The Role of Exoribonucleases in Human Cells using Osteosarcoma as a Model

Amy Louise Pashler

A thesis submitted in partial fulfilment of the requirements of the University of Brighton and the University of Sussex for the degree of Doctor of Philosophy

April 2019

Abstract

Post-transcriptional control of gene expression is a critical level of regulation in the Central Dogma. One of many layers is control of RNA turnover and metabolism which are vital to maintain cellular homeostasis. In RNA processing, the role of RNA stability is essential in determining how long a species of RNA is able to function within the cell. In addition, RNA stability is central to the process of translation of RNAs, as well as the ability of regulatory RNAs to inhibit or repress their target mRNAs. This thesis focuses on the function of exoribonucleases within human cells and their role in regulating gene expression. Exoribonucleases are key enzymes in RNA degradation, which degrade RNA in either the 5' – 3' direction, or in the 3' – 5' direction. There is particular focus on the role of the 5' – 3' exoribonuclease, XRN1, and how defects in 5' – 3' RNA decay can result in defective protein expression, and the onset of disease.

This thesis characterises the expression of XRN1 in human bone cancer cell lines, and compares expression to that of a non-cancer cell line control. In doing so, the elucidation of XRN1 as a potential tumour suppressor gene in the progression of the most common bone cancer, osteosarcoma is investigated, alongside characterisation of the expression of the 3' – 5' exoribonucleases, DIS3, DIS3L1 and DIS3L2. DIS3 and DIS3L2 have each been previously implicated in the progression of human disease: mutations in DIS3 have been associated with a variety of leukaemias, whereas mutations in DIS3L2 have been associated with congenital overgrowth syndromes and also Wilms' Tumour of the kidney. This thesis shows that alongside XRN1, DIS3L2 may DIS3L2 also changes in expression in osteosarcoma.

This thesis also elucidates phenotypic changes in cells where XRN1 has been knocked down by a lipid-based RNAi transfection system. Using the osteosarcoma cells as a model, XRN1 was knocked down over a period of time, with cellular effects being observed over a time course. With regards to fundamental cellular pathways of proliferation, viability, translation and apoptosis, phenotypic changes in the behaviour of the cell lines were not observed. In contrast, knock down of DIS3L2 in a human embryonic kidney cell line, HEK-293T, resulted in hyper-proliferation, but this effect was not observed in the osteosarcoma cell lines, demonstrating a tissue specific effect for this gene in humans. The existence of synergism between the major exoribonucleases was also studied in this thesis, which found that there may be co-ordinate regulation occurring between XRN1 and DIS3L2, but not with DIS3 and DIS3L1.

Finally, this thesis shows the transcriptional changes in SAOS-2 cells when XRN1 is knocked down, in comparison with a control, using RNA sequencing technology to ascertain which pathways this enzyme functions to regulate. RNA sequencing was performed in a global approach to identify transcriptional targets of XRN1 in osteosarcoma cells. In doing so, it was

shown that XRN1 is involved in a multitude of cellular processes, such as cellular migration and adhesion, and also may function in the EGF pathway, a known oncogenic pathway. It also became clear that XRN1 targets transcripts with similar motifs in the 3' UTR, and that transcript targets are conserved across more than one cell line.

This thesis marks the stepping stone to elucidating the role of XRN1 in human cells, and serves as an introduction to the function of XRN1 in human disease to pave the way for future research.

Table of Contents

Table of Figures

Commonly used abbreviations

Abbreviation	Definition
OS	Osteosarcoma
EWS	Ewing sarcoma
MSC	Mesenchymal Stem Cell
DIS3	Defective in sister chromatid rejoining 3
DIS3L1	Defective in sister chromatid rejoining 3 like 1
DIS3L2	Defective in sister chromatid rejoining 3 like 2
XRN1/Pacman	Exoribonuclease 1
XRN2	Exoribonuclease 2
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
PES1	Pescadillo Ribosomal Biogenesis Factor 1
AGO 1/2/3	Argonaute 1/2/3
GW182	Glycine-Tryptophan Protein of 182kDa
mRNA	messenger RNA
pre-mRNA	precursor messenger RNA
miRNA	micro RNA
pri/pre-miRNA	primary/precursor micro RNA
lncRNA	Long non-coding RNA
rRNA	Ribosomal RNA
siRNA	Short Interfering RNA
tRNA	Transfer RNA
5’ UTR	5’ Untranslated region
3’ UTR	3’ Untranslated region

CDS	Coding sequence
RISC	RNA Induced Silencing Complex
miRISC	microRNA Induced Silencing Complex
ARE	AU rich element
Poly(A) tail/region	Polyadenylated Tail/Region
Poly(U) tail/region	Polyuridylated Tail/Region
RBP	RNA Binding Protein
PABP	Poly(A) Binding Protein
DCP1/2	Decapping Protein 1/2
P-Bodies	Processing Bodies
NMD	Nonsense Mediated Decay
PBS	Phosphate Buffered Saline
Brd-U	Bromodeoxy-Uridine
WST-1	Water Soluble Tetrazolium Salts - 1
GFP	Green Fluorescent Protein
cDNA	complementary DNA
PCR	Polymerase chain reaction
qRT-PCR	quantitative reverse transcriptase PCR
RNAi	RNA interference
RNA-seq	RNA-sequencing
FPKM	Fragments per kilobase of transcript per million mapped reads
AU	Arbitrary Units
KD	Knock Down
SEM	Standard Error of the Mean

Dedicated to the memory of Michael Lawrence Pashler

12th November 1955 - 8th November 2016

'I will not say the day is done, nor bid the Stars farewell.'

Acknowledgements

First of all, I would like to thank my supervisor, Professor Sarah Newbury, for giving me the opportunity to undertake this PhD. I would like to thank her not only for the academic and scientific support she has offered me, but also the unwavering personal support I have received throughout the duration of my project, without which my PhD experience would not have been as rich as it has been.

In addition, I would like to extend my thanks to my co-supervisors, Dr Chris Jones and Dr Peter Bush, for their valuable input and advice over the past three and a half years. Specifically, I would like to thank Chris for all of his help with statistical analysis elements of this project and his efforts to proof-read my thesis, an undertaking which is not to be envied!

I would also like to thank both past and present members of the Newbury lab, who have provided advice, entertainment, and support over the last three and a half years. In particular, I would like to thank all the Masters students who have assisted me in driving the project forward and performing experiments which have contributed to our findings. In addition, a special thank you is extended to Clare Rizzo-Singh, without whom the lab would not run as smoothly as it does! Thank you for always being on hand to order everything I've needed to conduct the project!

I would also like to thank Dr Helen Stewart and Dr Stefano Caserta for their continued advice and support throughout the project with regards to human cell culture work, and also their personal support at times when I doubted myself. Thanks also goes to Inyaki Pueyo-Marques, for all of his help.

A very special thank you goes to Dr Benjamin Towler, whose dedication to helping me throughout my project is testament to the wonderful human being he is. Thank you so much for your support scientifically, and at home during my constant battles with lack of self-confidence. Without you I wouldn't be here today, and I will always be eternally grateful. Here's to the future!

Finally, and by no means least, I would like to thank my friends and family, specifically my Mum, Katie, and my Dad, Michael, to whom this thesis is dedicated. I wish, Dad, that you were here to see me finish what I started. Without their encouragement from the very beginning this would not have been possible. Thank you also to my sisters and my brother, Nana, and all of my aunties for all of their love.

Candidate's declaration

I declare that the research contained in this thesis, unless otherwise formally indicated within the text, is the original work of the author. The thesis has not been previously submitted to this or any other university for a degree, and does not incorporate any material already submitted for a degree.

Signed:

Dated:

Chapter 1

Introduction

1.1 The Central Dogma

The Central Dogma describes the evolutionarily conserved mechanism at the very heart of biology. At its most simplistic, it is the way in which sections of DNA are read at any given time by proteins in the nucleus, transcribed into an RNA message, and transported out of the nucleus into the cytoplasm, in order for it to be translated into an amino acid chain, later undergoing folding into the final protein form. At its most intricate and complex, it is a multi-layered, heavily controlled cellular mechanism, which is unique in its complexity when it is compared to other cellular mechanisms (Figure 1.1). The following section of the introduction will endeavour to explain the Central Dogma, from transcription through to translation.

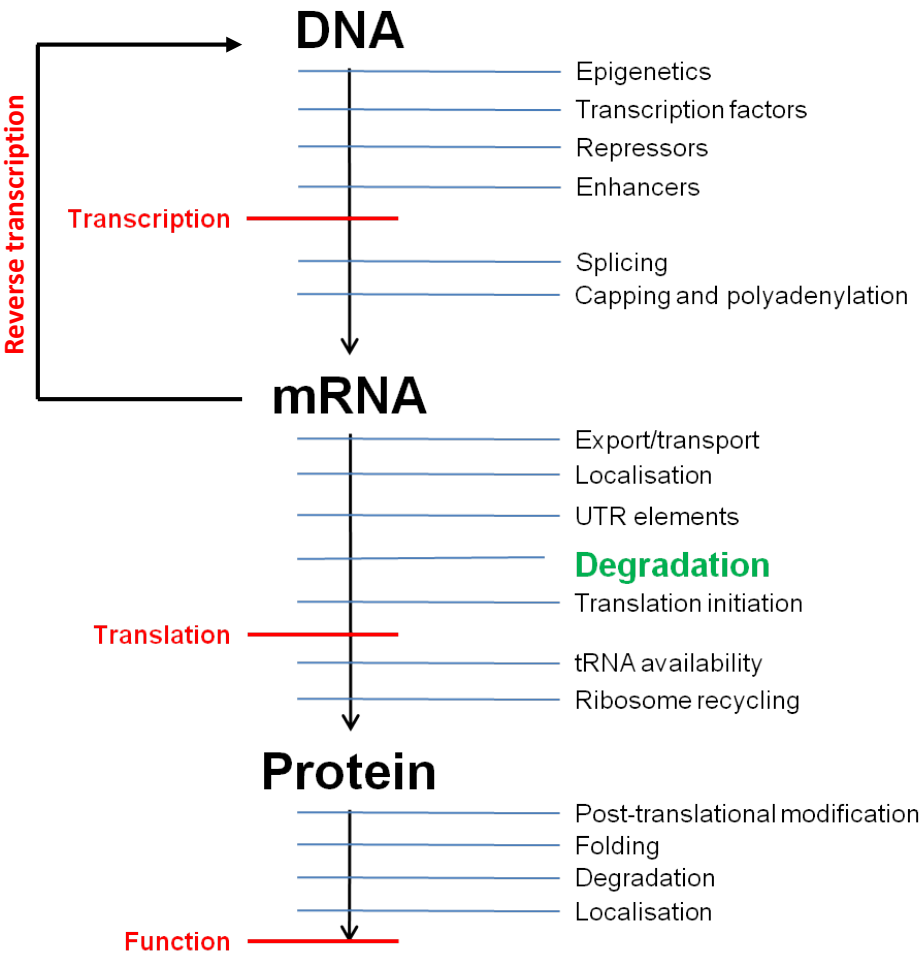


Figure 1.1. The Central Dogma. A brief flow chart of the Central Dogma, including the extensive layers of gene expression regulation. The aspect of gene regulation focused on in this project is RNA degradation, highlighted above.

1.1.2 A Brief Overview of Transcription

Transcription is the first level of the Central Dogma that needs to be completed correctly in order for normal gene expression to occur. The step begins with the catalysis of DNA unwinding by DNA helicase, which opens up the DNA into its respective strands: a coding strand and a template strand. This reaction promotes the recruitment of RNA polymerase, of which there are three (RNA polymerase I, II and III) to the promoter of the gene in order to start transcribing the selected gene along the template strand (Watson 2013). In eukaryotes, transcription initiation by RNA polymerase II occurs promiscuously, largely irrespective of annotated coding regions. This adds complexity to the transcriptome, and as such, sophisticated control mechanisms are required to prevent disruptive transcription events (Porrua and Libri 2015). Activator proteins bind to enhancers on the DNA, causing the DNA to bend, bringing the activator proteins into close proximity with promoters. Basal transcription factors bind to the promoter and subsequently recruit RNA polymerase keeping the RNA polymerase in place. Promoters usually have common sequences which recruit RNA polymerase, such as the TATA box motif.

The recruitment of RNA polymerase to the promoter leads to the formation of the transcription bubble, which is made up of a complex of proteins involved in ensuring correct transcription takes place, such as the Mediator complex, and a variety of transcription factors. The human Mediator complex is assembled from thirty different subunits, and is rich in disordered regions, which increase the potential for interactions with different regulatory proteins due to its malleable surface to its environment (Sierecki 2018). The Mediator complex undergoes a conformational change upon interaction with either RNA polymerase II or the C-terminal domain, and it is induced to form an extended structure, encompassing RNA polymerase, gene-specific regulators and the template strand.. The σ subunit of RNA polymerase stabilizes the transcription bubble when RNA polymerase binds to the unpaired bases, and the Mediator complex activates transcription through direct interactions with RNA polymerase and activators bound at the regulatory elements of target genes (Casamassimi and Napoli 2007)..

After transcription initiation, transcription of a single-stranded messenger RNA is continued along the length of the gene in the 3’ – 5’ direction, in a process known as transcription elongation. Elongation is a highly regulatory process, and is a repetitive but discontinuous formation of a phosphodiester bonds at every base pair. As such, specific DNA conformations, lesions and DNA binding proteins are critical in the study of elongation regulation (Imashimizu *et al.* 2014). Elongation takes places by translocation, and the addition

of nucleotides to the nascent mRNA strand. The polymerase moves forward by one base along the DNA using energy provided by ATP hydrolysis of the occupant NTP to a NMP, by the release of a pyrophosphate. Translocation is, therefore, synchronized to phosphodiester bond formation or pyrophosphate release (Imashimizu *et al.* 2014). An important characteristic of elongation is the ability of RNA polymerase to pause at specific sites along the template strand. In eukaryotes, RNA polymerase pausing at exon-intron junctions slows elongation, allowing for the formation of the spliceosome and efficient alternative splicing. Pausing also results in RNA polymerase back tracking, which plays a role in retaining polymerases in promoter-proximal regions to control specific gene transcription, and also serves to increase fidelity by providing a chance for proof-reading (Imashimizu *et al.* 2014).

The resulting RNA strand is produced in the 5’ – 3’ direction initially as a RNA:DNA hybrid strand, is complementary to the template strand, and almost identical to the coding strand, with the exception that all thymine residues are replaced with uracil residues. It is thought that uracil replaces thymine in the RNA strand because it is less stable than thymine, which means that metabolism of RNA is more energetically economical than metabolism using the thymine residues. Thymine, in contrast, is a better base in DNA because it is more stable, preventing mutations occurring.

Transcription termination occurs when RNA polymerase is released after it reaches a site in the DNA strand which signals for transcription to stop (Lykke-Andersen and Jensen 2007), and the newly formed mRNA strand completely dissociates from the RNA:DNA hybrid. These sites are recognized as processing and termination signals in the 3’ UTR of the nascent RNA by several components of the CPF-CF complex (Porrua and Libri 2015). This has been shown in some models to involve remodeling of RNA polymerase by an accessory helicase (SETX), responsible for the partial unwinding of the RNA polymerase transcription bubble during termination (Wang *et al.* 2019). In yeast, the interaction of the CPF-CF complex with both the RNA and RNA polymerase is thought to contribute redundantly to the recruitment of the complex to the 3’ end. At this point, the Ysh, an endonuclease of the CPF-CF complex, cleaves the RNA, and a poly(A) tail is added to the 3’ OH group (Porrua and Libri 2015). Termination is a vital step in transcription, because it prevents the occurrence of pervasive transcription. In protein coding genes, termination is coupled with cleavage and polyadenylation of the 3’ end of the transcript. In eukaryotic transcription termination, the process of termination is described in terms of 2 processes which each pertaining to how RNA polymerase is released from the RNA. In the allosteric model of transcription termination, the loss of elongation factors destabilizes the elongation complex, and the poly(A) signal causes a conformational change in the RNA polymerase, which is subsequently released. In the torpedo model, 5’ – 3’ exoribonuclease,

XRN2, is recruited to degrade nascent RNA cleavage, leading to the dissociation of the elongation complex. The subsequent conformational change of RNA polymerase causes its release from the RNA (Porrua and Libri 2015) (Figure 1.2). RNA polymerase can terminate at multiple sites beyond the poly(A) addition site, and the protein complex (the CPF-CF complex), which cleaves and polyadenylates the RNA strand, associates with the carboxyl-terminal domain of the RNA polymerase. Most RNA processing occurs co-transcriptionally (for example, splicing, polyadenylation and 5' end capping) and once transcription is complete, the pre-mRNA is released from the transcription bubble at RNA polymerase exit channels (Bentley 2005). This co-transcriptional processing increases the efficiency of transcription so that once processed, the mature mRNA is chaperoned out of the nucleus, for either translation or degradation at a faster rate. Non-coding RNAs are not always shuttled out of the nucleus; some remain in the nucleus for processing, degradation or to elicit their function.

Non-canonical transcription termination pathways include termination by a DNA 'roadblock', induced by a DNA binding protein. Here, RNA polymerase is stalled and targeted for degradation by the proteasome (Porrua and Libri 2015).

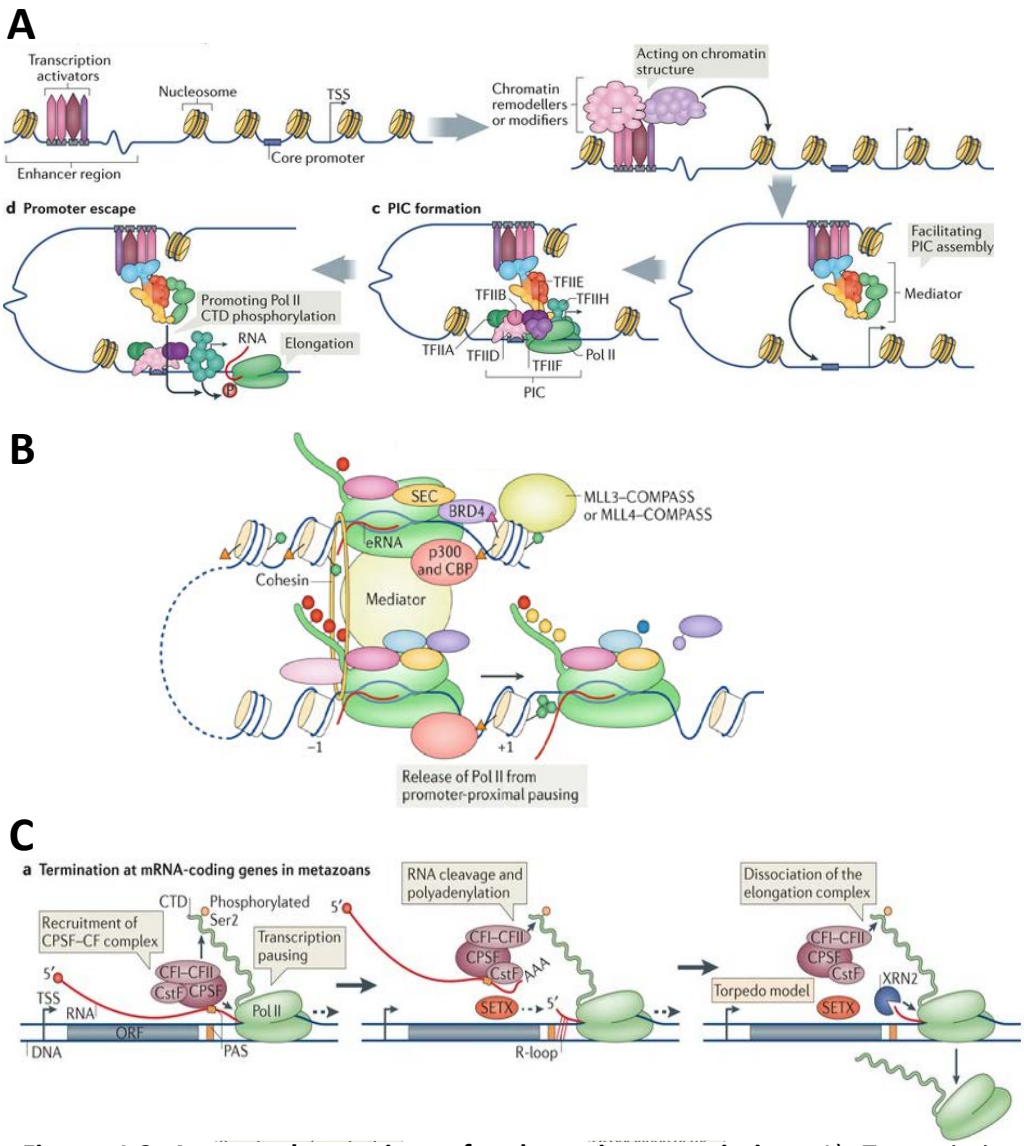


Figure 1.2 A general overview of eukaryotic transcription. A) Transcription Initiation: Transcription activation starts with the binding of transcription factors on enhancer regions. The transcription start site (TSS) is indicated by an arrow. Activators then recruit co-activator complexes that act as chromatin modifiers or remodellers to alter chromatin structure and to make it more accessible for other factors. Other co-activators are then recruited that act directly on the assembly of basal transcriptional machinery, the so-called preinitiation complex (PIC). Mediator is one of the key co-activator complexes. Adapted from (Soutourina 2017). B) Transcription Elongation: Elongation commences when the increased level of histone acetylation contributes to the binding of BRD4 and the super elongation complex (SEC) to chromatin through their bromodomains and YEATS domains, respectively. BRD4-containing and SEC-containing positive transcription elongation factor b (P-TEFb) phosphorylates negative elongation factor (NELF), DRB sensitivity-inducing factor (DSIF) and Ser2 residues of the Pol II CTD, which releases the paused Pol II into transcription elongation. Adapted from (Chen *et al.* 2018). C) Transcription Termination: In metazoans, the recruitment of the CPSF-CF complex leads to endonucleolytic RNA cleavage in the 3' UTR of the RNA and subsequent polyadenylation of the 3' end. Next, in the torpedo model, XRN2 (denoted above as Rat1) is recruited to degrade nascent RNA after cleavage, leading to the dissociation of the elongation complex. Conformational changes in RNA polymerase lead to its release from the complex (Porrua and Libri 2015).

1.1.3 A Brief Overview of Translation

Once out of the nucleus and having undergone extensive post-transcriptional modification (refer to section 1.2.2), the mRNA transcript is translated into a polypeptide, to be folded into a functional protein. There are three phases of translation (as in transcription): initiation, elongation, and termination.

Translation initiation can be dependent on the 5' methylguanine cap, or independent. Cap dependent translation initiation involves the interaction of the 5' cap with initiation factors (eukaryotic Initiation Factors (eIF) complex of proteins), which work together to bind the small 40S ribosomal subunit. The poly(A) binding protein also associates with the eIF4 complex, circularising the mRNA and creating a 43S preinitiation complex (PIC) at the 5' cap. This complex scans the mRNA strand towards the 3' end until it recognizes the start codon, which is usually encoded by AUG (methionine). The initiator tRNA (Met tRNA) is recruited to the AUG codon, and promotes the association of the large ribosomal subunit (60S) to the PIC (Figure 1.3). The association of both of the 43S PIC and large 60S ribosomal subunit completes the final 80S ribosome, and translation moves into the second stage, elongation (Hellen and Sarnow 2001, Hershey *et al.* 2012).

During IRES-dependent translation, scanning of the mRNA from the 5' end to the start codon does not occur. Instead, the ribosome is transported to the start site through the internal ribosome entry site (IRES) either by directly binding, or by IRES trans-acting factors (ITAFs). Elongation then commences (Hellen and Sarnow 2001).

The second stage of translation, elongation, describes the point at which the Met tRNA has occupied its site within the ribosome, and subsequent aminoacyl-tRNAs can be recruited to the mRNA in order to add a new amino acid to the growing polypeptide. The elongation factor, eEF1A, in complex with GTP, is responsible for delivering aminoacyl tRNAs to A site of the ribosome, and the recognition of codon: anticodon causes the hydrolysis of GTP>GDP, which subsequently releases eEAF1A from the ribosome. Aminoacyl-tRNAs are positioned in the correct place, first at the A site, a peptide bond is then formed between the previous amino acid and the new amino acid at the P site, and the mRNA is then shifted one codon down, releasing the tRNA at the E site. The complex of eEF1B, eEF1G and eEF1D promotes the exchange of the bound GDP for GTP, regenerating active eEF1A-GTP, allowing for the recruitment of the next aminoacyl-tRNA (Li *et al.* 2013).. This process continues, until the ribosome reaches the stop codon, and translation moves into the termination phase (Figure 1.4A).

Translation termination occurs when the cell's release factor, eRF1, recognizes a stop codon. At this point, hydrolysis of the complex, eRF1:eRF3:GTP occurs, releasing eRF3. eRF3 accelerates and increases termination efficiency. This allows ABCE1 to bind and facilitate the accommodation of eRF1 into an optimally active configuration, which promotes the hydrolysis of the ester bond linking the peptide to the tRNA. This hydrolysis leads to the release of the polypeptide (Dever and Green 2012). The ribosome is also disassembled. Once the polypeptide is released, it is shuttled to the endoplasmic reticulum to undergo protein folding, where isomerization of peptide bonds is catalyzed (Figure 1.4B). In a similar fashion to transcription, there are also multiple post-translational activities which modify new polypeptides, which will be discussed later in this chapter.

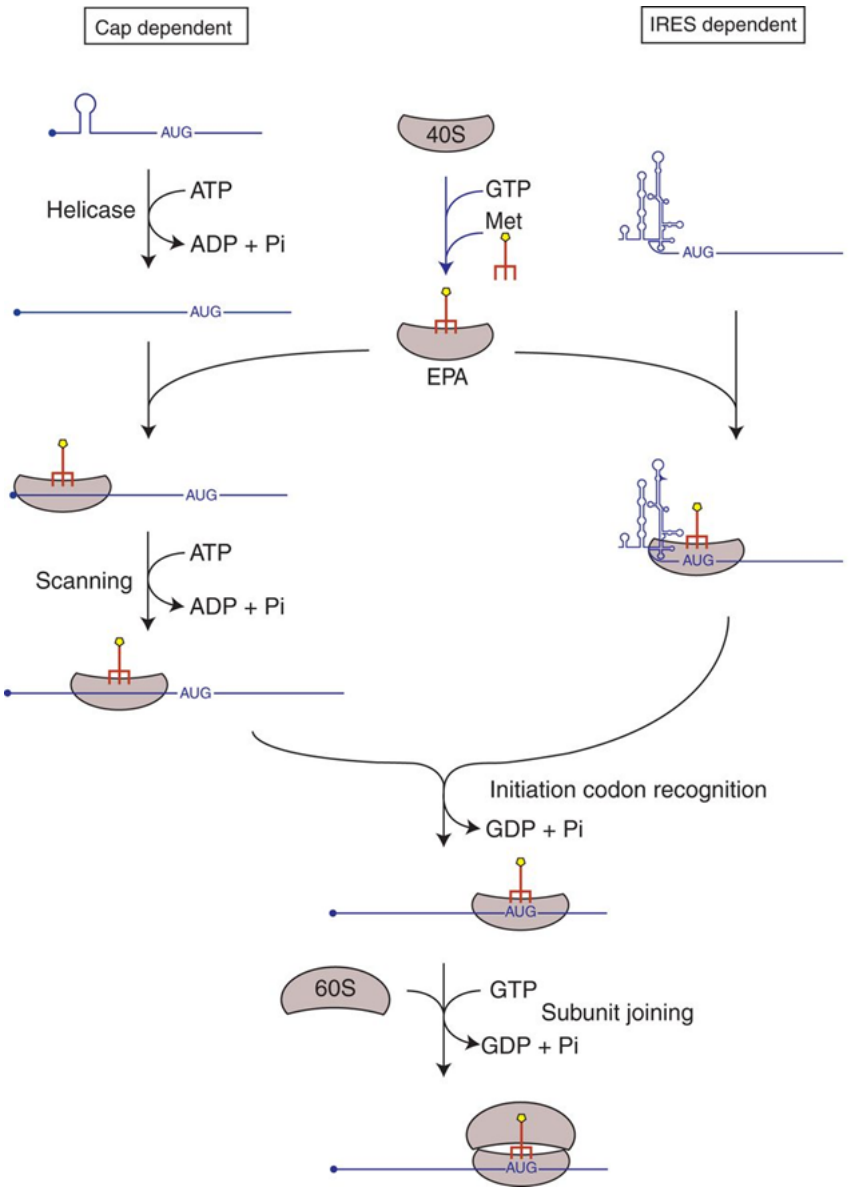


Figure 1.3. Eukaryotic Translation Initiation. A series of discrete steps of translation initiation, starting with the dissociation of the 80S ribosome into distinct subunits. This is followed by the binding of translation initiation factors, and the loading of the 43S preinitiation complex (PIC) onto an activated mRNP near the 5' cap. Scanning of the mRNA leads to the recognition of the start codon. The 60S ribosomal subunit joins the PIC, forming the final 80S ribosome, which then begins the elongation phase of protein synthesis. Adapted from (Hershey, Sonenberg *et al.* 2012).

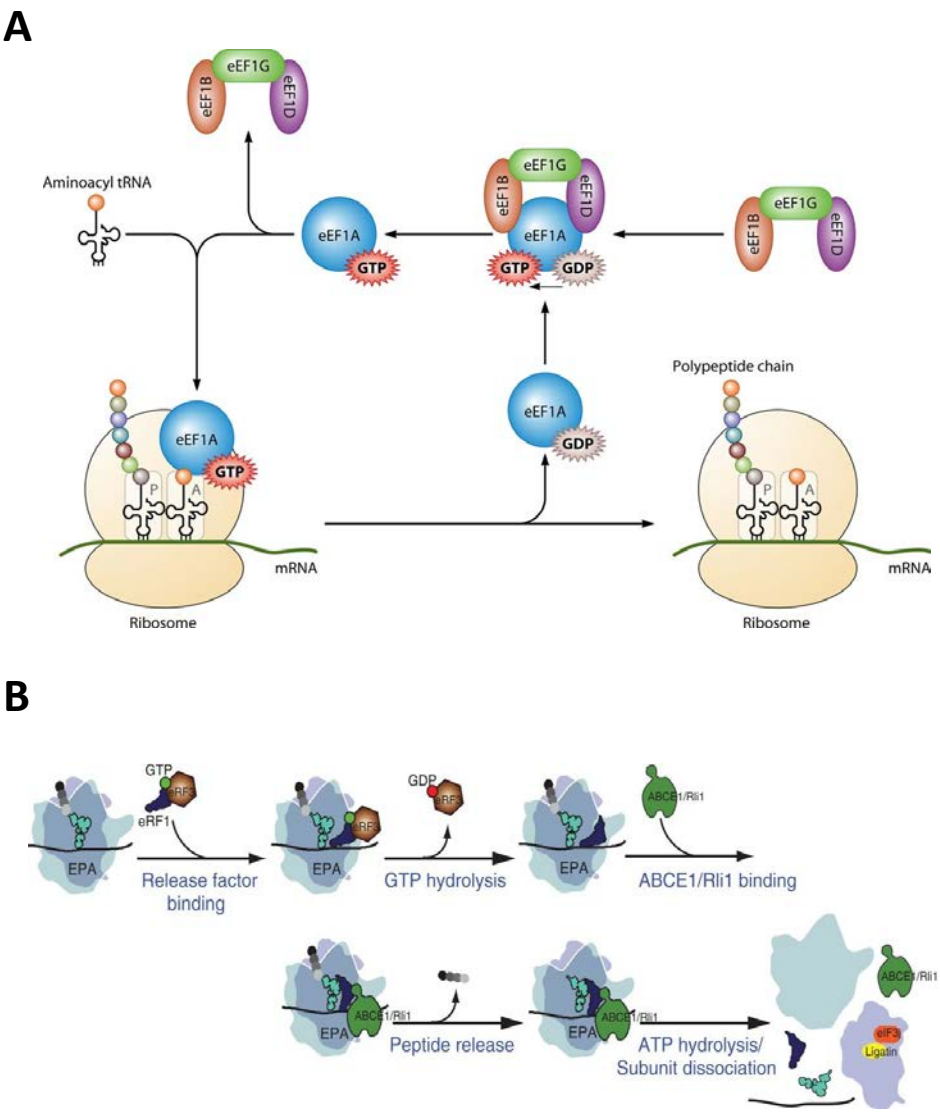


Figure 1.4. Eukaryotic Translation Elongation and Termination. A) Elongation: Elongation factor 1A (eEF1A), in complex with GTP, delivers an aminoacylated tRNA to the A site of the ribosome. The recognition of codon-anticodon causes hydrolysis of GTP > GDP, releasing eEAF1A from the ribosome. The eEF1B, eEF1G and eEF1D complex promotes the exchange of the bound GDP for GTP to regenerate active eEF1A-GTP, and promoting the recruitment of the next aminoacylated tRNA to the next codon. Adapted from (Li, Wei *et al.* 2013). B) Termination: When a stop codon is recognised, the eRF1:eRF3:GTP hydrolysis occurs, and eRF3 is released. eRF3 accelerates and increases termination efficiency. ABCE1 binds and facilitates the accomodation of eRF1 into an optimally active configuration, which then promotes the hydrolysis of the ester bond linking the peptide and the tRNA, leading to the release of the polypeptide. Adapted from (Dever and Green 2012).

1.2 Control of Gene Expression

Control of gene expression encompasses a vast array of mechanisms the cell uses in order for correct gene expression to take place. These mechanisms are fine-tuned and are highly conserved throughout evolution, highlighting the importance of these mechanisms in maintaining the correct balance of RNA and proteins within the cell. When these mechanisms fail, it can have detrimental effects on the cell, and can even lead to the onset of apoptosis, or at the other end of the spectrum, uncontrolled proliferation. The variety of control mechanisms the cell uses are present across all stages of the Central Dogma: transcriptionally, post-transcriptionally, translationally and post-translationally. This section will endeavour to introduce the concepts surrounding the control of gene expression.

1.2.1 Transcriptional control of gene expression

Transcriptional control of gene expression describes the vast expanse of both surveillance mechanisms used during transcription to ensure that correct transcription of pre-mRNA occurs, and also the mechanisms which ensure that the correct region of DNA is transcribed at any one time. There are many layers of transcriptional control of gene expression, from epigenetics, transcription factors, enhancers and repressors, through to the recognition of incorrect transcription by RNA polymerase, identified by other factors in the nucleus. It also describes how newly synthesized RNA can be processed in order for it to be subsequently degraded or translated into a new peptide.

The accessibility of RNA polymerase to the DNA template strand is one such way of controlling gene expression at transcription. This includes the way in which DNA is packaged prior to transcription and also during transcription. Histones are proteins responsible for packaging DNA, which therefore determine RNA polymerase accessibility. The cell modulates histones to allow or disallow access of RNA polymerase to specific loci. This ensures that the right genes are transcribed at the right time. Histones regulate transcription through specific modifications (such as acetylation) to their structure, which either promotes the unwinding of DNA for transcription, or represses transcription by promoting dense packaging of DNA (Eberharter and Becker 2002). Acetylation is usually associated with stimulating transcription, whereas methylation is conventionally associated with repressing transcription.

Transcription is also regulated by transcription factors, a large family of trans-acting gene regulatory molecules. By binding discrete *cis*-regulatory elements within the DNA, individual transcription factors have the potential to control hundreds of genes (Hobert 2008). Activity of transcription factors depends on the accessibility of their binding site, adding another

layer of regulatory control. Occupancy of the binding sites depends on nucleosome coverage of the site, and remodeling activities. Transcription factors are pivotal in determining cell-type specific gene expression, leading to the restriction of specific gene expression to a small subset of cells. Transcription factors can act in a double-negative way, in that the activating activity of one transcription factor can lead to the repression of expression of a transcriptional repressor. A classic example of how transcription factors display an intricate melody of expression in order to control specific gene expression is seen throughout many developmental processes, including bone development. In this case, expression of early differentiation transcription factor, Runx2, leads to the activation of subsequent differentiation transcription factors to promote correct osteoblastogenesis. Later on in the pathway, Runx2 becomes repressed, in order for osteoblast maturation to occur (Komori 2010).

Another layer of transcriptional control of gene expression includes how a transcript is processed before it is exported from the nucleus. Transcription results in a pre-mRNA structure which contains both introns and exons. In order for correct gene expression to occur, the non-coding regions, termed introns, need to be removed by splicing (Neugebauer 2002). Splicing is regulated by small nuclear RNAs (snRNAs), specifically, U1, U2, U4, U5 and U6 (Luhrmann *et al.* 1990), and occurs co-transcriptionally. The introns are excised from the pre-mRNA by the spliceosome, a protein complex which binds to a specific motif (usually AG) at each end of the intron. The spliceosome loops the intron into a circle, at which point the intron is cleaved out of the RNA strand, and the exons that were either side of the intron are ligated together (Will and Lührmann 2011).

A single pre-mRNA is capable of giving rise to a number of different mature mRNAs (isoforms), which generates the vast proteomic diversity in cells. The process of generating this diversity from a single pre-mRNA is called alternative splicing. It has been reported that up to 74% of multi-exon human genes utilise this process for the expression of a number of different proteins (Johnson *et al.* 2003). Alternative splicing is a highly regulated process, and errors in this process can have detrimental effects on the cell, and a number of diseases are associated with alternative splicing, including myotonic dystrophy (Tazi *et al.* 2009).

1.2.2 Post-transcriptional control of gene expression

This aspect of control of gene expression is the main focus of this project. Post-transcriptional control of gene expression describes post-transcriptional activities which determine the fate of a newly transcribed RNA strand after termination of transcription. This includes whether the transcript is tagged for degradation or translation, and how RNAs are capped by a 7-methyl guanosine cap at the 5’ end, and polyadenylated at the 3’ end. The

additions of these structures determine the stability of the RNA, and act to control the balance between RNA synthesis and degradation in the cell to prevent over production of proteins, or under production of proteins. In addition to post-transcriptional modifications, translation and decay, some RNAs are also sent for storage at sites, known as Processing bodies, in the cytoplasm, to be translated at a later time during the cell cycle, or as a response to cell stress. These processes are described in more detail later on this chapter.

The levels of stress that a cell is exposed to can dramatically affect cellular activity both transcriptionally and post-transcriptionally in order to protect the cell, or to induce programmed cell death. The changes in gene expression which occur during times of cell stress depends on the type of stress endured (Fulda *et al.* 2010). For example, stress caused by reactive oxygen species will cause a different response to nutrient depletion stress. It is important to bear this in mind when conducting experiments on cell lines (as in this project) as cells in culture will invariably undergo phases of stress which alter the gene expression landscape of the cell; this could influence the outcome of gene expression studies.

Other methods of post-transcriptional control of gene expression include microRNA (miRNA)-mediated gene expression which can occur both inside the nucleus and in the cytoplasm. A single miRNA can contribute to the repression of hundreds of genes through specific miRNA binding sites, and can occur through translational silencing of the RNA or through miRNA-mediated RNA degradation (Braun *et al.* 2012). This is discussed in more detail later on in this chapter.

1.2.3 Translational control of gene expression

After post-transcriptional processing of the RNA, it is either degraded or translated into protein. As in transcription, there are layers of control to regulate and ensure correct translation takes place. Translational control is important in defining the proteome, maintaining homeostasis and controlling vital cellular processes such as proliferation and subsequent organism growth and development. Features of translational control ensure that these processes function correctly and occur at different steps throughout translation. These include ribosome transit times (how quickly the ribosome moves along the RNA strand), though to specific binding of *trans* – acting proteins to mRNA in order to affect initiation rates (Hershey *et al.* 2012). For example, phosphorylation of eIF2α results in the inhibition of the initiator tRNA, iMet-tRNA, from binding to the ribosome, resulting in delayed translation. Indeed, phosphorylation of many initiation factors results in delayed translation.

Delayed translation initiation is not the only mechanism by which there is translational control of gene expression. Studies show that both the elongation and termination phases of translation can also be targets of translational control (Dever and Green 2012), via inhibition of elongation factors. This becomes rate limiting only when elongation is sufficiently slowed and affects the rate of protein synthesis. The rate of elongation can be slowed by the occurrence of rare codons and strong secondary structures in the coding region of the mRNA. It is thought that slowing elongation rate at specific regions can promote proper protein folding later on, if elongation is too fast, proteins can fail to fold properly (Zhang *et al.* 2009).

Translational termination can be regulated by suppression, either by read-through of the stop codon by the ribosome or by frame-shifting, which extends the carboxyl terminus of the protein. Read-through describes how the stop codon can be reprogrammed in order for the insertion of a seleno-cysteine residue, which causes translation to continue rather than terminate. Amendments to termination are often caused by features of the mRNA itself, rather than external proteins (Atkins and Gesteland 2002).

1.2.4 Post-translational control of gene expression

Post-translational control of gene expression is perhaps the last layer of gene expression control in the conventional Central Dogma. It is often described by a series of post-translational modifications (PTMs) which determine what happens to the peptide chain and newly formed protein. It can also include the phenomenon of proteolysis. PTMs are the covalent and enzymatic modification of proteins following synthesis which result in the formation of the mature protein. A very common form of PTM is phosphorylation, which is often associated with regulating the activity of enzymes. Proteins can also undergo glycosylation, which refers to the addition of a carbohydrate onto the protein. This can promote protein folding, as well as improving stability.

Post-translational control of gene expression generally describes the fine-tuning of newly synthesized proteins to ensure that they are ready for their designated function. This includes the removal of the initiator methionine residue, and also the formation of disulphide bonds from cysteine residues. After the formation of disulphide bonds, the resulting protein usually consists of two polypeptide chains connected by the bonds, which enhances protein folding (Robinson *et al.* 2017).

There are many ways in which a newly synthesized protein can be modified in order for its proper function to work. This section has referred to a few, however, it is important to emphasize that possible modifications to proteins is not limited to those mentioned here, and that there are an expanse of PTMs, such as ubiquitination as a means of regulating degradation,

which are discussed in other reviews (Mozzetta *et al.* 2015, Hendriks and Vertegaal 2016, Wang and Casey 2016, Murn and Shi 2017, Yang and Qian 2017).

1.3 A description of different RNA species

Over the past few decades, the level of knowledge about the structure of RNA has expanded exponentially. It has been increasingly shown that there are many different types of RNA species in existence within cells, all with their own specialised functions. The different species of RNA are discussed in the following subsections, which also include descriptions about their structures and functions. Each unique type of RNA has a specific role within the cell, whether this be as a direct regulator of gene expression, or utilized during translation to build new peptides, among others. This section will illuminate the diversity of eukaryotic RNAs, demonstrating how crucial RNA is to the viability of life across evolution.

1.3.1 RNAs involved in protein synthesis

1.3.1.1 Messenger RNA

Messenger RNA (mRNA) describes the RNA polymerized in the nucleus from a DNA template by RNA polymerase II. It is the intermediate between the genetic code of the DNA and the functional protein it encodes. Initially, pre-mRNA is synthesized, a preliminary strand of RNA which needs to be maturated to ensure it remains intact for protein synthesis. At this point, the pre-mRNA is spliced, capped and polyadenylated. Once fully processed, the mRNA is exported to the cytoplasm, where it is circularized for efficient translation into protein and subsequently degraded.

Structurally, mRNA has 3 sections: the 5' untranslated region (UTR), the coding sequence, and the 3' UTR (Figure 1.5A). The 5' and 3' UTRs are highly structured regions which contain regulatory elements and are not translated into protein. These sites are binding sites for *trans*-acting factors such as microRNAs or RNA binding proteins which act to regulate the stability of the RNA. The coding sequence is the code that is used to create the functional protein product. It is read as a triplicate code, each triplet codon recruits a specific transfer RNA (tRNA), which in turn provides the corresponding amino acid for the creation of a polypeptide chain, which is then processed into a functional protein.

1.3.1.2 Ribosomal RNA

Ribosomal RNAs are by far the most abundant species of RNA. They are fundamental components of the ribosome, the primary function of which is to read mRNA during translation. Approximately 60% of the ribosome is rRNA, and they generally contain 4 major rRNAs and around 50 component proteins. The 28S, 5.8S and 5S rRNAs, along with ribosomal proteins, constitute the larger ribosomal subunit and the 18S rRNA and other ribosomal proteins constitute the small ribosomal subunit. During translation, mRNA is fed between the two subunits and the ribosome catalyses the formation of a peptide bond between the two amino acids contained in the rRNA (Figure 1.5B).

The 28S, 18S and 5.8S rRNAs are co-transcribed by RNA polymerase I into a larger 47S precursor (pre-rRNA). The precursor, like pre-mRNA, undergoes a series of cleavage and trimming stages resulting in the three final rRNA structures. These are exported to the cytoplasm, where they interact with the larger and smaller subunit of the ribosome to effect their role in translation (Granneman and Baserga 2004).

The smaller 5S rRNA is transcribed by RNA polymerase III, and undergoes the same processing steps as the larger rRNAs. It too is then exported to the cytoplasm and complexes with the ribosome.

1.3.1.3 Transfer RNA

Transfer RNA (tRNA) is an adapter molecule composed of RNA, typically 76-90 nucleotides in length, with an L-shaped tertiary structure (Nguyen *et al.* 2019) (Figure 1.5C). These non-coding RNAs function during translation to provide the specific amino acid corresponding to the triplet codon in the mRNA for the progression of the peptide chain. tRNAs are transcribed by RNA polymerase III into precursor-tRNAs (pre-tRNAs), which undergo processing into a mature, highly structured RNA molecule.

The resulting mature tRNA becomes what is known as the cloverleaf structure. This structure composes of a coaxial stacking of RNA helices made up of a 5’ terminal phosphate, an acceptor stem, a 3’ CCA tail (used to attach the amino acid), and three arms (Pak *et al.* 2017). The three arms are the D arm, which contains dihydrouridine, the anticodon arm to provide directionality for the tRNA to read the mRNA in a Watson-Crick base pair manner (Nguyen *et al.* 2019), and the T arm, which contains a pseudouridine.

Before nuclear export, the CCA tail is charged with an amino acid by an aminoacyl tRNA synthetase. Each tRNA is charged with a unique amino acid for the elongation of the peptide chain.

1.3.2 RNAs involved in post-transcriptional modification

1.3.2.1 Small nuclear and nucleolar RNA

snRNAs are, as their name suggests, found within the nucleus, specifically localized to Cajal bodies. Cajal bodies are sites within the nucleus which are enriched in proteins and RNAs involved in mRNA processing. snRNAs are transcribed by RNA polymerase II or III and have regulatory features, such as facilitating the splicing of pre-mRNA.

Biogenesis of small nuclear RNAs (snRNAs) requires another class of non-coding RNA: small nucleolar RNA (snoRNA). These are a large subgroup of snRNAs. These RNAs reside in the nucleolus and mainly function to guide modification to other RNA structures.

1.3.3 RNAs involved in gene regulation

1.3.3.1 Long non-coding RNA

Long non-coding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides (Figure 1.6A). Historically, lncRNAs have not been readily studied, the genes encoding them having been thought of as mostly junk DNA. However, recently research into these RNAs has gained traction and they have gained notoriety as some of the most effective gene expression regulators.

LncRNAs are responsible for regulating many cellular processes, such as chromosome inactivation (Xist), transcription regulation and the ability to act as molecular sponges. The biogenesis of lncRNAs shows similarities with the biogenesis of mRNAs, and lncRNA transcripts have a 5' methylguanosine cap, and often undergo splicing and polyadenylation (Mercer and Mattick 2013). The similarities between mRNA and lncRNA indicate a dynamic evolutionary interface between coding and non-coding RNAs, and demonstrates that coding RNAs can lose their ability to encode a protein, and in the same respect, noncoding RNAs can acquire coding function (Zheng *et al.* 2007, Mercer and Mattick 2013). Many lncRNAs have been implicated in

the progression of cancer (such as MALAT-1), which will be discussed in more detail in section 1.8.4.2. During regulation of gene expression, lncRNAs are capable of both enhancing and repressing gene expression by modulating the activity of transcription factors and transcriptional co-regulators. These RNAs can also function post-transcriptionally in a mechanism similar to that of other regulatory RNAs. They are functionally able to base pair with target mRNAs in order to mask key elements within the mRNA which are required for binding *trans*-acting factors, affecting multiple steps during post-transcriptional gene expression.

1.3.3.2 microRNA

In contrast to lncRNAs, microRNAs (miRNAs) are small regulatory RNAs of approximately 18 - 22 nucleotides in length (Figure 1.6B). They generally act to suppress translation or promote RNA decay. Each miRNA has the potential to regulate specific target sites, mainly achieved through binding to complementary sites in the 5' and 3' UTRs of mRNAs. Since their discovery in 1993, miRNAs have been implicated in the regulation of almost every cellular process (Vishnoi and Rani 2017).

The structure of a miRNA changes throughout its biogenesis. The fully mature form of a miRNA, however, comes as a hairpin structure. miRNAs are transcribed by RNA polymerase II and are usually transcribed as part of one arm of an approximately 80 nucleotide RNA stem loop, which in turn forms part of a primary miRNA (pri-miRNA) which is several hundred nucleotides long. Pri-miRNAs contain miRNA precursors (pre-miRNAs). The DGCR8 protein associates with Drosha to cleave hairpins from the pri-miRNA, forming a pre-miRNA (Lee *et al.* 2003). Pre-miRNA hairpins are exported from the nucleus for cytoplasmic processing by the RNase, Dicer. Dicer trims down the hairpin bringing the overall length of the mature miRNA to approximately 22 nucleotides (Romero-Cordoba *et al.* 2014). Pre-miRNAs can also be produced from introns, which are then processed by Dicer into mature miRNAs. Biogenesis and mechanism of action of miRNAs is discussed in more detail in section 1.5.4.1.

1.3.3.3 Piwi-interacting RNA

Piwi-interacting RNA (piRNA) is the largest class of small non-coding RNA in higher eukaryotes. They form piRNA-protein complexes by interacting with piwi proteins. These complexes are mostly involved in epigenetic and post-transcriptional silencing of transposons, however, they can also be involved with other genetic elements in germline cells (Siomi *et al.* 2011). The length of piRNA varies from between 21 and 31 nucleotides, and they have been found both in the nucleus and in the cytoplasm.

piRNA biogenesis differs from other types of RNA in that they are derived from just one strand of DNA. Once a piRNA precursor is formed, the process of gene silencing by piRNA is initiated. This is known as the Ping-Pong cycle, and is used for both the maturation of piRNA and also targeted gene silencing, whereby the primary piRNA recognizes its complementary target and recruits piwi proteins. This results in a cleavage of the transcript, producing secondary piRNA (made from a target mRNA). The secondary piRNA is then targeted towards sequences with an adenine in the tenth position, for the silencing of a further transcript (Aravin *et al.* 2008).

Many aspects of the mechanism of action of piRNAs remain poorly understood.

1.3.3.4 Small interfering RNA

Small interfering RNA (siRNA) are double stranded RNA molecules approximately 20-25 nucleotides in length (Figure 1.6C). They function to regulate expression of specific genes with complementary sequences, directing them for post-transcriptional decay. Like miRNAs, siRNAs are processed by Dicer, which catalyses siRNA formation from long double-stranded RNAs or short hairpin RNAs. siRNAs differ from miRNAs in that they do not have a hairpin structure, and they typically induce gene silencing post-transcriptionally, rather than translationally (Mack 2007), however, once these RNAs are formed they can also be incorporated into the RNA-Induced Silencing Complex (RISC). When incorporated into the RISC, siRNAs are unwound, creating a less stable single stranded RNA to target complementary sequences on the mRNA. Once bound to its target, the siRNA induces endonucleolytic cleavage of the mRNA, directing the mRNA for degradation (Singh *et al.* 2018). In this project, siRNAs are utilised artificially for targeted knock down of both XRN1 and DIS3L2 mRNAs, demonstrating their strong therapeutic potential.

1.3.3.5 Circular RNA

Circular RNA (circRNA) are single stranded RNAs, commonly processed from pre-mRNA, which form a covalently closed loop, whereby by the 3’ and 5’ ends are joined together (which also means they are resistant to degradation by exoribonucleases) (Figure 1.6D). They are composed of between 1 and 5 exons, the length of which are longer than protein coding exons, which suggests that exon length is important in determining whether an exon is circularized or not (Memczak *et al.* 2013). These RNAs are usually formed during splicing events of protein coding RNAs (Figure 1.6D), and can be identified by the presence of scrambled exon junctions. The vast majority were shown to be processed from exons located within the middle of genes, coupling their processing to spicing events (Zhang *et al.* 2014). Back splicing of mRNA leads to the circularization of an exert mRNA, and a circRNA can be composed of purely a single exonic

mRNA, a multi-exon mRNA, a singular intron mRNA or a single intron combined with a single exon mRNA.

It is thought that circRNAs function as sponges for miRNAs, due to the presence of miRNA binding sites within their sequences. It is also thought that circRNAs may act as miRNA competitors, to suppress the ability of the miRNA to bind to its target (Ebert and Sharp 2010), thereby up regulating miRNA target expression (Qu *et al.* 2018). circRNAs are also implicated in the regulation of translation, including the suppression of translation of PABPN1 (important in mRNA processing) by circPABPN1 (Abdelmohsen *et al.* 2017)

Most circRNAs are aberrantly expressed in different cancer types, and some have been reported to play important roles in the development and progression of cancer. They can play either suppressive or oncogenic roles in multiple cancers. For example, reduced circRNA expression in colorectal cancer is negatively correlated with proliferation (Bachmayr-Heyda *et al.* 2015), and over expression of circITCH has been shown to suppress tumour growth by binding *miR-7*, *miR-17*, and *miR-124* in oesophageal squamous cell carcinoma (Li *et al.* 2015). A comprehensive list of current circRNAs known to be involved in the development and progression of cancer can be found in (Qu *et al.* 2018).

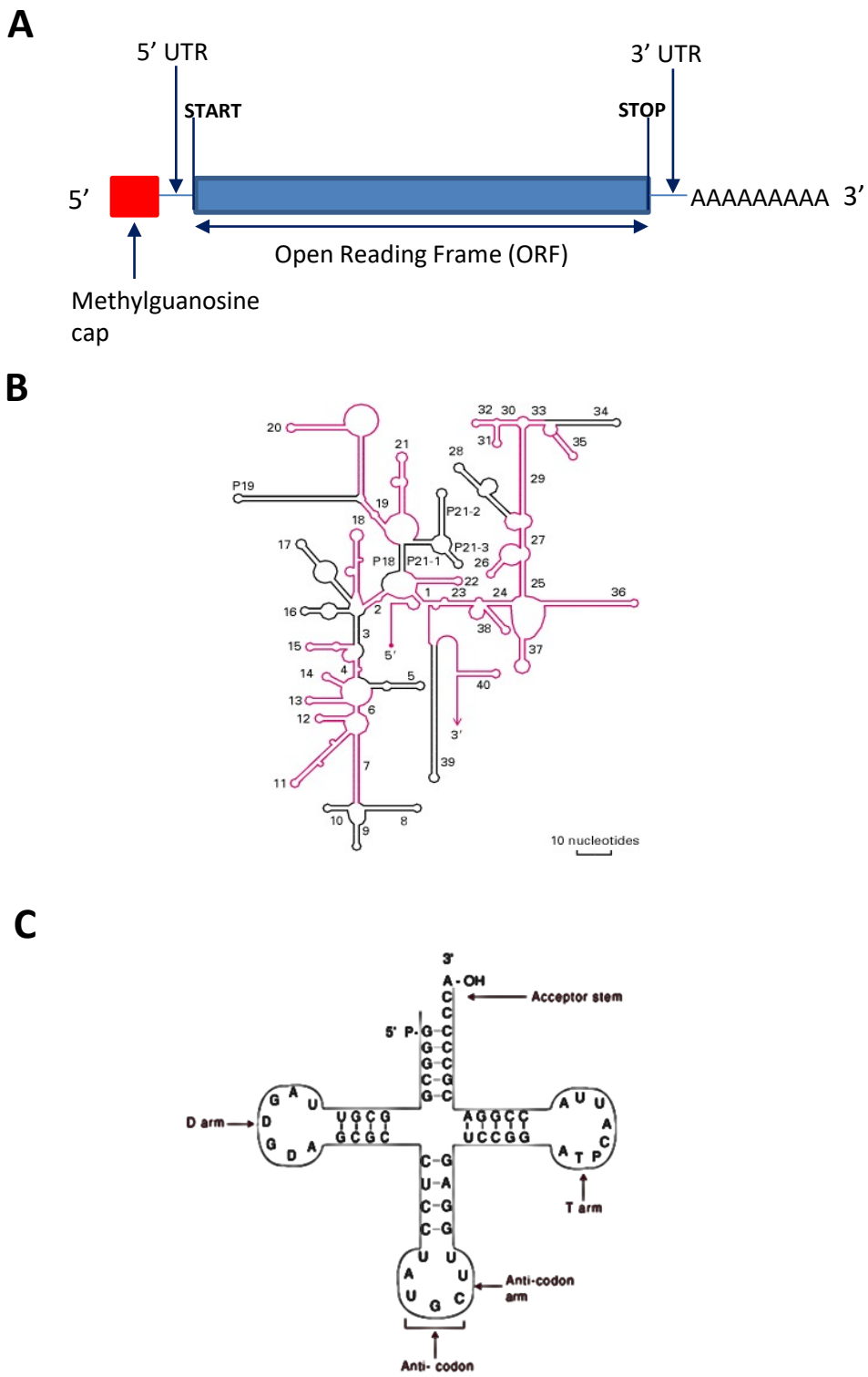


Figure 1.5. Structures of RNAs involved with protein expression. A) Messenger RNA (mRNA). B) The small 5S ribosomal subunit ribosomal RNA (rRNA). C) Transfer RNA (tRNA). Figures B and C were adapted from https://www.ncbi.nlm.nih.gov/books/NBK21603.

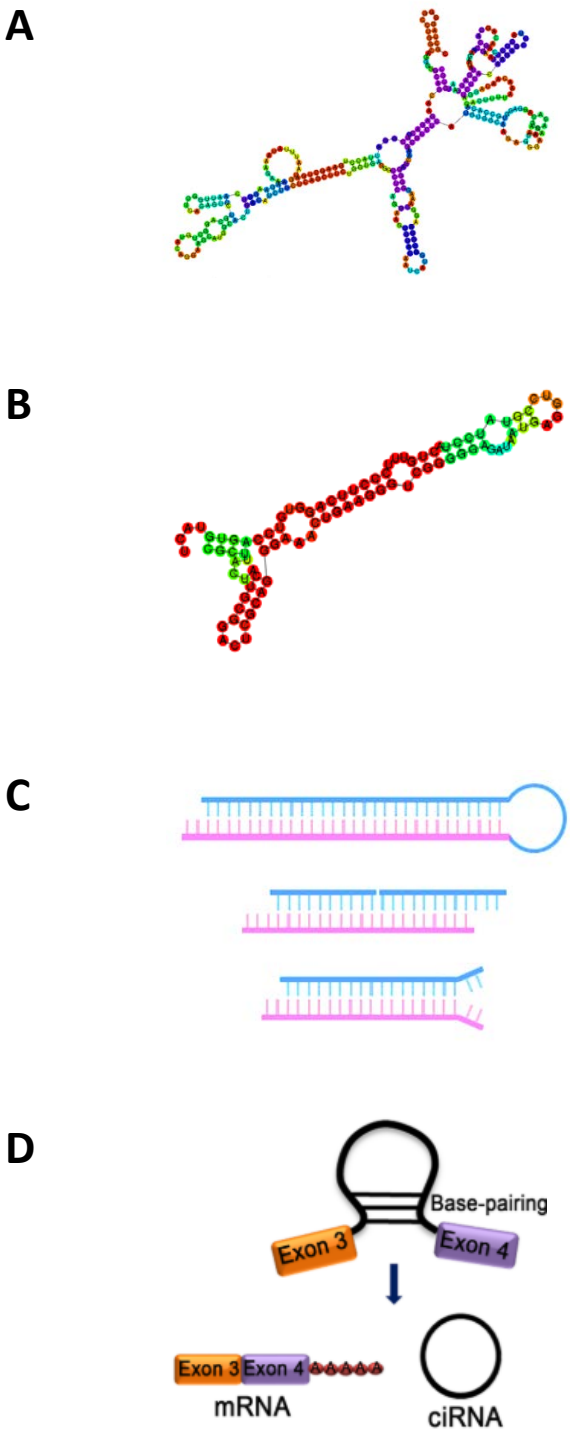


Figure 1.6 Structures of RNAs involved with regulation. A) Long noncoding RNA (lncRNA) adapted from http://www.bioquicknews.com/node/3573. B) MicroRNA (miRNA) adapted from http://cegg.unige.ch/mirortho/entry_details/28147. C) Structural variants of small interfering RNA (siRNA) as adapted from (Ku, Jo *et al*. 2016) D) Basic processing of circRNA, adapted from (Xu, Zhou 2018).

1.4 The Mechanisms and Roles of RNA Stability

The main focus of this project is post-transcriptional control of gene expression in the context of the study of RNA stability and decay in human osteosarcoma cells. This next section will describe in detail the pathways involved in determining RNA stability and the events occurring during RNA degradation by exoribonucleases.

1.4.1 RNA stability

RNA stability describes the stabilization of RNA molecules by the addition of specific molecular structures which protect them from degradation. The more protected the RNA, the longer its half-life will be in the cell. It also determines the rate of translation of the RNA. Stability can be conferred by specific structures or sequences at both the 3' and 5' ends of the RNA transcript, and removal of these structures signals for the decay of the RNA by leaving the 3' or 5' end vulnerable to decay by specific exoribonucleases. The sequences within the RNA which lead to the addition of these structures are referred to as the *cis*-acting elements. An example of the *cis*-acting elements are the highly conserved AU-rich elements (AREs) found in the 3' UTR. Around 5-8% of all human genes contain these regions which serve to decrease the stability of the transcript (Schoenberg and Maquat 2012). AREs interact with a number of binding proteins which also regulate transcript stability, including TTP, Auf1 and HuR. *Cis*-acting elements also contribute to the way mRNA is alternatively spliced to produce many different isoform products from one gene (Ghigna *et al.* 2008). These *cis*-acting elements interact with external *trans*-acting factors during RNA maturation. The *trans*-acting factors recognize specific regions in the RNA which determines the factor to be associated with the RNA, for example, specific RNA binding proteins recognize sequences within the RNA which are then responsible for the addition of the 7-methylguanosine cap, added to the 5' UTR of the transcript to stabilize the RNA. These regions also include binding sites for miRNAs, which often act to repress gene expression. *Cis*-acting elements are not only limited to the 5' UTR, the 3' UTR is also an important factor in determining the stability of a transcript. The 3' UTR provides binding sites for miRNAs and RBPs, which contribute to regulating the stability of the transcript.

Another aspect of RNA stability includes the phenomenon of codon optimality and bias. This describes the uneven use of synonymous codons in the transcriptome to guide the efficiency of protein production, translation fidelity and mRNA metabolism (Hanson and Coller 2018). Recent studies have demonstrated that there exists a tight coupling between ribosome

dynamics and stability of mRNA transcripts, revealing the effects of specific codons on translation efficiency.

1.4.2 Regulation of gene expression by 3' tailing

3' tagging of newly synthesized transcripts has been shown to promote both 5' – 3' and 3' – 5' degradation of RNA (Mullen and Marzluff 2008). It has been shown that around 80% of all mammalian transcripts harbour some 3' tagging, and a well-studied example is shown in the oligouridylation of histone RNAs (which lack poly(A) tails) to regulate their expression. Oligouridylation, performed by TUTases (Menezes *et al.* 2018), of histone RNAs promotes recruitment of the LSM1-7 complex, which subsequently leads to decapping and degradation of the transcript from both directions (Schmidt *et al.* 2011). The RNA binding protein, LIN28, has also been shown to interact with TUT4 to polyuridylate *pre-let-7*, in order to block biogenesis of the tumour-suppressor let-7 miRNA family (Menezes *et al.* 2018). Transcripts which undergo polyuridylation are historically known to be preferentially degraded by DIS3L2 (Chang *et al.* 2013, Gallouzi and Wilusz 2013, Malecki *et al.* 2013, Ustianenko *et al.* 2013).

As well as polyuridylation, polyadenylation is part of the maturation process of RNAs which contributes to the stability of the transcript (Colgan and Manley 1997). In this post-transcriptional modification of RNA, adenine residues are added to the 3' end of the pre-RNA by poly(A) polymerase, and polyadenylation ensures that only mature mRNA can be exported from the nucleus (Stewart 2019), presenting another mechanism by which gene expression is controlled. Poly(A) binding proteins then bind to the poly(A) tail which both facilitates translation and protects the transcript from degradation.

1.5 Mammalian exoribonucleases and how they function

Cells possess a number of exo and endoribonucleases, which function to degrade and cleave RNA in many regulatory capacities. Generally, there are more endoribonucleases than exoribonucleases, and they function both to cleave RNA and degrade RNA. Examples include: RNase H, which functions to cleave RNA in the DNA-RNA duplex to form single-stranded DNA (Nowotny 2009), RNase III, which cleaves double-stranded RNA to form mature RNAs such as miRNAs (Lee *et al.* 2019) and RNase P, a ribozyme which acts to cleave a precursor tRNA out of the 5' end of a single-stranded RNA (Xiao *et al.* 2001). Alongside these endoribonucleases, decapping enzymes also demonstrate RNA cleavage activities. Further information about known endoribonucleases can be found in reviews elsewhere (Li *et al.* 2010, Tomecki and Dziembowski

2010), however, given that the majority of eukaryotic RNA turnover is performed by exoribonucleases (Kushner 2004, Meyer *et al.* 2004), they will be the main focus of this section. This project focused on five of the known exoribonucleases in mammals which are involved in the RNA decay pathway. Two of these enzymes function to degrade RNA in the 5' – 3' direction: the nuclear XRN2 and cytoplasmic XRN1 enzymes. There are four enzymes which function to degrade RNA in the 3' – 5' direction. These are the members the DIS3 family, and the nuclear enzyme, RRP6 (Januszyk *et al.* 2011), which was not studied in this project. The three members of the DIS3 family which function in the 3' – 5' direction to degrade RNA include: the catalytic component of the nuclear exosome, DIS3, the catalytic component of the cytoplasmic exosome, DIS3L1, and the independent cytoplasmic exoribonuclease, DIS3L2. This section will discuss the biochemical mechanisms of action of each of the exoribonucleases.

1.5.1 XRN1: Structure and functions

XRN1 is a cytoplasmic exoribonuclease that degrades target mRNAs in the 5' – 3' direction. It is the only cytoplasmic 5' – 3' exoribonuclease in mammalian cells, highlighting its importance in maintaining the balance of RNA synthesis and decay in cells. XRN1 is highly conserved across evolution, with homologues in yeast, *Drosophila melanogaster* through to humans (Figure 1.7) (Jones *et al.* 2012). XRN1 protein is expressed ubiquitously throughout all human tissues, demonstrating the importance of this exoribonuclease (Figure 1.8).

XRN1 is responsible for catalyzing decay of target mRNAs in its RNase domain using the Brownian ratchet mechanism. This mechanism describes the mechanical motion produced during biochemical reactions such as this, whereby RNA is pulled through the enzyme based on the kinetics of the catalysis reaction.

During degradation by XRN1, RNA is fed through a narrow channel in the protein structure (specifically wide enough only for single-stranded RNA), through to the RNase domain basic pocket, where it is cleaved. The first 3 nucleotides are held in position by 2 conserved amino acid residues: His41 and Trp540. The basic pocket residues interact with the 5' phosphate bond of the first nucleotide, exposing it to 2 magnesium ions in the active site, at which point the bond undergoes hydrolysis (Jinek *et al.* 2011). His41 then uses the Brownian ratchet mechanism to move the RNA along for cleavage of the second nucleotide (Nagarajan *et al.* 2013). XRN1 then processively digests single stranded RNA. The basic pocket is only wide enough for 5' monophosphorylated RNAs, meaning that specificity for degradation by XRN1 is limited only to single-stranded RNAs which have had their 5' protective methyl-guanosine cap removed, as

additions such as these are too highly structured for XRN1 activity. Indeed, highly structured motifs stall XRN1, as seen in viral evasion of the host response mechanisms (Silva *et al.* 2010).

XRN1 has been shown to demonstrate substrate specificity beyond distinguishing RNAs which have or have not been decapped. Previous work in the Newbury lab has shown that in *Drosophila melanogaster* models XRN1 specifically targets pro-apoptotic RNAs in imaginal wing discs, suggesting a role for XRN1 in controlling growth and development (Waldron *et al.* 2015). It is not known, however, how substrate specificity at this level is regulated, and whether the targets of XRN1 are conserved across multiple tissue types, or even if it is specific to each organism. Indeed, RNA-sequencing of XRN1 mutant *Drosophila* wing imaginal discs showed that XRN1 specifically targets an insulin-like peptide, *dilp8*, which is associated with developmental delay in *Drosophila*. This evidence supports a role for XRN1 in regulating cellular growth and apoptosis (Jones *et al.* 2016). It is unknown if XRN1 differentially targets mRNAs through a mechanism of its own, or whether mRNAs are differentially targeted to it..

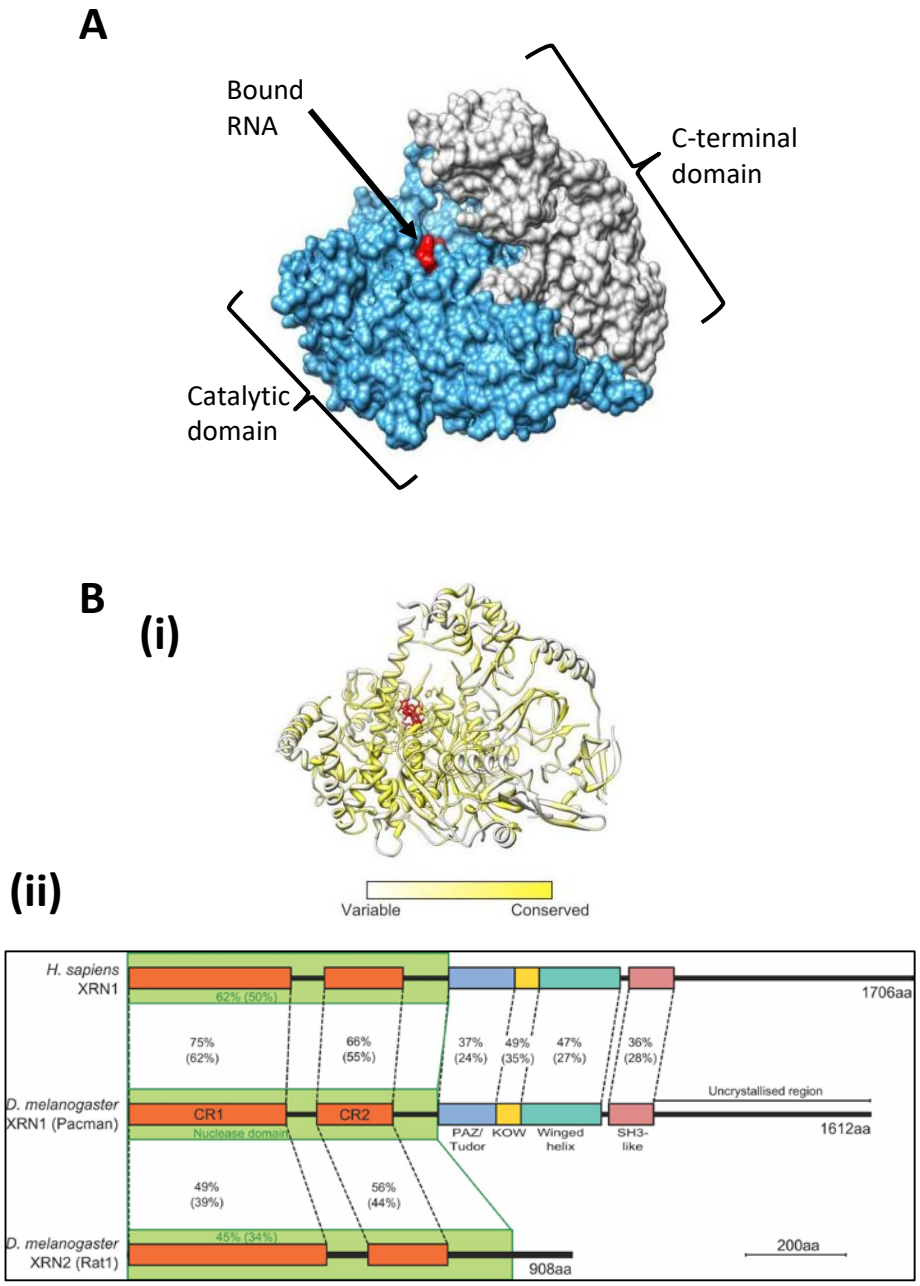


Figure 1.7. Structure of *Drosophila melanogaster* XRN1. A) Crystal structure of XRN1. The catalytic N-terminal domain is shown in blue. Bound RNA is shown by red in the catalytic domain. The C-terminal is shown in grey. B) (i) The degree of conservation of *Drosophila melanogaster* XRN1 across 196 eukaryotic XRN1s. The more highly conserved regions of the protein are denoted in yellow. In this schematic, it is clear that the C-terminal region of XRN1 is not as highly conserved as the N-terminal region (Jinek, Coyle *et al.* 2011). (ii) Numerical degree of conservation of XRN1 between *Homo sapiens* and *Drosophila melanogaster,* where percent similarity is shown together with percent identity (in parenthesis). These diagrams were adapted from (Nagarajan, Jones *et al*. 2013) and (Jones, Zabolotskaya *et al*. 2012).

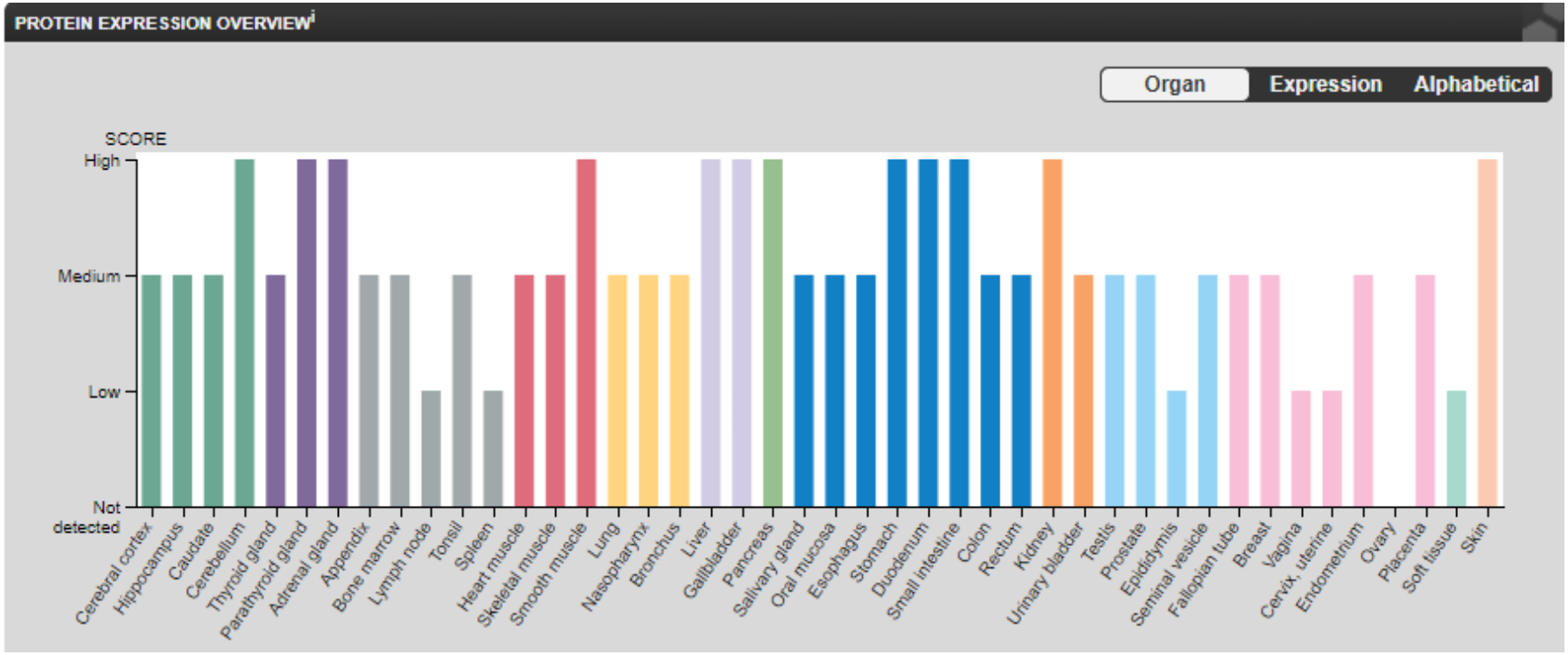


Figure 1.8. Ubiquitous expression of XRN1 in all human tissues. This figure shows how XRN1 is expressed in all human tissues, with the exception of the tissue in the ovaries. This figure was retrieved from the Human Protein Atlas (https://www.proteinatlas.org/ENSG00000114127-XRN1/tissue).

1.5.2 XRN2: Structure and Functions

XRN2 is the 5’ – 3’ exoribonuclease localized in the nucleus. XRN2 is also highly conserved throughout evolution, with its yeast paralogue named Rat1p. This exoribonuclease is a paralogue of XRN1, and there is a high degree of conservation between the 2 enzymes, most notably in their N-terminal and RNase domains (Figure 1.9). It is generally accepted that due to the high degree of conservation between XRN1 and XRN2, these enzymes likely metabolise RNA in very similar ways. Although both of these enzymes process ssRNA, the most relevant difference between these enzymes is that they function in different cellular compartments; whereas XRN1 is responsible for bulk mRNA turnover, XRN2 is heavily involved in pre-mRNA degradation, ribosomal RNA (rRNA) and microRNA (miRNA) processing, as well as noncoding RNA degradation (Luke *et al.* 2008, Zakrzewska-Placzek *et al.* 2010, Nagarajan *et al.* 2013). XRN2 also possesses a shorter C-terminal domain than XRN1. Another difference between the two enzymes is that XRN1 does not possess a nuclear localisation signal (NLS) in its C-terminal domain.

XRN2 is critical in correct transcription termination, to prevent the production of long mRNAs which can be deleterious to the cell (Proudfoot 2011). It is thought that XRN2 rapidly attacks the unprotected ends of cleavage products of nascent RNAs, after cleavage and polyadenylation of pre-RNAs. Alongside other cofactors, this rapid decay means that XRN2 catches up with RNA polymerase II and effectively collides with it, forcing it off the strand at the correct point, and terminating transcription (Dengl and Cramer 2009). In addition to its role in transcription termination, XRN2 acts as a quality control enzyme, which trims the 5’ ends of rRNA precursors (Geerlings *et al.* 2000).

XRN2 is also thought to be important in regulating telomere length, by promoting telomere elongation through degradation of a lncRNA which suppresses telomere lengthening. This implicates XRN2 in highly important cellular processes, such as cell ageing and senescence (Luke and Lingner 2009).

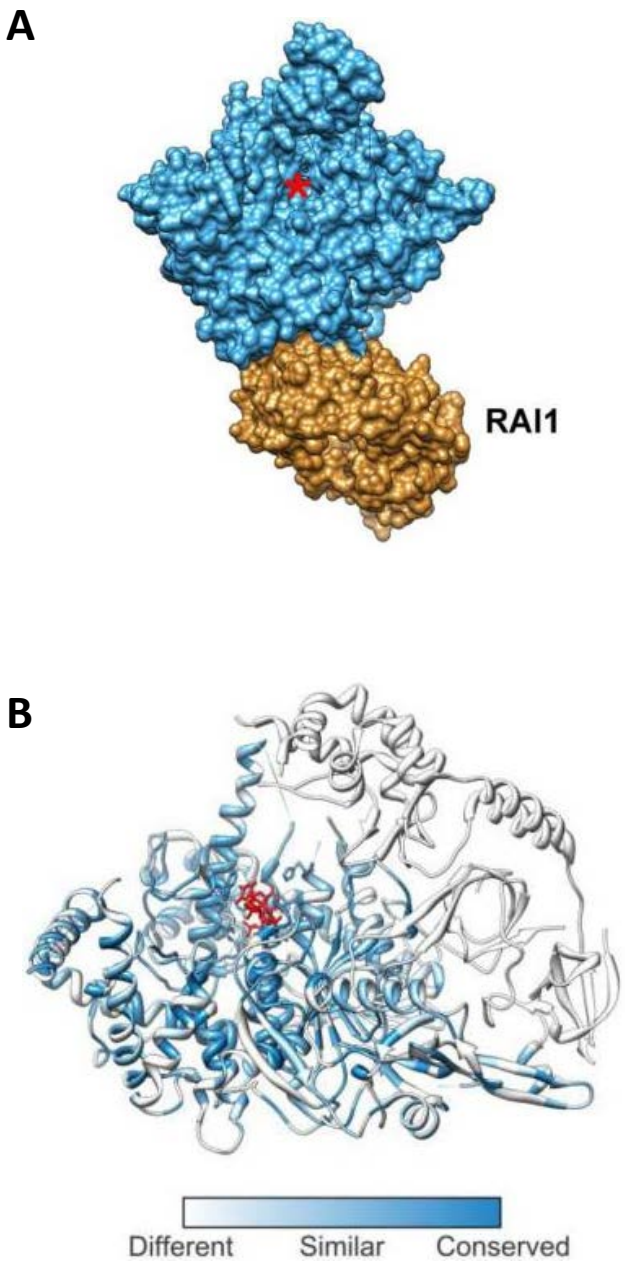


Figure 1.9. Structure of XRN2. A) Crystal structure of XRN2 in complex with RAI1 in *Saccharomyces pombe*. The blue area denotes XRN2, the brown area denotes its binding partner, RAI1. The active site of XRN2 is marked by the red asterisk (Nagarajan, Jones *et al.* 2013). B) Schematic showing the similarity of residues between XRN1 and XRN2. This shows that residues are more highly conserved around the active site, and less so towards the C-terminal, much of which is missing from XRN2 (Nagarajan, Jones *et al.* 2013).

