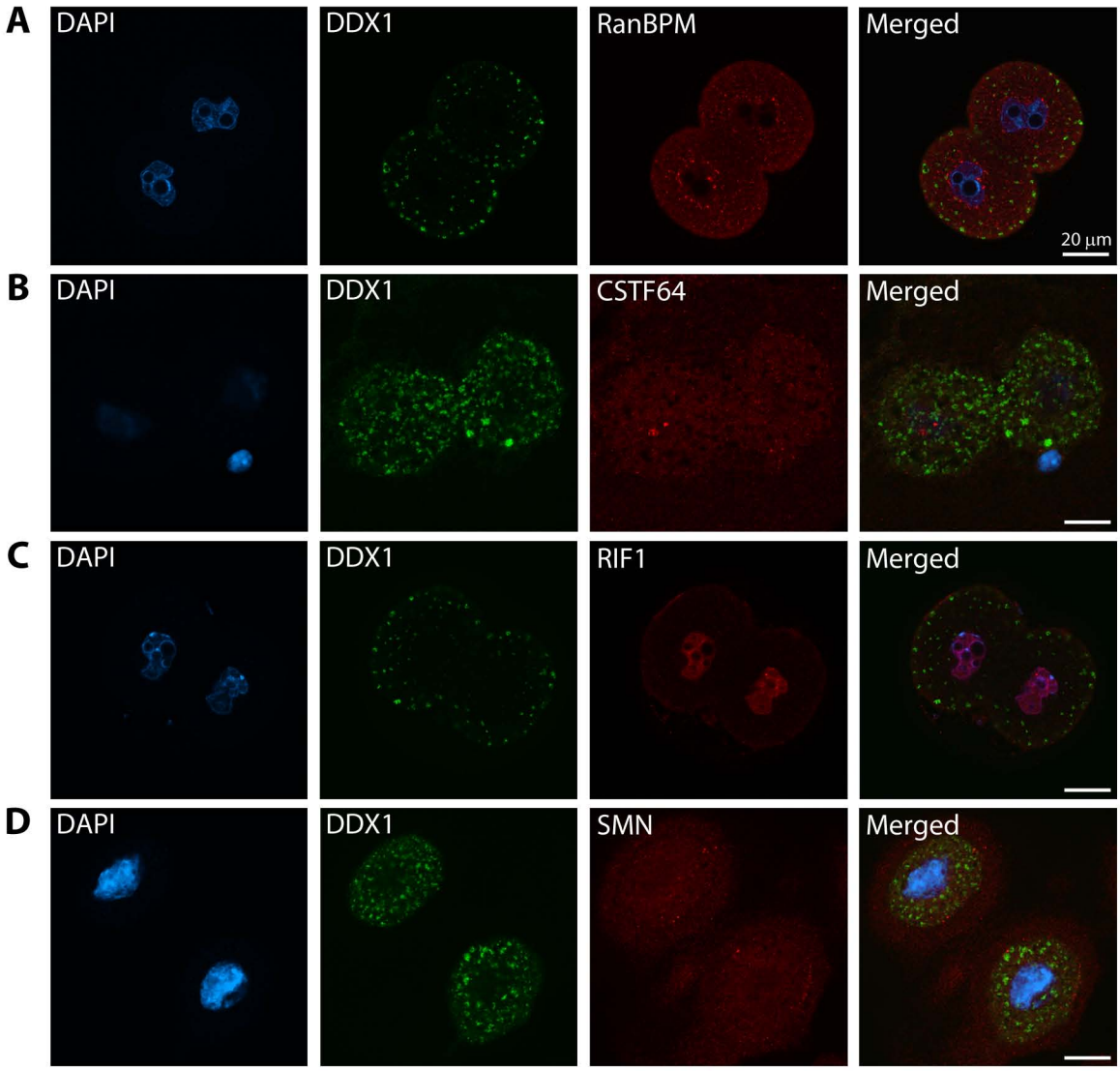


Figure 5.12 Immunostaining of DDX1 and known DDX1 interacting proteins in 2-cell stage embryos.

Embryos were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were co-immunostained with anti-DDX1 antibody and either anti-RanBPM (A), anti-CSTF64 (B), anti-RIF1 (C), or anti-SMN (D) antibodies. Nuclei were visualized with DAPI. Embryos were imaged as single optical sections with confocal microscopy. Scale bars = 20 µm.

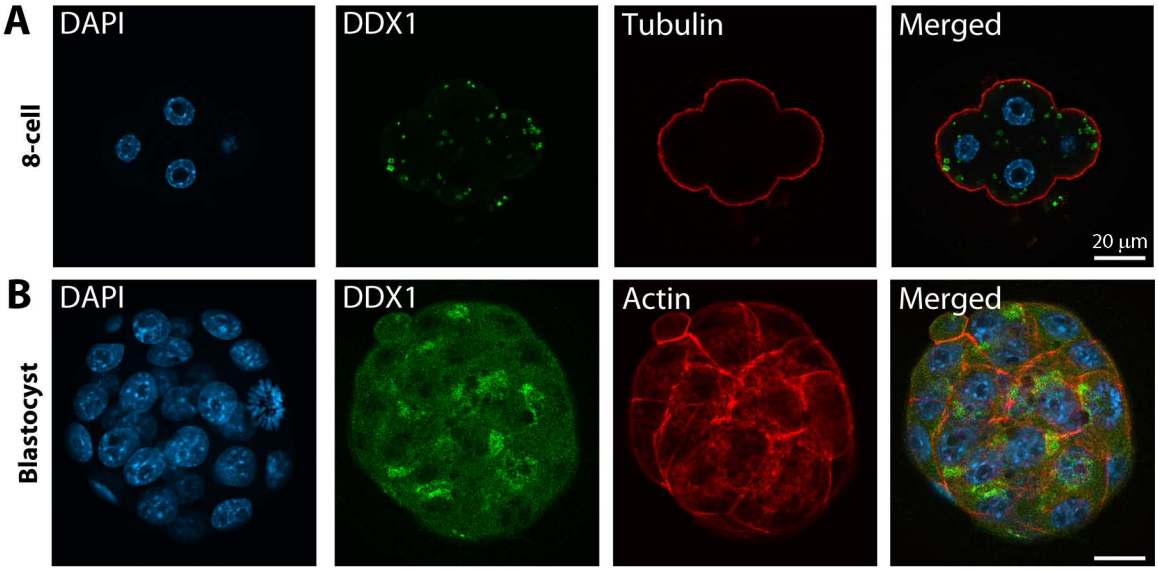


Figure 5.13 Immunostaining of DDX1 and cytoskeleton markers in E2.5 and E3.5 mouse embryos.

Embryos were collected and fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. (A) Embryos were co-immunostained with anti-DDX1 and anti-β-tubulin antibodies, followed by incubation with Alexa 488-conjugated goat anti-rabbit secondary and Alexa 555-conjugated goat anti-mouse secondary antibodies respectively. (B) Embryos were co-stained with anti-DDX1 antibody followed by Alexa 488-conjugated goat anti-rabbit secondary and Alexa 546 conjugated phalloidin. Nuclei were stained with DAPI. Embryos were imaged as either single optical sections (A) or as a 2D projection of Z-stack images (B) collected with a confocal microscope. Scale bars = 20 μm.

proteins might be present in DDX1 aggregates. To detect RNA, cells were stained with acridine orange. Although aggregates of RNA were identified using acridine orange, there was little if any overlap with DDX1 aggregates (Figure 5.14). DDX3 is a multifunctional DEAD box protein which is associated with mRNA processing (254-256). Like DDX1, DDX3 is involved in HIV-1 viral transcript transport as well as mRNA transport in neuronal granules (254,256,257). None of the embryos that were co-immunostained with anti-DDX1 and anti-DDX3 antibodies showed co-localization of these two proteins even though, like DDX1, DDX3 localized to numerous cytoplasmic foci at the blastocyst stage.

EXOSC5 is a component of the RNA exosome which plays a major role in all RNA degradation processes (258,259). EXOSC5 was detected in numerous foci throughout the embryo (Figure. 5.14). EXOSC5 foci were considerably smaller than DDX1 aggregates. Finally, we investigated the possibility that DDX1 might be associated with RNA degradation in early stage embryos. To address this possibility, we used an antibody to GW182, a component of RNA processing bodies involved with mRNA degradation (260-263). Although the antibody to GW182 detected aggregates within the embryos, there was no co-localization of these aggregates with DDX1 aggregates (Figure 5.14). Thus, we have yet to identify any protein that co-localizes with DDX1 in pre-implantation embryos.

5.5 Aggregation of DDX1 in embryos is RNA-dependent

Irradiation-induced DDX1 foci that form in the nucleus of cells that are exposed to γ-irradiation are dependent on the presence of RNA/DNA hybrids at sites of DNA double-strand breaks (124). Treatment of cells with RNAse H after irradiation results in loss of DDX1 foci at double-strand breaks. Furthermore, DDX1 has previously been shown to co-immunoprecipitate with stress granule components, an interaction that is dependent on the presence of single-strand RNA as the interaction is lost upon RNase A treatment (113). To determine whether the presence of DDX1 aggregates in early embryos is dependent on the presence of RNA, embryos were pre-treated with 2% Tween-20 prior to fixation followed by treatment with 150 μg/ml RNase A in PBS (PBS served as the negative control). Embryos were then fixed, immunostained using anti-DDX1 antibody and imaged as previously described. The detergent pre-treatment resulted in some cell shrinkage; however, the DDX1 aggregates were still present in control embryos (Figure. 5.15). Upon RNase A treatment, virtually all DDX1 aggregates disappeared resulting in a diffuse DDX1 staining pattern. These data suggest that DDX1 forms a complex with RNA during early stages of development.

5.6 DDX1 granules are unaffected by transcription or translation inhibition

Since the DDX1 aggregates were found to be RNA dependent, we treated embryos with inhibitors of transcription or translation to see what effect this might have on the size, shape or number of DDX1 foci/aggregates. We collected embryos at

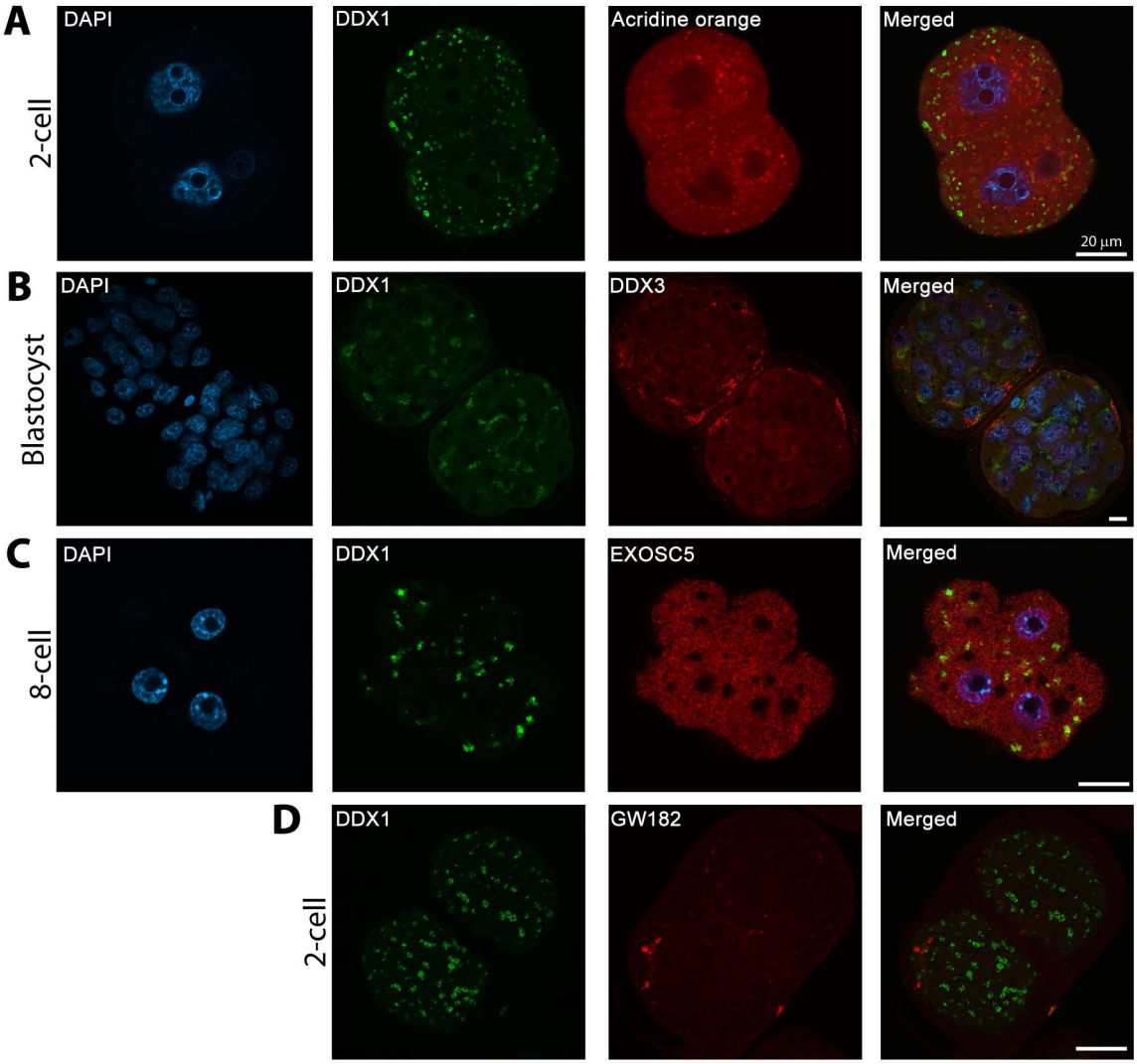


Figure 5.14 Co-staining of DDX1 and RNA or RNA binding proteins at various stages of pre-implantation development.

(A) Embryos were incubated with acridine orange at 2 μM for 20 minutes prior to fixation and staining for DDX1. (B-D) Embryos were fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were co-immunostained with anti-DDX1 antibody andand anti-DDX3 antibodies (B), anti-DDX1 and anti-EXOSC5 antibodies (C), or anti-DDX1 and anti-GW182 antibodies (D) antibodies. Nuclei were visualized with DAPI. Embryos were imaged as single optical sections with confocal microscopy. Scale bars = 20 μm.

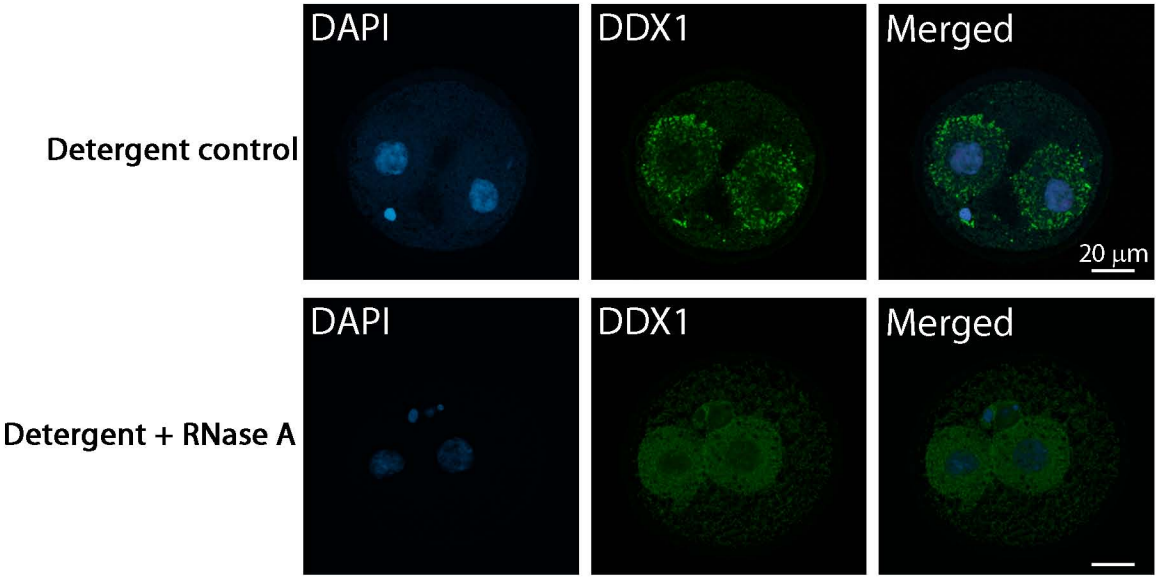


Figure 5.15 Formation of DDX1 aggregates at the 2-cell stage is RNA-dependent. E1.5 embryos were permeabilized prior to fixation in 0.1% saponin for 6 minutes followed by PBS treatment (control) or 150 μg/ml RNase A/PBS for 20 minutes at 37°C. Embryos were then fixed with 4% paraformaldehyde in PBS for 15 minutes and permeabilized for 5 minutes in 0.5% Triton-X-100/PBS. Embryos were immunostained with anti-DDX1 antibody. Nuclei were visualized with DAPI. Embryos were imaged as single optical sections with confocal microscopy. Scale bars = 20 μm.

E1.5 and incubated them in either 200 μM cordycepin for 8 hours to inhibit transcription or 150 μg/ml cyclohexamide for 8 hours to inhibit translation. The embryos were then fixed, labeled with anti-DDX1 antibody and imaged as described previously. There was no difference in the number or appearance of DDX1 aggregates following cyclohexamide treatment (Figure 5.16). Cordycepin treatment resulted in some disruption to the organization of the DDX1 aggregates with some clumping of aggregates observed. This clumping of aggregates was similar to the DDX1 staining pattern observed in embryos at the blastocyst stage.

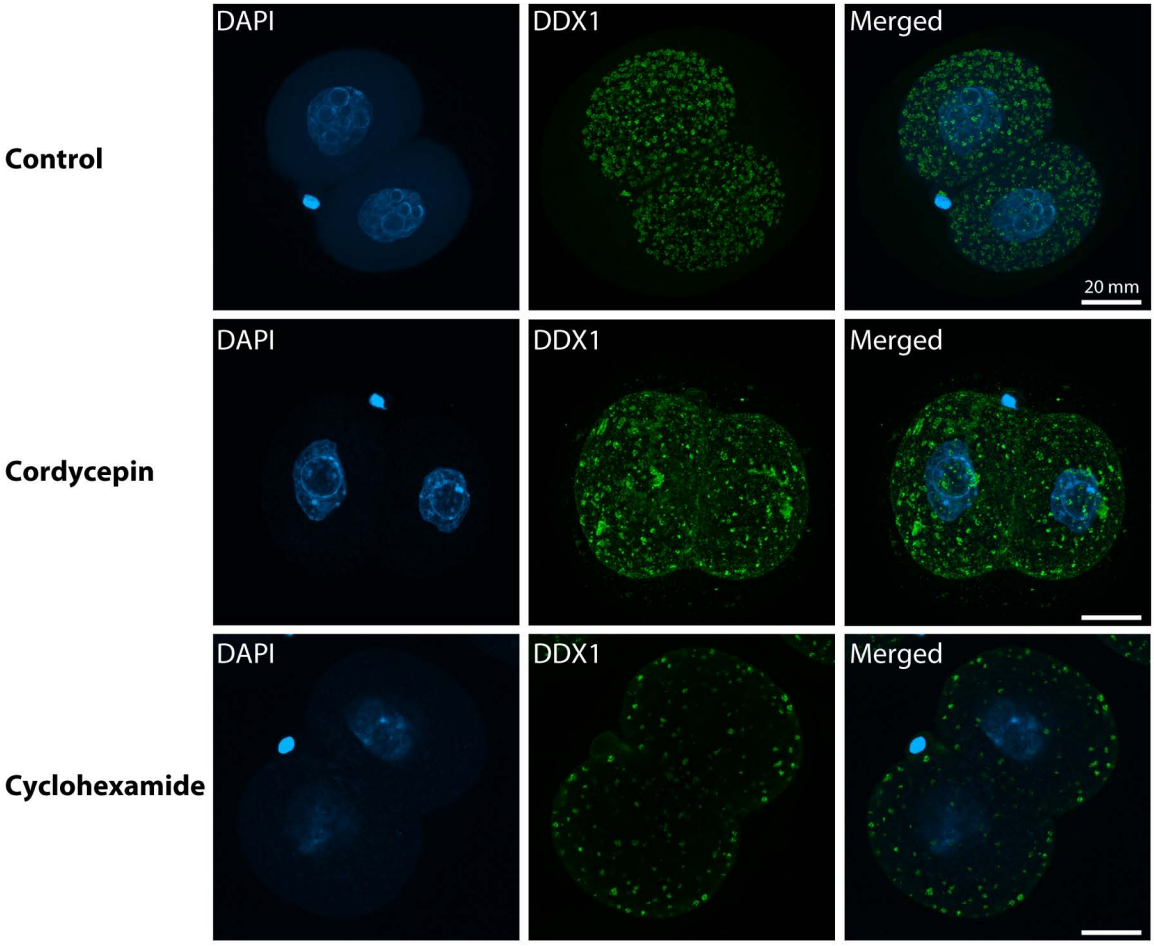


Figure 5.16 Effects of transcription or translation inhibition on DDX1 aggregates in E1.5 embryos.

E1.5 embryos were incubated in either 200 μM cordycepin for 8 hours to inhibit transcription or in 150 μg/ml cyclohexamide for 8 hours to inhibit translation. The embryos were then fixed and labeled with anti-DDX1 antibody. Nuclei were visualized with DAPI. Control and cordycepin embryos were displayed as 2D projections of Z-stacks while the cyclohexamide treated samples embryos were displayed as single optical slices using confocal microscopy. Scale bars = 20 μm.

Chapter 6

Discussion

6.1 Structural analysis of DDX1

Several DEAD box protein structures have been solved using X-ray crystallography (28,29,33). The two helicase domains fold to form a dumbbell-like shape with the two helicase domains separated by the interdomain region. Motifs from both helicase domains are required to form the ATP binding pocket, RNA binding, and domain interactions. ATP binding was also found to require a magnesium ion. Several structures were solved in both the presence and absence of RNA substrates as well as non-hydrolysable ATP analogues (37,264,265). Comparing these structures demonstrated the apparent flexibility of DEAD box proteins and how protein conformation is affected by substrate binding. The RNA binding motifs are found in both helicase domains and in the open state these motifs fail to form an RNA binding pocket and require additional conformational changes generally caused by ATP binding (33). Following ATP binding, RNA binding shifts the protein into its fully closed conformation (33). It has been proposed that the hydrolysis of ATP results in further conformational changes which alter the DDX-RNA interaction such that bound RNA is shifted to a strained position which promotes its remodeling by duplex destabilization and localized unwinding.

DEAD box protein structures have allowed a better understanding of how this family of proteins modifies RNA secondary structure and interacts with other proteins. DDX1 is unusual when compared to other DEAD box proteins in that it has a SPRY domain found within helicase domain I. Only one other DEAD box protein identified to date (DDX24) has a domain inserted into the helicase core (25). SPRY domains in other

proteins have been postulated to be involved in either protein-protein or protein-RNA interactions (67-69). The role of the SPRY domain in DDX1 function is not known. Our goal in generating the structure of DDX1 was to determine the structural impact of the SPRY domain on the overall folding pattern of the helicase domain. Ideally, co-crystals obtained in the presence of DDX1 and either RNA-RNA or RNA-DNA duplexes would be expected to generate additional information regarding the role of the SPRY domain in RNA binding.

Despite multiple attempts to crystallize DDX1, the structure of DDX1 still eludes us. Our attempts at generating protein for crystallography were met with challenges including protein solubility in the sub-10 mg/ml concentration range, and the presence of a doublet band in the purified full-length DDX1 protein suggesting the presence of a degradation product. Despite these problems, we were still able to grow some crystals in the small and large scale experiments. Unfortunately, the crystals that were produced failed to generate any diffraction patterns even at low resolution. We found that DDX1 tended to form crystals in the presence of Mg^{2+} and negative ions such as citrate and sulphate. DEAD box proteins are known to bind Mg^{2+} so this ion may be stabilizing the protein structure. Negative ions are likely associating with the RNA binding pocket which is positively charged due to exposed lysine and arginine residues and have been shown to be required for the generation of several other DEAD box protein crystals (33,37,38,266).

The inclusion of ATP was also found to help in the generation of DDX1 crystals. Some of the other DEAD box proteins have been shown to shift between open and

closed states and these conformational changes were influenced by ATP analogue binding or RNA binding (38,266). A number of labs have shown that crystallization of DEAD box proteins can also occur in the presence of RNA substrates, with the proteins taking on a more closed conformation (37,38,264). Our attempts to stabilize DDX1 by co-crystallizing DDX1 in the presence of an RNA-DNA duplex resulted in the formation of some crystals using small scale conditions; however, large scale conditions failed to generate crystals for diffraction.

The DDX1 protein is a relatively large protein, with a predicted molecular mass of 83 kDa. In order to increase protein yields and reduce degradation of the purified protein, we also tested large truncation mutants for their ability to crystallize. The N-terminal proteins were designed to divide the protein into its two helicase domains to yield 60 kDa and 30 kDa truncated proteins. Very low yields of the C-terminal DDX1 protein (30 kDa) were obtained, generating low protein concentrations. Despite this, we observed a few crystals in the small scale experiments; however, because of low yields, we were unable to proceed to the large scale experiments. The N-terminal DDX1 protein was much easier to produce and generated the highest concentrations of recombinant protein (although still below 10 mg/ml). N-terminal DDX1 protein generated crystals; however, they failed to produce a diffraction pattern. In retrospect, the design of the mutants may have resulted in excess linker sequences on the C-terminal end of helicase domain I. It is possible that truncated mutant protein with shorter linker sequences would produce more stable structures.

There are reports in the literature indicating that widely different results can be achieved using proteins from different species (267). For example, Stringent starvation protein A (SspA) is a bacterial protein that failed to generate quality diffraction data when the protein was prepared from *E. coli*; however, the same protein prepared from *Y. pestis* orthologues generated a diffraction pattern (268). We prepared recombinant DDX1 protein from 4 different species: human, chicken, mouse and *Drosophila*. *Ddx1* was cloned from cDNAs prepared from each of these four species and inserted into GST-tag expression vectors. Large quantities of full-length recombinant DDX1 protein were produced from human and *Drosophila*, but not mouse and chicken. *Drosophila* DDX1 is 68% identical to human DDX1 and is shorter by 20 amino acids. Human DDX1 protein generated more crystals than *Drosophila* DDX1 in the small scale experiments. The large scale attempts to grow *Drosophila* DDX1 crystals were unsuccessful; however, there are still several conditions that could be tested for growth.

Although a few attempts were made to test human DDX1 for large-scale crystal formation in the presence of RNA-DNA duplexes, additional experiments should be carried out using both human and *Drosophila* DDX1. Future experiments should focus on pursuing the small scale hits of human DDX1 supplemented with ATP, Mg^{2+}, and RNA-DNA duplex that were identified from the Nucleix suite. In particular, conditions containing spermine should be tested as this is a known nucleic acid binding protein and may assist in stabilizing the RNA-DNA duplex bound to DDX1. Additional focus on the N-terminal DDX1 protein in the presence of different concentrations of ATP and Mg^{2+} may also yield useful data. Furthermore, all crystals were grown at room

temperature and future work should include growing crystals at either reduced or increased temperatures to help stabilize the protein or protein-nucleotide complexes during crystallization (269-272). Decreasing the temperature may increase overall protein stability and decrease the rate of evaporation thereby prolonging the crystallization stage and reducing the formation of small crystals or precipitates. Increasing the temperature may lead to solubility differences which may increase the likelihood of nucleation events resulting in crystal formation.

The failure to generate a diffraction pattern could be due to a number of reasons including poor purity of the crystals or poor crystal conditions. The full-length DDX1 protein was not purified as a single band and the presence of multiple degradation products may have contributed to heterogeneous crystal formation and the failure to generate a diffraction pattern. Recently, there have been attempts using *in situ* proteolysis to generate crystals of proteins that have previously failed (273-277). This method includes a small amount of protease, such as trypsin or chymotrypsin, in addition to the mixture of protein and buffer in the crystallography drop. The *in situ* cleavage is thought to generate proteins more likely to generate crystals by cleaving flexible regions. This is similar to designing truncation mutants; however, in this case the truncations are formed during the crystallization step. An additional structure of the RIG-1 ATPase domain was recently solved using this method as it was hypothesized that the protein contained several flexible regions which prevented previous crystals from diffracting (278). Using this method, we may be able to improve

the quality of the crystals or generate alternative crystal forms that can be used to generate a complete or partial structure for DDX1.

6.2 DDX1 is a non-specific ribonuclease

Based on a previous publication from the lab, DDX1 is able to degrade ssRNA in an energy-independent and magnesium-dependent manner (124). In this paper, Li *et al*. report that DDX1 plays a role in the repair of DNA double-strand breaks and that retention of DDX1 at DNA double-strand breaks is dependent on DNA-RNA structures. Li *et al.* therefore proposed that DDX1 facilitates the removal or degradation of RNA from sites of double-strand breaks thereby allowing more efficient DNA repair. In keeping with the idea that DDX1 has ribonuclease activity, our *in vitro* data have shown that DDX1 is able to degrade ssRNA molecules made up of mixed nucleotides. I pursued these analyses by demonstrating that DDX1 can degrade ssRNA substrates made up of either pure purine or pyrimidine stretches. These data support the hypothesis that DDX1 does not have sequence specificity in regards to its ribonuclease activity and may degrade any RNA substrate, with the following caveats: (i) reactions were carried out using excess of DDX1 which may mask true substrate specificity, and (ii) DDX1 activity was low suggesting that DDX1 may interact with other proteins to optimize its ribonuclease activity.

Using the R29 (mixed nucleotides) ssRNA substrate labeled at the 3′ end, the 5′ end or internally-labeled, I found that DDX1 degrades ssRNAs to short oligonucleotides between 3-6 nucleotides in length. My results indicate that DDX1

does not have any nucleotide specificity. Furthermore, DDX1 may not be an exoribonuclease as a single nucleotide ladder would be expected if it had 3′-5′ exoribonuclease, and only single nucleotides if it had 5′-3′ exoribonuclease activity (using a 5′ end-labeled substrate). Single nucleotides were observed when internally labeled substrates were used which argues against endoribonuclease activity. In order to resolve whether DDX1 is an endoribonuclea, and exoribonuclease, or both, we will need to test additional substrates including R29 substrates labeled at the 3′end.

The minimum size of RNA that DDX1 can interact with is unknown, although it is likely an oligonucleotide of at least 6 nucleotides as one of the major degradation products was 6 nucleotides in length. These data are supported by crystallography data of protein-RNA co-crystals whereby the RNA binding pocket for DDX4, DDX19 and DDX48 were all shown to have 6 nucleotides directly interacting with the protein (38,266,279,280). Experiments could be carried out to determine the size of the smallest ssRNA that DDX1 can digest by testing progressively smaller oligonucleotides. Identifying the smallest ssRNA substrate for DDX1 ribonuclease activity would allow us to perform kinetics studies as there would be a single enzymatic event occurring on each RNA molecule.

Timed experiments were carried out in an attempt to identify intermediate digestion products of DDX1. Intermediate-size bands were observed when DDX1 was incubated with end-labeled R29. As R29 is made of 29 mixed nucleotides, it is possible that R29 is able to form secondary structures which can slow down or inhibit the ribonuclease activity of DDX1. The other substrates containing single stretches of a

single nucleotide were rapidly digested. AR29 (29 A's) was more rapidly and completely digested as compared to CR29 (29 C's). DDX1 has been previously shown to bind to poly(A) RNA and but not poly(C) RNA *in vitro* (281). Based on these data, DDX1 likely requires a ssRNA substrate for efficient degradation and may act preferentially on naked poly(A) RNA. Further analysis using decreasing DDX1 concentrations will be required to determine DDX1 sequence specificity and activity. As some of our reactions may have been within a few minutes of incubation, using less DDX1 per reaction would allow for kinetic analysis of the ribonuclease activity. Additionally, we could design longer or more complex oligonucleotides to study the effect of RNA secondary structure on the activity of DDX1. Longer oligonucleotides could be designed to fold into specific secondary structures, such as stem loops, to determine whether DDX1 can degrade these substrates. We could also design chimeric RNA-DNA oligonucleotides to further investigate how DDX1 cleaves RNA substrates.

The ribonuclease activity of DDX1 is weak *in vitro* and requires a significant molar excess of protein to substrate. It is possible that DDX1 requires co-factors or needs to be post-translationally modified to stimulate its activity. We have proposed that DDX1 plays a role in the clearance of RNA present at sites of DNA double-strand breaks. Proteins recruited to these sites may facilitate direct stimulation of DDX1's ribonuclease activity. It is important to note that different proteins may facilitate or inhibit the enzymatic activity of DDX1 depending on cellular context. For example, interaction between DDX1, Rev, and HIV genomic RNA facilitates the transport of full

length RNAs and prevents splicing, suggesting that DDX1's ribonuclease activity may be down-regulated in HIV-infected cells (117,282,283). DDX1 is also found in RNA transport granules involved in the transport of mRNAs to sites of translation (111,112). Again, the presence of a ribonuclease in RNA transport granules would seem counterintuitive to transporting full-length mRNA.