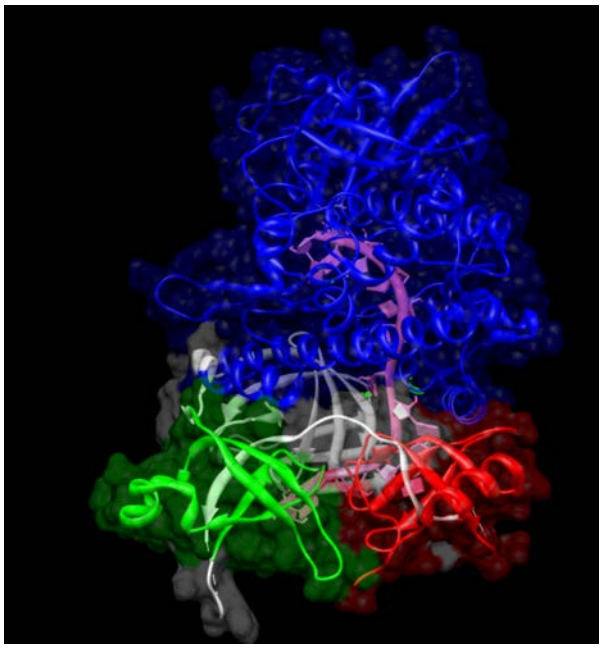
1.5.3 DIS3L2: Structure and functions

DIS3L2, like XRN1, is an independent cytoplasmic exoribonuclease which functions to degrade RNA in the 3' – 5' direction. DIS3L2 is highly conserved throughout evolution (Figure 1.10), and though its general structure has altered slightly across evolution, DIS3L2 in humans does not have an N-terminal domain which interacts with the exosome. The catalytic RNase domain, however, is highly conserved (Figure 1.12) (Malecki *et al.* 2013). DIS3L2 is thought to degrade ssRNA in a similar way to XRN1, utilizing the Brownian ratchet mechanism to rapidly degrade RNA, though in 3' – 5' direction. It has been shown to preferentially degrade RNAs which possess a highly uridylated 3' ends (Towler and Newbury 2018). In *Drosophila,* Dis3L2 interacts with the uridyl transferase (TUTase) called Tailor, which uridylates the 3' end and recruits Dis3L2 to degrade target RNA (Reimão-Pinto *et al.* 2016, Lin *et al.* 2017).

Work in the Newbury lab has shown that Dis3L2 specifically targets transcripts involved in growth and proliferation of specific tissues, using *Drosophila* models. Phenotypic analysis of wing imaginal discs where Dis3L2 had been knocked out showed that wings lacking Dis3L2 were significantly larger than wildtype wing imaginal discs (Towler *et al.* 2016). RNA-sequencing results of Dis3L2 mutant *Drosophila* supported this evidence by showing up regulation of growth factors when compared to the control. Targeting of growth factors by Dis3L2 contrasts with the targeting of pro-apoptotic transcripts by XRN1, suggesting that these enzymes may display some co-ordinate regulation to ensure correct tissue growth and development.

In humans, DIS3L2 mutations have been associated with the onset of Perlman's syndrome, a congenital overgrowth syndrome discussed in section 1.7.3. The fact that mutations result in overgrowth supports the notion that DIS3L2 is very important in regulating tissue growth, and that this activity is highly conserved. In addition, work in Chapter 4 of this project shows knock down of DIS3L2 in HEK-293T cells (a human embryonic kidney cell line) promotes proliferation of the cell line. This in line with the fact that DIS3L2 mutations are also linked to the development of Wilms' Tumour of the Kidney. Knock down of DIS3L2 in U-2 OS cells, on the other hand, do not result in cell line proliferation, suggesting that DIS3L2 has a tissue-specific effect, at least in humans.

A



B

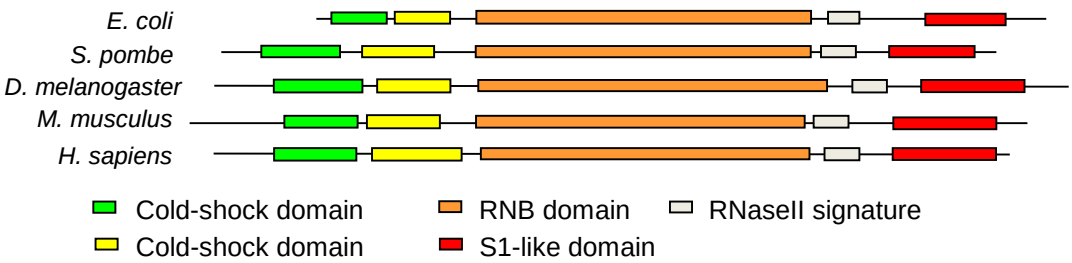


Figure 1.10. Structure and conservation of DIS3L2. A) A base-up view of the crystal structure of DIS3L2 in complex with poly(U) RNA substrate (pink). RNA is bound by two cold shock domains (red and green) at the base of the structure and fed through to the catalytic domain (blue) for degradation (adapted from Towler 2016). B) Conservation of DIS3L2 domains across evolution, demonstrating that DIS3L2 is highly conserved in all of its domains.

1.5.4 DIS3 and DIS3L1: Structures and functions

Given that DIS3 and DIS3L1 are catalytic paralogues of the nuclear (DIS3) and cytoplasmic (DIS3L1) exosome, these enzymes are discussed together. Both of these proteins display evolutionary conservation, with DIS3L2, as shown in Figure 1.12. The exosome is a barrel-shaped complex of proteins that degrades RNA in the 3' – 5' direction, through its core. Structurally, the exosome consists of 9 core proteins in a barrel-like structure. 6 proteins form the ring, and on top stands the cap of the exosome (Symmons and Luisi 2009, Fabre and Badens 2014). Enzymatic activity is catalyzed by either DIS3 or DIS3L1 (Figure 1.11) (or RRP6). The exosome proteins thread the RNA through the ring-like structure to the active site of DIS3 or DIS3L1 for 3' – 5' RNA decay or partial precursor RNA trimming (Kowalinski *et al.* 2016), depending on the whether it is nuclear or cytoplasmic RNA processing. The catalytic units of both the nuclear and cytoplasmic exosome differ from independent exoribonucleases in that they possess a Pilt N-terminal domain (PIN domain), which serves to tether the ribonucleases to the rest of the exosome (Schneider *et al.* 2009) and provides an endoribonuclease to DIS3.

1.5.4.1 DIS3

DIS3 is the catalytic unit of the nuclear exosome. The nuclear exosome functions to process global precursor RNAs into stable RNAs, as well as being pivotal in processing the 5.8S rRNA component of mature ribosomes (Houseley *et al.* 2006). It also acts as a surveillance mechanism to degrade defective mRNAs (such as those with premature stop codons) before they are exported out of the nucleus. The nuclear exosome, therefore, has a dual role in the nucleus: the first of which is to process RNAs post-transcriptionally to enable their maturation, and the second of which is to target and rapidly degrade defective mRNAs.

Despite their roles as catalytic units of the nuclear and cytoplasmic exosome respectively, DIS3 and DIS3L1 possess structural differences. For example, the PIN domain of DIS3 differs from that of its cytoplasmic paralogue, in that the PIN domain also enables DIS3 to catalyse endoribonucleolytic decay (Schneider *et al.* 2009). This characteristic marks DIS3 as functionally different from all of the other ribonucleases discussed throughout this project. DIS3 also, unsurprisingly, harbours a nuclear localisation domain, which is not present in DIS3L1.

The nuclear exosome requires additional protein cofactors in order for correct functionality. For example, in the nucleus, the TRAMP complex associates with the exosome. The TRAMP complex consists of a helicase, MTR4, which functions to first unwind structured RNAs prior to degradation by the exosome (LaCava *et al.* 2005, Slomovic and Schuster 2011). The TRAMP complex also consists of a poly(A) polymerase (TRF4/5) which is required for targeted

degradation of substrates. TRF4/5 polyadenylates cryptic unstable transcripts (CUTs), rRNAs and snoRNAs, which targets them for degradation by the nuclear exosome (LaCava *et al.* 2005).

DIS3 has been associated with human disorders including several types of blood cancer (discussed in section 1.7.1). The importance in cellular processes is also demonstrated by experiments in *Drosophila*, where it has been shown that DIS3 knock out results in a wingless phenotype (Towler *et al.* 2015).

1.5.4.2 DIS3L1

DIS3L1 is the catalytic unit of the cytoplasmic exosome (Staals *et al.* 2010, Tomecki *et al.* 2010). The activity of the exosome in the cytoplasm is in contrast with the role of the exosome in the nucleus. In the cytoplasm, the only known targets of degradation by the exosome are mRNAs. Here, the exosome regulates normal 3' – 5' RNA turnover after deadenylation, to help determine mRNA abundance (and subsequent translation rates) and to recognize and degrade cleavage products (van Hoof *et al.* 2000). In addition, the cytoplasmic exosome preferentially degrades mRNAs which have AREs (Mukherjee *et al.* 2002).

As in the nucleus, the cytoplasmic exosome also associates with other protein cofactors in order for correct RNA degradation to take place. In the cytoplasm, an example of these proteins includes the Ski complex (in yeast). The Ski complex associates with the cap of the exosome, and uses its helicase activity (via Ski2) to unwind RNA for feeding through the central channel of the exosome core for subsequent degradation (Halbach *et al.* 2013, Khemici and Linder 2018).

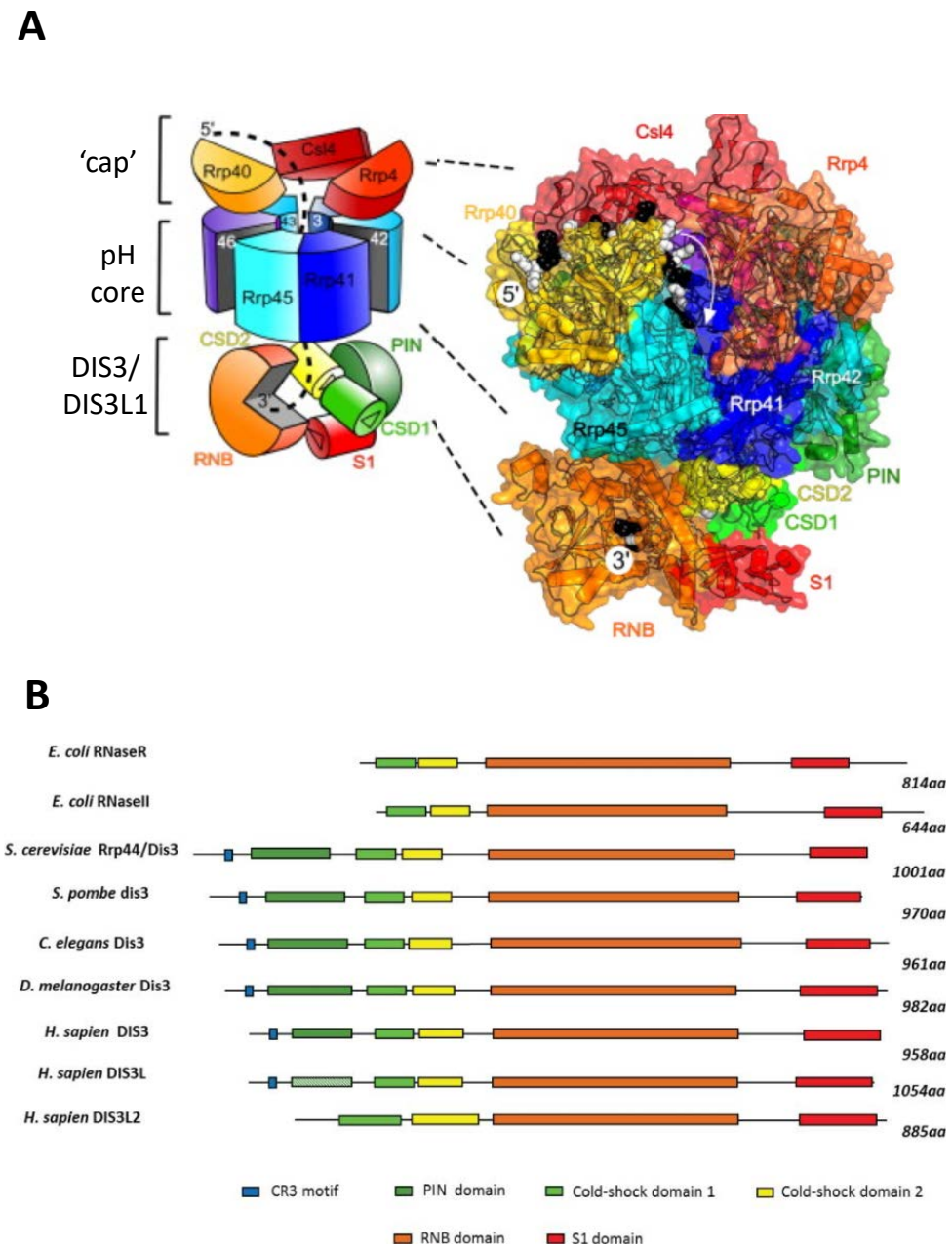


Figure 1.11. Structure of the exosome. A) A cartoon schematic of the full human exosome is shown, with its cap of RNA binding subunits (orange and red) atop a central channel RNase core of six subunits (light and dark blue). Below is the catalytic subunit of the exosome (nuclear subunit is DIS3, cytoplasmic is DIS3L1) (dark green through to orange and red). The structural arrangement crystal structure is also shown. This figure has been adapted from Symmons and Luisi 2009. B) A diagram to illustrate the high level of conservation of the DIS3 gene from *E. coli* through to *H. sapiens*, including the high level of conservation between *H. sapiens* DIS3, DIS3L1 (DIS3L in graphic) and DIS3L2. Adapted from (Robinson and Newbury 2015).

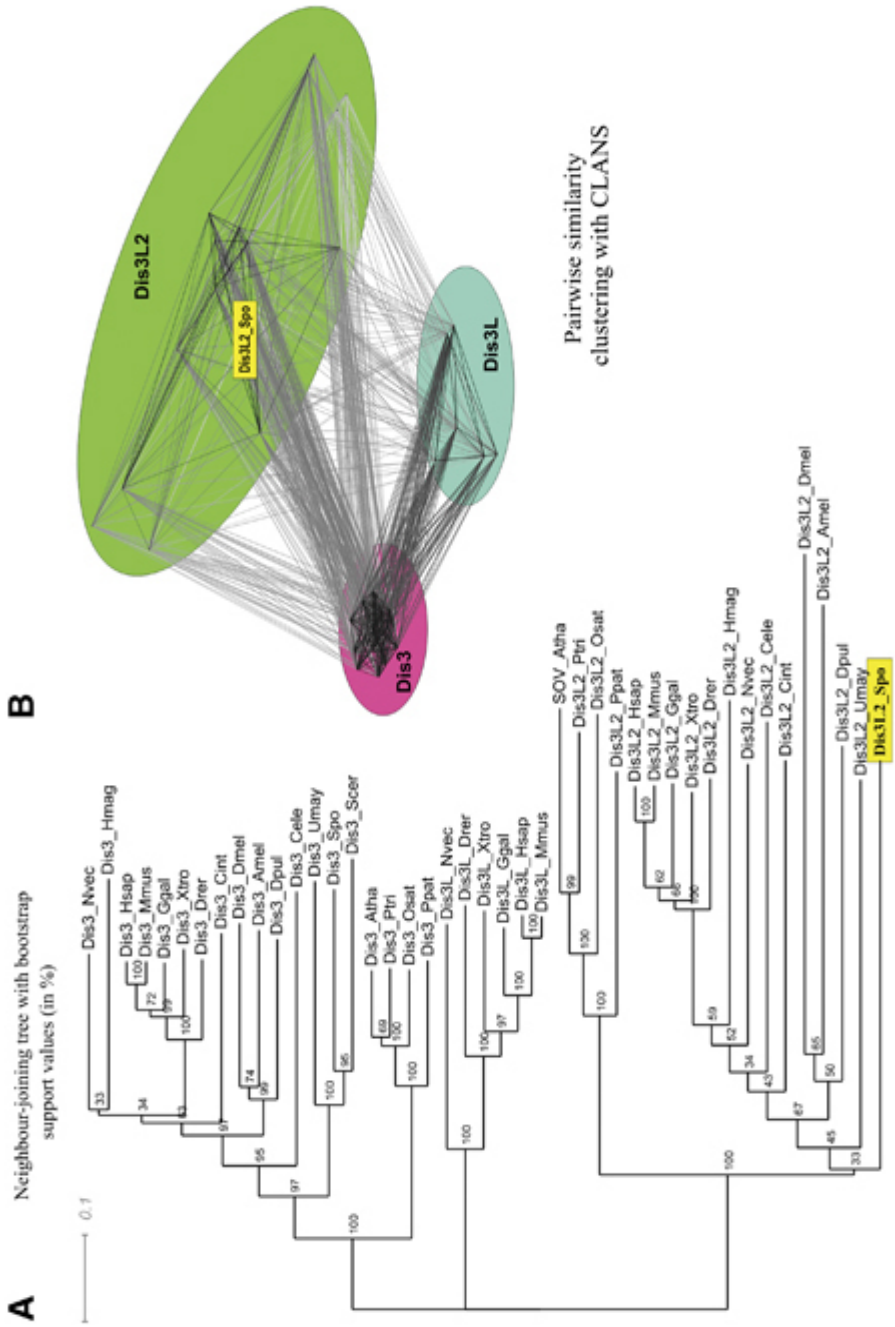


Figure 1.12. Evolutionary conservation of the DIS3 family member proteins. A) A neighbour joining tree of phylogenetic comparisons of 43 eukaryotic DIS3-like proteins. B) Clustering of the DIS3-like proteins on the basis of BLAST similarity usng CLANS, revealing distinct groups of DIS3, DIS3L1 and DIS3L2 sequences. Figure adapted from (Malecki, Viegas *et al.* 2013).

1.6 Pathways of RNA Decay

There exists multiple mechanisms by which specific RNAs are degraded. These include 5' – 3' RNA decay, 3' – 5' RNA decay, nonsense mediated decay (NMD) (and associated endonucleolytic decay) and degradation of long noncoding RNAs and miRNAs.

Once exported from the nucleus, mRNA is normally protected from degradation by the 5' methylguanosine cap, and the 3' poly(A) tail, which stabilizes the transcript. In order for mRNA to be degraded, these protective features must first be removed to allow access to the ends by exoribonucleases. In the case of 5' – 3' RNA decay this occurs in a two-step process, whereby first the poly(A) tail is removed by deadenylases (such as the CCR4-NOT and PAN2-PAN3 complexes), followed by the removal of the 5' methylguanosine cap by the decapping complex (including DCP1 and DCP2). The transcript is then primed for degradation by XRN1. Alternatively, for 3' – 5' decay, only the poly(A) tail needs to be removed in order for decay by the exosome or DIS3L2 to take place.

During NMD, mRNA decay is facilitated by endonucleolytic cleavage within the mRNA transcript, resulting in exposure of two vulnerable ends which are then targeted for degradation by XRN1 and the exosome or DIS3L2 respectively.

Although the pathways of RNA decay have remained highly conserved throughout evolution, there does exist some differences between yeast, *Drosophila* and human models. For example, in *S. cerevisiae* there is only one 3' – 5' exoribonuclease: DIS3 (known as Rrp44), whereas in *Drosophila*, there are two 3' – 5' paralogues: DIS3 and DIS3L2. In humans, this increases to three 3' – 5' paralogues: DIS3, DIS3L2 and DIS3L1 (reviewed in (Towler and Newbury 2018)). This divergence in evolution shows how eukaryotic RNA degradation became more complex, as organisms became more complex.

1.6.1 Summary of 5' – 3' RNA decay

1.6.1.1 Deadenylation

The first step in priming RNA transcripts for decay is the removal of the 3' poly(A) tail. As mentioned previously, the length of the poly(A) tail is crucial in determining the stability of the transcript, and it is likely that shorter poly(A) tails are less stable, because they are too short to be able to bind the protective Poly(A) binding proteins (PABPs), which need to be removed from longer poly(A) tails in order for degradation to take place. Because of protective PABPs (such as

the PAB1p-dependent poly(A) nuclease – PAN protein), deadenylation is usually a two-step process: removal of the protective PABP and recruitment of deadenylases to remove the poly(A) tail.

During the first step, the removal of the PABP is a major rate-limiting step (Brook and Gray 2012). PABP usually interacts with the translation initiation factor, eIF4G, and the 5' methylguanosine cap to form a closed loop mRNA conformation, to enhance ribosomal recruitment (Kahvejian *et al.* 2005). PABP1 is the most extensively studied PABP in mammals, though there are 5 known mammalian PABPs. PABP1 contains four RNA recognition motifs (RRMs), RRMs 1 and 2 bind the poly(A) tail with high affinity, and RRMs 3 and 4 bind AREs (Brook and Gray 2012). In the beginning, the poly(A) tails are shortened by the nuclease domain PAN2-PAN3 complex. This involves distributive trimming of the first approximately 110 nucleotides (there are 250 poly(A) nucleotides in mammals) by PAN2, stimulated by interaction between PAN3 and PABP1. This occurs until the poly(A) tail is too short for the PABP to bind (Wolf and Passmore 2014). PABP proteins are also thought to be displaced by cofactors recruited by both the GW182 protein and CCR4-NOT complex during decay of miRNA targets (Zekri *et al.* 2013). The paradoxical nature of the PABP protein is that the very act of protecting the mRNA can lead to the recruitment of deadenylation factors, and accelerated deadenylation.

At this point, the second stage of deadenylation occurs, whereby the CCR4-NOT complex removes the remaining A nucleotides. The CCR4-NOT complex consists of 2 highly conserved modules: the CNOT1/2/3 proteins (which provide a scaffold for the other subunits of the complex) and 2 deadenylases: CCR4 and CAF1 (Mayya and Duchaine 2019). As deadenylation continues, the CCR4-NOT complex recruits the decapping complex needed for removal of the 5' methylguanosine cap.

1.6.1.2 Decapping

Following deadenylation, an mRNA can be further processed so that it can undergo 5' – 3' decay by XRN1. The CCR4-NOT complex, utilised for deadenylation, is physically coupled to decapping by the DEAD-box helicase DDX6. This protein directly interacts with the CNOT1 subunit and decapping factors, as well as eIF4E-transporter, to increase the local concentration of decapping factors around the 5' cap (Nishimura *et al.* 2015). Decapping is initiated by the LSM1-7 complex and PAT1. PAT1 interacts with both the LSM1-7 complex and the decapping factor, DCP2, as well as DDX6. It has also been found to tightly associate with the CCR4-NOT complex, serving as the scaffold to bridge decapping and deadenylation (Wu *et al.* 2014). LSM1-7 complex members, Lsm2 and Lsm3, bridge the interaction of the LSM1-7 complex with PAT1 in order to stimulate decapping by the decapping complex. The LSM1-7/PAT1 complex initially

associates with a 10 nucleotide overhang left over by the deadenylation process at the 3' end (Chowdhury *et al.* 2007), protecting the 3' end from further trimming and 3' – 5' mediated degradation, as well as promoting decapping and 5' – 3' degradation (Tharun *et al.* 2005).

The decapping activators, PAT1, LSM1-7, DDX6 and DCP1 activate DCP2 decapping activity by inducing a conformational change in the C-terminal of DCP2, allowing it to adopt its active, closed conformation (She *et al.* 2008). This is triggered through association of DCP1 and DCP2 into the decapping complex, where DCP2 provides the catalytic activity, and hydrolysis of the methylguanosine cap is performed, via threading of the RNA through the large C-terminal domain of DCP2.

It is important to note that it has recently be observed that there are multiple decapping factors which differentially target specific RNAs depending on unique target features of individual RNAs (He *et al.* 2018). Here it was shown that decapping activators PAT1, LSM1, and DDX6 target specific subsets of yeast transcripts with overlapping substrate specificity, and RNA sequencing showed that different members of the decapping activator machinery differentially regulate subsets of transcripts. An earlier example showed that transcripts undergoing miRNA-mediated silencing recruit the DCP1/DCP2 decapping complex through GW182, a member of the RISC, rather than the LSM1-7/PAT1 complex (Rehwinkel *et al.* 2005).

1.6.1.3 Degradation is performed by XRN1

Once the methylguanosine cap has been removed, the transcript is vulnerable to degradation by XRN1 at the 5' end. To increase the efficiency of degradation, XRN1 associates with the decapping complex, specifically EDC4 (in humans) (DCP1 in *Drosophila*), at a proline-rich motif in the unstructured C-terminal region, coupling decapping and rapid 5' – 3' degradation (Braun *et al.* 2012) (see section 1.4.1 for XRN1 mechanism of action).

1.6.1.4 Decapping machinery is localized to P-bodies

Conventionally, it has been accepted for many years that the machinery dedicated to 5' – 3' degradation is compartmentalized into cytoplasmic Processing bodies (P-bodies). Traditionally, P-bodies are sites of mRNA storage, whereby translationally repressed mRNAs are stored until such time that they re-enter translation, or are degraded (this idea was first put forward by Roy Parker) (Sheth and Parker 2003). More recently, they have been described as a way for cells to adapt to stress quickly, through altering the gene expression landscape (Guzikowski *et al.* 2019).P-bodies, like other cellular granules, are a composition of mRNAs in complex with proteins associated with translational repression and 5' – 3' mRNA decay. They are conserved in eukaryotes, bearing similarities to other ribonucleoprotein (RNP) granules (such as Cajal bodies,

and stress granules). They depend on complex networks of protein-RNA interactions, and liquid-liquid phase separation in order for their formation (Luo *et al.* 2018).

P-bodies are a condensation of RNPs, resulting from a switch in state to coassemble into liquid or solid bodies. This phase transition has recently been reconstituted *in vitro* (Alberti and Hyman 2016), and more recently Hubstenberger *et al.* (2017) developed a fluorescence-activated particle sorting method to identify hundreds of proteins and thousands of mRNAs, displaying a dense network of interactions. They show how the condensation of RNPs into P-bodies is linked to the fate of large pools of RNA. This research also showed that more than one-fifth of cytoplasmic transcripts transit through P-bodies, however, there is a network of specific protein-RNA and protein-protein interations which separate mRNPs from other transcripts (Hubstenberger *et al.* 2017). This study supports Guzikowski *et al.,* in that P-body mRNAs generally encode regulatory processes, thereby providing a reservoir for quick adaptation of gene expression.

P-bodies sequester 5' – 3' degradation machinery and RNAs away from the cytoplasm both during cellular stress and during the normal cell cycle to provide a layer of post-transcriptional gene regulation to directly impact on mRNA levels, translation and cell survival (Ivanov *et al.* 2018, Guzikowski *et al.* 2019). Although it was first identified as a P-body marker in 1997 (Bashkirov *et al.* 1997), XRN1 has since been found to be not just present in P-bodies, but also distributed as a free enzyme in the cytoplasm for the clearance of endonucleolytically cleaved RNAs in the cytoplasm during nonsense mediated decay (discussed later) (Conti and Izaurralde 2005). In addition, P-bodies have also been shown to be co-translationally active, in that some deadenylated transcripts remain on polyribosomes (Hu *et al.* 2009). Given that XRN1 is involved in degradation due to nonsense-mediated decay and miRNA-mediated decay (discussed below), it is reasonable to assume that degradation machinery is localized, but not confined to, P-bodies, as was previously thought, and the issue around whether P-bodies are the site of translational repression and mRNA decay remains contentious (Luo *et al.* 2018)

1.6.2 Summary of 3' – 5' RNA decay

3' – 5' decay operates in a very similar manner to 5' – 3' decay, in that it requires deadenylation of the 3' poly(A) tail. It is not coupled with removal of the 5' methylguanosine cap. The removal of the 3' poly(A) tail is discussed above, however, it would be prudent at this point to mention that there are other deadenylases involved in deadenylation, aside from the PAN3-PAN3 and CCR4-NOT complexes. In addition to these complexes, humans have another deadenylase, poly(A) specific deadenylase complex (PARN), which could indicate that there is target specificity of the various deadenylases in humans. This has already been shown in mouse

models, where PARN targeted a discrete set of mRNAs involved in cell motility (Lee *et al.* 2012). More recently, PARN has been shown to be involved in regulating miRNA stability (Shukla *et al.* 2019).

The removal of the poly(A) tail leaves the 3' end of the transcript vulnerable to attack by the DIS3 family of exoribonucleases, whether this be by the exosome (DIS3/DIS3L1) or by independent cytoplasmic ribonuclease, DIS3L2. The catalytic activities of these enzymes is described above (sections 1.4.3 and 1.4.4).

Substrate targeting by the nuclear exosome is facilitated by its cofactor, MTR4, which links to RBP adapters. Examples of this include the nuclear exosome targeting complex (NEXT) and the poly(A) tail exosome targeting connection (PAXT). These complexes primarily target early, unprocessed transcripts for degradation by the nuclear exosome. Each of these mechanisms relies on the binding of zinc finger proteins and the activities of MTR4 and DExH/D box RNA helicase in order effective degradation to take place (Meola *et al.* 2016). The nuclear exosome can also be facilitated by the TRAMP complex (Lubas *et al.* 2011). In the cytoplasmic, the exosome is facilitated by the Ski complex, which is instrumental in threading RNA through the core of the exosome via its helicase activity (Zinder and Lima 2017).

In the cytoplasm, degradation by the exosome-independent ribonuclease, DIS3L2 (Lubas *et al.* 2013), is facilitated by the addition of uridine residues to the 3' end. This is catalyzed by TUTases, and provides a landing pad for DIS3L2 to engage in RNA degradation (Chang *et al.* 2013, Łabno *et al.* 2016).

Once degradation of the transcript has taken place, scavenger decapping enzymes (DCPS in mammals) degrades the remaining mRNA fragment with the 5' cap. DCPS, unlike the exoribonucleases, is capable of efficiently degrading capped RNA substrate providing the fragment is no longer than 10 nucleotides (Milac *et al.* 2014).

A summary schematic of the different RNA decay pathways are shown in Figure 1.13.

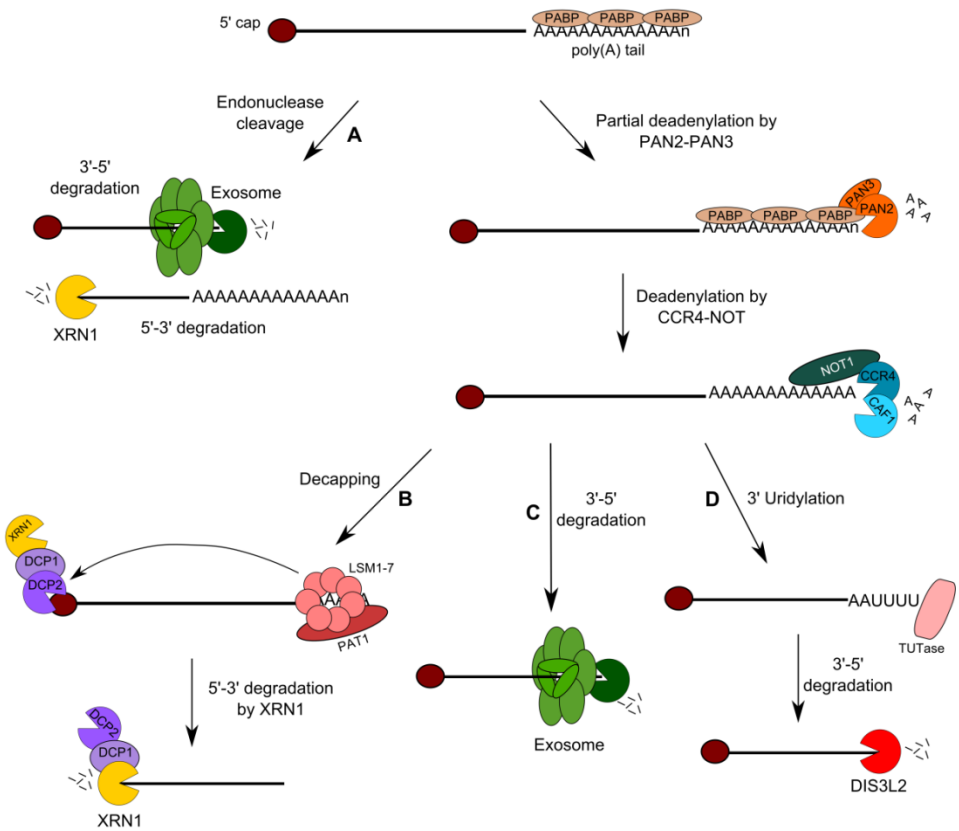


Figure 1.13: Cartoon depicting the major RNA degradation pathways. Bulk eukaryotic mRNA degradation is deadenylation dependent, which is predominantly carried out by the CCR4-NOT complex. A) Endonuclease cleavage of mRNA results in 2 unprotected strands, allowing degradation of the 3’ strand by the exosome and degradation of the 5’ strand by XRN1 to occur without the need for deadenylation or decapping. B) During 5'-3' degradation, the LSM1/PAT1 complex binds at the 3‘ end and activates the decapping complex. The unprotected 5’ ends are then degraded by XRN1. C) Following deadenylation, degradation can occur in the 3' - 5' direction by the exosome complex, which uses the exoribonucleases DIS3 and Rrp6 in the nucleus or DIS3L1 in the cytoplasm. D) Deadenylated mRNAs can also be uridylated by uridyl transferases (TUTases) which then target the RNA for degradation by the exosome independent 3'-5' exoribonuclease, DIS3L2. mRNAs can also be cleaved internally, leaving unprotected 3’ and 5’ ends which can be degraded by all three pathways (Braun, Truffault *et al.* 2012)(Fillman and Lykke-Andersen 2005).

1.6.3 Summary of Nonsense-Mediated Decay

Nonsense-mediated decay (NMD) describes an RNA surveillance mechanism designed to reduce errors in gene expression by eliminating RNA transcripts which contain premature stop codons. This occurs to avoid translation of aberrant mRNAs which could lead to deleterious gain-of-function protein expression (Conti and Izaurralde 2005). These errors occur as a result of errors made during transcription and defective splicing, or through damage to the RNA (Peccarelli and Kebaara 2014). It also leads to the degradation of transcripts with nonsense mutations in order to reduce the level of transcriptional noise (Mendell *et al.* 2004). Recent studies have also shown that NMD also regulates the stability of natural transcripts that play significant roles in cell functions (Han *et al.* 2018) via several NMD activating features, such as an exon-junction complex (EJC), upstream open reading frame and a UGA selenocysteine codon (in certain selenoprotein-encoding mRNAs), as well as a long 3' UTR (Kurosaki *et al.* 2019).

NMD targets multiple species of RNA, including mRNA and lncRNA, despite being often referred to as a protein-folding quality control mechanism, whereby premature termination codons (PTCs) are detected to eliminate mRNAs which will not fold properly in connection with translation termination (Popp and Maquat 2013). During the initiation of NMD, RNA helicases translocate along the RNA, aiming to unwind secondary structures and acting to remodel RNA-protein complexes. They can act as signaling molecules, to direct and recruit NMD machinery to the site (Rajkowitsch *et al.* 2007). The central component of the NMD machinery is the protein UPF1/SMG2. Phosphorylation of UPF1 is orchestrated by SMG1 (a PI3K-like kinase), and two additional subunits, SMG8 and SMG9 (Yamashita *et al.* 2009). Initially, UPF1 and SMG1 associate to form a clamp around the RNA, directly interacting with eukaryotic release factors, eRF1 and eRF3, which form the surveillance complex, SURF, which surveys the vicinity around the PTC (Kashima *et al.* 2006, Hug *et al.* 2016). Following a major conformational change of the SMG complex of proteins, the SURF complex interacts with UPF2, UPF3b and an EJC downstream of the PTC, forming the decay-inducing complex (DECID) (Deniaud *et al.* 2015). This triggers UPF1 phosphorylation and dissociation of the SURF complex, and UPF1 adopts its active helicase conformation, acting as an RNPase to unwind secondary structure and clear the way for the decay factors (Hug *et al.* 2016).

The activated NMD complex (formed of UPF1, UPF2, and UPF3b) translocates from its upstream position and towards the 3' end of the EJC, where UPF1 associates with SMG5, 6 and 7, and RNA decay factors. SMG7 recruits POP2, which catalyses deandenylation and initiates decapping for degradation by XRN1 (Unterholzner and Izaurralde 2004) after endonucleolytic decay of the RNA strand. Endonucleolytic decay takes place between the PTC and the STOP

codon, producing two separate RNA strands with a vulnerable 3’ and 5’ end, respectively. This leaves the strands open to degradation by XRN1 and the cytoplasmic exosome or DIS3L2 (Hug *et al.* 2016, Lloyd 2018) (Figure 1.14). It is important to note that NMD occurs by cleavage or deadenylation in humans, by cleavage only in *Drosophila* and deadenylation only in yeast, due to the absence of specific UPFs and SMGs in *Drosophila* and yeast, as a result of divergent evolution (Lloyd 2018).

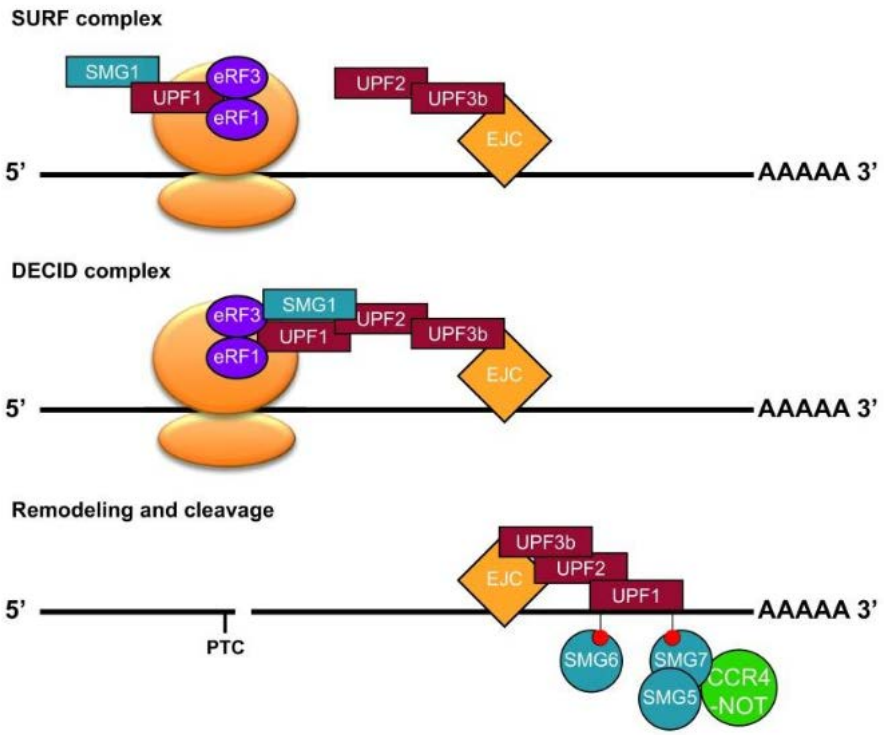


Figure 1.14. Brief overview of nonsense-mediated cleavage of mRNA. The activation of NMD begins with the recruitment of UPF1 and SMG1 to termination events by eRF1 and eRF3, leading to formation of the SMG-UPF1-eRF1-eRF3 (SURF) complex. Recruitment of UPF2 and UPF3 leads to the formation of the DECID complex, which in turn facilitates the phosphorylation of UPF1 by SMG1. The ribosome then dissociates from the transcript, and SMG5/6/7 are recruited to the transcript by UPF1 binding. The transcript is finally degraded after endonucleolytic cleavage by SMG6. The CCR4-NOT complex is recruited by SMG7/5 to the fragment harbouring the 3' UTR, for subsequent deadenylation and degradation by 3' – 5' exoribonucleases. Decapping enzymes are recruited to the 5' UTR fragment to promote degradation by XRN1. This figure has been adapted from Lloyd 2018.

1.6.4 Summary of microRNA-mediated decay

1.6.4.1 MicroRNA processing

microRNA (miRNA) genes, the majority of which are in intergenic regions of the genome, are first transcribed in the nucleus by RNA polymerase II (or, occasionally, RNA polymerase III). Clustered genes encoding miRNAs are often transcribed together as polycistronic transcripts, which then undergo individual processing into individual mature miRNAs. This is a multi-step process. First, miRNA genes are transcribed as capped, polyadenylated primary miRNAs (pri-miRNAs) by RNA polymerase II. Pri-miRNAs are generally kilobases long, with several stem-loop structures. Embedded within the stem loop structures are the 20-25 nucleotide sequences which form the mature miRNA. During the second step, the pri-miRNA is processed within the nucleus by the microprocessor complex of proteins, including Drosha, a class 2 RNase III enzyme, the double-stranded RNA binding protein (dsRBP), DGCR8 and several RNA helicases (DDX5 and DDX17) (Treiber *et al.* 2019). Drosha cleaves the pri-miRNA at the stem of a hairpin structure, and this releases a 60-70nt single hairpin structure precursor miRNA (pre-miRNA). Following nuclear processing by Drosha, the pre-miRNA is exported by exportin-5 (EXP5) into the cytoplasm for maturation. Maturation of the pre-miRNA in the cytoplasm is catalyzed by another RNase III-type protein, Dicer. Dicer is highly specific and cleaves the pre-miRNA structure exactly 22nt from the pre-existing pre-miRNA terminus. The mature miRNA duplex is finally loaded onto the Argonaute (AGO) proteins. At this point, AGO proteins loaded with miRNAs dissociate from Dicer, and form the miRISC (Wahid *et al.* 2010, Libri *et al.* 2013, O'Brien *et al.* 2018, Treiber *et al.* 2019) (Figure 1.15). One strand of the miRNA duplex (the guide strand) is preferentially retained in the miRISC by AGO in order for it to bind complementary seed sequences, located mainly in the 3' UTR of target mRNAs (Treiber *et al.* 2019), whereas the passenger strand is released, and thought to be degraded by ribonucleases (Pitchiaya *et al.* 2017). The preference for which strand serves as the guide strand can be dictated by the presence of U or C in the 5' most nucleotide of the miRNA, and for an A in the target sequences facing this position (Duchaine and Fabian 2019).

1.6.4.2 Mechanism of action of miRNAs

miRNAs have many roles within the cell. Some can be involved in cell-cell communication through extracellular secretion, and others can be utilized for gene expression regulation. This aspect of the introduction will focus on how miRNAs regulate gene expression. miRNAs can both inhibit gene expression, activate gene expression, and be involved in regulation of gene expression at multiple levels, including utilizing different subcellular compartments to regulate different aspects of gene expression.

The first aspect of gene regulation by miRNAs is how they can interact with newly synthesized mRNA before it leaves the nucleus. When acting as part of the miRNA Induced Silencing Complex (miRISC), where they form the nucleic acid core of the complex (Bartel 2004), the low molecular weight of the complex means it can interact with nuclear mRNAs via the miRNA response elements (MREs) which are complementary sequences on target mRNA, to facilitate their degradation within the nucleus (Pitchiaya *et al.* 2017). The miRISC constitutes a complex of Argonaute proteins and GW182. The interaction between AGO proteins and GW182 is essential for miRNA-mediated gene silencing, and recently, it has been shown that miRNA binding increases the affinity of AGO to GW182 via a hook structure in the N-terminus of GW182 (Elkayam *et al.* 2017). It was also shown that a single GW182 can recruit up to three copies of AGO using all three AGO-binding motifs (Figure 1.16A). This physical interaction is essential for facilitating downstream gene silencing. The docking of AGO onto GW182 was revealed as a gate-like interaction, and it was shown that both AGO1 and AGO2 have a single GW182-binding site within the Piwi domain, the affinity of which is enhanced by miRNA binding. This ensures that only the mature RISC can be recruited for silencing. A fully complementary MRE induces AGO2 endonuclease activity and cleavage of the mRNA (Jo *et al.* 2015). It is also thought that whilst in the nucleus, the miRISC may influence transcription directly. Studies have shown that when AGO2 is concentrated in the nucleus of senesced fibroblasts, it interacts with the miRISC to suppress the transcription of proliferation-promoting genes regulated by Rb/E2F (Benhamed *et al.* 2012). Indeed, miRNAs are also capable of influencing transcription when they are not part of the miRISC: for example, a recent study showed that *miR-522* is able to interact with a DNA cruciform structure within the promoter of CYP2E1 to suppress its transcription (Miao *et al.* 2016).

In the cytoplasm, the miRISC is involved with both endonucleolytic mRNA degradation, and translational inhibition when they form complexes with mRNAs associated with polysomes (Pillai *et al.* 2005). The miRISC inhibits translation initiation by interfering with eIF4E cap recognition, 40S small ribosomal subunit recruitment, or by preventing 80S ribosomal subunit formation through antagonizing 60S subunit joining (Figure 1.16B).

The Argonaute elements of the miRISC associate with the miRNA, which in turn associates via its seed sequence to the target mRNA, usually at sequences within the 3' UTR, where it shares base complementarity. At this point, cleavage, translational repression or destabilization of the transcript occurs (Wang *et al.* 2019). Cleavage of the RNA strand can occur by the slicing of miRNA targets by AGO1, which generates two truncated cleavage products (RISC 5' and RISC 3'), which are subsequently degraded by the RNA decay enzymes (Chantarachot and Bailey-Serres 2018).

The C-terminal domain of GW182 is essential for target mRNA silencing, and has been termed the silencing domain. The silencing domain is subdivided into a number of unstructured regions which flank an RNA-recognition motif- like domain, engendering gene silencing by interacting with a number of RBPs (Duchaine and Fabian 2019). The silencing domain of GW182 in the miRISC is thought to display most of the translational repression activity of the complex, and interacts with the CCR4-NOT and PAN2-PAN3 deadenylase complexes to instigate deadenylation of the RNA (Chekulaeva *et al.* 2011).The GW182 proteins also interact with the PABP, all of which suggests that GW182 proteins function as scaffold proteins in the miRISC (Braun *et al.* 2013), by interfering with the PABP-eIF4G association and mRNA circularization (Zekri *et al.* 2009). In turn, the interaction of GW182 with the deadenylation proteins leads to decapping by the DCP1/DCP2 complex, and facilitates degradation by XRN1 (Fabian and Sonenberg 2012) in 5' – 3' degradation (discussed above) (Figure 1.16). An elegant single-molecule imaging system showed in 2015 that the miRISC could slide along the 3' UTR of the target mRNA between miRNA binding sites, representing a scanning activity that could accelerate target recognition by reducing collisions along individual 3' UTRs (Chandradoss *et al.* 2015).

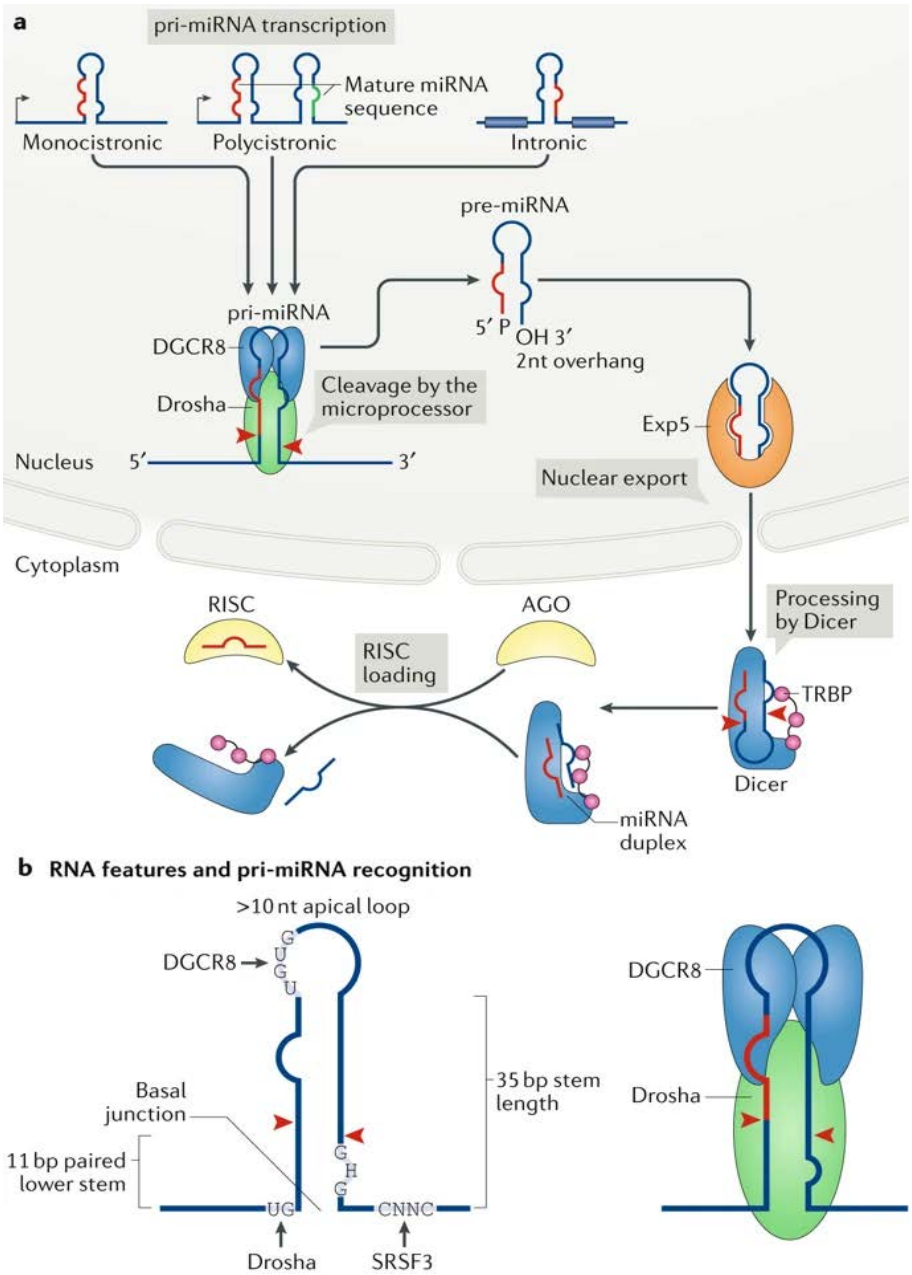
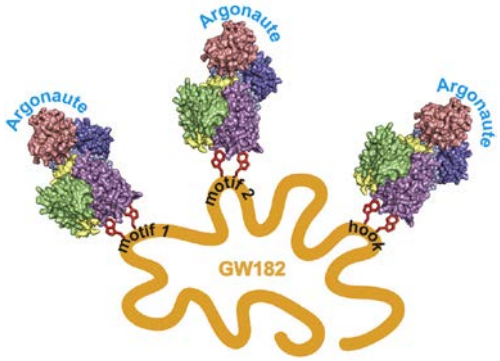


Figure 1.15. Overview of miRNA biogenesis. Adapted from Treiber *et al.*, this figure shows how miRNAs are generated from sequences within coding genes. A) Pri-miRNAs are first transcribed by RNA polymerase II in the nucleus. B) This pri-miRNA is then cleaved by Drosha and DGCR8 into a pre-miRNA. The pre-miRNA is exported out of the nucleus and into the cytoplasm where it is further processed by Dicer into the mature miRNA. Once the mature miRNA is loaded onto the RISC, it targets complementary sequences in mRNA to facilitate endonucleolytic cleavage and subsequent degradation by exoribonucleases (Treiber, Treiber *et al.* 2019).

A



B

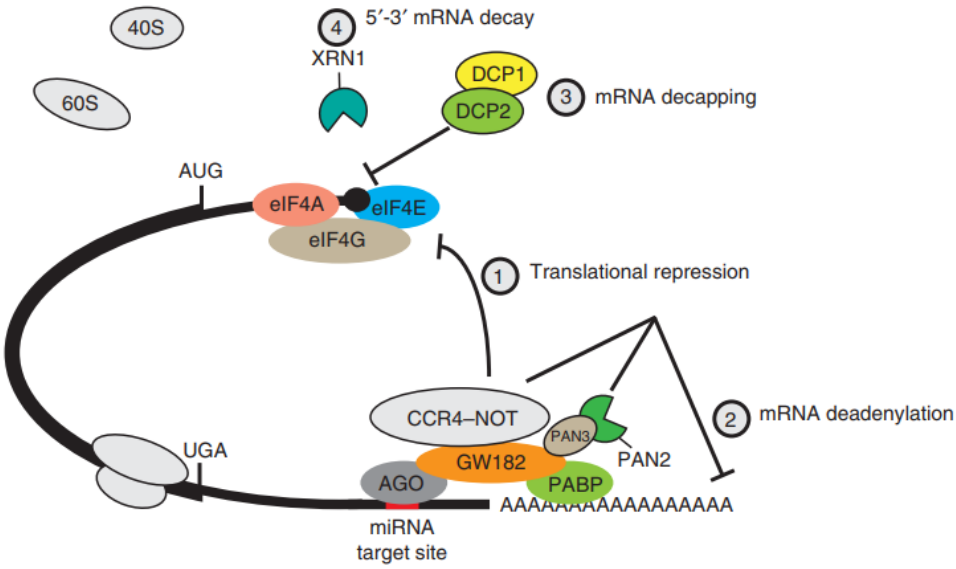


Figure 1.16. miRNA-mediated gene silencing. A) The association of multiple AGOs to a single GW182 protein enhances miRNA-mediated gene silencing, the affinity of this binding is strengthened by the presence of a miRNA, adapted from (Elkayam, Faehnle *et al.* 2017). B) Model of miRNA-mediated gene silencing, whereby in the first instance the miRISC represses cap-dependent translation at the initiation step in a deadenylation-independent manner by blocking mRNA circularization. In turn, the deadenylation complex, CCR4-NOT is recruited to deadenylate the mRNA. This leads to decapping by DCP1/2 and 5' – 3' RNA decay by XRN1. Adapted from (Duchaine and Fabian 2019).

1.6.4.3 miRNA degradation

As well as being regulators of gene expression, miRNAs are themselves targeted for degradation. A recent publication has identified that the RNase, Poly(A) Specific RNase (PARN), regulates selected miRNAs by deadenylating the 3' ends. PARN deadenylation prevents the degradation of miRNAs in the 3' – 5' direction by the cytoplasmic exosome and DIS3L2 (Shukla *et al.* 2019). The miRNAs targeted by PARN have been shown to be involved in the regulation of p53 in human cells, which highlights the clinical relevance of this fine-tuned mechanism of regulating gene expression by miRNAs.

PARN regulates miRNAs by stabilizing either mature or precursor miRNAs by removing the poly(A) tails added by the PAP protein, PAPD5 (Figure 1.17). RNA sequencing technology showed that knock down of PARN resulted in destabilization of 157 miRNAs, some of which were involved with repressing p53 translation, leading to an accumulation of this notorious oncogene (Shukla *et al.* 2019). These included mature forms of *miR-21-5p* and *miR-181b-5p*. This group identified a potentially novel therapeutic role for regulation of miRNA expression. Loss of PARN has also been associated with the onset of a severe form of dyskeratosis congenita (DC), a hereditary disease, whereby it was shown that mutations in PARN lead to the onset of a form of DC called Hoyeraal- Hreidarsson syndrome, which causes abnormally short telomeres and congenital defects (Burris *et al.* 2016). It is thought that this occurs as a result of defective regulation of miRNAs, as well as altered mRNA stability (Dhanraj *et al.* 2015).

Another recent publication noted the role of Tudor S/N-mediated endonucleolytic degradation of some miRNAs. In this study, they showed that the helicase, UPF1 (also involved in NMD), promotes Tudor S/N-mediated RNA decay (Elbarbary *et al.* 2017). Tudor S/N (TSN) is an evolutionarily conserved nuclease which contains a Tudor domain and five staphylococcal/micrococcal-like nuclease domains, and was the first constituent of the miRISC shown to have endonuclease activity (Caudy *et al.* 2003). In the present study, it was shown that UPF1 facilitates the dissociation of miRNAs from their mRNA targets while they are in complex with the miRISC, which makes the miRNA susceptible to Tudor-mediated miRNA decay (TumiD) by TSN (Figure 1.18) (Elbarbary *et al.* 2017).

Other work has shown that miRNAs can also be regulated in the nucleus by XRN2-mediated decay in *Caenorhabditis elegans* (Chatterjee and Großhans 2009). In this study, they showed that processing of precursor miRNA by Dicer and subsequent loading of mature miRNA onto the miRISC are coupled processes which precede degradation of mature miRNA. Efficient release of the miRNA from the miRISC exposes the miRNA to degradation by XRN2, which can be blocked by the addition of miRNA target RNA (Chatterjee and Großhans 2009).

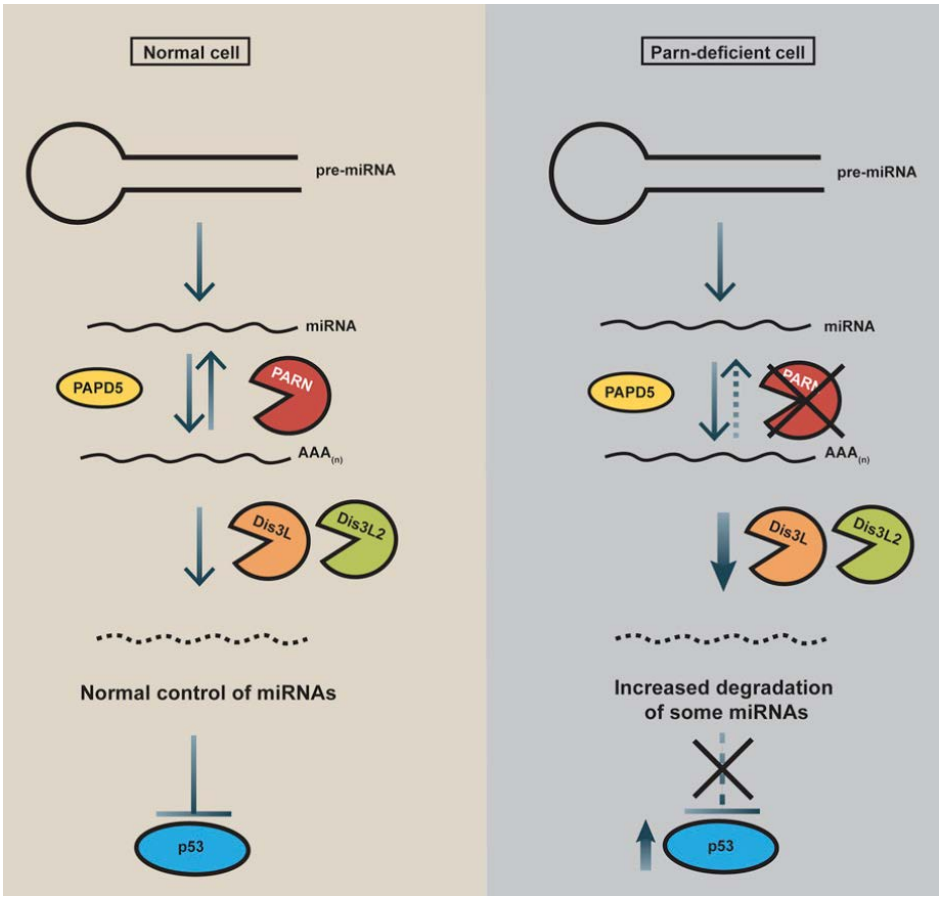


Figure 1.17. PARN-mediated stabilisation of miRNAs. PARN mediates stabilisation of miRNAs by deadenylating the 3' poly(A) tail. Stabilisation of miRNAs in this way leads to repression of p53 translation. Knock down of PARN results in increased degradation of miRNAs, leading to up regulation of p53 translation and accumulation of p53 protein in the cell. This figure has been adapted from Shukla *et al.* 2019.

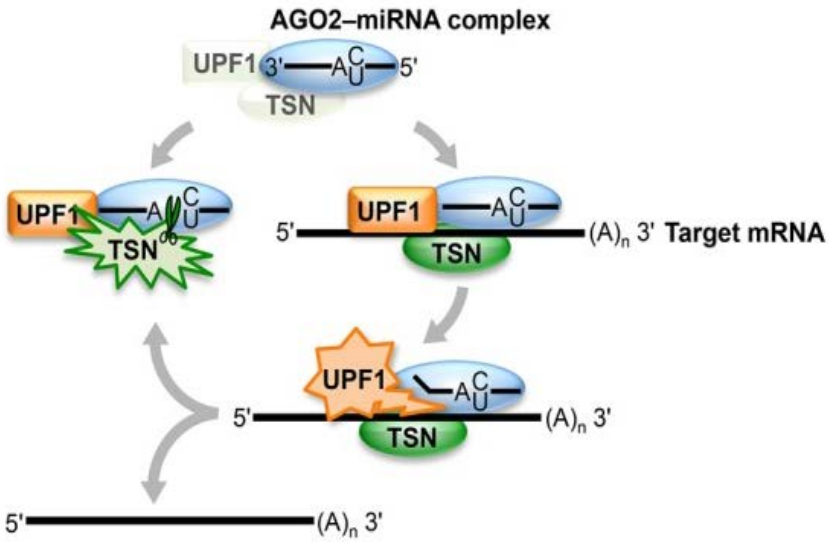


Figure 1.18. Proposed model for miRNA degradation by TumiD. Once TSN-UPF1 gains access to the miRNA in the miRISC (blue), UPF1 helicase activity frees the miRNA from the mRNA, making it susceptible to TSN-mediated endonucleolytic cleavage at accessible CA and UA dinucleotides. This figure has been adapted from Elbarbary, Miyoshi *et al.* 2017.

1.7 RNA stability: implications in human disease

Recently, RNA stability and RNA decay has been increasingly implicated in the onset of various human pathologies. This project focuses on the role of exoribonucleases in osteosarcoma cells due to a publication pertaining to the regulation of XRN1 in these cells, however, a general overview of diseases associated with defective RNA stability and RNA decay are discussed in the following subsections.

1.7.1 Pathologies generally associated with RNA stability and decay

Many members of the RNA stability and decay machinery are associated with the onset of disease. Of note, these include the development of motor neuron disease (MND) and related diseases, which has been linked with RNA processing, including defective AU-rich element-mediated decay, RNA export out of the nucleus, splicing and translation. AREs are present within many oncogenes, which owe their short half-life to the presence of AREs in the 3' UTR (Bakheet *et al.* 2001). As such, AREs have been implicated in the progression of many cancers, due to defective recruitment of appropriate RNA binding proteins to promote decay primarily by the exosome. This can mean either mislocalisation of RBPs to the incorrect cellular compartment, or defective expression of RBPs. Specifically, over-expression of the stability-promoting factor, HuR, leads to decreased decay of genes with AREs, consequently leading to the progression of cancers of the brain, liver, colon and breast (Khabar 2017). This effect is not limited to cancers, or to AREs, as the role of the RNA binding protein, TAR DNA binding protein (TDP-43) has been explicitly linked to patients displaying sporadic MND (Kolb *et al.* 2010).

The exoribonuclease DIS3 has also been shown to be involved with the onset of certain leukaemias. Specifically, whole genome/exome sequencing has shown DIS3 to have somatic mutations in 11% of multiple myeloma (MM) cases (Chapman *et al.* 2011), whereby mutations in its catalytic site lead to deficiencies in its exoribonuclease activity. Mutations in DIS3 have also been identified in other types of leukaemia, including acute myeloid leukaemia (AML) (Ding *et al.* 2012) and chronic lymphocytic leukaemia (CLL) (Tomecki *et al.* 2014). In addition, recent reports from the Newbury lab have shown that alongside mutations in the DIS3 gene, DIS3 isoforms are differentially expressed between different haematological cancers (Robinson *et al.* 2018). The interesting point here is that the two isoforms are differentiated from each other by the length of their PilT N-terminal domains (PIN). The PIN domain harbours the endonucleolytic activity of the enzyme, and also regulates the tethering of DIS3 to the exosome. Isoform 1 harbours a full length PIN domain, whereas isoform 2 is much shorter, missing a large segment of conserved amino acids. Despite this, isoform 2 shows greater endonuclease activity. This

study shows that isoform 1 is more highly expressed in multiple myeloma than isoform 2, perhaps explaining in part another method by which DIS3 activity is reduced (Robinson *et al.* 2018). In addition, DIS3 has shown itself to be implicated in the progression of solid tumour cancers, such as nodular melanoma, where it displayed significant up regulation (Rose *et al.* 2011).

Mouse models of ribonuclease knockouts are limited to one publication, which shows that loss of Dis3l2 partially phenocopies Perlman syndrome in mice and results in the up regulation of Igf2 in nephron progenitors, with no effect on levels of mature *let-7*. Nephron progenitors are potential cells of origin for Wilms tumours, which implies that loss of Dis3l2 promotes this pathology (Hunter *et al.* 2018).

Alongside the other RNA decay enzymes and RNA stability factors which contribute to different disease pathologies, it is important to mention the role of miRNAs in the development of certain disease conditions. With specific regard to mutations in the exonucleolytic activity of DIS3, RNA metabolism in the context of miRNAs appears to be directly affected. This is shown specifically during the maturation of miRNAs from the *let-7* family. This family of miRNAs have been well-characterised as tumour suppressors, and DIS3 is responsible for the degradation of *let-7* inhibitor, LIN28B. Reduced RNA degradation of LIN28B, caused by reduced activity of DIS3, directly inhibits the maturation of *let-7*. This in turn allows for the translation of *let-7* targets, including oncogenic factors such as MYC and RAS (Segalla *et al.* 2015). Work in *Drosophila* models identified that loss of DIS3 in a Ras overexpression background causes a severe overgrowth phenotype (Snee *et al.* 2016).

miRNAs have been linked to a wide variety of human diseases, including directly affecting many other cancers. For example, loss of *miR-16* has been associated with CLL (Sampath *et al.* 2012) and poor prognosis in prostate cancer patients (Bonci and De Maria 2016), in that depletion of *mir-16* was shown to cause up regulation of anti-apoptotic gene, BCL-2. In addition, *miR-21* has been shown to be over expressed in most solid tumours (Bullock *et al.* 2013). Tumours are also well documented as having mutations in the 3' UTR of oncogenes (such as single nucleotide polymorphisms), changing the sequence of miRNA binding sites, which prevents the suppression of oncogenes, and promotes tumour progression (Bhattacharya *et al.* 2012).

1.7.2 Human diseases associated with XRN1

XRN1 has been implicated in both the onset of osteosarcoma, an adolescent bone sarcoma, and the viral evasion of host responses. This section will aim to describe osteosarcoma, and how XRN1 is utilized by specific viruses for the evasion of the host response.

1.7.2.1 XRN1 has been implicated in osteosarcoma

Osteosarcoma is the third most common cancer affecting adolescents with a second smaller peak occurring between 60 to 80 years, particularly in people affected by Paget's disease of the bone – a disease which results in defective bone renewal and repair, leading to bone fragility and bone misshaping. Overall, osteosarcoma has a moderate incidence rate, with 10 to 26 per million new cases worldwide each year (Ando *et al.* 2013). Osteosarcoma often develops at either end of the long bones of the legs or arms, although tumours can also occur in the skull and pelvis. Tumours typically arise from osteoid-producing neoplastic cells adjacent to the long bone growth plate. The main cause of death in these patients is the spread of cancer cells to the lungs; these secondary tumours are present in 10-20% of cases at diagnosis (Ando *et al.* 2013).

In human osteosarcoma cells, XRN1 has been shown to be often down regulated (Zhang *et al.* 2002), raising the possibility that lower levels of XRN1 are involved in osteosarcoma development. However, null mutations and hypomorphic mutations of this conserved exoribonuclease in *Drosophila* result in apoptosis and compensatory proliferation which is not sufficient to rescue the organism (Waldron *et al.* 2015). If XRN1 activity is conserved between humans and *Drosophila*, one hypothesis to explain this apparent contradiction is that apoptosis is blocked in osteosarcoma cells, meaning that reduced expression of XRN1 causes an increase in proliferation of the tumour cells without the associated apoptosis. By this means, it may be that reduction of XRN1 confers a growth advantage on the tumour cells. Alternatively, downregulation of XRN1 may lead to increased expression of specific miRNAs and/or protein-coding RNAs which then promote the cancerous phenotype. Possible candidates are *IGF-1*, which is known to induce proliferation of osteoblasts (Govoni 2012, Ochiai *et al.* 2012, Zhang *et al.* 2012) and/or pro-inflammatory RNAs such as *FOS* and *MYC* which are widely known to be involved in cancer progression.

1.7.2.2 XRN1 and viral evasion of the host response

XRN1 also appears to play a pivotal role during the host response to viral infection and has been implicated in the pathogenesis of a number of viral families. Recent studies have shown

that the activity of XRN1 is specifically inhibited by certain Flaviviruses, which include Dengue fever virus, West Nile virus (Chapman *et al.* 2014), Yellow fever virus (Mukhopadhyay *et al.* 2005), Japanese Encephalitis virus (Moon *et al.* 2015) and the Zika virus. Encephalitis is symptomatic of each of these infections, alongside hemorrhagic fevers. The activity of XRN1 is also repressed in a similar manner by Hepatitis C (HCV) virus which primarily infects the liver, and can lay asymptomatic for many years. During this time, the virus can cause liver scarring, potentially leading to cirrhosis, liver cancer and liver failure (Boonstra *et al.* 2009). HCV, along with Hepatitis B, has also been tenuously linked with pancreatic cancer and leukaemia/lymphoma (Jarrett 2006, Fiorino *et al.* 2013, Fiorino *et al.* 2015). The targeting of XRN1 by these viruses indicates that XRN1 is normally vitally important in protecting the cell from these viral pathogens.

A series of publications have elucidated the molecular mechanisms whereby these viruses inhibit XRN1 as well as shedding light on the disease mechanisms. In the case of Dengue and West Nile virus (Kunjin strain), viral transcripts are able to inhibit the 5'-3' RNA degradation pathway by stalling XRN1 within the 3’ UTR of the viral RNA. The formation of RNA pseudoknots ("slipknots") (Chapman *et al.* 2014), as well as other RNA structures within the 3'UTR, stall the progression of XRN1 by blocking entry of the RNA into the active site (Chapman *et al.* 2014). These mechanisms are also present in the way other viruses stall and repress XRN1, including phlebovirus and arenavirus, which it is thought utilise G-quadruplexes to stall XRN1 (Charley *et al.* 2018), and Beet Necrotic Yellow Vein Virus (a phytovirus) which acts similarly to flaviviruses in stalling XRN1 to promote expression of noncoding RNA (Flobinus *et al.* 2018).

XRN1 degrades the viral RNAs from the 5’ end of the transcript, until it reaches these structures, leaving small flaviviral non-coding RNAs (sfRNAs), which accumulate in the cytoplasm. These small, structured RNAs bind to and sequester XRN1 thus reducing 5'-3' degradation of other cellular RNAs (Moon *et al.* 2012). For HCV, similar highly structured regions within the 5' UTRs of their transcripts stall and repress the activity of XRN1 (Figure 1.19). Using tissue culture cells infected with HCV, it was shown that the disruption of the 5’- 3’ mRNA decay pathway results in stabilisation of short-lived RNAs such as those encoding transcription factors involved in oncogenesis (e.g. MYC, FOS and JUN) and angiogenesis (e.g. VEGFA, HIF1A and CXCL2) (Moon *et al.* 2015). Since HCV infection is associated with the development of hepatocellular carcinoma, these factors might contribute to the development of the diseased state (Moon *et al.* 2012). This study also showed an increase in levels of capped and presumably functional RNAs suggesting that repression of XRN1 leads to shut down of the entire degradation pathway, possibly by sequestration of other decay components (Moon *et al.* 2015). In the case of Flaviviruses, the available evidence suggests that viral infection disproportionately affects

short-lived transcripts encoding cytokines and factors involved in innate immunity leading to the inflammatory symptoms observed (Moon *et al.* 2012).

The importance of XRN1 in limiting the pathogenesis of HCV is illustrated by the fact that this virus has evolved two ways to inhibit the action of XRN1. Recent publications show that 5' - 3' degradation of HCV transcripts is stalled by expression of *miR-122* in the host cell, which binds via its seed sequence to the 5'UTR of the viral transcript. The association of *miR-122*/RISC, including AGO2 (Shimakami *et al.* 2012), at two sites in the 5' UTR protects it from degradation by XRN1 (Thibault *et al.* 2015). *miR-122* is a major determinant of viral RNA replication, as it inhibits XRN1 from degrading the viral RNA, allowing accumulation, increased replication, and increased translation of viral RNA. Studies have shown that *miR-122* only needs to be bound at one of the two sites for it to have this inhibitory effect, and binding to both sites maintains a co-operative effect, rather than increasing the effectiveness (Thibault *et al.* 2015). In addition to the *miR-122* binding site, the 5' UTR HCV viral RNA includes four stem loops of which three (loops 2-4) appear to function as a highly structured internal ribosome entry site, allowing translation initiation of the viral RNA (Moon *et al.* 2015). Translation of the HCV core protein has been shown to activate the proto-oncogene, MYC, allowing the progression of hepatocellular carcinomas (Ray *et al.* 1995). While loops 1 and 2 are involved in viral replication, they also inhibit the activity of XRN1 by presenting structures which stall the progress of XRN1 as described above.

Another way in which viruses can inhibit the 5' - 3' RNA degradation pathway is by disruption of P-bodies, affecting the localisation of XRN1. Rotavirus infection is understood to disrupt P-bodies, the sites of concentration of the RNA degradation machinery. Rotavirus causes acute diarrhoea across many developing countries, mainly affecting young children, with its spread being attributable to poor sanitation and hygiene (Moe *et al.* 2009). This virus is a double-stranded, non-enveloped virus which utilises rotavirus protein NSP1 to degrade the deadenylase, PAN3, in the P-bodies, resulting in transcript stabilisation, an effect also seen in Poliovirus (Moon *et al.* 2012). Rotavirus also causes the relocalisation of XRN1 from the P-bodies to the nucleus, thereby inhibiting the degradation of viral RNA. However, the molecular details of this process are currently unknown. Viral proteins can specifically prevent host degradation of six structural transcripts and six non-structural transcripts, which are either translated into protein, or serve as templates for replication of viral progeny (Bhowmick *et al.* 2015). Recent work has also shown how rotavirus is capable of inducing remodeled stress granules alongside P-bodies to promote progeny production (Dhillon and Rao 2018).

Gamma-herpesviruses encode an endonuclease (SOX) which is capable of inducing cellular degradation by XRN1. This in turn leads to host cell transcriptional repression via reduced RNA Pol II recruitment (Abernathy *et al.* 2015). These induced transcriptional changes are a direct response of reduced RNA degradation by XRN1, which is evidence for the existence of a feedback loop between cytoplasmic degradation and nuclear transcription.

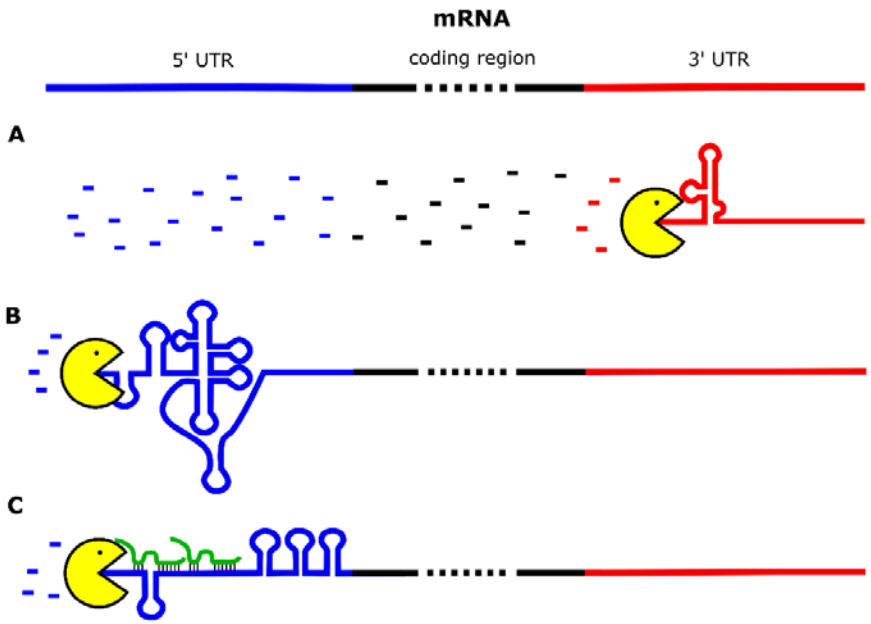


Figure 1.19. Inhibition of XRN1 in the 5′- and 3′-UTRs of viral transcripts. A) XRN1 is stalled at the site of the Flaviviral 3′-UTR due to the presence of pseudoknots or stem loops. It is blocked from continuing degradation, allowing the accumulation of RNA secondary structures that constitute the 3′-UTR of the flaviviral transcript. B) XRN1 is stalled at the 5′-UTR of Hepatitis C viral RNA due to the presence of stem loops and pseudoknots. In the case of A and B, XRN1 is bound to and sequestered by these strong secondary structures. C) XRN1 is stalled at the 5′-UTR due to binding of *hsa-miR-122* (green) at the seed sequence site and the presence of stem loops (1–4), during HCV infection. Figure from (Pashler, Towler *et al.* 2016).

1.7.2.3 Involvement of XRN1 in neurodegenerative disorders

Recent publications have shown that alongside its association with cancer, XRN1 is also implicated in a number of neurodegenerative disorders. In 2018, XRN1 and DIS3L2 were shown by immunocytochemistry to be sequestered into nuclear inclusions in the brains of patients with intranuclear inclusion body diseases (Mori *et al.* 2018). Also sequestered in these intranuclear inclusions were several proteins associated with motor neurone disease (MND). This is interesting because defective RNA processing has been reported to be associated with MND previously (Strong 2010). Although not seemingly to be directly because of XRN1, other aspects of the RNA processing pathway, including splicing events and translation initiation are involved. It could be that defects in these elements of RNA processing could contribute to differential gene expression, in that specific transcripts are not tagged appropriately for degradation, leading to up regulation of gene expression.

1.7.3 Human pathologies associated with DIS3L2

The clinical importance of DIS3L2 is demonstrated by its association with Perlman syndrome, a congenital overgrowth syndrome inherited in an autosomal recessive manner (Schilke *et al.* 2000). Affected children display foetal gigantism, abnormal enlargement of organs (e.g. kidneys), macrocephaly, facial abnormalities, neurodevelopmental delay and high neonatal mortality. Histological examinations reveal nephroblastomatosis, which is an important precursor for Wilms' tumour. Germline mutations in these children include deletions of either exon 6, 9 and/or 19, all of which encode regions of the catalytic domain thereby consistent with a loss of function of DIS3L2 (Astuti *et al.* 2012, Morris *et al.* 2013).

As briefly mentioned above, patients with Perlman syndrome harbour a strong predisposition to developing Wilms' Tumour. Mutations in DIS3L2 have also been associated with sporadic occurrence of Wilms' Tumour, a kidney cancer also referred to as a nephroblastoma (Schilke *et al.* 2000, Astuti *et al.* 2012, Morris *et al.* 2013). Wilms' Tumour results from the failure of groups of kidney cells to mature and differentiate, instead undergoing continuous proliferation. DIS3L2 is likely to be important in sporadic Wilms' Tumour as 30% of these tumours (6/20) show partial or complete DIS3L2 deletion (Astuti *et al.* 2012). This tumour is not only associated with aberrations in DIS3L2, but is also linked to mutations in *DROSHA*, leading to aberrant expression of oncogenic miRNAs (Wegert *et al.* 2015), such as *miR-562* (Drake *et al.* 2009), which promotes deregulation of EYA1. This is interesting because it suggests that this cancer is partially driven by extensive loss of the RNA degradation machinery itself, highlighting the importance of this pathway.

Other aberrations in the DIS3L2 gene have been associated with a Marfan-like syndrome with skeletal overgrowth (Tassano *et al.* 2013). Three patients carrying a chromosomal translocation at the chromosomal region 2q37.1 were shown to have a breakpoint within the DIS3L2 gene at intron 5 (2 patients) and intron 6 (1 patient) (Bocciardi *et al.* 2007, Moncla *et al.* 2007, Tassano *et al.* 2013). The symptoms of these patients included skeletal overgrowth and malformations including long, slender, and curved hands and feet (arachnodactyly), lack of apoptosis between digits (mild syndactyly), and abnormal curvature of the spine (severe scoliosis). However, for these patients, it cannot be ruled out that these phenotypes are also affected by mis-expression of the neighbouring gene NPPC (encoding C-type natriuretic peptide) or deletion of the gene encoding *hsa-miR-562*, which is located within intron 9 of *DIS3L2* (Tassano *et al.* 2013).

All information regarding OMIM numbers for all human congenital diseases and syndromes can be found in Table 1.1.

Disease/Syndrome	OMIM number
Myotonic Dystrophy	602668
Motor Neuron Disease	205100
Multiple Myeloma	254500
Acute Myeloid Leukaemia	601626
Chronic Lymphocytic Leukaemia	151400
Melanoma	155600
Prostate cancer	176807
Osteosarcoma	259500
Paget's Disease	167250
Hepatocellular Carcinoma	144550
Intranuclear Inclusion Body Disease	603472
Perlman Syndrome	267000
Wilms' Tumour	601583
Marfan Syndrome	154700
Ewing Sarcoma	612219
Non-Small Cell Lung Cancer	211980
Pancreatic Cancer	260350
Ovarian Cancer	167000
Colorectal Cancer	608812

Table 1.1. Table of diseases and syndromes discussed, with prospective OMIM numbers. Each disease or disorder is listed in chronological order of its appearance in the text.

1.8 An overview of Osteosarcoma and Ewing sarcoma

Sarcomas are malignancies which arise in cells of the mesenchyme, namely connective tissue, and other non-epithelial tissues, such as muscle and bone. There are three main types of sarcoma: soft tissue sarcoma, bone sarcoma and gastro-intestinal stromal tumours. This project focuses on two types of bone sarcoma: osteosarcoma and Ewing sarcoma. According to the charity, Sarcoma UK, there are 670 people diagnosed with bone sarcoma every year in the UK (https://sarcoma.org.uk/sarcoma-types/bone-sarcoma).

1.8.1 Treatments and pathology of osteosarcoma

This project was initially undertaken to ascertain whether there was a link between the development and progression of osteosarcoma (OS), and the level of XRN1 expression, using cell lines derived from OS patients. OS, also known as osteogenic sarcoma, is a malignancy of the bone which affects children and young adults usually between the ages of 10 and 21. In the United States, 3300 new cases of OS were diagnosed in 2016, and there were 1490 deaths from the disease in the same year. OS accounts for 2% of all paediatric cancers worldwide (Rickel *et al.* 2017). OS is generally a sporadic disease, however, it also occurs at increased rate in patients with Paget's disease of the bone. Symptoms of OS include bone pain, swelling and redness at the site of the tumour and limited movement of the affected joint. These types of tumour become hard calcified masses, as a result of calcium deposition by bone cells. If left undetected or untreated, OS metastasises to the lungs, becoming secondary lung cancer. 5 year survival rates for patients presenting a high grade OS in one location have remained static at between 60-70% since the 1970s, whereas patients who display secondary lung cancer in conjunction with the primary tumour have a survival rate of less than 40% over 5 years (Isakoff *et al.* 2015). OS can occur in any bone, however, most incidences arise at the metaphyses of long bones, such as the distal femur, proximal tibia and proximal humerus. This is also the site of mesenchymal stem cell differentiation into various bone cells and adipocytes, which means that often tumours can be fibroblastic or osteoblastic. Conventional OS, which accounts for around 85% of all OS cases is subdivided into these cell types. In individuals afflicted with Paget's disease (usually the elderly cohort of OS patients), tumours are often found in the axial skeleton and the skull. OS tumours are usually classified into histological divisions: conventional, telangiectatic, small cell, high-grade surface, secondary, low-grade central periosteal, and parosteal variants. The molecular basis for each subtype is not well understood (Martin *et al.* 2012).