DDX1 activity could also be regulated by post-translational modification. DDX1 has been shown to be phosphorylated by the ATM kinase (124) and 2D gel electrophoresis suggests that there are as many as 6 different post-translationally modified forms of DDX1 in HeLa cells (Lei Li, unpublished data). Other DEAD box proteins have been shown to undergo post-translational modification (6). For example, DDX5 and DDX17 (p68 and p72) are heavily regulated through various modifications including polyubiquitylation, sumoylation and phosphorylation. These modifications regulate protein-protein interactions can influence DEAD box protein localization and function.

RNA is a key biological molecule and it is tightly regulated in the cell. Degrading RNA is a rapid process and occurs through a variety of different mechanisms. One of the best characterized enzymes is the secreted endoribonuclease Ribonuclease A (RNase A), a common lab contaminant. The primary role of this family of proteins is to degrade foreign RNA (284,285). RNase A was one of the first proteins to have its structure determined by X-ray crystallography (286). RNA is by nature unstable due to the reactive 2′ hydroxyl group present in the sugar base. RNase A catalyzes the reaction of the 2′ hydroxyl group to form a 2′, 3′-cyclic phosphate which

breaks the backbone. RNase A has some specificity in that it cleaves ssRNA following pyrimidine residues. Histidine and lysine residues play key roles in the active site to catalyze the reaction. RNase A is an incredibly stable enzyme due primarily to its small size and high number of cysteine residues which form disulphide bridges (287). One of the approaches used to purify RNase A was to treat tissue with sulfuric acid and bring the reaction close to boiling temperatures, with the RNase A being the only surviving protein (288). RNase A's impressive stability make it an important lab contaminant. As a result, we have had to ensure that the ribonuclease activity associated with DDX1 is not due to RNAse A contamination. For this reason, we have gone to considerable trouble to ensure that the ribonuclease activity of DDX1 was different from that of RNase A both described above and previously published (124). In particular, DDX1's ribonuclease activity is destroyed by boiling, and unlike RNase A, DDX1's ribonuclease activity is magnesium-dependent.

There are several other ribonucleases, some which degrade dsRNA (e.g. Dicer, Drosha,) and some which degrade ssRNA in a general manner (e.g. RNase L) or in a specific manner (e.g. G3BP) (289). The mechanisms of RNA cleavage are known for some of these enzymes whereas others are still poorly understood. Structural studies have been key to determining the active sites of these enzymes. For example SMG6, a member of the nonsense: mediated mRNA decay pathway, has a similar structure to that of the RNase H family (290,291).

A number of ribonuclease domains have been identified including RNase A, H, P and III. Searching the DDX1 sequence for known ribonuclease domains was not

successful, with no putative domains identified (292,293). Identifying the different substrates that DDX1 is able to degrade may provide insight into which residues are essential for its ribonuclease activity as DDX1 may use similar mechanisms to that of other known ribonucleases. In addition, we used a truncation mutant strategy to identify possible ribonuclease domains of DDX1. We found that shortening the C-terminal region by 120 amino acids abolished DDX1's activity and shortening the protein by 95 amino acids reduced DDX1's ribonuclease activity. The role of the non-conserved C-terminal region of DDX1 is unknown with no identifiable domains located in this region. The C-terminal region of DDX1 could either play a role in its enzymatic activity or facilitate RNA binding to DDX1. Additional truncation and substitution mutations within this region may identify the minimal amino acid sequence required for ribonuclease activity. Electrophoretic mobility shift assays (EMSAs) could also be used to test the binding of RNA to wild-type DDX1 and mutant DDX1 proteins.

Another region of DDX1 that may contain a ribonuclease domain is the interdomain region between the two helicase domains. The interdomain region of DDX1 was targeted as a possible ribonuclease domain because it shares little homology with other DEAD box proteins and no other DEAD box protein has been shown to have ribonuclease activity. As the interdomain region resides between the two helicase domains, it is possible for it to fold in such a way as to interact with ssRNAs bound to DDX1. We performed sequence analysis of DDX1 from all known orthologues to identify highly conserved residues in the interdomain region. We

found that residues corresponding to human D430, D466, H470, S486, K490, and K493 were all highly conserved between the different orthologues of DDX1. These types of amino acids are also associated with the active sites of other ribonucleases although mutational analysis will be required to determine whether any of these residues play a role in DDX1-mediated RNA degradation (294-302).

The degradation of internally-labeled R29 ssRNA produced an interesting result in that we observed both single-labeled nucleotides and the larger oligonucleotides (3-6 nts) that were also present when end-labeled substrates were degraded. Previously, we had not observed any single nucleotide products. This may be due to the nature of the substrate as it was end-labeled. Perhaps DDX1 is able to cleave the RNA at multiple positions which in some cases would release single nucleotides. At this point we do not know the active site of DDX1 so additional experiments will be required to determine how DDX1 degrades RNA to single nucleotides or oligonucleotides 3-6 nts in length.

As previously discussed, most other ribonucleases catalyze the inherent instability of RNA by facilitating the interaction between the 2′ and 3′ hydroxyl groups to form a 2′, 3′-cyclic phosphate (289,303). To determine whether DDX1 degrades RNA in a similar manner, we could perform the following two experiments. First, to examine whether the 5′ end of the products generated by DDX1 have a hydroxyl group, we would degrade unlabelled ssRNA with DDX1 then perform a kinase labeling reaction on the products. RNA products retaining a 5′ hydroxyl group will be labeled with protein kinase. Second, to examine whether the 3′ end is modified, we would

subject the labeled ssRNA to periodate oxidation following degradation by DDX1 (304). The periodate oxidation breaks the bond between the 2′ and 3′ carbon in the ribose molecule which would a product visibly larger in size. Should no difference between the treated degradation products be observed, this would indicate that the 3′ end is likely modified through the formation of the 2′, 3′-cyclic phosphate (304).

6.3 *Ddx1*$^{-/-}$ mice die during pre-implantation development

Knock-out of several genes encoding DEAD box proteins has been found to be lethal during development (222-224,227-229). To date, only one other DEAD box gene knockout (*Ddx20)* has been shown to result in pre-implantation lethality (229). Although the function of DDX20 is still poorly understood, it is upregulated during MZT which suggests that DDX20 plays a key role in embryonic gene expression (229,230). Our data indicate that DDX1 is essential for pre-implantation in mice. Although we were unable to genotype embryos at the pre-blastocyst stage, the presence of developmentally arrested embryos in culture and the DDX1 immunofluorescence staining data lead us to hypothesize that the *Ddx1*$^{-/-}$ embryos stall at the 2- or 4-cell stage. The embryos stalling at different time points is likely due to residual DDX1 protein present as maternal complement. Based on sq-RTPCR we found that DDX1 is first expressed by the embryo at the 2-cell stage. The timing of embryonic lethality in mice suggests that DDX1 may play a role during the maternal to zygotic transition in the utilization of transcripts. DDX1 may facilitate the protection of some specific maternal RNAs, or perhaps it is associated with regulation of the newly

generated zygotic transcripts. The apparent lack of DDX1 in the nucleus of early stage embryos (1-cell to blastocyst stages) suggests that its cytoplasmic rather than nuclear role is essential for early development.

By placing 2- and 4-cell embryos in specialized culture medium, we were able to follow their developmental progress over the course of 48–72 hours. Stalled embryos were routinely observed upon crossing *Ddx1*$^{+/-}$ mice. Immunostaining analysis of the stalled embryos revealed low levels of DDX1, with few DDX1 aggregates remaining in the cytoplasm of embryonic cells after 72 hours in culture. Stalled embryos were also found to contain multiple nuclei suggesting that the cell cycle was able to progress despite the lack of cell division. Using live cell imaging we found that the embryos did not initiate cytokinesis and were stalled during interphase. The next step will be to carry out live cell imaging in conjunction with Hoechst staining of the nuclei and fluorescence imaging to determine whether stalled embryos have a normal cell cycle, other than absence of cytokinesis.

Ddx1 knock-out results in early embryonic lethality in mice, likely at the 2- to 4-cell stage. These results are different from those obtained using the *Drosophila melanogaster* model which is also being studied in the lab. *Ddx1* knockout flies are viable but infertile, indicating that there is no need for embryo-derived DDX1 during fly development. The early stages of development differ substantially between mice and flies and may account for the difference in survival of *Ddx1*$^{-/-}$ embryos. *Drosophila melanogaster* nuclei divide in a synchronous manner; however, they do not form cell membranes and the entire embryo is a syncytia until just prior to gastrulation

(approximately 13 mitotic divisions). Also unlike mice, *Drosophila* embryos retain many maternal transcripts and proteins during early stages of development and these early stages of development progress at a much faster rate than those in mice. It is therefore possible that there is no need for DDX1 in flies until the following generation when there is no longer any protection effect caused by maternal DDX1. Also, given that: (i) our stalled mouse embryos do not undergo cytokinesis and (ii) fly embryos form a syncytia thereby circumventing the need for cell division, the fundamental defect underlying *Ddx1* knock-out may be related to a role for DDX1 in cytokinesis in early stage mouse embryos.

The early embryonic lethality observed in *Ddx1*$^{-/-}$ mice limited our ability to study the biological role of DDX1 during development. To bypass this problem, we attempted to generate a tissue-specific *Ddx1* knockout using the Cre/*loxP* recombinase system. Although three embryonic stem cell lines carrying a mutant allele of *Ddx1* were tested, we were not able to obtain any chimeric mice with germline transmission of the targeted allele. Analysis of the genomic DNA isolated from the three embryonic stem cell lines revealed a mixed population of cells, with cells bearing the mutated allele being outcompeted in culture over time. It may be possible to re-establish the *Ddx1*$^{+/-}$ cell line by placing the cells under selection again. These selected cells could then be used to generate chimeras by microinjection or aggregation.

Generation of a *Ddx1*^{flox} mouse strain would allow examination of the role of DDX1 using two different approaches. First, we could use a oocyte specific promoter

to determine the role of DDX1 in oocyte development and what role the maternal complement of DDX1 plays in the very early stages of zygotic development. This approach might shed light on whether the *Ddx1*$^{-/-}$ embryos stall due to maternal complement RNAs being misregulated or whether DDX1 plays a key role in the maternal to zygotic transition. Second, we could examine the role of DDX1 at later stages of development by crossing the floxed mice to mice expressing Cre recombinase under the control of tissue-specific promoters such as the *Dkk3* (Dickkopf-3) promoter to knockout *Ddx1* specifically in retina progenitor cells (305). As shown in this thesis, DDX1 is highly expressed in the retina and then becomes restricted to specific cell lineages following differentiation such as the ganglion and amacrine cells. The loss of DDX1 in retinal progenitor cells would likely produce a mouse with retinal defects but otherwise normal. By studying the perturbations to the developing retina we would gain a better understanding of the role of DDX1 in differentiated cells.

6.4 Subcellular localization of DDX1

Based on previously published expression analysis, DDX1 is expressed in most if not all tissues, with highest levels found in early development and neuroectoderm-derived tissues (98). The results presented in this thesis for mouse tissues are in agreement with previous reports. We found DDX1 expression in all tissues tested, although not all cells were positive for DDX1. As expected based on cell line data (85,101), DDX1 was primarily localized to the nucleus of most cells. Interestingly, DDX1

was also found in the cytoplasm of some cells; e.g., in the intestinal crypts and the developing gametes. The cells of the intestinal crypt have high rates of proliferation as a constant supply of cells are required to replace the cells of the villi which are constantly being shed. Similarly, the highly proliferative spermatogonium cells of the testes have higher expression of DDX1 than mature spermatids. These data link rapid proliferation, at least in some cell types, to a requirement for cytoplasmic DDX1.

DDX1 expression in mouse cerebellum was quite different in P0 versus adult brain. At P0, when the cerebellum is still immature, DDX1 was expressed primarily in the Purkinje cell layer with no DDX1 detected in granule cell progenitors. These progenitor cells exit the cell cycle at approximately E10.5 so are not proliferating at P0 (306). The granule progenitor cells re-enter the cell cycle at P4 and greatly increase in number, ultimately making up 2/3 of the total number of neurons in the brain (307). In the adult mouse brain, DDX1 was found in the nucleus of the granule cell neurons in addition to the Purkinje cells. Intriguingly, granule cell progenitors are very susceptible to DNA damage (308-310). The absence of DDX1 in these cells at P0 may contribute to the susceptibility of these cells to DNA damage. It would be interesting to examine DDX1 expression at additional developmental stages to determine when DDX1 is first expressed in granule cells.

A striking difference in the localization of DDX1 was observed in the ovary. The cortical stroma cells were found to have both nuclear and cytoplasmic DDX1 while the cells of the theca externa (cells directly surrounding the follicles) showed either low or no DDX1 expression. The follicles containing the granulosa cells were strongly positive

for DDX1 in both the nucleus and cytoplasm. The granulosa cells of the follicle undergo multiple rounds of proliferation during follicle maturation, but the individual oocyte within each follicle only increases in size. The oocytes also had very high levels of DDX1, but in contrast to the granulose cells of the follicle, DDX1 was primarily observed in the cytoplasm in the cells. Large aggregates were observed in the developing oocyte and in the GV and M stages, with changes in numbers and sizes noted during oocyte maturation. In the future, earlier stage oocytes could be collected from young mice in order to determine whether DDX1 aggregates are present at all stages of oogenesis or only in the later stage oocytes when transcription is halted prior to ovulation. Collecting oocytes for immunoprecipitation or RNA immunoprecipitation studies followed by mass spectrometry or RNA sequencing would shed light on the protein-protein or protein-RNA interactions that DDX1 makes during oogenesis.

6.5 DDX1 forms large aggregates in early stage embryos

The localization of DDX1 in pre-implantation embryos was unexpected as all previous staining in cell lines and tissues, with the exception of *DDX1*-amplified retinoblastoma and neuroblastoma cells, showed a predominant localization of DDX1 to the nucleus with little cytoplasmic staining (85,96,101). Here, we report that DDX1 is primarily a cytoplasmic protein in early stage embryos with a very dynamic localization pattern as demonstrated by the localization changes from meiotically maturing oocytes to cells of the mature blastocyst. We found that DDX1 localizes to

cytoplasmic foci or aggregates in the maturing oocyte; however, DDX1 localization begins to change around the time of fertilization. For example, the number of DDX1 foci/aggregates goes down in the 2-cell embryo and DDX1 aggregates tend to be found closer to the cell membrane than the nucleus. The transition from the 4-cell embryo to the blastocyst is accompanied by further changes, with fewer but larger DDX1 aggregates. During this time, the localization of the DDX1-containing aggregates shifts from close to the cell membrane to where the bulk of the DDX1 is located on one side of the nucleus. It would be interesting to quantitate levels of DDX1 protein from fertilization to the blastocyst stage. For this experiment to be feasible, we would have to collect several hundred embryos.

As the role of DDX1 in early development is not known and DDX1 is found in large foci or aggregates in the cytoplasm, we examined whether DDX1 aggregates co-localized with cellular organelles. This analysis was carried out with markers for endoplasmic reticulum, ribosomes, mitochondria and Golgi. We were especially interested in the possibility that DDX1 might co-localize with mitochondria as the mitochondria in the pre-implantation embryo develop in a hypoxic environment, and are different in their appearance and activity from the mitochondria that develop in older embryos (311,312). Using Mitotracker Orange, we found that mitochondria in 2-cell stage embryos had a similar appearance to that of DDX1, forming aggregates throughout the cytoplasm; however, there were more mitochondria surrounding the nucleus than DDX1 aggregates. Similarly, despite similar staining patterns for DDX1 and mitochondria at the morula and blastocyst stages of development, we did not

observe any co-localization. The ribosomes were also of particular interest as DDX1 was previously shown to be associated with RNA granules and even RNA granules containing ribosomes (112). However, the ribosomes in pre-implantation embryos were found in small punctate foci which did not overlap with DDX1 aggregates. DDX1 also did not co-localize with endoplasmic reticulum and Golgi markers.