Current treatment options for OS have not changed dramatically since the introduction of neoadjuvant chemotherapy in the 1970s. Presently, treatments often include a combined course of chemotherapy, radiotherapy and can also include surgery. Usually, a patient will undergo a three month course of chemotherapy, followed by surgery. The chemotherapies used are most commonly high dose methotrexate (inhibitor of dihydrofolate reductase, leading to inhibition of DNA, RNA thymidylates and protein synthesis (Rajagopalan *et al.* 2002)), doxorubicin (anthracycline which results in cleavage of CREB3L1, and blocks topoisomerase 2, preventing proliferation (Patel and Kaufmann 2012)), and cisplatin (binds to the N7 reactive centre on purine residues to cause DNA damage and apoptosis (Dasari and Tchounwou 2014) (Isakoff *et al.* 2015). Most tumours can be safely removed whilst sparing the affected limb, however, in some cases where the tumour is large or its location is not suitable for limb-sparing surgery, the affected limb will be amputated. A tumour necrosis factor report is conducted by pathologists to determine the rate of tumour necrosis in response to chemotherapy, which directs clinicians as to the appropriate course of chemotherapy to continue treatment with.

1.8.2 Molecular pathology of the development of osteosarcoma

Over the last few decades, a focus has been on identifying oncogenic mutations in the case of inherited pre-dispositions to OS. At present, only tumour suppressor gene mutations have been implicated in familial syndromes, including *p53*, *Rb*, *RECQL4*, *BLM* and *WRN* (Rickel *et al.* 2017). Since OS is most often a sporadic disease, recent studies have focused on the non-familial traits of OS, including the mutations which occur randomly and lead to the onset of OS. Whole genome sequencing studies of OS samples have detected a distinct class of somatic mutations (including single base substitutions and indels) and DNA structural variations (including rearrangements and somatic copy number alterations) (Chen *et al.* 2014, Reimann *et al.* 2014). Chen *et al* also identified 122 cancer associated genes with at least one somatic variant break point. The identified genes are not implicated in RNA stability.

During the pre-malignant stage, the constant remodelling and replacement (via proliferation of progenitor osteoblasts) of bone cells during osteoclastogenesis and osteoblastogenesis results in the accumulation of multiple somatic mutations. Some of these cells accumulate enough mutations to eventually drive the development of OS. Indeed, one study has shown that OS has the highest somatic mutation rate of all childhood cancers, at 1.2 mutations per megabase, and is the same as in adult breast cancer (Lawrence *et al.* 2014). This same study identified the PI3K/mTOR pathway as particularly vulnerable in the development of OS, and showed that somatic mutations of *TP53* are also a causative factor of OS (Perry *et al.* 2014). In

another study, 14 genes were identified as the main drivers for the onset of OS, including (but not limited to) *BRCA 2*, *FANCA*, *PTEN* and *TP53*, which are all well-known cancer drivers (Kovac *et al.* 2015).

Alongside the identification of somatic mutations which drive OS, OS progression has also been associated with *FGFR1* overexpression. FGFR1 is a fibroblast growth factor receptor shown by whole exome sequencing to be amplified and highly expressed. It is also the site of gene fusion with *ZNF343* (Baroy *et al.* 2016). In addition, recurrent mutations of the IGF signalling genes have also been identified, this study also showed that there was amplification *IGF1-R* observed in 14% of tumours (Behjati *et al.* 2017). The role of the IGF pathway in OS is discussed in Chapter 5 of this project. In this case, targeting of the PI3K signalling pathway could act as a possible line of therapy, through inhibition of FGFR1.

It is evident within the literature that OS tumours display a high level of heterogeneity, which makes targeted treatment of these tumours particularly difficult. Identifying a role for the RNA decay pathways in the development of OS could be crucial to the development of new drug therapies. New therapies for this disease have not been developed since the 1970s, and survival rates have remained static since then. If this project can identify a role for XRN1 in the progression of OS, it would be a timely discovery.

1.8.3 A brief overview of Ewing sarcoma

Although this project focuses on OS, it was thought that dysregulation of the RNA decay pathways could have a role in multiple types of sarcoma. With this in mind, this project also includes preliminary studies into the expression of XRN1 in Ewing sarcoma, as well as OS.

Like OS, Ewing sarcoma (EWS) is a malignancy of the bone which predominantly affects children and young adults (usually between the ages of 10 – 24). It is the second most common bone tumour, next to OS. Current 5 year survival rates for EWS stand at 70% for localised tumours, and 30% for metastatic tumours (Kridis *et al.* 2017).

EWS tumours are highly malignant, and are characterised by undifferentiated small round cells. Histologically, EWS cells have little cytoplasm around the nucleus, and they do not exhibit extracellular matrix formation (Ozaki 2015), demonstrating the lack of differentiation in these cells. These tumours often express a balanced translocation involving the *EWS* gene on chromosome 22, and a member of the ETS transcription factors (Delattre *et al.* 1992). Specifically, EWS is diagnosed by the chromosomal translocation t(11;22)(q24;q12), which causes the fusion gene *EWS-FLI1* (Turc-Carel *et al.* 1988). This fusion gene is present in 85% of EWS cases, however, there are examples of less common fusion genes occurring in EWS patients,

such as *EWS-ERG* (10% cases) (Sorensen *et al.* 1994), and rarer still, *EWS-ETV1* (Jeon *et al.* 1995) and *EWS-FEV* (Peter *et al.* 1997). These genetic aberrations contribute to the formation of soft tissue tumours in and around the main bone structure. Unlike OS, EWS tumours are commonly found in the chest wall, pelvis, hip and retroperitoneum, as well as in lower and upper limb bones, whereas OS tumours are generally confined to the growth plates of the upper and lower limbs (Andreou *et al.* 2011).

In contrast to OS, EWS tumours can be found external to the main bone, especially in the soft tissue and adipose tissue surround in the main bone shaft, which makes it markedly different to OS pathology, which is confined to growth plates. Tumours which form within the bone or external to it, in other tissues, are indistinguishable from one another EWS does, however, metastasise primarily to the lungs, as in OS, and also the bone marrow and other soft tissues. Interestingly, the younger the age of the patient, the more likely the tumours are to be located on and within the bone, probably because younger patients are developing bone for growth. Soft tissue tumours are the most common form of EWS in older patients (Ozaki 2015).

Current treatments for EWS include combinations of the following six drugs: doxorubicin (see section 1.8.1) , cyclophosphamide (an alkylating agent which results in the addition of alkyl groups to DNA bases, leading to DNA fragmentation (Emadi *et al.* 2009)), vincristine (inhibition of microtubule formation in the mitotic spindle, arresting dividing cells in metaphase (Martino *et al.* 2018)), actinomycin-D (transcription inhibition by binding DNA at the transcription initiation complex, and preventing transcription elongation (Sobell 1985)), ifosfamide (similar to cyclophosphamide, see section 1.8.1), and etoposide (forms a ternary complex with DNA and topoisomerase II to prevent re-ligation of DNA strands, and causing DNA strand breaks (Montecucco *et al.* 2015)) (Ozaki 2015).Local lesions are often surgically resected or treated with radiotherapy.

1.8.4 Existing RNA studies of OS

1.8.4.1 miRNA expression analysis in OS

Whilst studies about the role of XRN1 in the progression of OS are limited to one published paper and this project, there is much published literature about how RNA affects the progression of this disease. Specifically, there are many publications which appertain to dysregulation of miRNAs in OS. It is hard to evaluate the role of miRNAs in the progression of OS, due to the lack of suitable normaliser genes for these RNAs (this is more of an issue in studies conducted in patient samples, such as serum), and so evaluation of the literature must be circumspect. Even

so, miRNA expression studies over the last few years have gained traction, with the identification of miRNAs which may be having pathological influences of the progression of OS.

There have been at least 24 miRNAs identified as down regulated in OS, with a further 11 identified as up regulated, including the notorious oncomiR, *miR-21* (Palmini *et al.* 2017). This miRNA in particular has been associated with increased cell proliferation and migration, and inhibition of apoptosis in OS cell lines (Vanas *et al.* 2016). Some of these dysregulated miRNAs have been collectively associated with the chemosensitivity and chemoresistance of OS tumours: *miR-513a-5p*, *miR-224*, *miR-21* and *miR-138* (Geng *et al.* 2016, Zhu *et al.* 2016, Dai *et al.* 2018). It has also been shown that *miR-140-5p* and *miR-184* are involved in chemoresistance, alongside the aforementioned miRNAs. One study identified that OS cells treated with doxorubicin and Cisplatin up regulated *miR-140-5p* in order to initiate the autophagy pathway, a cellular survival response to stress (Wei *et al.* 2016). Another study has identified that *miR-184* is increased upon treatment with doxorubicin, which inhibits the pro-apoptotic *BCL2L1*, leading to a reduction in apoptosis (Lin *et al.* 2016). It is clear from the expanse of the literature that miRNAs are heavily involved with the progression of OS, however, the true extent of the influence of miRNAs in OS remains unknown. This is due, in part, to the lack of standardised normalisation, but it should also be mentioned that observations of miRNA expression (excluding the miRNAs mentioned above) are usually not associated with a specific mechanism, and have been found to be observational only.

1.8.4.2 lncRNA expression analysis in OS

Like miRNAs, lncRNAs have also been extensively studied in OS. Again, reliable evidence for lncRNAs which specifically act to promote progression of OS is relatively rare, given that all cancers display dysregulation of at least some lncRNAs, including the infamous example of MALAT-1, named for its activity to promote metastasis of non-small cell lung adenocarcinomas (Gutschner *et al.* 2013). This particular lncRNA has been implicated in many types of cancer since its discovery as an instigator of metastasis in 2013, in a varying degree of tissue types, such as pancreatic cancer, ovarian cancer and colorectal cancer (Wu *et al.* 2018, Zhang *et al.* 2018, Zhuo *et al.* 2018). Of note, in the case of colorectal cancer, it was shown that MALAT-1 inhibits the expression of DCP1A, part of the decapping complex in 5' – 3' RNA degradation (Wu *et al.* 2018).

Nevertheless, the field of lncRNAs in OS has recently been flooded with publications about individual lncRNAs, which are suitably reviewed elsewhere (Chen *et al.* 2017). This section, therefore, will focus on two particular lncRNAs, one which has previously been implicated in the progression of EWS, with a recent discovery about a potential role in the progression of OS elucidated as well, and one which has been repeatedly linked to OS, named MEG3 (a potential

tumour suppressor). The first lncRNA is named EWSAT1 (Ewing sarcoma associated transcript 1), after its implication in EWS. In a recent study, gain-of-function and loss-of-function experiments revealed that EWSAT1 enhanced OS cell proliferation, migration and invasion (Sun *et al.* 2016). This study also showed mechanistically that this lncRNA was able to regulate the transcription of MEG3, which supports the notion for its role in the progression of OS, given that MEG3 repression has been previously associated with the onset of OS metastasis. It also suggests the EWSAT1-MEG3 expression axis may be a viable drug target for OS treatment.

In contrast to the oncogenic effects of EWSAT1, MEG3 has been identified as a potential tumour suppressor lncRNA in OS. MEG3 was found to be significantly down regulated in two OS cell lines when compared to a foetal osteoblast control. Over expression using the lentivirus transfection method lead to inhibition of a number of core processes, including cell proliferation and migration, through subsequent suppression of proteins such as NOTCH1, TGF-β, and N-cadherin (Zhang *et al.* 2017). Another study has suggested that MEG3 could be used as a prognostic marker of OS, by identifying that patient outcomes of those with lower levels MEG3 are seemingly worse than for those with higher levels of MEG3 (Tian *et al.* 2015). This study predicts that inhibition and lower levels of MEG3 are associated with metastasis.

Overall, there have been many lncRNAs implicated in the progression of OS, as summarised in the review by Chen *et al*. including some which have been identified as common markers of multiple types of cancer, for example, MALAT-1. Other lncRNAs associated with the onset of OS include (but are not limited to) TUG1, HOTTIP, HOTAIR, TUSC7 and SNHG12 (Chen *et al.* 2017). This shows that there is potential for the development of novel drugs, which specifically target lncRNAs, in the treatment of OS.

1.9 Preliminary work and previous publications

In 2002, a paper was published establishing a possible link between XRN1 expression and the progression of OS. Zhang *et al* tested the expression of XRN1 mRNA in 5 cell lines (HOb, HOS, U-2 OS, SAOS-2, and MG-63) using semi- quantitative RT-PCR and 9 patient biopsy samples using RT-PCR. Their results showed reduced expression of XRN1 in both OS cell lines and patient samples (Zhang *et al.* 2002), where expression in the HOS cell line was 20% of the control cell line, foetal osteoblasts, 44.60% in the U-2 OS cell line, 39.00% in the MG-63 cell line. The final cell line, SAOS-2 showed 23% higher expression of *XRN1*. This work was the founding basis of this project; the main aim of which has been to find out what effect lower XRN1 expression has on OS cells, and if it may be influencing the development of OS. Since 2002, there has been no

published work about how the RNA decay pathway may be defective during OS pathogenesis, and so research into this topic is timely.

A development in the understanding of the role of XRN1 in OS is not limited to this one type of cancer. This project will also seek to elucidate whether this defect is common among sarcomas, particularly those which stem from the mesenchymal stem cell differentiation pathway, such as Ewing Sarcoma (EWS). If defects in XRN1 are also present in EWS, this could indicate that these types of sarcoma occur due to defects in the mesenchymal stem cell differentiation pathway, meaning that these sarcomas could be developmental disorders. Not only would this reveal potential methods for better treatments, but it would also allude to a new way of better identifying the cause of childhood and adolescent cancers.

1.10 Project aims and brief results chapter overviews

This section will aim to set out a plan for the following project, including a brief description of the following results chapters. Initially, this project was designed to elucidate a role for XRN1 in the progression of OS. Over time, it became apparent that there was scope to investigate other members of the RNA decay machinery (specifically the other exoribonucleases) and how they might be influencing the progression of OS in OS cell lines as well as patient samples.

This project aims to elucidate the cellular effects of XRN1 in an OS model setting, as well as using bioinformatics to pinpoint specific RNA targets of XRN1 in the cell. It is hoped that this information will reveal possible pathway interactions, both in the progression of cancer and in a normal cellular environment that XRN1 is involved with, in conjunction to its role as a major 5’ – 3’ RNA decay enzyme.

There is a vast amount of literature appertaining to the role of XRN1 (and other exoribonucleases) in human cells, including in various disease states, however, there is not much known about the links between XRN1 and cancer. This project, therefore, slides neatly into a gap in present knowledge about the behaviour and expression of XRN1 in cancer cell lines, with specific regard to osteosarcoma, and introduces the concept that there may be co-ordinate regulation between XRN1 and other exoribonucleases, using these cells as a model.

1.10.1 Chapter 3: Characterisation of exoribonuclease expression in OS and EWS

The objectives of this first results chapter were to characterise the expression of XRN1 in OS cells and EWS cells, to identify how its expression is differential between cancer cells and control foetal osteoblasts. Alongside the expression of XRN1 in these cells, another objective was to characterise the expression of other RNA degradation enzymes in OS cells. This was used to

indicate whether there is a general role for the RNA decay pathway in the progression of cancer as a whole.

With this in mind, the main aims of this chapter were:

- To perform qRT-PCR, Western blotting and immunocytochemistry experiments across OS cells lines to gauge a multi-dimensional expression landscape for XRN1 in these cells and in patient samples.
- To perform qRT-PCR and Western blotting experiments across two EWS cell lines to find out whether XRN1 expression is different between two types of bone sarcoma, since patterns in XRN1 expression between these two types of cancer indicates high clinical importance.
- To perform qRT-PCR and Western blotting experiments to observe the expression of other RNA degradation enzymes in OS cells.

1.10.2 Chapter 4: The Effects of Changing Exoribonuclease Expression in OS Cells using siRNA

The primary objective of this chapter was to elucidate a specific cellular effect for both XRN1 and other ribonucleases after XRN1 knock down in the OS cells. Phenotypic studies were performed to identify the implication of lower XRN1 expression in the progression of cancer. It was thought that by knocking down XRN1 in SAOS-2 cells, it would be possible to enhance the cancer phenotype as it is seen in the more proliferative cell lines. Another objective for this chapter was to identify the potential existence of co-ordinate regulation between the ribonucleases in these cells.

The main aims of this chapter were:

- To optimize the knock down of XRN1 and DIS3L2 in OS cells using a lipid-based transfection system.
- To observe the phenotypic effects of XRN1 knock down in OS cells using assays specific for measuring cell line proliferation (and DNA synthesis), rate of apoptosis, cell viability (and metabolism), and translation.
- To perform dual knock down experiments, with qRT-PCR and Western blotting in order to elucidate the presence of co-ordinate regulation of the ribonucleases.
- To knock down DIS3L2 in OS cells and HEK-293T cells to observe whether DIS3L2 is acting with tissue specificity in human cells.

1.10.3 Chapter 5: Identifying Potential Transcript Targets of XRN1 using RNA Sequencing

The final results chapter moves away from elucidating a role for XRN1 in the progression of cancer towards using the OS cells as a model to try to find out the specific transcript targets of XRN1. From this, the objectives for this chapter were to ascertain the specific transcripts being targeted by XRN1, the cellular pathways XRN1 functions within, and how dysregulation of XRN1 could lead to disease states. It is currently unknown how the machinery of the Central Dogma decides which transcripts are to be translated, and which are to be degraded.

The main aims of this chapter were:

- To perform RNA sequencing on cells which have had XRN1 artificially knocked down, and compare the results to RNA sequencing results from control cells, where XRN1 had not been knocked down, about differential transcript expression.
- To study the gene ontology of the differentially expressed transcripts in order to draw conclusions about the cellular pathways XRN1 may be functioning within.
- To draw comparisons about the differentially regulated genes in terms of the presence of common RNA motifs and other transcript characteristics, such as the lengths of the 5’ UTRs and 3’ UTRs.
- To compare the RNA sequencing dataset with that from another publication where RNA sequencing was performed on a different cell line, to identify whether XRN1 targets the same genes across different tissues.

Chapter 2

Materials & Methods

2.1 Cell lines, patient samples and cell culture:

In total, eight human cell lines were employed for the undertaking of this work. For work relating to osteosarcoma: HOS (also known as TE85), SAOS-2, U-2 OS, MG-63 (these were discarded at a later date due to phenotypic abnormalities) and a non-cancerous control human foetal osteoblasts (HOb) were used. These cell lines were all obtained from the European Collection of Cell Cultures (ECACC), and each were regularly tested for mycoplasma infection using 5 mycoplasma specific primers and PCR. For work relating to Ewing Sarcoma: SK-ES-1 and RD-ES were employed. These were a kind gift from Professor. S. Burchill (University of Leeds). These cells were also regularly tested for mycoplasma using the same method. For work relating to the function of DIS3L2, HEK-293T cells were used, a kind gift from Dr. H. Stewart (Brighton & Sussex Medical School). These cells were also regularly tested for mycoplasma in the same way. All cell line information is available in Table 2.1.

Two canine osteosarcoma cell lines were studied alongside the human cell lines: D-17 and OSCA-8. D-17 was purchased from the American Type Culture Collection (ATCC) and OSCA-8 was purchased from Kerafast, Inc. Each was tested for mycoplasma in the same way as the human cell lines. 2 primary control osteoblast cell lines were used in this work: 11.11.16 Labradoodle and 18.11.16 Labrador/Retriever. These were a kind gift from Inês Perpétuo (Royal Veterinary College).

A total of twelve patient biopsy samples were used in this study (Table 2.2). All samples were primary osteosarcoma biopsy tissue which were routinely dissected from the patient for diagnostic purposes, and donated to research after informed consent of the patients. We obtained the samples from the Children's Cancer and Leukaemia Group (CCLG) in compliance with a Material Transfer Agreement (Brighton & Sussex Medical School HTA licence number: 12561), and in accordance with the ethics which this group uses.

Cell line	Bought from/ gifted by	Species	Cell type	Tissue	Properties	Type of cancer	Derived from
hFOB 1.19 (HOb)	ECACC	*Homo sapiens*	Osteoblasts	Bone	Adherent	N/A	Foetus -female
U-2 OS	ECACC	*Homo sapiens*	Epithelial	Bone	Adherent	Osteosarcoma	15 year old, Caucasian female
SAOS-2	ECACC	*Homo sapiens*	Epithelial	Bone	Adherent	Osteosarcoma	11 year old, Caucasian female
HOS	ECACC	*Homo sapiens*	Fibroblast/ epithelial-like	Bone	Adherent	Osteosarcoma	13 year old, Caucasian female
MG-63	ECACC	*Homo sapiens*	Fibroblast	Bone	Adherent	Osteosarcoma	14 year old, Caucasian male
SK-ES-1	Prof. S. Burchill, Univ. of Leeds	*Homo sapiens*	Epithelial	Bone	Adherent	Ewing Sarcoma	18 year old, Caucasian male
RD-ES	Prof. S. Burchill, Univ. of Leeds	*Homo sapiens*	Epithelial	Bone	Adherent	Ewing Sarcoma	19 year old, Caucasian male
HEK-293T	Dr. H. Stewart, Brighton & Sussex Medical School	*Homo sapiens*	Epithelial	Embryonic Adrenal glands (Kidney)	Adherent	N/A	Foetus - female

Table 2.1. Table to describe the cell lines used in this project. Cell lines used throughout the project, including their unique characteristics. All cells underwent rigorous mycoplasma testing throughout the project.

Sample ID	Sex (if disclosed)	Location of biopsy resection	Histological features
10/196	Unknown	Right scapula	Pleiomorphic cells, large cytoplasm and prominent nucleoli. Cells are irregular and osteoid formation is evident. Strong alkaline phosphatase staining – high levels of calcification.
10/10034	Female	Left femur	Dense population of nuclear pleiomorphic cells within bony trabecular. Osteoid deposition by abnormal cells, scattering of normal osteoblasts and osteoclasts within.
11/650 A, B & C	Unknown	Lung metastasis	All specimens show lung with nodules of tumour made up of spindle cells, marked nuclear pleiomorphism and mitoses. Marked osteoid matrix formation.
12/299	Unknown	Pelvis	Variable cellularity, with nests of cells displaying size and cytoplasm variability. Some pleiomorphic, multinucleated, giant cells. No evidence of osteoid. Described as chondroblastic OS.
16/496	Unknown	Right, proximal humerus	Severely pleiomorphic, hyperchromatic nuclei. Formation of malignant osteoid. Lytic lesion.
16/509	Unknown	Right distal femur	Some cells have intermediate-sized, round nuclei, others are pleiomorphic. Scattered giant cells also seen. Formation of osteoid.
16/528	Unknown	Right distal femur	Pleiomorphic cells with large amounts of cytoplasm. Large amount of osteoid production.
16/572	Unknown	Left distal femur	Trabecular bone infiltrated by bone-forming tumour. Viable tumour cells have plump pleiomorphic nuclei.
16/632	Unknown	Right proximal humerus	Fragments of bone-forming tumour. Small foci of cartilaginous matrix also identified. Cells have plump nuclei with mild pleiomorphism.
16/755	Unknown	Right femoral neck	Malignant neoplasm with sheets of pleiomorphic, ovoid and fusiform cells. Evidence of osteoid formation.
16/760	Unknown	Left proximal humerus	Trabecular bone infiltrated by tumour. Formation of malignant osteoid which is partly coarse, partly delicate, produced by plump pleiomorphic cells.

Table 2.2. Patient OS biopsy sample description. Samples 11/650, 12/299 and 16/755 displayed large necrosis of the sample, and so were not included in analysis. Details of sample 16/591 were not disclosed. Samples were released by the Children's Cancer and Leukaemia Group (CCLG).

2.1.1 Cell Culture

Osteosarcoma cell lines and the control cell line were cultured in Dulbecco's Modified Eagle's Medium/F12 (DMEM/F12 – Gibco, product code 21331-020) supplemented with 10% foetal bovine serum (FBS – PAN Biotech, product code P40-37100), 2mM L-Glutamine (Gibco, product code 25030-024) and antibiotics (100IU/mL penicillin, 100µg/mL streptomycin, Gibco, product code 15140-122), at 37^{o}C in a 5% CO_2 humidified incubator.

Ewing Sarcoma cell lines were cultured in McCoy's 5A (Modified) medium (Gibco, product code 26600-080) (SK-ES-1) and RPMI 1640 medium (Gibco, product code 12633-020) (RD-ES) respectively, each supplemented with 10% FBS, 2mM L-Glutamine and antibiotics (same concentration as osteosarcoma cell lines) and incubated in 5% CO_2 humidified incubator.

Canine osteosarcoma cell lines and primary osteoblast cell lines were cultured in Advanced DMEM (Gibco, product code 12491-023) supplemented with 10% FBS, 2mM L-glutamine and antibiotics. They were kept in the same conditions as the human cancer cell lines.

All cell lines were passaged using 0.25% Trypsin-EDTA (Trypsin: Gibco, product code 15090046, EDTA: Sigma Aldrich, product code E5134-500G) and cell viability was tested using 0.1% trypan blue staining (Gibco, product code 15250-061).

2.2 RNA Extraction, PCR, RT-PCR and qRT-PCR

2.2.1 RNA Extraction:

1x10^{6} cells were collected in a pellet and snap frozen in their physiological state before being lysed for RNA extraction by the miRNEasy mini-kit (Qiagen, product code 217004) according to the manufacturer's instructions, also incorporating an on-column DNase step to lessen the contamination by genomic DNA using the RNase-free DNase Set (Qiagen, product code 79254). Total RNA isolated was eluted in 14µL RNase-free water. 1µL of the sample was measured for RNA concentration using an ND-1000 spectrophotometer (Thermo Scientific), and later a Nano Drop One (Thermo Scientific).

2.2.2 RT-PCR:

RNA was diluted to 50ng/µL for reverse transcription PCR (RT-PCR). RT-PCR was performed using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, product code 4368814) according to manufacturer's instructions to a final volume of 10µL (5µL RNA: 5µL Master Mix). The final composition consisted of: 1µL 10X Buffer, 1µL random primers

(hexamers), 0.5µL Reverse Transcriptase, 0.4µL dNTPs, 2.1µL H_2O and 5µL diluted RNA. A 'No Template Control' was also created to control for non-specific probe binding during qRT-PCR. Samples were amplified in the Thermo Cycler following the below conditions for 35 cycles:

1. Annealing at 25˚C for 10 minutes
2. Incubation at 37˚C for 2 hours
3. Inactivation at 85˚C for 5 minutes
4. Hold at 4˚C for indefinite time

The same conditions were used for all RT-PCR experiments performed.

2.2.3 qRT-PCR:

XRN1 gene expression was analysed using the real-time TaqMan qRT-PCR system, which incorporates gene-specific primers and a fluorescent probe labelled with a FAM dye at the 5' end, and a quencher molecule on the 3' end. During the qRT-PCR, the target gene is amplified and a signal is generated when the probe undergoes cleavage by *Taq* polymerase, releasing the fluorophore having been previously quenched during the cDNA denaturation step. The signal from the fluorescent dye increases as the qRT-PCR commences through each cycle, reaching a detectable threshold value called the cycle threshold (C_t). It is the change in this value which is used to determine the level of gene expression as it gives the number of the cycle at which the fluorescence becomes above background, and so, the earlier in the reaction that the fluorescence becomes detectable, the higher the level of gene expression. For measuring the level of XRN1 expression in human samples and cell lines, 50 cycles of the qRT-PCR were run.

For subsequent qRT-PCR analysis of mRNA expression, a 2.5X reaction mix was made. 2.5µL of cDNA was used in conjunction with 12.5µL TaqMan Universal PCR MasterMix No AmpErase UNG (Applied Biosystems, product code 4324018) and 8.75µL RNase-free water. 1.25µL of specific TaqMan Gene Expression assay (Applied Biosystems – Table 2.3) was applied and qRT-PCR was performed in a ViiA 7 TM System for 50 cycles using the following conditions:

1. Activation at 95˚C for 10 minutes
2. Denaturation at 95˚C for 15 seconds
3. Annealing and Extension at 60˚C for 1 minute, fluorescence measured after each extension step.

For each biological replicate, two technical replicates were performed for the reverse transcription step. For the qRT-PCR step, 2 replicates for each of the replicates made during the reverse transcription step were created in order to identify any variation. All data were

referenced to appropriate housekeeper genes, *GAPDH* or *HPRT1*. Patient samples were normalised to *PES1*. Fold change in expression levels were calculated using the 2^ddCt method. Primer probe information can be found in Table 2.3.

2.2.4 PCR:

PCR for normal amplification of cDNA product was performed using 12.5µL AmpliTaq Gold 360 Master Mix (Applied Biosystems, product code 4398876), 10.5µL RNase-free H_2O, 1µL cDNA and 1µL of 5µM Forward and Reverse specific primers (primer details Table 2.4). A 'no Reverse Transcriptase' control was created to control for non-specific binding. Samples were run in a thermal cycler using the following conditions, with steps 2-4 being repeated for 35 cycles:

1. Activation at 95˚C for 10 minutes
2. Denaturation at 95˚C for 30 seconds
3. Annealing at 57˚C for 30 seconds
4. Extension at 72˚C for 90 seconds
5. Final extension at 72˚C for 10 minutes
6. Hold at 4˚C for indefinite time

Products were run on a 1.2% agarose gel (Fisher Bioreagents, product code BP1356-500) with 0.5µL of Gel Red stain (Biotium, product code 41003) incorporated (before the solution was heated). 2 sizes of agarose gel were made during experiments depending on the number of samples at any given time. Larger gels were made by adding 1.8g of agarose to 150mL of 1X TBE buffer (90mM Tris, 90mM Boric acid, 2mM EDTA), heating in a microwave until all agarose dissolved and pouring into gel cassette and left to set. The smaller gel was made by adding 0.6g of agarose to 50mL of 1X TBE, heating until all agarose had dissolved and pouring into smaller gel cassette, and left to set. Results were viewed on a UV Transilluminator (Fisher Scientific) for presence of specific DNA isoforms. Alongside the samples loaded into the gel, a 1kbp and 100bp ladder were loaded respectively onto the gel (New England Biolabs, product codes N3232 and N3231).

Primer/ siRNA	Assay ID	Sequences	Binding location	Application
Pre-XRN1	Lot: 1588077	Forward: ACGGGCCTGTCTT CAGAAAATATG Reverse: CCTGAACATAACA TGCAGCTTGATC	Unknown	TaqMan qRT-PCR
XRN1	Hs00943063	N/A	Exon 10-11 boundary, base 1289. Amplicon length = 71 bases	TaqMan qRT-PCR
DIS3	Hs0020014	N/A	Exon 15-16 boundary, base 2182. Amplicon length = 110 bases	TaqMan qRT-PCR
DIS3L1	Hs00370241	N/A	Exon 2-3 boundary, base 315. Amplicon length = 60	TaqMan qRT-PCR
DIS3L2	Hs04966835	N/A	Exon 3-4 boundary, base 86. Amplicon length = 86	TaqMan qRT-PCR
XRN2	Hs01082225	N/A	Exon 3-4 boundary, base 590. Amplicon length = 170	TaqMan qRT-PCR
NID1	Hs00915875	N/A	Exon 19-20 boundary, base 3729. Amplicon length = 75	TaqMan qRT-PCR
GLO1	Hs00198702	N/A	Exon 3-4 boundary, base 430. Amplicon length = 90	TaqMan qRT-PCR
EMP2	Hs00171315	N/A	Exon boundary 3-4, base 395. Amplicon length = 88	TaqMan qRT-PCR
NUDT15	Hs01087148	N/A	Exon boundary 1-2, base 341. Amplicon length - = 64	TaqMan qRT-PCR
EDIL3	Hs00964112	N/A	Exon boundary 3-4, base 814. Amplicon length = 62	TaqMan qRT-PCR
PPP1R148	Hs00261454	N/A	Within exon 1, base 1208. Amplicon length = 73	TaqMan qRT-PCR
GAPDH	Hs02786624	N/A	Exon 7, base 870. Amplicon length = 157	TaqMan qRT-PCR

Primer/ siRNA	Assay ID	Sequences	Binding location	Application
XRN1 siRNA	125199	N/A	Exon 11, base 1345	XRN1 knock down
DIS3L1 siRNA	141248	N/A	Exon boundary 3-4, base 449	DIS3L1 knock down
DIS3L2 siRNA	126074	Sense: GCAUUGGAAGG UAGUUAAAtt	Exon boundary 5-6, base 641	DIS3L2 knock down
Scrambled siRNA control	AM4611	N/A – no matched sequence to human genome	Unknown– no matched sequence to human genome	Knock down controls

Table 2.3. Gene expression assays and siRNAs used. The sequences were only available for custom-made primer-probes. Sequences inaccessible for off-the-shelf assays due to the primer-probes being commercial products.

Name	Sequence	Application
XRN1-001	Forward: CAACTATGCATTGGGGGCAG Reverse: ATGCAGGTTGAGCAATCGGA	PCR
Not1	Forward: TCGAGAGTCGCGGCGCTTCACTGCA Reverse: GTGAAGCGCCGCGACTC	Plasmid recircularisation

Table 2.4. Oligo and primer pairs for cloning and PCR.

2.3 Western Blotting:

$1x10^6$ cells previously snap frozen were lysed in 60µL of loading buffer (250mM Tris- Sigma Aldrich – product code T1503-1KG, 4% sodium dodecyl sulphate – Fisher Scientific – product code BP166-500, 20% Glycerol – Sigma Aldrich, product code G9012), supplemented with 2% β-mercaptoethanol (Sigma Aldrich, product code M7154-250ML) and 2% protease cocktail inhibitor (Roche, product code 04693124001). Suspensions were heated at 100^oC in a heating block for 7 minutes and centrifuged for 5 mins at full speed. Protein lysates were measured on a NanoDrop 1000. Lysates were resuspended in 20X Bromophenol Blue and 18µL was loaded onto a pre-case NuPAGE 7% Tris-Acetate gel (Thermo Fisher, product code EA03585BOX) with 1X Tris-Acetate SDS Running Buffer (Novex, product code LA0041). Electrophoresis was performed at 150V for 65 minutes. Color Prestained Protein Standard Ladder (11-245kDa) (New England Biolabs, product code P77125) was used to identify the correct sized bands. Proteins were transferred to a PVDF Immobilon-FL Transfer Membrane (Merck Millipore Ltd, product code IPFL20200) using a wet transfer blotting system (Biorad), which was run at 100V for 1 hour.

Transfer buffer was made up to 1L of 25mM Tris, 190mM Glycine (Fisher Scientific, product code G/P460/53), 0.05% SDS, 800mL dH_2O and 20% Methanol (VWR Chemicals, product code 20847-307). Membrane was blocked in 5% milk in PBS-T (Phosphate-buffered saline solution with Tween 20 – Sigma Aldrich, product code P1379-500ML) for 1 hour. Later western blotting utilised Odyssey Blocking Buffer (PBS) (LiCor, product code X3391). The membrane was then incubated overnight at 4^oC on a roller with primary antibody in wash buffer (0.5% milk in PBS-T, later Odyssey Blocking Buffer with 0.1% Tween 20): anti-XRN1 (1:2000, Bethyl Labs) anti-GAPDH (1:10,000, Abcam), anti-Tubulin (1:2000, Sigma Aldrich), anti-PCNA (1:1000, DAKO), anti-ATG7 (1:10,000, Abcam), anti-DCP2 (1:2000, Sigma Aldrich), anti-Puromycin (1:1000, Merck) (Table 2.5).

The membrane was washed 3x 0.1% PBS-T, and incubated at room temperature for 1 hour in fluorescent secondary antibodies (1:20,000, IRDye 680RD Goat anti-Rabbit and IRDye 800CW Goat anti-Mouse, Li-Cor) in Odyssey Blocking Buffer (PBS) (product number 927-40000) with 0.1% Tween 20 and 0.01% SDS (earlier Western blots were blocked in 5% Milk with 0.1% PBS Tween). Membrane was washed 3x in 0.1% PBS-T and 2x in PBS before being developed in an Odyssey Fc Imaging System (Li-Cor). Protein expression was normalised to GAPDH expression and quantified using Image Studio Lite software and Microsoft Excel.

Antibody	Dilution used at for Western blot	Dilution used at for immuno cytochemistry	Host species	Bought from	Product number
Anti-ATG7	1:10,000	N/A	Rabbit	Abcam	ab53255
Anti-BrdU	N/A	1:10	Mouse	Developmental Hybridoma bank	G3G4
Anti-Dcp2	1:2000	1:400	Mouse	Sigma Aldrich	SAB1408413
Anti-DIS3*	1:600	N/A	Mouse	Sigma Aldrich	SAB4200474
Anti-DIS3L1	1:500	N/A	Mouse	Abcam	ab89042
Anti-DIS3L2	1:500	N/A	Rabbit	Novus Biologicals	NBP1-84740
Anti-G3BP1	1:500	1:500	Mouse	Abcam	ab56574
Anti-GAPDH	1:10,000	N/A	Mouse	Abcam	ab8245
Anti-GW182	N/A	1:100	Mouse	Abcam	ab70522
Anti-PCNA	1:1000	N/A	Mouse	DAKO	M0879
Anti-puromycin	1:1000	N/A	Mouse	Merck	MABE343
Anti-tubulin	1:2000	1:800	Mouse	Sigma Aldrich	T6074
Anti-XRN1	1:2000	1:400	Rabbit	Bethyl Labs	A300-443A

Table 2. 5. List of primary antibodies used in Western blotting and immunocytochemistry. *Second antibody for DIS3 was also used: anti-DIS3, mouse, used at 1:500, bought from Thermo Fisher, product number: PA5-78427.

2.4 Cell Cycle Arrest:

Cells were plated at a density of $1x10^6$ in T-25 NUNC plastic flasks in 6mL culture medium. Cells were arrested in G_o phase by incubation for 3 days in nutrient deprivation medium (DMEM/F12 medium + 1% FBS) and corresponding flasks were arrested in M-phase using 1µg/mL colchicine in 6mL normal culture medium. Flasks incubated for 1.5hrs following colchicine addition. Cells were then harvested for Western blotting (see above) and snap frozen. GAPDH was used for normalisation.

2.5 XRN1 knock down transfections and subsequent assays:

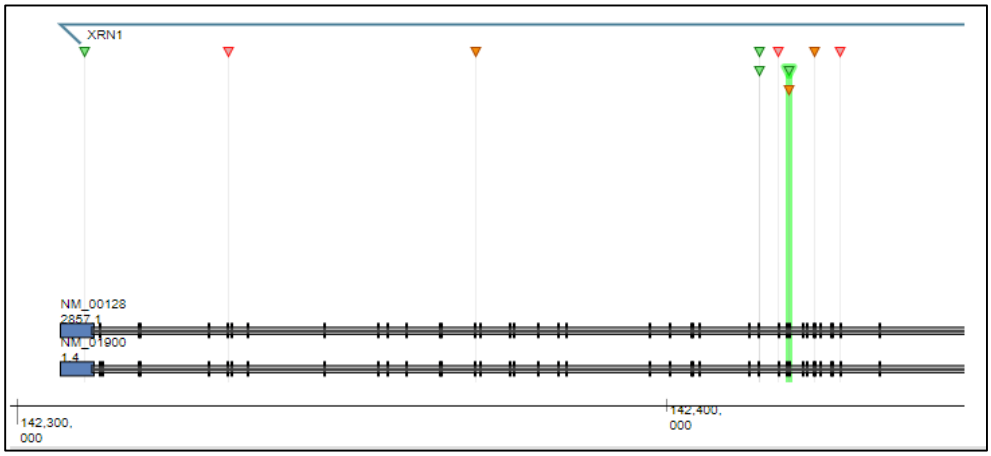
2.5.1 Transfection of siRNAs

Initial transfections took place in a 6-well plate with a cell density of $3x10^5$ per/well in 2ml culture media. After adherence, medium was replaced with antibiotic-free media for transfection. Lipofectamine RNAiMAX reagent (Thermo Fisher, product code 13778075) was used as a vector to transport the pre-designed siRNAs (Table 2.3) for transfection, mixed together with OptiMEM Reduced Serum Medium (Gibco, product code 31985070) in the following way: 9µL Lipofectamine RNAiMAX was added to 150µL OptiMEM Reduced Serum Medium. In a separate tube, siRNA was added to 150µL of OptiMEM Transfection Medium. The Lipofectamine- containing medium was mixed with the siRNA-containing transfection medium, and incubated for 5 minutes at room temperature. The resulting complex was then added (together with transfection medium) in a dropwise fashion to the cells. Cells were incubated at 37^oC in 5% CO_2 for 24hrs before medium was replaced with fresh, antibiotic-free full medium. 20pmol of XRN1 Silencer Pre-Designed siRNA (Fisher, product code AM16708) and XRN1 Silencer Negative Control #1 (scrambled sequence control) (Fisher, product code AM4611) were used in initial investigations, with parameters changing according to the size of the wells used per experiment. Knockdown was confirmed using qRT-PCR and Western blotting (see above). DIS3L1 knockdowns were performed in the same way. DIS3L2 knockdowns in HEK-293T and U-2 OS cells were performed using 30pmols of DIS3L2 Silencer Pre-Designed siRNA (Fisher, product code AM16708) and DIS3L2 Silencer Negative Control #1 (Fisher, product code AM4611) according to the results from concentration optimisation experiments. Binding sites for each siRNA are shown in Figure 2.1.

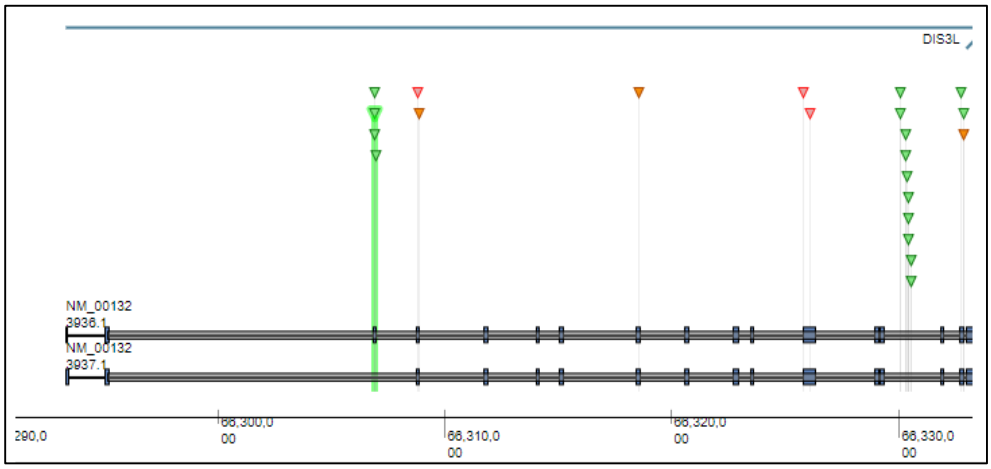
For growth curves, cells were plated at a density of $1.5x10^5$ or $3x10^5$ (depending on the cell line) in a 6- well plate and counted in triplicate post-transfection every 24 hrs for 144hrs

(excluding 120hrs), their viability was tested with 0.1% trypan-blue stain (in 1X PBS) on a Neubauer Chamber.

XRN1 siRNA



DIS3L1 siRNA



DIS3L2 siRNA

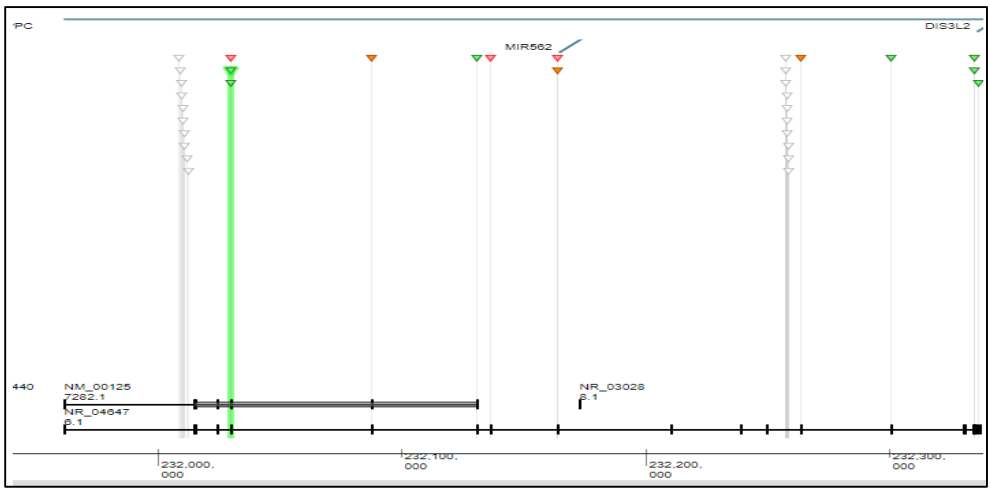


Figure 2.1. siRNA binding sites. A visual representation of siRNA binding sites during knock down experiments. All genes are transcribed from left to right. Exons are denoted by small vertical lines and the introns are denoted by the linking horizontal lines between the exons. Splice variants are also given. The red and green arrows show the variety of binding sites of the siRNAs offered by Thermo Fisher, the highlighted green arrows show the binding site of the particular siRNA used in this project.

2.5.2 Cell Viability

To assess cell viability, Cell Proliferation Reagent WST-1 (4-[3-(4-Iodophenyl)-2-(4-nitrophenyl)-2H-5-tetrazolio]-1,3-benzene disulfonate, Sigma Aldrich, product code lot no. 18993700) assay was employed (Figure 2.2). Cells were plated in duplicate at a density of $2x10^4$ in 96-well plates with 90µL medium. At 24 hour intervals following transfection (with 5pmol siRNA) up to 144hrs, 10µL of WST-1 was added to respective wells. Absorbance was recorded after 1.75hrs incubation with WST-1 and 10 seconds of shaking. Absorbance was measured at 450nm using a plate-reader (Bio TEK) and KC4 software.

2.5.3 Cell proliferation

To assess cell proliferation, a BrdU (Bromodeoxyuridine, Sigma Aldrich, product code B5002-100MG) DNA synthesis assay was employed. Cells were plated at a density of $5x10^4$ on coverslips in 24-well plates with 400µL full medium. BrdU was incorporated at a concentration of 10µM for 6 hours (24hrs after initial transfection with 10pmol siRNA) before cells were washed in PBS and fixed using 4% paraformaldehyde (PFA, Sigma Aldrich, product code P6148). Cells were permeabilised in 0.3% PBS-Triton X-100 (Sigma Aldrich, product code 93443) for 45 minutes before incubation with 4M HCl for 30 minutes. Cells were then incubated in 0.1M sodium borate (Sigma Aldrich, product code 221732) for 10 minutes and washed 3x with 0.3% PBS-Triton X-100. Cells incubated overnight at 4^oC in Mouse anti-BrdU antibody (1:20, Developmental Studies Hybridoma Bank) with 0.3% PBS-Triton X-100, supplemented with 10% FBS. Cells were washed in 0.3% PBS-Triton X-100 and then incubated with Cy3 conjugated Donkey anti-Mouse IgG (1:350, Jackson Immuno Research) for 2 hours at room temperature. Cells were washed again before being mounted onto Apex Superior Adhesive Slides (Leica Biosystems, product code 3800075) using VECTASHIELD with DAPI (4', 6-diamidino-2-phenylindole) mounting medium (Vector Laboratories, Inc. product code H-1200). Slides were sealed with clear nail polish. Slides analysed using confocal microscopy (Leica SP5 laser scanning microscope), which utilised a spectral head and Acousto-Optical Beam Splitter (AOBS). Results were analysed using ImageJ software, whereby the Dead_Easy Mitoglia Plug-In was employed to perform unbiased mitotic cell counts. Detection threshold was set to 60 for BrdU staining and 20 for DAPI staining. Cell volume threshold was set to 50. Quantification was a comparison of the percentage BrdU-positive cells against total number of cells (DAPI positive regions) between controls and XRN1 knockdown cells.

2.5.4 Apoptosis

Apoptosis was measured using a Caspase-Glo 3/7 assay (Promega, lot #0000239042) (Figure 2.3). $2x10^4$ cells were plated in triplicate wells, one triplicate per 24hr timepoint. Cells were incubated overnight and transfected with 5pmol siRNA for 24hrs. Transfection reagent was removed and replaced with antibiotic-free medium. In the wells to be tested, 25µL of medium was mixed with 25µL of Caspase-Glo 3/7 assay, shaken for 30 seconds, and incubated for 1.5hrs at 37^{o}C. Luminescence was measured using a plate reader (Bio TEK) and KC4 software.

2.5.5 SUnSET Labelling:

The rate of translation was measured using the incorporation of the antibiotic, puromycin (Merck, product code 540411), into nascent polypeptide chains. This was then detected using an antibody specific to puromycin (anti-puromycin, 1:1000, Merck, product code MABE343). Cells were plated at a density of $4x10^5$ cells/well in a 6-well plate with 2mL of fully supplemented culture medium. Cells were incubated overnight at 37^{o}C in 5% CO_2. Cells were then transfected using 20pmols XRN1 siRNA alongside the scrambled siRNA control (20pmols) and a Lipofectamine-only control. Cells were then further incubated for 24hrs in antibiotic-free medium (2mLs), and the medium was changed and cells washed in PBS. A further 24hr incubation was performed. 2.5µg/mL of puromycin was added to the wells respectively and the wells were incubated for 1 hour at 37^{o}C. Following this, cells were harvested ready for Western blotting.

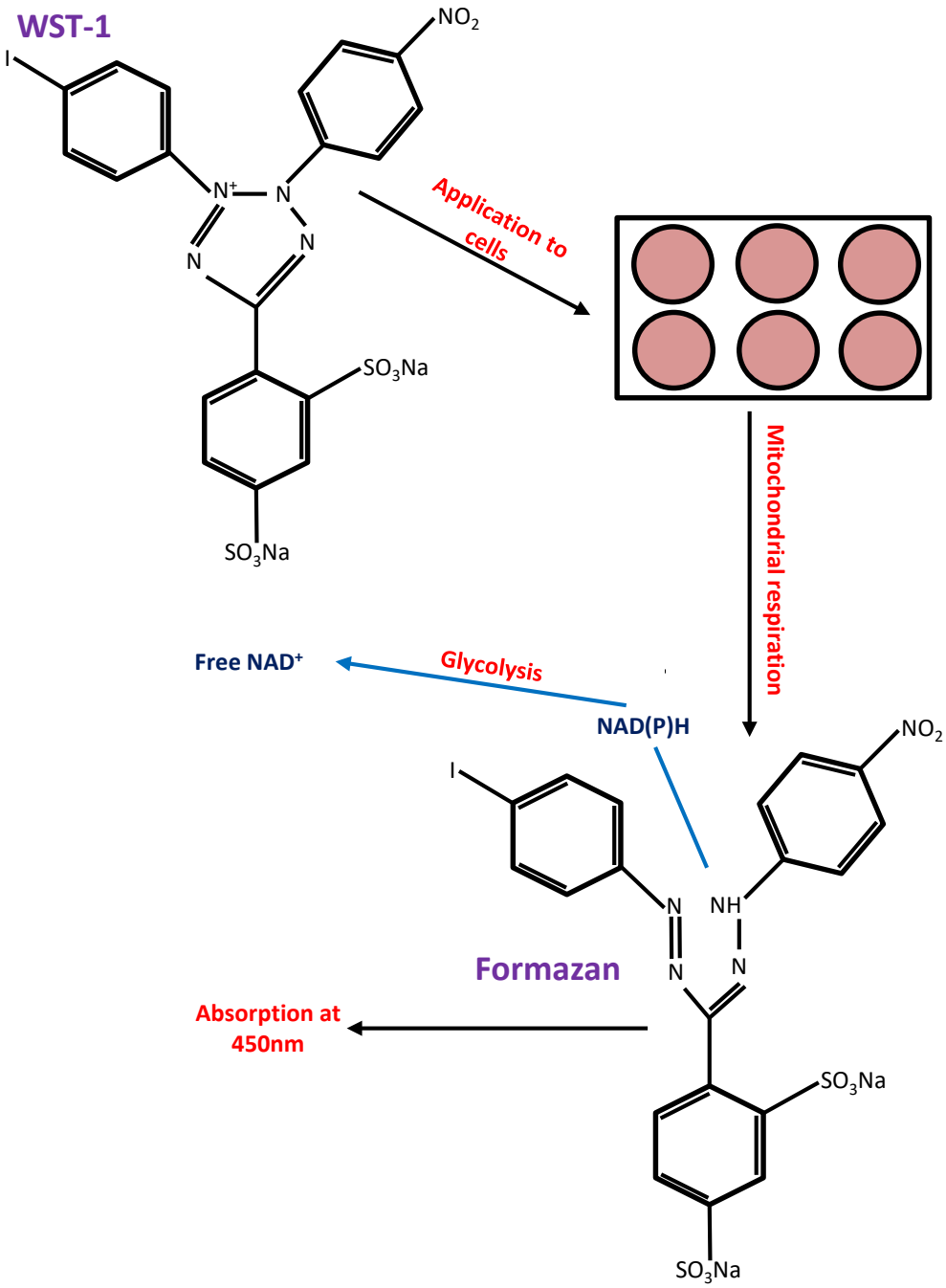


Figure 2.2. WST-1 cell viability assay mechanism of action. WST-1 is applied to cells, whereby it is cleaved during mitochondrial respiration. This can only occur through the utility of glycolytic production of NAD(P)H, and therefore only happens in viable cells, and is represented by the blue arrow in this schematic. The absorbance is measured at 450nm (orange dye), giving a measure for viable cell activity in the well (the higher the absorbance, the higher the level of viable cells).

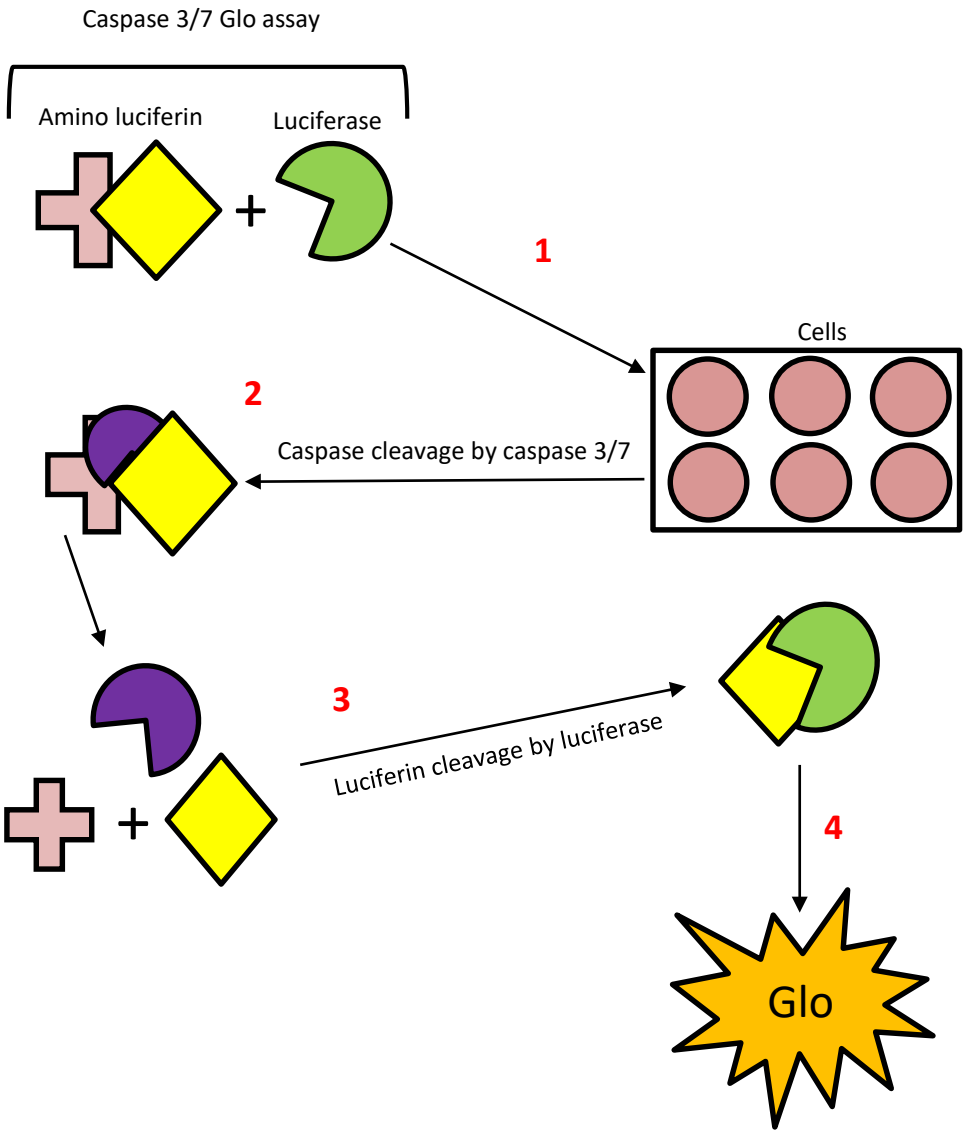


Figure 2.3. Mechanism of Caspase 3/7 Glo assay activity. 25µL of Caspase 3/7 Glo assay is applied to the wells, a dilution factor of 2 with cell culture medium (1). The plate is then incubated at 37°C for 1.5 hours. During this time, the aminoluciferin component of the assay is cleaved by caspase 3/7, releasing the luciferin (2). The luciferin in turn is catalyzed by the enzymatic component of the assay, luciferase (3). This reaction causes electron excitation and this is emitted as the 'Glo' component (4). The emitted light is detected on the plate reader, and is directly proportional to the amount of caspase activity, therefore giving an indication as to the level of apoptosis occurring.

2.6 Immunohistochemistry:

Bone samples used for the optimisation of immunohistochemistry were rat femur, kindly donated by Dr. A. Dilley (Brighton & Sussex Medical School) for this work, in compliance with the UK Animals (Scientific Procedures) Act (1986). Femurs were initially cut using bone-specific scissors, first laterally to remove the femoral head from the rest of the femur. The femoral head was then cut longitudinally to expose the centre of the head without destroying the structure. The sections were then fixed in 4% PFA for 24hrs. Sections were decalcified in 10% formic acid (Dr. L. Dixon, Royal Sussex County Hospital) until femoral heads were flexible. Samples were wax-embedded for 20hrs in a tissue processor (Leica TP1050) using the following parameters:

Reagent	**Time Hr:Min**
Formalin	01:00
IMS 50%	00:30
IMS 70%	01:00
IMS 90%	03:00
IMS 100%	01:00
IMS 100%	01:00
IMS 100%	01:00
Xylene	00:45
Xylene	02:00
Xylene	02:00
Histowax	01:00
Histowax	03:00
Histowax	03:00

Wax–embedded samples were then transferred to wax blocks using a wax station (Tissuetek III). Sections were cut using a microtome (Leica RM2135) and MX35 Premier Microtome Blades (Thermo Scientific) and mounted onto Apex Superior Adhesive Slides (Leica,

Biosystems, product code 3800075). Deparaffinisation of the slides was performed using these parameters:

Reagent	**Time (min)**
2 x 100% Xylene (Sigma Aldrich, product code 108-38-3) washes	05:00 each
1 x Xylene 1:1 with 100% Ethanol (Sigma Aldrich, product code 51976) wash	05:00
2 x 100% ethanol washes	05:00
1 x 70% ethanol wash	05:00
1 x 50% ethanol wash	05:00
1 x 20% ethanol wash	05:00
1 x dH_2O wash	01:00

Antigen retrieval was performed by incubating the slides in a humidity chamber overnight at 60^oC in a Memmert oven in Tris-EDTA buffer (10mM Tris, 1mM EDTA), 1L dH_2O, 0.5mL Tween-20) at pH9. Samples were rinsed in dH_2O and blocked in PBTA (0.3% PBS-Tween, 1%FBS) for 2 hours at room temperature. Samples were incubated overnight at 4^oC in primary antibodies: anti-sclerostin (1:50, Abcam, product code ab63097), anti-osteocalcin (1:500, Abcam, product code ab13420), anti-Vimentin (1:150, Abcam, product code 8978) and anti-RANKL (1:100, Abcam, product code ab9957). Samples were incubated in secondary antibody (1:200, FITC-conjugated Goat anti-Mouse IgG, Jackson Immuno Research) for 2 hours and then washed in PBTA. 200µL of VECTASHIELD (with DAPI) mounting medium was applied and coverslip laid on top. Samples would have been analysed using confocal microscopy (Leica SP5) and ImageJ software.

2.7 Immunocytochemistry:

Cells were plated at a density of $5x10^4$ on coverslips in 24-well plates, and were cultured overnight in 400µL full medium. Cells were fixed in 4% paraformaldehyde for 15 minutes at room temperature, washed in PBS and permeabilised in saponin (200µg/mL) (Sigma, product code 47036) for 10 minutes. Cells were then blocked in 2% Fish gelatin (Sigma, product code G7041-

100G) for 20 minutes. Cells were washed in PBS and incubated overnight in primary antibodies for XRN1 (1:400, Bethyl labs) and TUBULIN (1:800, Sigma), an antibody against DCP2 (1:400, Sigma) was used for colocalisation studies (Table 2.5). Antibodies were diluted in 2% Fish gelatin in PBS. The cells were washed with PBS and incubated for one hour with secondary antibodies in 2% Fish gelatin in PBS: Cy3-conjugated Goat anti-Rabbit IgG (1:400, Jackson ImmunoResearch) and FITC-conjugated Goat anti-Mouse IgG (1:200, Jackson ImmunoResearch) (Table 2.6). Cells were washed in PBS, and mounted onto Apex Superior Adhesive Slides (Leica, Biosystems) using Vectashield with DAPI mounting resin.

Images were taken using a Leica SP5 laser scanning microscope and analysed in ImageJ using the Dead_Easy mitoglia PlugIn.

For colocalisation studies, images were analysed using the Colocalisation PlugIn.

Antibody	Dilution used at for immuno cytochemistry	Bought from	Product number
Cy3-conjugated Goat anti-Rabbit IgG	1:400	Jackson ImmunoResearch	111-165-144
FITC-conjugated Goat anti-Mouse	1:200	Jackson ImmunoResearch	115-096-072

Table 2.6. List of secondary antibodies used in immunocytochemistry.

2.8 Cloning

2.8.1 Creation of a control plasmid (See Appendix Figures 4-7)

To create an appropriate control for overexpression of XRN1, steps were taken to remove XRN1 sequences from the original plasmid (kindly donated by Dr E. Izaurralde – Max Planck Institute, Tübingen) (Figure 2.4), and then to reanneal the vector backbone with a sequence of DNA with no cut sites in the rest of the plasmid.

2.8.1.1 Plasmid digestion and gel purification:

Plasmid digestion was performed using 1μg plasmid, 1μL of Xho1 restriction enzyme (NEB, product code R0146), 1μL of PstI restriction enzyme (NEB, product code R0140), 10X Buffer 3 (NEB, product code B7003), and made up to 30μL dH_2O. Samples were incubated at 37˚C for 2 hours. After 30 minutes incubation, 1μL of calf intestinal phosphatase (CIP) (NEB, product code M0290S) was added to the mix for the duration of the incubation. A heat denaturation step was performed at 65˚C for 20 minutes to denature the restriction enzymes. Products were then run on a 0.8% agarose gel at 110V for 70 minutes. It was run for a further 35 minutes at 120V for better band separation.

The correct bands were then cut out using a dark reader transilluminator (Clare Chemical Research) and scalpel. The DNA was then purified from the gel using the QIAquick Gel Extraction Kit (Qiagen, product code 28704) according to manufacturer's instructions. DNA concentration was then measured on a NanoDrop One.

2.8.1.2 Recircularisation of linearised plasmid:

To recircularise the plasmid, a sequence of DNA including the Not1 restriction site was chosen because the Not1 enzyme did not have a cut site within the rest of the vector backbone. As a control for the cloning, digesting the recircularised plasmid with Not1 would digest the new plasmid, but not the old one, therefore confirming successful recircularisation. A Not1 oligo pair was ordered (Merck) to ligate into the vector backbone (Table 4). The oligos were annealed to each other in the following reaction: 1μL Forward oligo, 1μL Reverse oligo, 1μL T4 DNA Ligase 10X Buffer (NEB, product code B0202S), 0.5μL T4 polynucleotide kinase (NEB, product code M0201S) and 6.5μL dH_2O (total 10μL reaction). This was placed in the thermal cycler under the following protocol:

1. 37˚C for 30 minutes
2. 95˚C for 5 minutes

3. Reduction of temperature by 5˚C every 1 minute for steady cooling to 25˚C
4. 4˚C for ∞

Following the annealing of the Not1 oligos, a ligation reaction was performed to insert the annealed oligos into the vector backbone in the following way: 50ng of the DNA extracted from the gel in step 1, 1.5µL T4 DNA Ligase 10X Buffer, 1µL T4 DNA Ligase (NEB, product code M0202S), 1.4µL of 1:200 diluted annealed and phosphorylated oligos, made up to 15µL total with dH_2O. This mix was then incubated overnight in the thermal cycler at 16˚C.

Recircularisation of the plasmid with Not1 sequence was confirmed by gel electrophoresis. 20ng/µL of product and 20ng/µL unligated control sample were diluted in 1.5µL of 5X Gel Loading Buffer and loaded onto a 0.8% agarose gel, which was run at 110V for 1 hour and 20 minutes. The gel was then imaged in a UV transilluminator to show successful supercoiling and recircularisation of the plasmid.

The control plasmid was then ready for transformation into DH5-α *E. coli* cells.

2.8.2 Bacterial transformation of *E.coli*:

50µL of competent DH5-α *E.coli* (kindly donated by Dr. H. Jagatia, University of Sussex, later bought from NEB, product code C2988) were thawed, and kept on ice. 1µL of XRN1 plasmid (100ng/µL) (as a control, plasmid map Figure 2.2) and 1µL of newly generated plasmid was added to a bacteria aliquot respectively. Bacteria were then incubated for 30 minutes on ice. Bacteria were subjected to heat shock for 1 minute in a water bath at 42^{o}C. 950µL of SOC nutrient-rich medium was added to the bacteria. The suspension was then incubated for 1 hour at 37^{o}C on a rocking incubator (250rpm). Cells were plated by streaking across an LB-agar plate (60mm petri-dish). Plates were incubated at 37^{o}C overnight in a Memmert oven. Single colonies were picked from around the plates and each single colony was placed into a Universal tube containing 4mLs of LB broth (containing Kanamycin for selection at a concentration of 50mg/mL). The Universal tubes were then incubated at 37^{o}C overnight in a rocking incubator (220rpm). Bacterial growth was indicated by turbidity.

Successful transformation was tested by DNA extraction and PCR. This was performed using the QiaPREP Spin MiniPrep kit (cat. no. 27106) and later extractions were performed using the Qiagen Midi-prep kit. DNA was measured on a NanoDrop One. PCR was performed using XRN1 specific primers (forward and reverse, spanning isoform-001).

The following PCR mix was used: 5µL AmpliTaq Gold 360, 3.6µL water, 0.4µL of the primer pair, and 1µL of the DNA extracted (diluted to 1:100). After the PCR, 2µL of Gel Loading

Dye were added to each product. PCR products were run on a 1.2% agarose gel, containing 1µL of Gel Red (Biotium), alongside a 100bp and a 1kbp DNA ladder (New England BioLabs). The gel was run for 1.25hrs at 120V and visualised in a UV Transilluminator.

2.8.3 Transfection of XRN1 and control DNA:

$3x10^5$cells/well were plated in a 6-well plate in 2mLs of fully supplemented medium. After 24 hrs incubation at 37^oC in 5% CO_2, the media was changed to antibiotic-free full medium. DNA was transfected via Lipofectamine LTX with PLUS reagent (Thermo Fisher, product code A12621) vector according to manufacturer's instructions: 500µL Opti-MEM reduced serum medium was mixed with 2.5µg DNA. 2.5µL PLUS reagent was added to the diluted DNA and incubated for 5 minutes at room temperature. 6.25µL of Lipofectamine LTX was added to the diluted DNA. The complex was incubated for 30 minutes at room temperature. The entire content was then added to its respective well, which was then incubated at 37^oC in 5% CO_2 for 24 hrs. Cells were then harvested for confirmation of XRN1 rescue by Western blotting.

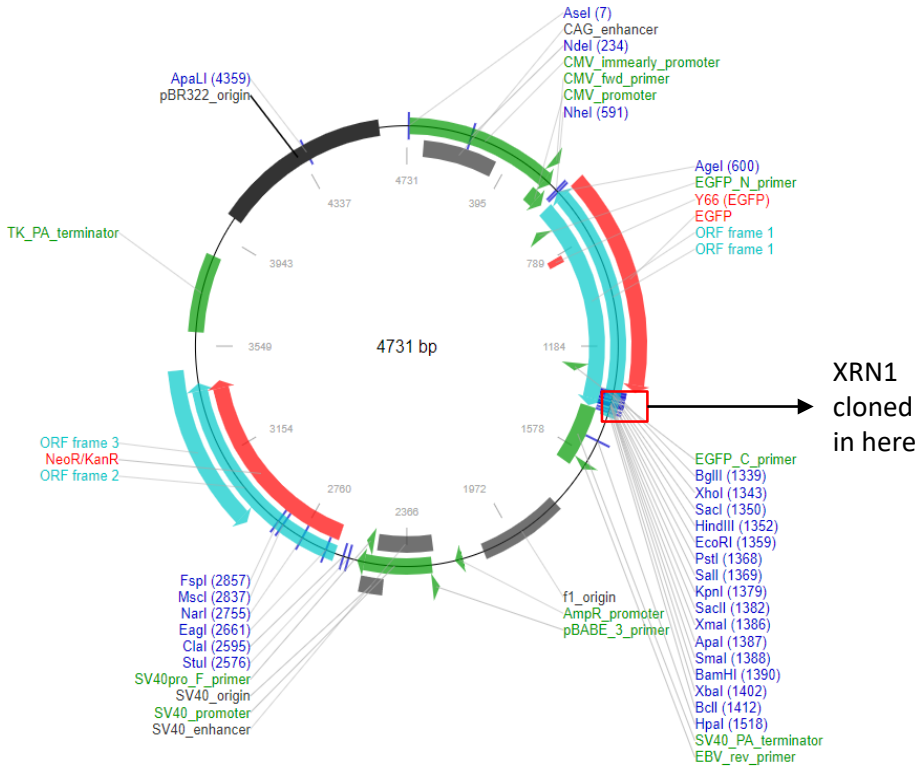


Figure 2.4. pEGFP-C1 plasmid map for XRN1 rescue experiments. XRN1 was cloned into this plasmid at base 1348 of the plasmid sequence, at the site of the Xho1 restriction enzyme, and between Xho1 and a second restriction enzyme, Pst1. With XRN1 cloned in, the plasmid was approximately 10kbp in size.

Chapter 3

Characterisation of exoribonuclease expression in Osteosarcoma and Ewing Sarcoma cells

3.1 Introduction

XRN1 is a 5'- 3' exoribonuclease, and critical member of the RNA stability pathway which acts to degrade mRNA in the cytoplasm. It functions post-transcriptionally to clear the cell of aberrant mRNAs and fine-tune post-transcriptional regulation of gene expression in all tissues (see Introduction Figure 1.7). Recent studies have suggested that activity of XRN1 may be co-translationally linked, however, this is a relatively new idea for the field (Hu *et al.* 2009, Pelechano *et al.* 2015) and the structure for XRN1 receiving mRNA from the translating ribosome has only just been resolved, an exciting development in this field (Tesina *et al.* 2019). XRN1 targets transcripts which have undergone removal of the protective 5' 7-methylguanosine cap by decapping enzymes, which renders the transcript unstable, and open to degradation by XRN1. XRN1 is well documented to be present in punctate granules called Processing Bodies (P-bodies) (Wang *et al.* 2018), where it co-localises with other members of the degradation machinery, including the decapping enzymes necessary for access to the 5' end of the transcript for degradation (Decker and Parker 2012) (it is important to note that XRN1 localisation to P-bodies is not exclusive, and XRN1 can be localised in other cellular compartments (Eulalio *et al.* 2007)).

P-bodies have historically been thought of as the site of mRNA storage prior to translation, and also the site of RNA degradation in the cytoplasm. There are a number of P-body markers in the cell (such as DCP2 and GW182), many of which constitute other facets of RNA degradation machinery (the decapping enzymes), evidence for active RNA degradation at specific points in the cytoplasm (Decker and Parker 2012). However, research from the Izaurralde group has shown that RNA degradation can still take place even when P-bodies are not present, (Eulalio *et al.* 2007), supported by the resolution of XRN1 degrading RNA co-translationally in complex with the translating ribosome, which does not take place in P-bodies (Tesina *et al.* 2019). In addition, there is new information about the existence of a molecular shuttling service, whereby different nucleic acids and proteins can be shuttled between these different molecules depending on the activity of the cell at any given time. Storage of mRNA in P-bodies prior to translation has also been disputed, with theories emerging about storage of mRNA for other reasons, such as storing mRNA as part of the stress response, or as an RNA

degradation stockpile when RNA degradation is suboptimal. This is supported by recent evidence which suggests that P-bodies allow for the phase transition of soluble RNPs into a liquid or solid state, thereby interacting with cytoplasmic RNAs in order to facilitate quick adaptations to the gene expression landscape, perhaps in response to cellular stress, by condensing repressed mRNA regulons (Hubstenberger *et al.* 2017). In this chapter, the localisation of XRN1 with regards to its presence in P-bodies is discussed. Immunocytochemistry experiments will aim to elucidate how localisation of XRN1 in P-bodies may change in cancer phenotypes, from being situated in speckles across the cytoplasm in the normal control, to being much larger, fewer foci gathered at the edges of the nucleus in the cancer cells, suggesting a different role for P-bodies.

RNA degradation has often been overlooked in the study of disease, although interest in the RNA field with regard to its role in various pathologies is increasing. This includes the association of other exoribonucleases, DIS3 and DIS3L2 in blood and kidney cancers (Astuti *et al.* 2012, Robinson *et al.* 2015), as well as evasion the host response by stalling of XRN1 in viral pathogenesis (Moon *et al.* 2012, Chapman *et al.* 2014, Moon *et al.* 2015, Moon *et al.* 2015).

Osteosarcoma (OS) is the fourth most common primary malignancy in childhood and adolescence. OS is a disease which develops at the growth plates of long bones, most commonly, the femur. It is associated with defects in adolescent development, during periods of sustained and fast growth. A major problem with diagnosing this disease stems from the fact that pains in the joints of the limb bones are often attributed to growing pains, and so further investigation is not undertaken. If undiagnosed, OS metastasises to the lungs, and it is this secondary cancer which increases mortality rates in patients. Though its incidence rate is relatively rare compared to other high profile cancers, the effects of treatment of this cancer on patients are long-lasting, and include amputation and the legacy effects of chemotherapy. In addition, 5-year survival rates have remained static since the introduction of neoadjuvant chemotherapies and tumour resection surgeries in the 1970s (Isakoff *et al.* 2015, Rogers and Conran 2019). Survival rates after metastasis remain exceptionally low, at less than 40%, and so investigation into the potential for new therapies is timely.

Ewing sarcoma (EWS) is another type of bone sarcoma which derives from the same lineage of cells (mesenchymal stem cells), although these tumours can be found along the entire bone shaft, and even in extra osseous tissue, such as the surrounding muscle and adipocyte cells, suggesting that EWS occurs earlier on the differentiation pathway (Ozaki 2015). The prognosis for patients with EWS tends to be worse than for OS because this type of sarcoma is generally more aggressive, due to the rapid proliferation of these immature bone progenitor cells. Having

both OS and EWS develop from cells of the same lineage means that there is potential for the development of treatments using the same drug targets for both cancers.

The differential expression of XRN1 in osteosarcoma cell lines and patient samples was first identified in 2002 (Zhang *et al.* 2002). Using semi-quantitative RT-PCR, Zhang *et al* showed that there was reduced expression of XRN1 (denoted as hSEP1 in this paper) mRNA in three out of four primary OS cell lines, and in eight out of nine OS biopsy specimens when compared to foetal osteoblast control cell line, alongside the identification of mutations within the gene in two of the OS cell lines (HOS and U-2 OS) (Zhang *et al.* 2002).

Since 2002, there have been no further publications appertaining to the function of XRN1 in the progression of OS, and the idea of XRN1 being a novel candidate tumour suppressor gene has been lost over time. This may be because its original assigned gene name notation as XRN1 was originally known as the human homologue of the yeast protein, SEP1. Due to the nature of OS as a particularly aggressive and painful disease, along with current treatment options, it seems timely to take the helm and investigate the potential role of XRN1 in the progression of OS, using more up-to-date techniques, equipped with higher levels of sensitivity.

The purpose of this chapter is to illustrate a more complete characterisation of XRN1 expression and localisation in not only OS cell lines, but also in cell lines of another type of bone sarcoma: Ewing sarcoma (EWS). If XRN1 shows the same characteristics of expression in EWS as in OS cells, this would support the idea that XRN1 is a novel tumour suppressor in different types of bone sarcoma, and increases its credibility as a potential drug target. In addition to this, this chapter will elucidate the point within the Central Dogma at which XRN1 expression becomes differential, using a *pre-XRN1* mRNA assay, complementary to the premature *XRN1* mRNA, before post-transcriptional modifications and circularisation occurs. If XRN1 is being differentially regulated at a specific point during the Central Dogma, this will show that there may be other factors involved causing dysregulation, such as possible transcription factors or inhibition by microRNAs, and it could also point to the possible involvement of other members of the transcriptional or translational machinery, as well as the degradation machinery, acting to promote the progression of OS.

Alongside observational studies of the expression of XRN1 in OS, expression of the other exoribonucleases, DIS3, DIS3L1, DIS3L2 and XRN2 in OS cell lines will also be discussed in this chapter. DIS3 is the catalytic member of the nuclear exosome machinery which degrades RNA in the 3'-5' direction. DIS3L1 is the catalytic member of the cytoplasmic exosome which degrades RNA is the 3'-5' direction. DIS3L2 is, like XRN1, an independent 3'-5' exoribonuclease which functions in the cytoplasm. DIS3L2 has been shown to target highly uridylated transcripts

associated with growth (Cheng *et al.* 2019), marking it as a particularly interesting enzyme. The 3'-5' exoribonucleases act after the removal of the poly(A) tail of the transcript at the 3' end by deadenylases, which destabilises the transcript, and makes the transcript vulnerable to degradation by the exosome or DIS3L2. Both DIS3 and DIS3L2 have been implicated in blood cancers and Wilms' Tumour of the kidney respectively, and so it is reasonable to assume that they could be involved with the progression of OS, alongside XRN1. The activity of DIS3L1 remains relatively unknown in most tissues, and so if a potential function for DIS3L1 can be elucidated in OS cells, this will be a novel finding for the field. In contrast, XRN2 is a 5' – 3' exoribonuclease which functions in the nucleus to degrade RNA and promote the termination of transcription. Diseases associated with defective XRN2 decay include amyotrophic lateral sclerosis (ALS), a disease XRN1 has also been implicated in. XRN2 is thought to be involved with the progression of ALS via the formation of DNA/RNA structures (R-loops) (Skourti-Stathaki *et al.* 2011). The expression of each of these exoribonucleases has never been studied before in OS cells, and so if there seems to be dysregulation of these enzymes, this will be a novel finding for research into this disease, and may go some way into explaining how XRN1 is being regulated in OS.

This chapter will also endeavour to clarify whether XRN1 expression is lower in both OS cell lines and in OS patient samples compared to a control cell line. This is a fundamental concept during this project; if XRN1 expression is lower in the patient samples compared to controls, this could have major clinical implications with regard to developing a novel therapy for both OS and other types of sarcoma.

3.2 Aims & Hypothesis:

The aims of this chapter were designed to recapitulate the results seen from a previous publication, and to add a new depth of observation to the expression of XRN1 in OS cells. They were also designed to expand studies of the potential role of XRN1 in OS to another type of bone sarcoma, Ewing sarcoma, and to elucidate the expression profiles of other exoribonucleases in OS. These aims add different scopes to the characterisation of XRN1 in both different settings, as well as its molecular capacity using various molecular biological techniques.

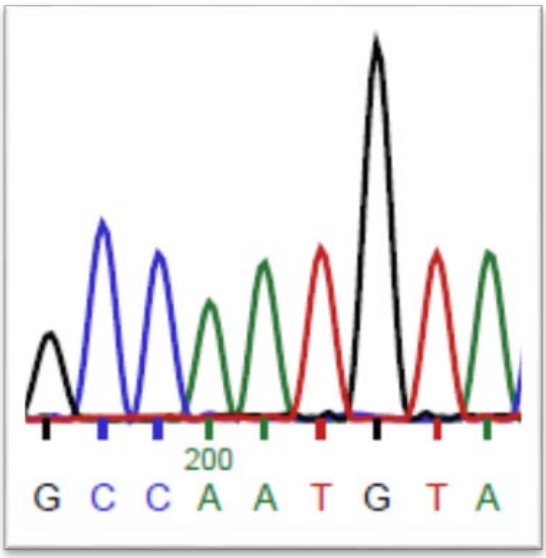
The hypothesis for the work in this chapter was that XRN1 expression is reduced and XRN1 is differentially localised in OS and EWS cell lines compared to a foetal osteoblast control. Therefore, the aims of the chapter were to:

1. Characterise mRNA and protein expression of XRN1 in osteosarcoma cell lines.
2. Characterise mRNA and protein expression of XRN1 in Ewing sarcoma cell lines.
3. Determine whether XRN1 is being regulated at the transcriptional, post-transcriptional or translational level.
4. Determine the localisation of XRN1 in osteosarcoma cells, and determine if there is colocalisation of XRN1 with P-bodies.
5. Analyse *XRN1* mRNA expression in patient samples.
6. Characterise the expression of the 3'-5 exoribonucleases, DIS3, DIS3L1, DIS3L2 and 5'-3' exoribonuclease, XRN2 in the OS cell lines.