DDX1 has previously been shown to co-localize with proteins associated with nuclear bodies, such as CstF64 (cleavage bodies) and RIF1 (DDX1 bodies) (101,102,124) (Lei Li, unpublished data) and to reside adjacent to Cajal bodies and gems (SMN). Cleavage bodies are sites of 3' mRNA processing and polyadenylation (101-103). The Cajal bodies are sites of snRNP and snoRNP biogenesis as well as pre-rRNA and histone mRNA processing (104-106). We were interested in determining whether components of cleavage bodies, Cajal bodies and gems might also relocate to the cytoplasm during early embryonic development. To address this possibility, we carried out co-immunostaining analysis using anti-DDX1 antibody in conjunction with antibodies to proteins normally associated with cleavage bodies, Cajal bodies and gems. We found no evidence of co-localization of DDX1 with CstF64 (cleavage bodies), SMN (Cajal bodies, gems), or Rif1. CstF64 was located in the nucleus of pre-implantation embryos as well as the scaffolding protein Rif1. These results indicate that the cytoplasmic localization of DDX1 in early stage embryos is unlikely to be related to the roles that it plays in nuclear bodies such as nuclear mRNA processing.

A significant amount of mRNA is stored in the developing oocyte (80 picograms). mRNA in the oocyte is highly regulated, both in translational activation

and degradation (313). RNP complexes found during mouse development, such as P-bodies, germ cell granules and stress granules, are poorly understood unlike their counterparts in *Xenopus* and *Drosophila* mainly due to the difficultly in acquiring sufficient material for study (314,315). Only recently, it was found that P-bodies are not present at all stages of mouse development and are actually lost during meiotic maturation and only reappear at the blastocyst stage (316). Some of the components of P-bodies were found to localize to another RNA-containing body in the cortex of the oocyte during early stages of oogenesis when P-bodies are lost (316). The primary difference between these 'alternative' P-bodies and P-bodies is the absence of proteins associated with degradation such as the decapping protein DCP1A. Based on these data, it has been proposed that the alternative P-bodies (also known as subcortical RNP domains or SCRD) are likely sites of RNA storage rather than RNA degradation. By staining RNAs in the 2-cell embryo with acridine orange, we observed large foci; however, these foci did not co-localize with DDX1 and were found primarily near the nucleus of the cell.

The DDX1 staining pattern in early-stage embryos is unlike that of other known RNA binding proteins expressed during embryonic development such as MSY2 (317). We also did not observe any co-localization between DDX1 and either the exosome (EXOSC5) or P-bodies (GW182). The staining of GW182 has not been previously reported in the 2-cell embryo and other markers may be required to show the lack of co-localization with P-bodies, such as DCP1A (316,318). Although there was no co-localization of DDX1 and GW182 in 2-cell stage embryos, it remains possible that these

two proteins do co-localize at later stages of development when P-bodies first appear. It would also be worthwhile to carry out co-immunostaining experiments with markers for SCRD or alternative P-bodies (DDX6) as well as for other RNP structures found during embryonic development such as those containing Zar1-like protein. Zar1-like protein has been shown to associate with P-body proteins during preimplantation development and its overexpression results in developmental arrest at the 2-cell stage (319). Zar1-l foci are quite large and are also found throughout the cytoplasm, with a staining pattern that is similar to that of DDX1 at the 2-cell stage. Zar1-like protein has been proposed as another alternative mRNA storage granule. At this time, we cannot exclude the possibility that DDX1 is part of one or more of these alternative mRNA storage complexes.

We also examined whether DDX1 aggregates might be affected by inhibition of transcription or translation. Translation inhibition, which had been shown to affect P-body size and formation (318), had no effect on the localization pattern of DDX1. Treating the cells with cordycepin for 12 hours to inhibit transcription caused the aggregates to increase in size and lose their even distribution pattern. This supports the hypothesis that DDX1 interacts with some species of RNA at early developmental stages. Note that the long incubation time with cordycepin was to ensure that transcription was actively inhibited as MZT is occurring at the 2-cell stage with transcription taking place in waves during this stage of development (179,320,321). However, a caveat of doing such a long incubation with cordycepin is that it could affect overall levels of RNA as well as prevent additional synthesis of DDX1 which may

be required to maintain aggregate structure as we have found that DDX1 is first expressed at the 2-cell stage of development. The shift towards larger granules upon cordycepin treatment suggests that transcription inhibition alters the composition of the DDX1 aggregates. Perhaps an even longer incubation with cordycepin would result in a DDX1 staining pattern resembling that observed in the 4-cell embryo.

Our most striking results were obtained upon digesting 2-cell embryos with RNase A. This treatment almost completely abolished the DDX1 aggregates. Additional experiments at both later and earlier time points are required to determine whether all DDX1 foci/aggregates have an RNA component. If so, this may suggest that DDX1 is playing a RNA-protecting role during early development, perhaps by stabilizing or extending the half-life of specialized RNAs in the aggregates. The suggestion that DDX1 plays a specialized RNA-protecting role comes from the fact that DDX1 aggregates are present post-fertilization when most maternal mRNAs are already degraded. If this is the case, one may postulate that the interaction partners of DDX1 in the aggregates inhibit its ribonuclease activity. In this regard, it is interesting that the size and number of DDX1 aggregates in the cytoplasm undergoes dramatic changes from the maturing oocyte to the blastocysts, with the gradual increase in the size of the aggregates accompanied by a concomitant decrease in the number of aggregates suggesting that the amount of DDX1 may remain the same as development progresses.

A dramatic change in the localization of DDX1 occurs during morula and blastula formation. The DDX1 protein found in the blastomeres assembles in much

larger aggregates as the embryo approaches late E3.5. In many cases, the DDX1 signal appears to be predominantly a single large complex located adjacent to the nucleus. This complex may contain DDX1 which is being targeted for degradation as DDX1 may no longer be required in as high quantities as during pre-implantation development. Alternatively, DDX1 may play a specific role involving general RNA metabolism. We have yet to carry out co-localization experiments to identify any proteins that may be associated with the very large DDX1 aggregates in E3.5 embryos. It would be interesting to compare the proteins associating with DDX1 in different stage embryos to determine whether DDX1 plays different roles in 2-cell stage embryos versus blastocysts. In addition, cell lines that predominantly express cytoplasmic DDX1 could be used to identify proteins that bind cytoplasmic DDX1. Any identified partners could then be examined in pre-implantation embryos. Although not conclusive, these data suggest that DDX1 is neither associated with RNA degradation in early-stage embryos, nor is it associated with sites of active transcription or translation.

6.6 Wild-type lethality observed in the *Ddx1* gene-trapped mouse line

Using our *Ddx1* gene-trapped mouse line, we found that both knockout and wild-type mice displayed lethality. Intriguingly, only wild-type mice that were derived from second generation or later heterozygous parents displayed this lethality (Figure 6.1C), as well as a portion of wild-type mice derived from abnormal heterozygote backcrosses (Figure 6.1D). The wild-type embryos die between E4.5 and E6.5. Strain

background is irrelevant as similar observations were made in both the FVB and C57BL/6 strains.

Our observations regarding the inheritance pattern of the wild-type lethality indicate that it occurs in a non-Mendelian fashion. Two types of non-Mendelian inheritance have been previously reported: paramutation and imprinting (322,323). Paramutation is a type of epigenetic inheritance where the two alleles of a single gene interact to affect overall gene expression. In some cases, the process occurs through an intermediate such as an RNA molecule (see *Kit* knockout mouse described below) (165). Imprinting is the most widely studied mechanism of non-Mendelian inheritance. Imprinted genes are generally clustered and expressed from a single allele due to epigenetic silencing of the other allele. Silencing is primarily the result of DNA methylation and the regulation of the methylation pattern is influenced by either the maternal or paternal genome. Based on our data, *Ddx1* cannot be classified as an imprinted gene as it is biallelically expressed and wild-type lethality is inherited from either the male or female parent. Thus, the mode of inheritance resembles that of paramutation; however, unlike previously described paramutations, we observe two types of heterozygote mice indicating a two stage inheritance mechanism or a more complex inheritance mechanism.

While we were able to clearly delineate the inheritance pattern underlying the lethality associated with the *Ddx1* wild-type genotype, we have yet to find the cause of the wild-type embryonic lethality. Furthermore, we have yet to determine by what mechanism the trait is passed on from one generation to the next. The fact that there

are two different types of heterozygote mice in our *Ddx1* mouse strain suggests that the mechanism of inheritance is directly associated with the DNA. The inheritance pattern that we observe is different from the paramutation-like effect found in the *c-Kit* mouse strain where a miRNA species was being transmitted through multiple generations resulting in gene silencing even in the wild-type mice (165). This miRNA caused all heterozygotes to transmit the effect (white spots in their fur, white tail tips) and the affected wild-type mice in turn continued to pass on this miRNA for several generations. Since the inheritance pattern of the $Ddx1^{S}$ allele appears to be cis-acting in nature, we analysed the methylation status of the *Ddx1* gene in progeny derived from different *Ddx1* heterozygote crosses. Preliminary methylation analysis of the *Ddx1* gene showed no significant difference between any of the genotypes tested. Other types of base modifications may explain the observed wild-type lethality. Alternatively, there may be changes in base methylation which reside outside the region selected for analysis. The unusual phenomenon observed upon crossing *Ddx1* heterozygote mice needs to be analysed further as it may provide insight into another mechanism of non-Mendelian inheritance.

Previous data in the lab suggest that overexpression of DDX1 may have detrimental effects on development. For example, all attempts to generate transgenic mice overexpressing DDX1 have failed, with DDX1 levels in any surviving lines being similar to that found in wild-type mice (our unpublished data). Similarly, while it is possible to transiently overexpress DDX1 in cell lines, stable lines generated by transfecting DDX1 expression constructs consistently express similar levels of DDX1 as

their non-transfected counterparts (our unpublished data). Thus, it is possible that wild-type embryos from heterozygote crosses may be dying because of DDX1 overexpression. Analysis of DDX1 RNA and protein levels in *Ddx1* heterozygote mice showed similar levels of DDX1 despite these mice having a single normal *Ddx1* allele. These mice were also found to express β-gal from the gene-trapped allele. The presence of products from both the wild-type and fusion alleles indicates that DDX1 is bi-allelically expressed. Together, these data suggest that DDX1 levels are tightly regulated and the single wild-type allele is compensating for absence of product from the mutant allele in the heterozygote mice, such that similar levels of DDX1 are produced in heterozygote and wild-type mice.

We propose that the first generation heterozygotes (derived from a backcross) have undergone an initial epigenetic change on the *Ddx1* wild-type allele which regulates overexpression from this single allele. This mark is not permanent going to the next generation as wild-type offspring survive at the expected rates and likely remove the mark in order to produce the normal levels of DDX1 from both alleles. However, the heterozygote offspring in the second generation are in some way altered so that they are no longer reset to normal DDX1 levels. Thus, when these second generation heterozygote mice are mated to other second generation heterozygotes they produce virtually no wild-type mice as these mice inherit two copies of the compensating allele.

Wild-type ($Ddx1^{S/S}$) mice are predicted to overexpress DDX1 which somehow affects embryonic development at the peri-implantation stage leading to death of

wild-type embryos. Wild-type mice generated by back-crosses of *Ddx1*$^{S/-}$ mice are also affected but at a reduced rate as *Ddx1*$^{S/+}$ mice only inherit a single compensating allele. The survival of *Ddx1*$^{S/+}$ wild-type mice may be due to an acceptable range of variability in DDX1 levels. This concept would explain why we only recover 50% of the expected wild-type mice from *Ddx1*$^{S/-}$ backcrosses as compared to *Ddx1*$^{+/-}$ backcrosses. In order to pursue this hypothesis, we would need to culture blastocysts generated from *Ddx1*$^{S/-}$ intercrosses and *Ddx1*$^{+/-}$ intercrosses up to the hatching stage. Hatched blastocysts would then be harvested for RNA and DNA isolation. The embryos would be genotyped and RT-PCR used to determine whether DDX1 is overexpressed in the *Ddx1*$^{S/S}$ mice. Immunofluorescence staining would also be performed to study the subcellular localization of DDX1 in the different categories of wild-type embryos.

Trophoblast cells play a major role in hatching and implantation of mouse embryos. There are several different types of trophoblast cells which are formed at different stages of development. As we observe no resorbed embryos and normal numbers of wild-type mice at E3.5, we hypothesize that the trophoblast cells are affected in the *Ddx1*$^{S/S}$ mice in a manner which results in their inability to implant or induce decidua swelling of the uterus. Knockout of the transcription factor TEAD4 results in failure of trophoblast growth (324). TEAD4 was found to regulate a subset of specific trophectoderm genes including *Cdx2* and *esomesodermin* (*Eomes*). Knockout of either *Cdx2* or *Eomes* in mice also results in peri-implantation lethality and cultured knockout embryos fail to form trophectoderm outgrowths (209,325). By culturing embryos derived from *Ddx1*$^{S/-}$ intercrosses, we could determine whether the *Ddx1*$^{S/S}$

embryos are able to adhere to the dish and form outgrowths. We could also stain embryos for markers of trophoblast development including TEAD4, Cdx2, Eomes. We could also study whether other epiblast-specific markers are perturbed, such as Oct4 and Nanog. Oct4 is believed to repress the expression of genes associated with trophoblast development, whereas Nanog promotes expression of epiblast-specific genes (211,220,326). Further analysis of both the *Ddx1*$^{-/-}$ and *Ddx1*$^{S/S}$ embryonic lethality may lead to a better understanding of the different roles that DDX1 plays during development.

6.7 Concluding remarks

In conclusion, based on the cytoplasmic localization of DDX1 at early stages of development, the embryonic lethality resulting from *Ddx1* knock-out in mouse, and the interaction of DDX1 with RNA in early stage embryos, we propose that DDX1 plays an essential role in RNA metabolism in pre-implantation embryos. The change in the appearance of the DDX1 aggregates from small aggregates in oocytes to the larger organized aggregates at the 2-cell stage suggests that DDX1 may be associated with the specific regulation of some RNAs between the 1- and 4-cell stages. Disruption of DDX1 expression at these stages results in developmental arrest at the 2- or 4-cell stage. The pre-implantation embryo undergoes a significant shift in its RNA pool following fertilization, with most of the maternal mRNAs being degraded and the initiation of embryonic transcription (177,180,183). DDX1 may play a role in regulating

specific transcripts or even specific types of RNAs, by either facilitating or preventing their destruction.

While we have shown that DDX1 functions as an ribonuclease *in vitro*, we have no direct evidence that DDX1 has ribonuclease activity *in vivo*. It will be important to determine under what conditions DDX1 might use this activity *in vivo*. If DDX1 can indeed cleave RNAs at specific sites, or degrade small segments of RNA which in turn facilitates digestion by the major degradation complexes such as the RNA exosome or processing bodies, this could have broad implications for our understanding of RNA protection and RNA degradation in the developing embryo.

An unexpected result of this thesis is the wild-type lethality associated with *Ddx1* wild-type allele when inherited from second generation intercrossed heterozygote mice (Figure 4.11). The pattern of inheritance precludes imprinting as wild-type embryonic lethality can be inherited from both parents. Furthermore, paramutations described to date do not conform to our observed pattern of inheritance. The identification of a novel non-Mendelian inheritance pattern resulting in wild-type embryo death is a unique observation that may lead to the identification of a new mechanism for the transmission of information from parent to offspring.

Despite the growing body of work regarding DDX1, we still have more questions than answers regarding its functions and roles in development. Based on its various subcellular localizations in different tissues and cell types, DDX1 likely plays multiple roles in the cell. Previous work supports an important role for DDX1 in the nucleus of the cell, particularly as related to the repair of DNA double-strand breaks.

The work described in this thesis suggests an equally important role for cytoplasmic DDX1 during development. The unusual distribution pattern of DDX1 in pre-implantation embryos and our inability to co-localize DDX1 with a variety of proteins previously associated with DDX1 and RNA metabolism suggests that we have identified a novel DDX1-RNA containing structure which may shed light on how RNA is processed and utilized in early stage embryos.