3.3 Characterisation of XRN1 mRNA and protein expression in osteosarcoma cell lines

Forming the basis of this project was the identification of a paper published in 2002, which showed the reduction of *hSEP1* mRNA in 3 out of 4 osteosarcoma cell lines and eight out of nine osteosarcoma patient biopsy samples. SEP1 is the yeast homologue for XRN1, and at the time of publication in 2002, little was known about the function of this protein in cells. The authors detected this change in regulation when comparing mRNA expression to that of the foetal osteoblast control cell line, using semi-quantitative PCR techniques. This paper also identified mutations in both the HOS cell line (referred to as TE85 in the paper) (GAT>AAT, which lead to an amino acid change D1137N), and the U-2 OS cell line (V1484A). The mutation in the HOS cell line was a homozygous mutation in DNA encoding the C-terminus, and the mutation in U-2 OS was shown to be a missense mutation in the DNA encoding the RNase domain which therefore may affect catalytic activity. The authors suggest that these mutations may have accounted for reduced expression of *hSEP1*, despite the mutation in HOS being in a particularly flexible part of the protein, which has yet to be solved by crystal structure (Nagarajan *et al.* 2013). Prior to the beginning of this project, DNA extraction from each of these 2 cell lines was performed, and sequencing spanning these regions showed that the mutation in the HOS cell line was indeed present (Figure 3.1), resulting in the same amino acid change as shown by Zhang *et al* (Zhang *et al.* 2002). However, the mutation in the U-2 OS cell line was not present, and no mutation was detected in the RNase domain of XRN1 in this cell line. It is therefore likely that this mutation has derived as a result of replication errors in the specific line used by the group, probably due to continuous culture.

HOb



HOS

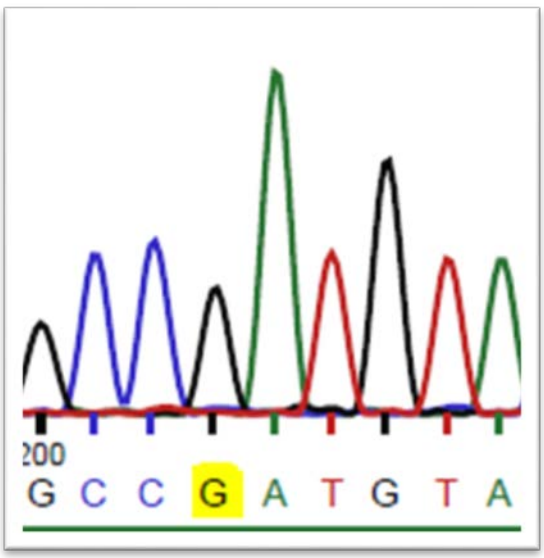


Figure 3.1. Mutations in the C-terminal domain of the XRN1 gene in the HOS cell line. Previous research conducted in the lab shows that, contrary to previous publications, the only mutation found in XRN1 was in the C-terminal domain in the HOS cell line, resulting in a silent mutation. The other mutation tested for was in the RNase domain of XRN1 in the U-2 OS cell line. Results from this DNA sequencing showed no such mutation.

3.3.1 *XRN1* mRNA expression is lower in OS cells

To begin with, the findings by Zhang *et al* were validated using a more up-to-date, and sensitive technique. This involved using the TaqMan qRT-PCR system to detect changes in mRNA expression. Alongside qRT-PCR, growth curves were conducted to observe the rate of growth of each of the osteosarcoma cell lines being tested. These cell lines were: HOS (TE85), MG-63, U-2 OS, and SAOS-2, with human foetal osteoblasts (HOb) utilised as a primary cell line control (See Materials and Methods, Table 2.1). The HOb cells were chosen as the reference control cell line because at the time of investigation, they were the only osteoblast cell type originating in the growth plate that could have been used as the 'healthy' control. The HOb cell line is a primary cell line, and as such does not display a carcinogenic phenotype, allowing comparisons to be made between these cells and the cancer cells, which are themselves immature osteoblasts.

XRN1 gene expression was analysed using the real-time TaqMan qRT-PCR system. The TaqMan qRT-PCR system was chosen because it holds several advantages over other systems. These advantages include the overriding benefit that TaqMan probes label specific target cDNA sequences, rather than monitoring the amplification of any dsDNA sequence, which removes the need for post-PCR processing. TaqMan qRT-PCR is highly sensitive and because each probe is specifically designed for each target gene, the chance of generating false-positives is extremely low.

Results were normalised to *HPRT1* expression. *HPRT1* expression was selected as the reference gene in this experiment because it was judged to be the most consistently expressed at the level of RNA in these cells during previous work in the lab by a Post- Doctoral researcher. Subsequent experiments utilised *GAPDH* as the normaliser gene as this was also consistently expressed across the cell lines.

qRT-PCR showed that *XRN1* expression was lower in three out of four OS cell lines: HOS, U-2 OS, and MG-63 (Figure 3.2B), compared to the foetal osteoblast control. It was also observed that the level of expression of *XRN1* mRNA correlated with the rate of growth of the cell line, whereby the cell line displaying the most reduced expression, HOS (x4-fold lower), was shown to be the most proliferative cell line in culture. This pattern was true of all cell lines, with the only cell line to show no change in *XRN1* expression, SAOS-2, also being the cell line with the longest doubling time (Figure 3.2A). Figure 3.2A shows that the HOS, U-2 OS and MG-63 cell lines demonstrate a much higher proliferative potential than SAOS-2, which does not proliferate as quickly, despite a healthy growth cycle being observed for this cell line. This observation is in line with the hypothesis that XRN1 could be acting as a tumour suppressor in human cells, where the lower the expression, the greater the proliferative potential. It is also important to note that

the cells were an almost even split between which sex karyotype was expressed, indicating that lower expression of XRN1 is not sex-linked.

The results of the qRT-PCR experiment correlated with the growth of each of the cell lines, as shown by cell line growth analysis. Growth curves were conducted to analyse the rate of growth of each cell line over a 9-day period (Figure 3.2A), whereby cells (plated at a density of 80,000 per well in a 6-well plate) were cultured in the same medium over this time period and were counted in triplicate every 24 hours in order to gain a complete, natural profile of the growth of these cells. The mean total for each count was then plotted. Previous studies across 22 OS cell lines have shown that each OS cell line displays differing rates of growth, of which the HOS and U-2 OS cell lines were among the most proliferative in culture, whereas the SAOS-2 cell line did not display such high levels of proliferation, recapitulated in the data presented here. This pattern was also true during investigations of the potential of the cell lines to form tumours *in vitro*, where, again, HOS displayed the most tumourigenic potential, and had the highest colony-forming potential (Lauvrak *et al.* 2013).

The observations seen here are interesting because it gives strength to the idea that reduced XRN1 expression is implicated in cell line growth, supporting the notion that XRN1 could be acting as a tumour suppressor gene. Reduced expression of XRN1 could be affecting degradation of transcripts involved in growth and apoptosis.

3.3.2 XRN1 protein expression is variable across the OS cell lines

The RNA levels give a good insight into the regulation of specific genes in the cell, however, the level of RNA does not always reflect subsequent protein expression. This is because there are many regulatory levels (including RNA degradation) post-transcriptionally and post-translationally which dictate how much RNA is translated into protein. These also include (but are not limited to) regulatory elements in the 5' UTR (uORFs and IRES), differential RNA half-lives of the same isoforms of RNA and protein modification. It is also clear that the correlation between RNA and protein differs between cell lines, and this may especially be the case with the OS cell lines as some are epithelial, and others are fibroblastic (See Table 2.1). RNA to protein ratios can be compared using the Pearson's coefficient, and research suggests that the average correlation between RNA and protein can be measured at 0.6 (where 1 = complete correlation) (Edfors *et al.* 2016). This shows a moderate to strong positive correlation when RNA and protein levels are compared without taking into account gene-specific differences across many different cell lines and tissue types. In addition, it has been shown that this ratio does not necessarily mean that this correlation is always true, for example, the stomatin gene gives a high

Pearson coefficient across tissues, however, relative RNA to protein levels independent of the tissue type of the cell shows that protein expression is relatively low (Edfors *et al.* 2016).

Because of this, protein expression of XRN1 in the OS cell lines was measured to find out whether XRN1 protein level parallels the RNA expression level. Western blotting was used to analyse XRN1 protein level using a widely available XRN1 antibody, shown by a previous post-doctoral Researcher to be very specific to XRN1, giving a band of the correct size (181kDa). Tubulin served as a loading control because it was consistently expressed across both the control cell line and the OS cell lines. XRN1 was shown to be variable across the replicates, leading to results which showed that XRN1 expression was comparatively lower at the level of protein expression (Figure 3.2C), however, statistical significance could only be seen in the U-2 OS cell line (approximately 78% reduction observed), due to the large variation of detected expression. Protein expression was quantified using ImageJ software whereby expression thresholds were determined by eye, and compared to XRN1 expression in the foetal osteoblast control cell line. Depleted protein expression of XRN1 was not expected in the SAOS-2 cell line after results from mRNA expression analysis, and this was confirmed at the protein level. It is important to note that the cell line MG-63 was excluded from protein expression analysis and all subsequent experiments because the cell line demonstrated uncharacteristic behaviours in culture, possibly as a result of culture mineralisation (OS tumour cells often show excessive mineralisation – see section 1.8). Results from this cell line were not trusted, and so the cell line was dropped from all further investigations.

Overall, protein expression of XRN1 in these cell lines has been shown to be variable, however, lower XRN1 protein expression was shown to be statistically significant in the U-2 OS cell line, in line with qRT-PCR RNA expression data, and unchanged in SAOS-2, also in line with qRT-PCR data. Previous work has shown statistically significant lower expression of XRN1 in the HOS cells also (see Appendix, Figure A.1), however, this was not recapitulated by the author.

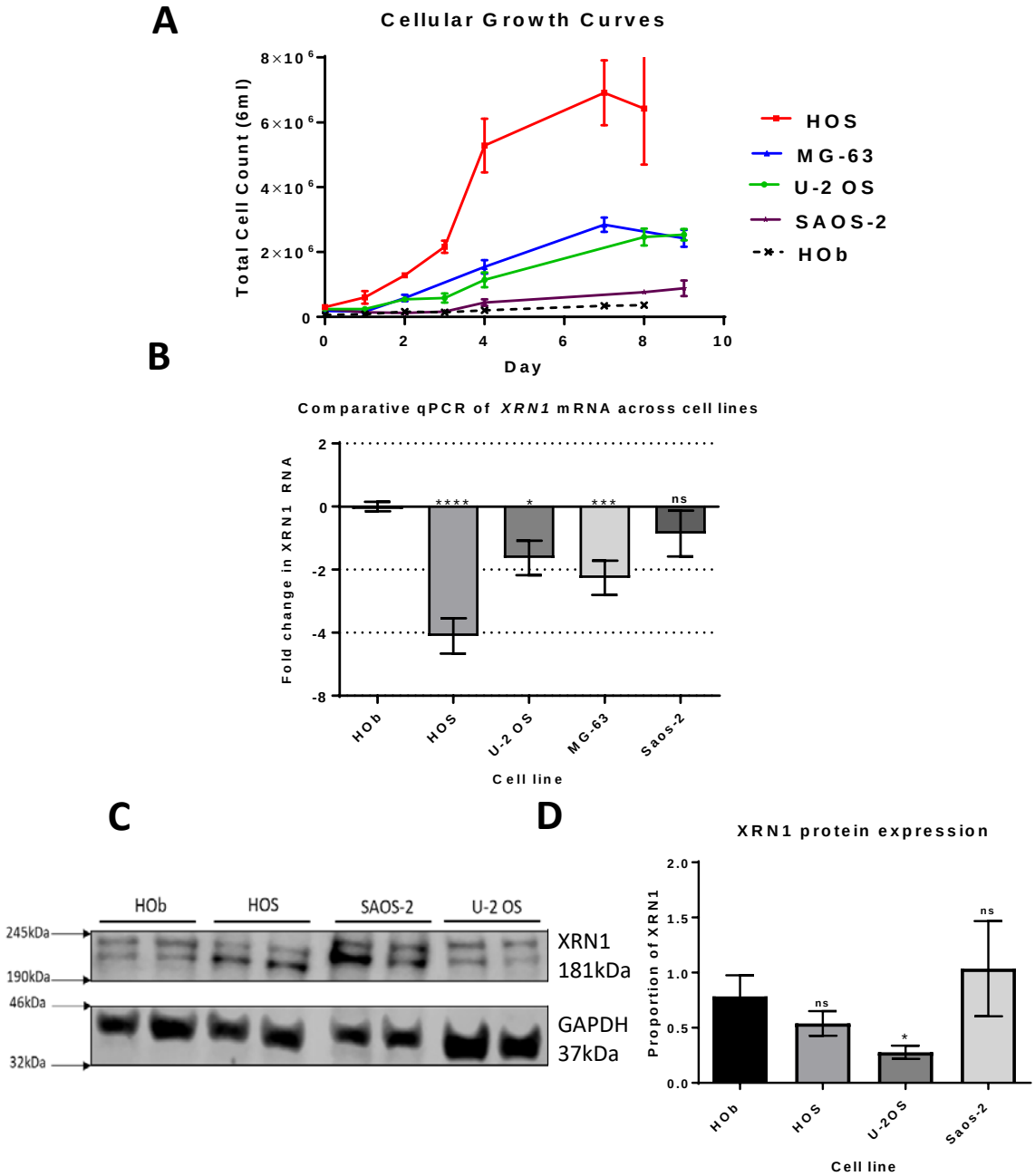


Figure 3.2. Preliminary data confirming lower expression of XRN1 in OS cell lines. A) Linear regression growth curves characterising the rate of growth of each OS cell line in comparison to non-OS foetal osteoblast (HOb) control, error bars = SEM. B) qRT-PCR quantification of XRN1 mRNA expression across OS cell lines in comparison to the HOb control cell line showing the $2^{(ddC_t)}$, normalised to HPRT1. Error bars = SEM, based on n=7 for HOb, MG-63 and U-2 OS, n=5 for HOS and SAOS-2. Statistical quantification performed by unpaired t-test where: HOb vs. HOS p=<0.0001, HOb vs. U-2 OS p=0.0140, HOb vs. MG-63 p=0.0009 and HOb vs. SAOS-2 showed no significant down regulation. C) Gel visualisation of XRN1 expression in OS cell lines, samples loaded in duplicate. D) Western blot quantification of relative XRN1 protein expression in OS cell lines compared to HOb, normalised to α-Tubulin, where error bars = SEM, based on n=7 for HOb, n=8 for U-2 OS and n=4 for HOS and SAOS-2. Statistical quantification performed by unpaired t-test, where HOb vs. HOS p=0.3947, HOb vs. U-2 OS p=0.0198 and HOb vs. SAOS-2 p=0.5514.

3.4 Characterisation of XRN1 mRNA and protein expression in Ewing Sarcoma

Ewing Sarcoma (EWS) is another bone cancer common in adolescents. Though there are some differences between OS and EWS, these cancers primarily evolve from the same mesenchymal stem cell pathway, and genetic differences between the two is limited to the chromosomal translocation of EWS-Fli1 and *EWS-ERG*, the presence of which is only seen in EWS. If XRN1 is being dysregulated during stem cell differentiation of this pathway, leading to oncogenesis of immature osteoblasts, it is reasonable to hypothesise that the same pattern of expression would be seen in EWS, adding strength to the argument that XRN1 could be critical to growth in normal development.

The cell lines utilised for this set of experiments were SK-ES-1 and RD-ES. These were a kind gift from Prof. S. Burchill (University of Leeds). Again, the human foetal osteoblast cell line was used as the control cell line due to the fact that this cancer is derived from the same stem cell lineage. A mesenchymal stem cell line (MSC) was first tested, however, these cells had already senesced before the cell line could expand enough to retrieve adequate levels of RNA for testing. Ideally, the MSCs would have been a better control, because they are not differentiated at all, and so are more like the Ewing sarcoma cell lines than the osteoblasts.

3.4.1 XRN1 mRNA expression is lower in EWS cells

XRN1 expression in EWS was analysed using the TaqMan qRT-PCR system in the same way as previous qRT-PCR, with a total of 10 biological replicates. qRT-PCR showed that *XRN1* mRNA expression is significantly lower in EWS cell lines, SK-ES-1 and RD-ES, relative to the HOb control cell line (Figure 3.3). Cell proliferation rates were determined by growth curves conducted over a 9 day period, with counts being taken every 24hrs (Figure 3.3A), in line with previous growth curves. This showed that, compared to the foetal osteoblasts, SK-ES-1 were the most proliferative cell line, and RD-ES was less proliferative. As seen in the OS cell lines, the most proliferative EWS cell line, SK-ES-1, also displayed the most reduced expression of *XRN1*, with a reduction of x8 fold. The less proliferative cell line, RD-ES, showed a reduction of x3.25 fold (Figure 3.3B) when compared to the HOb control cell line.

3.4.2 XRN1 protein expression is lower in EWS cells

In line with other experiments, the protein expression of XRN1 was also tested in these cell lines, which also showed lower expression at this level. Analysis showed that again, the most proliferative cell line displayed the most reduction of XRN1 protein, with 76.66% less protein in

SK-ES-1 cells than HOb cells, while RD-ES showed 66.22% less protein (Figure 3.3C and D). Protein expression was normalised to the loading control, GAPDH.

Ewing sarcoma presents in cells which are less differentiated than those in OS, in which all cancer cells are defective osteoblasts. These results suggest that dysregulation of XRN1 is occurring at an earlier point in the mesenchymal stem cell differentiation lineage, due to the level of cell immaturity. If lower levels of XRN1 is a common defect in these immature osteoblast cells, it would be reasonable to assume that expression could be lower in the progression of other sarcomas from the same stem cell lineage, such as chondrosarcoma and liposarcoma.

The work on Ewing sarcoma cells was planned and directed by the author and carried out by a full time MSc student, who was under full supervision of the author.

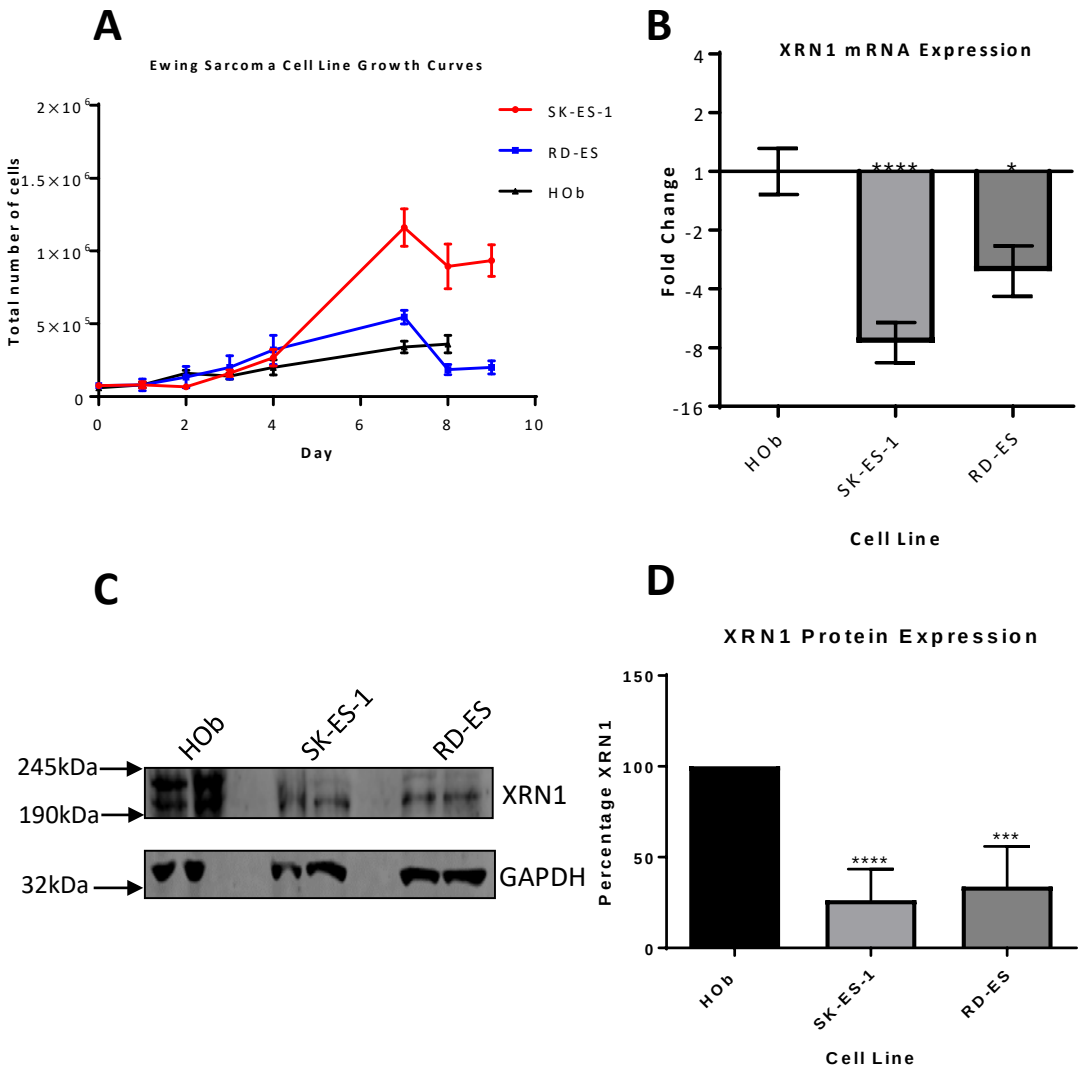


Figure 3.3. XRN1 expression in Ewing Sarcoma. A) Growth curves showing the rate of growth of 2 EWS cell lines relative to the foetal osteoblast (HOb) control over 9 day growth period, whereby cell counts were conducted in triplicate (in conjunction with 0.1% Trypan-blue staining), and the average number plotted. Error bars represent SEM. B) qRT-PCR quantification of XRN1 mRNA expression in EWS shows significant reduction of XRN1 mRNA in EWS cell lines when compared to the HOb control (normalised to GAPDH). Statistical analysis performed by unpaired t-test where: HOb vs. SK-ES-1 p=<0.0001 and HOb vs. RD-ES p=0.0183. Error bars represent SEM, based on n=10 C) Gel image of XRN1 protein expression compared to the HOb (control) cell line, samples loaded in duplicate. D) Western blot quantification of XRN1 protein expression in EWS showing significant reduction of XRN1 protein when compared to the HOb control (normalised to GAPDH), concentration of loaded protein = 3mg/mL in 18µL. Statistical analysis performed by unpaired t-test where: HOb vs. SK-ES-1 p=0.0002 and HOb vs. RD-ES p=0.0012. Error bars represent SEM, based on n=8. Work performed by MSc student, Thomas Burgess under supervision.

3.5 *XRN1* is being regulated post-transcriptionally in OS cells

An important question to ask when assessing the regulation of XRN1 in OS was at which point during the Central Dogma XRN1 was being regulated. This could help to answer how the levels of XRN1 are regulated in the cell, whether XRN1 is self-regulatory or being regulated transcriptionally, or whether it is being regulated by another factor in the cell which hasn't as yet been considered, such as another exoribonuclease targeting the *XRN1* RNA for degradation. It could also be regulated by a translation-acting factor during protein synthesis. It may be that reduced XRN1 expression is actually an indirect consequence of a different mechanism in the cell and finding out at which level of control it is being dysregulated can help to answer this.

A TaqMan primer-probe assay was designed to detect the levels of pre-mRNA by qRT-PCR. This pre-mRNA is defined as mRNA which has not undergone post-transcriptional processing to become mature mRNA. Pre-mRNAs are those which have not yet been spliced to remove the introns, or had modifications added to the 3' or 5' end of the transcript, and are also not yet circularised. Because primers were designed to span the exon-intron junctions and mature mRNA only has exon-exon junctions, the primer-probes assay could not bind to mature *XRN1* mRNA, thereby only picking up pre-mRNA for *XRN1* (see Materials & Methods Table 2.3).

GAPDH was again utilised as a suitable normaliser as in all cell line qRT-PCR and Western blotting because it was consistently expressed across all cell lines, both at mRNA and protein levels. qRT-PCR showed that *XRN1* pre-mRNA expression is not lower prior to post-transcriptional modification. The levels of *XRN1* pre-mRNA showed no reduced expression, in either HOS or SAOS-2 cell lines, compared to the HOb control. Surprisingly, the levels of *XRN1* pre-mRNA were in fact significantly higher (Figure 3.4) in these cell lines. This finding was very interesting, because it means that the notion of XRN1 being dysregulated at the level of transcription can be ruled out. *XRN1* must, therefore, be dysregulated by another factor, post-transcriptionally, and this could potentially be by itself, or other exoribonucleases. The implication of the roles of other exoribonuclease being involved in the reduction of *XRN1* mRNA could be an interesting aspect of carcinogenesis, leading to aberrant targeting of XRN1 for an as yet unknown reason. Other factors targeting *XRN1* also cannot be ruled out, and there is evidence to suggest that XRN1 can be inhibited by miRNA binding at motifs within the mRNA sequence (Table 3.1). If miRNA binding is responsible for *XRN1* dysregulation, it will be necessary in the future to quantify the expression of miRNAs which might be inhibiting *XRN1* in OS cells.

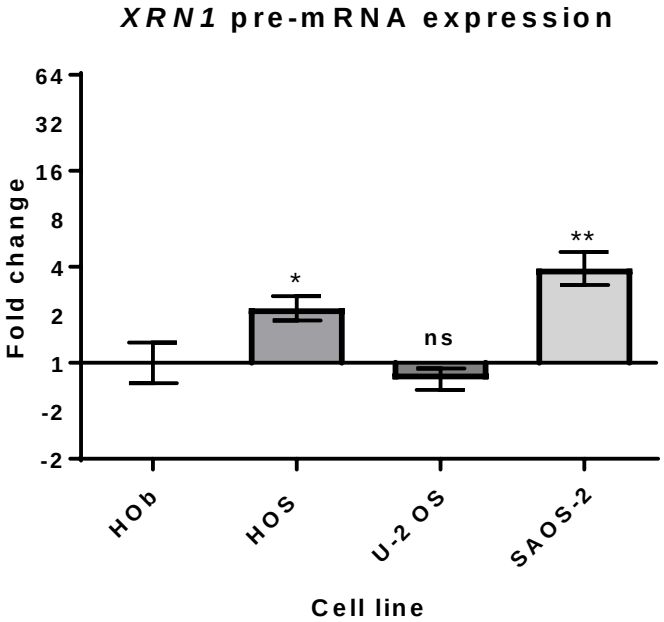


Figure 3.4. Expression of *XRN1* pre-mRNA in OS cell lines relative to the HOb control cell line. Bar graph quantification of the level of *XRN1* pre-mRNA expression showing higher *XRN1* expression before post-transcriptional modification to mature *XRN1* mRNA in 2 of 3 OS cell lines (normalised to GAPDH). This is in contrast to post-transcriptional mature mRNA of *XRN1*, which shows lower expression. Statistical analysis performed by unpaired t-test, where HOb vs. HOS p=0.0447, HOb vs. U-2 OS = ns, and HOb vs. SAOS-2 p=0.0048, based on n=6 biological replicates. Expression was normalised to GAPDH. Error bars represent SEM.

miRNA	Binding site within 3' UTR of XRN1
miR-550	TCAGGCA
miR-202	ATACCTC
miR-20a & miR20b	AGCATTTT
miR-637	GCCTCCAG
miR-891a & miR-1273	TTGTTGCC
miR-10a & miR-10b	TACAGGGT
miR-1289	TCCACTCC
miR-542-3p	TCTGTCAT
miR-422a & miR-373	AAGTTCAG
miR-920	CAGTTCCC
miR-632 & miR-654-3p	AGTAGACAT
miR-940	CCCTGTCT
miR-421	TCTGTTGA
miR-608	ACCACCCC
miR-1255a & miR-1255b	GCTCATTC
miR-373	GAGCACTT
miR-1300	CCTTTTCA
miR-98	ATTACCTC
Let-7*a,****b****,c,d,g,i & miR-181a*	TTGAATGT
miR-936	CTCTACTG
miR-943	AACAGTTA
miR-578	ACAAGGA
miR-549	TGGTTGTC
miR-214	GCTTGCTG
miR-1270	TATTTCCA
miR-768-5p	TTCTCCAA
miR-92a *& miR-25*	AGTGCAAT
miR-539	GATTTCTC
miR-1294	AGCCTCA
miR-577	TGTGCAAA

Table 3.1. Known binding sites of miRNAs within the 3' UTR of XRN1. Those miRNAs with associated pathologies are highlighted in bold. Information taken from http://www.targetscan.org/cgi-bin/targetscan/vert_72/view_gene.cgi?rs=ENST00000264951.4&taxid=9606&members=&showcnc=0&shownc=0&showncf1=&showncf2=&subset=1. Continued over page.

miRNA	Binding site within 3' UTR of XRN1
miR-17-5p, miR-20-5p, miR-93-5p, miR-106-5p, miR-519-3p	GCACUUUA
miR-338-3p	AUGCUGGU
miR-499a-5p	AGUCUUAA
miR-205-5p	UGAAGGA
miR-145-5p	AACUGGAA
miR-137	AGCAAUA
miR-124-3p	GUGCCUU
miR-204-5p % miR-211-5p	AAAGGGAA
miR-29-3p	UGGUGCU
miR-155-5p	AGCAUUA
miR-138-5p	ACCAGCA

Table 3.1. Known binding sites of miRNAs within the 3' UTR of XRN1. Those miRNAs with associated pathologies are highlighted in bold. Information taken from http://www.targetscan.org/cgi-bin/targetscan/vert_72/view_gene.cgi?rs=ENST00000264951.4&taxid=9606&members=&showcnc=0&shownc=0&showncf1=&showncf2=&subset=1.

3.6 Localisation of XRN1 in the cell

3.6.1 XRN1 localisation is differential between OS cells and control cells

Alongside studies for measuring the relative expression of XRN1 mRNA and protein within the cells, the localisation of XRN1 particles within the cell was also studied. Immunocytochemistry was used in order to confirm both mRNA and protein expression profiles of XRN1 in the OS cell lines, and also to confirm the cellular localisation of XRN1. This could offer corroborative evidence for a change in protein expression, and also indicate whether localisation of XRN1 is differential in cancer cells when compared to the HOb cells, which also could suggest differential functions of XRN1 or cellular sequestration. The number of XRN1 foci visible throughout the cell was measured using Z-stack imaging confocal microscopy, and colocalisation of these foci with the P-body marker, DCP2, was also measured. XRN1 was seen to be widely distributed in the HOb cell line, in smaller speckles, compared to the OS cell lines, the more aggressive of which displayed larger punctate foci, smaller in number and consistent with the previous data showing a reduction in XRN1 within the cell (Figure 3.5). The least aggressive cell line displayed a similar phenotype to the HOb cell line, as expected, as it also showed no significantly lower expression of XRN1 at both mRNA and protein level. The foci in this cell line, SAOS-2, displayed a more uniform, speckled scatter of XRN1 foci throughout the cytoplasm, in line with that observed in the HOb cell line, and not the condensation of XRN1 foci into fewer, larger, and more punctate foci as seen in the HOS cell line. This could be indicative of the level of stress experienced by each of the cell lines, as the cell lines which are more aggressive are generally in a state of higher stress, given the environment they are cultured in (more competition and less availability of nutrients). It would be interesting to stain the cells for a stress granule marker, such as G3BP1, in the future to ascertain whether this is the reason for the change in localisation.

XRN1 was not detected in the nucleus, as has been observed in a previous publication (Medina *et al.* 2014). The relative amount of XRN1 foci seen in the cells was normalised to cell size to account for the vast difference in size between the cell lines. The control line cells were typically 6x larger than the cancer cells, probably because the cancer cells were much less differentiated. By normalising to cell size, the quantification of XRN1 foci matched the results seen in the qRT-PCR data for *XRN1* mRNA expression in the HOS cell line, and also the SAOS-2 cell line. Results for U-2 OS showed more XRN1 foci than the control, which is not in line with both the mRNA expression data and protein expression data for this cell line.

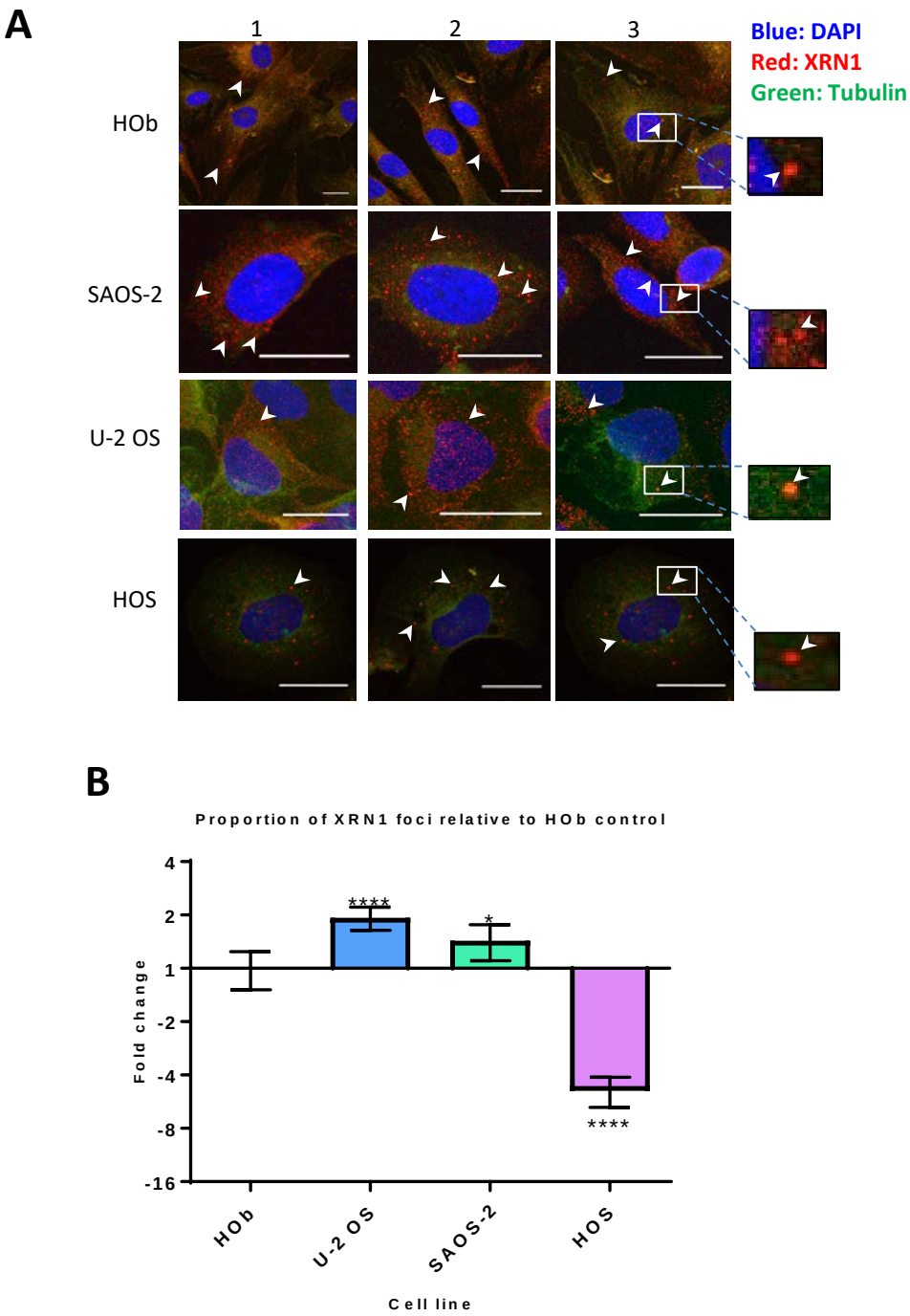


Figure 3.5. Immunofluorescent labelling of XRN1 to show its distribution within the cell. A) Endogenous XRN1 is localised within the cytoplasm of all tested cell lines. XRN1 is shown in red, Tubulin is shown in green. Scale bar is equal to 19µM, and nuclei are stained with DAPI. Images taken at 100X objective. White arrows denote exemplary XRN1 foci, insets show magnified examples (x20 zoom). B) Bar graph quantification of the count of XRN1 foci visible in each cell line in comparison to the HOb control cell line normalised to account for difference in cell size (HOb cells are approximately 6x larger than the cancer cells) (see Chapter 2, section 2.7). Statistical analysis was performed using an unpaired t-test where: HOb vs. U-2 OS p=<0.0001, HOb vs. SAOS-2 p= 0.0336 and HOb vs. HOS p=<0.0001 based on n=20-35. Error bars represent the 95% CI of the mean. Work performed by MSc student, Liza Faustino, under supervision of the author.

3.6.2 XRN1 colocalises with P-bodies differentially in OS cells and control cells

Alongside studies of XRN1 localisation in OS cells, the colocalisation of XRN1 with a P-body marker was also tested. Overall, XRN1 displayed complete co-localisation with the P-bodies (as detected by DCP2) in the HOb cell line, with at most partial co-localisation seen in the OS cell lines (Figure 3.6). The Pearson coefficient was used to measure co-localisation in terms of the relationship between XRN1 and the P-body marker, DCP2, whereby the complete co-localisation is equal to 1, and anything less than 1 demonstrates partial co-localisation. This observation showed that reduced co-localisation (where the Pearson coefficient of XRN1:DCP2 is <1) was observed in all the cancer cell lines when compared to the HOb control. This suggests that not only is XRN1 expression lower in the OS cell lines, as shown by qRT-PCR, but it is also differentially localised and therefore unlikely to be partaking in correct RNA degradation.

It is important to take into account the current hypotheses about P-bodies, and the emerging concept about how RNA degradation can still occur in the absence of these granules. The fact that XRN1 is differentially localised does not mean that 5'- 3' RNA degradation in the cytoplasm does not happen, however, it does suggest that XRN1 is acting independently of other RNA degradation machinery, with specific regard to the decapping enzymes. This is line with published literature which suggests that XRN1 activity is not limited to the P-bodies (Eulalio *et al.* 2007). It is also possible that other aspects of the 5' – 3' degradation machinery are not uniquely sequestered in P-bodies. If colocalisation studies were performed on other members of this pathway, it might be observed that XRN1 still demonstrates colocalisation with some aspects of the pathway, but not conventionally in P-bodies. Schools of thought now regard P-bodies as a mechanism of phase separation to protect mRNAs from being translated.

XRN1 localisation studies were designed and carried out by the author and a part-time Masters student, Liza Faustino, under direct supervision of the author.

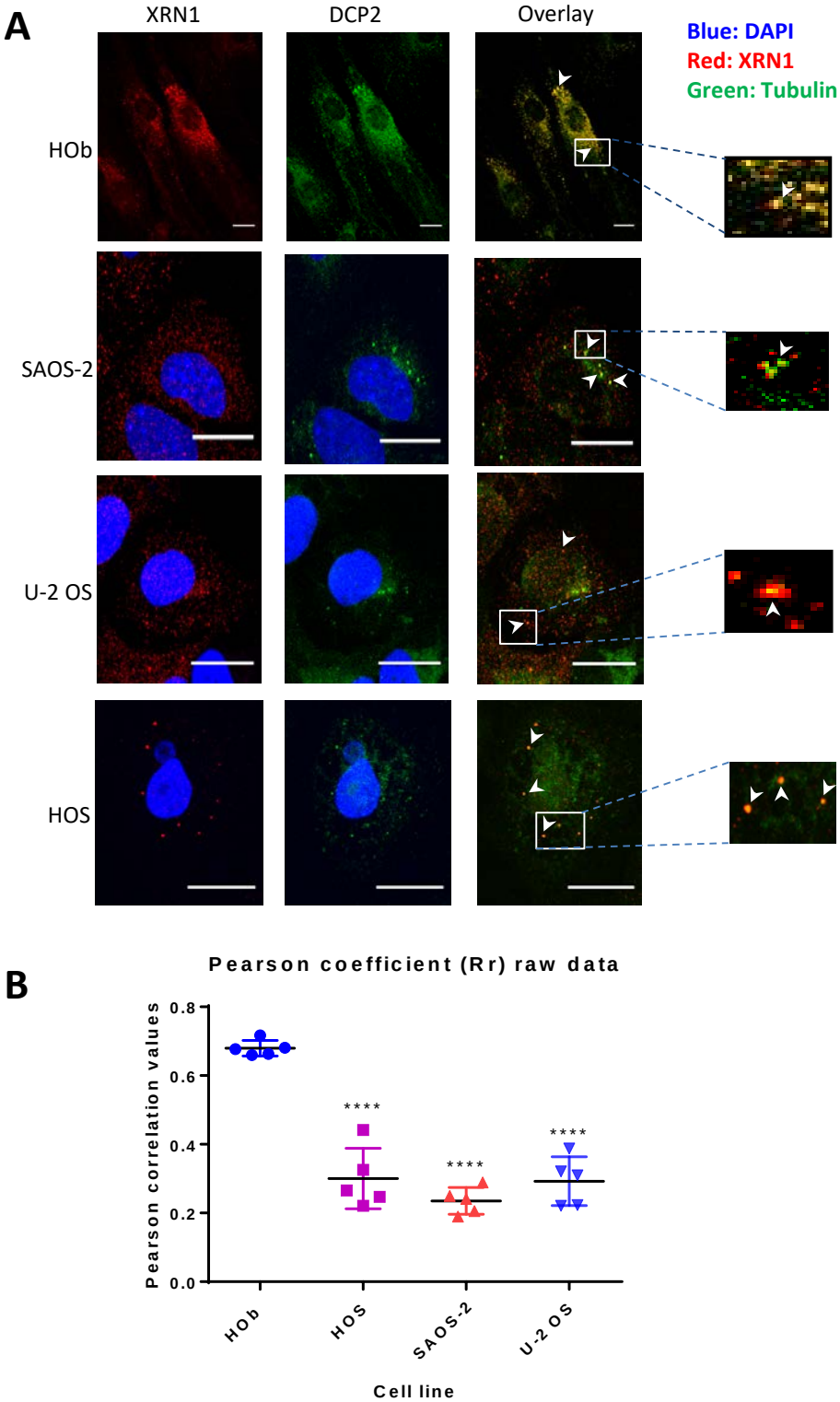


Figure 3.6. Co-localisation of XRN1 with P-body marker, DCP2. A) Immunofluorescent staining of OS cell lines and the HOb control cell line with α-XRN1 (red) and α-DCP2 (green). Scale bars represent 19µM, nuclei are stained with DAPI (blue). Images taken at 100X objective. Insets show magnified (20x zoom) examples of colocalisation. B) Pearson coefficient scale of XRN1:DCP2 colocalisation, where complete colocalisation = 1. Statistical analysis performed by unpaired t-test, where HOb vs. HOS p=<0.0001, HOb vs. SAOS-2 p=<0.0001 and HOb vs. U-2 OS p=<0.0001, based on n=5. Error bars represent S.D.

3.7 Expression of *XRN1* mRNA in human osteosarcoma biopsy samples

After the finding that *XRN1* mRNA expression is lower in OS cell lines, and that XRN1 protein expression was variably lower, it was prudent to investigate the levels of *XRN1* in OS biopsy samples, to find out whether lower *XRN1* expression may have a clinical implication as a novel drug target. This was a major clinical aspect of the project and was determined by studying the levels of *XRN1* mRNA in 14 human OS samples. RNA extraction of small pieces of biopsy was undertaken, and qRT-PCR was performed. The biopsy samples were received from the tissue bank of the Children's Cancer and Leukaemia Group (CCLG), and consisted of both primary OS tumour sections from the growth plate of femurs, tibias and humeri, and also samples of lung metastases. These were received as small, frozen tissue sections (refer to Chapter 2, section 2.1, Table 2.2).

It is interesting to investigate both primary and secondary samples because if lower XRN1 expression is seen in the primary, rather than the secondary cancer, it could indicate a mechanism for the onset of metastasis. Likewise, if XRN1 was even lower in the secondary cancer as opposed to the primary tumour, this would also support a novel notion for the involvement of defective RNA degradation in cancer progression. Of the 14 samples obtained, only 9 were studied by qRT-PCR due to poor quality of the other samples, unfortunately this included the samples obtained from the metastatic lung lesions. More detailed information about the samples is shown in Table 2.2.

3.7.1 Confirmation of a suitable normaliser gene

Because this work involved 9 individual patient samples of different tissue origin, the level of cellular heterogeneity was vast. This is because the biopsies received consisted of a multitude of cell types, including blood cells and possibly cells from other tissues, each with their own expression patterns, and each very different between patients. It was also difficult to confirm how the samples were collected, for these samples the information was not available. Changes in the way samples are collected could lead to differential gene expression due to stress, or degradation of important RNAs and proteins. Cell lines do not provide a landscape of heterogeneity in this way, and so using patient samples to analyse the expression of XRN1 is needed to verify the clinical importance of this work. For these reasons, finding a suitable normaliser gene was very important. Conventional genes such as *HPRT1* and *GAPDH* which were used previously with the homogeneous cell lines were found to be unsuitable for this experiment due to inconsistent gene expression. A preliminary test of alternative control assays was conducted by a post-doctoral researcher (Dr Chris Jones) in the lab, which suggested

that *PES1*, an rRNA processing factor, could be used as a housekeeping gene for the patient samples because it was consistently expressed in patient samples. After performing qRT-PCR assessing the consistency of all three potential housekeepers, *PES1* was indeed found to be the most consistent (Figure 3.7).

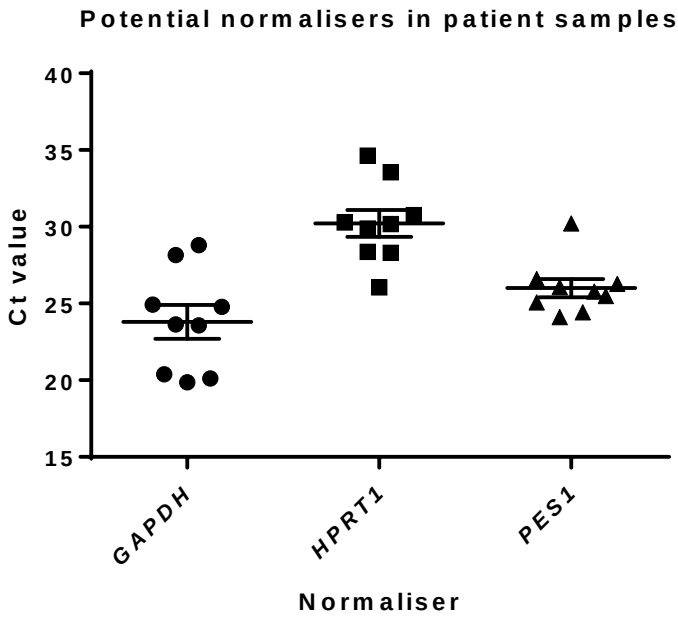


Figure 3.7. qRT-PCR analysis of 3 potential normalisers. Assessment of *GAPDH*, *HPRT1* and *PES1* raw C_t consistency showed that *PES1* was the most consistently expressed normaliser gene across 9 patient samples. Error bars represent SEM.

3.7.2 *XRN1* mRNA is lower in patient samples

qRT-PCR showed that *XRN1* mRNA expression was lower by x16 fold in patient biopsy samples in comparison with the control HOb cell line (Figure 3.8). It is also interesting that the biopsies of tumours resected from the leg (red) tended to show lower levels of *XRN1* than biopsies from tumours in the upper body (blue). This is true for all points except one. At this time it is not possible to speculate why this might be the case, but does provide a basis for future investigation into the differences between tumours arising in different places, and also whether excessive down reduction of *XRN1* in these tumours is pushing higher proliferation rates.

Using the HOb cell line as a control is a limiting factor of this investigation because it is not representative of *in vivo* conditions, and may exhibit differential gene expression as a result of being in culture. However, it has proven extremely difficult to find a more appropriate control at this time, such as healthy tissue from patient-matched samples, for obvious ethical reasons. More suitable controls for this experiment include using age-matched healthy bone, or biopsies taken from the non-cancerous growth plate of the same bone. It would be inappropriate to use bone material removed during replacement therapies of the hip or knee due to arthritis, as has been proposed in the past, because these bones are also diseased, and so may display altered genotypes and gene expression profiles. Samples such as these may also be undergoing necrosis at the time of resection. The use of an adult osteoblast cell line, rather than a foetal osteoblast cell line as a control would also be preferable, because of the differences in gene expression at different stages of development (refer to Chapter 6, section 6.8.1).

qRT-PCR analysis of *XRN1* expression in patient samples has shown that *XRN1* expression is lower both *in vivo* and *in vitro*, providing substantial evidence that there may be an interesting role for this gene during the development of osteosarcoma, given the vital regulatory activity of XRN1 in the RNA decay pathway.

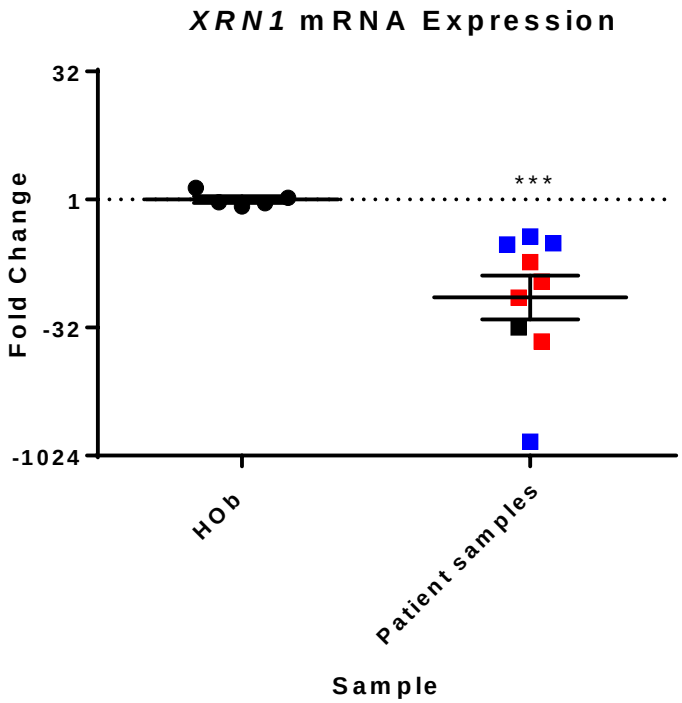


Figure 3.8. *XRN1* expression in patient samples. Dot plot representing the distribution of the fold change of *XRN1* mRNA expression individually in patient biopsy samples, where red = tumour origin is the hip and femur, and blue = tumour origin is the scapula or humerus. Black is denoted to the sample of unknown origin. *XRN1* expression in these samples was measured in comparison to the foetal osteoblast cell line (HOb), normalised to *PES1*. Data shown is fold change plotted $2^{(\Delta\Delta C_t)}$, error bars represent SEM obtained from 9 patient samples and 5 biological HOb replicates. Statistical analysis was performed by a Mann-Whitney test where p=0.0010.

3.8 Expression of DIS3, DIS3L1, DIS3L2 and XRN2 in OS cells

The lower expression of XRN1 in OS cell lines and patient samples shows that there may be a role for the pathways involved in RNA degradation in the progression of cancer. With this, and to make the work more novel, it was decided to characterise the expression of the other exoribonucleases in the OS cells: DIS3, DIS3L1, DIS3L2 and XRN2. DIS3 has already been implicated in the progression of acute myeloid leukaemia, chronic monocytic myeloid leukaemia and multiple myeloma, whereas DIS3L2 has been implicated in the progression of Wilms' Tumour of the kidney. Although DIS3L2 is implicated in Wilms' Tumour by genetic aberrations (via partial or complete deletion of the gene) (Astuti *et al.* 2012) rather than post-transcriptional dysregulation, this may also be the case in OS as well, and defective DIS3L2 could, in fact, lead indirectly to the decreased expression of XRN1.

3.8.1 mRNA expression analysis of *DIS3*, *DIS3L1* and *DIS3L2*

DIS3 is the catalytic member of the nuclear exosome, which acts to degrade target transcripts in the 3'- 5' direction. DIS3L1 is the catalytic member of the cytoplasmic exosome which also targets transcripts in the 3'- 5' direction. DIS3L2 is the independent exoribonuclease in the cytoplasm which degrades transcripts in the 3'- 5' direction in the same way that XRN1 degrades transcripts in the 5'-3' direction, and has a preference for heavily uridylated transcripts. Both qRT-PCR and Western blotting were utilised to characterise the expression of these exoribonucleases in OS cell lines. mRNA expression data showed that expression of both *DIS3* and *DIS3L2* is significantly higher in the HOS and SAOS-2 cell line, and expression of *DIS3L1* is significantly higher in both HOS and U-2 OS cell lines (Figure 3.9) when compared to the HOb cell line and using *GAPDH* as a normaliser. Interestingly, *DIS3L1* expression seems to mirror *XRN1* mRNA expression (Figure 3.2), in that its expression is highest (more than 4x fold) in HOS, and not significantly higher in SAOS-2. The fact that *DIS3L1* mirrors *XRN1* expression suggests that it may be involved in the post-transcriptional dysregulation of XRN1, either because DIS3L1 is dysregulating XRN1, or because XRN1 is dysregulated by another factor, leading to compensatory up regulation of DIS3L1.

Further analysis of *DIS3L1* in patient samples showed that mRNA expression of this exoribonuclease shows that expression of *DIS3L1* is highly variable in these samples, with a large range of expression, however, overall expression is not significantly different in patient samples when compared to the HOb control, and when using *PES1* as a normaliser (Figure 3.10).

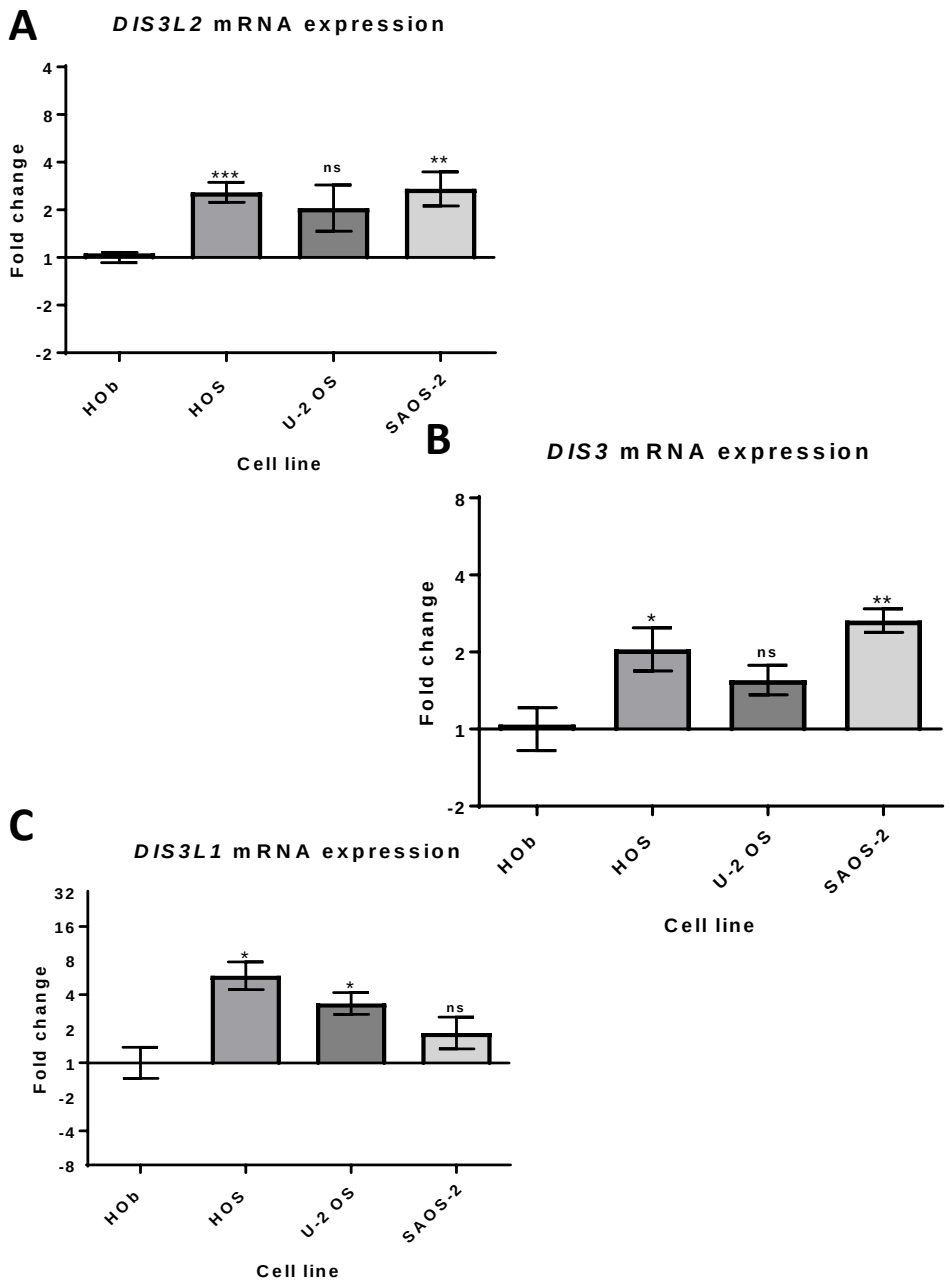


Figure 3.9. mRNA expression of 3'-5' exoribonucleases in OS cell lines compared to HOb. A) Bar graph quantification of *DIS3L2* mRNA expression, showing statistically significant higher expression of *DIS3L2* in 2 of 3 OS cell lines. Statistics performed by unpaired t-test, where HOb vs. HOS p=0.004, HOb vs. U-2 OS is ns and HOb vs. SAOS-2 p=0.0049, based on n=5. Expression normalised to *GAPDH*, error bars represent SEM. B) Bar graph quantification of *DIS3* mRNA expression showing statistically significant higher expression of *DIS3* in 2 of 3 OS cell lines. Stats performed by unpaired t-test, where HOb vs. HOS p=0.0302, HOb vs. U-2 OS is ns, and HOb vs. SAOS-2 p=0.0022, based on n=5. Expression normalised to *GAPDH*, error bars represent SEM. C) Bar quantification of *DIS3L1* mRNA expression, showing statistically significant higher expression of *DIS3L1* in 2 of 3 OS cell lines. Stats performed by unpaired t-test, where HOb vs. HOS p=0.0138, HOB vs. U-2 OS p= 0.0355 and HOb vs. SAOS-2 is ns, based on n=3. Expression normalised to *GAPDH*, error bars represent SEM.

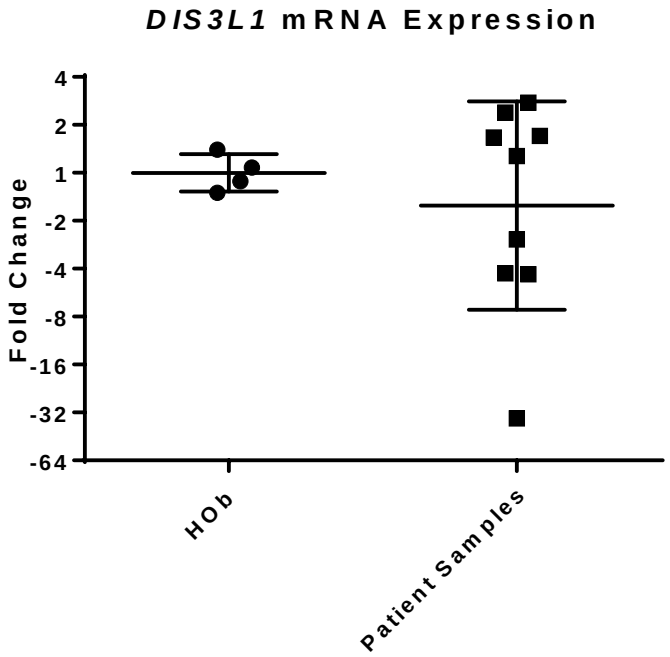


Figure 3.10. Expression of *DIS3L1* mRNA in patient samples. Dot plot representing the distribution of *DIS3L1* expression in each patient sample, showing only one significant point of lower expression when normalised to *PES1*. Statistical analysis performed by Mann-Whitney test, based on n=9, error bars = 95% CI of the mean.

3.8.2 Protein expression analysis of DIS3, DIS3L1 and DIS3L2

Western blot analysis was attempted for each of the 3'- 5' exoribonucleases, however this proved relatively inconclusive due to the quality of available antibodies. DIS3 protein expression was tested using 2 different antibodies from 2 companies (Sigma Aldrich and Thermo Fisher) (see Chapter 2, Table 2.5), and each was very non-specific (Fig 3.11B). One such example even appeared to show binding in protein lysates taken from *Drosophila melanogaster*, thus protein expression for DIS3 was unquantifiable. In contrast, the antibody for DIS3L2 worked very specifically, and protein expression data showed that expression of this enzyme was significantly higher in HOS, U-2 OS and SAOS-2 cell lines, compared the foetal osteoblast control (Figure 3.11A), in partial agreement with the mRNA expression data, when normalised to GAPDH. This result was highly interesting, because it also suggested the involvement of DIS3L2 in the dysregulation of XRN1. Protein expression of DIS3L2 in U-2 OS showed that it was expressed at higher levels, which was not the case when *DIS3L2* mRNA was tested, suggesting that there might be a change in translation of DIS3L2 in this cell line.

Protein expression data for DIS3L1 showed that this enzyme is expressed at extremely low levels in this tissue type. Western blot attempts to try to optimise experimental conditions to see if this helped in the detection of DIS3L1 were unsuccessful, so it was decided to search the Human Protein Atlas, to see if DIS3L1 protein is detectable in bone tissue. The Human Protein Atlas verified that protein expression in bone cells is very low (Figure 3.12B), which explains why it was hard to detect in the OS cell lines. As with a lot of techniques, there is always the risk of sample (in this case, protein) degradation when performing Western blots, and so this might help to explain why detection of DIS3L1 in the OS cell lines was particularly difficult. DIS3L1 was, however, detected in another human cell line (H929) at quantifiable levels which meant that there was no problem with the antibody (Figure 3.12A).

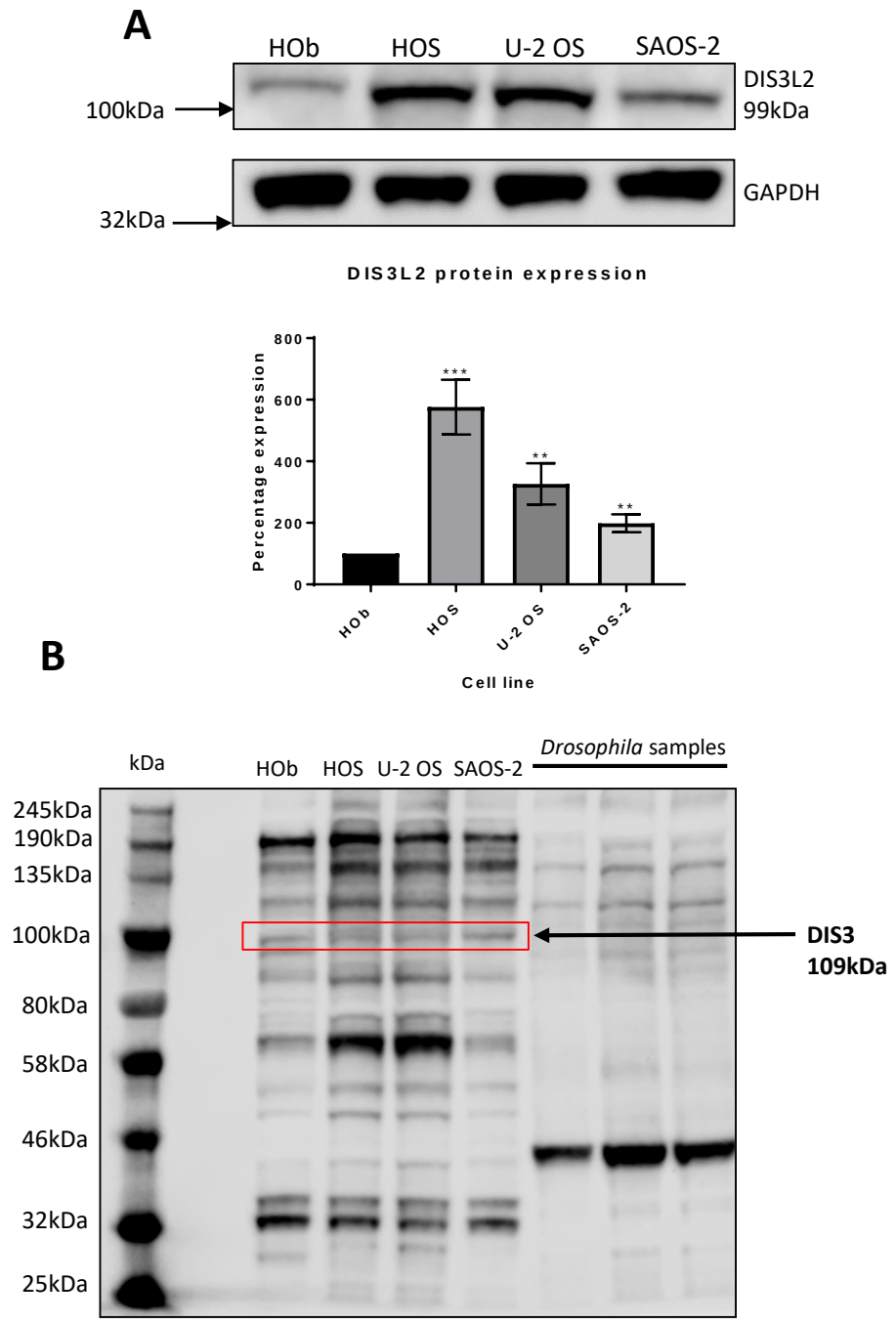


Figure 3.11. Protein expression of DIS3L2 and DIS3 in OS cell lines. A) Gel visualisation and graphical quantification of DIS3L2 in OS cell lines, showing significantly higher expression in protein expression compared to the HOb control. Statistical analysis was performed by an unpaired t-test, based on n=7, where HOb vs. HOS p=0.0002. HOb vs. U-2 OS p=0.0055 and HOb vs. SAOS-2 p=0.0055. Error bars = SEM. Expression was normalised to GAPDH. B) Gel visualisation of OS cell line protein incubated with an antibody for DIS3, however, this image depicts a very non-specific antibody, which appears to also bind in non-human (*Drosophila*) samples as well. Bands for DIS3 are unquantifiable due to the ambiguity of the antibody (expected size 109kDa), based on n=3.

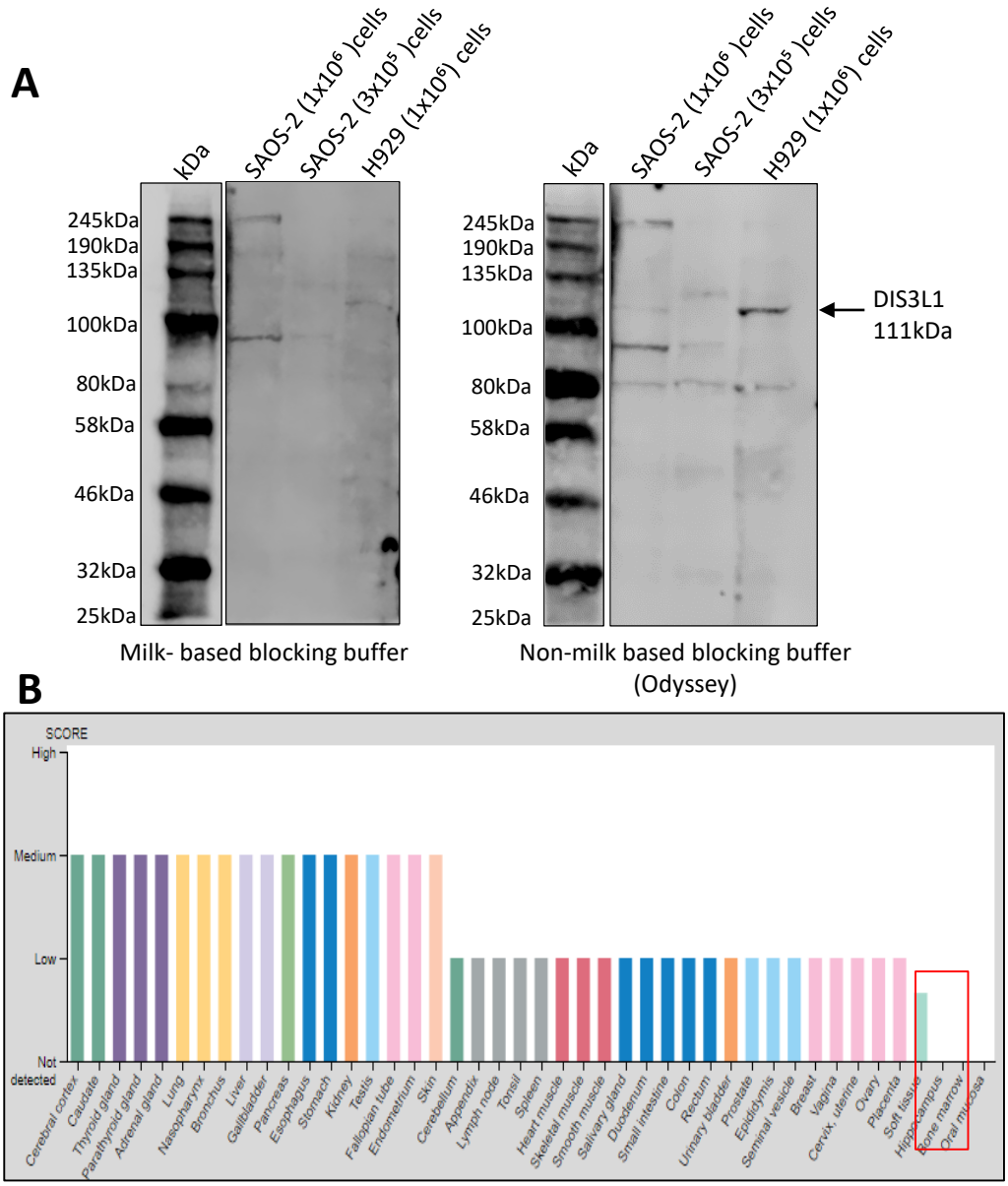


Figure 3.12. DIS3L1 protein expression comparison in SAOS-2 and H929, and across two blocking conditions. A) Gel visualisation of DIS3L1 (111kDa) expression in SAOS-2 cells comparing two different blocking conditions. A leukaemia cell line (H929) was used to ensure that the antibody specific to DIS3L1 was working. DIS3L1 was too low to detect in the SAOS-2 cells (despite lysing higher numbers of cells) in both conditions, though the appearance of 2 bands in lane 1 suggest the expression of different isoforms in bone. It has also been shown that a salmon sperm-based blocking buffer is a cleaner method with less auto-fluorescence than using a milk based buffer, probably due to the increased presence of immunoglobulins in the milk**.** B) Data from the Human Protein Atlas which shows very low expression of DIS3L1 in bone tissue as denoted by the red box (retrieved from https://www.proteinatlas.org/ENSG00000166938-DIS3L/tissue). Data generated using integration of various omics technologies: antibody-based imaging, mass spectrometry-based proteomics, transcriptomics and systems biology.

3.8.3 mRNA expression analysis of *XRN2*

In contrast to DIS3, DIS3L1 and DIS3L2, XRN2 is a 5'- 3' exoribonuclease which functions to degrade RNA in the nucleus. Like XRN1 and DIS3L2, XRN2 is an independent degradation enzyme, and does not exist as part of a large macromolecular complex, such as the exosome. It is generally thought that the main function of XRN2 is its involvement in the termination of transcription by RNA polymerase II. Cleavage of the polyadenylation site liberates the 5' fragment of the newly synthesised RNA, in order for it to be processed and become mature RNA, which is then shuttled out of the nucleus for translation. The remaining 3' fragment remains attached to RNA polymerase II, and XRN2 is subsequently recruited to degrade the 3' fragment in a 5'- 3' direction, promoting transcription termination (Eaton *et al.* 2018).

For the purposes of this project, XRN2 expression was characterised to the level of mRNA, but not protein. It was hypothesised that because the function of XRN2 is primarily for the termination of transcription, rather than the maintenance of cellular homeostasis through the balance of RNA synthesis and degradation, that it would be unlikely that changes in expression of XRN2 would contribute to dysregulation of XRN1, especially since *XRN1* pre-mRNA was unchanged in the cancer cell lines. XRN2 was characterised to complete the characterisation of the set of human exoribonucleases, and to observe if changes in expression of the exoribonucleases was global across all members. TaqMan qRT-PCR was again used to measure the levels of *XRN2* mRNA, which found that *XRN2* expression was significantly higher in all three OS cell lines when compared to the HOb control (Figure 3.13). The highest expression of *XRN2* was in the U-2 OS cell line, whereby expression was x3.15 fold higher than the HOb control. Expression was x2.64 fold higher in the HOS cell line, and x2.18 fold higher in SAOS-2. All results were normalised to *GAPDH*. The regulation of mRNA does not necessarily mean that protein expression will also be higher, however, the up regulation of *XRN2* mRNA suggests that transcription termination events may be disrupted, which could hint as to the dysregulation of XRN1. It would be interesting in the future to measure the protein levels of this enzyme, to confirm disruption of transcription termination.

In summary, the expression of all exoribonucleases (with the exception of XRN1) assessed was higher in at least two out of three OS cell lines at the level of mRNA expression, and expression of DIS3L2 is higher in all three OS cell lines at protein level, compared to the foetal osteoblast control. DIS3 and DIS3L1 were unquantifiable at protein level, however, *DIS3L1* mRNA expression was not significantly changed in the patient samples. *XRN2* expression was higher at the level of mRNA in all three OS cell lines. Taken together, this chapter suggests, at

least at the level of RNA, that RNA decay enzymes are variably expressed in OS cell lines, which could be contributing to the pathogenic phenotype.

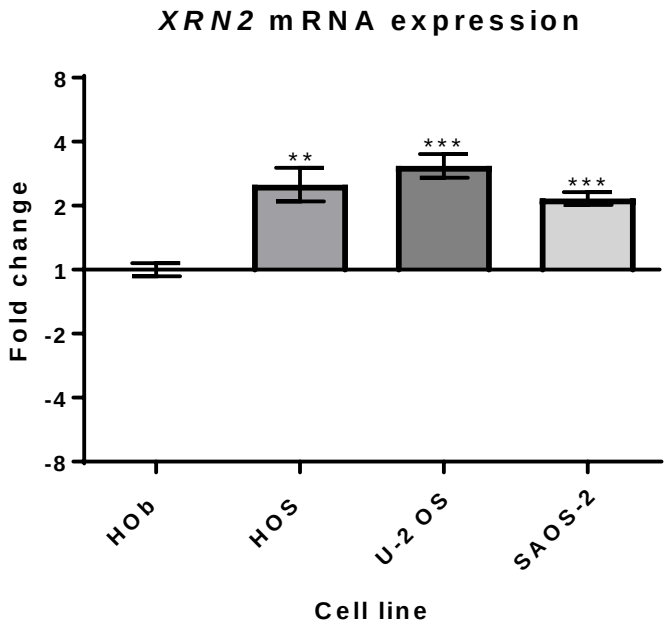


Figure 3.13. Expression of *XRN2* mRNA in OS cells. Graph to show statistically higher expression of the 5' -3' nuclear exoribonuclease, *XRN2*, as quantified by qRT-PCR. Statistical test performed were unpaired t-tests, where HOb vs. HOS p=0.0093, HOb vs. U-2 OS p=0.0009 and HOb vs. SAOS-2 p=0.0006. Error bars = SEM, based on n=3. Expression normalised to *GAPDH.*

3.9 Discussion

As can be observed from the studies presented in this chapter, the expression of XRN1 is at lower levels in OS cell lines when compared to its expression in the foetal osteoblast control cell line, and this could be associated with the rate of growth of the cells. It has been shown that the cell lines with the lowest levels of *XRN1* mRNA expression were also the most proliferative with regard to cell growth. Lower XRN1 expression was also seen in EWS cell lines, adding strength to the argument for XRN1 as pathologically lower in at least these two types of sarcoma. More work will need to be performed in order to get a better idea as to the protein expression landscape of XRN1: Figure 3.2 shows that protein expression of XRN1 is variable across the cell lines, with the only statistically significant reduction in levels observed in the U-2 OS cell line. Expression of XRN1 protein in SAOS-2 also remains consistent with mRNA expression, however, protein expression of XRN1 in the HOS cell line is not significantly lower. It is, however, comparatively lower, and previous experimental work before this project showed that XRN1 protein expression was lower in this cell line, despite this not being observed by the author. The fact that fewer XRN1 foci are seen in the HOS immunocytochemistry images also suggests that XRN1 protein expression is reduced in this cell line despite the variability seen in the Western blotting. Protein expression of XRN1 was consistently lower in the EWS cell lines when compared to the HOb control cell line, which also correlated with the mRNA expression and rate of growth of these cell lines.

Although the use of foetal osteoblast cells is a limitation to the work in this chapter, it is arguably more of a limitation when analysing the expression of XRN1 in EWS cells, because EWS cells are histologically classified as being composed of uniformly undifferentiated cells, which is not a classification for OS cells (Tu *et al.* 2017). The heterogeneity of a EWS tumour could be vast. Osteoblast cells are most commonly found in the growth plate, whereas EWS cancers can form anywhere along the bone shaft, and indeed, can also be identified in the extra osseous tissue surrounding the bone (for example, in the adipose tissue). This suggests that it would have been prudent to have used other available non-immortalised cells of the MSC lineage, such as early adipocytes, in addition to foetal osteoblasts, to ensure that the same expression pattern of XRN1 was observed. Never the less, the decreased expression of XRN1 is still highly significant when EWS cells are compared with foetal osteoblasts, and so remains an interesting line of further enquiry in the future.

With regard to the immunocytochemistry data, it would be useful to know if XRN1 is co-localising with any of the RNA degradation machinery in the cell, other than the decapping enzyme, DCP2, to rule out experimental error. On the other hand, these results suggest that

defective RNA degradation is taking place in the cancer cells. The fact that there are significantly fewer XRN1 foci in the HOS cells when compared to the HOb cell line supports that XRN1 expression is lower in the OS cell lines, despite the variation in the Western blot protein expression analyses.

XRN1 has historically been associated with cytoplasmic foci which have since been termed P-bodies (Ingelfinger *et al.* 2002, Long and McNally 2003, Cougot *et al.* 2004). P-bodies are known to be sites of mRNA storage, for mRNAs awaiting translation or degradation, however, it is unclear whether reduced degradation by XRN1 leads to an accumulation of mRNA in storage, and whether this could present a stress to the cell. Recent research reported at the RNA UK conference in January 2018 (unpublished) suggests that XRN1 localisation is not limited only to P-bodies, but is also localised in many different particles throughout the cell, the majority of which have yet to be fully identified with regards to functionality. It is thought that these particles may be playing an active role in translation, however, further research is needed to elucidate the exact role of these particles, and the role of XRN1 within these particles. An argument for the disparity in phenotypes of the localisation of XRN1 between HOb cells and the cancer cell lines is that the larger punctate foci could be as a result of cellular stress during tissue culture, caused by the competition for nutrients and hyper-proliferation. It would be interesting to see whether the granules harbour markers of stress granules, such as G3BP1, to ascertain the nature of the granules, as it is plausible that they could be stress granules.

Although XRN1 did not display complete co-localisation with DCP2 in the OS cells compared to the HOb control, this is not indicative of RNA degradation levels, and suggests that XRN1 may be acting independently of the decapping enzymes, which has not been reported before. Generally, it is accepted that XRN1 functions with the decapping enzymes during 5'- 3' decay, however, this data suggests they are not always functioning together, meaning that either RNA degradation is defective, or XRN1 is not functioning together with the decapping enzymes for the catalysis of RNA degradation. There is evidence to suggest that when XRN1 is reduced, a resulting change in the size of P-bodies is observed (Eulalio *et al.* 2007, Lubas *et al.* 2013), supporting the results observed in this chapter.

There is the notion that the differences in XRN1 localisation could be as a result of a different type of RNA degradation taking place in various cellular compartments. For example, it would not be unreasonable to assume that rather than 'normal' 5' – 3' degradation taking place, it could be that nonsense-mediated RNA decay or indeed, miRNA-mediated decay is facilitating this change in localisation of XRN1. It would be an interesting observation if NMD RNA decay is occurring at higher levels in the cancer cells compared to the control cell line, and

could explain why XRN1 may be localising in different cellular compartments. NMD machinery has been reported to be recruited to P-bodies for mRNA degradation (Chantarachot and Bailey-Serres 2018), so it is possible that (in cancer cells) when XRN1 is recruited for RNA degradation in P-bodies, it is acting as part of the NMD pathway. In the same way, it is possible that, in cancer cells, general 5' – 3' RNA decay does not necessarily happen in P-bodies. New data has suggested that XRN1 is capable of degrading mRNA co-translationally (Tesina *et al.* 2019), supporting the notion that XRN1 does not always localise to P-bodies. It is reasonable to hypothesise that because cancer cells proliferate faster, they also undergo higher levels of translation, requiring more XRN1 for co-translational decay than XRN1-mediated decay in P-bodies. There is currently no evidence to suggest that the P-body colocalisation marker used in this experiment, DCP2, does not localise outside of the P-bodies. However, given that XRN1 acts co-translationally, this might suggest that DCP2 does in fact localise elsewhere in the cytoplasm, because in order for XRN1 to degrade transcripts they must first be decapped. This could also indicate the presence of unknown decapping factor. If DCP2 does colocalise to areas outside of P-bodies, it can no longer be used as a P-body marker.

As well as XRN1 being shown to be lower in OS and EWS cells, it has also been shown that *XRN1* is also lower in patient biopsy samples. Again, the use of foetal osteoblasts as a control for these samples is not ideal, however, due to the nature of the ethics, gaining patient matched healthy tissue for controls proved impossible within the time frame of this project. Arguably, bone tissue from necrotic hips joints of older patients under gong hip replacement surgery could have been used as controls, however, it was felt that these would also not have been very good controls, due to a number of factors:

1. The patients would not be age-matched, and so XRN1 expression may have been differential due to the life-stage of the patient, rather than because of carcinogenesis.
2. Tissue from necrotic hip joints would not show the normal level of XRN1, purely because they are necrotic, and so display a totally different genotype than those in tumours. Indeed, XRN1 has been shown to have a role in apoptosis (Waldron *et al.* 2015) and so these tissues may also have differentially expressed XRN1.
3. Necrotic hip joints are diseased tissues, usually due to arthritis, and so the expression profile of the tissue will probably differ from both the osteosarcoma tissue and also healthy hip joints.

It is significant that *XRN1* has been found to be expressed at lower levels in the OS patient samples, because these samples are very heterogeneous in their cellular, and therefore

genetic, make up. It is well known that tumours can recruit the tissue's healthy cells in order to promote its own survival, and so it would be expected that the patient samples are made up of both tumour and healthy cells. Levels of XRN1 in healthy osteoblast cells from adolescents are not known, so it is difficult to make a comparison.

Given the heterogeneity of the patient biopsy samples, it is remarkable that *XRN1* was found to be consistently lower, albeit with much variation, strengthening the evidence that XRN1 expression is lower *in vivo*, at least to the level of mRNA expression, compared to the HOb control. It was not possible to conduct protein expression studies in the patient samples due to the limited amount of material, and the fact that XRN1 is a highly unstable protein, the freeze-thaw cycles of these samples may have led to decreased sample integrity. There was also no information provided with the samples which pertained to how quickly the samples were initially frozen from the point of resection, which could also have contributed to decreased sample quality. Protein expression studies on such samples may not show the true state of protein expression at the time of harvest.

Another limitation to this part of the chapter is that the ratio of normal cells to cancerous cells is unknown. One way of finding this out would be to separate the cells in the samples using flow cytometry according to fluorescence of cell-type specific markers. Unfortunately, the small sample size used in this project prohibited the use of material for flow cytometry. If more OS samples could be accrued, then this is something that could be undertaken in the future, and *XRN1* mRNA expression could then be measured in both the healthy and cancerous cells separately.

It would be interesting in the future to characterise the expression of XRN1 in other sarcomas, and indeed, in other types of cancer, such as carcinomas, to identify whether the dysregulation of XRN1 is in fact a global manifestation in cancer cells generally, or whether it is specific to sarcomas of the bone. In addition to this, it would be interesting to see if the phenotype observed in EWS cells is observed in EWS patient samples as well, as this would add weight to the idea that XRN1 is indeed dysregulated in bone cancer patient samples. All sarcomas are derived from the mesenchymal stem cell lineage, so theoretically one could expect that XRN1 expression is lower in most sarcomas, if this is a product of defective differentiation. There were attempts made throughout the project to culture and characterise the expression of XRN1 in SKUT-1 cells (gifted by Dr. An Coosemans, KU Leuven, Belgium), which are cells derived from uterine sarcomas, however, these cells unfortunately tested positive for mycoplasma, and so were not used in these experiments.

Although it is recognised that the OS cell lines are undifferentiated, over proliferative, immature osteoblast progenitor cells, it is not known how far through the differentiation process these cells were before becoming malignant. It is also unknown how differentiated the HOb control cell line is, though it is generally acknowledged to be more differentiated than the cancer cell lines. It would be important to verify that differential expression of XRN1 in the OS cell lines was due to carcinogenesis, and not a product of stage of differentiation. It is possible that the function of exoribonucleases during differentiation would change throughout the process, as different genes are expressed at specific times, and this could possibly explain why XRN1 is reduced in OS cells at this point. To do this, mesenchymal stem cells could be pushed into the osteogenic differentiation pathway using OsteoMAX-XF medium, and over the differentiation time period (approximately 1 month), the expression of early (Runx2), mid (osteocalcin and osteopontin) and late (alkaline phosphatase) (Ducy *et al.* 1999) differentiation markers could be analysed alongside XRN1 expression, to see how XRN1 is expressed over time in these cells. If XRN1 expression changes throughout differentiation, this could indicate that the OS cells may be at a different stage of differentiation to the HOb cells, and so another aspect to this experiment would be to test the stage of differentiation of the OS cell lines and the HOb control cell line. This would give a definitive answer as to whether the regulation of XRN1 in these cells is due to oncogenesis, or because they are at a different stage of differentiation to the control.

Preliminary work to investigate this has been started, though no conclusions can yet be made.

Ultimately, the observation that XRN1 expression is lower at the level of mRNA in both patient samples and in OS cell lines provides scope for further investigation of a potential role for XRN1 in the progression of OS. In addition, the finding that XRN1 is being post-transcriptionally regulated, rather than transcriptionally or translationally, suggests that there is another factor involved in the regulation of XRN1. This could be because of a number of different reasons:

1. There is involvement of the other exoribonucleases in the cells, aberrantly degrading XRN1 mRNA when it becomes mature.
2. Factors known to inhibit XRN1, such as miRNAs which have been identified by Next Generation Sequencing, may be inhibiting XRN1 expression. It may be that these factors are up regulated, which may lead to aberrant XRN1 inhibition. These factors include a miRNA of the let-7 family, *let7b-5p*, among others such as *miR-410-3p*, and *miR-92a-3p* (Chou *et al.* 2018). There are many miRNAs which have been

implicated in the progression of OS, which are discussed later on this project (Chapter 5).

3. XRN1 may be self-regulating. Though there is no current evidence to suggest that XRN1 is self-regulatory and degrades its own mRNA in response to the maintenance of cellular homeostasis, it is reasonable to suggest that in certain instances XRN1 could degrade its own mRNA. This is with particular regard to cancer cell lines, as these cancer cell lines are infamous for their chromosomal instabilities and genetic aberrancies.

This chapter has already alluded to the potential role of other factors which may be influencing XRN1 expression, specifically DIS3L1 and DIS3L2. Unfortunately, although DIS3L1 may be expressed at a low level, it is virtually impossible to detect in OS cell lines, making it very hard to work with. DIS3L2, however, does exhibit aberrant expression alongside XRN1 (higher expression as opposed to lower expression), and is readily detectable in human OS cells, marking it as another member of the RNA degradation machinery which may be involved in the progression of OS. Interestingly, the only other 5'- 3' exoribonuclease in human cells, *XRN2*, is expressed at significantly higher levels in OS cells when compared to the control cells at the level of mRNA. XRN2 has a function in fine-tuning transcriptional termination in the nucleus, and does not act to degrade mRNA in the same way as XRN1 in the cytoplasm. For this reason, XRN2 was not taken forward with regard to the role of RNA stability in human OS cells, as it is not directly involved in the translation of mRNAs into protein further down the gene expression pathway. It was not felt that XRN2 could offer a reason for lower XRN1 expression through interplay or dysregulation, because XRN1 was found to be dysregulated post-transcriptionally. On the other hand, there may be a different effect of up regulation of XRN2 in the cancer cells, and this may include aberrant transcription termination, which would be interesting to investigate in the future.

XRN1 is known for targeting pro-apoptotic transcripts, and DIS3L2 is known for limiting growth, so the phenomenon observed in this chapter is completely contradictory to the data already published for these exoribonucleases. It could be that the observations made in this chapter are cell-type specific to cells in this tissue, and further investigation is needed into the function of the enzymes to advance the understanding of how these enzymes interact on both a 'normal' and pathological level. It seems that from what is already known about these enzymes, their differential expression in OS cell lines should be detrimental to cancer cells, however, it is not. Perhaps the cancer cells have managed to overcome apoptosis, but have retained the benefits of compensatory proliferation. After all, the OS cells are driven by very different objectives than the control cells.

Subsequent chapters in this thesis will endeavour to uncover the potential role of XRN1 and the other exoribonucleases in the progression of OS, using a siRNA-based transfection system to knock down activity and various phenotypic assays to assess which pathways may be affected by the dysregulation of the exoribonucleases.

Chapter 4

The effects of changing the gene expression exoribonucleases in OS cells using siRNA

4.1 Introduction

The observation that XRN1 levels are lower in OS and EWS cell lines, as well as in patient samples compared to the foetal osteoblast control, led to the natural progression of the project to try to find out why XRN1 is expressed at lower levels in some sarcoma cell lines, and whether it offered a functional advantage. By knocking down XRN1 in SAOS-2 cells, which did not display significantly reduced expression levels of XRN1, the aim of this chapter was to observe whether reducing XRN1 expression could enhance the cancer phenotype in this cell line, given that in cell lines where XRN1 was significantly reduced, proliferation rates were much higher. In addition, the regulation of the exoribonucleases, DIS3, DIS3L1 and DIS3L2 was also analysed. This showed that *DIS3* expression was higher in two out of three OS cell lines to the level of mRNA expression, *DIS3L1* expression was higher at the level of mRNA in all OS cell lines, and DIS3L2 expression was higher at both mRNA and protein level in two out of three OS cell lines. This associative observation could suggest that there may be interplay between the exoribonucleases to account for the reduction of XRN1 expression levels.. Therefore, the knock down of XRN1, DIS3L1 and DIS3L2 is discussed in this chapter, in an attempt to understand the synergism between these enzymes, and to understand the role of XRN1 within the cell.

Cancer cells are well documented as being able to function very well despite molecular defects in pathways that are usually considered as fundamental to cell survival and propagation, for example, defective DNA repair as a result of BRCA mutations (Mahdavi *et al.* 2019). As such, cancer cells become very good at using alternative pathways to overcome these defects, and it is through this theory that chemotherapies can be very successful. An example of this is seen in the use of the PARP-inhibitor drug, Olaparib, to treat BRCA-related breast cancers (Rytelewski *et al.* 2016). Cancers of this nature are highly dependent on the poly(ADP-ribose) polymerase (PARP) pathway of DNA repair, due to defects in BRCA 1 or BRCA 2, which mean that DNA repair by homologous recombination can no longer function properly. By developing a drug to target PARP member proteins, cancer cells can no longer undergo DNA repair via this pathway. During periods of massively increased rates of proliferation and mitosis, the cells undergo apoptosis due to defective DNA repair. This medication has proven to show high efficacy in these cancers because it has the ability to completely knock out the only DNA repair pathway relied on by these cancer cells.

Another example of a drug being used to target specific molecular fingerprints is the second generation drug, Afatinib. This drug has recently been accepted for use within the NHS for treatment of non-small cell lung cancer (NSCLC) in patients, where the pathology of the NSCLC includes a mutation in the EGFR gene. It is a tyrosine-kinase inhibitor which functions to block cell growth by irreversibly inhibiting EGF receptor signalling. Afatanib may also help overcome resistance to first generation treatments for this particular type of cancer, including erlotinib and gefitinib. By targeting the EGFR signaling pathway, the drugs are also effectively inhibiting the multiple downstream growth pathways in NSCLC (Wirth 2015).

The Afatinib drug is preferable to conventional chemotherapy, as non-cancerous cells are less susceptible to treatment, meaning that side effects of drugs like Afatanib are limited. The treatment of patients is therefore more targeted and less severe.

By investigating the roles of XRN1 in OS, and later the potential roles of the other exoribonucleases, it could be determined whether XRN1 is a suitable candidate for future drug target therapies in cancer, or if there are other proteins which are responsible for regulating XRN1, which once targeted, could lead to XRN1 rescue, and increased apoptosis. Previous work conducted in the lab showed that there were no mutations in the XRN1 RNase domain to explain its lower expression levels, and work presented in Chapter 3 shows that XRN1 expression is not lower at the level of transcription; i.e. it is reduced post-transcriptionally. The experiments in this current chapter were designed to find out if lower XRN1 expression had an impact on a number of processes fundamental to cancer cell survival: proliferation, apoptosis, metabolism/viability, or translation, the so-called hallmarks of cancer. Experiments where the expression of DIS3L1 and DIS3L2 was knocked down were designed to investigate whether these exoribonucleases work synergistically with XRN1, and if this is a possible cause for lower XRN1 expression, and higher expression of DIS3L1 and DIS3L2 in OS.

4.2 Aims & Hypothesis:

Given that lower XRN1 expression is observed in OS cell lines and patient samples, the aims were as follows:

1. To optimize the knock down of XRN1 and DIS3L2 using siRNA in both U-2 OS and SAOS-2 cell lines.
2. To observe the effects of XRN1 knock down on SAOS-2 cell line expansion by performing growth curves.

3. To look at how the rate of proliferation changes upon XRN1 knock down using Brd-U staining.
4. To measure the effect of XRN1 knock down on caspase activity using a caspase assay.
5. To measure the effect of XRN1 knock down on cell viability using a WST-1 assay to assess metabolism in the cells.
6. To utilize SUnSET labelling to investigate the effect of XRN1 knock down on the rate of translation.
7. To knock down DIS3L2 in U-2 OS cells to observe the effects on U-2 OS cell line proliferation.
8. To investigate the potential for synergistic regulation between XRN1 and DIS3L1, by performing depletion and co-depletion experiments of XRN1 and DIS3L1, and to investigate the effects this has on the expression of DIS3 and DIS3L2. This will enable the identification of possible co-ordinated regulation between these ribonucleases.

The overall aim for this chapter was to see if knocking down XRN1 confers a growth advantage to cancer cells and to see if this is caused by up regulation of DIS3, DIS3L1 and DIS3L2.

4.3 Optimisation of XRN1 knock down using siRNA

Following on from the consistent observation that XRN1 expression is significantly lower in both OS and EWS cell lines, it was suggested that knock down of XRN1 may be conferring an advantage to the cancer cells. To assess the effect of XRN1 reduction, XRN1 was synthetically down regulated in order to establish a particular role for XRN1 activity in relation to a specific pathway involved in cancer progression. SAOS-2 cells were chosen to knock down XRN1, for the following reasons:

1. They have no significant, natural reduction of XRN1, and so observable knock down using siRNA would be more easily quantified by qRT-PCR.
2. They are cancer cells.
3. They are the slowest growing cell line used in this project, and so any increase in their rate of proliferation would be more observable
4. It would be harder to observe effects in the other cell lines because they already harbour a natural knock down of XRN1, and so may already be growing optimally.

Knock down of XRN1 was chosen over complete knock out techniques (such as CRISPR) because previous work in the lab has shown that complete knock out of XRN1 is developmentally lethal in the *Drosophila* model. Indeed, attendance at specific conferences

(FASEB: Post-transcriptional Control of Gene Expression; Mechanisms of RNA Decay -2018, and EMBO: mRNA Turnover: Mechanisms, Regulation and their implication in Infectious and Age-related Diseases - 2019) showed that knock down of XRN1 by siRNA is widely accepted as the most appropriate mechanism for XRN1 depletion when studying its role in the cell, given that complete knock out is lethal in multiple organismal models (Waldron *et al.* 2015). Optimisation of XRN1 knock down was conducted to establish the optimal concentrations of siRNA necessary to convey an efficient (70-80%) knock down of XRN1 expression. Time courses were also performed to establish the time at which XRN1 is most effectively knocked down, and for how long it stays knocked down for in the cell lines. Both U-2 OS and SAOS-2 were utilized for optimization in order to evaluate the consistency of knock down across cell lines, and to check the efficacy of the technique. Introduction of foreign material to any cell line will convey a certain degree of toxicity, however, it is important to limit the effect of toxicity whilst also achieving efficient knock down.

The use of RNAi by siRNA was selected to knock down XRN1 because it was the technique which involved adding the least amount of foreign material to the cells. An alternative method for achieving 100% knock out of XRN1 is to use CRISPR-Cas9, however, previous work in the Newbury lab has shown that this results in organism lethality (Waldron 2014). Electroporation had also been previously tested in cell lines, but this method resulted in substantial cell death. For these reasons, it was decided that transfecting an XRN1 siRNA to reduce XRN1 expression using a lipid vector system would be best.

Using siRNA in conjunction with Lipofectamine RNAiMAX allows the siRNA to breach the phospholipid bilayer without affecting the integrity of the cell, overcoming electrostatic repulsion. This system incorporates siRNA, which is negatively charged, into a cationic lipid vector (Figure 4.1). Once inside the cell, the RNA Induced Silencing Complex is assembled, and degrades the mRNA with the target sequence. The RISC is comprised of proteins involved in targeting and cleavage of RNA: Argonaute and Dicer. The Piwi domain of the Argonaute protein serves to endolytically cleave the target mRNA, which is then degraded in the 3'-5' direction by the exosome, and the 5'-3' direction by XRN1.

This mechanism of RNAi does not require the use of electroporation to order to make the cells competent to take up these molecules, which can damage cells due to the creation of an electric field to increase cell line permeability. This in itself removes the most of the risk associated with RNAi. Since previous work in the lab (not shown) where a system of electroporation to create pores in the cell membrane for siRNA to enter was shown to have

devastating effects on cell viability, it was thought that using this more gentle technique would be less harmful to the culture. The cell lines used are also adherent cells, which are widely known to be easier to transfect than cells in suspension which often require harsher transfection protocols. Controls used included a +liposome/-siRNA transfection to control for effects caused by the RNAiMAX system and a randomised-sequence scrambled siRNA to control for effects as a direct consequence of XRN1 knock down rather than incorporation of a siRNA. The siRNA used to silence XRN1 targeted exon 11 of the *XRN1* RNA transcript, from base pair 1345.

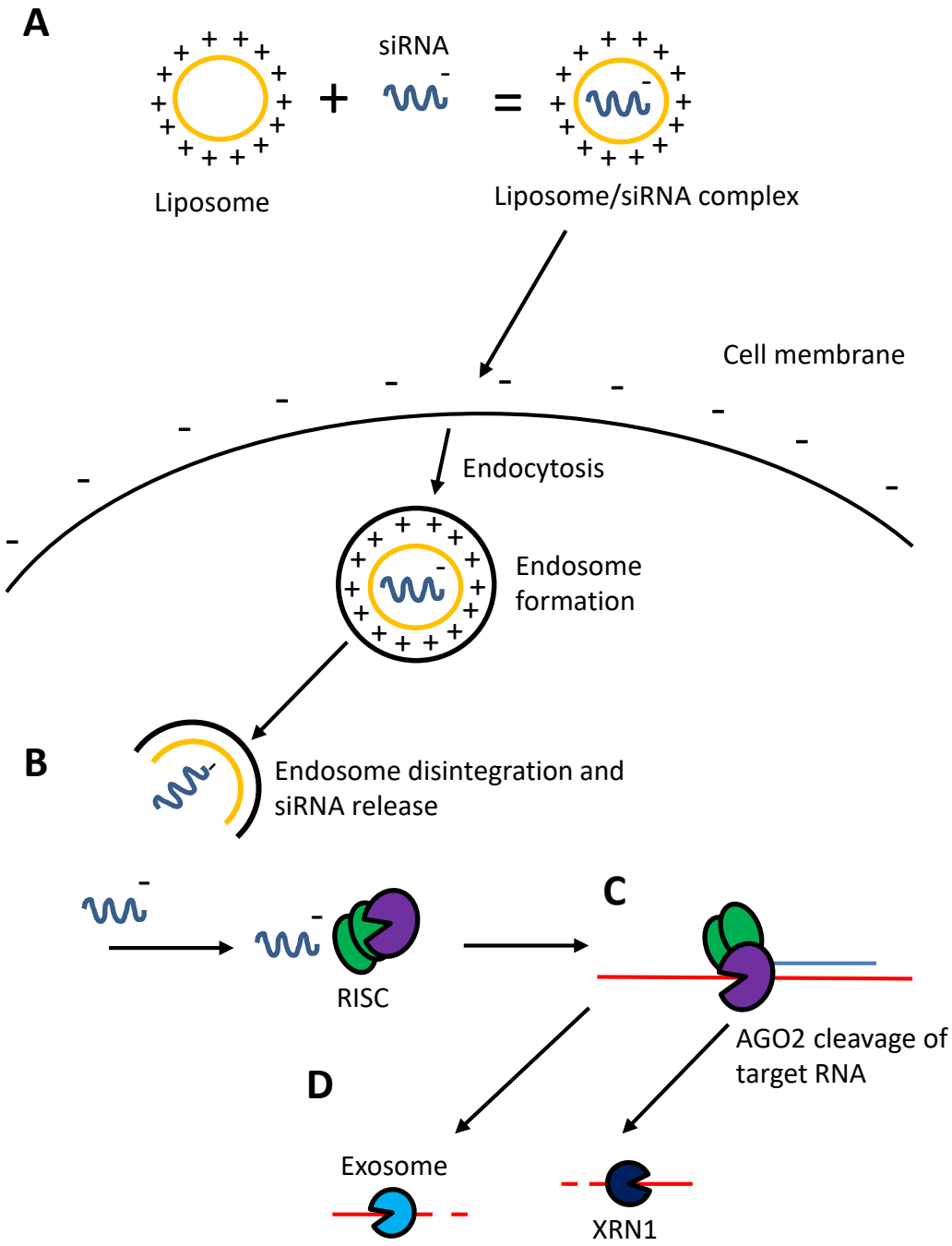


Figure 4.1. Mechanism of action for siRNA induced silencing using Lipofectamine RNAiMAX. A) Liposomes complex to siRNA, ensuring passage across the phospholipid bilayer into the cell. B) After endocytosis, the endosome:liposome:siRNA complex disintegrates and the siRNA complexes with the RISC. C) The complex is complementary to the target RNA, and the endonuclease, AGO2, endolytically cleaves the target strand via it's Piwi domain, leaving the RISC:siRNA to bind to the next complementary strand. D) mRNA fragments of the target strand are degraded by the exosome and XRN1.

4.3.1 XRN1 siRNA concentration optimisation

First of all, to find out at which concentration the siRNA worked most efficiently, several different concentrations were used (30pmol, 20pmol, 10pmol and 5pmol in 2mLs of medium). XRN1 mRNA and protein expression was evaluated and the efficiency of knock down compared at the different concentrations (Figure 4.2). Given the results of this titration experiment, consequent experiments were conducted using 20pmols of siRNA for transfections in 6-well plates, and 5pmols was used for transfections in 96-well plates. This was decided because the level of down regulation achieved with 20pmols of siRNA was similar to 30pmols, and much better than lower concentrations, and therefore it was used for future experiments.

4.3.2 XRN1 siRNA time point optimisation

Having demonstrated that 20pmols was the optimal concentration of siRNA to induce XRN1 knock down, a time course was performed to investigate how long XRN1 remains at lower levels post-transfection. The siRNA was removed from the culture after 24 hours and replaced with fresh medium. Time courses were conducted over a period of 144 hours, and this showed that XRN1 stays knocked down for up to 144 hours, with most knock down occurring at between 24 and 48 hours (Figure 4.3). Further western blot analysis showed that XRN1 protein is knocked down by a mean percentage of 81.8% after 24 hours exposure to siRNA (Figure 4.4).

After these rounds of optimisation, it was decided that 20pmols of XRN1 siRNA was sufficient to knock down XRN1 protein expression to a mean percentage of 18.2% of the original when compared to the scrambled siRNA control for up to 144 hours post transfection.

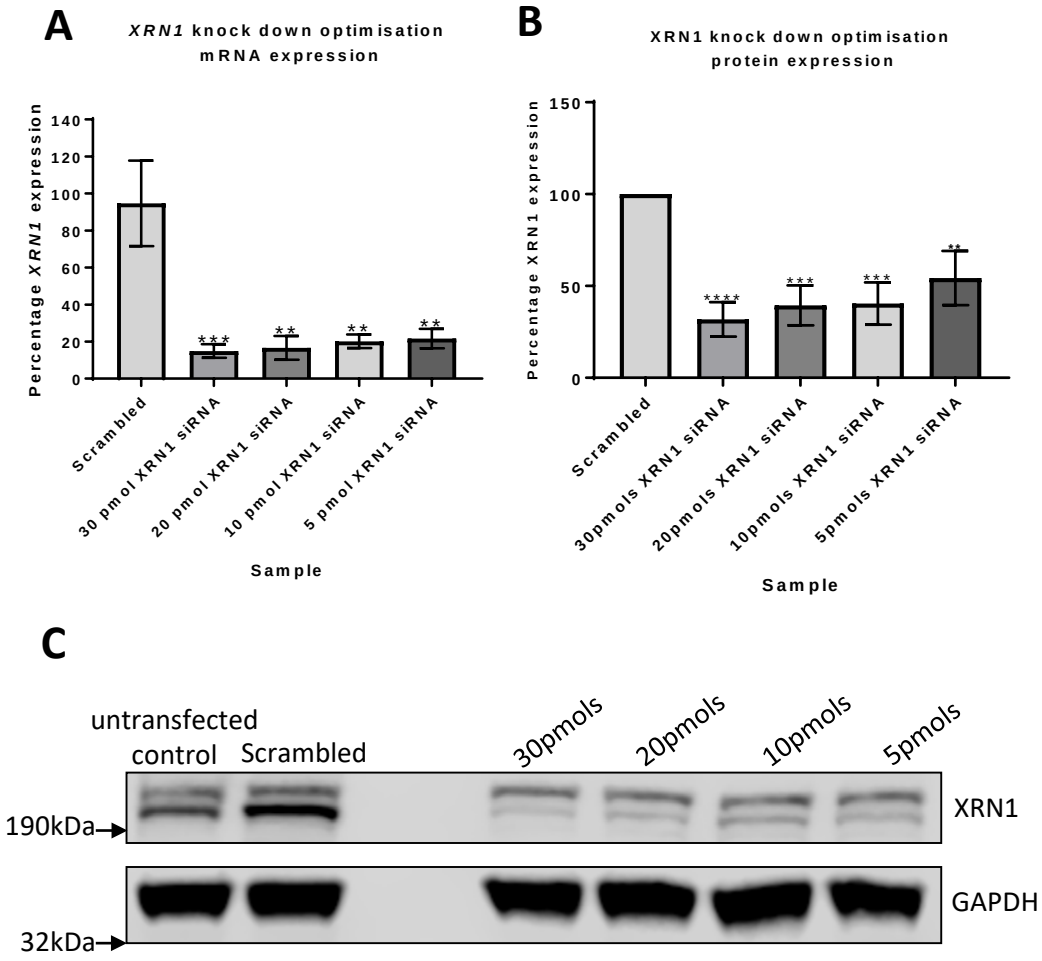


Figure 4.2. Optimisation of XRN1 siRNA concentration for knockdown in SAOS-2 cells. A) Bar graph to show quantification of *XRN1* mRNA to confirm knock down by qRT-PCR. Statistical analysis shown by unpaired t-test where: scrambled vs. 30pmols siRNA p=0.0007, scrambled vs. 20pmols siRNA p=0.0014, scrambled vs. 10pmols siRNA p=0.0017 and scrambled vs. 5pmols siRNA p=0.0031. Error bars = SEM. n=5. B) Bar graph quantification to show decreased XRN1 protein expression compared to the scrambled siRNA control at different concentrations, both bands were included in quantification as both decreased during knock down. Statistical analysis compared each siRNA titre to the scrambled control, as shown by unpaired t-test, where: scrambled vs. 30pmols siRNA p=<0.0001, scrambled vs. 20pmols siRNA p=0.0002, scrambled vs. 10pmols siRNA p=0.0003, and scrambled vs. 5pmols siRNA p=0.0075. Error bars = SEM. n=5. C) Gel image showing the confirmation of XRN1 knockdown using siRNA in a Lipofectamine RNAiMAX vector, in comparison to constant level of GAPDH.

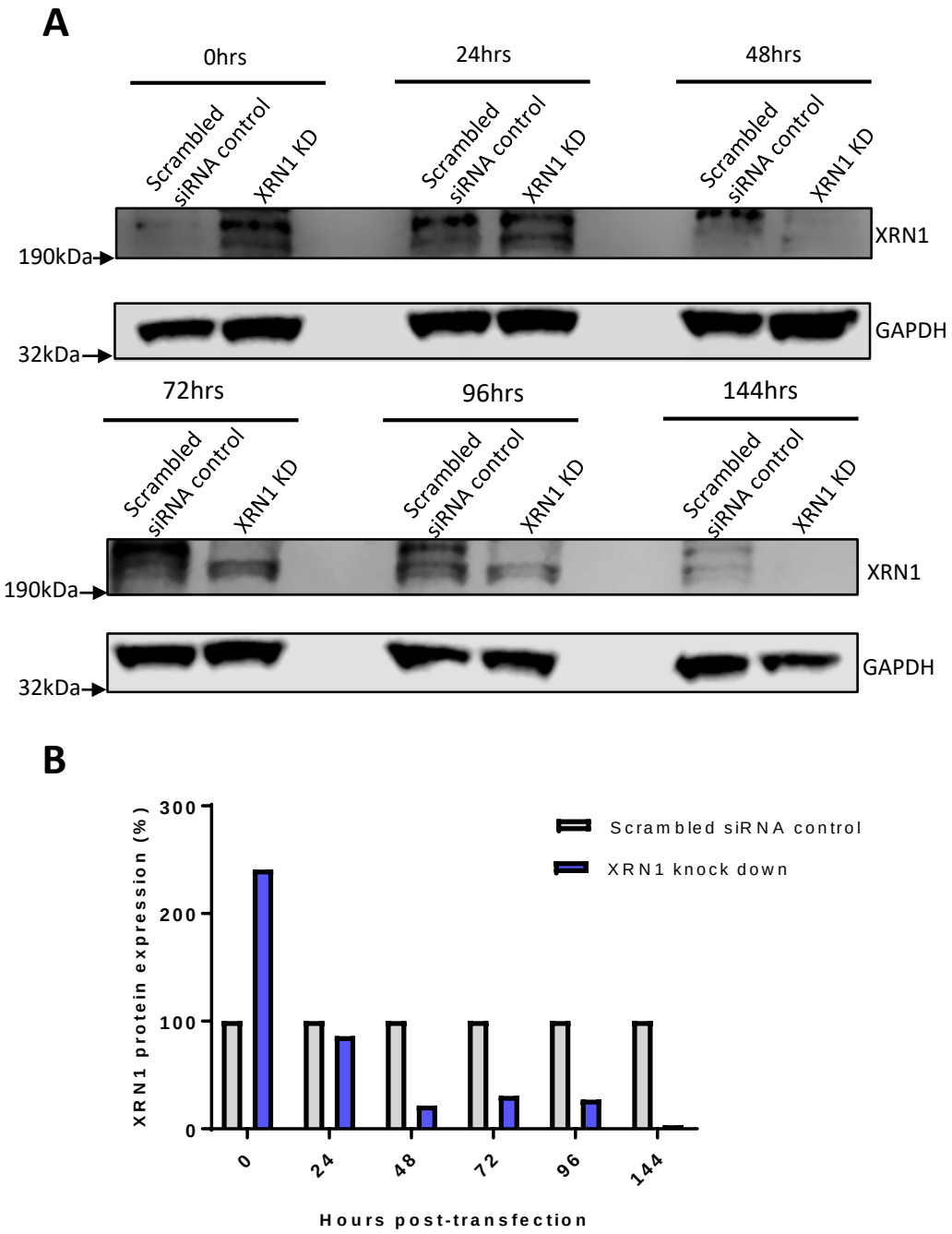


Figure 4.3. Length of transient siRNA knock down of XRN1. A) Gel visualisation of how long the transient transfection of XRN1 siRNA is effective in cell culture over a 7-day period. 20pmols of siRNA were utilised according to previous optimisation step (Figure 4.2). All samples loaded onto the same gel. B) Bar graph to show the quantification of XRN1 knock down over a 7-day period compared to the scrambled control at each time point, where XRN1 protein was measured using Li-Cor imaging software, normalised to GAPDH, and measured against the scrambled siRNA control. Cell pellets were taken at 0hrs post transfection and every 24hrs thereafter, n=2.

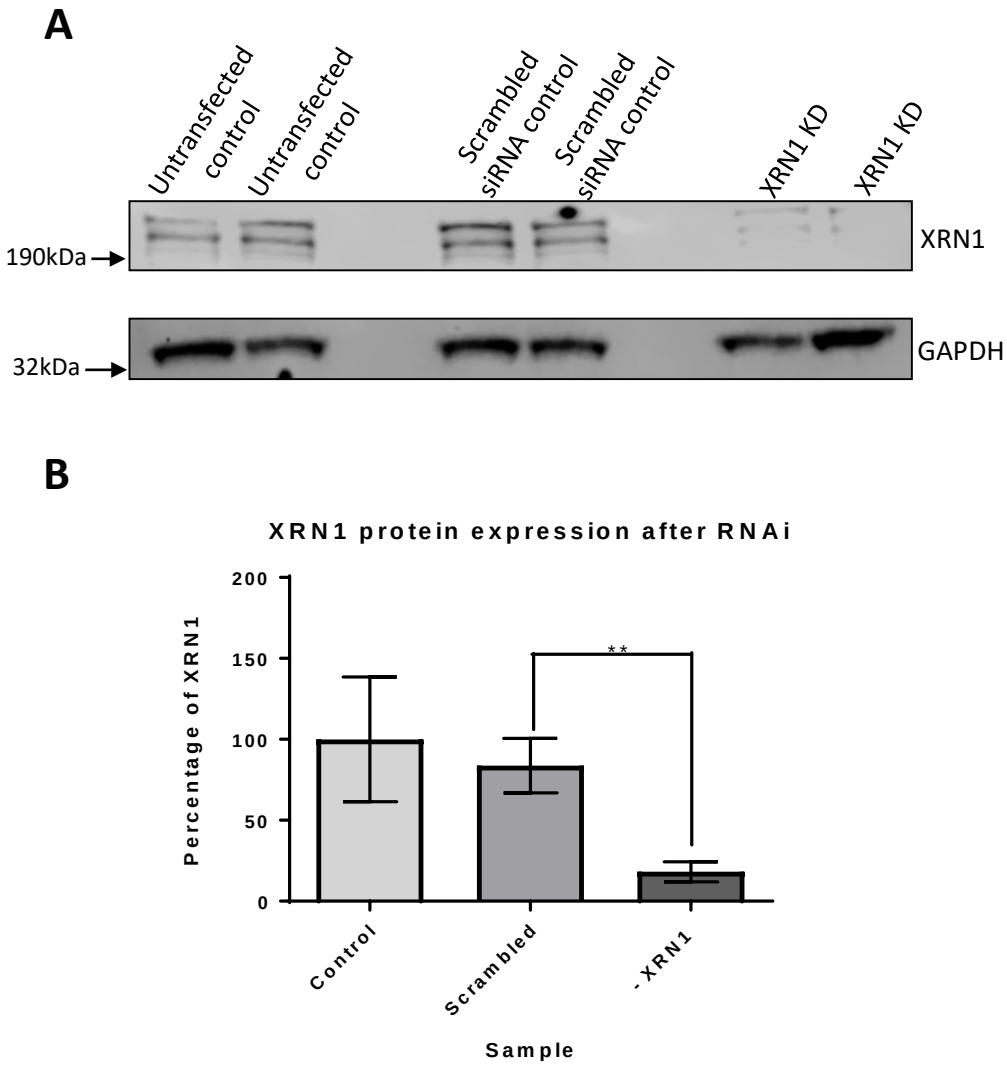


Figure 4.4. Confirmation of XRN1 knockdown in SAOS-2 cells using Lipofectamine RNAiMAX. A) Western blot gel image depicting visual down regulation of XRN1 in SAOS-2 cells, after 24hrs transfection, protein concentrations were not diluted to the same concentration due to low yield, protein quantification was normalised to GAPDH. Samples loaded in duplicate B) Graphical quantification of protein after XRN1 knockdown, showing an average of 81.8% protein knock down after 24hrs transfection, n=6. Error bars represent SEM. Statistical analysis performed by unpaired t-test where scrambled vs. XRN1 knockdown cells p=0.0044.

4.4 Effects of XRN1 knock down on cell line proliferation

Given that XRN1 has been shown to target pro-apoptotic transcripts in *Drosophila melanogaster*, it stands to reason that reduction in XRN1 expression would lead to higher rates of apoptosis. However, the major hallmark of cancer is the ability of cancer cells to proliferate uncontrollably, effectively overriding the apoptotic pathways. The fact that XRN1 expression is lower in three out of four OS cell lines shows a contradiction to this theory, whereby the most proliferative cell line, HOS, also displays the most significant decrease of XRN1. SAOS-2 showed no significant decrease of XRN1 and this cell line was also the least proliferative, with the longest doubling time. By knocking down XRN1 in SAOS-2 it was thought that this proliferative phenotype of the HOS cell line might be recapitulated in SAOS-2, thus marking this apparent contradiction as a novel function of XRN1 in human cells. If XRN1 knock down causes an increase in the rate of proliferation, this gene would become a novel candidate drug target.

Proliferation was initially measured by growth curves. These were performed using the optimal conditions of knock down as found in section 4.3, and cells were analysed over a 144 hour period. Cell counts were taken in duplicate every 24 hours (excluding the 120 hour time point). 0.1% trypan-blue stain was utilized to remove the chance of counting cells which had already undergone apoptosis or necrosis. The average of the duplicate cell counts was plotted for the growth curve (Figure 4.5). This was repeated three times.

Figure 4.5 shows that knock down of XRN1 seems to reduce the rate of proliferation of the SAOS-2 cell line, in contradiction to the original hypothesis that lower XRN1 expression confers a growth advantage to cancer cells. This could be because of stress conferred by the transfection technique. The growth curve shows that the RNAiMAX system of transfection may be affecting the growth of the cells, as transfection with the scrambled control siRNA seems to negatively affect cell line proliferation. It may be that due to the toxic nature of the RNAiMAX system XRN1 knock down may be more significantly detrimental to cell growth, however, due to the fact that the scrambled siRNA control is also detrimental, results show that XRN1 knock down is not affecting cell line growth when compared to the siRNA control. If XRN1 depletion is reducing the rate of growth, this would be similar to the phenotype seen in *Drosophila* during XRN1 depletion. The degree of variation within these samples is also noted, as it is quite large. Another possibility could be that there is compensatory regulation of RNA degradation in certain proliferation pathways by other cytoplasmic exoribonucleases, masking the effects that XRN1 knock down could be having. The presence of residual XRN1 (as this transfection technique is not 100% efficient) may also be sufficient to rescue any phenotype conferred by loss of XRN1.

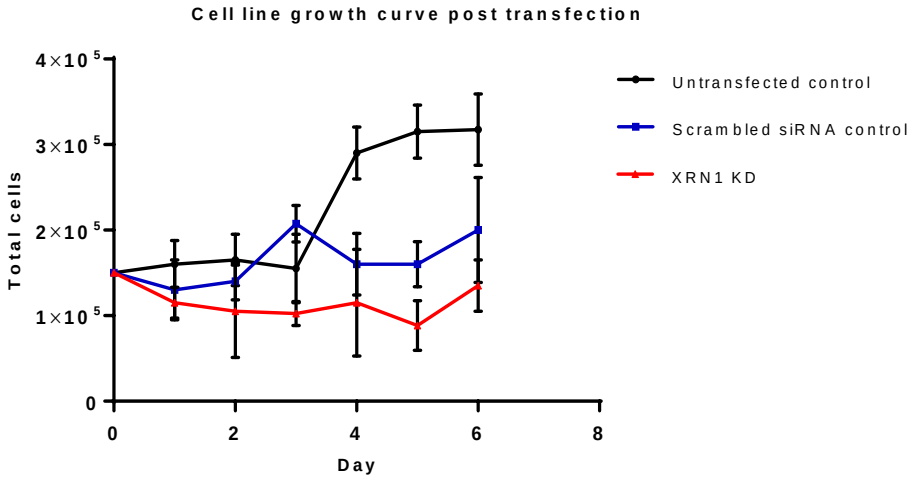


Figure 4.5. Growth curves showing the difference between growth with and without transfection with XRN1 siRNA. These curves show the differential growth of SAOS-2 cells over a 144hr period after 24hr transfection compared to non-transfected and scrambled siRNA controls. Counts were taken in duplicate every 24hrs and the averages plotted of cell counts in 1.5mL. Error bars represent SEM, based on n=3 biological replicates.

4.5 Assessment of SAOS-2 cell line proliferation by Brd-U staining

Due to the growth curves being variable, Brd-U staining was used to measure the rate of proliferation in a different way. Brd-U (5-Bromo-2'-deoxyuridine) staining is an immunocytochemistry method used to decipher the proportion of actively replicating cells by incorporation of Brd-U into replicating DNA. It is, first and foremost, a method used to measure the rate of DNA synthesis, however, this can be translated into a measure for proliferation. Brd-U is a synthetic nucleoside which is analogous to thymidine. It is incorporated into newly synthesised DNA of replicating cells during S- phase, and it can be detected by immunofluorescence upon application of α-Brd-U antibody (Wojtowicz and Kee 2006). The percentage of cells positively stained for Brd-U was calculated against the total number of cells indicated by DAPI staining, using ImageJ software with the DeadEasy_Mitoglia plugin to count cells fluorescing above a certain threshold (in this case it was set to 60 arbitrary units (AU)). 60AU was chosen because it allowed for the fluorescence of the stained cells to be measured without being affected by background fluorescence as observed when using higher thresholds. A limit was also set for cell size to minimise the plugin counting sporadic fluorescence. The protocol was optimised for use in SAOS-2 cells, where cell permeabilisation using 0.3% Triton X100 was adjusted to avoid cell structure disintegration. Original protocols also suggested using 0.6% Triton X100 to permeabilise cells, however, this caused widespread cell disintegration, and so 0.3% was used instead.

Results showed that, consistent with growth curves, there is no significant difference between the rate of proliferating cells between control cells and those which have had XRN1 knocked down (Figure 4.6). The Brd-U staining supports the evidence shown by the growth curves (Figure 4.5) that knocking down XRN1 does not affect the rate of proliferation, both by cell counting and by measuring the proportion of cells undergoing DNA replication. There was a slightly reduced rate of proliferation in both the scrambled siRNA control and XRN1 knock down cells, though this could have been an effect of the introduction of siRNA into the cells, leading to cell stress and reduced proliferation. It is also possible that there may be compensatory regulation being performed by other cytoplasmic exoribonucleases to offset RNA degradation not being performed due to the reduced expression of XRN1.

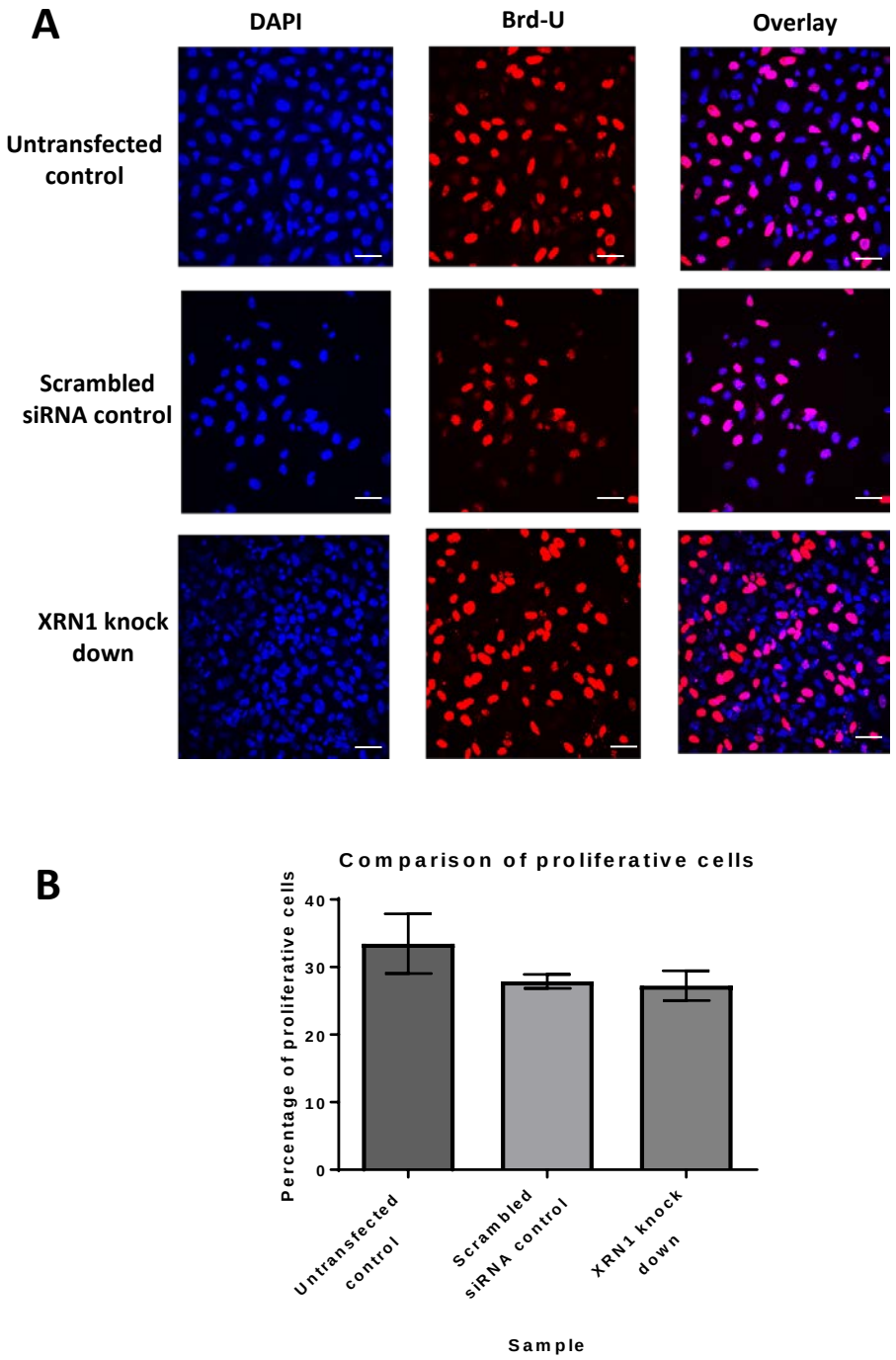


Figure 4.6. Brd-U incorporation experiment assessing the percentages of proliferating cells in SAOS-2 cells under 3 different conditions. A) Confocal microscopy images depicting cells in S-phase. Nuclei are stained with DAPI, proliferating cells shown by continuous Brd-U staining. Scale bar represents 50µM, images taken at x40 objective. B) Bar graph quantification of the percentage of proliferating cells against the total number of cells, comparing the percentage of proliferating cells at 24 hours post-transfection in XRN1 knock down samples to controls. Results show there is no significant difference in the percentage of proliferating cells after XRN1 is down regulated. Statistical analysis was performed by an unpaired t-test, based on n=3. Error bars represent SEM.

4.6 Effects of XRN1 knock down on the rate of apoptosis

After the assessment of the effect of XRN1 knock down on proliferation, an apoptosis assay was conducted, in order to investigate whether XRN1 knock down has an effect on the rate of apoptosis. Given that XRN1 is known to target pro-apoptotic transcripts, it would be reasonable to assume that the rate of apoptosis would increase in response to XRN1 knock down, despite this being in contrast to the rate of growth of the cell lines with a natural reduction in XRN1 expression.

In order to study the effects of XRN1 knock down on SAOS-2 cells, a Caspase-Glo 3/7 (Promega) assay was used. Caspase-Glo 3/7 is composed of aminoluciferin and a thermostable luciferase, optimised to detect caspase 3/7 activity upon application (Payne *et al.* 2013) (refer to Chapter 2, Figure 2.3). Adding the assay to the cells in a 96-well plate results in cell lysis, and caspase cleavage of the substrate, luciferin, generating free aminoluciferin, which is catalysed by luciferase. This results in the emission of a luminescent glow, detected by a plate reader. The luminescent signal detected is proportional to the rate of apoptosis. This assay was used in accordance with the manufacturer's instructions.

This assay was applied to SAOS-2 cells under the 3 experimental conditions (negative control, scrambled control and XRN1 knock down) according to the manufacturer's instructions, and showed that XRN1 knock down does not affect the rate of apoptosis (Figure 4.7). However, as seen in the growth curves and DNA synthesis assays, the transfection technique may be causing low-level apoptosis compared to the negative control. The rate of apoptosis in XRN1 knock down cells overlapped with the rate of apoptosis in the scrambled control, suggesting that introduction of a foreign nucleic acid causes low-level apoptosis. As expected, the untransfected control exhibited more apoptosis towards the end of the growth cycle, this could be because as cells become more confluent they begin to experience elevated levels of stress, competition for nutrients, and high density conditions.

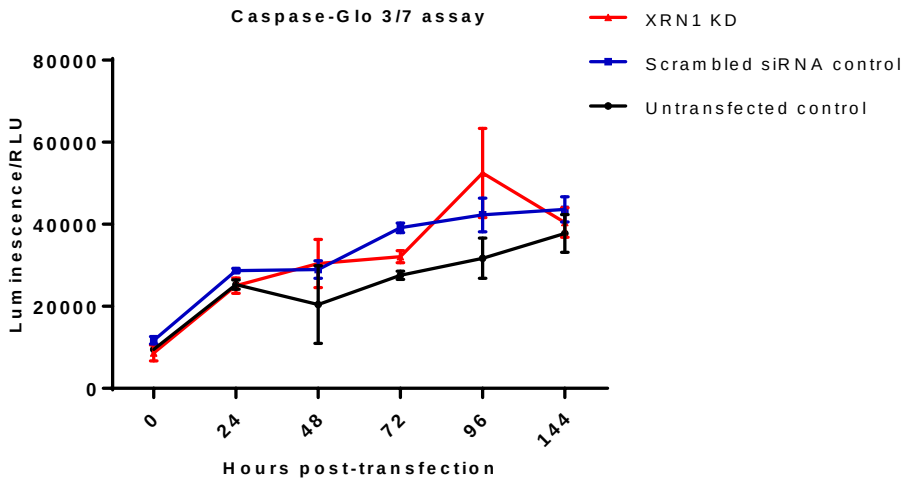


Figure 4.7. Knocking down XRN1 does not significantly affect the rate of apoptosis. Line graph represents the levels of luminescence after 1.5hrs of Caspase- Glo assay incubation (including 30s shake) at 24hr time intervals post- transfection with XRN1 siRNA compared to controls. Luminescence positively correlates with the amount of apoptotic cells, these results demonstrating that knocking down XRN1 does not significantly affect the rate of apoptosis, based on n=3 biological replicates, error bars represent SEM.

4.7 Effects of XRN1 knock down on cell line viability

Another major hallmark of cancer is the way cancer cells utilize cellular metabolism in response to stress, and indeed undergo metabolic reprogramming in order to promote cancer cell survival and growth (Boroughs and DeBerardinis 2015). This particular area of research has been extensively studied for many years, and altered metabolic features of cancer cell metabolism compared to healthy cells is considered a hallmark of cancer. This includes the infamous Warburg Effect, first identified over 90 years ago. The Warburg Effect describes how cancer cells rewire their metabolism to increase glucose uptake and subsequent fermentation of glucose to lactate, even in the presence of oxygen, to give them a growth advantage over non-tumour cells (Liberti and Locasale 2016). Differential cell metabolism in cancer pathologies serve to improve cellular fitness during stressful conditions (such as hypoxia), or allow cells to proliferate uncontrollably, and there are 3 major elements to the study of cancer cell metabolism (DeBerardinis and Chandel 2016):

1. Altered bioenergetics – to support anabolic growth during nutrient deprivation (Pavlova and Thompson 2016).
2. Enhanced biosynthesis – use of aerobic glycolysis to increase glycolytic rate and provide precursors to the anabolic pathways (Pavlova and Thompson 2016).
3. Redox balance – Positive regulation of cell proliferation and adaptation to metabolic stress (Finkel 2012).

Although knock down of XRN1 did not appear to affect the rate of proliferation or apoptosis, this enzyme is critical to cellular function, and it may be that the knock down of XRN1 could be advantageous to the ability of cancer cells to survive through the adaptation of metabolic pathways during stress. To observe whether the cancer cells survived better or worse after XRN1 knock down, cellular survival was measured using a WST-1 cell viability assay, which was performed in the same way as the Caspase-Glo 3/7 assay, to gain an insight into the possible effects of XRN1 knock down on cellular viability through measuring mitochondrial metabolic rate as a proxy for viability of mitochondria.

The WST-1 assay comprises of WST-1, a stable tetrazolium salt which is cleaved to soluble formazan by the succinate-tetrazolium reductase system in mitochondrial respiration (Yin *et al.* 2013) (refer to Chapter 2, Figure 2.2). This mechanism is reliant on the glycolytic production of NAD(P)H in viable cells, therefore, the amount of formazan produced directly correlates to the number of viable cells, as measured by absorbance. Maximum absorbance is measured at wavelength 450nm.

Figure 4.8 shows that knock down of XRN1 does not significantly affect cell viability when compared with the scrambled control, however, the general trend seems to again show the same effect: the introduction of siRNA onto cells induces toxicity, and so any additional toxicity caused by knockdown of XRN1 is not seen as significant. Multiple t-tests were performed on the data to compare growth between the multiple time points, of which none were significantly different. These data, taken together with previous data, suggests that loss of XRN1 in SAOS-2 cells has limited effects on cellular proliferation, viability or apoptosis. It also indicates that residual XRN1 expression, despite being at lower levels, is sufficient to rescue any phenotype that might be observed during complete knock out of XRN1 (such as lethality). This effect has been observed in other work in the Newbury lab, whereby knock down of DIS3 (to 11% of wildtype levels) did not result in an observable phenotype, however, complete knock out resulted in lethality.

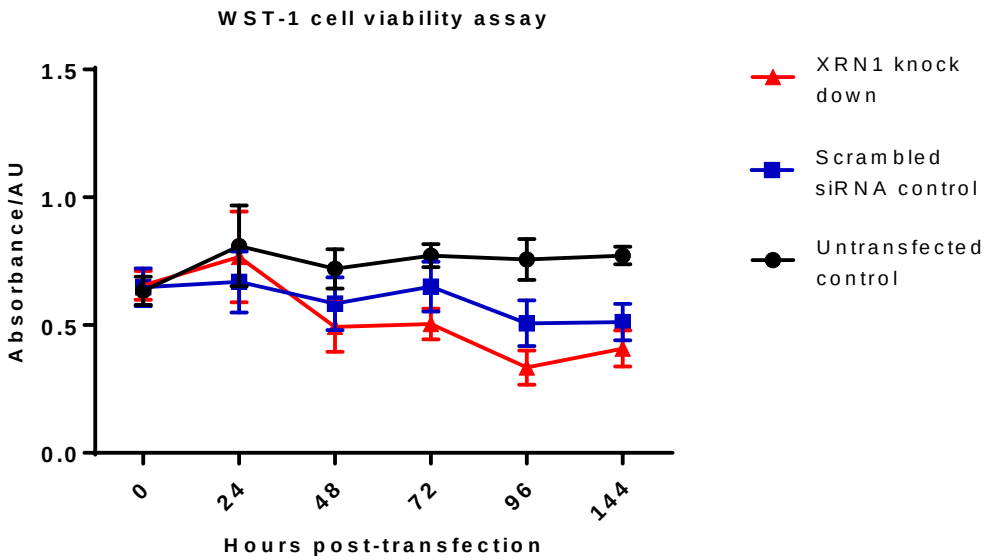


Figure 4.8. Knocking down XRN1 does not affect cell viability in SAOS-2 cells. Line graph represents the levels in absorbance after 1.5-2hrs of WST-1 assay application at 24hr time intervals post- transfection with XRN1 siRNA compared to controls. Absorbance positively correlates with the amount of viable cells, these results demonstrate that knocking down XRN1 negatively affects absorbance, suggesting that there are less metabolically active cells in these samples, however, this was not significant. Error bars represent SEM, based on n=3.

4.8 Effects of XRN1 knock down on the rate of translation

At the EMBO conference for Eukaryotic RNA Turnover (2017), unpublished data was presented showing that the homologue for XRN1 in *Arabidopsis thaliana*, XRN4, is involved in the degradation of RNA encoding the translational initiation factor eIF-4A. This protein is a DEAD-BOX helicase critical for correct functional translation. The significance for this is that XRN1 may function as a regulatory enzyme in translation, and by identifying a conserved mechanism for regulation of translation initiation between plants and humans there may be scope for the assumption that defects in this regulation could lead to pathologies such as cancer. It is well known that cancer cells produce excessive amounts of protein, and oncogenic signaling pathways often enhance translation initiation, through stimulation of the eIF-4F complex (Truitt and Ruggero 2016). Lowering expression of XRN1 could be one of the mechanisms used by cancer cells to promote excessive translation, in turn driving protein production to almost toxic levels, which could help to explain why cancer cells go through the cell cycle faster, thereby promoting proliferation. Given that recently published data showed that XRN1 degrades mRNA as it is released by the ribosome (Tesina *et al.* 2019), and therefore XRN1 acts co-translationally, it could be that reduction of XRN1 promotes increased translation, as there are more transcripts in the cell which are not being degraded.

To investigate this theory, SUnSET labelling was used to measure the amount of global translation occurring in XRN1 knock down cells, to see if translation was enhanced. This method utilizes the antibiotic puromycin to incorporate into the nascent peptide chain, which can then be detected by an anti-puromycin antibody during western blotting. Puromycin is an aminonucleoside antibiotic produced in *Streptomyces alboniger*, and is a structural analogue of the 3’ end of aminoacyl tRNA carrier for tyrosine and phenylalanine. Puromycin inserts into the peptide chain instead of the tRNA via the formation of a peptide bond. Unlike tRNA, puromycin has a non-hydrolysable amide bond, meaning that it cannot be removed from the chain. This causes truncation of elongation, and the release of the truncated peptide from the ribosome (Schmidt *et al.* 2009) (Figure 4.9 (Goodman and Hornberger 2013)). In usual circumstances, the application of puromycin would result in cell death (due to the blocking of translation), however, using it at low concentrations for limited periods of time means that translation is not fully blocked, and cells survive beyond the time of application. This results in a protein smear on a western blot, whereby peptides of different sizes are tagged with puromycin. Quantification of translation was measured by analyzing the area of the smear and normalizing it to the background fluorescence in ImageStudio. This value was in turn normalized to GAPDH for a relative percentage of puromycin incorporation.

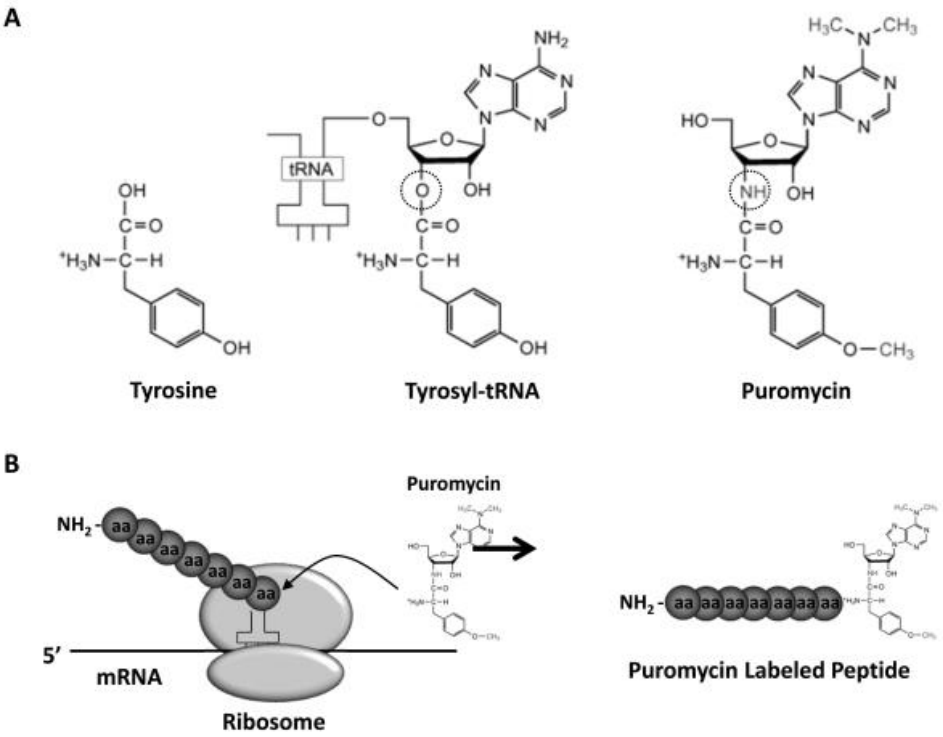


Figure 4.9. Puromycin mechanism of action. A) Puromycin molecular structure as compared to tyrosine and it's associated tyrosyl-tRNA. The circles represent the hydrolysable ester bond in the tyrosyl-tRNA, and the non-hydrolysable amide bond in puromycin, which causes puromycin to stay locked onto the elongating peptide chain. B) Mechanism of action of puromycin: when puromycin is incorporated into the elongating peptide chain by the formation of a peptide bond, the labelled peptide can no longer undergo elongation, and is released from the ribosome. This figure has been adapted from (Goodman and Hornberger 2013).

4.8.1 SUnSET labelling optimisation

Due to the toxicity of puromycin, it was important to optimize this method to prevent cell death during puromycin incorporation. Initial concentrations of 2.5, 5.0 and 10.0μg/mL were applied to cells over time points of 10 minutes, 20 minutes, and 45 minutes. Puromycin incorporation was then quantified, whereby it all detectable signal within the lane on the gel was measured and normalized to GAPDH expression. After this round of optimization, it was clear that puromycin needed to be applied for longer time points: 45 minutes, 1 hour, 1.5 hours and 2 hours. This showed that 2.5μg/mL of puromycin was a sufficient concentration of puromycin to be incorporated into peptide chains over 1.5 hours. A final optimization step was undertaken on a gradient gel (4-12% Bis-tris gel), to ensure measurable protein stacking, and to double check the parameters of the experiment (Figure 4.10).

4.8.2 The rate of translation is unchanged in OS cells

XRN1 was then knocked down in SAOS-2 cells, and translation was measured against the scrambled siRNA control. In this experiment, it was clear that the addition of a siRNA was not affecting the rate of translation in the same way that it affected proliferation, apoptosis and cell viability, which meant that any result observed could be attributed to knock down of XRN1 on its own. Figure 4.11 shows successful puromycin incorporation and XRN1 knock down in SAOS-2 cells. Figure 4.13 shows that in SAOS-2 cells, knock down of XRN1 does not affect the rate of translation when compared to the scrambled control. Samples were normalized to GAPDH expression, whereby GADPDH expression was analysed using the same protein lysate electrophoresed on a 7% Tris-acetate gel. After normalization to GAPDH, puromycin incorporation in each knock down sample was compared to puromycin incorporation in each parallel scrambled siRNA control sample, to ascertain whether there was any difference in the amount of protein translated between the two samples. XRN1 knock down confirmation was also performed across each individual replicate to ensure that knock down was successful.

Across the replicates, the level of translation was maintained at close to 100% when compared to the scrambled control. The general trend showed that translation may be slightly reduced, however, after repeating the experiment 5 times no significant change could be detected. This experiment was also conducted in U-2 OS cells (Figure 4.12), and the same result was observed (Figure 4.13). Interestingly, the protein smear seemed to be more prominent towards the bottom of the gel, perhaps indicating that these cells are producing higher quantities of smaller proteins than larger proteins. It is unclear whether this observation is indicative of these proteins being defective due to truncation by incorporation of puromycin

into the peptide chain, or if smaller proteins are generally higher in quantity than larger proteins in these cells. If the former is the case, small peptides would be over-represented compared to larger peptides using this method, because large peptides can be truncated as small, medium or large peptides, medium peptides can be truncated as only medium or small, and small peptides can only be truncated as small peptides.

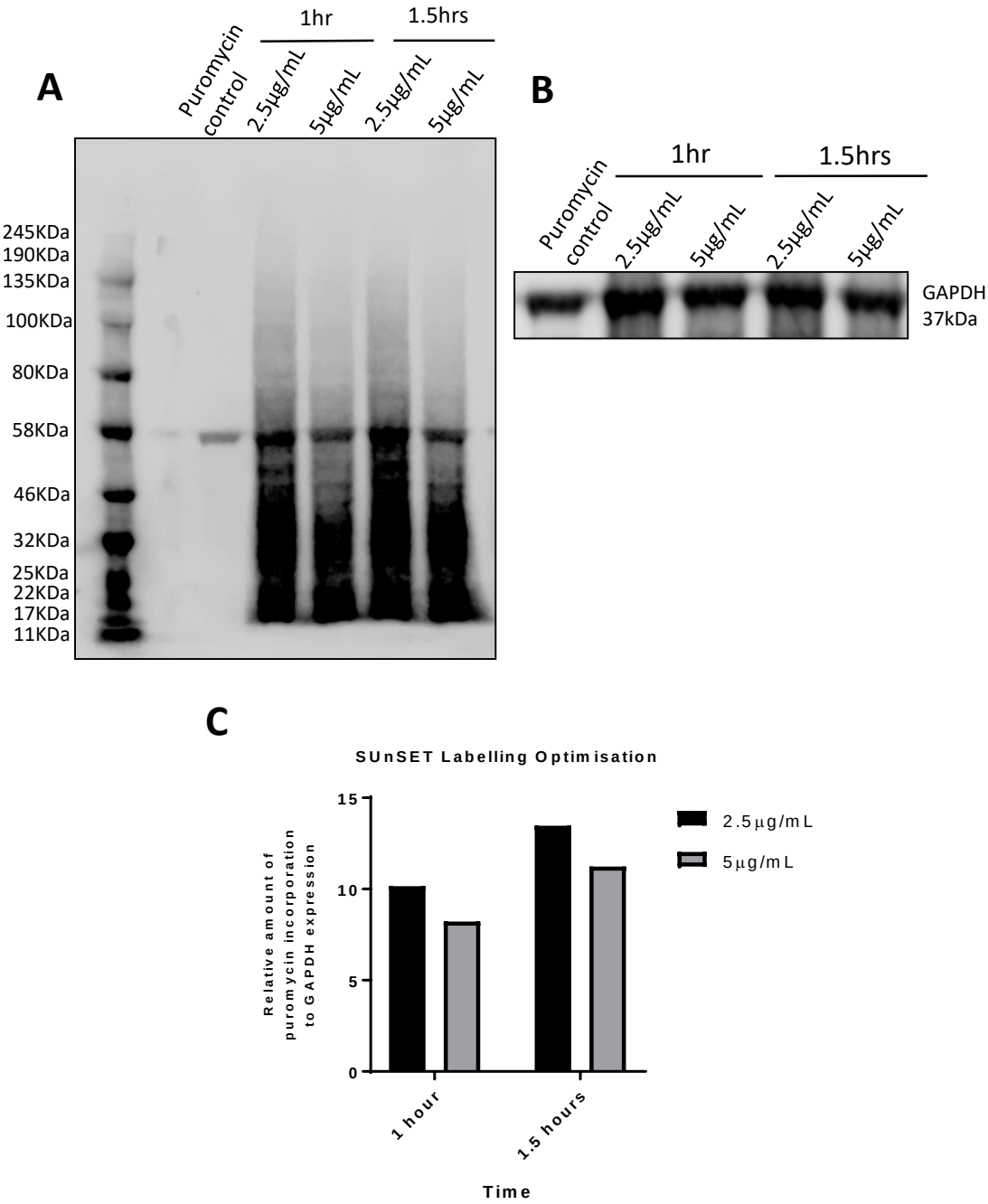


Figure 4.10. SUnSET labelling optimisation. A) Gel visualisation of puromycin incorporation in the fifth and final round of SUnSET labelling optimisation. Protein lysates were run on a 4-12% Bis-Tris gradient gel. Puromycin control sample represents puromycin incorporated into lysis buffer, which shows that it is not detected when there are no peptides present. B) Gel visualisation of GAPDH for normalisation. C) Bar graph quantification of the relative amount of puromycin incorporation after the application of 2.5µg/mL and 5µg/mL of puromycin at 2 different time points, respectively. Quantification was calculated by measuring the fluorescence of fluorescent secondary antibodies using Li-Cor Image Studio software. Incorporation was normalised to GAPDH expression.

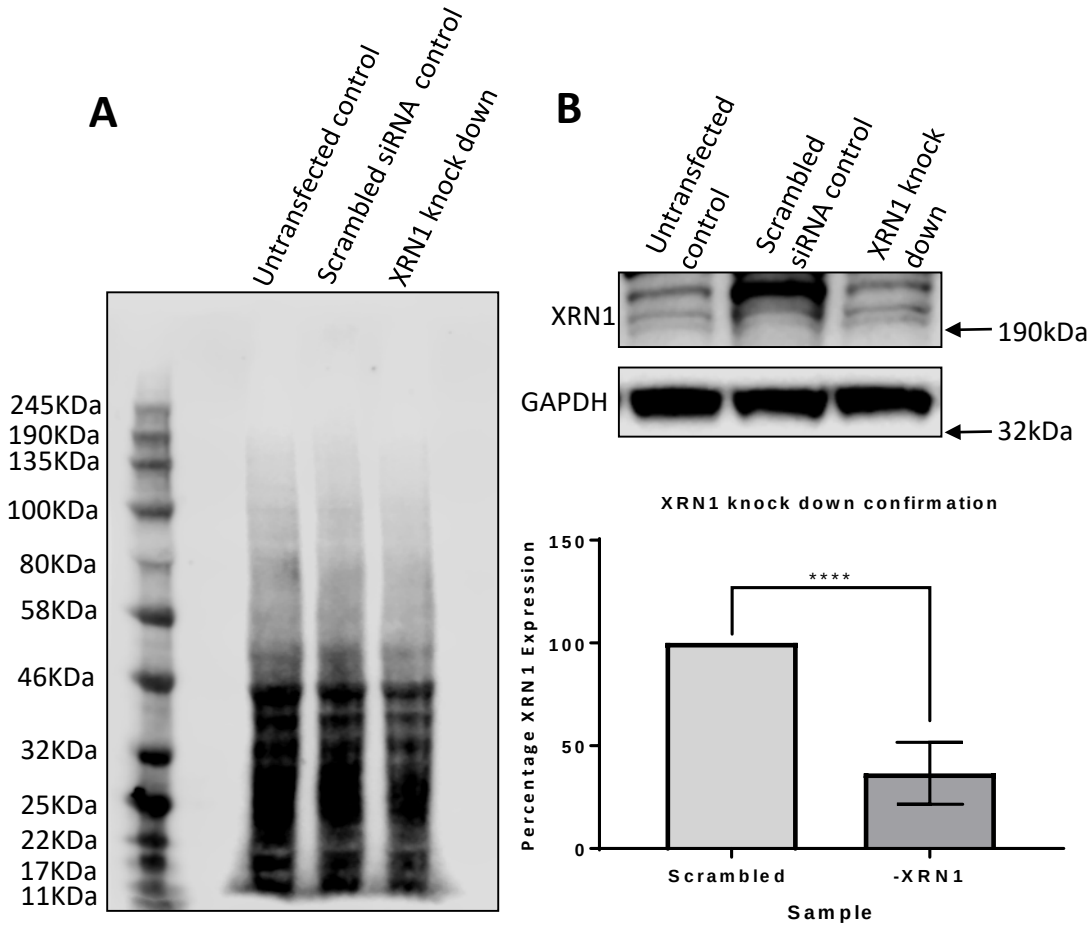


Figure 4.11. Results of SUnSET Labelling Experiment in SAOS-2 cells. A) Gel visualisation of puromycin incorporation into nascent peptide chains in an untransfected control (+ Lipofectamine), scrambled siRNA control and the XRN1 knock down sample, using the chosen concentration previously described. B) Gel visualisation of XRN1 knock down compared to the controls, relative to GAPDH, and bar graph quantification of XRN1 knock down, error bars = SEM. An unpaired t-test was performed where scrambled vs. –XRN1 p= <0.0001, based on n=6. Quantification of rate of translation shown in Figure 4.15.

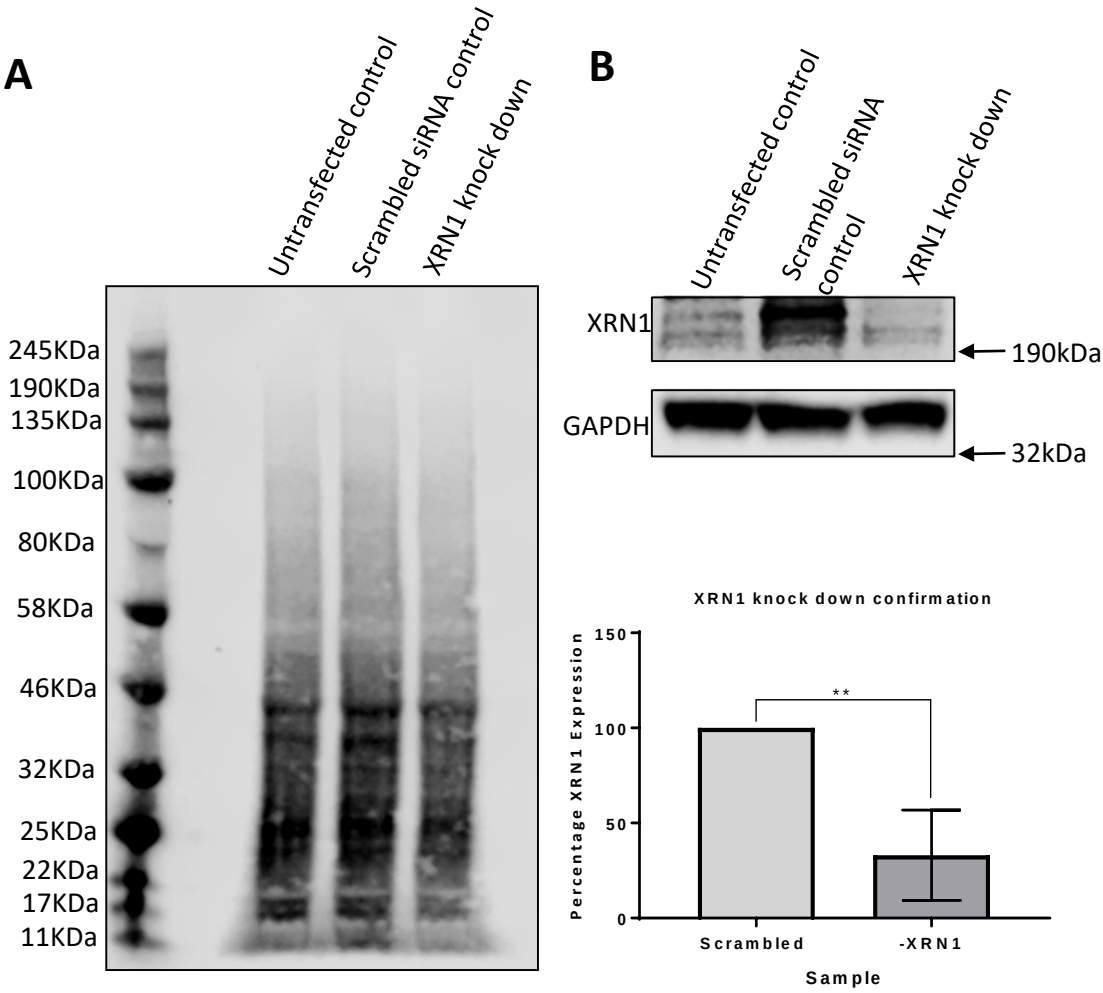


Figure 4.12. Results of SUnSET Labelling Experiment in U-2 OS cells. A) Gel visualisation of puromycin incorporation into nascent peptide chains in an untransfected control (+ Lipofectamine), scrambled siRNA control and the XRN1 knock down sample, using the chosen concentration previously described. B) Gel visualisation of XRN1 knock down compared to the controls, relative to GAPDH, and bar graph quantification of XRN1 knock down, error bars = SEM. An unpaired t-test was performed where scrambled vs. –XRN1 p= 0.0081, based on n=3. Quantification of rate of translation shown in Figure 4.15

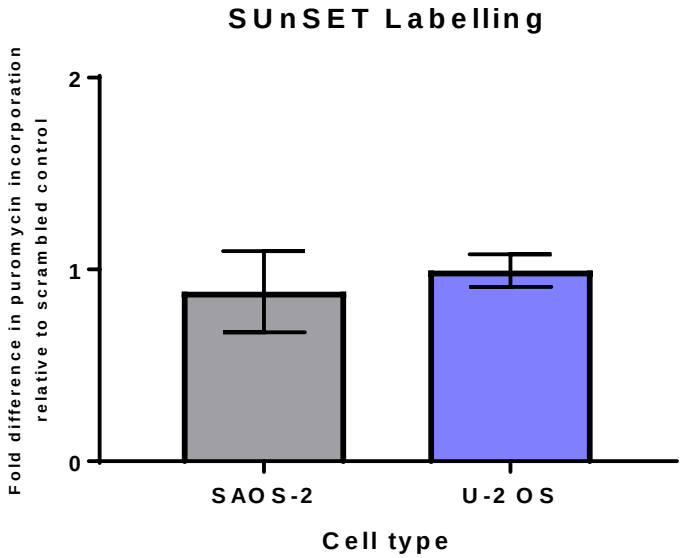


Figure 4.13. There is no relative change in translation after XRN1 knock down in SAOS-2 and U-2 OS cell lines. Bar graph quantification of puromycin incorporation as fold difference between the scrambled siRNA control and the XRN1 knock down samples in both SAOS-2 and U-2 OS cell lines, normalised to GAPDH expression. Error bars = SEM. Statistical analysis was performed by unpaired t-test where scrambled vs. XRN1 showed no significant difference in fold change of translation in both SAOS-2 (based on n=5) and U-2 OS (based on n=3) cell lines.

4.9 Optimisation of DIS3L2 knock down using siRNA

Following on from the optimization of the Lipofectamine transfection system for XRN1, the same was performed for DIS3L2 siRNA. Knock down of DIS3L2 was optimised in HEK-293T cells and performed in U-2 OS cells, because U-2 OS showed statistically significant up regulation of DIS3L2 and these cells had already proved to be highly transfectable during optimisation of XRN1 siRNA. HEK-293T cells were chosen for optimization because they are an embryonic kidney cell line, and DIS3L2 has previously been implicated in Wilms' Tumour of the kidney.

Optimisation tests for this siRNA included concentration optimization, where DIS3L2 protein expression was quantified by Western blot after cells were exposed to 30pmols, 20pmols, 10pmols and 5pmols of siRNA, and showed that 30pmols of siRNA produced the only statistically significant knock down of DIS3L2 (Figure 4.14). A time course showed that knock down of DIS3L2 became evident after 48 hours post exposure to siRNA (4.15A). The siRNA used was a pre-designed siRNA which targeted exons 5 and 6 of DIS3L2 at base pair 641. It silences both isoforms of DIS3L2.

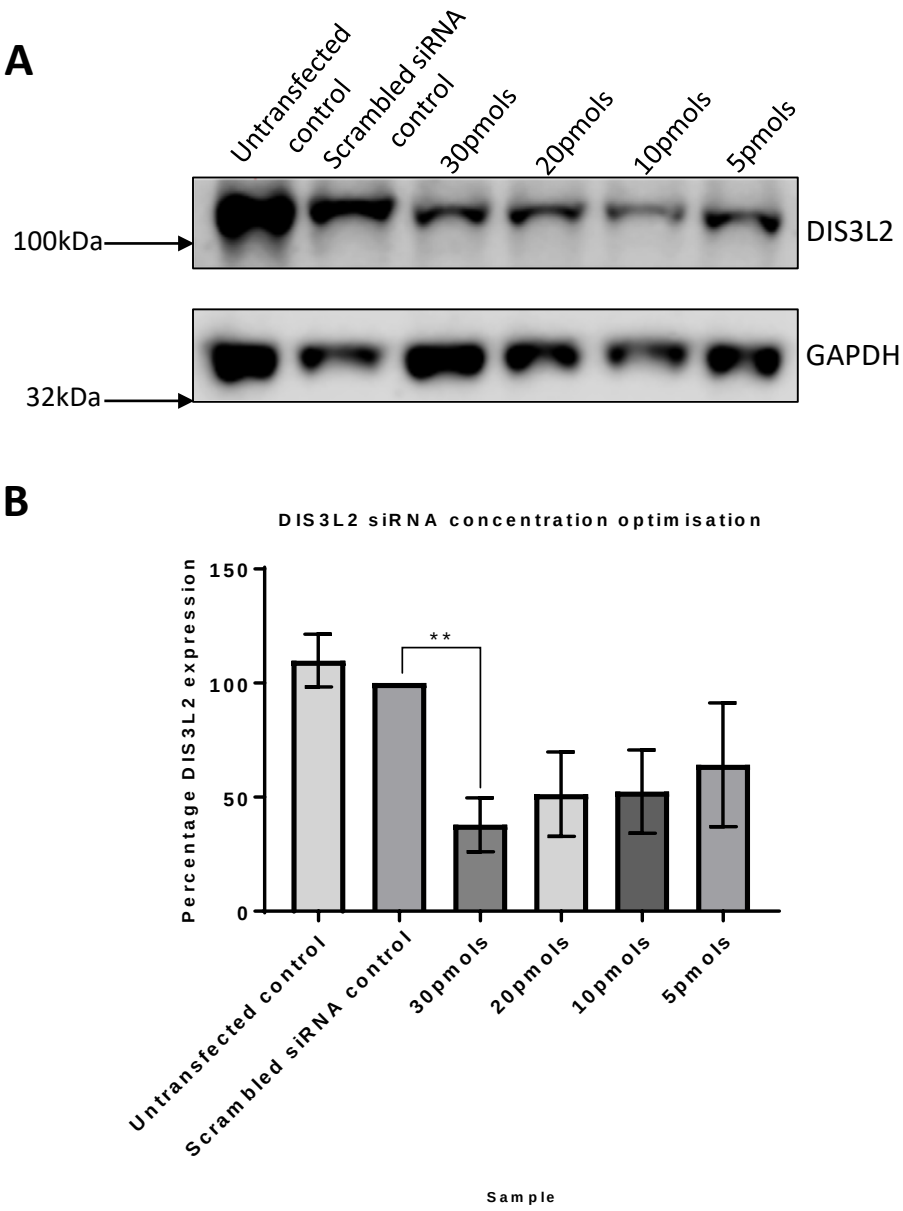


Figure 4.14. Optimisation of DIS3L2 siRNA concentration for knock down in HEK-293T cells. A) Gel visualisation of different siRNA concentrations compared to the untransfected and scrambled siRNA controls using GAPDH as a normaliser. B) Graphical quantification of DIS3L2 protein expression after 24hrs knock down with differential titres of DIS3L2 siRNA, relative to the scrambled control. Knock down performed for 48 hours. Error bars = SEM, based on n=3 biological replicates, statistical analysis performed by unpaired Student t-test, whereby scrambled vs. 30pmols siRNA p=0.0062, no other siRNA concentrations produced statistically significant differences in DIS3L2 protein expression.

4.10 Effects of DIS3L2 knock down on U-2 OS cell line proliferation

Both XRN1 and DIS3L2 are differentially expressed in OS cells (Chapter 3): XRN1 expression is lower and DIS3L2 expression is higher compared to expression in foetal osteoblast cells. To identify a potential role for DIS3L2, the Lipofectamine-based siRNA transfection system was utilized to knock down DIS3L2 in both U-2 OS cells and HEK-293T cells.

DIS3L2 was knocked down in HEK-293T cells in order to test whether the effect of DIS3L2 knock down was the same as in *Drosophila melanogaster* tissues, whereby down regulation of DIS3L2 causes proliferation of developing tissues, specifically in the wing imaginal discs, as part of another study being undertaken in the lab. Previous studies have also implicated DIS3L2 in the progression of various kidney disorders, and HEK-293T cells are a human embryonic kidney line (Astuti *et al.* 2012). A growth curve was performed, in which cells were exposed to DIS3L2 siRNA for 24 hours, medium was removed and cells were cultured as normal for 144 hours. Cell counts were taken in triplicate every 24 hours from 0 hours post-transfection in 6-well plates. Results showed that DIS3L2 knock down in HEK-293T cells causes an increase in proliferation in HEK-293T cells (Figure 4.15) when compared to the scrambled siRNA control. Knock down was confirmed by Western blot (Figure 4.15A) and all protein detected was normalized to the housekeeper gene, GAPDH.

Growth curve studies of DIS3L2 knock down in U-2 OS cells were performed in exactly the same way, and successful knock down was observed between 72 and 144 hours post transfection (Figure 4.16A). These studies showed that, unlike in HEK-293T cells, knock down of DIS3L2 did not affect proliferation of this cell line compared to the scrambled siRNA control (Figure 4.16B). This indicates a tissue-specific function for DIS3L2, which is not maintained in U-2 OS cells. This could also indicate that there may be a difference in DIS3L2 activity in embryonic cells versus adolescent, or in the cell line, where cancer cells respond differently to those virally transformed. These cells also responded to the application of scrambled siRNA better than the other cell lines, possibly due to the remarkable capabilities of cancer cells to overcome induced stress.