References

1. Liu, F., Putnam, A. A., and Jankowsky, E. (2014) DEAD-Box Helicases Form Nucleotide-Dependent, Long-Lived Complexes with RNA. *Biochemistry* **53**, 423-433

2. Pyle, A. M. (2011) RNA helicases and remodeling proteins. *Current opinion in chemical biology* **15**, 636-642

3. Jarmoskaite, I., and Russell, R. (2011) DEAD-box proteins as RNA helicases and chaperones. *Wiley interdisciplinary reviews. RNA* **2**, 135-152

4. Pan, C., and Russell, R. (2010) Roles of DEAD-box proteins in RNA and RNP Folding. *RNA biology* **7**, 667-676

5. Jankowsky, E., and Putnam, A. (2010) Duplex unwinding with DEAD-box proteins. *Methods in molecular biology* **587**, 245-264

6. Gustafson, E. A., and Wessel, G. M. (2010) DEAD-box helicases: posttranslational regulation and function. *Biochem Biophys Res Commun* **395**, 1-6

7. Jankowsky, E., and Fairman, M. E. (2007) RNA helicases--one fold for many functions. *Current opinion in structural biology* **17**, 316-324

8. Linder, P. (2006) Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res* **34**, 4168-4180

9. Abdelhaleem, M. (2005) RNA helicases: regulators of differentiation. *Clinical biochemistry* **38**, 499-503

10. Abdelhaleem, M. (2004) Do human RNA helicases have a role in cancer? *Biochim Biophys Acta* **1704**, 37-46

11. Abdelhaleem, M., Maltais, L., and Wain, H. (2003) The human DDX and DHX gene families of putative RNA helicases. *Genomics* **81**, 618-622

12. Lorsch, J. R. (2002) RNA chaperones exist and DEAD box proteins get a life. *Cell* **109**, 797-800

13. Tanner, N. K., and Linder, P. (2001) DExD/H box RNA helicases: from generic motors to specific dissociation functions. *Mol Cell* **8**, 251-262

14. Fuller-Pace, F. V. (1994) RNA helicases: modulators of RNA structure. *Trends Cell Biol* **4**, 271-274

15. Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J., and Slonimski, P. P. (1989) Birth of the D-E-A-D box. *Nature* **337**, 121-122

16. Caruthers, J. M., and McKay, D. B. (2002) Helicase structure and mechanism. *Current opinion in structural biology* **12**, 123-133

17. Umate, P., Tuteja, N., and Tuteja, R. (2011) Genome-wide comprehensive analysis of human helicases. *Communicative & integrative biology* **4**, 118-137

18. Linder, P., and Jankowsky, E. (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nature reviews. Molecular cell biology* **12**, 505-516

19. Anantharaman, V., Koonin, E. V., and Aravind, L. (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res* **30**, 1427-1464

20. Prakash, K., and Tuteja, R. A novel DEAD box helicase Has1p from Plasmodium falciparum: N-terminal is essential for activity. *Parasitol Int* **59**, 271-277

21. Shimada, Y., Fukuda, W., Akada, Y., Ishida, M., Nakayama, J., Imanaka, T., and Fujiwara, S. (2009) Property of cold inducible DEAD-box RNA helicase in hyperthermophilic archaea. *Biochem Biophys Res Commun* **389**, 622-627

22. Trubetskoy, D., Proux, F., Allemand, F., Dreyfus, M., and Iost, I. (2009) SrmB, a DEAD-box helicase involved in Escherichia coli ribosome assembly, is specifically targeted to 23S rRNA in vivo. *Nucleic Acids Res* **37**, 6540-6549

23. Rudolph, M. G., and Klostermeier, D. (2009) The Thermus thermophilus DEAD box helicase Hera contains a modified RNA recognition motif domain loosely connected to the helicase core. *RNA* **15**, 1993-2001

24. Marchat, L. A., Orozco, E., Guillen, N., Weber, C., and Lopez-Camarillo, C. (2008) Putative DEAD and DExH-box RNA helicases families in Entamoeba histolytica. *Gene* **424**, 1-10

25. Zhao, Y., Yu, L., Fu, Q., Chen, W., Jiang, J., Gao, J., and Zhao, S. (2000) Cloning and characterization of human DDX24 and mouse Ddx24, two novel putative DEAD-Box proteins, and mapping DDX24 to human chromosome 14q32. *Genomics* **67**, 351-355

26. Godbout, R., Hale, M., and Bisgrove, D. (1994) A human DEAD box protein with partial homology to heterogeneous nuclear ribonucleoprotein U. *Gene* **138**, 243-245

27. Schutz, P., Bumann, M., Oberholzer, A. E., Bieniossek, C., Trachsel, H., Altmann, M., and Baumann, U. (2008) Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci U S A* **105**, 9564-9569

28. Benz, J., Trachsel, H., and Baumann, U. (1999) Crystal structure of the ATPase domain of translation initiation factor 4A from Saccharomyces cerevisiae--the prototype of the DEAD box protein family. *Structure* **7**, 671-679

29. Story, R. M., Li, H., and Abelson, J. N. (2001) Crystal structure of a DEAD box protein from the hyperthermophile Methanococcus jannaschii. *Proc Natl Acad Sci U S A* **98**, 1465-1470

30. Caruthers, J. M., Johnson, E. R., and McKay, D. B. (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc Natl Acad Sci U S A* **97**, 13080-13085

31. Hilbert, M., Karow, A. R., and Klostermeier, D. (2009) The mechanism of ATP-dependent RNA unwinding by DEAD box proteins. *Biological chemistry* **390**, 1237-1250

32. Banroques, J., Doere, M., Dreyfus, M., Linder, P., and Tanner, N. K. (2010) Motif III in superfamily 2 "helicases" helps convert the binding energy of ATP into a high-affinity RNA binding site in the yeast DEAD-box protein Ded1. *Journal of molecular biology* **396**, 949-966

33. Schutz, P., Karlberg, T., van den Berg, S., Collins, R., Lehtio, L., Hogbom, M., Holmberg-Schiavone, L., Tempel, W., Park, H. W., Hammarstrom, M., Moche, M., Thorsell, A. G., and Schuler, H. (2010) Comparative structural analysis of human DEAD-box RNA helicases. *PloS one* **5**

34. Cordin, O., Tanner, N. K., Doere, M., Linder, P., and Banroques, J. (2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J* **23**, 2478-2487

35. Tanner, N. K. (2003) The newly identified Q motif of DEAD box helicases is involved in adenine recognition. *Cell Cycle* **2**, 18-19

36. Banroques, J., Cordin, O., Doere, M., Linder, P., and Tanner, N. K. (2008) A conserved phenylalanine of motif IV in superfamily 2 helicases is required for cooperative, ATP-dependent binding of RNA substrates in DEAD-box proteins. *Mol Cell Biol* **28**, 3359-3371

37. Mallam, A. L., Del Campo, M., Gilman, B., Sidote, D. J., and Lambowitz, A. M. (2012) Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* **490**, 121-125

38. Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., and Yokoyama, S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell* **125**, 287-300

39. Andersen, C. B., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L., Pedersen, J. S., Seraphin, B., Le Hir, H., and Andersen, G. R. (2006) Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* **313**, 1968-1972

40. Bowers, H. A., Maroney, P. A., Fairman, M. E., Kastner, B., Luhrmann, R., Nilsen, T. W., and Jankowsky, E. (2006) Discriminatory RNP remodeling by the DEAD-box protein DED1. *RNA* **12**, 903-912

41. Tran, E. J., Zhou, Y., Corbett, A. H., and Wente, S. R. (2007) The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. *Mol Cell* **28**, 850-859

42. Del Campo, M., Mohr, S., Jiang, Y., Jia, H., Jankowsky, E., and Lambowitz, A. M. (2009) Unwinding by local strand separation is critical for the function of DEAD-box proteins as RNA chaperones. *Journal of molecular biology* **389**, 674-693

43. Liu, F., Putnam, A., and Jankowsky, E. (2008) ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc Natl Acad Sci U S A* **105**, 20209-20214

44. Yang, Q., and Jankowsky, E. (2005) ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry* **44**, 13591-13601

45. Gururajan, R., Mathews, L., Longo, F. J., and Weeks, D. L. (1994) An3 mRNA encodes an RNA helicase that colocalizes with nucleoli in Xenopus oocytes in a stage-specific manner. *Proc Natl Acad Sci U S A* **91**, 2056-2060

46. Liang, L., Diehl-Jones, W., and Lasko, P. (1994) Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities. *Development* **120**, 1201-1211

47. Andreou, A. Z., and Klostermeier, D. (2014) eIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. *Journal of molecular biology* **426**, 51-61

48. Chen, Y., Potratz, J. P., Tijerina, P., Del Campo, M., Lambowitz, A. M., and Russell, R. (2008) DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proc Natl Acad Sci U S A* **105**, 20203-20208

49. Seal, S. N., Schmidt, A., and Marcus, A. (1983) Eukaryotic initiation factor 4A is the component that interacts with ATP in protein chain initiation. *Proc Natl Acad Sci U S A* **80**, 6562-6565

50. Edery, I., Humbelin, M., Darveau, A., Lee, K. A., Milburn, S., Hershey, J. W., Trachsel, H., and Sonenberg, N. (1983) Involvement of eukaryotic initiation factor 4A in the cap recognition process. *J Biol Chem* **258**, 11398-11403

51. Rogers, G. W., Jr., Komar, A. A., and Merrick, W. C. (2002) eIF4A: the godfather of the DEAD box helicases. *Prog Nucleic Acid Res Mol Biol* **72**, 307-331

52. Rogers, G. W., Jr., Richter, N. J., and Merrick, W. C. (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J Biol Chem* **274**, 12236-12244

53. Montpetit, B., Thomsen, N. D., Helmke, K. J., Seeliger, M. A., Berger, J. M., and Weis, K. (2011) A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature* **472**, 238-242

54. Godbout, R., and Squire, J. (1993) Amplification of a DEAD box protein gene in retinoblastoma cell lines. *Proc Natl Acad Sci U S A* **90**, 7578-7582

55. Kuroda, H., White, P. S., Sulman, E. P., Manohar, C. F., Reiter, J. L., Cohn, S. L., and Brodeur, G. M. (1996) Physical mapping of the DDX1 gene to 340 kb 5' of MYCN. *Oncogene* **13**, 1561-1565

56. Pandita, A., Godbout, R., Zielenska, M., Thorner, P., Bayani, J., and Squire, J. A. (1997) Relational mapping of MYCN and DDXI in band 2p24 and analysis of

amplicon arrays in double minute chromosomes and homogeneously staining regions by use of free chromatin FISH. *Genes Chromosomes Cancer* **20**, 243-252

57. Ponting, C., Schultz, J., and Bork, P. (1997) SPRY domains in ryanodine receptors (Ca(2+)-release channels). *Trends in biochemical sciences* **22**, 193-194

58. Rhodes, D. A., de Bono, B., and Trowsdale, J. (2005) Relationship between SPRY and B30.2 protein domains. Evolution of a component of immune defence? *Immunology* **116**, 411-417

59. Wang, D., Li, Z., Messing, E. M., and Wu, G. (2005) The SPRY domain-containing SOCS box protein 1 (SSB-1) interacts with MET and enhances the hepatocyte growth factor-induced Erk-Elk-1-serum response element pathway. *J Biol Chem* **280**, 16393-16401

60. Yap, M. W., Nisole, S., and Stoye, J. P. (2005) A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Current biology : CB* **15**, 73-78

61. Perron, M. J., Stremlau, M., and Sodroski, J. (2006) Two surface-exposed elements of the B30.2/SPRY domain as potency determinants of N-tropic murine leukemia virus restriction by human TRIM5alpha. *Journal of virology* **80**, 5631-5636

62. Diaz-Griffero, F., Perron, M., McGee-Estrada, K., Hanna, R., Maillard, P. V., Trono, D., and Sodroski, J. (2008) A human TRIM5alpha B30.2/SPRY domain mutant gains the ability to restrict and prematurely uncoat B-tropic murine leukemia virus. *Virology* **378**, 233-242

63. Sivaramakrishnan, G., Sun, Y., Rajmohan, R., and Lin, V. C. (2009) B30.2/SPRY domain in tripartite motif-containing 22 is essential for the formation of distinct nuclear bodies. *FEBS letters* **583**, 2093-2099

64. Ikeda, M., Inoue, F., Ohkoshi, K., Yokoyama, S., Tatemizo, A., Tokunaga, T., and Furusawa, T. (2012) B-box and SPRY domain containing protein (BSPRY) is associated with the maintenance of mouse embryonic stem cell pluripotency and early embryonic development. *The Journal of reproduction and development* **58**, 691-699

65. D'Cruz, A. A., Babon, J. J., Norton, R. S., Nicola, N. A., and Nicholson, S. E. (2013) Structure and function of the SPRY/B30.2 domain proteins involved in innate immunity. *Protein science : a publication of the Protein Society* **22**, 1-10

66. Kovalskyy, D. B., and Ivanov, D. N. (2014) Recognition of the HIV capsid by the TRIM5alpha restriction factor is mediated by a subset of pre-existing conformations of the TRIM5alpha SPRY domain. *Biochemistry*

67. Perfetto, L., Gherardini, P. F., Davey, N. E., Diella, F., Helmer-Citterich, M., and Cesareni, G. (2013) Exploring the diversity of SPRY/B30.2-mediated interactions. *Trends in biochemical sciences* **38**, 38-46

68. Yuan, Y., Fu, C., Chen, H., Wang, X., Deng, W., and Huang, B. R. (2006) The Ran binding protein RanBPM interacts with TrkA receptor. *Neuroscience letters* **407**, 26-31

69. Styhler, S., Nakamura, A., and Lasko, P. (2002) VASA localization requires the SPRY-domain and SOCS-box containing protein, GUSTAVUS. *Developmental cell* **3**, 865-876

70. Park, E. Y., Kwon, O. B., Jeong, B. C., Yi, J. S., Lee, C. S., Ko, Y. G., and Song, H. K. (2010) Crystal structure of PRY-SPRY domain of human TRIM72. *Proteins* **78**, 790-795

71. Godbout, R., Packer, M., and Bie, W. (1998) Overexpression of a DEAD box protein (DDX1) in neuroblastoma and retinoblastoma cell lines. *J Biol Chem* **273**, 21161-21168

72. Villegas, V. M., Hess, D. J., Wildner, A., Gold, A. S., and Murray, T. G. (2013) Retinoblastoma. *Current opinion in ophthalmology* **24**, 581-588

73. Temming, P., Lohmann, D., Bornfeld, N., Sauerwein, W., Goericke, S. L., and Eggert, A. (2012) Current concepts for diagnosis and treatment of retinoblastoma in Germany: aiming for safe tumor control and vision preservation. *Klinische Padiatrie* **224**, 339-347

74. Schaiquevich, P., Ceciliano, A., Millan, N., Taich, P., Villasante, F., Fandino, A. C., Dominguez, J., and Chantada, G. L. (2013) Intra-arterial chemotherapy is more effective than sequential periocular and intravenous chemotherapy as salvage treatment for relapsed retinoblastoma. *Pediatric blood & cancer* **60**, 766-770

75. Pichi, F., Lembo, A., De Luca, M., Hadjistilianou, T., and Nucci, P. (2013) Bilateral retinoblastoma: clinical presentation, management and treatment. *International ophthalmology* **33**, 589-593

76. Luo, C., and Deng, Y. P. (2013) Retinoblastoma: concerning its initiation and treatment. *International journal of ophthalmology* **6**, 397-401

77. Brualla, L., Mayorga, P. A., Fluhs, A., Lallena, A. M., Sempau, J., and Sauerwein, W. (2012) Retinoblastoma external beam photon irradiation with a special 'D'-shaped collimator: a comparison between measurements, Monte Carlo simulation and a treatment planning system calculation. *Physics in medicine and biology* **57**, 7741-7751

78. Honavar, S. G., Shields, C. L., Shields, J. A., Demirci, H., and Naduvilath, T. J. (2001) Intraocular surgery after treatment of retinoblastoma. *Archives of ophthalmology* **119**, 1613-1621

79. Moshfeghi, D. M., Wilson, M. W., Grizzard, S., and Haik, B. G. (2005) Intraocular surgery after treatment of germline retinoblastoma. *Archives of ophthalmology* **123**, 1008-1012

80. Graeber, C. P., Gobin, Y. P., Marr, B. P., Dunkel, I. J., Brodie, S. E., Bornfeld, N., Char, D. H., Folberg, R., Imhof, S. M., Lin, A. Y., Berry, J. L., Al Mesfer, S., Moll, A. C., and Abramson, D. H. (2011) Histopathologic findings of eyes enucleated after treatment with chemosurgery for retinoblastoma. *The open ophthalmology journal* **5**, 1-5

81. Rushlow, D. E., Mol, B. M., Kennett, J. Y., Yee, S., Pajovic, S., Thériault, B. L., Prigoda-Lee, N. L., Spencer, C., Dimaras, H., Corson, T. W., Pang, R., Massey, C., Godbout, R., Jiang, Z., Zacksenhaus, E., Paton, K., Moll, A. C., Houdayer, C., Raizis, A., Halliday, W., Lam, W. L., Boutros, P. C., Lohmann, D., Dorsman, J. C., and Gallie, B. L. (2013) Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *The lancet oncology* **14**, 327-334

82. Squire, J. A., Thorner, P. S., Weitzman, S., Maggi, J. D., Dirks, P., Doyle, J., Hale, M., and Godbout, R. (1995) Co-amplification of MYCN and a DEAD box gene (DDX1) in primary neuroblastoma. *Oncogene* **10**, 1417-1422

83. Cheung, N. K., and Dyer, M. A. (2013) Neuroblastoma: developmental biology, cancer genomics and immunotherapy. *Nature reviews. Cancer* **13**, 397-411

84. Nakagawara, A., and Ohira, M. (2004) Comprehensive genomics linking between neural development and cancer: neuroblastoma as a model. *Cancer Lett* **204**, 213-224

85. Godbout, R., Li, L., Liu, R. Z., and Roy, K. (2007) Role of DEAD box 1 in retinoblastoma and neuroblastoma. *Future Oncol* **3**, 575-587

86. Kitada, K., Aida, S., and Aikawa, S. (2012) Coamplification of multiple regions of chromosome 2, including MYCN, in a single patchwork amplicon in cancer cell lines. *Cytogenet Genome Res* **136**, 30-37

87. Reiter, J. L., and Brodeur, G. M. (1996) High-resolution mapping of a 130-kb core region of the MYCN amplicon in neuroblastomas. *Genomics* **32**, 97-103

88. Weber, A., Starke, S., Bergmann, E., and Christiansen, H. (2006) The coamplification pattern of the MYCN amplicon is an invariable attribute of most MYCN-amplified human neuroblastomas. *Clinical cancer research : an official journal of the American Association for Cancer Research* **12**, 7316-7321

89. Adamovic, T., Trosso, F., Roshani, L., Andersson, L., Petersen, G., Rajaei, S., Helou, K., and Levan, G. (2005) Oncogene amplification in the proximal part of chromosome 6 in rat endometrial adenocarcinoma as revealed by combined BAC/PAC FISH, chromosome painting, zoo-FISH, and allelotyping. *Genes Chromosomes Cancer* **44**, 139-153

90. Fruhwald, M. C., O'Dorisio, M. S., Rush, L. J., Reiter, J. L., Smiraglia, D. J., Wenger, G., Costello, J. F., White, P. S., Krahe, R., Brodeur, G. M., and Plass, C. (2000) Gene amplification in PNETs/medulloblastomas: mapping of a novel amplified gene within the MYCN amplicon. *J Med Genet* **37**, 501-509

91. George, R. E., Kenyon, R., McGuckin, A. G., Kohl, N., Kogner, P., Christiansen, H., Pearson, A. D., and Lunec, J. (1997) Analysis of candidate gene co-amplification with MYCN in neuroblastoma. *Eur J Cancer* **33**, 2037-2042

92. Karlsson, A., Helou, K., Walentinsson, A., Hedrich, H. J., Szpirer, C., and Levan, G. (2001) Amplification of Mycn, Ddx1, Rrm2, and Odc1 in rat uterine endometrial carcinomas. *Genes Chromosomes Cancer* **31**, 345-356

93. Defferrari, R., Tonini, G. P., Conte, M., Papio, F., Sementa, A. R., Valent, A., Schena, F., Perri, P., and Mazzocco, K. (2007) Concomitant DDX1 and MYCN gain in neuroblastoma. *Cancer Lett* **256**, 56-63

94. Weber, A., Imisch, P., Bergmann, E., and Christiansen, H. (2004) Coamplification of DDX1 correlates with an improved survival probability in children with MYCN-amplified human neuroblastoma. *J Clin Oncol* **22**, 2681-2690

95. Balko, J. M., and Arteaga, C. L. (2011) Dead-box or black-box: is DDX1 a potential biomarker in breast cancer? *Breast cancer research and treatment* **127**, 65-67

96. Germain, D. R., Graham, K., Glubrecht, D. D., Hugh, J. C., Mackey, J. R., and Godbout, R. (2011) DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. *Breast cancer research and treatment* **127**, 53-63

97. Taunk, N. K., Goyal, S., Wu, H., Moran, M. S., Chen, S., and Haffty, B. G. (2012) DEAD box 1 (DDX1) expression predicts for local control and overall survival in early stage, node-negative breast cancer. *Cancer* **118**, 888-898

98. Godbout, R., Packer, M., Katyal, S., and Bleoo, S. (2002) Cloning and expression analysis of the chicken DEAD box gene DDX1. *Biochim Biophys Acta* **1574**, 63-71

99. Zinsmaier, K. E., Eberle, K. K., Buchner, E., Walter, N., and Benzer, S. (1994) Paralysis and early death in cysteine string protein mutants of Drosophila. *Science* **263**, 977-980

100. Rafti, F., Scarvelis, D., and Lasko, P. F. (1996) A Drosophila melanogaster homologue of the human DEAD-box gene DDX1. *Gene* **171**, 225-229

101. Bleoo, S., Sun, X., Hendzel, M. J., Rowe, J. M., Packer, M., and Godbout, R. (2001) Association of human DEAD box protein DDX1 with a cleavage stimulation factor involved in 3'-end processing of pre-MRNA. *Mol Biol Cell* **12**, 3046-3059

102. Li, L., Roy, K., Katyal, S., Sun, X., Bleoo, S., and Godbout, R. (2006) Dynamic nature of cleavage bodies and their spatial relationship to DDX1 bodies, Cajal bodies, and gems. *Mol Biol Cell* **17**, 1126-1140

103. Murthy, K. G., and Manley, J. L. (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes & development* **9**, 2672-2683

104. Darzacq, X. J. B. E. V. C. K. A. M. B. E. K. T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *The EMBO Journal* **21**, 2746-2756

105. Jády, B. E. D. X. T. K. E. M. A. G. B. E. K. T. (2003) Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *The EMBO Journal* **22**, 1878-1888

106. Schul, W., van Der Kraan, I., Matera, A. G., van Driel, R., and de Jong, L. (1999) Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. *Mol Biol Cell* **10**, 3815-3824

107. Eliceiri, G. L., and Ryerse, J. S. (1984) Detection of intranuclear clusters of Sm antigens with monoclonal anti-Sm antibodies by immunoelectron microscopy. *Journal of cellular physiology* **121**, 449-451

108. Velma, V., Broome, H. J., and Hebert, M. D. (2012) Regulated specific proteolysis of the Cajal body marker protein coilin. *Chromosoma* **121**, 629-642

109. Andrade, L. E., Tan, E. M., and Chan, E. K. (1993) Immunocytochemical analysis of the coiled body in the cell cycle and during cell proliferation. *Proc Natl Acad Sci U S A* **90**, 1947-1951

110. Liu, Q., and Dreyfuss, G. (1996) A novel nuclear structure containing the survival of motor neurons protein. *EMBO J* **15**, 3555-3565

111. Kanai, Y., Dohmae, N., and Hirokawa, N. (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* **43**, 513-525

112. Miller, L. C., Blandford, V., McAdam, R., Sanchez-Carbente, M. R., Badeaux, F., DesGroseillers, L., and Sossin, W. S. (2009) Combinations of DEAD box proteins distinguish distinct types of RNA: protein complexes in neurons. *Molecular and cellular neurosciences* **40**, 485-495

113. Onishi, H., Kino, Y., Morita, T., Futai, E., Sasagawa, N., and Ishiura, S. (2008) MBNL1 associates with YB-1 in cytoplasmic stress granules. *J Neurosci Res* **86**, 1994-2002

114. Kedersha, N., and Anderson, P. (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem Soc Trans* **30**, 963-969

115. Kedersha, N., and Anderson, P. (2007) Mammalian stress granules and processing bodies. *Methods in enzymology* **431**, 61-81

116. Xu, L., Khadijah, S., Fang, S., Wang, L., Tay, F. P., and Liu, D. X. (2010) The cellular RNA helicase DDX1 interacts with coronavirus nonstructural protein 14 and enhances viral replication. *Journal of virology* **84**, 8571-8583

117. Edgcomb, S. P., Carmel, A. B., Naji, S., Ambrus-Aikelin, G., Reyes, J. R., Saphire, A. C., Gerace, L., and Williamson, J. R. (2012) DDX1 is an RNA-dependent ATPase

involved in HIV-1 Rev function and virus replication. *Journal of molecular biology* **415**, 61-74

118. Li, Y. P., Handberg, K. J., Juul-Madsen, H. R., Zhang, M. F., and Jorgensen, P. H. (2007) Transcriptional profiles of chicken embryo cell cultures following infection with infectious bursal disease virus. *Arch Virol* **152**, 463-478

119. Sunden, Y., Semba, S., Suzuki, T., Okada, Y., Orba, Y., Nagashima, K., Umemura, T., and Sawa, H. (2007) DDX1 promotes proliferation of the JC virus through transactivation of its promoter. *Microbiol Immunol* **51**, 339-347

120. Sunden, Y., Semba, S., Suzuki, T., Okada, Y., Orba, Y., Nagashima, K., Umemura, T., and Sawa, H. (2007) Identification of DDX1 as a JC virus transcriptional control region-binding protein. *Microbiol Immunol* **51**, 327-337

121. Tingting, P., Caiyun, F., Zhigang, Y., Pengyuan, Y., and Zhenghong, Y. (2006) Subproteomic analysis of the cellular proteins associated with the 3' untranslated region of the hepatitis C virus genome in human liver cells. *Biochem Biophys Res Commun* **347**, 683-691

122. Yang, Q., Del Campo, M., Lambowitz, A. M., and Jankowsky, E. (2007) DEAD-box proteins unwind duplexes by local strand separation. *Mol Cell* **28**, 253-263

123. Yang, Q., and Jankowsky, E. (2006) The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nature structural & molecular biology* **13**, 981-986

124. Li, L., Monckton, E. A., and Godbout, R. (2008) A role for DEAD box 1 at DNA double-strand breaks. *Mol Cell Biol* **28**, 6413-6425

125. Rozen, F., Edery, I., Meerovitch, K., Dever, T. E., Merrick, W. C., and Sonenberg, N. (1990) Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol* **10**, 1134-1144

126. Nicholson, A. W. (1999) Function, mechanism and regulation of bacterial ribonucleases. *FEMS microbiology reviews* **23**, 371-390

127. Davis, A. J., and Chen, D. J. (2013) DNA double strand break repair via non-homologous end-joining. *Translational cancer research* **2**, 130-143

128. Jasin, M., and Rothstein, R. (2013) Repair of strand breaks by homologous recombination. *Cold Spring Harbor perspectives in biology* **5**, a012740

129. Anderson, P., and Kedersha, N. (2009) RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* **10**, 430-436

130. Gallo, C. M., Munro, E., Rasoloson, D., Merritt, C., and Seydoux, G. (2008) Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos. *Dev Biol* **323**, 76-87

131. Kotaja, N., and Sassone-Corsi, P. (2007) The chromatoid body: a germ-cell-specific RNA-processing centre. *Nat Rev Mol Cell Biol* **8**, 85-90

132. Anderson, P., and Kedersha, N. (2006) RNA granules. *The Journal of cell biology* **172**, 803-808

133. Anderson, P., and Kedersha, N. (2009) Stress granules. *Curr Biol* **19**, R397-398

134. Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M. J., Scheuner, D., Kaufman, R. J., Golan, D. E., and Anderson, P. (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* **169**, 871-884

135. Kindler, S., Wang, H., Richter, D., and Tiedge, H. (2005) RNA transport and local control of translation. *Annual review of cell and developmental biology* **21**, 223-245

136. Willis, D. E., and Twiss, J. L. (2010) Regulation of protein levels in subcellular domains through mRNA transport and localized translation. *Molecular & cellular proteomics : MCP* **9**, 952-962

137. Gibbings, D. J., Ciaudo, C., Erhardt, M., and Voinnet, O. (2009) Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. *Nature cell biology* **11**, 1143-1149

138. Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007) P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* **27**, 3970-3981

139. Franks, T. M., and Lykke-Andersen, J. (2008) The control of mRNA decapping and P-body formation. *Mol Cell* **32**, 605-615

140. Orban, T. I., and Izaurralde, E. (2005) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* **11**, 459-469

141. Weston, A., and Sommerville, J. (2006) Xp54 and related (DDX6-like) RNA helicases: roles in messenger RNP assembly, translation regulation and RNA degradation. *Nucleic Acids Res* **34**, 3082-3094

142. Flemr, M., Ma, J., Schultz, R. M., and Svoboda, P. P-body loss is concomitant with formation of a messenger RNA storage domain in mouse oocytes. *Biol Reprod* **82**, 1008-1017

143. Aizer, A., Brody, Y., Ler, L. W., Sonenberg, N., Singer, R. H., and Shav-Tal, Y. (2008) The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol Biol Cell* **19**, 4154-4166

144. Mendez, R., and Richter, J. D. (2001) Translational control by CPEB: a means to the end. *Nat Rev Mol Cell Biol* **2**, 521-529

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

146. Anderson, P., and Kedersha, N. (2008) Stress granules: the Tao of RNA triage. *Trends in biochemical sciences* **33**, 141-150

147. Kedersha, N. L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999) RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J Cell Biol* **147**, 1431-1442

148. Mollet, S., Cougot, N., Wilczynska, A., Dautry, F., Kress, M., Bertrand, E., and Weil, D. (2008) Translationally repressed mRNA transiently cycles through stress granules during stress. *Mol Biol Cell* **19**, 4469-4479

149. Hess, R. A., and Renato de Franca, L. (2008) Spermatogenesis and cycle of the seminiferous epithelium. *Advances in experimental medicine and biology* **636**, 1-15

150. Braun, R. E. (1998) Post-transcriptional control of gene expression during spermatogenesis. *Seminars in cell & developmental biology* **9**, 483-489

151. Kleene, K. C. (1993) Multiple controls over the efficiency of translation of the mRNAs encoding transition proteins, protamines, and the mitochondrial capsule selenoprotein in late spermatids in mice. *Dev Biol* **159**, 720-731

152. Schmidt, E. E., Hanson, E. S., and Capecchi, M. R. (1999) Sequence-independent assembly of spermatid mRNAs into messenger ribonucleoprotein particles. *Mol Cell Biol* **19**, 3904-3915

153. Braun, R. E. (1990) Temporal translational regulation of the protamine 1 gene during mouse spermatogenesis. *Enzyme* **44**, 120-128

154. Cho, C., Willis, W. D., Goulding, E. H., Jung-Ha, H., Choi, Y. C., Hecht, N. B., and Eddy, E. M. (2001) Haploinsufficiency of protamine-1 or -2 causes infertility in mice. *Nature genetics* **28**, 82-86

155. Shirley, C. R., Hayashi, S., Mounsey, S., Yanagimachi, R., and Meistrich, M. L. (2004) Abnormalities and reduced reproductive potential of sperm from Tnp1- and Tnp2-null double mutant mice. *Biology of reproduction* **71**, 1220-1229

156. Meistrich, M. L., Brock, W. A., Grimes, S. R., Platz, R. D., and Hnilica, L. S. (1978) Nuclear protein transitions during spermatogenesis. *Federation proceedings* **37**, 2522-2525

157. Balhorn, R., Gledhill, B. L., and Wyrobek, A. J. (1977) Mouse sperm chromatin proteins: quantitative isolation and partial characterization. *Biochemistry* **16**, 4074-4080

158. Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E. J., Roloff, T. C., Beisel, C., Schubeler, D., Stadler, M. B., and Peters, A. H. (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nature structural & molecular biology* **17**, 679-687

159. Gatewood, J. M., Cook, G. R., Balhorn, R., Schmid, C. W., and Bradbury, E. M. (1990) Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. *J Biol Chem* **265**, 20662-20666

160. Banerjee, S., and Smallwood, A. (1998) Chromatin modification of imprinted H19 gene in mammalian spermatozoa. *Molecular reproduction and development* **50**, 474-484

161. Gardiner-Garden, M., Ballesteros, M., Gordon, M., and Tam, P. P. (1998) Histone- and protamine-DNA association: conservation of different patterns within the beta-globin domain in human sperm. *Mol Cell Biol* **18**, 3350-3356

162. Wykes, S. M., and Krawetz, S. A. (2003) The structural organization of sperm chromatin. *J Biol Chem* **278**, 29471-29477

163. Dadoune, J. P. (2009) Spermatozoal RNAs: what about their functions? *Microscopy research and technique* **72**, 536-551