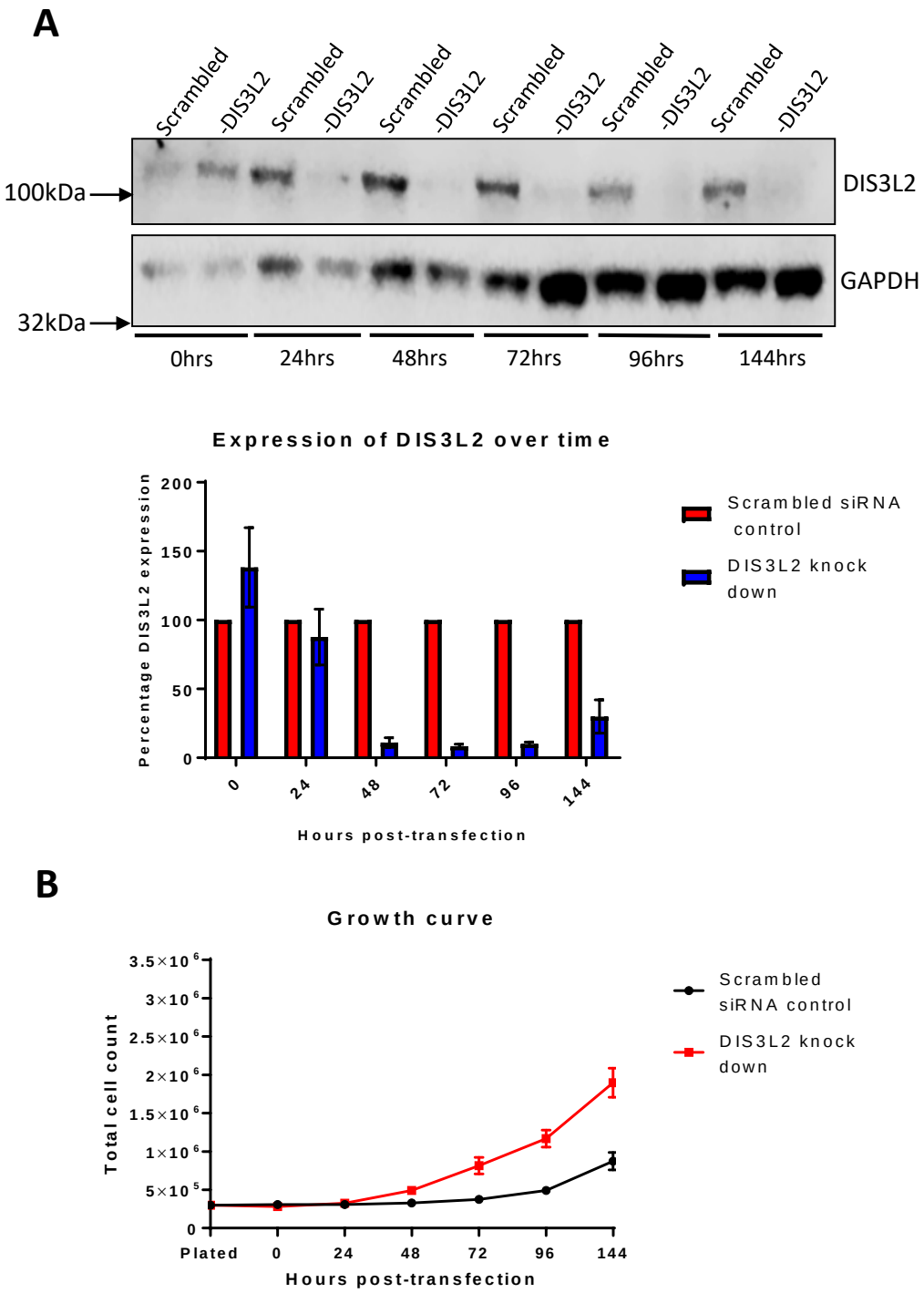


Figure 4.15. Effect of DIS3L2 knock down on growth of HEK-293T cells. A) Gel visualization and graphical quantification of knock down of DIS3L2 protein over the time period of the growth curve. Error bars = SEM, based on n=4. B) Growth curve of HEK-293T cells after DIS3L2 knockdown, over the period of 144hrs. Error bars = SEM, based on n=4.

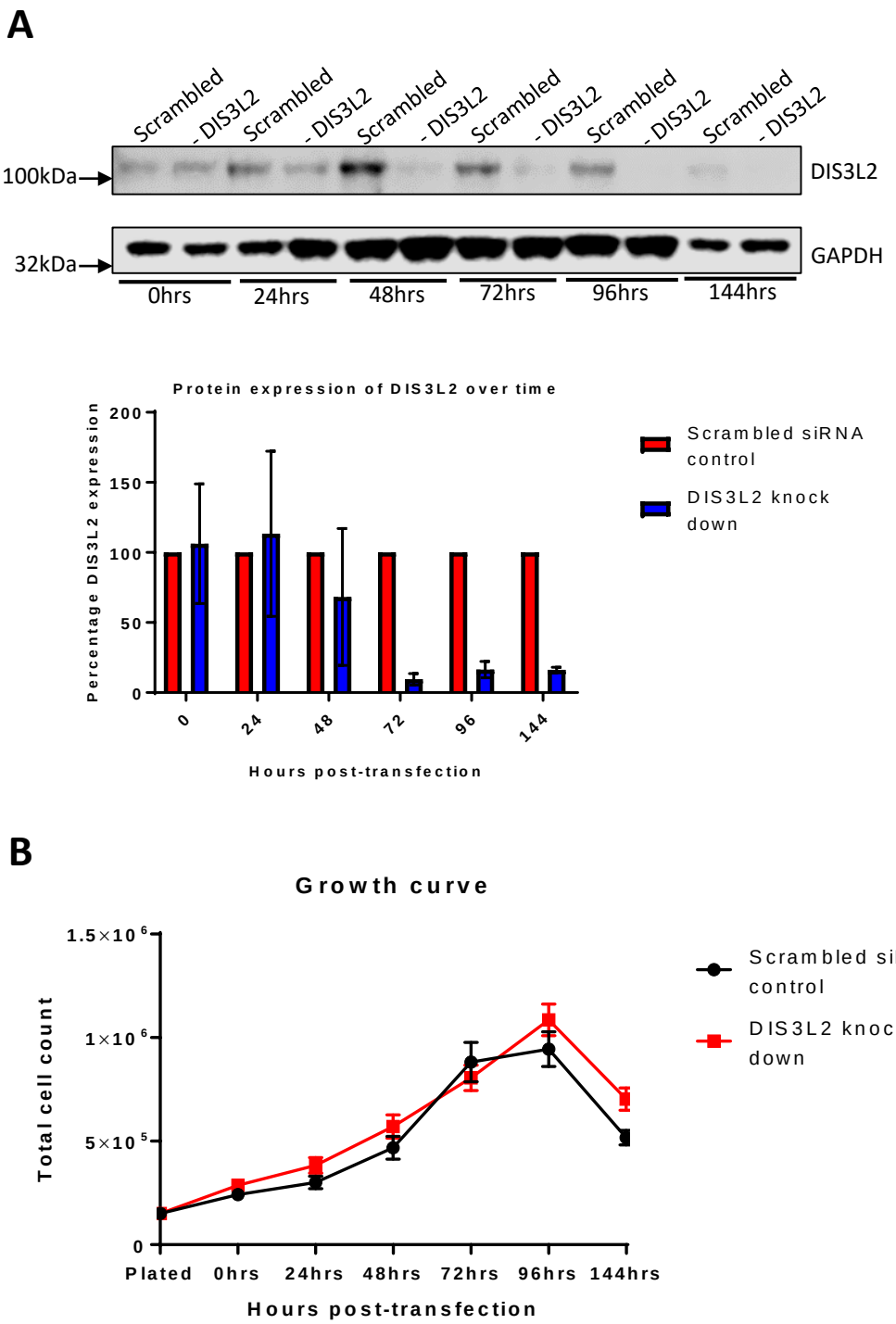


Figure 4.16. Effect of DIS3L2 knock down on growth of U-2 OS cells. A) Gel visualization and graphical quantification of knock down of DIS3L2 protein over the time period of the growth curve. Error bars = SEM, based on n=4. B) Growth curve of U-2 OS cells after DIS3L2 knockdown, over the period of 144hrs. Error bars = SEM, based on n=4.

4.11 Knocking down DIS3L1 and XRN1 shows no link between the activities of the 2 exoribonucleases

When the expression of each of the exoribonucleases was analysed (Chapter 3), it was clear that *DIS3L1* mRNA expression displayed significant up regulation in OS cells when compared to the HOb control cell line. Protein expression of DIS3L1 was undetectable in OS cells, possibly because of the ultra-low level of expression of DIS3L1 in bone tissue (Human Protein Atlas, Figure 3.14B). Based on the idea of interplay between the exoribonucleases, and if XRN1 and DIS3L1 are responsible for regulating each other, it may be that DIS3L1 becomes more up regulated when XRN1 is more down regulated. If this is the case, it could be expected that upon XRN1 knock down in cells with a natural reduction of XRN1, DIS3L1 could be up regulated to the point that protein detection is possible. In addition, this chapter has shown that knock down of XRN1 in SAOS-2 cells does not result in an observable phenotype. This could be because of compensatory regulation of target mRNAs being performed by other exoribonucleases in the cell, which could be differentially expressed in response to lower expression of XRN1.

With this in mind, it was decided to knock down XRN1 and DIS3L1 separately and to observe the effect the knock down of one has on the expression of the other exoribonuclease. Knock down was performed in SAOS-2 cells with siRNAs in the same way as XRN1 knock down studies were initially performed, to keep experiments consistent. A scrambled siRNA control sample was created for each knock down and results were also compared to a wildtype control.

Successful knock down was confirmed by qRT-PCR in each experiment (Figure 4.17A&B) and the mRNA expression of each exoribonuclease was analysed by qRT-PCR to observe the expression each of the exoribonucleases in response to the other being knocked down (Figure 4.17C&D). XRN1 protein levels were also assessed in DIS3L1 knock down cells (Figure 4.18). The expression of other exoribonucleases, DIS3 and DIS3L2, was also analysed in response to double knock down of XRN1 and DIS3L1, under the assumption that they may be conferring compensatory expression and in that case would be up regulated during artificial knock down of XRN1 and DIS3L1 (Figure 4.19).

Results showed that knocking down XRN1 has no significant effect on the mRNA expression of *DIS3L1* (Figure 4.17C), and that knocking down DIS3L1 has no effect on the mRNA expression of *XRN1* (Fig 4.17D). These results suggest that XRN1 and DIS3L1 are not regulating each other, and the cell does not appear to up regulate one gene in response to the down regulation of the other gene. Given that DIS3L1 is naturally expressed at very low levels in this type of tissue, it is not surprising that small changes in expression of this particular exoribonuclease might not detectable, and will also help to explain some of the variability in

these results. XRN1, on the other hand, is readily detectable in this tissue, and so expression changes in XRN1 would be detectable if down regulation of DIS3L1 was having an effect on XRN1. Protein expression of XRN1 in response to DIS3L1 knockdown shows a comparative up regulation of XRN1 protein (Figure 4.18), however, the variability of the results after 5 repeats show that the result is not statistically significant. Together, the results suggest that XRN1 and DIS3L1, both cytoplasmic exoribonucleases, do not function synergistically with each other, nor is one solely responsible for targeting the other for degradation.

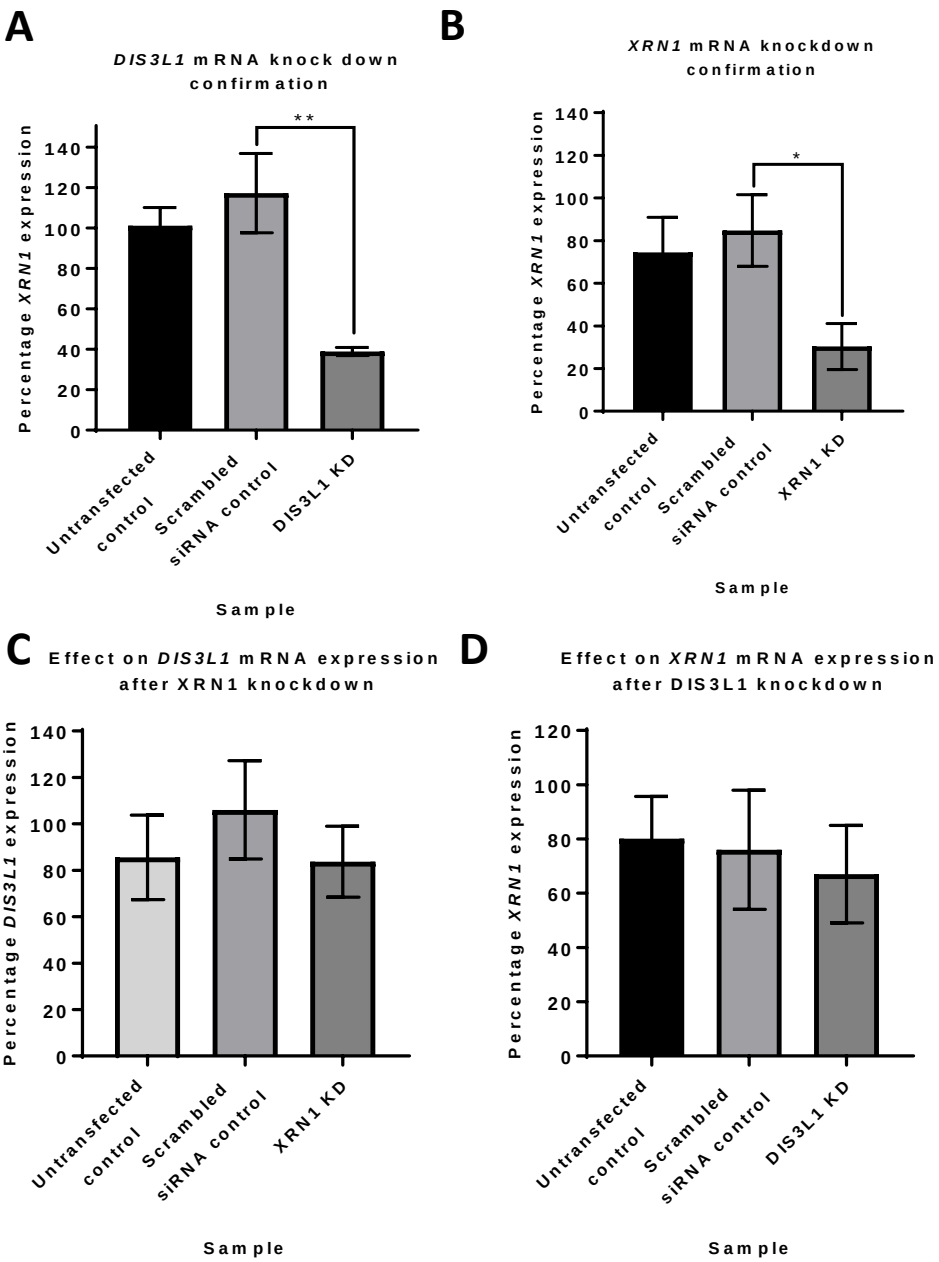


Figure 4.17. Effect of XRN1 and DIS3L1 knock down on the expression of each other in SAOS-2 cells. A) Confirmation of XRN1 knock down using qRT-PCR, error bars = SEM, n=4. Statistics performed by unpaired t-test where scrambled vs. XRN1 KD p=0.0350. B) Confirmation of DIS3L1 knock down using qRT-PCR, error bars represent SEM, n=4. Statistics performed by unpaired t-test, where scrambled vs. XRN1 KD p=0.0073. C) Quantification of *DIS3L1* mRNA expression after XRN1 knock down, no significant changes in expression were observed against the scrambled controls using unpaired t-tests. Error bars = SEM, n=4 D) Quantification of *XRN1* mRNA expression after knock down of DIS3L1, no significant changes in expression were observed against the scrambled controls using unpaired t-tests. Error bars = SEM, n=3. All expression data was normalised to *GAPDH*.

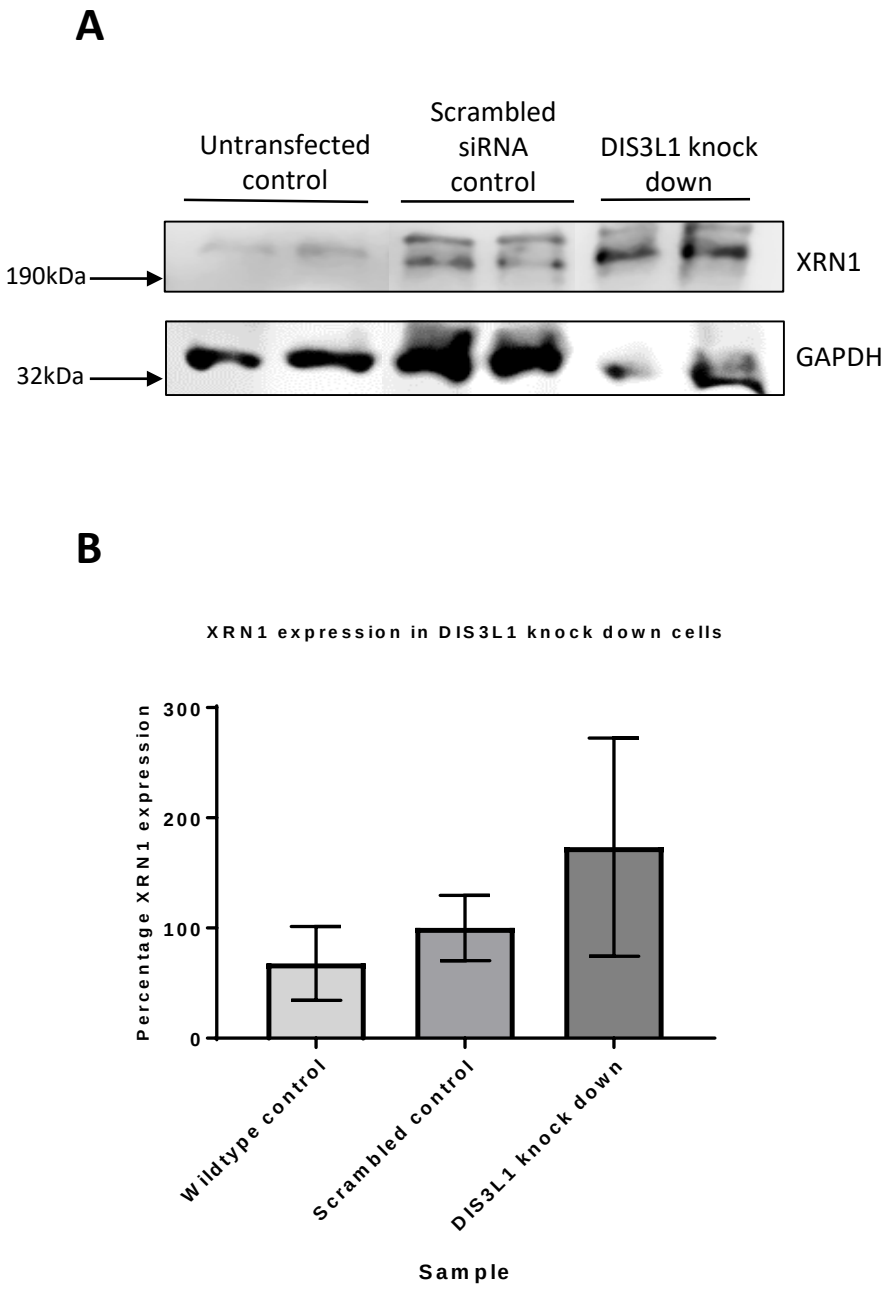


Figure 4.18. XRN1 protein expression in response to knock down of DIS3L1 in SAOS-2 cells. A) Gel visualisation of XRN1 protein expression after 48hrs of DIS3L1 knock down, loaded in duplicate. B) Quantification of XRN1 protein expression after knock down of DIS3L1 was measured by normalizing to GAPDH. Statistical analysis was performed by unpaired t-test, where scrambled control vs. DIS3L1 knock down samples were insignificant, based on n=5. Error bars = SEM.

4.12 Knocking down DIS3L1 and XRN1 results in differential expression of DIS3 and DIS3L2

4.12.1 Expression of DIS3 after DIS3L1 and XRN1 knock down

Expression of *DIS3* mRNA was measured alongside *DIS3L2* mRNA in response to XRN1 and DIS3L1 knock down. These results showed that *DIS3* mRNA is significantly 2-fold down regulated in response to XRN1 knock down, but RNA expression is not affected by DIS3L1 knock down (Figure 4.19A). This suggests that DIS3 may be a target for degradation by another factor in the cell, and XRN1, but not DIS3L1, is responsible for regulating this particular factor, which is not possible after XRN1 knock down. It could also be the result of a knock on effect resulting in decreased transcription of DIS3. It also suggests a previously unknown way that the exoribonucleases might be regulating each other and how different exoribonucleases have specific targets in human cells, depending on the main function of the cell within the specific tissue. Due to the lack of specificity of the DIS3 antibody (as seen in Chapter 3, Figure 3.11), protein levels for DIS3 could not be determined.

What could be interesting is to look at the transcripts that are targeted by DIS3, to find out whether there is a possibility that DIS3 becomes differentially regulated due to the expression of its specific targets. For example, do cells down regulate certain regulatory proteins of the RNA degradation pathway in response to an increased need for the transcripts which are normally targeted? If so, can members of the pathway be regulated by other members of the pathway to ensure this happens? It is known that dysregulation of DIS3 promotes global accumulation of Promoter Upstream Transcripts (PROMPTs) and enhancer RNAs (eRNAs), both of which have regulatory roles in the cell (Szczepińska *et al.* 2015). Indeed, DIS3 targeting eRNAs would change transcriptional regulation, alongside the argument for eRNAs being a product of 'leaky' transcription, another reason for being targeted for degradation by DIS3, which is the catalytic member of the nuclear exosome.

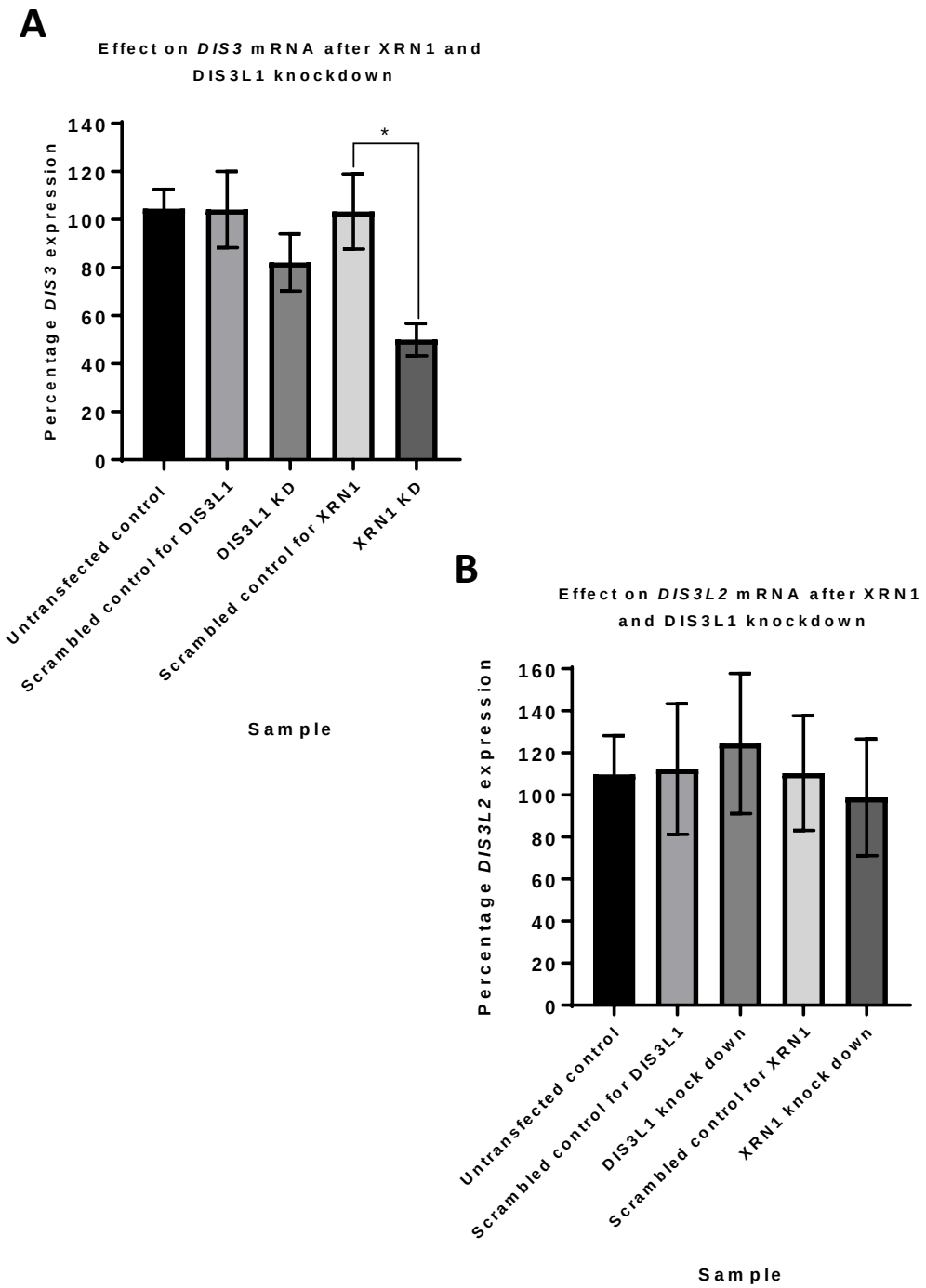


Figure 4.19. Effect of knock down of XRN1 and DIS3L1 on expression of *DIS3* and *DIS3L2* in SAOS-2 cells as shown by qRT-PCR. A) Quantification of *DIS3* mRNA after DIS3L1 and XRN1 knock down respectively, expression was normalised to *GAPDH*, unpaired t-test was performed where: no significant changes seen during DIS3L1 knock down, and scrambled vs. –XRN1 p=0.0120. Error bars= SEM, n=4. B) Quantification of *DIS3L2* mRNA expression after respective DIS3L1 and XRN1 knock down. Expression was normalised to *GAPDH* where no significant changes in expression were observed against the scrambled controls using unpaired t-tests. Error bars = SEM, n=4.

4.12.2 Expression of DIS3L2 after DIS3L1 and XRN1 knock down

DIS3L2 mRNA expression, on the other hand, was not affected significantly by knock down of XRN1 or DIS3L1, however, protein expression of DIS3L2 was somewhat up regulated in knock down samples (both DIS3L1 and XRN1) when compared to the controls (Figure 4.20), however, this up regulation was not significant. The fact that *DIS3L2* mRNA expression is not affected by knock down of either of these enzymes is somewhat surprising given that protein expression of DIS3L2 was higher in OS cells in the same way that XRN1 is lower, however, this is another example of the non-correlative cellular mRNA to protein ratio.

The trend of up regulation of DIS3L2 protein in response to separate knock downs of DIS3L1 and XRN1 suggests the existence of potential synergism between DIS3L2 and both XRN1 and DIS3L1, supporting the notion of compensatory expression of these enzymes in the cell, however, more clarification is needed to confirm the extent of up regulation of DIS3L2. Compensatory expression between these enzymes would not be wholly surprising, given that they are all present in the cytoplasm and performing cytoplasmic RNA degradation (as opposed to DIS3 in the nucleus), however, it would be an unusual phenomenon in the sense that DIS3L2 and XRN1 have been shown to target transcripts in opposite pathways (growth and apoptosis). Next steps to take with regard to compensatory expression of the exoribonucleases would be to investigate whether the transcripts being targeted by the exoribonuclease which is down regulated become targets of the exoribonucleases which are up regulated, in order to maintain the same cellular homeostasis.

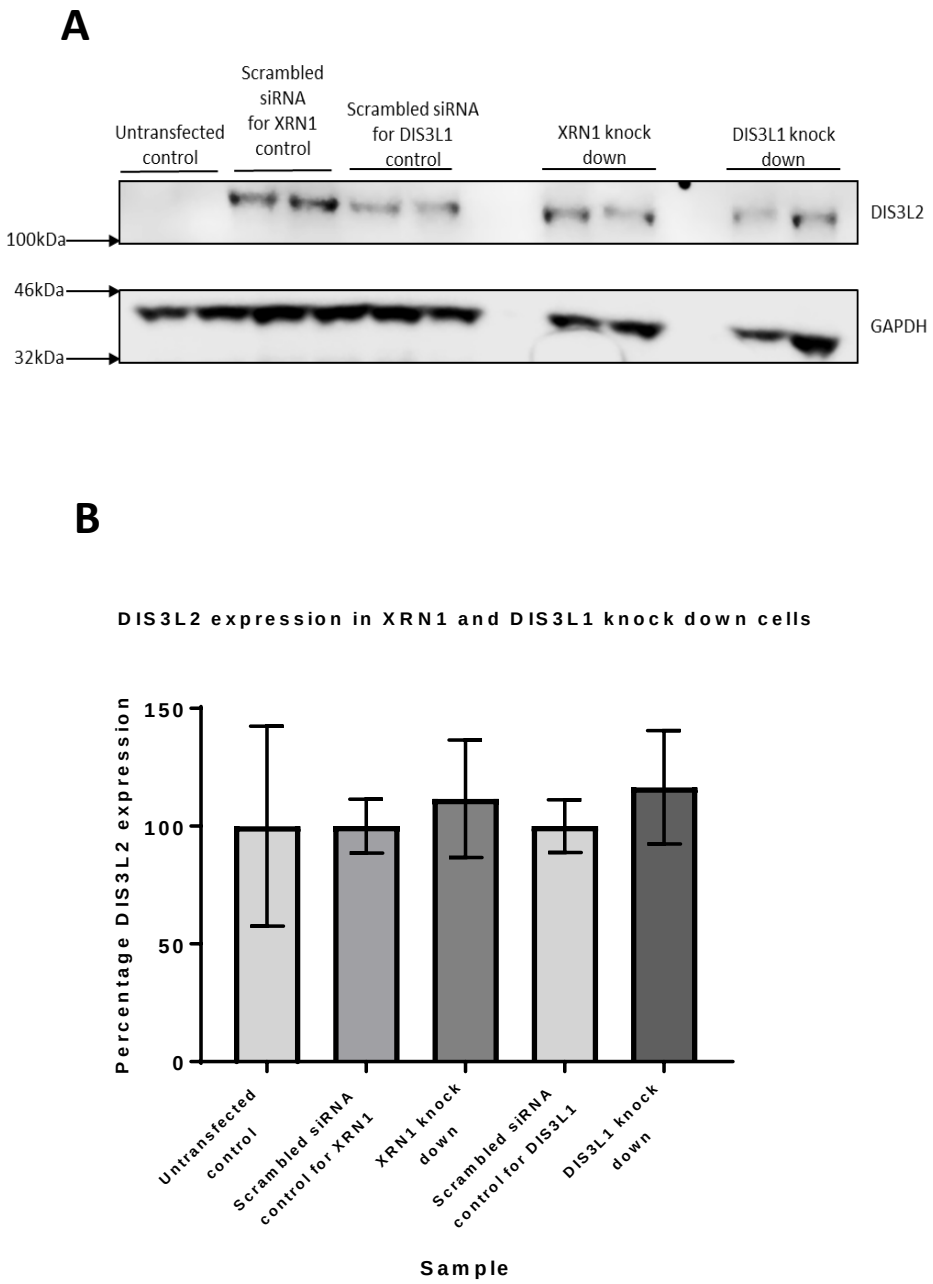


Figure 4.20. Protein expression of DIS3L2 in response to XRN1 and DIS3L1 knock down in SAOS-2 cells. A) Gel visualisation of DIS3L2 protein expression after 48hr knock down of XRN1 and DIS3L1 respectively, loaded in duplicate. B) Quantification of the percentage of DIS3L2 protein expression after XRN1 and DIS3L1 knock down, normalized to GAPDH. Error bars = SEM, based on n=3. Statistics were performed by unpaired t-test, where all comparisons were made to the scrambled siRNA controls for XRN1 and DIS3L1 respectively. All results were not significant.

4.13 Discussion

In this chapter, it has been shown that knock down of XRN1 using the Lipofectamine RNAiMAX system in SAOS-2 cells does not result in any significant impact on proliferation (as shown by growth curves and a DNA synthesis assay), apoptosis (as shown by the Caspase-Glo 3/7 assay) or cell viability (as shown by the WST-1 assay), nor does it show an effect on the rate of translation (as shown by SUnSET labelling). It has also been shown that knocking down DIS3L2 in U-2 OS cells does not influence proliferation, unlike the knock down in HEK-293T cells, where proliferation is increased. Furthermore, results in this chapter suggest that when XRN1 and DIS3L1 are down regulated respectively, DIS3L2 is up regulated to some extent compared to controls. This is interesting because it serves to show that there may be some form of compensatory regulation occurring between the ribonucleases, or that DIS3L1 and XRN1 are responsible for regulating DIS3L2.

Limitations in this chapter show that introduction of siRNA appears to have adverse effects on the behaviour of the cell line, however, it was thought to be the least invasive way of achieving high XRN1 knock down. Previous work in the lab had shown that other methods, such as electroporation, had led to an increased risk of cell death and so Lipofectamine was judged to be the best way forward, however, the risk of inducing adverse effects on the cell line was not totally eliminated. All of the experiments were inclusive of a negative control, whereby Lipofectamine reagent was added to cells which were not exposed to siRNA. These cells grew very well under this condition, and so it was concluded that the introduction of a foreign siRNA was the main cause for limitation in this procedure. Given that this was the best option available for knock down of XRN1 (see section 4.3), and that a scrambled control is absolutely necessary for this work, this limitation was unavoidable. It has meant distinguishing between true and false positive results was possible.

The efficiency of knock down of XRN1 was generally very good, with between 70% and 85% knock down consistently. This could only have been improved upon by creating a CRISPR knockout cell line for XRN1, a technology which was not widely available to the researcher at the time this series of experiments was performed. This could have achieved a knock out cell line, and so shown a true representation of the effect of repression of XRN1 on OS cells. Another advantage to using a CRISPR cell line is that the transfection is not transient, maintaining XRN1 depletion throughout the life of the cell line, providing the knock out is not lethal. Information obtained by the author after attendance at relevant conferences also confirmed that complete knock out of XRN1 is lethal to cells, and so this technology was not appropriate, regardless of its availability at the time. It is not clear whether the same cells with a high percentage of knock

down are consistently repressing XRN1, or whether some of these stop repressing XRN1, and new cells start repressing XRN1. Though this issue is somewhat resolved by using immunocytochemistry to fluorescently tag cells which have successful knockdown consistently, cells migrate during cell culture, and live cell imaging would also be needed to keep tabs on the movement of the cells. A 70-85% knock down of XRN1 does recapitulate the conditions of HOS cells, which have approximately 75% less XRN1 expression than the control, and so in this way the method of transfection was more like that of wildtype osteosarcoma cells.

With hindsight, a better way to represent the data obtained by the Caspase 3/7 Glo assay would be to create an enzyme kinetics graph. This is because it would be easier to interpret higher enzyme kinetics of caspase 3/7 by activity, rather than by using units of RFU (relative fluorescence units). It would be preferable to measure activity at the beginning of Caspase Glo 3/7 incubation, and then at various time points throughout the incubation period, to determine whether enzyme saturation has taken place. Although it is clear that overall caspase activity is not significantly changed, it is not clear how the rate of enzyme activity is changing throughout incubation. Further optimization is needed to obtain appropriate assay incubation periods, as the results so far show only assumed maximal caspase activity at the end of the test.

It is possible that although transfection efficiency was good, the remaining XRN1 in the cell was still enough to not cause an obvious functional phenotype. It could also be explained by compensatory expression of other exoribonucleases in the cell. It would not be unreasonable to assume that the cell may induce up regulation of other exoribonucleases in response to down regulation of XRN1. This is especially true with regard to the function of DIS3L2, a 3'- 5' exoribonuclease which has been shown to have a major role in proliferation in the *Drosophila melanogaster* wing disc model (Towler *et al.* 2016). XRN1 and DIS3L2 may work antagonistically in the cell (XRN1 promotes proliferation, DIS3L2 limits proliferation), however, repression of one could lead to up regulation of the other, as shown in Figure 4.20. This could result in compensatory proliferation by activity of DIS3L2.

It is unknown why DIS3L2 displays what appears to be very fine-tuned cell line and tissue specificity, although it is worth noting that HEK-293T cells are human embryonic kidney cells, taken from a developing foetus. This is a better match for the wing discs because, although a different species, these are also embryonic tissues, whereas U-2 OS cells are not embryonic. It could be argued that DIS3L2 is highly functional during early development, with decreased activity during the rest of the lifespan of the organism. Studies in *Drosophila* have shown that when DIS3L2 knock down in the early larval stage, the overgrowth phenotype is observed, whereas after this stage, no phenotype is observed (Towler *et al.* 2016). Indeed, partial and

whole deletions of DIS3L2 are implicated in the progression of Wilms' Tumour of the kidney. Wilms' Tumour of the kidney is also a consigned cancer of early childhood, and arises during key developmental years. Deletions of DIS3L2 are also pathologically relevant to Perlman's syndrome, a foetal overgrowth syndrome where phenotypes include macroencephaly and enlargement of key organs such as the kidneys. In this chapter, it has been shown that knock down of DIS3L2 in human embryonic cells, HEK-293T, does indeed increase proliferation of the cell line. This identifies DIS3L2 as a potential drug target for kidney cancers, however, when DIS3L2 is knocked down in the OS cell line, U-2 OS, no change in proliferation is observed. This is indicative of a tissue-specific phenotype, an interesting observation which is not covered in this thesis, but is discussed in a publication from the Newbury lab which is currently under review. It could be argued that DIS3L2 becomes less active as childhood commences, hence why there is no obvious phenotype when knocked down in U-2 OS, which is a cell line originally immortalized from an adolescent. Despite DIS3L2 knock down not giving an obvious phenotype in U-2 OS cells, OS is still a developmental disorder, so exploring this notion could be interesting.

Although there was no observable change in mRNA expression of DIS3L2 in response to knock down of both XRN1 and DIS3L1, there was a trend of up regulation of DIS3L2 protein expression, although this wasn't significant. Given that DIS3L2 may be more active in kidney tissue, it would be an idea in the future to see if it is more highly up regulated in HEK-293T cells after knock down of XRN1 and DIS3L1, due to the tissue specificity argument.

The fact that DIS3 shows down regulation in response to XRN1 knock down suggests that there may be some synergism between the two enzymes, and due to its targeting of regulatory RNAs this could help to explain why the proliferation and apoptosis pathways are not functioning in the normal way in response to regulation by XRN1 and DIS3L2 respectively. However, the functional significance of this is arguable, given that XRN1 and DIS3 act in different cellular compartments. Alternatively, it could suggest that by knocking down XRN1, this promotes up regulation of another factor, which could, in turn, lead to inhibition of DIS3. This could be a transcription factor, or an RNA binding protein, which leads to the down regulation of DIS3.

In the future it would be advantageous to investigate how XRN1 might be involved with cellular migration. Migration is an important factor in tumour development, and measuring the migration and invasion potential of the cell line would give an idea into how invasive the cancer is. If XRN1 is involved in targeting motility RNAs, it would make sense for it to be down regulated in cancer cells. A technique to achieve this utilizes the use of the Oris 96-well plate system, where cells are seeded around a small bung, of known diameter, and allowed to adhere, before the

bung is removed. Measuring the distance of migration of cells over the space left by the bung at different time points would give a rate for migration of the cells over the time period of the experiment. This experiment would include the knock down condition, the scrambled siRNA control condition, and the wildtype condition to assess whether cells migrate at a faster rate during the knock down condition compared to the controls.

As alluded to in Chapter 3, another theory to explain why XRN1 does not seem to have a high impact on the proliferation of cancer in these cells is that XRN1 may be indirectly down regulated in OS. There is an argument to make that there is in fact another candidate molecule in the cell affecting the expression of *XRN1* RNA, and that it is not XRN1 itself that is contributing to the OS phenotype and subsequent pathology. There are many genes that could be responsible for the low expression of XRN1 indirectly, as it is well known that to become malignant many cancers require the activation of multiple oncogenes. If XRN1 is being targeted, directly or indirectly, for the advancement of the cancer, then this presents *XRN1* as a novel tumour suppressor gene, whereby its suppressed activity is leading to defective apoptosis, and increased proliferation. This chapter shows that only DIS3L2 might appear to change its regulation pattern in response to both XRN1 and DIS3L1 knock down (whereas DIS3 only changes in response to XRN1 knock down) suggesting that DIS3L2, as the only other cytoplasmic exoribonuclease independent of the exosome, could show interplay with XRN1, and also the exosome. It has been shown that XRN1 and DIS3L2 co-precipitate together (Lubas *et al.* 2013), suggesting that they are able to target the same RNAs supporting the notion that DIS3L2 may displayed altered expression to compensate for reduced XRN1.

The notion that there is interplay between DIS3L2 and XRN1 is somewhat contradictory, because DIS3L2 has been shown to target transcripts encoding growth proteins. It seems unusual that XRN1, a gene that targets pro-apoptotic transcripts, should display lower expression and DIS3L2, which targets proliferative transcripts, should display higher expression compared to the HOb cell line, and yet the cancer cells are still able to proliferate uncontrollably. This paradox may suggest that there are other regulatory factors in the cell which might be involved, however, candidates for these remain unknown. Another hypothesis could be that perhaps, for some reason, cancer cells redirect the targeting of the exoribonucleases in some way, maybe by differential tagging of transcripts destined for degradation in order that they are not then degraded, allowing for proliferation to continue. Or, perhaps the cancer cells have developed a way of evading the RNA degradation pathway, or a way to utilize it to the advantage of the cancer cell. Alternatively, the small amount of up regulation of DIS3L2 might not be enough to have an effect on the cell.

The next chapter in this thesis will endeavour to find out which transcripts XRN1 is responsible for targeting in SAOS-2 cells, by knocking down XRN1 and utilizing RNA sequencing to decipher which transcripts appear to be differentially regulated compared to the scrambled control. This will give an insight into the functionality of XRN1 in OS, and will shed light on what is really happening inside the cancer cells with regard to XRN1.

Chapter 5

Identifying potential transcript targets of XRN1 using RNA sequencing

5.1 Introduction

RNA sequencing is a well-established technique used to identify RNA expression from biological samples. In this chapter, RNA sequencing was utilized to measure the transcriptomic differences between SAOS-2 cells which do not show a reduction in XRN1 expression (compared to control osteoblast cells (Figure 3.2)), and those which have had XRN1 artificially depleted by RNAi. RNA sequencing is not only used to measure the differential expression between mRNAs, but can also be used to identify expression, and differential expression, of other types of transcripts, such as lncRNAs, miRNAs, snoRNAs and tRNAs. The results from this type of deep sequencing analysis generally provide a highly intricate landscape of cellular processes across multiple contexts, and this technique is applicable for a broad range of molecular biology experiments. The strength of RNA sequencing technology overcomes the limitations of more traditional molecular biology tools used to measure transcriptomic expression differences, such as microarrays, because it has a better dynamic range and is not limited to a particular subset of transcripts – it can detect all transcripts. RNA sequencing is also advantageous because of its high processivity, and the sheer amount of metadata obtained from one experiment. This technique allows for analysis of the entire transcriptome, and so it also has the capability of identifying novel transcripts. This extent of deep analysis is simply not possible with more conventional techniques, which rely on current knowledge of the transcriptome in order to produce a dataset.

Over recent years, RNA sequencing technology has become much more accessible and much more affordable, another reason why it is often a preferable technique to use compared to other conventional techniques. It is for these reasons that RNA sequencing was chosen to perform the experiment in this chapter. Additionally, another advantage to using RNA sequencing is the scope for the detection of transcripts expressed at low levels. In the past, detection of transcripts has been subject to the level of background ‘noise’; an artifact which, though present, is much less prevalent in RNA sequencing technology. The fact that RNA sequencing can allow for the identification of extremely lowly expressed transcripts marks a major change in the way scientists view the transcriptome, offering up a wealth of previously undiscoverable data regarding gene expression.

Given that the previous work identified a cell-specific depletion of XRN1 in OS and EWS cell lines, it was felt that RNA sequencing would give an in-depth analysis that could penetrate much deeper into the landscape of molecular pathways of these cells. RNA sequencing can give a much more specific idea about the types of pathways XRN1 is involved in during the progression of OS, and perhaps why XRN1 expression appears to be lower in these cells, given that the previous results chapter did not show a clear phenotype. There is a difficulty in detecting phenotypes within cell lines, and this experiment should help to resolve this. Alternatively, it could provide an indication as to the functional relationships occurring, which would guide further analysis about why XRN1 displays reduced expression in these cells, and whether it is an indirect effect of another major phenomenon going on in the cells.

This RNA sequencing experiment will strive to elucidate a role for XRN1 in human cells, and it will also allow for the identification of XRN1 sensitive transcripts from the point of view of basic science. The information obtained by this experiment will help to fill current gaps in the knowledge about the activity of XRN1 in human cells. It is well known that morphological phenotypes, such as rate of growth and apoptosis, migration and the appearance of the cells within a culture, arevariable. Thus, an RNA sequencing experiment can be used to identify changes in the gene expression landscape of the cells over several replicates, and consistent changes could be plausibly attributed to the experiment, rather than external environment. In the same way, a cell line may not demonstrate any phenotypic change in response to expression manipulation experiments (such as transfection using siRNA), however, there still may be global or specific transcriptomic changes happening at the molecular level. If this is the case, the results of the RNA sequencing experiment in this chapter should show that XRN1 depletion does have an effect on the expression of specific transcripts in the cell, even though no obvious cellular phenotypes were previously observed.

5.2 Aims and hypothesis

The hypothesis in this chapter was that XRN1 is responsible for regulating specific transcripts involved in pathways associated with the progression of cancer. The aims of this chapter were to:

1. Perform an RNA sequencing experiment to determine differential expression of transcripts between SAOS-2 control cells and SAOS-2 cells where XRN1 has been knocked down synthetically using RNAi.

2. Analyse and identify differentially expressed transcripts using bioinformatics, with particular focus on up regulated transcripts, as these are more likely to be direct XRN1 targets.
3. Use gene ontology tools to uncover gene enrichment to identify cellular pathways which may be being regulated by XRN1

5.3 Knock down confirmation of XRN1 in SAOS-2 and U-2 OS cells

The first step in the RNA sequencing experiment was to create the samples which were to be analysed. Initially, this included performing XRN1 knock down on two OS cell lines: SAOS-2 and U-2 OS, in parallel with transfections of scrambled siRNA for use as a control, however, U-2 OS was later dropped from the experiment due to poor RNA integrity of the samples, possibly as a result of transfection of XRN1 siRNA. The purpose of using the scrambled siRNA of random sequence was to control for the cells natural response to the presence of a foreign RNA molecules and the transfection conditions used within the experiment. This therefore would give greater confidence in distinguishing differential gene expression following XRN1 depletion.

Knock down was performed in exactly the same way as described in previous chapters, whereby 20pmols of XRN1 siRNA or scrambled control siRNA was transfected to cells at a density of $3x10^5$ cells in a 6-well plate using the Lipofectamine RNAiMAX system. Cells were exposed to the siRNA for 24 hours, at which point the siRNA was removed, and fresh medium was added to the cells. After another 24 hours (48 hours post transfection), the cells were harvested, and RNA extraction was performed using the Qiagen miRNeasy Kit. 48 hours exposure to siRNA was performed as it previously showed the greatest down regulation of XRN1 (Figure 4.2). RNA concentrations were measured on a NanoDrop One spectrophotometer for initial concentration measurements (Table 5.1), and then 500ng of the RNA was used for subsequent qRT-PCR to check that the knock down had been effective (Figure 5.1). The NanoDrop One measurements also showed that the RNA samples were of high purity given the 260/280 and 260/230 ratios, meaning that there was only low contamination with solvents and organic materials. Only the samples where *XRN1* was knocked down greater than 2-fold (>50% depletion) when compared to the scrambled control and normalized to *GAPDH* expression were considered for RNA sequencing, Samples with knock down of greater than 4-fold were generally used for RNA sequencing analysis as they were thought to indicate supreme levels of XRN1 down regulation, essentially making the RNA sequencing worthwhile (the average percentage of XRN1 expression after knock down was 21% compared to 100% for the scrambled siRNA controls). In total, 6 samples were created for XRN1 knock down and 6 samples were selected for the scrambled

siRNA control. Each of the knock down samples and control samples were matched numerically, and performed in parallel with each other to allow for the detection of batch effects which may have been caused by variations in culture conditions week to week.

Treated sample	RIN	260/280	260/230	Yield (ng/µL)
Scrambled siRNA control n1	9.7	2.13	2.14	488.5
XRN1 Knock Down n1	8.9	2.10	1.91	417.8
Scrambled siRNA control n2	9.3	2.08	2.07	636.5
XRN1 Knock Down n2	7.5	2.12	1.40	480.4
Scrambled siRNA control n3	9.7	2.11	2.07	341.4
XRN1 Knock Down n3	6.2	2.11	1.92	364.0
Scrambled siRNA control n4	8.7	2.09	2.03	317.4
XRN1 Knock Down n4	8.3	2.08	2.09	225.1
Scrambled siRNA control n5	9.5	2.08	1.96	441.7
XRN1 Knock Down n5	8.4	2.11	2.00	411.1
Scrambled siRNA control n6	9.6	2.09	1.52	344.2
XRN1 Knock Down n6	8.8	2.09	1.95	315.6

Table 5.1. List of RNA sequencing samples sent off for sequencing. 12 samples in total were sent for RNA sequencing, where n denotes the replicate number. Each knock down sample was complemented in parallel with a scrambled siRNA control sample. RIN = RNA Integrity Number, this value denotes the quality of the RNA to be sequenced, where any number above 6 is considered of high enough quality. It was measured using the Agilent 2100 Bioanalyzer, in parallel with 2100 Expert software. The 260/280 and 260/230 ratios display low contamination with solvents and organic material such as proteins, and these were measured (alongside RNA yield) on a NanoDrop One.

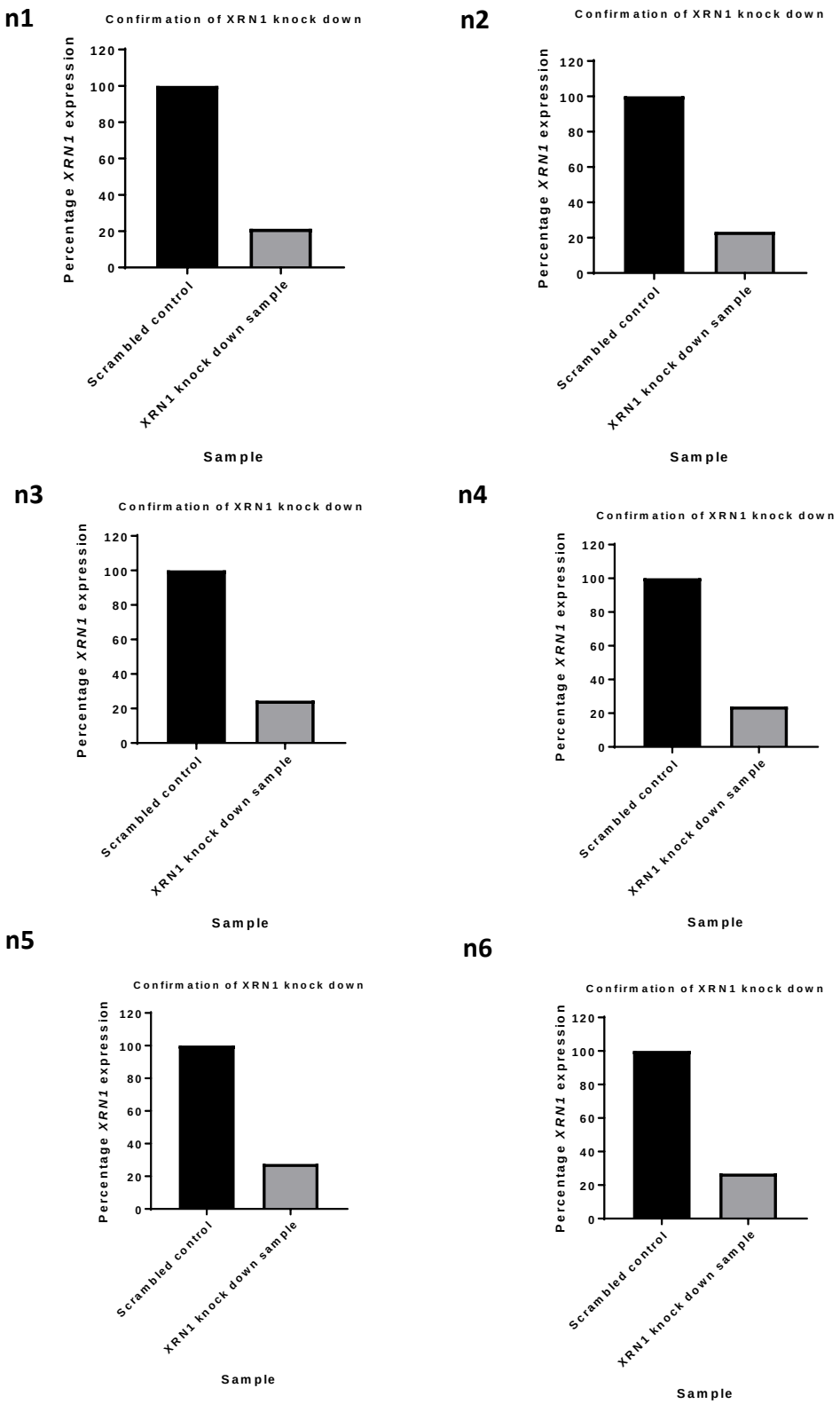


Figure 5.1. Confirmation of XRN1 knock down in SAOS-2 RNA sequencing samples. XRN1 knock down was confirmed in all 6 samples for RNA sequencing using qRT-PCR. Expression was normalised to *GAPDH*, individual statistical analysis was not performed due to 1 replicate per sample.

5.4 Bioanalyzer and Qubit confirmation of RNA integrity

Prior to samples being sent for sequencing, 2 further checks were made on the RNA to fully ensure that the quality and integrity of the RNA was such that good quality sequencing data could be obtained.

The first of these was the use of a Bioanalyzer to measure RNA integrity. This method utilizes a gel slurry to measure RNA integrity, which is detected by the fluorescence of dye which intercalates into nucleic acid. The nucleic acid samples are run through a microfluidic plate of microchannels, and the nucleic acid is electrophoretically separated. The intercalated dyes are then detected by fluorescence and translated into electropherograms (the trace). The resulting trace should, if RNA integrity is good, show 2 peaks: one for the 18S ribosomal RNA (which comes up between 40S and 45S, where S= the sedimentation coefficient, Svedberg unit) and one for the 28S ribosomal RNA (which comes up between 45S and 50S). Generally, the more pronounced the peaks with less interference, the better the quality of the RNA. Interference indicates that there is at least some partially degraded RNA in the sample, which can occur for a number of reasons, such as repeated freeze-thaw cycles of the RNA, or poor sample preparation. If the RNA is highly degraded, the ribosome peaks often shift to the left of the expected markers. The Bioanalyzer gives an RNA integrity number (RIN) as a score for RNA integrity, which operates on a 1-10 scale, whereby 1 = very poor RNA integrity and 10 = excellent RNA integrity. For the purposes of the RNA sequencing experiment, any RNA sample submitted to the Bionanalyzer which gave a RIN of <6 was discarded. The samples submitted for analysis from the U-2 OS cell line knock downs mostly had very poor RIN values, and it was for this reason that they were discarded from the RNA sequencing experiment. The Bioanalyzer traces for the samples which were sent off for sequencing are shown in Figure 5.2; the RIN values (from the Bioanalyzer), 260/230 and 260/280 ratios, and overall RNA yield (as measured on a NanoDrop One) are shown in Table 5.1.

As a third layer of control for RNA concentration, a Qubit was used to assess the RNA samples as it provides a more accurate quantification than the NanoDrop. The Qubit, like the Bioanalyzer, can measure fluorescence of DNA, RNA and protein, and molecule-specific dyes are used to incorporate into nucleic acid. These dyes have very low fluorescence when they are not bound to their target, however, the intensity of fluorescence is magnified intensely when bound. Fluorescence is directly proportional to the concentration of the molecule being analysed, and the Qubit then converts the reading into a measurement for concentration against known standards for the molecule. In this case, the Qubit measured RNA concentrations of between 200 and 400ng per sample in 20μL (this amount was required by Leeds Genomics in order for

them to perform library preparation) after samples had been diluted in preparation for RNA sequencing. These concentrations were within the range specified by Leeds Genomics as optimal for library preparation and RNA sequencing. RNA samples were sent for RNA sequencing (including library preparation) at the Leeds Genomics facility at the University of Leeds.

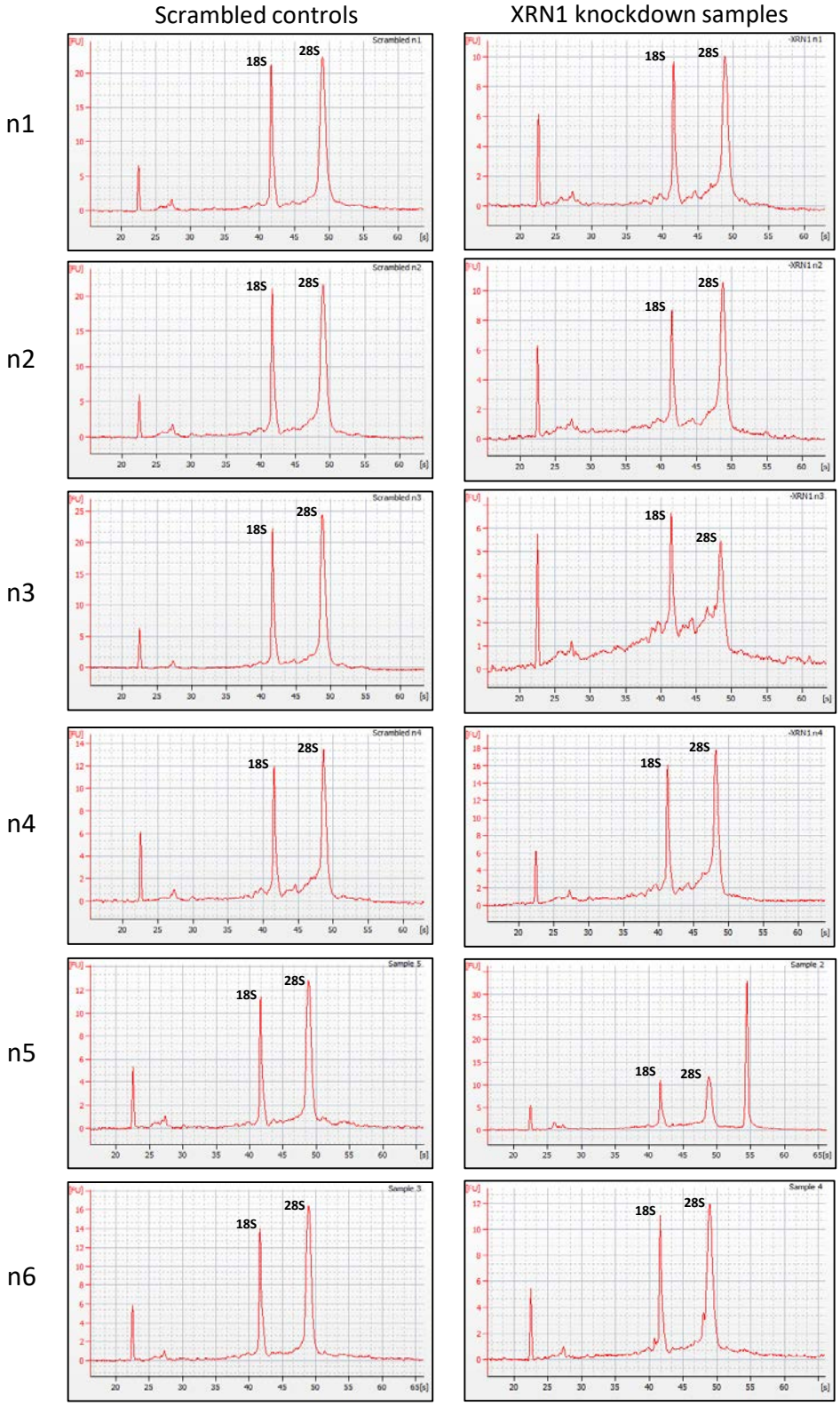


Figure 5.2. Bioanalyser traces for RNA sequencing samples. 2 columns of bioanalyser traces performed to evaluate the quality of RNA sent off for sequencing, using the Agilent 2100 Bioanalyser. All RNA with a RIN of >6 were sent off for sequencing, according to recommendations from Leeds Genomics.

5.5 RNA sequencing raw data analysis

After the RNA samples were deemed of high enough quality for RNA sequencing, they were sent by courier to the Leeds Genomics sequencing facility for Next Generation Sequencing (NGS) using the NextSeq Illumina technology. Key applications for this platform include whole-transcriptome profiling, which is not a key application of other benchtop sequencers. Indeed, the NextSeq is the most advanced of the benchtop sequencers offered by Illumina, and provides scope for analysis for a multitude of applications, more than any other of the benchtop sequencers available. The output of the NextSeq is also superior, with the output potential for 400 million reads, more than any other benchtop sequencer currently offered by Illumina.

NGS works by performing sequencing by synthesis allowing the tracking of the addition of labelled nucleotides as the cDNA is copied (Illumina). The first step in this process is library preparation of the RNA sent for sequencing. Library preparation was performed using the Illumina TruSeq kit. For these samples, rRNA depletion (using RiboZero) was used rather than poly(A) selection in the library preparation step to allow for the assessment of the whole transcriptome, rather than just RNAs which have a poly(A) tail. Library preparation involves both rRNA depletion (due to the extensive amount of rRNA), and fragmentation of the RNA, which is followed by cDNA synthesis, adapter ligation and, lastly, PCR amplification. It is important that PCR amplification is not extensive, to avoid PCR bias. Once the cDNA has been hybridized to the flow cell, the sequencing reaction can take place. DNA polymerase allows the catalysis of the incorporation of fluorescently labelled dNTPs into a DNA template strand during repeated cycles of DNA synthesis. Each cycle produces its own pattern of excitation as single nucleotides are identified via their fluorophore excitation. In contrast to other sequencing methods, NGS is able to conduct this process across millions of fragments of DNA in parallel, rather than sequencing just one single strand at a time. This massively improves the output of the experiment, and the speed with which sequencing can occur. In this experiment, single-end sequencing (as opposed to paired-end sequencing) was utilized because it provides accurate read alignment whilst also generating a dataset which is much more manageable, and is less expensive. Single-end sequencing involves sequencing the RNA from just one end of the transcript, whereby paired-end sequencing sequences the transcript from both ends. Although it does not provide an analysis as in-depth as paired-end sequencing, it is the simplest and quickest RNA sequencing method to generate high quality data.

Once the raw data files, in a FastQ format, had been received from the Leeds Genomics facility, it was important to process the data in order compare gene expression between the scrambled siRNA control samples and the XRN1 knock down samples. This involved using a

pipeline adapted by a post-doctoral researcher in the lab (Dr Ben Towler) for the purposes of analyzing RNA sequencing data. Data processing was carried out by the author using this pipeline as described in Section 5.6.

5.6 RNA sequencing analysis pipeline

When the physical aspect of RNA sequencing has been completed, the data needs to be processed in order for it to be analysed. The file received from the sequencing centre gave an indication as to the quality of the sequencing, which in this case was very good (Figure 5.3). The reads were mostly in line with expected distributions and the quality scores were high. The files were then put through an array of computational processing to ensure optimal quality of the data was reached before mapping the sequences to the human genome, so that gene expression could be quantified. There are many RNA-sequencing pipelines that can be used to analyse the results of an RNA sequencing experiment, and this particular pipeline was followed due to existing expertise in the lab:

1. **Quality assessment** – This is achieved using a program called FastQC.
2. **Adapter removal and trimming** – Specific adapter sequences are added during library preparation which are then sequenced alongside the fragments, and these sequences need to be removed. The quality assessment step identifies over-represented adapter sequences, and the specific index used can be found using the online document published by Illumina. Adapter sequences are then removed by a program called Scythe. Adapter removal will improve the quality of the data, however, this can be further improved by a program called Sickle, which removes poor quality reads (reads below a quality score of 30) to enhance mapping efficiency, by quality trimming. It also keeps paired reads together if paired-end sequencing was performed. All data referring to the adapter sequences and quality trimming can be viewed in Table 5.2.
3. **Building the reference genome** – The reads are aligned to a supplied genome, to identify where in the genome they are derived from. In this experiment, the human genome GRCh38.93 was obtained from Ensembl. This step was completed using the index function of HiSat2.
4. **Aligning the sequences to the reference genome** – Once built, the reads need to be mapped to the reference genome. This is also achieved by HiSat2. The output of this process is a .sam file: Sequence Alignment/ Map file. In order to assemble the aligned reads into transcripts a binary .sam file (.bam) is required. This conversion from .sam to .bam was achieved using samtools.

5. **Assembling the sequences of transcripts** – This step is performed by the programme, Cufflinks. This programme assembles reads into transcripts, forms initial abundance estimates and performs initial normalisations (transcript length, total number of reads per sample).
6. **Creating the reference assembly** – The programme, Cuffmerge, creates an optimal transcript assembly, in all senses a 'master transcriptome', which is useful for both lowly expressed genes and downstream quantification of novel transcripts that might be sample specific.
7. **Quantification of the abundances of each transcript** – Cuffquant uses the Cuffmerge file to allow quantification of novel transcripts. Cuffquant also re-assesses the normalization performed by Cufflinks using transcript length determined by Cuffmerge. Normalised abundance is then presented as fragments per kilobase per million reads (FPKM). This is then used to measure differential expression of transcripts.
8. **Differential expression of transcript analysis** – Comparison of each transcript in the control samples versus the knock down samples was undertaken by Cuffdiff. This programme calculates the average FPKM for each condition and then calculates the fold change and statistical significance for every transcript, which can then be further analysed and graphically represented in other analysis programmes, for example, R and Microsoft Excel.

A flowchart outlining the sequence of events is shown in Figure 5.4 and a summary of all the software used in the RNA sequencing analysis pipeline can be found in Table 5.3.

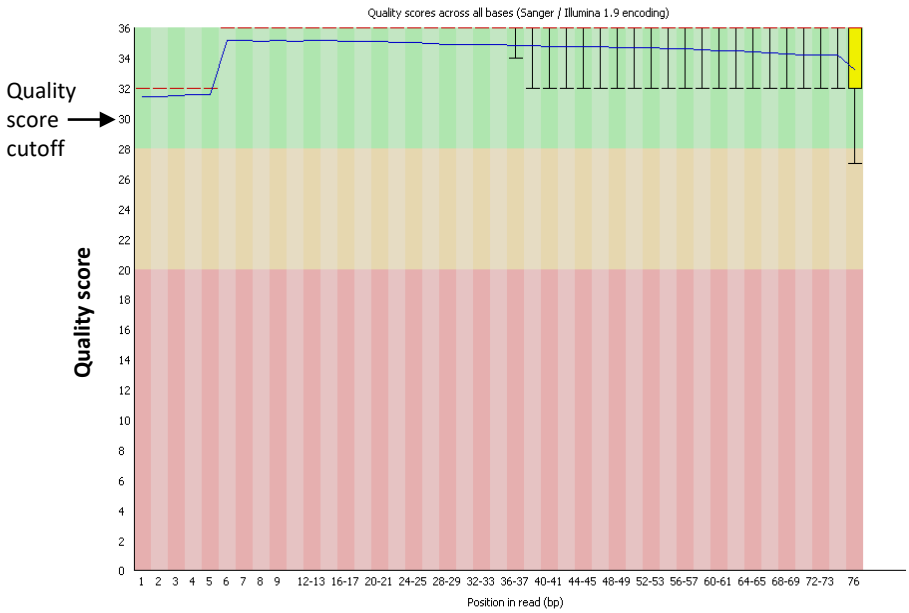


Figure 5.3. Example Quality Score Chart. A quality score is calculated for each base of each read. The line and error bars show the average quality of each base. Scores in the green area are considered excellent, yellow acceptable, and red poor. This particular chart shows the read quality of the Scrambled control replicate n1, and is a typical representation of the quality of the reads obtained from the other samples.

Sample	Illumina Adapter index	Adapter identifier sequence	Number of reads before adapter removal	Adapter-containing reads	Reads without adapter	Records kept after sickle quality trimming	HiSat2 percentage of reads aligned 0 times	HiSat2 number of reads aligned 0 times	HiSat2 percentage of reads aligned 1 time	HiSat2 number of reads aligned 1 time	HiSat2 percentage of reads aligned >1 time	HiSat2 number of reads aligned >1 time	HiSat2 overall percentage alignment	HiSat2 total number of reads aligned
Scrambled n1	2	CGATGT	47673024	311026	47361998	175000	2.85%	1354504	84.37%	40076130	12.77%	6067390	97.15%	46143520
Scrambled n2	6	GCCAAT	41741998	243399	41598000	143998	2.92%	1216387	85.62%	35615662	11.46%	4765951	97.058%	40381613
Scrambled n3	12	CTTGTA	44223933	224980	44076762	147171	2.85%	1257637	85.93%	37874720	11.22%	4944405	97.15%	42819125
Scrambled n4	4	TGACCA	44163001	238557	43708038	143691	2.92%	1275490	86.39%	37758235	10.69%	4674313	97.08%	42432548
Scrambled n5	15	ATGTCA	43962878	234365	43806644	156234	2.93%	1282297	85.61%	37503084	11.46%	5021263	97.07%	42524347
Scrambled n6	21	GTTTCG	41699876	225379	41556893	142983	2.88%	1194834	85.53%	35542932	11.6%	4819127	97.12%	40362059
XRN1 kd n1	5	ACAGTG	46013238	342179	45836738	176500	3.07%	1407080	83.67%	38349374	13.27%	6080284	96.93%	44429658
XRN1 kd n2	7	CAGATC	43277224	245214	43136018	141206	3.19%	1375124	85.18%	36744437	11.63%	5016457	96.81%	41760894
XRN1 kd n3	19	GTGAAA	41207596	224980	41053919	153677	3.19%	1308622	85.52%	35108920	11.29%	4636377	96.81%	39745297
XRN1 kd n4	13	AGTCAA	44163001	240981	44005909	157092	3.24%	1427007	85.33%	37548819	11.43%	5030083	96.76%	42578902
XRN1 kd n5	20	GTGGCC	40582918	205478	40448478	134440	3.16%	1279055	85.11%	34426934	11.72%	4742489	96.84%	39169423
XRN1 kd n6	22	CGTACG	38326048	212383	38196075	129973	3.19%	1218412	84.76%	32375348	12.05%	4602315	96.81%	36977663

Table 5.2. Read quality processing steps, including the before and after numbers per processing step.

Figure 5.4. Flowchart showing the sequence of analysis and the programmes used for processing raw RNA sequencing data. This flowchart shows the exact sequence used in the processing of the raw RNA sequencing, from initial quality checking through to measuring differential gene expression between the knock down samples and the scrambled siRNA controls.

Programme	Version	Use
FastQC	0.11.7	Quality assessment and adapter sequence identification on raw data.
Scythe	0.993b	Adapter removal.
Sickle	1.29	Quality trimming after adapter removal to enhance mapping efficiency.
HiSat2	2.1.0	Aligning reads to the reference genome.
SAMtools	0.1.19	To manipulate the alignments from the SAM format into the BAM format for indexing.
Cufflinks	2.2.1	Assembling reads into transcripts and initial normalisations.
Cuffmerge	2.2.1	Creates a 'master' transcriptome, an optimal transcript assembly.
Cuffquant	2.2.1	Final quantification of the abundance of each transcript.
Cuffdiff	2.2.1	Comparisons of transcript abundance between specific conditions for differential expression analysis.
R	I386 3.5.1	Graphic analysis of RNA seq data.
CummeRbund	3.8	An R visualization package for Cufflinks high throughput sequencing data
ggplot2	3.1.0	A script to create elegant data visualisations within cummeRbund.
FeatureCounts	1.5.2	Ultrafast and accurate read summarization programme.
MEME Suite	5.0.4	Discovers novel, ungapped motifs in sequences.

Table 5.3. Overview of the programs used during the RNA sequencing analysis.

5.7 RNA sequencing identifies XRN1-sensitive transcripts

Following on from the file created by Cuffdiff, Microsoft Excel was used to filter the results into a shortlist of genes which were significantly up regulated upon XRN1 knock down corrected for multiple comparisons using a false discovery rate of 0.05 (known as the q-value). A series of data filtering was performed because the statistical output from Cuffdiff can be biased at times by a single replicate, which could skew the output due to the amount of multiple comparisons it has to make. The total number of genes which were significantly differential in their expression was 1647. This first filtering selection involved obtaining all up regulated transcripts based on consistent up regulation of the transcripts across all replicates as calculated by Cuffdiff. Analysis of down regulated genes is described later on in this chapter (see section 3.8.5, Table 5.4). The total number of up regulated genes was 766, and the total number of down regulated genes was 881. The smallest fold change up regulation was 1.23 fold.

The second stage was to compare each individual replicate for consistency, which also served to remove any genes with unrealistic expression patterns, such as genes which were up regulated by infinity fold (i.e. only detected in knock down samples), or which were highly variable across replicates. The issue surrounding some of the genes which were identified as being differentially expressed by infinity fold is that for very lowly expressed genes the infinite fold change does not distinguish between 0-1000 fold expression, and 0-0.001 fold expression, and so although genes being up regulated infinitely was an interesting observation, they were excluded from the analysis because of uncertainty. The next selection criterion was to filter the results so that only transcripts which were up regulated >1.5 fold across each individual replicate were considered for validation. This meant comparing the n1 scrambled control to n1 knock down only, and so on, to ensure there was consistency across all replicates. This step reduced the total number of up regulated transcripts to 331.

The last selection process was to filter the results based on the FPKM (**F**ragments **P**er **K**ilobase of transcript per **M**illion mapped reads) values of all knock down samples being >1. Expression is measured by FPKM, and this allows for assessment of technical variation between the replicates at this stage. RNA sequencing has technical variation at the lowest level, and so using a FPKM cutoff value of 1 means that there can be more confidence in the fold change of the transcript. An FPKM value cut off of 0.3 has previously been used (Ramskold *et al.* 2009), and more recent studies have shown that 1 is also a reasonable cut off (Hart *et al.* 2013). The higher the FPKM, the more confidence there is in the up regulation of the transcript, because more of the transcript has been detected. After first selecting the knock down sample results for FPKM values of >1, the total number of up regulated genes was reduced to 264. The results

were then filtered again based on the FPKM values of each of the Scrambled control replicates being >1. This brought the final number of confidently up regulated, expressed genes to 212. A summary table of the filtering rules and gene totals is shown in Table 5.4. The top 20 up regulated genes and their functions are represented in Tables 5.5A and 5.5B.

Following the stringent filtering criteria outlined above, R was used to visualize the final RNA sequencing results from the above pipeline in the form of a dispersion plot of all detected genes, whereby all up regulated transcripts are plotted in pink above the line x=y (Figure 5.5).

Rule number	Rule description	Subsequent gene total for up regulated genes	Subsequent total for down regulated genes
1	Significantly differential according to Cuffdiff	766	881
2	Fold change >1.5	331	668
3	FPKM >1 in knock down samples	264	366
4	FPKM >1 in scrambled control samples	212	318

Table 5.4. Summary of rules for filtering the RNA sequencing data. These rules were applied in order to reduce the chance of false positives from being included in the data analysis. Cuffdiff can produce biased results across multiple comparisons, which is why it was important to filter the results based on the fold change being >1.5 and FPKM being >1 between each replicate.

Gene	Full name	Molecular Function	Biological process	Log2 Fold change
NID1	Nidogen-1	Ca2+ binding, collagen binding, laminin binding, proteoglycan binding.	Cell-matrix adhesion, basement membrane organisation.	2.39
C1QL1	Cq1-related factor	Signalling receptor binding.	Locomotory behaviour, maintenance of synapse structure, motor learning.	2.19
GDF11	Growth/ differentiation factor 11	Cytokine activity, growth factor activity, TGF-b receptor binding.	Mesoderm development, regulator of: apoptosis, cell proliferation (negative), MAPK cascade, SMAD signal transduction.	2.18
ENHO	ENHO (protein name: Adropin)	Hormone activity.	Insulin signalling, glucose homeostasis. Positive regulator of lipid synthesis, metabolism, Notch signalling, oxidoreductase, RNA pol II DNA binding.	2.16
HMOX1	Haem oxygenase 1	Haem oxygenase activity.	Angiogenesis, apoptosis, iron ion homeostasis, cellular response to hypoxia.	1.92
GLO1	Lactoylglutathione lyase	Zinc ion binding.	Cellular metabolism, negative regulator of apoptosis, osteoclast differentiation.	1.84
NUDT15	Nucleotide triphosphate diphosphatase	Catalyse the hydrolysis of NTPs and NDPs.	May have a role in DNA synthesis and cell cycle progression by stabilising PCNA.	1.81
KDELC2	KDEL motif-containing protein 2	Glucosyltransferase activity.	Thought to be located in the ER lumen.	1.78
CPNE2	Copine-2	Calcium-dependent phospholipid binding.	Cellular response to calcium ion.	1.77
LRRC20	Leucine-rich repeat-containing protein 20	N/A	Implicated in cognitive development, and linked to OCD and autism.	1.71

Table 5.5A. List of functions for the top 20 up regulated genes, top 1-10 featured here. Genes continued in Table 5.5B.

Gene	Full name	Molecular Function	Biological process	Log2 Fold change
IL6ST	Interleukin-6 receptor subunit beta	Cytokine binding/receptor activity, growth factor binding.	Of note, involvement with osteoblast differentiation and most types of immune response inc inflammation and T cell proliferation. Also postive regulator of Notch and VEGF.	1.70
SEMA6B	Semaphorin-6B	Chemorepellent activity and semaphorin receptor binding.	Negative regulation of axon extension in axon guidance and positive regulation of cell migration.	1.68
EMP2	Epithelial membrane protein 2	Integrin binding and kinase binding.	Angiogenesis, cell migration, proliferation, cell-matrix adhesion, cell death, vasculogenesis and kinase activity.	1.68
CORO1A	Coronin-1A	Actin binding, RNA binding, PI3K binding.	Heavy involvement in the immune response, and involved in the neuron apoptotic process, nerve growth factor signalling, vesicle fusion and cell migration.	1.68
PPP1R14B	Protein phosphatase 1 regulatory subunit 14B	Serine/threonine phosphatase inhibitor activity.	Innate immune response and regulation of phosphorylation.	1.66
EDIL3	EGF-like repeat and discoidin I-like domain containing protein 3	Ca2+ ion binding, ECM structural constituent and integrin binding.	Positive regulator for cell adhesion, organismal development.	1.64
IGFBP5	Insulin-like growth factor-binding protein 5	Fibronectin, IGF-1 and IGF-2 binding.	Glucose metabolism, postive regulation of the insulin-like growth factor receptor signalling pathway and osteoblast differentiation and muscle development.	1.64
LHX2	Lim/ homeobox protein Lhx2	Chromatin binding, transcription factor. RNA pol II proximal promoter sequence-specific DNA binding.	Axon development, mesoderm development, regulation of trancription.	1.63
SCAMP1	Secretory carrier-associated membrane protein 1	Post-Golgi recycling pathways, it acts as a recycling carrier to the cell surface.	Neutrophil degranulation, post-Golgi vesicle-mediated transport.	1.63
MPP2	MAGUK p55 subfamily member 2	N/A	Synaptic potentiation.	1.60

Table 5.5B. Top 11-20 up regulated genes as shown in the RNA sequencing.

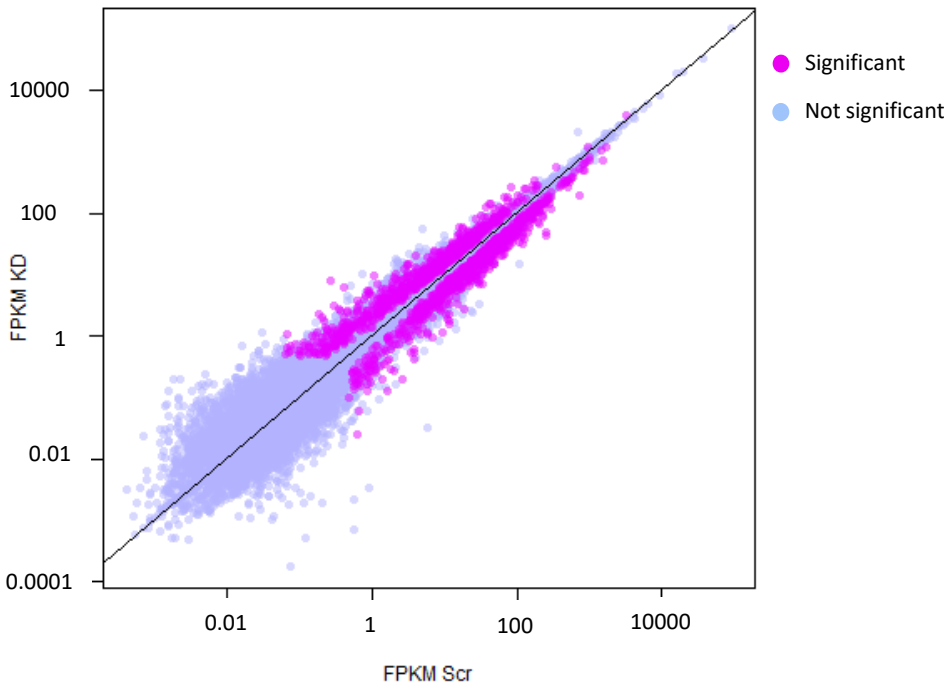


Figure 5.5. Graph to show differential expression between XRN1 knock down cells and control cells treated with scrambled siRNA, as detected by RNA sequencing. All genes detected by RNA sequencing are plotted on the graph, all gene transcripts which are differentially expressed significantly are highlighted with pink. As can be seen by the graph, most detected genes are not expressed differentially in the knock down samples when compared to the scrambled control samples. Knock down samples are denoted by KD on the y-axis and scrambled control samples are denoted by Scr on the x-axis.

5.8 Gene Ontology analysis identifies an enrichment of cellular pathways

Gene ontology (GO) analysis was also performed in order to find out if the differentially expressed genes were associated with any specific cellular pathway. This involved finding the Ensembl gene IDs for each gene, and inputting the Ensembl gene ID list into a web-based functional annotation tool: the Database for Annotation, Visualisation and Integrated Discovery (DAVID) (Huang da *et al.* 2009, Huang da *et al.* 2009). DAVID allows for gene enrichment to be quantified using its extensive knowledgebase of annotations and is constantly being updated with information from other databases. DAVID searches against the background genome in order to allocate genes to categories based on gene list submitted and how likely they are to be a part of a particular pathway. Each category is assigned as enrichment score (calculated as –log10(p-value)), and a significant enrichment score is defined as any score >1.3 when using p=0.05 as a cut-off. GO analysis was performed based on the functional annotation of the gene, using the default settings. During the GO analysis it was apparent that using human genes was an advantage due to the fact that this particular genome has been highly annotated.

A limitation of GO analysis by DAVID is that it relies on previous annotations of biological functions for genes, therefore, any previously unidentified transcripts cannot be assigned a category, which is problematic when trying to identify transcripts targeted by exoribonucleases. These transcripts may be very short-lived within the cell, and so a high proportion of the transcripts picked up the highly sensitive nature of RNA sequencing may not have been annotated before. Nevertheless, GO analysis identified 32 clusters of enriched genes. Of these, 11 had enrichment scores >1.3, suggesting a role for XRN1 in regulating transcripts from these pathways. Annotation cluster 1, with the highest enrichment score of 2.61, showed that the highest level of enrichment of up regulated genes in the RNA sequencing dataset were heavily involved in the Epidermal Growth Factor (EGF) pathway (GO cluster analysis shown in Table 5.6). This is interesting because EGF is a growth pathway, and so depletion of XRN1 seems to suggest an enrichment of transcripts involved this major oncogenic pathway. EGF is associated with many different types of cancer, for example, non-small cell lung cancer, and defects in this pathway contribute to uncontrolled proliferation (Chen *et al.* 2019). Previous growth experiments (Chapter 4) suggested that XRN1 knock down did not have a significant effect on the rate of cell growth, however, these results suggest that knock down of XRN1 results in the increased expression of a selection of transcripts which have the potential to contribute to a higher rate of cell growth.

The most enriched gene within the EGF-like cluster is the Insulin-like growth factor binding protein (IGF-BP5). Although this particular member of the IGF pathway has not been

validated by qRT-PCR for confirmation (due to assay expense), other members of the pathway have previously been tested for dysregulation in the OS cell lines, based on previous links of the IGF pathway with the progression of OS (Herzlieb *et al.* 2000, Mao *et al.* 2017): IGF-1 and its receptor, IGF-1R. Of these, *IGF-1* showed a significantly higher expression in all three cell lines, and *IGF-1R* showed a significantly higher expression in all three cell OS cell lines when compared to the HOb control (Figure 5.6). This suggests another effect is impacting on the amount in intracellular IGF-1 in the cells. The levels of *IGF-1* and *IGF-1R* were also tested by qRT-PCR in the patient samples, which showed that *IGF-1* is significantly lower in the patient samples as well, however, *IGF-1R* is much more variable, and no significant change in expression was observed (Figure 5.7). These changes are, therefore, likely to be caused by variation within cell lines rather than by a reduction in XRN1. Strengthening this argument is the fact that neither IGF-1 nor IGF-1R were detected as differentially expressed in the RNA sequencing dataset when XRN1 was knocked down. The patient samples were normalized to the HOb control cell line, the limitations of which have been previously discussed, which could account for the disparity in results between the patient and cell line data.

The RNA sequencing dataset shows that IGFBP5 is the 23rd most up regulated gene (fold change x3.12), and IGFBP4 is also up regulated (by a fold change of x1.82), at position 184 out of 212 up regulated genes. These may be candidates responsible for the dysregulation of the IGF pathway upon XRN1 knock down, and so qRT-PCR will be performed in the future to validate their expression. IGFBP4 binds to IGF proteins to prolong their half-life, and alters their interaction with cell surface receptors. IGFBP5, on the other hand, is a potent inhibitor of IGF-1, and functions within the mTORC1 pathway as a potential tumour suppressor (Ding *et al.* 2016). Alongside this, it has been shown that up regulation of TGF-β can also suppress the expression of IGF-1, preventing the differentiation of osteoblast progenitors into mature osteoblasts (Ochiai *et al.* 2012). It would be interesting in the future to characterize the expression of TGF-β in the OS cell lines, as up regulation may be causing lower *IGF-1* expression in these cells, preventing terminal differentiation.

The gene clusters were each sorted into six clusters to group the genes into related clusters. The only clusters taken forward were those with a significant enrichment score of >1.3, and the genes within each cluster were tabulated in Table 5.6. Interestingly, several of the genes appear across multiple categories, suggesting that XRN1 is having an effect on multiple pathways through regulating a few consistent genes.

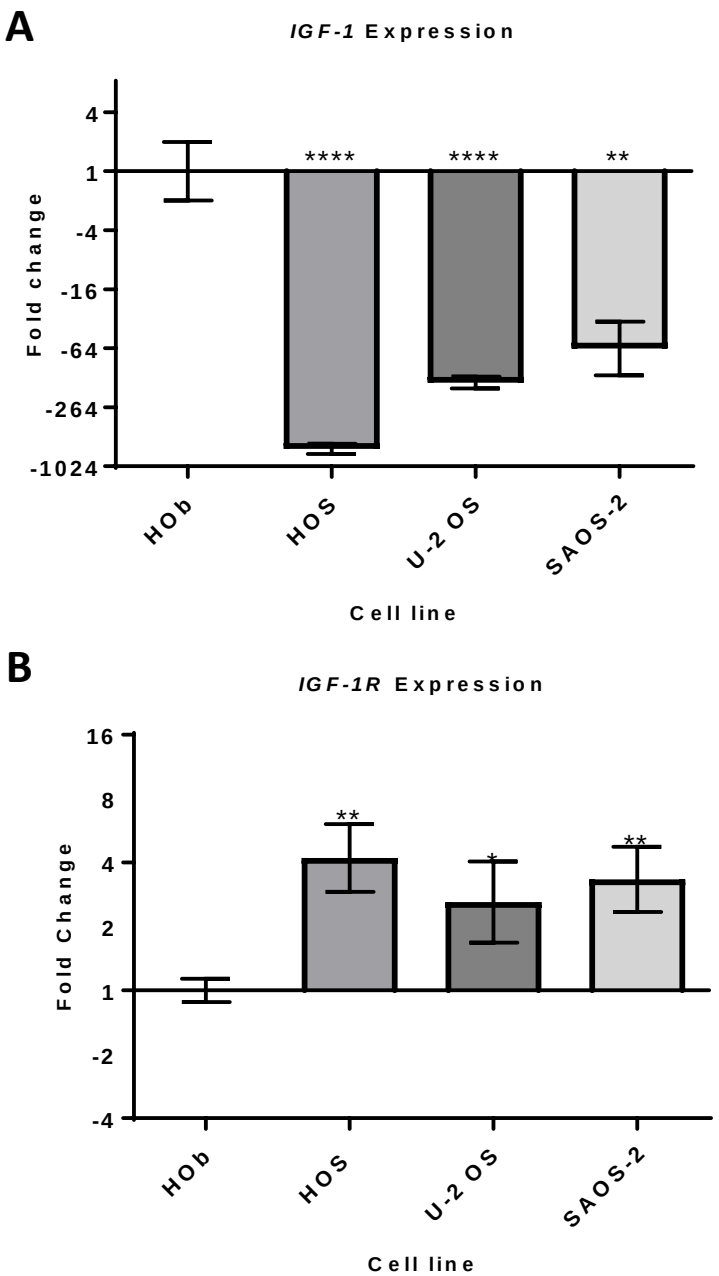


Figure 5.6. Expression of *IGF-1* and *IGF-1R* mRNA in OS cell lines relative to the HOb control cell line. A) Bar graph quantification of *IGF-1* mRNA, showing significantly lower expression of *IGF-1* in cancer cell lines, statistical analysis was performed by unpaired t-test, where HOb vs. HOS p=<0.0001, HOb vs. U-2 OS p=<0.0001 and HOb vs. SAOS-2 p=0.001, based on n=5-7. Expression was normalised to *GAPDH*, error bars represent SEM. B) Bar graph quantification showing significantly higher expression of *IGF-1R* mRNA in cancer cell lines, statistical analysis was performed by unpaired t-test, where HOb vs. HOS p=0.0017, HOb vs. U-2 OS is 0.0351, and HOb vs. SAOS-2 p=0.0044, based on n=5. Expression was normalised to *GAPDH*, error bars represent SEM. Attempts to knock down XRN1 in the primary HOb cell line to gauge characterisation of *IGF-1* and *IGF-R* expression were not successful.

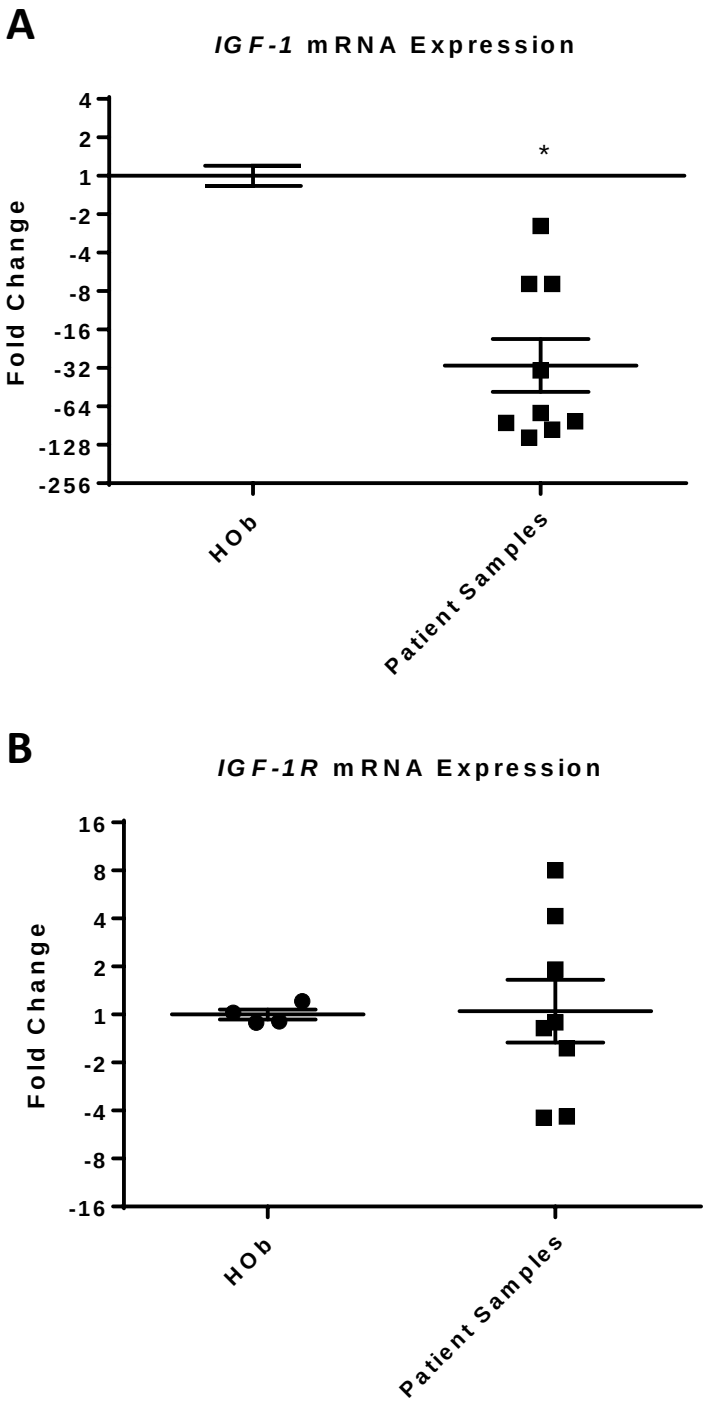


Figure 5.7. Quantification of *IGF-1* and *IGF-1R* mRNA expression in patient samples. A) Quantification of expression of *IGF-1* mRNA in patient samples showing lower expression when compared to HOb. Statistical analysis was performed by a Mann-Whitney test, HOb vs. patient samples p=0.0364, based on n=9. Expression was normalised to *Pes1*, error bars represent SEM. B) Quantification of *IGF-1R* mRNA expression in patient samples, showing no significant change compared to HOb. Statistical analysis was performed by a Mann-Whitney test, HOb vs. patient samples, based on n=9. Expression was normalised to *Pes1*, error bars represent SEM.

Cluster 1: EGF-like Enrichment score: 2.61	Cluster 2: PI3K and Notch signaling Enrichment scores: 1.94 & 1.61	Cluster 3: Golgi apparatus Enrichment score: 1.61	Cluster 4: Cell adhesion Enrichment score: 1.53	Cluster 5: Extracellular signaling Enrichment score: 1.50	Cluster 6: Cell migration Enrichment score: 1.47
NID1, EDIL3, IGFBP5,MEG F6, NOTCH3, EPHB2, TNC, RCN3, ADAMTS10, ITPR3, CELSR1, FBN2,VCAN, ITGA11, OLFML2A, FOXF2, NPTX2, CAPNS1, TGFB1I1, SLC25A6,CA CNB3, EHD3, IGFBP4, COL6A1, THBS2, EGFR, HSP90AA1	MYB, CSF1, BCL2, VEGFB, TNC, ITGA11, COL6A1, THBS2, EGFR, HSP90AA1, NEURL1B, MDK, DTX4, RBPJ, NOTCH3	EMP2, SCAMP1, ATP1B1, CLEC2D, SAR1B, SCARA3, B4GAT1, CHST14, KIF20A, HS6ST1, GLCE, GOLGA4, PJA2, CHST3, CHPF, MMGT1, NOTCH3, CSF1, EGFR	PTPRS, ALCAM, EDIL3, VCAN, TGFB1L1, EMP2, NEURL1B, NID1, TNC, ITGA11, COL6A1, THBS2,	C1QL1, GDF11, ENHO, GLO1, KDELC2, IL6ST, SEMA6B, SLC7A2, VASH2, APCDD1, TMEM170B, LYPD6, GPR137C, SEMA6C, ERMP1, GPR161, OLFML3, HHIPL2, RNF165, NRTN, SEMA4F, FGFRL1, FNDC10, VASH1, CD83, ABCA2, PANX3, SHISA5, PTPRS, ALCAM, MEG6F, EPHB2, RCN3, ADAMTS10, CELSR1, FBN2, OLFML2A, NPTX2, CAPNS1, EDIL3, IGFBP5, VCAN, IGFBP4, EMP2, CSF1, NID1, NOTCH3, TNC, ITGA11, COL6A1, THBS2, EGFR, B4GAT1	CORO1A, LHX2, TLE2, RDX, HAS2, HES4, SEMA6B, SEMA6C, NRTN, SEMA4F, EPHB2, B4GAT1, EGFR, CSF1

Table 5.6. Table of clusters of enriched, up regulated genes. The statistical cut off for enriched genes is when the enrichment score = 1.3, this table includes all clusters of genes for which the enrichment score is >1.3.

5.9 Bioinformatic prediction of transcriptional versus post-transcriptional changes in expression

Because the Cufflinks analysis extended only to the cumulative change of expression of the transcript as a whole, it was not discrete enough to indicate at what level the transcript was differentially expressed: transcriptionally or post-transcriptionally. FeatureCounts (Table 5.3) was used to sort the gene expression data by using an algorithm to pin point the location of each read to an individual intron and exon of each gene within the genome. In doing so, it was possible to gauge the expression changes of introns and exons within a specific gene, and by extension elucidate whether the differential up regulation of the top 20 genes was transcriptional or post-transcriptional, which would indicate targeting by XRN1. If the expression of the intron changed consistently with the expression of the exon, that would indicate a transcriptional change. If the expression of the intron did not change in line with the exon, that would indicate a post-transcriptional change, i.e. by sorting the data by up regulated pre-mRNA and up regulated mature mRNA, it was possible to tell which genes were strong candidates for targets of XRN1, an exoribonuclease responsible for the degradation of mature mRNA.

To control for this in the analysis, the expression of XRN1 was used because this gene was known to be significantly down regulated post-transcriptionally, because the mature mRNA was knocked down by siRNA. FeatureCounts calculated that the amount of intron mapping reads did not change (fold change was -1.06), whilst the fold change of exon mapping reads was -2.70 fold, showing that XRN1 is indeed down regulated post-transcriptionally, and is exactly as expected (Figure 5.8). Following this, FeatureCounts was used to determine the transcriptional and post-transcriptional changes in expression for all up regulated transcripts.

The specific fold changes of the mature (exonic reads) and pre-mRNA (intronic reads) of the top 20 genes were selected for analysis. In order to determine whether the expression of the intron was different enough from the expression of the exon to be classed as either transcriptionally or post-transcriptionally dysregulated, a threshold fold change difference was used. Any fold changes which were <1.5 different between intron and exon were classed as transcriptional changes, and any fold change differences between intron and exon that were >1.5 were classed as post-transcriptional changes. The threshold value of 1.5 was chosen because that is the fold change cut off used throughout this chapter to describe notable changes in expression. The results showed that of the top 20 up regulated genes, 11 were indicative of post-transcriptional up regulation, suggesting that of the top 20 genes, 11 are targets for degradation by XRN1, and therefore, by knocking down XRN1, these transcripts become up regulated. Of note, the most up regulated gene, NID1, was indeed post-transcriptionally up

regulated. One limitation to this is the possibility that because genes have introns, some reads may be detected in the introns as well as the exons. The full list of the top regulated genes and their intron and exon fold changes are listed in Table 5.7.

Down regulated genes were not analysed in the same way because any transcriptional or post-transcriptional down regulation of these genes are likely to be indirect effects which could conceivably be because a target of XRN1 is being up regulated, whether it be a transcription silencer or another exoribonuclease, the like of which has not been deduced by the current dataset, however, it would be possible in the future to look into this further.

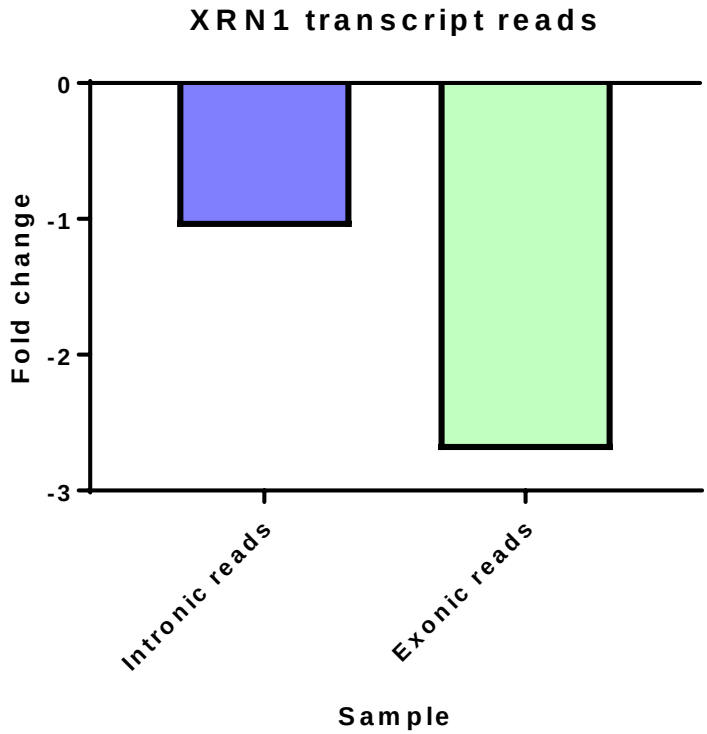


Figure 5.8. XRN1 is post-transcriptionally silenced by siRNA as expected. As a control for the analysis of transcriptional vs. post-transcriptional regulation of the up regulated transcripts, the algorithm was first run on XRN1, already known to be down regulated due to post-transcriptional repression by siRNA. The fold change of the introns was -0.94 (approximately 6% knock down) and the fold change of the exons was -2.70 (approximately 63% knock down). This is in line with the siRNA transfection technique used, and shows correct function of the algorithm.

Gene	Fold change of exons	Fold change of introns	Fold change difference	Transcriptional dysregulation	Post-transcriptional dysregulation
NID1	4.86	2.18	2.68		Yes
C1QL1	4.18	5.02	-0.84	Yes	
GDF11	2.47	2.21	0.26	Yes	
ENHO	4.08	0	4.08	N/A due to no intron mapping reads.	N/A due to no intron mapping reads.
AC008013.1	3.40	2.27	1.13	Yes	
HMOX1	3.57	3.60	-0.03	Yes	
GLO1	3.37	1.26	2.11		Yes
NUDT15	3.42	1.08	2.34		Yes
KDELC2	2.97	2.09	0.88	Yes	
CPNE2	3.61	2.43	1.18	Yes	
LRRC20	3.29	1.78	1.51		Yes
IL6ST	2.81	2.19	0.62	Yes	
SEMA6B	2.96	2.69	0.27	Yes	
EMP2	2.96	1.24	1.72		Yes
CORO1A	3.52	3.25	0.27	Yes	
PPP1R148	2.94	1.02	1.92		Yes
EDIL3	3.32	1.43	1.89		Yes
IGFBP5	3.01	2.26	0.75	Yes	
LHX2	2.39	2.65	-0.26	Yes	
SCAMP1	2.99	1.37	1.62		Yes

Table 5.7. Intron and exon fold change analysis on the top 20 upregulated genes from the RNA sequencing data as obtained by the Cufflinks pipeline.

5.10 Assessment of the read distribution to determine direction of decay in control cells

This part of the analysis was to assess the read distribution across the 5' UTR, coding sequence (CDS) and 3' UTRs of the genes in the dataset, which entailed identifying the proportion of the reads across the samples which mapped to each segment. The theory behind the following analysis was that it adds another layer of control to the overall analysis, in that transcripts being degraded by XRN1 could be distinguished by the proportion of reads mapping to the three segments of the transcript, due to the fact that XRN1 is a 5' exoribonuclease. In doing so, it was then sequentially possible to find out whether there were any differences between the amount of reads mapping to 3' UTRs, CDSs, and 5' UTRs in transcripts, which would identify whether the whole transcript changes during XRN1 knock down or whether only part of it is differentially regulated. To do so, FeatureCounts was used (as in section 5.9) to determine the number of reads which mapped to defined regions across each sample. The read count mapping to 3' UTRs was relatively consistent across all samples and ranged from between 13.4% to 16.9% of all reads within the sample mapping to 3' UTRs. FeatureCounts was then used to determine the percentage of reads which mapped to the 5' UTR, which showed that read mapping to the 5' UTR was also relatively consistent across all samples, where the percentage of reads assigned to the 5' UTR ranged from 1.9% - 2.4%. Interestingly, the percentage of reads mapping to the 5' UTR was marginally lower in XRN1 knock down samples (between 1.9 and 2.3%) than in the control samples (between 2.1 – 2.4%).

The same computational processing was then applied to find the proportions of reads in each sample which mapped to the CDS of the transcript. This showed that the percentage of assigned reads mapping to the CDS ranged between 18.6% - 22.9%. Again, the percentage of mapped reads in the knock down samples was lower (between 18.6% - 20.8%) than in the control sample (21.0% - 22.9%).

Following on from the mapping of reads to the genomic locations of the transcript, the proportion of reads mapping to the 5' UTR, CDS and 3' UTR in the control samples compared to the knock down samples was found, indicating whether or not there was a change between the proportion of reads mapping to each segment of the transcript upon knock down of XRN1. This analysis was undertaken on the top 10 genes which had previously been shown to be up regulated post-transcriptionally (see section 5.9). As such, it was assumed that transcripts in the process of being targeted by XRN1 would have a lower proportion of reads mapping to the 3' UTR in the control samples, than in the knock down samples, due to lack of degradation intermediates. Of the top 10 post-transcriptionally up regulated genes, there was proportional

mapping differences in 8 genes (Table 5.8), whereby the 5’ UTR seemed to have more reads mapping to it in the knock down samples than in the control samples. This offers a promising prospect for these genes as identifiable targets of XRN1, as they are not being degraded from the 5’ end when XRN1 is absent. There was also an example (EDIL3) where the CDS had more reads mapped to it than any other portion of the transcript in the knock down sample, compared to the control sample, in which the 3’ UTR showed the highest proportion of reads mapping to it.

Gene	% reads mapping in the 5'UTR (control)	% reads mapping in the CDS (control)	% reads mapping in the 3'UTR (control)	% reads mapping in the 5'UTR (knock down)	% reads mapping in the CDS (knock down)	% reads mapping in the 3'UTR (knock down)	Is there an increase in 5' UTR and/or CDS reads?
NID1	0.36	67.26	32.39	1.42	63.83	34.76	No
ENHO	13.46	33.33	53.21	39.48	21.13	39.39	Yes
GLO1	0.73	33.22	66.05	4.33	27.48	68.19	Yes
NUDT15	0.11	33.85	66.04	8.59	23.48	67.92	Yes
CPNE2	11.58	73.74	14.69	34.48	46.48	19.04	Yes
LRRC20	12.32	21.53	66.15	21.64	14.73	63.62	Yes
EMP2	2.78	17.74	79.48	6.93	9.36	83.72	Yes
PPP1R148	N/A	N/A	N/A	16.37	66.43	17.20	Poor UTR annotation
EDIL3	5.64	56.93	37.43	8.87	30.48	60.65	Yes
SCAMP1	6.52	36.21	57.66	10.87	20.08	69.05	Yes

Table 5.8. UTR and CDS percentage read mapping. Proportional read mapping conducted on the top 10 post-transcriptionally up regulated genes to decipher whether there was differential degradation taking place, as well as stabilization of the 5' end.

5.11 Analysis of the 3' UTR showed that there are common motifs among the up regulated genes

The 3' UTR (UnTranslated Region) of a transcript contains a number of regulatory sequences and structures, and the length of the 3' UTR can play key roles in the stability of the RNA. Regulatory sequences of the 3' UTR include (but are not limited to) miRNA binding sites, and RNA motifs to which RNA binding proteins bind to regulate the efficiency of a transcript's degradation. For example, substrates enriched with guanosine regions are associated with slower degradation of the substrate by DIS3L2 in *Drosophila* models (Reimão-Pinto *et al.* 2016). This is thought to work in tandem with the uridyltransferase, Tailor, which demonstrates high uridylation activity upon recognition of 3' G (Cheng *et al.* 2019). In contrast, highly uridylated transcripts are associated with rapid degradation by DIS3L2 (Menezes *et al.* 2018). Another example is adenylate-uridylate (AU)-rich elements. AU-rich elements (AREs) are present in 58% of the human transcriptome and are often found in intronic regions (Bakheet *et al.* 2018). AREs interact with RNA binding proteins, such as HuR, to regulate the stability of the transcript. HuR binding, for example, leads to stabilization of the transcript for gene expression. 3' UTR processing is an extremely important aspect of gene expression regulation, as this can determine exactly how a gene is regulated and may determine whether a transcript is destined for degradation by a specific pathway.

It was decided that due to the highly important regulatory nature of the 3' UTR, it would be sensible to look for commonalities between the 3' UTRs of the up regulated genes. This included additional analysis to identify any enriched regulatory motifs which may direct XRN1 mediated decay. For this, it was important to extricate all post-transcriptionally up regulated genes from the total amount of up regulated genes, for which differential expression fold change was >1.5. This selection, using FeatureCounts, resulted in the identification of 59 genes showing predicted post-transcriptional increases in expression. Using Ensembl, the 3' UTR sequences for each gene was found. In the case that there were multiple potential sequences available, the chosen sequence was selected based on concurrence between three reference databases: CCDS, UniProt and RefSeq. The selected sequence had to concur with all three, and in the case of multiple sequences concurring, the longest sequence was selected.

Each sequence was measured, which showed that the average 3' UTR length of all post-transcriptionally up regulated genes was 3017.76 bases. In comparison to the average length of 3' UTRs across the human genome (1200 nucleotides), the average length of the 3' UTR in post-transcriptionally up regulated genes is much longer, which suggests that they are more susceptible to post-transcriptional regulation (Gruber *et al.* 2016). This might not be surprising

given that these genes are predicted to be regulated post-transcriptionally by XRN1. The sequences were then analysed for common motifs using the online tool, MEME Suite (Bailey *et al.* 2009), to find out whether the potential targets of XRN1 have a common signature for degradation, or specific binding sites. The sequences were submitted to the programme based on analysis with stringent selection criteria. A motif was only deemed as common amongst the 3' UTRs if it appeared >10 times. In addition, the number of motif repetitions within one 3' UTR was calculated, to identify sequences which had more than one repetition of the same motif, which is to be expected within UTRs of longer lengths. The 5 most common motifs are shown in Figure 5.9. The concurrence between the 5 motifs are listed below:

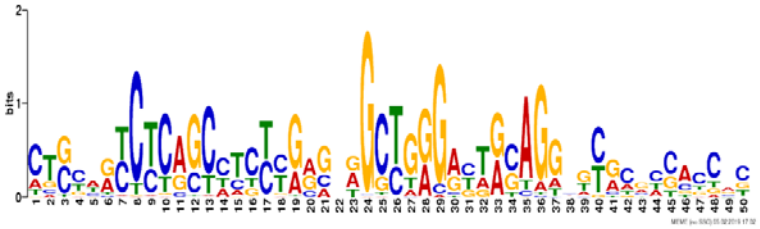
- Motif 1 was shown to be concurrent between 21 genes,
- Motif 2 was shown to be concurrent between 26 genes,
- Motif 3 was shown to be concurrent between 34 genes,
- Motif 4 was shown to be concurrent between 34 genes,
- Motif 5 was shown to be concurrent between 33 genes.

The motifs identified include motifs with long regions of repeated bases. In the top two motifs with common regions identified in 34 genes, the repeated regions were regions of polyuridyl and poly-cytosine regions. It was interesting that a polyuridyl motif was the joint most common motif because this motif is known to promote turnover for many types of mRNA which have terminally polyadenylated regions (Mayr 2016), and therefore it would be expected to be present genes targeted for degradation. DIS3L2 is known to target highly uridylated transcripts and so this supports the notion that because XRN1 and DIS3L2 co-precipitate (Lubas *et al.* 2013), they are capable of targeting the same transcripts. Polyadenylation was also a very common motif, and whilst this is to be expected within 3' UTRs (to influence the stability of the transcript), the fact that most of the genes identified displayed both motifs of polyuridylation and polyadenylation suggests that these particular transcripts were perhaps pre-disposed for degradation, which simply wasn't occurring due to knock down of XRN1. Indeed, a large proportion of the genes were not detected as having common motifs, and the same genes were identified repeatedly throughout the analysis as having these 5 motifs. These genes included, and were not limited to, GLO1 and EMP2, which were the 4th and 7th most post-transcriptionally up regulated genes extricated from the data respectively. It is therefore possible that these motifs are involved in targeting the transcript to XRN1-mediated decay. Indeed, the majority of the genes which are identified concurrently are the same across all identified motifs, with specific reference to SMAD2. This is probably due to the extremely long length of the 3' UTR of SMAD2 (32.870kb), making specific motifs more likely to be present.

Another interesting motif identified here is the presence of G-rich regions. Early work conducted within the RNA stability field showed that XRN1 is inhibited by G-rich regions due to the structure of the G-quadruplexes, however, more recently it has been shown, using computational studies, that XRN1 actually preferentially degrades RNAs which have a G-rich region (Huppert *et al.* 2008). The fact that the G-rich motif is seen in this RNA sequencing dataset supports this notion. Indeed, this motif was present in 26 out of 59 post-transcriptionally up regulated genes, providing strong evidence that these genes are being targeted by XRN1.

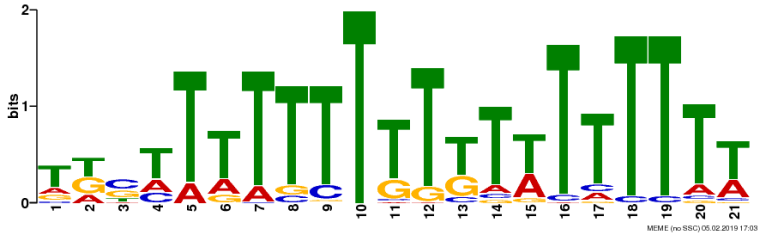
The variety of the motifs identified in this section suggest that there are multiple motifs which could promote degradation of the transcript by XRN1. It would be interesting to analyse the RNA binding proteins (RBPs) associated with these motifs, as they are also important in regulating RNA stability. For example, the RBP, HNRNCPC, is known to bind to polyuridylated regions, whilst also peaking at polyadenylated regions in order to regulate 3' end processing (Gruber *et al.* 2016). It would be interesting to see if this protein displays interaction with the exoribonucleases to regulate RNA stability. In addition, the RBP, HuR, which is a protein which shuttles between the nucleus and the cytoplasm to bind RNA in response to stress, is known to interact with poly(U) tracts. Interestingly, HuR has been shown during Hepatitis C virus (HCV) infection to displace the poly(U) tract binding protein (PTB) in order to facilitate viral replication (Shwetha *et al.* 2015). This is particularly interesting because XRN1 has also been implicated in HCV as way for the virus to evade the host response. This suggests that there is a deeper interaction between XRN1 and RBPs than originally thought, and more analysis should be conducted to map this.

Motif 1

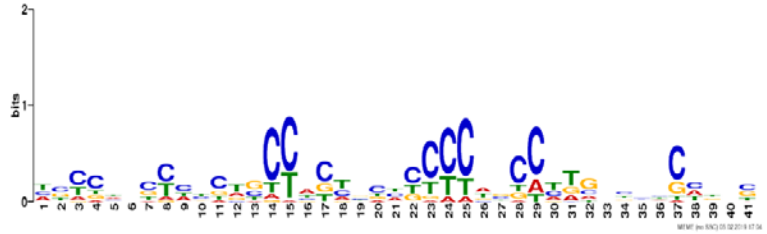


Motif 2

Motif 3



Motif 4



Motif 5

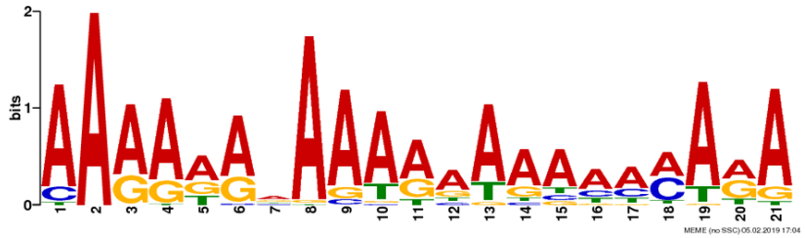


Figure 5.9. Common motifs in 3' UTRs found on MEME Suite. The above motifs were identified (using MEME Suite) as common regions between the 3' UTRs of the 59 post-transcriptionally up regulated genes in the RNA sequencing data.

5.12 Analysis of down regulated genes

The same filtering system used to sort the differentially up regulated genes was also applied to sort the down regulated. Genes which were >1.5 fold down regulated were first sorted, and then sorted by consistency between the individual replicates, before being finally sorted by FPKM. All genes with a FPKM value of <1 were excluded from the analysis (see Table 5.4). Out of the whole RNA sequencing dataset for this experiment, 668 genes were significantly down regulated. Of these, 366 were down regulated by >1.5 fold change, and of these 318 had a FPKM value of >1. The functions of the top 20 most down regulated genes are displayed in Tables 5.9A and 5.9B. GO analysis of the 318 genes showed 7 significantly enriched clusters (an enrichment score of >1.3). The genes in these 7 clusters were sorted into 4 categories: Cell adhesion and motility, ER and Golgi apparatus, Transmembrane and Lipid metabolism. In line with the qRT-PCR confirmation of XRN1 knock down, the RNA sequencing dataset output from Cuffdiff shows a cumulative down regulation of XRN1 in these cells as well, of 1.69 fold change. However, intronic versus exonic expression data shows that the down regulation of exons is -2.7 fold (Figure 5.8), in line with the siRNA transfection technique, which inhibits cytoplasmic mature mRNA, rather than nuclear pre-mRNA. In the top 20 most down regulated genes, the most frequent functions involve the regulation of cell adhesion and motility, which suggests a role for XRN1 in affecting cell migration.

Interestingly, there are two clusters of enriched genes which feature in both the up regulated gene set and the down regulated gene set, suggesting that regulation of these aspects of cell maintenance is in part due to XRN1. The categories are: A) the categories for cell adhesion and cell migration in up regulated genes, and cell adhesion and motility in down regulated genes, B) Extracellular signaling in up regulated genes and Transmembrane in down regulated genes. Genes involved with the Golgi apparatus are also consistently mis-expressed between the up regulated and down regulated genes. Figure 5.10 shows the separation of the enriched gene clusters between the up and down regulated genes

The GO analysis shows an enrichment of down regulated genes encoding transmembrane proteins. This supports a role of XRN1 in cell-cell signaling, and corroborates recent reports in which XRN1 has been localized to the eisosome granules in yeast (Grousl *et al.* 2015, Vaskovicova *et al.* 2017). Eisosomes are plasma membrane domains, which concentrate lipids, transporters and signaling molecules, and mark the site of endocytosis (Olivera-Couto *et al.* 2015). The genes in each category are tabulated in Table 5.10. If XRN1 is involved in extracellular signaling in SAOS-2 cells, this adds another conserved role for XRN1 from at least single-celled eukaryotes through to humans. It is important to remember, however, that

although XRN1 could be inducing these effects, it could still be indirect, and not caused by XRN1 down regulation itself.

Gene	Full name	Molecular Function	Biological process	Log2 Fold change
ENO1	Alpha-enolase	Involved in glycolysis, specifically synthesizing pyruvate from D-glyceraldehyde.	Growth control, hypoxia intolerance, and allergic responses. Stimulates immunoglobulin production.	-1.05
H3F3B	Histone H3.3	Constitutes the predominant form of H3 in non-dividing cells and incorporated into chromatin independently of DNA synthesis.	Represents an epigenetic imprint of transcriptionally active chromatin. Transcriptional regulation, DNA repair, DNA replication and chromosomal stability.	-0.74
RPL11	60S ribosomal protein L11	Large component of the ribosome. Part of the 5S/RNP it Is essential for the large subunit activity of formation and maturation of rRNAs.	Couples ribosome biogenesis to p53 activation. As part of the 5S RNP it accumulates in the nucleoplasm and inhibits MDM2 to mediate activation of p53.	-0.74
ATP5PB	ATP synthase F (0) complex subunit B1, mitochondrial	Produces ATP from ADP during respiration, and has proton transmembrane transporter activity.	ATP biosynthesis and hydrolysis.	-0.78
CFL1	Cofilin-1	Regulates actin skeleton dynamics, binds F-actin.	Cytoskeleton organization, required for neural tube morphogenesis and cell migration.	-2.19
CERS2	Ceramide synthase 2	DNA binding, sphingosine N-acyltransferase activity.	Suppresses growth of cancer cells and may be involved in sphingolipid synthesis.	-0.98
BZW1	Basic leucine zipper and W2 domain-containing protein 1	Cadherin binding and RNA binding.	Enhances histone H4 transcription, but does not bind DNA directly.	-0.73
MYL12B	Myosin regulatory light chain 12B	Binds calcium.	Important role in regulation of both smooth muscle and non-muscle cell contractility via phosphorylation.	-0.84
PRSS23	Serine protease 23	Serine-type endopeptidase activity.	Cellular metabolism and post-translational modification.	-0.78
S100A2	Protein S100-A2	May function as a calcium sensor and modulator, contributing to calcium signaling.	Roles in many physiological processes, including to possibly suppress tumour growth.	-1.32

Table 5.9A. List of functions for the top 20 down regulated genes, top 1-10 featured here. Genes continued in Table 8B.

Gene	Full name	Molecular Function	Biological process	Log2 Fold change
PLAC8	Placenta-specific gene 8 protein	Chromatin binding.	Positive regulation of proliferation, cold-induced thermogenesis, transcription. Negative regulation of apoptosis.	-0.94
HMGN1	Non-histone chromosomal protein HMG-14	Alters interaction between DNA and histone octamer and maintains active genes in a unique chromatin conformation.	Chromatin organization, positive regulation of transcription and elongation. Regulates development and proliferation and response to UV-B and C.	-0.90
TNPO1	Transportin-1	Nuclear transport receptor for NLS, thought to mediate docking of the importin to the nuclear pore complex.	Involved in cilium assembly, protein > nucleus import, translocation and regulating mRNA stability.	-0.85
UXT	Protein UXT	Involved in transcription regulation. Regulates androgen-receptor mediated transcription, and NF-kappa-B.	Microtubule cytoskeleton organization, and transport, apoptosis, negative regulation of G0>G1 transition and transcription.	-0.98
HAPLN1	Hyaluronan and proteoglycan link protein 1	ECM constituent, hyaluronic acid binding.	Cell adhesion, CNS development and ECM organization.	-0.80
CCDC80	Coil-coil domain-containing protein 80	Fibronectin binding and heparin binding.	Promotes cell adhesion and matrix assembly, ECM organization.	-0.78
SSR3	Translocon-associated protein subunit gamma	Binds calcium to the ER membrane	SRP-dependent co-translational protein targeting to membrane.	-1.15
YARS	Tyrosine-tRNA ligase	Catalyses the attachment of tyrosine to tRNA. Also, IL-8 receptor binding.	Apoptosis and tRNA aminoacylation.	-1.30
SLC9A3R2	Na(+)/H(+) exchange regulatory cofactor NHE-RF2	Scaffold protein, to connect plasma membrane proteins to help link them to the actin cytoskeleton.	Regulates surface expression of plasma membrane proteins.	-0.92
YWHAZ	14-3-3 protein zeta/delta	Adapter protein implicated in many signaling pathways and has many binding partners.	Maintenance of the Golgi apparatus, platelet activation, apoptosis, mRNA stability, signal transduction and synapse maturation.	-1.58

Table 5.9B. Top 11-20 down regulated genes as shown in the RNA sequencing.

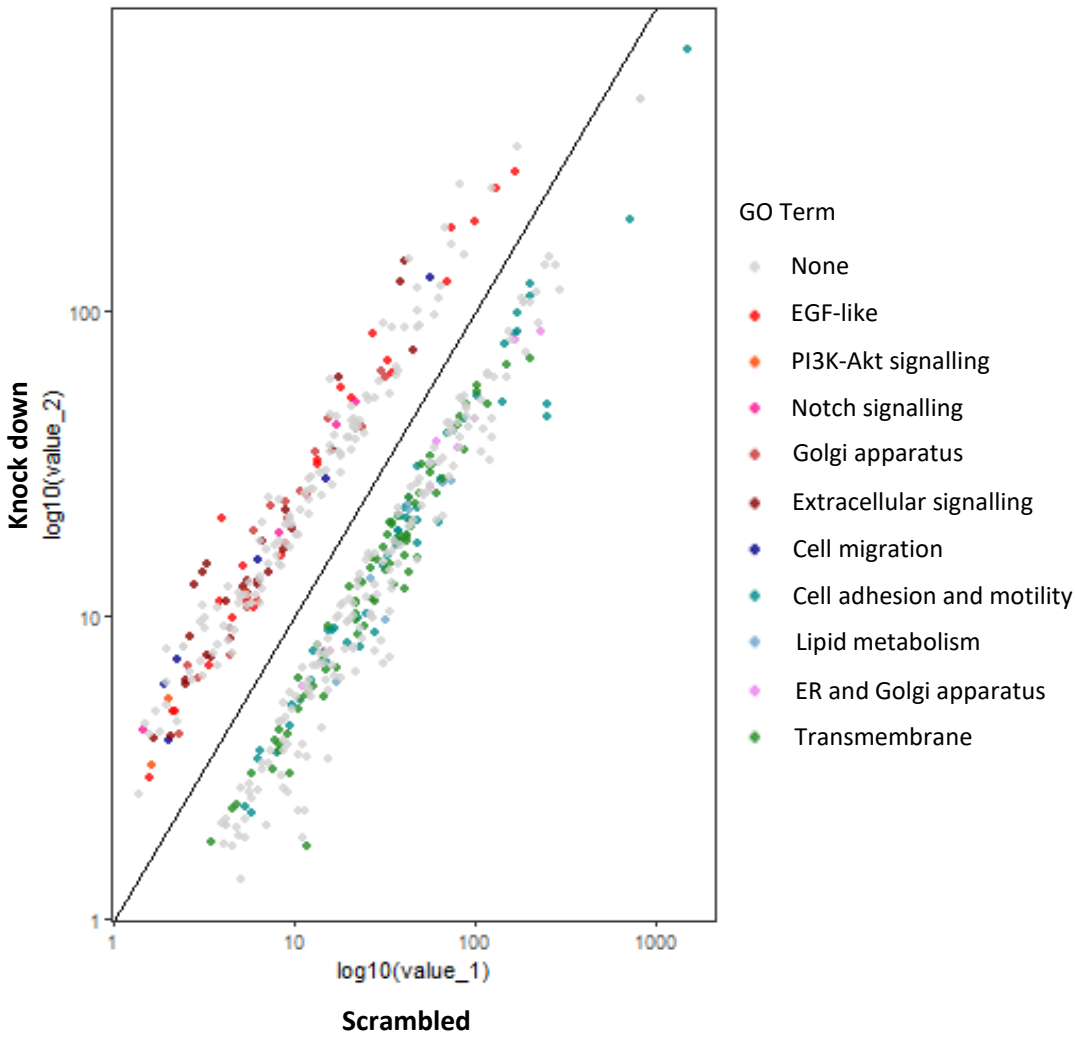


Figure 5.10. Graph to show the separation of enriched gene clusters. Genes were first sorted by selecting only those whose differential expression was >1.5 fold change. Next, they were sorted by consistency between their individual replicates to remove those which were not consistent. Lastly, the genes were sorted based on their FPKM value being >1. The final lists of up and down regulated genes were submitted for GO analysis, and enriched gene clusters were obtained. The graph shows the enrichment of specific gene clusters in both up (above the xy line) and down (below the xy line) regulated genes.

Cluster 1: Cell adhesion & motility Enrichment score: 4.32 & 1.67	Cluster 2: Lipid metabolism Enrichment score: 2.00	Cluster 3: ER & Golgi apparatus Enrichment score: 1.62	Cluster 4: Transmembrane Enrichment score: 1.32
ENO1,CFL1, BZW1, MYL12B, CCDC80, SLC9A3R2, ITGA3, CAPZA1, CD164, CAPN2, GNG5, FERMT2, SH3GL1, ITGA4, CYR61, STAT1, FAM129B, ZC3H15, CLINT1, ITGAV, PPP2R1B, CAV2, TWF1, ELK1, GNA13, GNG12, STK38L, CLMP, CD36, CTNNAL1, TFPI, CXADR, PXN, SDC4, SDCBP, TES, KRAS, EHD4, GNG2, IL7R, ARHGAP18, HBEGF, PERP, VAMP4, ITGB8, CTGF, SLC7A7, ULBP2, DSG2, PRKAB2, PRKAA1	PRKAB2, HACD1, PRKAA1, DEGS1, MTMR2, SCD, PTGES3P1, AGPAT3, ADIPOR1, LYPLA, GM2A	COPB2, DYNLL1, BCAP31, SAR1A, TMED2, LMAN1, RAB2B, STX17, VAMP4, BNIP1	ENO1, CFL1, CCDC80, SLC9A3R2, ITGA3, CD164, CAPN2, GNG5, FERMT2, SH3GL1, ITGA4, FAM129B, CLINT1, ITGAV, GNA13, GNG12, STK38L, CLMP, CD36, CTNNAL1, TFPI, CXADR, SDC4, SDCBP, KRAS, EHD4, GNG2, IL7R, HBEGF, PERP, VAMP4, ITGB8, SLC7A7, ULBP2, DSG2, HACD1, DEGS1, MTMR2, SCD, AGPAT3, ADIPOR1, SSR3, AP2S1, SEC11A, MFSD1, SLC25A39, SHMT2, ADD3, TM9SF2, SNX4, SEC11C, RNF167, ARHGAP28, KARS, EFR3A, CCNY, KCTD20, TRAM1, DPY19L1, MRS2, GLIPR1, PGRMC1, PEF1, UQCC1, PEX6, MIGA1, TOMM22,TMEM167B, TMEM263, NECAP2, VPS25, ASAH1, SNX12, TMEM33, OSBPL8, GDE1, CLIP1, RIC1, SLC2A1, AP1M2, TM2D1, FAM120A, CPD, TMEM208, SCYL2, NOC4L, SERINC3, IL31RA, TMEM60, GPR183, ERLEC1, PHLDA3, TMTC3, SLC35F2, PORCN, PLEKHA8, GLT8D2, CRNP1, UNC13B, LIFR, SYBU, APOL6, GLIPR2, E2F5, DIRC2, SLC17A5, SLC31A1, IFIT5, FAM234B, RAB22A, MOSPD1, TNFRSF21, RPIA, SLC46A3, ANO2, PIGM, ACVR2B, GPR157, SYNPO

Table 5.10. Table of clusters of enriched, down regulated genes. The statistical cut off for enriched genes is when the enrichment score = 1.3, this table includes all clusters of genes for which the enrichment score is >1.3.

5.13 Comparisons of differential gene expression after XRN1 knock down in two different cell lines

Alongside filtering the results by fold change and FPKM values, it was decided to compare the data set with another data set from another paper in which the authors had knocked down XRN1 in a different cell line. To identify specific targets of XRN1 which may be conserved between mammalian cells, a comparison between differentially expressed transcripts in SAOS-2 cells and those in HeLa cells was performed. Supplemental data (raw RNA sequencing fold change) from a publication by the Dziembowski group showed the differential expression of 1400 genes upon XRN1 knock down in HeLa cells (Lubas *et al.* 2013). There are limitations to this comparison, including experimental differences in the way that XRN1 was depleted in the HeLa cells. This paper uses a much higher concentration of siRNA to deplete XRN1, which may contribute to differences between the datasets, as well as differences in data processing (see section 5.15).

The differential fold changes from the SAOS-2 cells were directly compared to those in the HeLa cells to look for consistently mis-expressed genes across two different cell types. The result of this comparison showed that there were 101 genes in total that were differentially expressed in both cell lines after XRN1 knock down. There were also examples of some genes being up regulated in SAOS-2 cells which were down regulated in HeLa cells, which could, arguably, be attributed to differences in cell type. Of the 101 total genes, there were 71 genes which were consistently up regulated after XRN1 knock down in both cell lines. These were then filtered based on fold change and FPKM value to increase the confidence in the up regulation, which showed that of the 71 up regulated genes, there were 19 up regulated genes which displayed high confidence based on their fold change and FPKM (Figure 5.11). These genes could represent more 'globally' sensitive transcripts whilst the others are likely to demonstrate an element of tissue/cell type specificity. Of note, the presence of EMP2 in this dataset shows that some of the post-transcriptionally regulated genes found in the OS cells are also dysregulated in HeLa cells.

The same was applied to down regulated genes, which showed that of the 101 consistently expressed genes, 4 were consistently down regulated after XRN1 knock down in both SAOS-2 and HeLa cell lines (Figure 5.12). Again, the genes which were most down regulated between the 2 cell lines differed. In SAOS-2 cells, the most down regulated gene which was also down regulated in HeLa was SDCBP, whereas the most down regulated gene in HeLa that was also down regulated in SAOS-2 was DSG2.

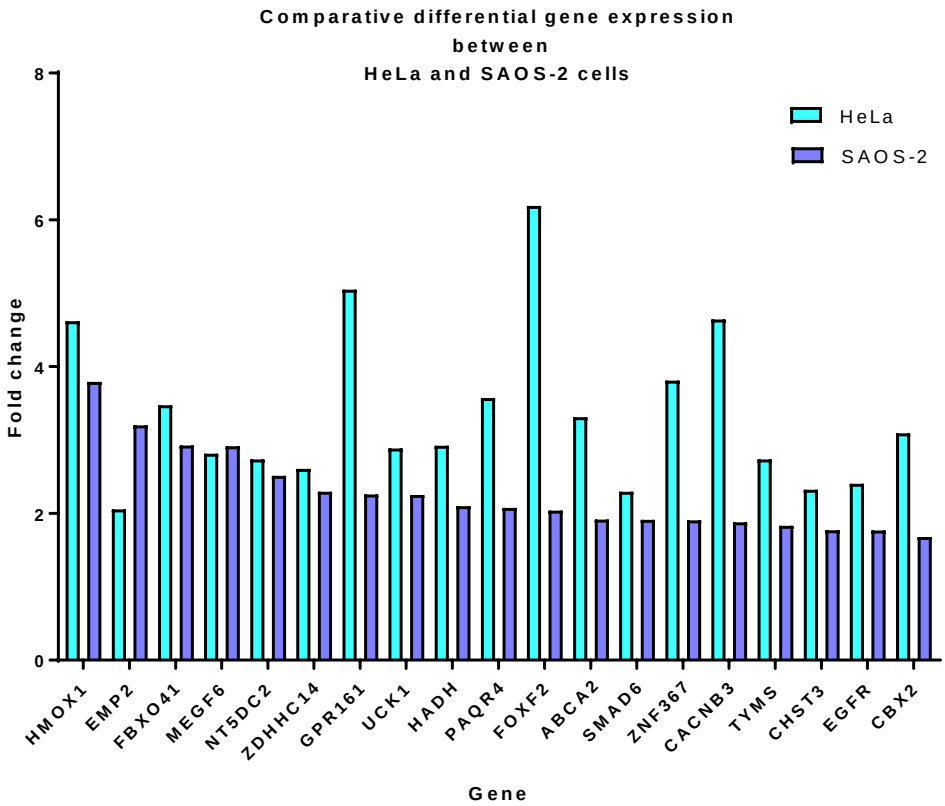


Figure 5.11. Comparative differential up regulation in gene expression between HeLa and SAOS-2 cells. Bar graph to show the raw fold change up regulation of genes which are consistently up regulated after XRN1 knock down in both HeLa cells and SAOS-2 cells in RNA sequencing experiments performed by 2 different labs (Lubas, Damgaard *et al*. 2013).

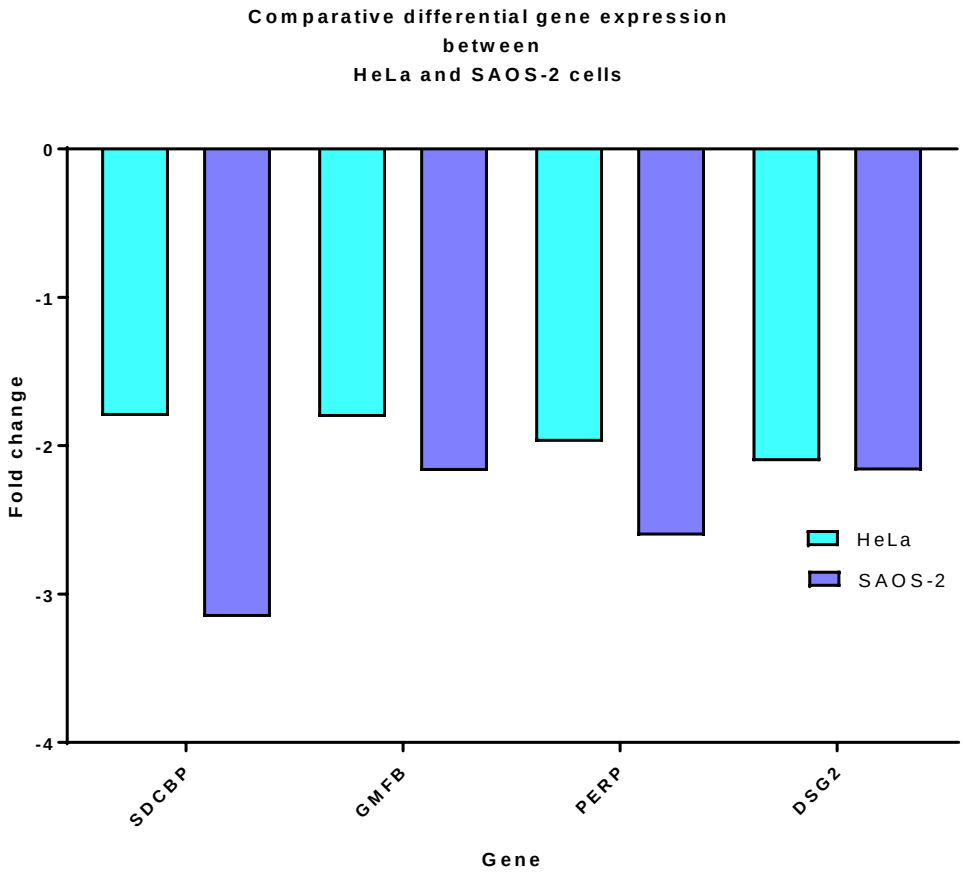


Figure 5.12. Comparative differential down regulation in gene expression between HeLa and SAOS-2 cells. Bar graph to show the raw fold change down regulation of genes which are consistently down regulated after XRN1 knock down in both HeLa cells and SAOS-2 cells in RNA sequencing experiments performed by 2 different labs (Lubas, Damgaard *et al.* 2013).

5.14 Validation of up regulated mRNAs using qRT-PCR

After all the RNA sequencing analysis had been undertaken, it was initially decided to validate six of the differentially up regulated genes. This was, in part, due to assay expense. The six genes were chosen based on the following criteria:

1. Extent of up-regulation,

2. Whether this up regulation was as a result of accumulation of exon or intron reads (i.e is the gene transcriptionally or post-transcriptionally regulated),

3. Fold change of the up regulation (i.e. <1.5 at the intron level and >2 at the exon level). Based on the above (Table 6), the following genes were selected for initial validation:

NID1: The most up regulated gene in response to XRN1 knock down and is post-transcriptionally regulated.

GLO1: Post-transcriptionally regulated, where intron fold change is <1.5 and exon fold change is >2.

NUDT15: Post-transcriptionally regulated, where intron fold change is <1.5 and exon fold change is >2.

EMP2: Post-transcriptionally regulated, where intron fold change is <1.5 and exon fold change is >2. It is also differentially up regulated in HeLa cells upon knock down of XRN1.

PPP1R148: Post-transcriptionally regulated, where intron fold change is <1.5 and exon fold change is >2. This also needs validating due to the poor annotation of the UTRs (Table 5.7).

EDIL3: Post-transcriptionally regulated, where intron fold change is <1.5 and exon fold change is >2. This gene was also of interest because of its EGF-like activity as identified by the GO analysis.

These genes were validated to ensure that they were not false positives as a result of the bioinformatics analysis. Another level of control to ensure that the analysis of intron and exon fold change also means that the pre-mRNA of these genes will also be looked at during the validation step. If the pre-mRNA changes in contradiction to how the bioinformatics has suggested, and any of these genes is differentially expressed at the level of transcription, the gene in question will be eliminated as a potential target of XRN1.

Validation of these genes using qRT-PCR was undertaken during the writing of this project and initial results are shown in Figure 5.13. Preliminary results show that at least five out of the six chosen genes are up regulated in the XRN1 knock down samples when compared to the scrambled control. The gene which did not show significant up regulation was PPP1R148. There are two possible explanations for this: 1. There could be a problem with the assay used during validation or 2. It is likely that this gene is not being targeted by XRN1. In the event of the latter being true, PPP1R148 should be excluded as a potential target of XRN1. These results will be further validated in the future by observing the change in expression of the potential targets when XRN1 is over expressed. If the potential targets are being regulated by XRN1, it would be expected that the expression of these genes would decrease when XRN1 is over expressed. These genes will also be tested in HEK-293T cells after XRN1 knock down, to elucidate whether XRN1 is targeting the same transcripts in different cell types, or whether it functions to degrade different transcripts in different cell types.

Validation was directed by the author and performed by an MSci student, under full supervision of the author during the writing of this thesis.

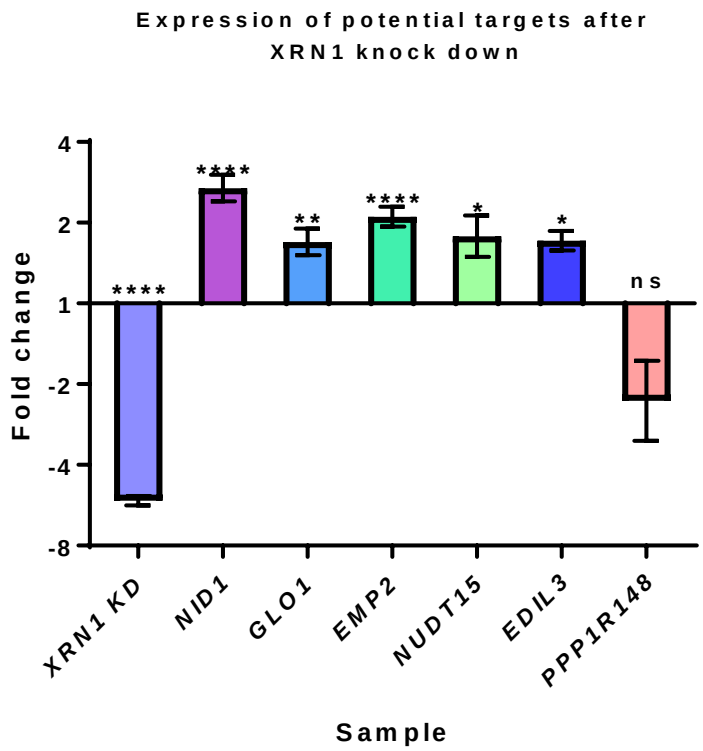


Figure 5.13. Preliminary validation of potential transcript targets of XRN1 in SAOS-2 cells. Bar graph to show qRT-PCR validation of potential targets identified by RNA sequencing. Samples were normalized to *GAPDH*. Statistical analysis was performed by unpaired t-test, where *XRN1* knock down vs. scrambled control p=<0.0001, *NID1* vs. scrambled control p=<0.0001, *GLO1* vs. scrambled control p=0.0028, *EMP2* vs. scrambled control p=<0.0001, *NUDT15* vs. scrambled control p=0.0435, *EDIL3* vs. scrambled control p=0.0105 and *PPP1R148* vs. scrambled control p=0.0597. Error bars show SEM, based on n=6.

5.15 Discussion

This RNA sequencing experiment has shown that the data obtained by RNA sequencing has an enormous potential for extensive data mining. The work in this chapter has demonstrated that when XRN1 is knocked down in SAOS-2 cells, substantial differential expression of up regulated and down regulated genes occurs when compared to the scrambled siRNA control. It has been shown that genes which are up regulated are involved in cancer associated pathways, including the EGF pathway, whilst those that are down regulated display high enrichment in cell adhesion and extracellular signaling pathways.

Genes which were up regulated were also analysed for potential transcriptional or post-transcriptional up regulation, which showed that 11 of the top 20 differentially expressed genes were regulated at the level of post-transcription, adding strength to argument for their targeting by XRN1. In addition, the RNA sequencing dataset obtained for this work was compared to the RNA sequencing dataset published by the Dziembowski group on HeLa cells which had undergone XRN1 knock down. This comparison identified genes which were simultaneously dysregulated in both cell lines, suggesting that these genes are regulated by XRN1 in multiple cell types. One of the limitations of this comparison include the fact that the dataset used had been processed by the Dziembowski group, and their exact method of data processing was not the same as the method used in this project. For example, in the experiment described in this thesis, Cufflinks algorithms were used to process all the data, but the bioinformatics approaches used in the Lubas *et al* publication included the use of the program, Galaxy, to process the RNA reads, and DESeq to identify differentially expressed genes. They also used the program TopHat, which is an older, less accurate alignment tool than HiSat2, to align genes to the reference human genome. In addition, a different version of the human genome was used as a reference to map and align reads (hg19/GRCh37), and so there is a chance that there is some disparity between the two sets of results (Lubas *et al.* 2013).

Ideally, paired-end sequencing would be used for RNA sequencing, because it allows for a more in-depth analysis of the RNA sequences present. Paired-end sequencing involves sequencing both the 5’ end and the 3’ end of the DNA fragments prepared in the library preparation. Another advantage to paired-end sequencing is the removal of PCR duplicates created during the library preparation, which are common artifacts resulting from PCR amplification during this step. Paired-end sequencing is also preferential because it allows for the detection of common DNA rearrangements (such as insertions, deletions and inversions) and gene fusions in cancer, alongside novel splice isoforms. In addition, paired-end sequencing

produces high quality alignment across repetitive sequences. However, paired-end sequencing on the NextSeq was too expensive for this project.

The limitations in this chapter include the power of the replicates. Generally, to increase the power of any experiment, as many replicates as possible should be conducted. In this case, due to expense, only 6 replicates were used for RNA sequencing. This is many more replicates than is usual in most published experiments, which usually use two or three, however, if more replicates were used, more genes would have been filtered out of the selection process due to inconsistencies between the individual replicates, leaving a more refined selection of differentially expressed genes.

It would also have been preferable to conduct this experiment on multiple cell types, to determine if XRN1 is targeting the same transcripts in a variety of tissues, or if XRN1 targets different transcripts in different tissue types. It is perfectly plausible that XRN1 has different targets in different cells in order to maintain the correct function of those cells in their respective tissues. As displayed in Chapter 4 by DIS3L2, the potential for exoribonucleases to have tissue-specific function, such as DIS3L2, may also be reflected by XRN1. By analyzing differential expression of transcripts after XRN1 knock down across multiple cell types, it would be possible to determine whether this is indeed the case. As seen in Figure 5.11, XRN1 knock down in both SAOS-2 cells and HeLa cells results in 19 shared differentially expressed genes between the two cell types, however, it is well documented that HeLa cells are not ideal for this kind of experiment, due to their extensive chromosomal abnormalities (Mittelman and Wilson 2013). Therefore, it would be better to have conducted this experiment on cell lines of a more stable nature, and then seen how the results align. To start with, it is possible to test the misexpressed genes in the other OS cell lines. If the genes are differentially expressed in the other OS cell lines which display natural down regulation of XRN1, this would confirm that the effects of XRN1 knock down are not artifacts of the experiment, and are, in fact, most likely targets of XRN1. This could also be assessed by qRT-PCR in other cell types to gain an idea about how targets might differentiate between cell types.

Another limitation of this experiment which has recently come to light is the possibility of the introduction of siRNA affecting the expression of miRNAs in the cell. It has been reported that the introduction of high levels of siRNA during knock down experiments can compete with endogenous miRNAs (Khan *et al.* 2009), which could affect how efficient the knock down of XRN1 is in this case. This group also showed that siRNA transfection could result in the up regulation of certain genes, which could occur as a result of competition between the siRNA and endogenous miRNA for intracellular machinery responsible for processing small RNAs. In turn

this leads to impaired gene repression by miRNAs, and gene up regulation. It is unknown how this might have affected the outcome of this RNA sequencing experiment, although knock down was optimised to ensure that the lowest concentration of XRN1 siRNA was used in this experiment. It could also be assumed that the scrambled siRNA control would also have this effect, and so there can be confidence that the differentially expressed genes are as a result of defective degradation by XRN1, and not that they are usually subject to regulation by miRNA-mediated decay.

Ideally, more of the genes identified as being potential targets for degradation by XRN1 would have been validated by qRT-PCR. Any data obtained bioinformatically should be tested physiologically to ensure that any false positives are identified. This validation is currently being undertaken in the lab, however, the results of the validation are preliminary. Alongside validation of the potential targets, it is also possible to design pre-mRNA assays in order to make sure the intron and exon data analysis performed does not also harbour false positives.

In addition to this work yet to be fulfilled, also in line for testing are the RNA decay rates of those potential targets for XRN1. This would show, using a different method, that the targets are being degraded by XRN1. By measuring the decay rates of the potential targets with and without XRN1 knock down, it will give support to the hypothesis that these targets are normally degraded by XRN1. It would be expected that those RNAs being targeted by XRN1 would have longer half-lives in the XRN1 knock down cells compared to the control cells, due to the lack of degradation. The experiment would first utilize transfection of XRN1 siRNA, followed by blocking transcription in the cells using actinomycin D (which inhibits initiation of transcription by intercalating into the minor groove of DNA and preventing binding of topoisomerases (Koba and Konopa 2005), or α-amanitin (which inhibits transcription by binding and inhibiting RNA polymerase II (Casse *et al.* 1999)). This allows for the measurement of RNA stability. Transcription inhibition is then followed by performing time-course dependent PCR, to detect the RNAs of interest. This allows the measurement for RNA half-life, and so the decay rate can be calculated (Chen *et al.* 2008).

Given that there are more differentially down regulated genes than differentially up regulated genes, it would suggest that a main target for XRN1 is part of the transcription machinery, perhaps a transcription factor. The idea that XRN1 itself is a transcription factor also cannot be ruled out due to the extensive number of down regulated transcripts in the knock down samples. This supports the notion described by other publications that XRN1 is a transcriptional activator in yeast (Medina *et al.* 2014).

Alongside the analysis of the up regulated genes, it was interesting that no autophagy genes were found to be up regulated among the dataset. Recent reports describe a relatively novel role for XRN1 in post-transcriptional control of autophagy, a major pathway associated with cancer progression. Previous publications suggest that XRN1 acts as a negative regulator of autophagy, whereby depletion of XRN1 in both *Saccharomyces cerevisiae* and HeLa cells results in up regulation of autophagy genes and autophagosome formation (Delorme-Axford *et al.* 2018). Autophagy is the cellular process in which during starvation conditions the cell degrades and recycles dispensable organelles and other cellular components in an orderly fashion in order to promote cell survival by maintaining homeostasis. It can also act as another programmed cell death pathway (PCD II) in which it works to induce cell death when the cell does not have the capacity to induce canonical apoptosis (Yonekawa and Thorburn 2013). During cancer, autophagy has been shown to be both involved in tumour suppression by inducing endocytic degradation of growth factors, as well as promoting tumour growth by endocytic degradation of apoptotic factors during stress conditions (such as that presented by the hypoxic environment of the inner tumour) (Yang *et al.* 2011). Both down regulated and up regulated gene lists were searched and no autophagy genes were present in the differentially expressed gene lists, indicating that XRN1 may not be not regulating autophagy, positively or negatively, in SAOS-2 cells. This is in contradiction to published work (Delorme-Axford *et al.* 2018), however, it may suggest that the effects of XRN1 on autophagy may be cell-type specific. It is also possible that the extent of knock down of XRN1 in SAOS-2 cells was not sufficient enough to induce changes in the levels of expression of autophagy genes. These genes are usually expressed in response to cell stress by starvation. Looking at gene expression in starvation conditions was not analysed in this project, which may explain why no autophagy genes are identified as differentially expressed in this experiment.

It was interesting to analyse the RNA sequencing dataset for the proportion of reads mapping to the various portions of each transcript. This acted as another layer of control during the identification of potential targets of XRN1, by distinguishing whether the transcript may have been differentially degraded between the two conditions. This showed that in the top 10 up regulated genes, there was evidence of a lack of degradation occurring in these transcripts in the knock down samples, as more reads were mapping to the 5' UTR and the CDS than in the control samples. The fact that more reads were mapping to the 3' UTR in the control samples suggests that the 5' UTR and CDS were being degraded constitutively from the 5' end by XRN1. This was very promising in identifying possible targets of XRN1 as it showed stabilization of the 5' end of the transcript, which would be expected in transcripts not being degraded in this fashion.

The main limiting factor to this type of analysis is that at present the computer programmes available are not efficient at being able to distinguish between the different isoforms of the same gene, and so differentiating the specific isoforms of differentially expressed genes could not be undertaken Although there are programmes available which allow the analysis of different isoforms, the type of deep analysis required in this context is still primitive. Cuffdiff does perform an isoform estimation, however, for this dataset it was not successful. The analysis conducted in this chapter does not take into account transcript differences of the different isoforms, and so the differences seen are a generalization for each gene as a whole, rather than in their different entities. It may be possible to go further into this analysis using programmes to find differential alternative splicing events, however, it would be easier in the future to assess differential expression of isoforms using qRT-PCR.

This chapter also discusses the types of motifs found within the 3' UTRs of the differentially expressed genes. This analysis identified 5 common motifs between the top 59 genes which were up regulated, with significant commonalities found amongst 34 genes, which were recurring in the motif analysis. The fact that 34 genes harboured the same motifs (notably the polyuridylated tracks and G-rich regions) suggests that these genes are the most likely to have been targeted for degradation by the 5' end. It is curious that there are regions of polyuridylation, given that this is characteristic of tagging for degradation by DIS3L2, suggesting that there may be some co-ordinate degradation between XRN1 and DIS3L2, despite there being little evidence to show that this might be the case. Although commonly associated with degradation by DIS3L2, poly(U) tracts (which are added to the 3' UTR and are not an intrinsic part of the sequence) are generally thought to stimulate RNA degradation, and even a single U nucleotide can be recognized by the LSM1-7 complex. This has been shown in mammalian cells to lead to decapping and degradation of the RNA from the 5' end by XRN1 (Song and Kiledjian 2007). The presence of G-rich regions in the 3' UTR also supports previous work alluding to the preferential degradation of targets with this motif by XRN1. Future work into analysis of the 3' UTRs of the up regulated transcripts could shed light into the interaction between XRN1 and RNA binding proteins, which bind to specific regions in the 3' UTR to influence RNA stability and post-transcriptional control of gene expression.

It will be interesting in the future to validate the expression of the six genes that were chosen in a more complete way, including analyzing the expression of these genes during XRN1 over expression and also the levels of pre-mRNA present for each gene using qRT-PCR. By doing so, the confirmation of these transcripts as targets of XRN1 will be better supported experimentally. It would also be beneficial to extend the list of genes for validation to get a more

complete overview of the types of genes XRN1 may be targeting, however, this depends highly on the expense of the assays.

Chapter 6

Discussion

6.1 Summary of findings

6.1.1 XRN1 is post-transcriptionally expressed at lower levels in Osteosarcoma cell lines and patient samples

The work presented in the Chapter 3 showed that *XRN1* mRNA expression is lower across 3 out of 4 OS cell lines (HOS, U-2 OS and MG-63). It was also shown that the level by which *XRN1* was lower correlated with the proliferative rate of the cell line, i.e. the faster the doubling time of the cell line, the lower the level of *XRN1* they appeared to have. In addition, it was shown that XRN1 also displays lower expression in the related cancer, Ewing sarcoma, an important commonality between the two cancers. The fact that *XRN1* mRNA expression is lower in cell lines derived from both of these cancers (compared to a foetal osteoblast control cell line) supports the notion of a role for XRN1 in bone cancer progression. Adding strength to this idea is that *XRN1* was also shown to be lower in OS patient samples, showing that lower *XRN1* expression is unlikely to be an artifact of the OS cell lines.

The levels of *XRN1* pre-mRNA were determined using a primer-probe pair designed to capture *pre-XRN1* RNA in order to determine whether changes in expression of XRN1 were as a result of changes during transcription or if levels were being post-transcriptionally regulated. The results suggested that XRN1 is being post-transcriptionally dysregulated when compared to the foetal osteoblast control. This implies that there is an external factor affecting the stability of *XRN1* mRNA to such a degree that it becomes significantly reduced in 3 out of 4 OS cell lines. These external factors could include miRNAs, or RNA binding proteins. A large amount of miRNAs have been shown to be differentially expressed in OS, and so any of these could be a candidate for regulating XRN1 (discussed in section 6.5). Motif analysis of the 3' UTR in Chapter 5 showed that post-transcriptionally up regulated transcripts had several common motifs, including C, U and A-rich regions, which could determine whether specific transcripts are destined for degradation by XRN1. In addition, current work is in progress to identify whether there are any specific miRNA binding sites in these transcripts.

It is important to note that the major limiting factor in this work is that the levels of XRN1 were directly compared to a foetal osteoblast cell line, rather than to an osteoblast cell line derived from an age-matched source. This could mean that there are expression differences

between the cell lines due to the gene expression at the time of bone development, rather than due to carcinogenesis. At the time the lab work was being conducted, there were no other available cell lines more appropriate for control, however, since the writing of this thesis, a new cell line has been released which is an adult osteoblast cell line. Though not perfect because the cell lines and patient samples were from adolescents, this would be a more appropriate control cell line to use in the future to confirm true XRN1 expression.

6.1.2 XRN1 is differentially localized in two out of three OS cell lines

Immunocytochemistry was used to observe whether the cellular localization of XRN1 was different in the OS cells when compared to the foetal osteoblasts. These experiments showed that in the control cell line XRN1 is uniformly distributed throughout the cytoplasm in a speckled manner. In contrast, XRN1 is seen in larger, punctate foci which are fewer in number in the more proliferative cell line, HOS. This was in line with the decreased protein expression observed by Western blot prior to this project (Appendix Figure A.1). These foci were also observed (but not quantified) to be mainly situated close to the nuclear envelope, rather than distributed throughout the cell. This phenotype was also true of XRN1 localisation in the U-2 OS cell line. In line with an insignificant change in expression of XRN1 in the SAOS-2 cell line, this cell line displayed XRN1 in a similar way to the control cell line, which was to be expected.

6.1.3 Other exoribonucleases in the cells are also dysregulated

Alongside expression studies for XRN1 in the OS cells, the cells were also tested for aberrant expression of other exoribonucleases in the cell. These included: DIS3, DIS3L1, DIS3L2 and XRN2. It was shown that at the level of mRNA, all exoribonucleases were higher in at least two out of three OS cell lines when compared to the foetal osteoblast control, however, not all of them showed higher expression in the same cell lines. *DIS3* and *DIS3L2* expression levels were higher in both the HOS and SAOS-2 cell lines, whereas *DIS3L1* expression was higher in HOS and U-2 OS cell lines, mirroring the changes observed for *XRN1*. *XRN2* expression was shown to be higher in all three OS cell lines.

Although protein expression data of DIS3 and DIS3L1 was difficult to obtain due to the lack of suitable antibodies (as well as very low expression in bone-derived cells in the case of DIS3L1), the data suggesting that other exoribonucleases in the cell were dysregulated in the opposite way to XRN1 supported the notion for co-ordinate regulation of exoribonucleases across the cell, especially in the case of DIS3L2 (the antibody for which was very good), which was higher in the OS cells. Chapter 4 discusses the lack of co-ordinate regulation occurring between XRN1 and the catalytic members of the exosome. It also shows how there is evidence

for some potential co-ordinate regulation occurring between XRN1 and DIS3L2, however, these subtle correlations were not statistically significant.

6.1.4 Protein expression data for the regulation of all tested exoribonucleases is not complete

An important point to make within Chapter 3 is that the protein expression data is not wholly complete (see Figures 3.12 - 3.13), in that there were issues in obtaining protein expression data due to the lack of specific antibodies. In addition, protein expression data for XRN1 only partially correlated with the results seen in the mRNA expression analysis. This could be due to several reasons: A) that the changes seen at the level of mRNA were not recapitulated at the level of translation, and so XRN1 protein was not downregulated, supporting the concept that global mRNA levels do not necessarily capitulate the levels of subsequent protein expression. B) That Western blotting as a technique is not as sensitive as TaqMan qRT-PCR, and so changes in expression at the level of protein cannot be reliably obtained.

Protein expression data for the other exoribonucleases was difficult to obtain. In the case of DIS3, this was because all the commercially available antibodies to human DIS3 were non-specific, and so quantifying the correct band was a real problem. For DIS3L1, protein expression data was almost impossible to obtain because of the extremely low levels of DIS3L1 expression in the OS cells. Again, the fact that DIS3L1 was difficult to analyse pertains to the need for a more sensitive technique than Western blotting in order to analyse protein expression in cell lines.

6.1.5 Artificial knock down of XRN1 does not result in an observable phenotype

Chapter 4 investigated the effect of XRN1 knock down in SAOS-2 cells (and U-2 OS) in order to determine if lower levels of XRN1 cause major phenotypic changes in the cell. After performing growth curves and a Brd-U assay to assess the rate of proliferation in response to XRN1 knock down, it was shown that no significant change in proliferation was caused by XRN1 knock down. In addition, a WST-1 cell viability assay was performed which also showed that XRN1 knock down does not cause any observable change in cell metabolism in these cells. Alongside the viability assay, a caspase assay was also performed to observe any changes in the rate of apoptosis. Again, no observable differences were seen. SUnSET labelling was used to assess the rate of translation in the cells after XRN1 knock down, which also did not show any observable phenotypes. This could be because knock down of XRN1 is not 100% efficient, and so any residual XRN1 expression might be enough to mask any phenotype in these cells.

Whilst this set of experiments did not show any phenotypic changes to the cell lines, this does not mean that there are no phenotypic changes occurring. Indeed, the use of cell lines in this kind of work may not always lead to a significant phenotypic change, because ultimately cell lines are not indicative of 'normal' wildtype behaviour, and more changes may be seen if these cells had remained *in situ*. Another explanation for not seeing obvious phenotypes may also be that there was compensatory regulation of other exoribonucleases in the cells which ensured the maintenance of characteristic behaviour of the cell lines. It could also be that during immortalisation of the cell line, the pathways regulated by XRN1 have changed, and no longer operate in a normal way. Other groups have also not seen a phenotype in cell lines upon XRN1 knock down, as confirmed by attendance at relevant conferences. This has also been published since the undertaking of the work in a PhD thesis (Garcia-Moreno *et al.* 2019).

6.1.6 DIS3L2 has a tissue-specific function

Knock down of DIS3L2 across two different cell types confirmed the existence of a potential tissue-specific activity for DIS3L2. Knock down of DIS3L2 in U-2 OS showed that there was no effect of cellular proliferation in this cell line, whereas knock down of DIS3L2 in HEK-293T cells showed an increase in cellular proliferation. This was an important observation because DIS3L2 was expressed at higher levels in the OS cell lines compared to the foetal osteoblasts, which suggested a role for DIS3L2 in OS progression. However, given that DIS3L2 has been implicated in Wilms' Tumour of the kidney, and HEK-293T cells are derived from embryonic kidney cells, it is more likely that DIS3L2 is constitutively active in these cells rather than the OS cells.

It is also possible that DIS3L2 expression changes throughout development of these tissues. In particular, it is plausible that DIS3L2 expression in the OS cell lines displays different developmental timing to its expression in HEK-293T cells, and that expression of DIS3L2 in adolescent OS cell lines does not affect cell line proliferation, however, its expression in embryonic kidney cells does affect cell line proliferation.

6.1.7 There may be synergistic activity between the exoribonucleases but this is not confirmed

Chapter 4 also includes looking at the expression of the exoribonucleases when one of them is knocked down. This was designed to identify the potential for synergistic expression between the exoribonucleases in order to maintain cellular homeostasis. Given that DIS3L1 displayed a mirror image increase in expression in the OS cell lines when compared to reduced

XRN1 expression, it was thought that DIS3L1 may be compensating for lower expression of XRN1. However, when XRN1 was knocked down this was not shown to be the case.