164. Amanai, M., Brahmajosyula, M., and Perry, A. C. (2006) A restricted role for sperm-borne microRNAs in mammalian fertilization. *Biology of reproduction* **75**, 877-884

165. Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006) RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* **441**, 469-474

166. Lalancette, C., Miller, D., Li, Y., and Krawetz, S. A. (2008) Paternal contributions: new functional insights for spermatozoal RNA. *Journal of cellular biochemistry* **104**, 1570-1579

167. Wagner, K. D., Wagner, N., Ghanbarian, H., Grandjean, V., Gounon, P., Cuzin, F., and Rassoulzadegan, M. (2008) RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse. *Developmental cell* **14**, 962-969

168. Worch, S., Hansmann, I., and Schlote, D. (2008) Paramutation-like effects at the mouse scapinin (Phactr3) locus. *Journal of molecular biology* **377**, 605-608

169. Johnson, G. D., Lalancette, C., Linnemann, A. K., Leduc, F., Boissonneault, G., and Krawetz, S. A. (2011) The sperm nucleus: chromatin, RNA, and the nuclear matrix. *Reproduction* **141**, 21-36

170. Kiani, J., and Rassoulzadegan, M. (2013) A load of small RNAs in the sperm - how many bits of hereditary information? *Cell research* **23**, 18-19

171. De Leon, V., Johnson, A., and Bachvarova, R. (1983) Half-lives and relative amounts of stored and polysomal ribosomes and poly(A) + RNA in mouse oocytes. *Dev Biol* **98**, 400-408

172. Li, L., Zheng, P., and Dean, J. (2010) Maternal control of early mouse development. *Development* **137**, 859-870

173. Gardner, A. J., Williams, C. J., and Evans, J. P. (2007) Establishment of the mammalian membrane block to polyspermy: evidence for calcium-dependent and -independent regulation. *Reproduction* **133**, 383-393

174. Gardner, A. J., and Evans, J. P. (2006) Mammalian membrane block to polyspermy: new insights into how mammalian eggs prevent fertilisation by multiple sperm. *Reproduction, fertility, and development* **18**, 53-61

175. Tosti, E. (2006) Calcium ion currents mediating oocyte maturation events. *Reproductive biology and endocrinology : RB&E* **4**, 26

176. Telford, N. A., Watson, A. J., and Schultz, G. A. (1990) Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Molecular reproduction and development* **26**, 90-100

177. Schultz, R. M. (1993) Regulation of zygotic gene activation in the mouse. *BioEssays : news and reviews in molecular, cellular and developmental biology* **15**, 531-538

178. Schultz, R. M. (2002) The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Human reproduction update* **8**, 323-331

179. Tadros, W., and Lipshitz, H. D. (2009) The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033-3042

180. Piko, L., and Clegg, K. B. (1982) Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Dev Biol* **89**, 362-378

181. Lykke-Andersen, K., Gilchrist, M. J., Grabarek, J. B., Das, P., Miska, E., and Zernicka-Goetz, M. (2008) Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. *Mol Biol Cell* **19**, 4383-4392

182. Alizadeh, Z., Kageyama, S., and Aoki, F. (2005) Degradation of maternal mRNA in mouse embryos: selective degradation of specific mRNAs after fertilization. *Molecular reproduction and development* **72**, 281-290

183. Hamatani, T., Carter, M. G., Sharov, A. A., and Ko, M. S. (2004) Dynamics of global gene expression changes during mouse preimplantation development. *Developmental cell* **6**, 117-131

184. Perreault, S. D. (1992) Chromatin remodeling in mammalian zygotes. *Mutation research* **296**, 43-55

185. Worrad, D. M., Ram, P. T., and Schultz, R. M. (1994) Regulation of gene expression in the mouse oocyte and early preimplantation embryo: developmental changes in Sp1 and TATA box-binding protein, TBP. *Development* **120**, 2347-2357

186. Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000) Active demethylation of the paternal genome in the mouse zygote. *Current biology : CB* **10**, 475-478

187. Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000) Demethylation of the zygotic paternal genome. *Nature* **403**, 501-502

188. Masui, Y., and Wang, P. (1998) Cell cycle transition in early embryonic development of Xenopus laevis. *Biology of the cell / under the auspices of the European Cell Biology Organization* **90**, 537-548

189. Howe, J. A., and Newport, J. W. (1996) A developmental timer regulates degradation of cyclin E1 at the midblastula transition during Xenopus embryogenesis. *Proc Natl Acad Sci U S A* **93**, 2060-2064

190. Howe, J. A., Howell, M., Hunt, T., and Newport, J. W. (1995) Identification of a developmental timer regulating the stability of embryonic cyclin A and a new somatic A-type cyclin at gastrulation. *Genes & development* **9**, 1164-1176

191. Lu, X., Li, J. M., Elemento, O., Tavazoie, S., and Wieschaus, E. F. (2009) Coupling of zygotic transcription to mitotic control at the Drosophila mid-blastula transition. *Development* **136**, 2101-2110

192. Kimelman, D., Kirschner, M., and Scherson, T. (1987) The events of the midblastula transition in Xenopus are regulated by changes in the cell cycle. *Cell* **48**, 399-407

193. Nothias, J. Y., Majumder, S., Kaneko, K. J., and DePamphilis, M. L. (1995) Regulation of gene expression at the beginning of mammalian development. *J Biol Chem* **270**, 22077-22080

194. Nothias, J. Y., Miranda, M., and DePamphilis, M. L. (1996) Uncoupling of transcription and translation during zygotic gene activation in the mouse. *EMBO J* **15**, 5715-5725

195. Martinez-Salas, E., Linney, E., Hassell, J., and DePamphilis, M. L. (1989) The need for enhancers in gene expression first appears during mouse development with formation of the zygotic nucleus. *Genes & development* **3**, 1493-1506

196. Zheng, P., and Dean, J. (2007) Oocyte-specific genes affect folliculogenesis, fertilization, and early development. *Seminars in reproductive medicine* **25**, 243-251

197. Wu, X., Viveiros, M. M., Eppig, J. J., Bai, Y., Fitzpatrick, S. L., and Matzuk, M. M. (2003) Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nature genetics* **33**, 187-191

198. Torres-Padilla, M. E., and Zernicka-Goetz, M. (2006) Role of TIF1alpha as a modulator of embryonic transcription in the mouse zygote. *The Journal of cell biology* **174**, 329-338

199. Morison, I. M., Ramsay, J. P., and Spencer, H. G. (2005) A census of mammalian imprinting. *Trends in genetics : TIG* **21**, 457-465

200. Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., and Li, E. (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**, 3195-3205

201. Li, E., Bestor, T. H., and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915-926

202. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257

203. Brink, R. A. (1958) Paramutation at the R locus in maize. *Cold Spring Harbor symposia on quantitative biology* **23**, 379-391

204. Goff, S. A., Klein, T. M., Roth, B. A., Fromm, M. E., Cone, K. C., Radicella, J. P., and Chandler, V. L. (1990) Transactivation of anthocyanin biosynthetic genes following transfer of B regulatory genes into maize tissues. *EMBO J* **9**, 2517-2522

205. Brzeski, J., and Brzeska, K. (2011) The maze of paramutation: a rough guide to the puzzling epigenetics of paramutation. *Wiley interdisciplinary reviews. RNA* **2**, 863-874

206. Chandler, V. L. (2007) Paramutation: from maize to mice. *Cell* **128**, 641-645

207. de Vanssay, A., Bouge, A. L., Boivin, A., Hermant, C., Teysset, L., Delmarre, V., Antoniewski, C., and Ronsseray, S. (2012) Paramutation in Drosophila linked to emergence of a piRNA-producing locus. *Nature* **490**, 112-115

208. Larue, L., Ohsugi, M., Hirchenhain, J., and Kemler, R. (1994) E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc Natl Acad Sci U S A* **91**, 8263-8267

209. Strumpf, D., Mao, C. A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005) Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**, 2093-2102

210. Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & development* **17**, 126-140

211. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655

212. Frum, T., Halbisen, M. A., Wang, C., Amiri, H., Robson, P., and Ralston, A. (2013) Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Developmental cell* **25**, 610-622

213. Fujimori, T. (2010) Preimplantation development of mouse: a view from cellular behavior. *Development, growth & differentiation* **52**, 253-262

214. Cockburn, K., and Rossant, J. (2010) Making the blastocyst: lessons from the mouse. *The Journal of clinical investigation* **120**, 995-1003

215. O'Sullivan, C. M., Liu, S. Y., Karpinka, J. B., and Rancourt, D. E. (2002) Embryonic hatching enzyme strypsin/ISP1 is expressed with ISP2 in endometrial glands during implantation. *Molecular reproduction and development* **62**, 328-334

216. Seshagiri, P. B., Sen Roy, S., Sireesha, G., and Rao, R. P. (2009) Cellular and molecular regulation of mammalian blastocyst hatching. *Journal of reproductive immunology* **83**, 79-84

217. Gordon, J. W., and Dapunt, U. (1993) A new mouse model for embryos with a hatching deficiency and its use to elucidate the mechanism of blastocyst hatching. *Fertility and sterility* **59**, 1296-1301

218. Morrisey, E. E., Tang, Z., Sigrist, K., Lu, M. M., Jiang, F., Ip, H. S., and Parmacek, M. S. (1998) GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes & development* **12**, 3579-3590

219. Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C. A., Narita, N., Saffitz, J. E., Simon, M. C., Leiden, J. M., and Wilson, D. B. (1995) Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. *Development* **121**, 3877-3888

220. Silva, J., Nichols, J., Theunissen, T. W., Guo, G., van Oosten, A. L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009) Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-737

221. Kwon, G. S., Viotti, M., and Hadjantonakis, A. K. (2008) The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Developmental cell* **15**, 509-520

222. Tanaka, S. S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., and Noce, T. (2000) The mouse homolog of Drosophila Vasa is required for the development of male germ cells. *Genes & development* **14**, 841-853

223. Tsai-Morris, C. H., Sheng, Y., Lee, E., Lei, K. J., and Dufau, M. L. (2004) Gonadotropin-regulated testicular RNA helicase (GRTH/Ddx25) is essential for spermatid development and completion of spermatogenesis. *Proc Natl Acad Sci U S A* **101**, 6373-6378

224. Kato, H., Sato, S., Yoneyama, M., Yamamoto, M., Uematsu, S., Matsui, K., Tsujimura, T., Takeda, K., Fujita, T., Takeuchi, O., and Akira, S. (2005) Cell type-specific involvement of RIG-I in antiviral response. *Immunity* **23**, 19-28

225. Wang, Y., Zhang, H. X., Sun, Y. P., Liu, Z. X., Liu, X. S., Wang, L., Lu, S. Y., Kong, H., Liu, Q. L., Li, X. H., Lu, Z. Y., Chen, S. J., Chen, Z., Bao, S. S., Dai, W., and Wang, Z. G. (2007) Rig-I-/- mice develop colitis associated with downregulation of G alpha i2. *Cell research* **17**, 858-868

226. Lamm, G. M., Nicol, S. M., Fuller-Pace, F. V., and Lamond, A. I. (1996) p72: a human nuclear DEAD box protein highly related to p68. *Nucleic Acids Res* **24**, 3739-3747

227. Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'Malley, B. W., and Kato, S. (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nature cell biology* **9**, 604-611

228. Inoue, A., Li, T., Roby, S. K., Valentine, M. B., Inoue, M., Boyd, K., Kidd, V. J., and Lahti, J. M. (2007) Loss of ChlR1 helicase in mouse causes lethality due to the accumulation of aneuploid cells generated by cohesion defects and placental malformation. *Cell Cycle* **6**, 1646-1654

229. Mouillet, J. F., Yan, X., Ou, Q., Jin, L., Muglia, L. J., Crawford, P. A., and Sadovsky, Y. (2008) DEAD-box protein-103 (DP103, Ddx20) is essential for early embryonic development and modulates ovarian morphology and function. *Endocrinology* **149**, 2168-2175

230. Zeng, F., and Schultz, R. M. (2005) RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. *Dev Biol* **283**, 40-57

231. Nagy, A. (2003) *Manipulating the mouse embryo : a laboratory manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

232. Kovalchuk, I., and Zemb, F. J. (2010) *Plant epigenetics : methods and protocols.*,

233. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning : a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

234. Rohde, C., Zhang, Y., Reinhardt, R., and Jeltsch, A. (2010) BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. *BMC bioinformatics* **11**, 230

235. Tseng, S. S., Weaver, P. L., Liu, Y., Hitomi, M., Tartakoff, A. M., and Chang, T. H. (1998) Dbp5p, a cytosolic RNA helicase, is required for poly(A)+ RNA export. *EMBO J* **17**, 2651-2662

236. Raghava, G. P. S. (2000) APSSP2 : A combination method for protein secondary structure prediction based on neural network and example based learning. *CASP5*, A-132

237. Hodgson, J. G., Yeh, R. F., Ray, A., Wang, N. J., Smirnov, I., Yu, M., Hariono, S., Silber, J., Feiler, H. S., Gray, J. W., Spellman, P. T., Vandenberg, S. R., Berger, M. S., and James, C. D. (2009) Comparative analyses of gene copy number and mRNA expression in glioblastoma multiforme tumors and xenografts. *Neuro-oncology* **11**, 477-487

238. Amler, L. C., Schurmann, J., and Schwab, M. (1996) The DDX1 gene maps within 400 kbp 5' to MYCN and is frequently coamplified in human neuroblastoma. *Genes Chromosomes Cancer* **15**, 134-137

239. Buis, J., Wu, Y., Deng, Y., Leddon, J., Westfield, G., Eckersdorff, M., Sekiguchi, J. M., Chang, S., and Ferguson, D. O. (2008) Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. *Cell* **135**, 85-96

240. Brown, E. J., and Baltimore, D. (2000) ATR disruption leads to chromosomal fragmentation and early embryonic lethality. *Genes & development* **14**, 397-402

241. Li, L. C., and Dahiya, R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**, 1427-1431

242. Puverel, S., and Tessarollo, L. (2013) RanBPM, a scaffolding protein for gametogenesis. *Current topics in developmental biology* **102**, 357-384

243. Wei, J. D., Kim, J. Y., Kim, A. K., Jang, S. K., and Kim, J. H. (2013) RanBPM protein acts as a negative regulator of BLT2 receptor to attenuate BLT2-mediated cell motility. *J Biol Chem* **288**, 26753-26763

244. Zhang, J., Ma, W., Tian, S., Fan, Z., Ma, X., Yang, X., Zhao, Q., Tan, K., Chen, H., Chen, D., and Huang, B. R. (2014) RanBPM interacts with TbetaRI, TRAF6 and curbs TGF induced nuclear accumulation of TbetaRI. *Cell Signal* **26**, 162-172

245. Chapman, J. R., Barral, P., Vannier, J. B., Borel, V., Steger, M., Tomas-Loba, A., Sartori, A. A., Adams, I. R., Batista, F. D., and Boulton, S. J. (2013) RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol Cell* **49**, 858-871

246. Daley, J. M., and Sung, P. (2013) RIF1 in DNA break repair pathway choice. *Mol Cell* **49**, 840-841

247. Di Virgilio, M., Callen, E., Yamane, A., Zhang, W., Jankovic, M., Gitlin, A. D., Feldhahn, N., Resch, W., Oliveira, T. Y., Chait, B. T., Nussenzweig, A., Casellas, R., Robbiani, D. F., and Nussenzweig, M. C. (2013) Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. *Science* **339**, 711-715

248. Feng, L., Fong, K. W., Wang, J., Wang, W., and Chen, J. (2013) RIF1 counteracts BRCA1-mediated end resection during DNA repair. *J Biol Chem* **288**, 11135-11143

249. Shi, T., Bunker, R. D., Mattarocci, S., Ribeyre, C., Faty, M., Gut, H., Scrima, A., Rass, U., Rubin, S. M., Shore, D., and Thoma, N. H. (2013) Rif1 and Rif2 shape telomere function and architecture through multivalent Rap1 interactions. *Cell* **153**, 1340-1353

250. Yamazaki, S., Hayano, M., and Masai, H. (2013) Replication timing regulation of eukaryotic replicons: Rif1 as a global regulator of replication timing. *Trends in genetics : TIG* **29**, 449-460

251. Zimmermann, M., Lottersberger, F., Buonomo, S. B., Sfeir, A., and de Lange, T. (2013) 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* **339**, 700-704

252. Zou, T., Yang, X., Pan, D., Huang, J., Sahin, M., and Zhou, J. (2011) SMN deficiency reduces cellular ability to form stress granules, sensitizing cells to stress. *Cellular and molecular neurobiology* **31**, 541-550

253. Hua, Y., and Zhou, J. (2004) Rpp20 interacts with SMN and is re-distributed into SMN granules in response to stress. *Biochem Biophys Res Commun* **314**, 268-276

254. Choi, Y. J., and Lee, S. G. (2012) The DEAD-box RNA helicase DDX3 interacts with DDX5, co-localizes with it in the cytoplasm during the G2/M phase of the cycle, and affects its shuttling during mRNP export. *Journal of cellular biochemistry* **113**, 985-996

255. Pek, J. W., and Kai, T. (2011) DEAD-box RNA helicase Belle/DDX3 and the RNA interference pathway promote mitotic chromosome segregation. *Proc Natl Acad Sci U S A* **108**, 12007-12012

256. Soto-Rifo, R., and Ohlmann, T. (2013) The role of the DEAD-box RNA helicase DDX3 in mRNA metabolism. *Wiley interdisciplinary reviews. RNA* **4**, 369-385

257. Soto-Rifo, R., Rubilar, P. S., and Ohlmann, T. (2013) The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Res* **41**, 6286-6299

258. Lykke-Andersen, S., Tomecki, R., Jensen, T. H., and Dziembowski, A. (2011) The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. *RNA biology* **8**, 61-66

259. Chlebowski, A., Lubas, M., Jensen, T. H., and Dziembowski, A. (2013) RNA decay machines: the exosome. *Biochim Biophys Acta* **1829**, 552-560

260. Jagannath, A., and Wood, M. J. (2009) Localization of double-stranded small interfering RNA to cytoplasmic processing bodies is Ago2 dependent and results in up-regulation of GW182 and Argonaute-2. *Mol Biol Cell* **20**, 521-529

261. Li, S., Lian, S. L., Moser, J. J., Fritzler, M. L., Fritzler, M. J., Satoh, M., and Chan, E. K. (2008) Identification of GW182 and its novel isoform TNGW1 as translational repressors in Ago2-mediated silencing. *Journal of cell science* **121**, 4134-4144

262. Shen, X. H., Han, Y. J., Cui, X. S., and Kim, N. H. (2010) Ago2 and GW182 expression in mouse preimplantation embryos: a link between microRNA biogenesis and GW182 protein synthesis. *Reproduction, fertility, and development* **22**, 634-643

263. Yao, B., Li, S., Lian, S. L., Fritzler, M. J., and Chan, E. K. (2011) Mapping of Ago2-GW182 functional interactions. *Methods in molecular biology* **725**, 45-62

264. Fan, J. S., Cheng, Z., Zhang, J., Noble, C., Zhou, Z., Song, H., and Yang, D. (2009) Solution and crystal structures of mRNA exporter Dbp5p and its interaction with nucleotides. *Journal of molecular biology* **388**, 1-10

265. Mallam, A. L., Jarmoskaite, I., Tijerina, P., Del Campo, M., Seifert, S., Guo, L., Russell, R., and Lambowitz, A. M. (2011) Solution structures of DEAD-box RNA chaperones reveal conformational changes and nucleic acid tethering by a basic tail. *Proc Natl Acad Sci U S A* **108**, 12254-12259

266. Bono, F., Ebert, J., Lorentzen, E., and Conti, E. (2006) The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126**, 713-725

267. Savchenko, A., Yee, A., Khachatryan, A., Skarina, T., Evdokimova, E., Pavlova, M., Semesi, A., Northey, J., Beasley, S., Lan, N., Das, R., Gerstein, M., Arrowmith, C. H., and Edwards, A. M. (2003) Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches. *Proteins* **50**, 392-399

268. Andrykovitch, M., Routzahn, K. M., Li, M., Gu, Y., Waugh, D. S., and Ji, X. (2003) Characterization of four orthologs of stringent starvation protein A. *Acta crystallographica. Section D, Biological crystallography* **59**, 881-886

269. McPherson, A. (1990) Current approaches to macromolecular crystallization. *European journal of biochemistry / FEBS* **189**, 1-23

270. Benvenuti, M., and Mangani, S. (2007) Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. *Nature protocols* **2**, 1633-1651

271. Christopher, G. K., Phipps, A. G., and Gray, R. J. (1998) Temperature-dependent solubility of selected proteins. *Journal of Crystal Growth* **191**, 820-826

272. Zhu, D. W., Garneau, A., Mazumdar, M., Zhou, M., Xu, G. J., and Lin, S. X. (2006) Attempts to rationalize protein crystallization using relative crystallizability. *Journal of structural biology* **154**, 297-302

273. Wernimont, A., and Edwards, A. (2009) In situ proteolysis to generate crystals for structure determination: an update. *PloS one* **4**, e5094

274. Dong, A., Xu, X., Edwards, A. M., Midwest Center for Structural, G., Structural Genomics, C., Chang, C., Chruszcz, M., Cuff, M., Cymborowski, M., Di Leo, R., Egorova, O., Evdokimova, E., Filippova, E., Gu, J., Guthrie, J., Ignatchenko, A., Joachimiak, A., Klostermann, N., Kim, Y., Korniyenko, Y., Minor, W., Que, Q., Savchenko, A., Skarina, T., Tan, K., Yakunin, A., Yee, A., Yim, V., Zhang, R., Zheng, H., Akutsu, M., Arrowsmith, C., Avvakumov, G. V., Bochkarev, A., Dahlgren, L. G., Dhe-Paganon, S., Dimov, S., Dombrovski, L., Finerty, P., Jr., Flodin, S., Flores, A., Graslund, S., Hammerstrom, M., Herman, M. D., Hong, B. S., Hui, R., Johansson, I., Liu, Y., Nilsson, M., Nedyalkova, L., Nordlund, P., Nyman, T., Min, J., Ouyang, H., Park, H. W., Qi, C., Rabeh, W., Shen, L., Shen, Y., Sukumard, D., Tempel, W., Tong, Y., Tresagues, L., Vedadi, M., Walker, J. R., Weigelt, J., Welin, M., Wu, H., Xiao, T.,

Zeng, H., and Zhu, H. (2007) In situ proteolysis for protein crystallization and structure determination. *Nature methods* **4**, 1019-1021

275. Bai, Y., Auperin, T. C., and Tong, L. (2007) The use of in situ proteolysis in the crystallization of murine CstF-77. *Acta crystallographica. Section F, Structural biology and crystallization communications* **63**, 135-138

276. Mandel, C. R., Gebauer, D., Zhang, H., and Tong, L. (2006) A serendipitous discovery that in situ proteolysis is essential for the crystallization of yeast CPSF-100 (Ydh1p). *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 1041-1045

277. Tong, Y., Dong, A., Xu, X., and Wernimont, A. (2014) Salvage or recovery of failed targets by in situ proteolysis. *Methods in molecular biology* **1140**, 179-188

278. Civril, F., and Hopfner, K. P. (2014) Crystallization of Mouse RIG-I ATPase Domain: In Situ Proteolysis. *Methods in molecular biology* **1169**, 27-35

279. Weirich, C. S., Erzberger, J. P., Flick, J. S., Berger, J. M., Thorner, J., and Weis, K. (2006) Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. *Nature cell biology* **8**, 668-676

280. Collins, R., Karlberg, T., Lehtio, L., Schutz, P., van den Berg, S., Dahlgren, L. G., Hammarstrom, M., Weigelt, J., and Schuler, H. (2009) The DEXD/H-box RNA helicase DDX19 is regulated by an {alpha}-helical switch. *J Biol Chem* **284**, 10296-10300

281. Chen, H. C., Lin, W. C., Tsay, Y. G., Lee, S. C., and Chang, C. J. (2002) An RNA helicase, DDX1, interacting with poly(A) RNA and heterogeneous nuclear ribonucleoprotein K. *J Biol Chem* **277**, 40403-40409

282. Robertson-Anderson, R. M., Wang, J., Edgcomb, S. P., Carmel, A. B., Williamson, J. R., and Millar, D. P. (2011) Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element. *Journal of molecular biology* **410**, 959-971

283. Fang, J., Kubota, S., Yang, B., Zhou, N., Zhang, H., Godbout, R., and Pomerantz, R. J. (2004) A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev. *Virology* **330**, 471-480

284. Beintema, J. J., and Kleineidam, R. G. (1998) The ribonuclease A superfamily: general discussion. *Cellular and molecular life sciences : CMLS* **54**, 825-832

285. Dyer, K. D., and Rosenberg, H. F. (2006) The RNase a superfamily: generation of diversity and innate host defense. *Molecular diversity* **10**, 585-597

286. Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Johnson, L. N., and Richards, F. M. (1967) The structure of ribonuclease-S at 3.5 A resolution. *J Biol Chem* **242**, 3984-3988

287. Klink, T. A., Woycechowsky, K. J., Taylor, K. M., and Raines, R. T. (2000) Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *European journal of biochemistry / FEBS* **267**, 566-572

288. Raines, R. T. (1998) Ribonuclease A. *Chemical reviews* **98**, 1045-1066

289. Li, W. M., Barnes, T., and Lee, C. H. (2010) Endoribonucleases--enzymes gaining spotlight in mRNA metabolism. *The FEBS journal* **277**, 627-641

290. Eberle, A. B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T. H. (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature structural & molecular biology* **16**, 49-55

291. Glavan, F., Behm-Ansmant, I., Izaurralde, E., and Conti, E. (2006) Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J* **25**, 5117-5125

292. Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998) SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* **95**, 5857-5864

293. Letunic, I., Doerks, T., and Bork, P. (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40**, D302-305

294. Gaur, D., and Batra, J. K. (2005) Role of aspartic acid 121 in human pancreatic ribonuclease catalysis. *Molecular and cellular biochemistry* **275**, 95-101

295. Stern, M. S., and Doscher, M. S. (1984) Aspartic acid-121 functions at the active site of bovine pancreatic ribonuclease. *FEBS letters* **171**, 253-256

296. Inoue-Ito, S., Yajima, S., Fushinobu, S., Nakamura, S., Ogawa, T., Hidaka, M., and Masaki, H. (2012) Identification of the catalytic residues of sequence-specific and histidine-free ribonuclease colicin E5. *Journal of biochemistry* **152**, 365-372

297. Lou, Y. C., Huang, Y. C., Pan, Y. R., Chen, C., and Liao, Y. D. (2006) Roles of N-terminal pyroglutamate in maintaining structural integrity and pKa values of catalytic histidine residues in bullfrog ribonuclease 3. *Journal of molecular biology* **355**, 409-421

298. Huyghues-Despointes, B. M., Thurlkill, R. L., Daily, M. D., Schell, D., Briggs, J. M., Antosiewicz, J. M., Pace, C. N., and Scholtz, J. M. (2003) pK values of histidine residues in ribonuclease Sa: effect of salt and net charge. *Journal of molecular biology* **325**, 1093-1105

299. Park, C., Schultz, L. W., and Raines, R. T. (2001) Contribution of the active site histidine residues of ribonuclease A to nucleic acid binding. *Biochemistry* **40**, 4949-4956

300. Walter, B., and Wold, F. (1976) The role of lysine in the action of bovine pancreatic ribonuclease A. *Biochemistry* **15**, 304-310

301. Huang, Y. C., Lin, Y. M., Chang, T. W., Wu, S. J., Lee, Y. S., Chang, M. D., Chen, C., Wu, S. H., and Liao, Y. D. (2007) The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. *J Biol Chem* **282**, 4626-4633

302. Trautwein, K., Holliger, P., Stackhouse, J., and Benner, S. A. (1991) Site-directed mutagenesis of bovine pancreatic ribonuclease: lysine-41 and aspartate-121. *FEBS letters* **281**, 275-277

303. Boix, E., Blanco, J. A., Nogues, M. V., and Moussaoui, M. (2013) Nucleotide binding architecture for secreted cytotoxic endoribonucleases. *Biochimie* **95**, 1087-1097

304. Gesner, E. M., Schellenberg, M. J., Garside, E. L., George, M. M., and Macmillan, A. M. (2011) Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nature structural & molecular biology* **18**, 688-692

305. Sato, S., Inoue, T., Terada, K., Matsuo, I., Aizawa, S., Tano, Y., Fujikado, T., and Furukawa, T. (2007) Dkk3-Cre BAC transgenic mouse line: a tool for highly efficient gene deletion in retinal progenitor cells. *Genesis* **45**, 502-507

306. Morales, D., and Hatten, M. E. (2006) Molecular markers of neuronal progenitors in the embryonic cerebellar anlage. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**, 12226-12236

307. Wechsler-Reya, R. J., and Scott, M. P. (1999) Control of neuronal precursor proliferation in the cerebellum by Sonic Hedgehog. *Neuron* **22**, 103-114

308. Roussel, M. F., and Hatten, M. E. (2011) Cerebellum development and medulloblastoma. *Current topics in developmental biology* **94**, 235-282

309. Kisby, G. E., Olivas, A., Park, T., Churchwell, M., Doerge, D., Samson, L. D., Gerson, S. L., and Turker, M. S. (2009) DNA repair modulates the vulnerability of the developing brain to alkylating agents. *DNA repair* **8**, 400-412

310. Katyal, S., and McKinnon, P. J. (2008) DNA strand breaks, neurodegeneration and aging in the brain. *Mechanisms of ageing and development* **129**, 483-491

311. Jansen, R. P., and de Boer, K. (1998) The bottleneck: mitochondrial imperatives in oogenesis and ovarian follicular fate. *Molecular and cellular endocrinology* **145**, 81-88

312. Acton, B. M., Jurisicova, A., Jurisica, I., and Casper, R. F. (2004) Alterations in mitochondrial membrane potential during preimplantation stages of mouse and human embryo development. *Molecular human reproduction* **10**, 23-32

313. Piko, L., and Clegg, K. B. (1982) Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. *Developmental biology* **89**, 362-378

314. Lin, M. D., Jiao, X., Grima, D., Newbury, S. F., Kiledjian, M., and Chou, T. B. (2008) Drosophila processing bodies in oogenesis. *Dev Biol* **322**, 276-288

315. Radford, H. E., Meijer, H. A., and de Moor, C. H. (2008) Translational control by cytoplasmic polyadenylation in Xenopus oocytes. *Biochim Biophys Acta* **1779**, 217-229

316. Flemr, M., Ma, J., Schultz, R. M., and Svoboda, P. (2010) P-body loss is concomitant with formation of a messenger RNA storage domain in mouse oocytes. *Biology of reproduction* **82**, 1008-1017

317. Yu, J., Hecht, N. B., and Schultz, R. M. (2003) Requirement for RNA-binding activity of MSY2 for cytoplasmic localization and retention in mouse oocytes. *Dev Biol* **255**, 249-262

318. Swetloff, A., Conne, B., Huarte, J., Pitetti, J. L., Nef, S., and Vassalli, J. D. (2009) Dcp1-bodies in mouse oocytes. *Mol Biol Cell* **20**, 4951-4961

319. Hu, J., Wang, F., Zhu, X., Yuan, Y., Ding, M., and Gao, S. (2010) Mouse ZAR1-like (XM_359149) colocalizes with mRNA processing components and its dominant-negative mutant caused two-cell-stage embryonic arrest. *Developmental dynamics : an official publication of the American Association of Anatomists* **239**, 407-424

320. Li, L., Lu, X., and Dean, J. (2013) The maternal to zygotic transition in mammals. *Molecular aspects of medicine* **34**, 919-938

321. Li, X. Y., Cui, X. S., and Kim, N. H. (2006) Transcription profile during maternal to zygotic transition in the mouse embryo. *Reproduction, fertility, and development* **18**, 635-645

322. Liebers, R., Rassoulzadegan, M., and Lyko, F. (2014) Epigenetic Regulation by Heritable RNA. *PLoS genetics* **10**, e1004296

323. Leighton, P. A., Saam, J. R., Ingram, R. S., and Tilghman, S. M. (1996) Genomic imprinting in mice: its function and mechanism. *Biology of reproduction* **54**, 273-278

324. Yagi, R., Kohn, M. J., Karavanova, I., Kaneko, K. J., Vullhorst, D., DePamphilis, M. L., and Buonanno, A. (2007) Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* **134**, 3827-3836

325. Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A., Carlton, M. B., Pearce, J. J., Barton, S. C., Surani, M. A., Ryan, K., Nehls, M. C., Wilson, V., and Evans, M. J. (2000) Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature* **404**, 95-99

326. Pesce, M., and Scholer, H. R. (2001) Oct-4: gatekeeper in the beginnings of mammalian development. *Stem cells* **19**, 271-278