In addition, when XRN1 was knocked down, DIS3 expression was down regulated, suggesting a relationship between the two enzymes. It is unlikely that XRN1 is regulating DIS3 directly, because DIS3 is down regulated when XRN1 is down regulated, however, there is an argument to suggest that if XRN1 is acting as a transcription factor (discussed in section 6.3) this might account for down regulation of DIS3. Again, there also does not appear to be any synergism between XRN1 and DIS3L2, as expression of DIS3L2 does not change upon knock down of XRN1. There was an increase in protein expression of DIS3L2, however, this was not significant, and so a relationship between XRN1 and DIS3L2 cannot be confirmed at this time.

6.1.8 Knocking down XRN1 leads to differential gene expression in SAOS-2 cells

Chapter 5 shows that despite there being no observable phenotypic changes during XRN1 knock down, there are substantial gene expression changes. RNA sequencing data showed that there were a total of 766 genes which were up regulated, 59 of these genes were bioinformatically predicted to be post-transcriptionally regulated genes which were differentially up regulated compared to the scrambled siRNA control dataset. It was important to identify the genes which were post-transcriptionally up regulated (by comparing reads mapping to introns and reads mapping to exons) in order to be able to predict that these genes are normally degraded by XRN1, thereby separating direct and indirect effects of XRN1 knock down. Gene ontology analysis showed that there was a high enrichment of differentially expressed genes in the EGF pathway, a growth pathway associated with many cancers. It was interesting that in the top 10 post-transcriptionally up regulated genes NUDT15 was identified. *NUDT15* was later verified by qRT-PCR as being significantly up regulated upon XRN1 knock down. NUDT15 is a member of the Nudix family of proteins (Nudix Hydrolase 15), and is known to possess mRNA decapping activity (Song *et al.* 2013). Decapping activity is a critical determinant of mRNA decay, and up regulation of *NUDT15* suggests an existence of a feedback mechanism between XRN1 and the decapping factors to ensure tight regulation of 5’ – 3’ RNA degradation. It is suggested that several of the Nudix proteins, including NUDT15, catalyse cleavage in conjunction with DCP2, in a pleiotropic manner to generate both m^7GMP and m^7GDP (Song *et al.* 2013). It is important in the future to validate the protein expression of NUDT15 (and all other potential targets) as mRNA expression is not always recapitulated in protein expression.

It was also shown in the RNA sequencing dataset that in the majority of post-transcriptionally up regulated transcripts there was an increase in reads mapping to the 5’ end, comparatively to the control samples. This showed that less RNA degradation from the 5’ end

of the transcript was occurring, probably due to depleted XRN1, and that the 5' end was being stabilized.

The number of down regulated genes identified was greater than the number of up regulated genes, which may support the notion for XRN1 being a transcription factor (Haimovich *et al.* 2013). This data suggests that XRN1 has an activator role more often than it has a repressor role. Alternatively, it could be as a result of indirect effects caused by an increase in expression of XRN1 targets, in turn leading to up or down regulation of gene expression in other genes. If XRN1 is acting as a transcription factor this would make sense as to why many genes are down regulated. It also supports recent data pertaining to the presence of XRN1 in the nucleus, which has only been reported once before (Medina *et al.* 2014), as this is where transcription takes place.

6.2 Is XRN1 involved in the progression of osteosarcoma?

At this stage of enquiry, it would be appropriate to suggest that a role for XRN1 involvement in the progression of OS remains unknown. It is reasonable to assume that given *XRN1* is reduced in three out of four OS cells tested, as well as patient biopsy samples, and it is also lower in at least two Ewing sarcoma cell lines, compared to the foetal osteoblast control, that there is a role for XRN1 in the progression of two types of sarcoma. However, given that most characterisation assays have resulted in no substantial phenotype it becomes tempting to suggest that any specific role for XRN1 is not prominent enough to affect cellular behaviour by itself. Indeed, reports at the FASEB conference for post-transcriptional control of gene expression (Scottsdale, Arizona, 2018) suggested that complete XRN1 knock out in different cell types (such as HEK-293T cells) have also yielded no phenotype, which would suggest the XRN1 behaves redundantly. This is somewhat paradoxical in that XRN1 remains the only known 5' – 3' cytoplasmic exoribonuclease, and it has been shown that mutant *Drosophila melanogaster* are not viable when XRN1 is mutated. This suggests that XRN1 is not completely redundant, so the question becomes: what else is there in the cell which might cause this? One concept is that some transcripts normally targeted by XRN1 may be being redirected for 3' – 5' decay, which is consistent with XRN1 and DIS3L2 co-precipitating in an RNA dependent manner (Lubas *et al.* 2013). Alternatively, this may suggest that using these cell lines to study XRN1 is not a good model.

It is possible, of course, that lower levels of XRN1 leads to aberrations in the RNA decay pathway other than at the point of degradation, including the possibility that a lack of XRN1 may

lead to problems of the recruitment of other decay factors to the 5' cap, given that XRN1 directly interacts with the decapping complex (Braun *et al.* 2012). This was also shown in the RNA sequencing data which demonstrated a stabilisation of the 5' ends of predicted XRN1 target transcripts in XRN1 knock down cells. Immunocytochemistry experiments would show how the localization of the 5' decapping machinery differs between cells which have normal levels of XRN1 compared to those which have a depletion of XRN1, and whether defective degradation is occurring due to multiple members of the degradation pathway. The RNA sequencing data did not reveal significant dysregulation of expression of the 5' decapping machinery, and so it is more likely that differential localization due to XRN1 depletion is happening. In Chapter 3 it was shown that there is a level of differential localisation of XRN1 in the cell line where XRN1 is the most reduced in expression (HOS), supporting this notion.

It would make sense that reduced XRN1 expression could be a by-product of another fundamental change in gene expression, for example, if a major transcription factor is being affected, this could lead to the transcription of another protein/ncRNA in the cell which is responsible for down regulating XRN1 in OS cells. This then may have a downstream effect on the pathways usually regulated in part by XRN1 activity. Indeed, if either the IGF or EGF pathways are involved in the progression of OS, then the fact that XRN1 is lower suggests it has an indirect effect on growth.

For the time being, the OS cell lines can be used as a good model to identify the roles of XRN1 within human cells, however, it is too early to suggest that dysregulation of XRN1 is contributing to the progression of osteosarcoma itself. It may be that, given the results of the RNA sequencing experiment in Chapter 5, XRN1 is involved in regulating the expression of extracellular proteins, and possibly cell migration, a pathway which has not been studied in this project but should be analysed in the future. This will be discussed in more detail in section 6.8.2.

6.3 XRN1 as a transcription factor

Although paradoxical, the notion of XRN1 as a transcription factor is a controversial theory which has been widely circulated in the field, and has been the subject of 3 publications since 2013, each from the same lab. Given the extensive list of down regulated genes in the RNA sequencing data, the idea that XRN1 could be acting transcriptionally needed to be addressed. Despite this theory contradicting what is already known about XRN1, and its strictly cytoplasmic localization, there has been a publication showing the presence of XRN1 in the nucleus, and its transcription factor activity (Medina *et al.* 2014). In this publication, the research group show,

using genome-wide assays, that depletion of XRN1 in *Saccharomyces cerevisiae* not only decreases the decay rate of a specific subset of RNAs, but also that the RNA synthesis rate of these same subset of RNAs is also decreased.

The group has also shown, in a previous publication, that XRN1 is capable of shuttling between the nucleus and the cytoplasm to play a role as a transcriptional activator (Haimovich *et al.* 2013). In this work, the group show that members of the 5'-3' decay pathway shuttle between the nucleus and the cytoplasm in a manner dependent on proper mRNA degradation by XRN1. They state that in the nucleus the components (including XRN1) associate with chromatin to stimulate transcription initiation and elongation, as shown by ChIP-exo analysis. Further analysis showed that the components of the degradation machinery seemed to bind the chromatin as a complex, rather than as independent factors. This is in contrast to other published work, which has shown that XRN1 co-sediments with polysomes in both human cell lines (Lubas *et al.* 2013) and *Drosophila* models (Antic *et al.* 2015), showing that XRN1 performs mRNA degradation co-translationally in the cytoplasm. This is in addition to a recent publication resolving the crystal structure of XRN1 in complex with the ribosome in order to degrade mRNA as translation progresses along the transcript (Tesina *et al.* 2019).

Further work from this group has since shown that whilst acting as an RNA decay enzyme and a transcription factor, XRN1 can also act as a translation initiator. This was demonstrated, as in the previous publications from this group, in *S. cerevisiae.* Here they showed that XRN1 promotes translation of a specific subset of membrane proteins, dependent on poor RNA structural components (Blasco-Moreno *et al.* 2019) in the endoplasmic reticulum. They also suggest that XRN1 activates translation of transcripts which have highly structured and longer 5' UTRs (as shown by poly ribosome sequencing), which occurs at the site of the 80S ribosome (XRN1 was shown to co-sediment with the 80S ribosome).

Immunocytochemistry experiments about the localisation of XRN1 in mammalian cells performed in this project do not correlate with the data presented by the above publications which suggest that XRN1 acts a transcription factor. XRN1 was not shown to be present in the nucleus in either the foetal osteoblast control cells or the OS cell lines, in agreement with most other publications in the field, therefore refuting the theory that XRN1 is a transcription factor, at least in mammalian cells. This is equally compounded by the fact that XRN1 does not possess a Nuclear Localisation Signal.

The RNA sequencing data in this project alludes to a possible role for XRN1 in the regulation of membrane proteins, specifically those involved in cell motility and adhesion, in agreement with the paper previously described (Blasco-Moreno *et al.* 2019), however, it is

unlikely that XRN1 is initiating translation of these RNAs. It is more likely that XRN1 depletion leads to an up regulation in translation of these transcripts purely because there is a lack of degradation, and as a result, higher transcript availability. It is also widely accepted that XRN1 cannot processively degrade transcripts with highly structured 5' UTRs, and at these structures the activity of XRN1 is stalled. RNAs with highly structured 5' UTRs may do so because they are not destined to be degraded, and are, therefore, being translated instead. It has been shown that XRN1 activity is co-translational, and structural studies in *Saccharomyces cerevisiae* have resolved protein structures to show that XRN1 associates with the 80S ribosome. In this study, they showed that newly translated mRNA is fed directly into the active site of XRN1 in order for it to be degraded, which does not support a role for XRN1 in initiating translation (Tesina *et al.* 2019) (Figure 6.1). This structural resolution confirms previous reports that mRNA degradation takes place in the vicinity of the ribosomes in both *Saccharomyces cerevisiae* (Hu *et al.* 2009) and *Drosophila melanogaster* (Antic *et al.* 2015), and that mRNA degradation is not exclusive to P-bodies (see section 6.4). Alongside RNA sequencing and localisation studies in this project, Chapter 4 also shows that when XRN1 is knocked down in human OS cells there is no significant impact on the rate of translation, which is again contradictory to the work recently published.

Various other articles describe a tight coupling for transcriptional processing in the nucleus and RNA decay at the boundary of the nucleus, at the site of the exit channels (Braun and Young 2014). It seems that this would be a more plausible phenomenon than XRN1 shuttling between the nucleus and the cytoplasm, and there is no current evidence to suggest that nuclear shuttling of XRN1 takes place in humans. The data presented in this project showed that XRN1 does accumulate in larger foci around the edge of the nucleus in the cancer cells, supporting the idea of tight transcriptional processing between the cytoplasm and the nucleus.

Indeed, it would make sense that there was some nuclear signaling taking place between the nucleus and the cytoplasm to account for these changes in transcription upon a depletion of XRN1. It would be advantageous for the cell both from an energy-saving point of view and as a result of signaling feedback that transcription of certain genes would be decreased if the mRNA previously transcribed is not being degraded, as there would be no need for further transcription. Perhaps the idea of XRN1 being a transcription factor is not conserved from yeast to higher eukaryotes. It has been reported that not all members of the degradation machinery harbour the capacity to activate transcription through a lack of an activating domain, so perhaps over the course of evolution, the activation domain has been lost from more members of the degradation pathway, resulting in a role which is purely cytoplasmic, and likely that this phenomenon is not conserved. There is no evidence of this occurring in higher eukaryotes.

Interestingly, studies in viruses have shown that increased cellular activity of XRN1 can generate a transcriptional response. It has been shown that the Gamma-herpesviruses (such as Kaposi's sarcoma-associated herpesvirus and Epstein-Barr virus) (Covarrubias *et al.* 2009) encode a cytoplasmic nuclease (SOX), which cleaves mRNAs and leaves them vulnerable to degradation by XRN1 (Gaglia *et al.* 2012). Accelerated degradation by XRN1 leads to repression of transcription by RNA polymerase II in the host, suggesting that mammalian cells can sense changes in the mRNA degradation. The virus benefits from this in that it evades the host defence mechanisms and escapes decay-induced transcriptional repression of viral mRNAs (Abernathy *et al.* 2015), allowing transcription of viral mRNA to take place. This work presents an alternative mechanism for viruses being able to evade the host response, in conjunction with work by the Wilusz lab, who have shown that secondary structures on viral RNA stall the activity of XRN1 to promote the expression of viral RNAs, bypassing the host response, as discussed in the Introduction of this thesis.

In all, the evidence for mammalian cell signaling feedback between cytoplasmic degradation and the rate of transcription seems to be a much more substantial and plausible concept than the idea that XRN1 is being shuttled between the cytoplasm and the nucleus to act on both pathways. This notion is supported by localization studies in this project, as well as the RNA sequencing dataset, in which the GO analysis did not suggest a role for XRN1 in transcription. It is more likely that XRN1 interplays with another cellular factor to control gene expression, and that when XRN1 is artificially knocked down, this results in a subset of genes which are post-transcriptionally up regulated (as would be expected) and a large subset of genes which are differentially down regulated, possibly due to an indirect link with XRN1 activity. This concept is supported by the fact that there are 19 commonly up regulated genes when XRN1 is knocked down in HeLa cells, as well as at least 4 genes which are commonly down regulated.

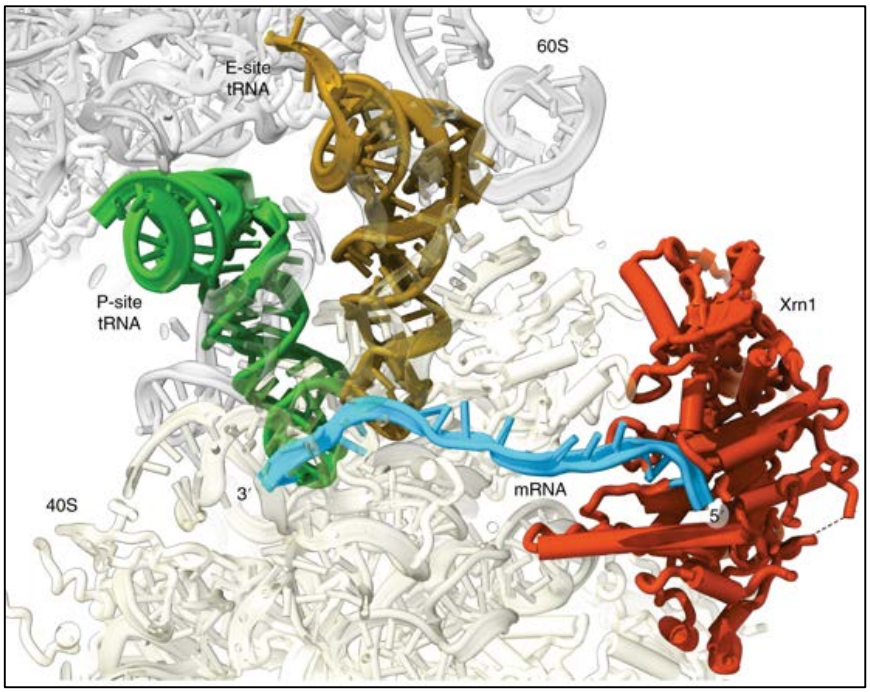


Figure 6.1. Path of mRNA from the decoding site of the ribosome to the degradation site within the catalytic core of XRN1 as adapted from (Tesina, Heckel *et al.* 2019).

6.4 The role of XRN1 in P-bodies and stress granules

Conventionally, Processing bodies (P-bodies) are described as cytoplasmic granules that couple mRNA degradation with transcript storage. Though present in lower levels in cells not undergoing stress, the granules are further induced with the onset of stress, as a response which leads to the inhibition of translation initiation. Translation initiation is generally thought of as a competitor to RNA degradation, in that they both work in opposite directions to maintain cellular homeostasis and balance the rate of mRNA translation and mRNA degradation. Translation initiation can be inhibited by strong structures in the 5' cap, whilst degradation can be inhibited by the cap binding protein, eIF4E, which blocks the access of decapping enzymes (Decker and Parker 2012). P-bodies are comprised mainly of 5' – 3' degradation machinery (including the CCR4-NOT deadenylation complex, decapping activators and XRN1) and mRNAs in complex with proteins associated with translational repression (Luo *et al.* 2018). In contrast, another cellular granule associated with translation repression are stress granules, which are comprised of predominantly messenger ribonucleoproteins (mRNPs) which form from mRNAs stalled in translation initiation (Protter and Parker 2016). Stress granules are traditionally viewed as being formed, and associated with, during times of cellular stress and the induction of autophagy (Monahan *et al.* 2016). Both P-bodies and stress granules are formed of mRNPs, derived from pools of untranslating mRNA.

There are current discussions in the field about how to class these two granules. Previous work has shown that P-bodies and stress granules work in tandem as a response to stress to inhibit translation and conserve cellular activity, and whilst doing so, remain as distinctly separate compartments, with no compartmental mixing (Hubstenberger *et al.* 2017). However, recent work disclosed at the RNA UK conference in January 2018 showed that in yeast there are many classes of granules, each pertaining to differential functions within the cell. The type of granule induced may depend on the metabolic needs of the cell at any one time, and need to be seen as more discrete factors. There were several types of granule proposed at the meeting, including granules termed Afe (an amalgamation of glycolytic RNAs) and Translation Factor (a liquid body where translation takes place). It is also known that there are many different types of granules in human cells. The role of P-bodies being exclusive sites of 5' – 3' mRNA decay has been widely disputed, and it has been accepted that XRN1 does not exclusively localize to P-bodies (Eulalio *et al.* 2007), because 5' – 3' RNA decay still occurs in the absence of P-bodies. This phenomenon has been shown both in this thesis (Chapter 3) and in a recent publication pertaining to XRN1 localisation in granules known as eisosomes, which are plasma membrane-associated particles which may function during extracellular signaling in yeast during

glucose exhaustion stress (Grousl *et al.* 2015). XRN1 localisation outside of P-bodies has also been confirmed with the resolution of XRN1 associating with the 80S ribosome (Tesina *et al.* 2019) (see section 6.3). It is also possible that XRN1 localises to GW-bodies, which act separately to P-bodies, however, both are involved with miRNA/siRNA silencing (Gibbings *et al.* 2009). The localisation of XRN1 to GW-bodies was not explored in this project.

There are many common factors which are present in both P-bodies and stress granules (such as decapping activator, Dhh1, and translation initiation factor, eIF4E), including some mRNAs, however, the presence of translation initiation factors is unique to stress granules (Luo *et al.* 2018). The most important factor relative to both P-body and stress granule is XRN1 (Kedersha *et al.* 2005). Under certain stresses, P-bodies can be known to dock with stress granules, which can lead to P-body/stress granule fusion (Stoecklin and Kedersha 2013). In section 6.3, the idea of XRN1 as a transcription factor was discussed as an unlikely phenomenon going on in the cells, given the results of the localisation experiments in Chapter 3 and the RNA sequencing experiment in Chapter 5. Here, it could be viewed that the feedback signaling linking XRN1 to transcription comes as a response to its activity in both P-bodies and stress granules. If it is influencing the inhibition of translation in stress granules, this could explain the relative response of the cell to adapt the rate of transcription accordingly.

A potential theory to be explored could be that separately these granules are dormant in the cytoplasm until certain stresses occur, which induces them to become docked to one another for the inhibition of translation, hence why XRN1 is present in each of them. It could be that the mRNAs stored in the granules are so stored in order that they are quickly accessible in times of stress. Alternatively, they could be sites of storage induced during stress conditions, which allow for rapid adaptation and reinstatement of normal translation once the stress has been removed. Once the stress is detected, these granules come together to inhibit translation of stored mRNAs, subsequently repressing cellular translation. The docking of the two granules may also induce specific XRN1-mediated RNA degradation in response to cellular stress.

It is difficult to point to a specific role for XRN1 within P-bodies and stress granules because the mechanisms by which they act remain elusive. Indeed, a recent paper alluded to how repressed mRNAs may return from P-bodies to the cytosol to re-enter polysomes for translation (Hubstenberger *et al.* 2017) and that translation repressors, but not decay machineries, are essential for P-body formation (Ayache *et al.* 2015). There is also the fact that there are multiple examples of P-bodies exhibiting remodeling capabilities (Ernoult-Lange *et al.* 2012), which shows that simplifying the role of P-bodies to XRN1-mediated RNA degradation and mRNA storage is not sufficient to explain the role of P-bodies in the cell, or the role of XRN1

within P-bodies. This is especially true since XRN1 localises to both P-bodies and stress granules, identifying a role for this 5'-3' exoribonuclease in both of these granules. It is likely that XRN1 localises to P-bodies to regulate the expression of the specific subgroup of RNAs which are stored in them. These generally include regulatory transcripts, suggesting that P-bodies provide a reservoir for quick adaptation to gene expression (Hubstenberger *et al.* 2017), rather than being the exclusive site of 5' – 3' RNA degradation.

This project has shown that XRN1 displays complete colocalisation with P-bodies in the foetal osteoblasts, but only partial colocalisation with P-bodies in the OS cells. Arguably, the OS cells were under more cellular stress due to the nature of cancer cell culture, and so it could be argued that where XRN1 was not colocalising to P-bodies, it might actually have been localising to stress granules as well. If so, is the environmental nature of the stress applied to cancer cells enough to cause P-body/stress granule docking? This could be another defect in cancer cells which needs to be explored further.

6.5 Interactions between XRN1 and miRNAs in osteosarcoma

One aspect which has not been addressed in this project is the possibility that differential expression of miRNAs may be responsible for the decreased expression of XRN1 in OS cells, in order for the progression of osteosarcoma to occur. These miRNAs, which could also be referred to as an oncomirs, may be participating in XRN1 inhibition by binding to the 3'UTR of mature *XRN1* mRNA, preventing its translation. The following section will postulate possible ways in which they could be enhancing the progression of OS in humans by inhibiting XRN1.

In addition, the possibility that XRN1 itself could be regulating miRNAs will also be discussed. Lower expression of XRN1 could lead to defective regulation of a subset of thus far unidentified miRNAs, leading a cascade of defects in post-transcriptional control of gene expression. Up regulation of specific miRNAs normally degraded by XRN1 could be responsible for repression of tumour suppressor transcripts, or even activating oncogenes as a direct result.

6.5.1 Are miRNAs responsible for regulating expression of XRN1 during the progression of OS?

Given that miRNAs are highly effective non-coding RNAs which work to regulate gene expression at multiple levels, it is feasible that one or more miRNAs could be responsible for the decreased expression of XRN1 in OS. It was important to discuss this possibility because it was not explored during the undertaking of the project, yet it remains a possible reason why XRN1

expression is lower in OS, and it is widely acknowledged that miRNAs are involved in the progression of cancer. Specific miRNAs have been shown in the literature to target XRN1 expression. In particular, *miR-204* has been shown to repress the expression of XRN1 in prostate cancer cells, in a growth regulatory manner (Ding *et al.* 2015).

In Chapter 3, Table 3.1 shows that in fact there are many miRNAs which are known to target XRN1, including many oncogenic miRNAs. This list includes*miR-92b-3p* and *miR-92a-3p*, which have been shown to be involved in the progression of OS. It would be interesting to observe the expression of these miRNAs in the OS cell lines, because if they are dysregulated in OS, it may help to explain why XRN1 is also dysregulated in this cancer. There are many publications appertaining to the expression of miRNAs in OS, including the identification of 182 miRNAs in the OS cell line, SOSP-9607, which were elucidated through DNA sequencing and database searching (Gao *et al.* 2007).

In the future it would be prudent to conduct miRNA sequencing of the OS cells, to get a fine-tuned dataset to describe the expression of miRNAs in OS. This type of deep sequencing was recently conducted on patient samples of both primary tumour and lung metastases, which identified 65 differentially expressed miRNAs between patient samples and normal controls (Xie *et al.* 2018). This supports the idea that XRN1 is being regulated by these miRNAs.

6.5.2 Is lower XRN1 expression responsible for defective regulation of miRNA expression in OS?

Another school of thought is that because XRN1 expression is lower compared to the HOb control, this could mean that specific miRNAs become up regulated, and it is this that then leads to the dysregulation of cellular pathways and cancer progression. It has been shown across the field that XRN1 is both responsible for degrading miRNAs, as well as being inhibited by them. XRN1 has been shown to degrade miRNAs in *C. elegans* through to human cell lines, and two examples of miRNA degradation by XRN1 are *miR-241* (*C. elegans*) and *miR-382* (HEK-293T cells) (Zhang *et al.* 2012). XRN1 has also been shown to rapidly degrade extracellular miRNAs to promote the epithelial-mesenchymal transition in cancer cells (Zangari *et al.* 2017). Interestingly, *mir-382* has been implicated in protecting patients from OS metastases (Xu *et al.* 2015) and tumour growth and chemoresistance in U-2 OS cells (Xu *et al.* 2014), demonstrating clinical importance for regulation of this miRNA by XRN1.

There are a number of miRNAs which have been shown to associate with XRN1. It has been shown that extracellular *miR-233-3p* is targeted for degradation by XRN1, and that silencing XRN1 promotes the epithelial-mesenchymal transition directly through allowing *ex-*

miR-233-3p to bind to its target gene, *FOXO1* (Zangari *et al.* 2017), and prolonging invasive phenotypes. It has also been shown in *Drosophila* models that XRN1 affects the expression of particular miRNAs which could indirectly regulate the expression of mRNAs in the wing imaginal discs. One such miRNA was *miR-277-3p*, which was shown to be affected by XRN1 at the maturation step of miRNA processing (Jones *et al.* 2013). It was hypothesised in this study that *miR-277-3p* targeted Hsp67Bc (thought to be important in the autophagy pathway) because they both showed altered expression in *Drosophila* mutants, however, there was no obvious link between the two, therefore this remains speculative.

It is also possible that XRN1 targets miRNAs at different levels of their maturation process. It is probable that there is some regulation of pre-miRNAs by XRN1. One study has shown that there are a number of RNA binding proteins which interact with pre-miRNAs to regulate their expression, including mRNA capping and 3' end processing factors (Treiber *et al.* 2017), indicating that interaction with XRN1 is possible. Despite this, other studies have shown that it is more likely that post-transcriptional regulation of miRNAs can occur via protein – or RNA-mediated pathways. One of these pathways encompasses both XRN1 and XRN2 complexes to directly mediate miRNA turnover (Chatterjee and Großhans 2009), however, it is currently disputed whether targets of miRNAs promote or inhibit turnover of mature miRNA. It has been reported that extensive complementarity between miRNAs and seed sequences of target mRNAs promotes target mediated miRNA decay in human cells (Baccarini *et al.* 2011). This is not necessarily conserved across evolution, as in *Drosophila* models, mRNAs have been shown to confer protection to miRNAs from targeted degradation (Ameres *et al.* 2010). Most recently, it was shown that seed-matched targets enhanced miRNA stability and nuclear retention U-2 OS cells, utilised by Ago1 (Pitchiaya *et al.* 2017).

It would be useful to conduct a miRNA sequencing experiment on the SAOS-2 cells in the same way as was conducted in Chapter 5, whereby wildtype SAOS-2 cells transfected with a scrambled siRNA control are compared to SAOS-2 cells which have had XRN1 knocked down by a siRNA targeting XRN1. This would confirm whether XRN1 is responsible, in turn, for regulating the expression of specific miRNAs in OS cells. Not only would these experiments show the relationship between XRN1 and miRNAs, it would also help to inform future work into potential drug targets for this disease, and perhaps even diagnostic or prognostic biomarkers.

6.6 XRN1 and its role in the autophagy pathway

As discussed in Chapter 5, XRN1 has been implicated in the autophagy pathway. Recent publications have shown that during times of stress, XRN1 acts as a negative regulator of autophagy (Delorme-Axford *et al.* 2018) in both *Saccharomyces cerevisiae* and mammalian cells. This publication shows how chromosomal deletion of XRN1 enhances autophagy and up regulates autophagosome formation, as well as up regulating autophagy-related (ATG) transcripts, during nitrogen deprivation conditions. This group and another have also previously identified decapping enzyme DCP2 (Hu *et al.* 2015) and DEAD-box helicase, Dhh1, as post-transcriptional regulators of autophagy, and so it follows suit that XRN1 would also be involved in regulating autophagy.

Another paper has illustrated that other aspects of the RNA degradation pathway are involved in the regulation of autophagy, including the Pat-1-LSM1-7 complex. During nitrogen starvation-induced autophagy, this complex is able to bind and stabilize a specific subset of ATG transcripts that would otherwise undergo 3'- 5' degradation by the cytoplasmic exosome (Gatica and Klionsky 2019). This group shows that both XRN1 and exosome-mediated *ATG* mRNA degradation are tightly regulated to prevent excessive autophagy induction during nutrient-rich conditions.

Despite current literature suggesting a role for XRN1 in the induction of autophagy, the results of the RNA sequencing experiment in Chapter 5 did not show up regulation of any autophagy related transcripts upon XRN1 knock down. Genes involved in autophagy were also not found in the list of down regulated transcripts, suggesting that either XRN1 is not a regulator of autophagy, or that the regulation of autophagy is only active during times of cell stress (this scenario would seem more likely). The cells were not under stress conditions during the XRN1 knock down experiments, and so the lack of identification of autophagy transcripts suggests that this is not because XRN1 does not regulate autophagy transcripts, but rather that cells were not undergoing induction of autophagy during this experiment. It could also suggest that the effect of XRN1 on the regulation of autophagy is cell-type specific.

6.7 Implications of defective exoribonucleases in human disease

6.7.1 XRN1 is utilized by viruses to evade the host-response

XRN1 has been implicated thoroughly in the mechanism by which viruses evade the host response. It is extraordinary how viruses have evolved ways to stall XRN1 specifically in order to promote the propagation of its own RNAs. Stalling XRN1 on RNA lariats and stem loop structures to avoid degradation is a highly effective way for short viral RNAs to hijack cellular systems and establish a productive infection. It has been shown to be a conserved mechanism across an increasing number of viral families, including flaviviruses, arenaviruses, phleboviruses and Hepatitis C virus (HCV) (Chapman *et al.* 2014, Charley *et al.* 2018, Guo *et al.* 2018). Some of these viruses lead to pathologies such as encephalitis, providing another link between XRN1 and neuropathologies, although in a different context. HCV can, if undetected, lead to the onset of hepatocellular carcinoma, however, a link with XRN1 has not been established so far. It is thought that HCV manipulates host-gene expression to bypass antiviral responses, promote cell growth and prevent apoptosis to establish a chronic infection (Huang *et al.* 2005), though this is through mRNA stabilization. The RNA-binding protein, NS5A, binds specifically to G and U- rich elements within HCV RNA to protect it from degradation, possibly by XRN1. NS5A has also been implicated in regulating cell growth and apoptosis (Huang *et al.* 2005). It is important to note that the recognition and binding of NS5A to G and U-rich elements in HCV RNA is consistent with motifs on transcripts identified as up regulated when XRN1 is knocked down in Chapter 5, and therefore, it is possible that NS5A targets transcripts for XRN1-mediated degradation.

6.7.2 Mutations in DIS3 and DIS3L2 are linked to the progression of cancer

Despite more work needed to conclusively confirm a role for XRN1 dysregulation in cancer, 3'-5' exoribonucleases have been implicated in the progression of specific cancers.

The catalytic member of the nuclear exosome, DIS3, has been shown to be mutated in multiple myeloma, a rare blood disorder which affects the plasma cells in the bone marrow. Mutations in DIS3 are not limited to this blood disorder, and have been further implicated in acute myeloid leukaemia, where 4% of patients displayed mutations. Whole genome/exome sequencing analysis showed that there are multiple somatic mutations in DIS3, which interfere with exonucleolytic activity of the enzyme, in 11% of patients (Lionetti *et al.* 2015).

DIS3L2, an independent cytoplasmic 3'-5' exoribonuclease has also been implicated in the progression of a subset of cancers. Like DIS3, it has been shown that mutations in DIS3L2 are associated with the onset of the childhood cancer, Wilms' tumour of the kidney. In addition,

these mutations are also implicated in the congenital overgrowth syndrome, Perlman's syndrome. Children born with Perlman's syndrome have a pre-disposition to Wilms' tumour. Likewise to XRN1, this syndrome also displays neurological effects, with survivors suffering with macrocephaly (seen in various virus pathologies – section 6.7.1), and severe developmental delay (Morris *et al.* 2013). Clinical features of Perlman's syndrome results from defective RNA metabolism, specifically RNA metabolism which affects transcripts involved in growth and division. DIS3L2 has also been shown to suppress the growth of HeLa cells when overexpressed, identifying tumour suppressor activity for this gene (Astuti *et al.* 2012). This is consistent with work in *Drosophila melanogaster* which has shown that organisms mutant for DIS3L2 show increased proliferation in the wing imaginal disc, identifying a role for DIS3L2 in regulating proliferation and tissue growth (Towler *et al.* 2016).

Given that mutations in DIS3 and DIS3L2 are both associated with the development of certain cancers, and that XRN1 could be involved in the development of hepatocellular carcinoma arising from defective RNA degradation in HCV-affected patients, there is a strong argument to suggest that dysregulation of the exoribonucleases is highly clinically important. The implications for human disease are greatly increased in disorders and viral assaults where it has been shown that RNA metabolism is affected. This is not surprising, given that RNA degradation is a surveillance pathway dedicated to maintaining the correct balance of RNA synthesis and subsequent translation. This fine tuning of the Central Dogma is highly evolutionarily conserved, demonstrating the importance of the pathway, meaning that defects within the pathway are likely to cause detrimental effects to the cell. It is somewhat surprising that currently there are no known therapies in development to combat this aspect of disease, particularly in neurological disorders. In an age where neurodegenerative disorders are one of the leading causes of death in an ageing population, it seems pivotal to try to develop therapies which target the aggregation of protein deposits due to defective RNA processing, including decay. One such target for therapy, for example, is TDP-43, an RNA binding protein which is involved in splicing, stability, and transport (Tollervey *et al.* 2011, Liu *et al.* 2017). Mutations the gene encoding TDP-43, *TARDBP*, are associated with the progression of MND, currently incurable, and fatal.

6.8 Future work

This section will be split into future work which needs to be conducted on each results chapter in this project:

6.8.1 Characterisation of XRN1 in OS cells

In this chapter, the mRNA and protein expression of XRN1 in both OS and EWS cells, as well as OS biopsy samples, was analysed using qRT-PCR, Western blotting and immunocytochemistry. To complete the characterization of XRN1 in OS cells, the next steps will be to reanalyze the expression of XRN1 in OS and EWS using a more appropriate control. So far, the best available control for this analysis has been the foetal osteoblast cells. This control is not ideal for a number of reasons:

A) The cells are foetal cells, and so are not representative of the age of the OS cell donors, or the patient biopsies.

B) With particular regard to the patient biopsies, the foetal osteoblasts do not recapitulate the environment in which the cells were taken. The biopsies were taken from a tumour microenvironment, made up of several different cell types. The proportion of different types of cells in these biopsies is unknown, and are being directly normalized to a cell line (this in itself is not appropriate) where 100% of the cells are foetal osteoblasts.

C) In addition to the issues of using a foetal osteoblast control cell line to analyse XRN1 expression in patient biopsies, there is also the issue that the cell line is not an age-matched control, and therefore is gene expression in these cells really representative of the gene expression landscape at this different stage of development.

D) It is not entirely appropriate to use foetal osteoblasts as the only control to analyse expression of XRN1 in EWS cell lines. EWS is a highly variable cancer, and is capable of developing in a multitude of different cell types in the bone environment. It is not limited to affecting just osteoblasts. This project does not investigate the expression of XRN1 in other bone cell types, such as those which can be affected by EWS (adipocytes, osteocytes), and so it is unknown whether XRN1 expression can be truly described as lower in this cancer. It is also unknown what type of cell the EWS cell lines were originally derived from, the only information available is that both the SK-ES-1 cell line and the RD-ES cell line are epithelial cell lines.

An appropriate set of controls for this chapter include the use of the relatively new osteoblast cell line, which is derived from a more-appropriately age-matched (adult) source. For the biopsy analysis, it would be appropriate to gather healthy tissue from the same microenvironment as the tumour. In cases where limb amputation has been a course of treatment for OS, usually the affected limb is removed intact, and so it would be possible to take tissue from the opposite end of the limb, in the parallel 'healthy' growth plate environment. This would be the best control for this experiment because it would give the opportunity for an exact

age-matched sample from the same tissue, with the original gene expression landscape. Another control for this analysis would be to retrieve bone from children undergoing bone surgery for non-cancerous reasons. This could be achieved by collecting the dust created by surgeons drilling into the bone, which could then be utilised for RNA extraction and subsequent qRT-PCR analysis of the expression of XRN1. This is also the case for testing the expression of the other exoribonucleases explored throughout this project: DIS3, DIS3L1, DIS3L2 and XRN2.

Other areas of future work for this chapter would be to obtain patient biopsies and appropriately matched controls in order to gain a more complete insight into the expression of XRN1 in EWS. So far, only 2 cell lines have been tested, however, there are more cell lines which need to be tested in order to assume that levels of XRN1 are lower compared to the control cell line in the majority of EWS cell lines.

It would seem sensible to check XRN1 expression using appropriate controls across a variety of sarcoma cell lines and biopsy samples. If XRN1 expression can be shown to be lower in a multitude of sarcomas, this would offer scope for further investigation into a potential role for XRN1 in cancer progression, including the potential for novel drug development in these types of cancer which can specifically target the anomaly of lower XRN1 expression.

6.8.2 The effects of changing the regulation of the exoribonucleases in OS cells using siRNA

Although many cancer-specific pathways were looked at in this chapter, as well as potential synergism between the exoribonucleases, one hallmark of cancer was not completed: the effect of XRN1 knock down on cellular migration and invasion into external tissue (that which is not already affected by the tumour).

Cell migration and cell invasion can be described as 2 different phenomena. Cell migration is described as the propensity of cells to be able to move, including the reorganization of the cytoskeleton in order to push the cell in certain directions in response to extracellular stimuli, in an orchestrated manner. This can occur during organismal development and during wound healing, and cells normally migrate in a very specific direction. Cell invasion, on the other hand, whilst related to cell migration, defines the ability of the cell to become motile and navigate throughout the ECM in order to infiltrate its own or neighbouring tissues. Cell invasion is particularly interesting to look at with regard to cancer because it is one of the ways that cancer cells directly penetrate into surrounding tissue during tumour growth and metastasis. Indeed, the first step for tumour metastasis is the invasion of tumour cells into the stroma, and stromal collagen is dramatically remodeled to accommodate this (Clark and Vignjevic 2015). The

way a tumour begins invasion depends a lot on the microenvironment, which influences the mode and dynamics of the way cancer cells invade. In order to acquire the characteristics needed for cells to detach from the primary tumour and invade the basement membrane, cells undergo epithelial-mesenchymal transition (EMT), whereby cells lose their polarity and E-cadherin mediated cell-cell adhesion. Evidence has shown that post-EMT cells have a tendency to become more migratory, though it is important to note that increased migration is not an inexorable consequence of EMT (Schaeffer *et al.* 2014).

XRN1 has been linked with cell migration in several publications, identifying another pathway which may be being dysregulated when XRN1 expression is lower. Indeed, previous work in the Newbury lab has shown that XRN1 mutant *D. melanogaster* show defects in wound healing and epithelial closure (Nagarajan *et al.* 2013), and in *C. elegans* mutations in XRN1 cause defects in ventral epithelial enclosure during embryogenesis (Newbury and Woollard 2004). More recently, it has been shown that the contents of extracellular vesicles, which play an important role in intercellular communication, can be subject to rapid decay by XRN1 once within the recipient cell. Silencing XRN1 in lung adenocarcinoma cells showed slowed degradation of extracellular miRNAs (specifically ex-*miR-223-3p*), which led to the induction of prolonged invasiveness of the recipient cell (Zangari *et al.* 2017). Invasiveness was initially only transient during wildtype XRN1 expression.

It would be sensible to observe the consequences of XRN1 knockdown with regards to cellular migration rates, to see if the effect seen in (Zangari *et al.* 2017) can be recapitulated in these cells. If this is the case, then there would more evidence to suggest that lower XRN1 expression is having a major effect on cancer progression. Initial migration studies could be conducted by using a bung system to observe cell migration over a period of days. The bung of known diameter allows for a space to be left in a well where cells do not adhere. Once seeded and adherence has taken place, the bung can be removed and each day migration can be measured. If there is a difference in migration over time in cells where XRN1 is knocked down when compared to those of a scrambled siRNA control, then this would indicate that XRN1 may be having an effect in this pathway.

Invasion studies would be performed using a cell invasion assay. Several methods designed to measure cell invasion utilize fluorometric or colorimetric detection methods to detect invasion of the cell layer into a matrix layer of gelatin, collagen I, fibronectin or extracellular matrix. The idea would be to determine differential invasion between cells with and without XRN1 knock down. A transwell assay could also be used to measure cellular invasion. A transwell assay is designed to assess permeability of cells through a membrane

(usually a matrix of collagen to promote both adherence and migration) within cell culture wells, usually a Boyden chamber (Senger *et al.* 2002). Following an incubation period, cells which have migrated through the membrane are stained and counted.

Other future work in this area includes observing any physiological changes in cells when XRN1 expression is rescued or over-expressed. Initial studies looking into XRN1 rescue were performed, however, due to unforeseen circumstances it was very difficult to obtain an appropriate control for these experiments. An XRN1-expressing plasmid was kindly donated by Eliza Izaurralde, however, there was a discrepancy between the plasmid map which was published and what was published on the plasmid database. This discrepancy made it difficult to create a control plasmid, where XRN1 had been removed.

Initial experiments seemed to show that by rescuing XRN1 expression in U-2 OS cells the growth rate of the cells increased when compared to an empty eGFP vector control (Appendix Figure 4). However, it was decided that the promoter in this eGFP control plasmid (kindly donated by Dr. L. Mullen, Brighton & Sussex Medical School) was incredibly strong, which could have caused a detrimental effect on the growth of the cell line. Cloning experiments were undertaken to cleave the XRN1 sequence out of the original plasmid to allow the sticky ends of the plasmid backbone to ligate back together without XRN1, leading to the creation of an appropriate control plasmid. The major problem with this was that the XRN1 sequence made up almost half the entire size of the plasmid, and so it proved very difficult to remove (see Appendix Figures 5-8), added to the fact that it wasn't entirely clear which plasmid map to trust. Over many repeats testing various cut sites between the maps the correct map was eventually identified (see Materials & Methods Figure 2.2). Following this, the sequence has since been successfully removed by a post-doctoral researcher in the lab; however, the over expression assays could not be conducted within the remaining time frame of the project.

It will be important to conduct the same assays as presented in the chapter to see if XRN1 rescue affects proliferation, apoptosis, translation and cell viability. This will support the notion for XRN1 having a role in cancer progression.

This project has only investigated the cellular effects of XRN1 by knocking down XRN1 using a lipid-based siRNA transfection system. Naturally, this method does not result in 100% depletion of XRN1, and any residual expression may explain why no phenotypes could be observed. Future work to determine whether this is the case includes the use of a CRISPR-Cas9 knock out cell line, whereby XRN1 expression can be fully depleted. At the time of testing, this method was considered by the author, however, work previously published by the Newbury lab had shown that XRN1 knock out in *Drosophila* pupae resulted in organismal lethality (Waldron

et al. 2015). In this case, the use of CRISPR-Cas9 to knock out XRN1 in cell lines was decided against. At present, it would be advantageous to perform knock out experiments in human cell lines, because lethality may not be conserved between these two models. Equally, a catalytically inactive mutant XRN1 could also be employed to achieve 100% activity depletion. Studies in yeast have shown that various point mutations in the active site, including D208A, results in a catalytically inactive XRN1, whilst also maintaining organism viability (Medina *et al.* 2014). Overall, it is clear that given there is no observable phenotype whilst there is still residual XRN1 activity, there needs to be phenotypic investigations performed when XRN1 activity is completely ablated.

6.8.3 Identifying potential transcript targets of XRN1 using RNA sequencing

Chapter 5 shows that when XRN1 is knocked down in human SAOS-2 cells there is evidence of post-transcriptional up regulation of a subset of genes. It also showed that some genes were consistently differentially expressed in HeLa cells after XRN1 knock down (Lubas *et al.* 2013). GO analysis showed enrichment of genes involved in major cell pathways, such as the EGF pathway, general cellular organelle maintenance and cell adhesion. For future work in this aspect of the project, it is vitally important to perform qRT-PCR validation of the up regulated genes, as this will show conclusively that XRN1 is involved in the control of gene expression of these genes. It will back up the evidence elucidated by the bioinformatical algorithms to show that specific genes are being post-transcriptionally up regulated, and this can be also be shown by detection of the pre-mRNA of these genes.

Alongside validating predicted targets using qRT-PCR, it would be useful to conduct RNA sequencing experiments (or targeted qRT-PCR) on a number of other cell lines in the future, to see if the same genes are being up regulated when XRN1 is knocked down, and to rule cell-type specific effects. The RNA sequencing data obtained by knocking down XRN1 in HeLa cells was very interesting as it showed that XRN1 may be targeting the same transcripts in different tissue types, however, HeLa cells are notoriously difficult to trust. Their high number of chromosomal aberrations mean that they are not typical human cells, and so not a reliable source to obtain this type of information. By conducting RNA sequencing on other cell types where XRN1 is knocked down (such as Ewing sarcoma cells – see Chapter 3), it will be easier to identify if XRN1 is targeting similar or the same types of transcripts in different cell types. If XRN1 is targeting the same transcripts, it would suggest that its activity is not cell-type specific, and that it acts ubiquitously throughout all tissues. If it is acting ubiquitously, then it would suggest that XRN1 has the same effect in multiple tissues. If there is a role for XRN1 in cancer progression, then it

would make sense that XRN1 could be dysregulated in multiple types of cancer. This could be important in the future development of cancer therapies.

It would also be useful to conduct qRT-PCR experiments in cell lines where XRN1 expression is naturally lower, to ascertain whether the genes identified as potential targets of XRN1 are differentially expressed compared to the control. This would provide more evidence that these genes are true targets of XRN1, added to the bioinformatic predictions. Alongside this, it would be interesting to observe the decay rates of the potential targets of XRN1 in both knock down cell types versus control and in the cell lines where XRN1 expression is naturally lower. It would be presumed that if the targets are truly being degraded by XRN1, then the half-lives of the particular mRNAs would increase. This can be shown by using a transcriptional inhibitor, such as actinomycin-D or α-amanitin and subsequent qRT-PCR.

6.8.4 Work to be continued

Overall, this section has identified areas of work which need to be carried out in the future in order to characterise the functions of XRN1 in human cells more fully with regards to its activity in the RNA degradation pathway, such as how it regulates specific RNAs to maintain cellular homeostasis. It is important for future publications that this work is carried out in order to ensure the correct information is being published, and could be started by the creation (using CRISPR technology) of a complete XRN1 knock out cell line, or a cell line where XRN1 is catalytically dead, to ensure complete efficient depletion of XRN1. Incomplete knock down may not be sufficient to see the overall phenotype in these already immortalized cell lines, and so this step would offer the first confirmations on the results shown in this thesis.

6.9 Concluding Remarks

This project has shown that to elucidate the role of a single enzyme in human cells using both traditional biochemical techniques, alongside more modern bioinformatical approaches, is a difficult task. This project has shown that XRN1 displays lower expression in both OS and EWS cells when compared to the HOb control cell line, as well as in OS patient samples. It has shown that although there is a reduction of XRN1 in these cells, the artificial knock down of XRN1 using siRNA does not result in an observable phenotype in human cell lines. RNA sequencing has shown that despite there being no observable phenotype, there was a large increase in enriched genes that were involved in the EGF pathway, and other fundamental cellular pathways.

It is clear from the body of work presented that there is still scope for investigation to be done with regard to the activity of XRN1 in OS cells, however, it is also clear that across the literature there is also debate about the activity of XRN1 in many different types of cell and organism. This project has tried to elucidate a general role for XRN1 in human cells, given that there was no observable phenotype, particularly one that is associated with the hallmarks of cancer: proliferation, apoptosis, cell viability and later with the RNA sequencing data, autophagy.

From this, it is not clear whether XRN1 has a role in cancer progression. It is also not clear that OS cells are a suitable model to look for the cellular effects of XRN1. There are conflicting reports across the literature about the role of XRN1 across both tissue type and organism, with different phenotypes being observed in yeast species compared to *Drosophila* models and human cell lines. It is probable that XRN1 targets different RNAs in different cell types according to the function of the cell, and what the cell needs in order to function as part of the tissue it belongs to. In addition, this is also the case for other exoribonucleases in the cell. This was shown both by growth curve experiments in two different cell types after DIS3L2 knock down, and the exceptionally low expression of DIS3L1 in bone cells compared to other cells as shown by the Human Protein Atlas.

Greater analysis of the targets of XRN1, and indeed the other exoribonucleases, in human cells will answer many questions about RNA stability in humans. Thus far, there have been many associations of RNA stability with the onset of disease, including how mutations in DIS3 and DIS3L2 can lead to different disorders, however, it is not altogether clear how the dysregulation of the different elements of the RNA stability pathway leads to these pathologies. Greater analysis of this needs to be undertaken in the future, alongside the work discussed in this chapter.

Bibliography

Abdelmohsen, K., A. C. Panda, R. Munk, I. Grammatikakis, D. B. Dudekula, S. De, J. Kim, J. H. Noh, K. M. Kim, J. L. Martindale and M. Gorospe (2017). "Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1." RNA Biol **14**(3): 361-369.

Abernathy, E., S. Gilbertson, R. Alla and B. Glaunsinger (2015). "Viral Nucleases Induce an mRNA Degradation-Transcription Feedback Loop in Mammalian Cells." Cell Host Microbe **18**(2): 243-253.

Alberti, S. and A. A. Hyman (2016). "Are aberrant phase transitions a driver of cellular aging?" Bioessays **38**(10): 959-968.

Ameres, S. L., M. D. Horwich, J.-H. Hung, J. Xu, M. Ghildiyal, Z. Weng and P. D. Zamore (2010). "Target RNA–Directed Trimming and Tailing of Small Silencing RNAs." Science **328**(5985): 1534.

Ando, K., M. F. Heymann, V. Stresing, K. Mori, F. Redini and D. Heymann (2013). "Current therapeutic strategies and novel approaches in osteosarcoma." Cancers (Basel) **5**(2): 591-616.

Andreou, D., S. S. Bielack, D. Carrle, M. Kevric, R. Kotz, W. Winkelmann, G. Jundt, M. Werner, S. Fehlberg, L. Kager, T. Kuhne, S. Lang, M. Dominkus, G. U. Exner, J. Hardes, A. Hillmann, V. Ewerbeck, U. Heise, P. Reichardt and P. U. Tunn (2011). "The influence of tumor- and treatment-related factors on the development of local recurrence in osteosarcoma after adequate surgery. An analysis of 1355 patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols." Ann Oncol **22**(5): 1228-1235.

Antic, S., M. T. Wolfinger, A. Skucha, S. Hosiner and S. Dorner (2015). "General and MicroRNA-Mediated mRNA Degradation Occurs on Ribosome Complexes in Drosophila Cells." Mol Cell Biol **35**(13): 2309-2320.

Aravin, A. A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, K. F. Toth, T. Bestor and G. J. Hannon (2008). "A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice." Mol Cell **31**(6): 785-799.

Astuti, D., M. R. Morris, W. N. Cooper, R. H. Staals, N. C. Wake, G. A. Fews, H. Gill, D. Gentle, S. Shuib, C. J. Ricketts, T. Cole, A. J. van Essen, R. A. van Lingen, G. Neri, J. M. Opitz, P. Rump, I. Stolte-Dijkstra, F. Muller, G. J. Pruijn, F. Latif and E. R. Maher (2012). "Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility." Nat Genet **44**(3): 277-284.

Atkins, J. F. and R. Gesteland (2002). "Biochemistry. The 22nd amino acid." Science **296**(5572): 1409-1410.

Ayache, J., M. Benard, M. Ernoult-Lange, N. Minshall, N. Standart, M. Kress and D. Weil (2015). "P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes." Mol Biol Cell **26**(14): 2579-2595.

Baccarini, A., H. Chauhan, Thomas J. Gardner, Anitha D. Jayaprakash, R. Sachidanandam and Brian D. Brown (2011). "Kinetic Analysis Reveals the Fate of a MicroRNA following Target Regulation in Mammalian Cells." Current Biology **21**(5): 369-376.

Bachmayr-Heyda, A., A. T. Reiner, K. Auer, N. Sukhbaatar, S. Aust, T. Bachleitner-Hofmann, I. Mesteri, T. W. Grunt, R. Zeillinger and D. Pils (2015). "Correlation of circular RNA abundance with proliferation--exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis, and normal human tissues." Sci Rep **5**: 8057.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant, L. Clementi, J. Ren, W. W. Li and W. S. Noble (2009). "MEME SUITE: tools for motif discovery and searching." Nucleic Acids Res **37**(Web Server issue): W202-208.

Bakheet, T., M. Frevel, B. R. Williams, W. Greer and K. S. Khabar (2001). "ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins." Nucleic Acids Res **29**(1): 246-254.

Bakheet, T., E. Hitti, M. Al-Saif, W. N. Moghrabi and K. S. A. Khabar (2018). "The AU-rich element landscape across human transcriptome reveals a large proportion in introns and regulation by ELAVL1/HuR." Biochim Biophys Acta Gene Regul Mech **1861**(2): 167-177.

Baroy, T., C. S. Chilamakuri, S. Lorenz, J. Sun, O. S. Bruland, O. Myklebost and L. A. Meza-Zepeda (2016). "Genome Analysis of Osteosarcoma Progression Samples Identifies FGFR1 Overexpression as a Potential Treatment Target and CHM as a Candidate Tumor Suppressor Gene." PLoS One **11**(9): e0163859.
Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-297.
Bashkirov, V. I., H. Scherthan, J. A. Solinger, J. M. Buerstedde and W. D. Heyer (1997). "A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates." J Cell Biol **136**(4): 761-773.
Behjati, S., P. S. Tarpey, K. Haase, H. Ye, M. D. Young, L. B. Alexandrov, S. J. Farndon, G. Collord, D. C. Wedge, I. Martincorena, S. L. Cooke, H. Davies, W. Mifsud, M. Lidgren, S. Martin, C. Latimer, M. Maddison, A. P. Butler, J. W. Teague, N. Pillay, A. Shlien, U. McDermott, P. A. Futreal, D. Baumhoer, O. Zaikova, B. Bjerkehagen, O. Myklebost, M. F. Amary, R. Tirabosco, P. Van Loo, M. R. Stratton, A. M. Flanagan and P. J. Campbell (2017). "Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma." Nat Commun **8**: 15936.
Benhamed, M., U. Herbig, T. Ye, A. Dejean and O. Bischof (2012). "Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells." Nat Cell Biol **14**(3): 266-275.
Bentley, D. L. (2005). "Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors." Curr Opin Cell Biol **17**(3): 251-256.
Bhattacharya, A., J. D. Ziebarth and Y. Cui (2012). "Systematic analysis of microRNA targeting impacted by small insertions and deletions in human genome." PLoS One **7**(9): e46176.
Bhowmick, R., A. Mukherjee, U. Patra and M. Chawla-Sarkar (2015). "Rotavirus disrupts cytoplasmic P bodies during infection." Virus Res **210**: 344-354.
Blasco-Moreno, B., L. de Campos-Mata, R. Böttcher, J. García-Martínez, J. Jungfleisch, D. D. Nedialkova, S. Chattopadhyay, M.-E. Gas, B. Oliva, J. E. Pérez-Ortín, S. A. Leidel, M. Choder and J. Díez (2019). "The exonuclease Xrn1 activates transcription and translation of mRNAs encoding membrane proteins." Nature Communications **10**(1): 1298.
Bocciardi, R., R. Giorda, J. Buttgereit, S. Gimelli, M. T. Divizia, S. Beri, S. Garofalo, S. Tavella, M. Lerone, O. Zuffardi, M. Bader, R. Ravazzolo and G. Gimelli (2007). "Overexpression of the C-type natriuretic peptide (CNP) is associated with overgrowth and bone anomalies in an individual with balanced t(2;7) translocation." Hum Mutat **28**(7): 724-731.
Bonci, D. and R. De Maria (2016). "miR-15/miR-16 loss, miR-21 upregulation, or deregulation of their target genes predicts poor prognosis in prostate cancer patients." Mol Cell Oncol **3**(4): e1109744.
Boonstra, A., L. J. van der Laan, T. Vanwolleghem and H. L. Janssen (2009). "Experimental models for hepatitis C viral infection." Hepatology **50**(5): 1646-1655.
Boroughs, L. K. and R. J. DeBerardinis (2015). "Metabolic pathways promoting cancer cell survival and growth." Nat Cell Biol **17**(4): 351-359.
Braun, J. E., E. Huntzinger and E. Izaurralde (2012). "A molecular link between miRISCs and deadenylases provides new insight into the mechanism of gene silencing by microRNAs." Cold Spring Harb Perspect Biol **4**(12).
Braun, J. E., E. Huntzinger and E. Izaurralde (2013). "The role of GW182 proteins in miRNA-mediated gene silencing." Adv Exp Med Biol **768**: 147-163.
Braun, J. E., V. Truffault, A. Boland, E. Huntzinger, C. T. Chang, G. Haas, O. Weichenrieder, M. Coles and E. Izaurralde (2012). "A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation." Nat Struct Mol Biol **19**(12): 1324-1331.
Braun, K. A. and E. T. Young (2014). "Coupling mRNA Synthesis and Decay." Molecular and Cellular Biology **34**(22): 4078.
Brook, M. and N. K. Gray (2012). "The role of mammalian poly(A)-binding proteins in co-ordinating mRNA turnover." Biochem Soc Trans **40**(4): 856-864.
Bullock, M. D., K. M. Pickard, B. S. Nielsen, A. E. Sayan, V. Jenei, M. Mellone, R. Mitter, J. N. Primrose, G. J. Thomas, G. K. Packham and A. H. Mirnezami (2013). "Pleiotropic actions of miR-

21 highlight the critical role of deregulated stromal microRNAs during colorectal cancer progression." Cell Death Dis **4**(6): e684.
Burris, A. M., B. J. Ballew, J. B. Kentosh, C. E. Turner, S. A. Norton, N. Giri, B. P. Alter, A. Nellan, C. Gamper, K. R. Hartman and S. A. Savage (2016). "Hoyeraal-Hreidarsson Syndrome due to PARN Mutations: Fourteen Years of Follow-Up." Pediatr Neurol **56**: 62-68.e61.
Casamassimi, A. and C. Napoli (2007). "Mediator complexes and eukaryotic transcription regulation: an overview." Biochimie **89**(12): 1439-1446.
Casse, C., F. Giannoni, V. T. Nguyen, M. F. Dubois and O. Bensaude (1999). "The transcriptional inhibitors, actinomycin D and alpha-amanitin, activate the HIV-1 promoter and favor phosphorylation of the RNA polymerase II C-terminal domain." J Biol Chem **274**(23): 16097-16106.
Caudy, A. A., R. F. Ketting, S. M. Hammond, A. M. Denli, A. M. Bathoorn, B. B. Tops, J. M. Silva, M. M. Myers, G. J. Hannon and R. H. Plasterk (2003). "A micrococcal nuclease homologue in RNAi effector complexes." Nature **425**(6956): 411-414.
Chandradoss, S. D., N. T. Schirle, M. Szczepaniak, I. J. MacRae and C. Joo (2015). "A Dynamic Search Process Underlies MicroRNA Targeting." Cell **162**(1): 96-107.
Chang, H. M., R. Triboulet, J. E. Thornton and R. I. Gregory (2013). "A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway." Nature **497**(7448): 244-248.
Chantarachot, T. and J. Bailey-Serres (2018). "Polysomes, Stress Granules, and Processing Bodies: A Dynamic Triumvirate Controlling Cytoplasmic mRNA Fate and Function." Plant Physiol **176**(1): 254-269.
Chapman, E. G., D. A. Costantino, J. L. Rabe, S. L. Moon, J. Wilusz, J. C. Nix and J. S. Kieft (2014). "The structural basis of pathogenic subgenomic flavivirus RNA (sfRNA) production." Science **344**(6181): 307-310.
Chapman, E. G., S. L. Moon, J. Wilusz and J. S. Kieft (2014). "RNA structures that resist degradation by Xrn1 produce a pathogenic Dengue virus RNA." Elife **3**: e01892.
Chapman, M. A., M. S. Lawrence, J. J. Keats, K. Cibulskis, C. Sougnez, A. C. Schinzel, C. L. Harview, J. P. Brunet, G. J. Ahmann, M. Adli, K. C. Anderson, K. G. Ardlie, D. Auclair, A. Baker, P. L. Bergsagel, B. E. Bernstein, Y. Drier, R. Fonseca, S. B. Gabriel, C. C. Hofmeister, S. Jagannath, A. J. Jakubowiak, A. Krishnan, J. Levy, T. Liefeld, S. Lonial, S. Mahan, B. Mfuko, S. Monti, L. M. Perkins, R. Onofrio, T. J. Pugh, S. V. Rajkumar, A. H. Ramos, D. S. Siegel, A. Sivachenko, A. K. Stewart, S. Trudel, R. Vij, D. Voet, W. Winckler, T. Zimmerman, J. Carpten, J. Trent, W. C. Hahn, L. A. Garraway, M. Meyerson, E. S. Lander, G. Getz and T. R. Golub (2011). "Initial genome sequencing and analysis of multiple myeloma." Nature **471**(7339): 467-472.
Charley, P. A., C. J. Wilusz and J. Wilusz (2018). "Identification of phlebovirus and arenavirus RNA sequences that stall and repress the exoribonuclease XRN1." J Biol Chem **293**(1): 285-295.
Chatterjee, S. and H. Großhans (2009). "Active turnover modulates mature microRNA activity in Caenorhabditis elegans." Nature **461**: 546.
Chekulaeva, M., H. Mathys, J. T. Zipprich, J. Attig, M. Colic, R. Parker and W. Filipowicz (2011). "miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs." Nat Struct Mol Biol **18**(11): 1218-1226.
Chen, C. Y., N. Ezzeddine and A. B. Shyu (2008). "Messenger RNA half-life measurements in mammalian cells." Methods Enzymol **448**: 335-357.
Chen, M., Y. Xu, J. Zhao, W. Zhong, L. Zhang, Y. Bi and M. Wang (2019). "Concurrent Driver Gene Mutations as Negative Predictive Factors in Epidermal Growth Factor Receptor-Positive Non-Small Cell Lung Cancer." EBioMedicine.
Chen, R., G. Wang, Y. Zheng, Y. Hua and Z. Cai (2017). "Long non-coding RNAs in osteosarcoma." Oncotarget **8**(12): 20462-20475.
Chen, X., A. Bahrami, A. Pappo, J. Easton, J. Dalton, E. Hedlund, D. Ellison, S. Shurtleff, G. Wu, L. Wei, M. Parker, M. Rusch, P. Nagahawatte, J. Wu, S. Mao, K. Boggs, H. Mulder, D. Yergeau, C. Lu, L. Ding, M. Edmonson, C. Qu, J. Wang, Y. Li, F. Navid, N. C. Daw, E. R. Mardis, R. K. Wilson, J. R. Downing, J. Zhang and M. A. Dyer (2014). "Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma." Cell Rep **7**(1): 104-112.

Cheng, L., F. Li, Y. Jiang, H. Yu, C. Xie, Y. Shi and Q. Gong (2019). "Structural insights into a unique preference for 3' terminal guanine of mirtron in Drosophila TUTase tailor." Nucleic Acids Res **47**(1): 495-508.
Chou, C. H., S. Shrestha, C. D. Yang, N. W. Chang, Y. L. Lin, K. W. Liao, W. C. Huang, T. H. Sun, S. J. Tu, W. H. Lee, M. Y. Chiew, C. S. Tai, T. Y. Wei, T. R. Tsai, H. T. Huang, C. Y. Wang, H. Y. Wu, S. Y. Ho, P. R. Chen, C. H. Chuang, P. J. Hsieh, Y. S. Wu, W. L. Chen, M. J. Li, Y. C. Wu, X. Y. Huang, F. L. Ng, W. Buddhakosai, P. C. Huang, K. C. Lan, C. Y. Huang, S. L. Weng, Y. N. Cheng, C. Liang, W. L. Hsu and H. D. Huang (2018). "miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions." Nucleic Acids Res **46**(D1): D296-d302.
Chowdhury, A., J. Mukhopadhyay and S. Tharun (2007). "The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs." Rna **13**(7): 998-1016.
Clark, A. G. and D. M. Vignjevic (2015). "Modes of cancer cell invasion and the role of the microenvironment." Current Opinion in Cell Biology **36**: 13-22.
Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." Genes Dev **11**(21): 2755-2766.
Conti, E. and E. Izaurralde (2005). "Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species." Curr Opin Cell Biol **17**(3): 316-325.
Cougot, N., S. Babajko and B. Séraphin (2004). "Cytoplasmic foci are sites of mRNA decay in human cells." J Cell Biol **165**(1): 31-40.
Covarrubias, S., J. M. Richner, K. Clyde, Y. J. Lee and B. A. Glaunsinger (2009). "Host shutoff is a conserved phenotype of gammaherpesvirus infection and is orchestrated exclusively from the cytoplasm." J Virol **83**(18): 9554-9566.
Dai, N., Y. Qing, Y. Cun, Z. Zhong, C. Li, S. Zhang, J. Shan, X. Yang, X. Dai, Y. Cheng, H. Xiao, C. Xu, M. Li and D. Wang (2018). "miR-513a-5p regulates radiosensitivity of osteosarcoma by targeting human apurinic/apyrimidinic endonuclease." Oncotarget **9**(39): 25414-25426.
Dasari, S. and P. B. Tchounwou (2014). "Cisplatin in cancer therapy: molecular mechanisms of action." Eur J Pharmacol **740**: 364-378.
DeBerardinis, R. J. and N. S. Chandel (2016). "Fundamentals of cancer metabolism." Sci Adv **2**(5): e1600200.
Decker, C. J. and R. Parker (2012). "P-bodies and stress granules: possible roles in the control of translation and mRNA degradation." Cold Spring Harb Perspect Biol **4**(9): a012286.
Delattre, O., J. Zucman, B. Plougastel, C. Desmaze, T. Melot, M. Peter, H. Kovar, I. Joubert, P. de Jong, G. Rouleau and et al. (1992). "Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours." Nature **359**(6391): 162-165.
Delorme-Axford, E., E. Abernathy, N. J. Lennemann, A. Bernard, A. Ariosa, C. B. Coyne, K. Kirkegaard and D. J. Klionsky (2018). "The exoribonuclease Xrn1 is a post-transcriptional negative regulator of autophagy." Autophagy **14**(5): 898-912.
Dengl, S. and P. Cramer (2009). "Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro." J Biol Chem **284**(32): 21270-21279.
Deniaud, A., M. Karuppasamy, T. Bock, S. Masiulis, K. Huard, F. Garzoni, K. Kerschgens, M. W. Hentze, A. E. Kulozik, M. Beck, G. Neu-Yilik and C. Schaffitzel (2015). "A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation." Nucleic Acids Res **43**(15): 7600-7611.
Dever, T. E. and R. Green (2012). "The elongation, termination, and recycling phases of translation in eukaryotes." Cold Spring Harb Perspect Biol **4**(7): a013706.
Dhanraj, S., S. M. Gunja, A. P. Deveau, M. Nissbeck, B. Boonyawat, A. J. Coombs, A. Renieri, M. Mucciolo, A. Marozza, S. Buoni, L. Turner, H. Li, A. Jarrar, M. Sabanayagam, M. Kirby, M. Shago, D. Pinto, J. N. Berman, S. W. Scherer, A. Virtanen and Y. Dror (2015). "Bone marrow failure and developmental delay caused by mutations in poly(A)-specific ribonuclease (PARN)." J Med Genet **52**(11): 738-748.

Dhillon, P. and C. D. Rao (2018). "Rotavirus Induces Formation of Remodeled Stress Granules and P Bodies and Their Sequestration in Viroplasms To Promote Progeny Virus Production." J Virol **92**(24).
Ding, L., T. J. Ley, D. E. Larson, C. A. Miller, D. C. Koboldt, J. S. Welch, J. K. Ritchey, M. A. Young, T. Lamprecht, M. D. McLellan, J. F. McMichael, J. W. Wallis, C. Lu, D. Shen, C. C. Harris, D. J. Dooling, R. S. Fulton, L. L. Fulton, K. Chen, H. Schmidt, J. Kalicki-Veizer, V. J. Magrini, L. Cook, S. D. McGrath, T. L. Vickery, M. C. Wendl, S. Heath, M. A. Watson, D. C. Link, M. H. Tomasson, W. D. Shannon, J. E. Payton, S. Kulkarni, P. Westervelt, M. J. Walter, T. A. Graubert, E. R. Mardis, R. K. Wilson and J. F. DiPersio (2012). "Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing." Nature **481**(7382): 506-510.
Ding, M., R. K. Bruick and Y. Yu (2016). "Secreted IGFBP5 mediates mTORC1-dependent feedback inhibition of IGF-1 signalling." Nat Cell Biol **18**(3): 319-327.
Ding, M., B. Lin, T. Li, Y. Liu, Y. Li, X. Zhou, M. Miao, J. Gu, H. Pan, F. Yang, T. Li, X. Y. Liu and R. Li (2015). "A dual yet opposite growth-regulating function of miR-204 and its target XRN1 in prostate adenocarcinoma cells and neuroendocrine-like prostate cancer cells." Oncotarget **6**(10): 7686-7700.
Drake, K. M., E. C. Ruteshouser, R. Natrajan, P. Harbor, J. Wegert, M. Gessler, K. Pritchard-Jones, P. Grundy, J. Dome, V. Huff, C. Jones and M. A. Aldred (2009). "Loss of heterozygosity at 2q37 in sporadic Wilms' tumor: putative role for miR-562." Clin Cancer Res **15**(19): 5985-5992.
Duchaine, T. F. and M. R. Fabian (2019). "Mechanistic Insights into MicroRNA-Mediated Gene Silencing." Cold Spring Harb Perspect Biol **11**(3).
Ducy, P., M. Starbuck, M. Priemel, J. Shen, G. Pinero, V. Geoffroy, M. Amling and G. Karsenty (1999). "A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development." Genes Dev **13**(8): 1025-1036.
Eaton, J. D., L. Davidson, D. L. V. Bauer, T. Natsume, M. T. Kanemaki and S. West (2018). "Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity." Genes Dev **32**(2): 127-139.
Eberharter, A. and P. B. Becker (2002). "Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics." EMBO Rep **3**(3): 224-229.
Ebert, M. S. and P. A. Sharp (2010). "MicroRNA sponges: progress and possibilities." Rna **16**(11): 2043-2050.
Edfors, F., F. Danielsson, B. M. Hallström, L. Käll, E. Lundberg, F. Pontén, B. Forsström and M. Uhlén (2016). "Gene-specific correlation of RNA and protein levels in human cells and tissues." Mol Syst Biol **12**(10): 883.
Elbarbary, R. A., K. Miyoshi, O. Hedaya, J. R. Myers and L. E. Maquat (2017). "UPF1 helicase promotes TSN-mediated miRNA decay." Genes Dev **31**(14): 1483-1493.
Elkayam, E., C. R. Faehnle, M. Morales, J. Sun, H. Li and L. Joshua-Tor (2017). "Multivalent Recruitment of Human Argonaute by GW182." Mol Cell **67**(4): 646-658.e643.
Emadi, A., R. J. Jones and R. A. Brodsky (2009). "Cyclophosphamide and cancer: golden anniversary." Nat Rev Clin Oncol **6**(11): 638-647.
Ernoult-Lange, M., S. Baconnais, M. Harper, N. Minshall, S. Souquere, T. Boudier, M. Benard, P. Andrey, G. Pierron, M. Kress, N. Standart, E. le Cam and D. Weil (2012). "Multiple binding of repressed mRNAs by the P-body protein Rck/p54." Rna **18**(9): 1702-1715.
Eulalio, A., I. Behm-Ansmant and E. Izaurralde (2007). "P bodies: at the crossroads of post-transcriptional pathways." Nat Rev Mol Cell Biol **8**(1): 9-22.
Eulalio, A., I. Behm-Ansmant, D. Schweizer and E. Izaurralde (2007). "P-body formation is a consequence, not the cause, of RNA-mediated gene silencing." Mol Cell Biol **27**(11): 3970-3981.
Fabian, M. R. and N. Sonenberg (2012). "The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC." Nature Structural &Amp; Molecular Biology **19**: 586.
Fabre, A. and C. Badens (2014). "Human Mendelian diseases related to abnormalities of the RNA exosome or its cofactors." Intractable Rare Dis Res **3**(1): 8-11.

Finkel, T. (2012). "From sulfenylation to sulfhydration: what a thiolate needs to tolerate." Sci Signal **5**(215): pe10.
Fiorino, S., L. Bacchi-Reggiani, D. de Biase, A. Fornelli, M. Masetti, A. Tura, F. Grizzi, M. Zanello, L. Mastrangelo, R. Lombardi, G. Acquaviva, L. di Tommaso, A. Bondi, M. Visani, S. Sabbatani, L. Pontoriero, C. Fabbri, A. Cuppini, A. Pession and E. Jovine (2015). "Possible association between hepatitis C virus and malignancies different from hepatocellular carcinoma: A systematic review." World J Gastroenterol **21**(45): 12896-12953.
Fiorino, S., E. Chili, L. Bacchi-Reggiani, M. Masetti, G. Deleonardi, A. G. Grondona, T. Silvestri, E. Magrini, N. Zanini, A. Cuppini, R. Nardi and E. Jovine (2013). "Association between hepatitis B or hepatitis C virus infection and risk of pancreatic adenocarcinoma development: a systematic review and meta-analysis." Pancreatology **13**(2): 147-160.
Flobinus, A., N. Chevigny, P. A. Charley, T. Seissler, E. Klein, C. Bleykasten-Grosshans, C. Ratti, S. Bouzoubaa, J. Wilusz and D. Gilmer (2018). "Beet Necrotic Yellow Vein Virus Noncoding RNA Production Depends on a 5'→3' Xrn Exoribonuclease Activity." Viruses **10**(3).
Fulda, S., A. M. Gorman, O. Hori and A. Samali (2010). "Cellular stress responses: cell survival and cell death." Int J Cell Biol **2010**: 214074.
Gaglia, M. M., S. Covarrubias, W. Wong and B. A. Glaunsinger (2012). "A common strategy for host RNA degradation by divergent viruses." J Virol **86**(17): 9527-9530.
Gallouzi, I. E. and J. Wilusz (2013). "A DIStinctively novel exoribonuclease that really likes U." Embo j **32**(13): 1799-1801.
Gao, J., T. T. Yang, X. C. Qiu, B. Yu, J. W. Han, Q. Y. Fan and B. A. Ma (2007). "[Cloning and identification of microRNA from human osteosarcoma cell line SOSP-9607]." Ai Zheng **26**(6): 561-565.
Garcia-Moreno, M., M. Noerenberg, S. Ni, A. I. Jarvelin, E. Gonzalez-Almela, C. E. Lenz, M. Bach-Pages, V. Cox, R. Avolio, T. Davis, S. Hester, T. J. M. Sohier, B. Li, G. Heikel, G. Michlewski, M. A. Sanz, L. Carrasco, E. P. Ricci, V. Pelechano, I. Davis, B. Fischer, S. Mohammed and A. Castello (2019). "System-wide Profiling of RNA-Binding Proteins Uncovers Key Regulators of Virus Infection." Mol Cell **74**(1): 196-211.e111.
Gatica, D. and D. J. Klionsky (2019). "Towards understanding mRNA-binding protein specificity: lessons from post-transcriptional regulation of ATG mRNA during nitrogen starvation-induced autophagy." Curr Genet.
Geerlings, T. H., J. C. Vos and H. A. Raue (2000). "The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-->3' exonucleases." Rna **6**(12): 1698-1703.
Geng, S., L. Gu, F. Ju, H. Zhang, Y. Wang, H. Tang, Z. Bi and C. Yang (2016). "MicroRNA-224 promotes the sensitivity of osteosarcoma cells to cisplatin by targeting Rac1." J Cell Mol Med **20**(9): 1611-1619.
Ghigna, C., C. Valacca and G. Biamonti (2008). "Alternative splicing and tumor progression." Curr Genomics **9**(8): 556-570.
Gibbings, D. J., C. Ciaudo, M. Erhardt and O. Voinnet (2009). "Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity." Nat Cell Biol **11**(9): 1143-1149.
Goodman, C. A. and T. A. Hornberger (2013). "Measuring protein synthesis with SUnSET: a valid alternative to traditional techniques?" Exerc Sport Sci Rev **41**(2): 107-115.
Govoni, K. E. (2012). "Insulin-like growth factor-I molecular pathways in osteoblasts: potential targets for pharmacological manipulation." Curr Mol Pharmacol **5**(2): 143-152.
Granneman, S. and S. J. Baserga (2004). "Ribosome biogenesis: of knobs and RNA processing." Exp Cell Res **296**(1): 43-50.
Grousl, T., M. Opekarova, V. Stradalova, J. Hasek and J. Malinsky (2015). "Evolutionarily conserved 5'-3' exoribonuclease Xrn1 accumulates at plasma membrane-associated eisosomes in post-diauxic yeast." PLoS One **10**(3): e0122770.
Gruber, A. J., R. Schmidt, A. R. Gruber, G. Martin, S. Ghosh, M. Belmadani, W. Keller and M. Zavolan (2016). "A comprehensive analysis of 3' end sequencing data sets reveals novel

polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation." Genome Res **26**(8): 1145-1159.
Guo, L., I. Vlasova-St Louis and P. R. Bohjanen (2018). "Viral manipulation of host mRNA decay." Future Virol **13**(3): 211-223.
Gutschner, T., M. Hammerle, M. Eissmann, J. Hsu, Y. Kim, G. Hung, A. Revenko, G. Arun, M. Stentrup, M. Gross, M. Zornig, A. R. MacLeod, D. L. Spector and S. Diederichs (2013). "The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells." Cancer Res **73**(3): 1180-1189.
Guzikowski, A. R., Y. S. Chen and B. M. Zid (2019). "Stress-induced mRNP granules: Form and function of processing bodies and stress granules." Wiley Interdiscip Rev RNA: e1524.
Haimovich, G., D. A. Medina, S. Z. Causse, M. Garber, G. Millan-Zambrano, O. Barkai, S. Chavez, J. E. Perez-Ortin, X. Darzacq and M. Choder (2013). "Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis." Cell **153**(5): 1000-1011.
Halbach, F., P. Reichelt, M. Rode and E. Conti (2013). "The yeast ski complex: crystal structure and RNA channeling to the exosome complex." Cell **154**(4): 814-826.
Han, X., Y. Wei, H. Wang, F. Wang, Z. Ju and T. Li (2018). "Nonsense-mediated mRNA decay: a 'nonsense' pathway makes sense in stem cell biology." Nucleic Acids Res **46**(3): 1038-1051.
Hanson, G. and J. Coller (2018). "Codon optimality, bias and usage in translation and mRNA decay." Nat Rev Mol Cell Biol **19**(1): 20-30.
Hart, T., H. K. Komori, S. LaMere, K. Podshivalova and D. R. Salomon (2013). "Finding the active genes in deep RNA-seq gene expression studies." BMC Genomics **14**: 778.
He, F., A. Celik, C. Wu and A. Jacobson (2018). "General decapping activators target different subsets of inefficiently translated mRNAs." Elife **7**.
Hellen, C. U. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." Genes Dev **15**(13): 1593-1612.
Hendriks, I. A. and A. C. Vertegaal (2016). "A comprehensive compilation of SUMO proteomics." Nat Rev Mol Cell Biol **17**(9): 581-595.
Hershey, J. W., N. Sonenberg and M. B. Mathews (2012). "Principles of translational control: an overview." Cold Spring Harb Perspect Biol **4**(12).
Herzlieb, N., B. W. Gallaher, A. Berthold, R. Hille and W. Kiess (2000). "Insulin-like growth factor-I inhibits the progression of human U-2 OS osteosarcoma cells towards programmed cell death through interaction with the IGF-I receptor." Cell Mol Biol (Noisy-le-grand) **46**(1): 71-77.
Hobert, O. (2008). "Gene regulation by transcription factors and microRNAs." Science **319**(5871): 1785-1786.
Houseley, J., J. LaCava and D. Tollervey (2006). "RNA-quality control by the exosome." Nat Rev Mol Cell Biol **7**(7): 529-539.
Hu, G., T. McQuiston, A. Bernard, Y. D. Park, J. Qiu, A. Vural, N. Zhang, S. R. Waterman, N. H. Blewett, T. G. Myers, R. J. Maraia, J. H. Kehrl, G. Uzel, D. J. Klionsky and P. R. Williamson (2015). "A conserved mechanism of TOR-dependent RCK-mediated mRNA degradation regulates autophagy." Nat Cell Biol **17**(7): 930-942.
Hu, W., T. J. Sweet, S. Chamnongpol, K. E. Baker and J. Coller (2009). "Co-translational mRNA decay in Saccharomyces cerevisiae." Nature **461**(7261): 225-229.
Hu, W., T. J. Sweet, S. Chamnongpol, K. E. Baker and J. Coller (2009). "Co-translational mRNA decay in Saccharomyces cerevisiae." Nature **461**: 225.
Huang da, W., B. T. Sherman and R. A. Lempicki (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Res **37**(1): 1-13.
Huang da, W., B. T. Sherman and R. A. Lempicki (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc **4**(1): 44-57.
Huang, L., J. Hwang, S. D. Sharma, M. R. Hargittai, Y. Chen, J. J. Arnold, K. D. Raney and C. E. Cameron (2005). "Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein." J Biol Chem **280**(43): 36417-36428.
Hubstenberger, A., M. Courel, M. Bénard, S. Souquere, M. Ernoult-Lange, R. Chouaib, Z. Yi, J.-B. Morlot, A. Munier, M. Fradet, M. Daunesse, E. Bertrand, G. Pierron, J. Mozziconacci, M. Kress

and D. Weil (2017). "P-Body Purification Reveals the Condensation of Repressed mRNA Regulons." Molecular Cell **68**(1): 144-157.e145.
Hubstenberger, A., M. Courel, M. Benard, S. Souquere, M. Ernoult-Lange, R. Chouaib, Z. Yi, J. B. Morlot, A. Munier, M. Fradet, M. Daunesse, E. Bertrand, G. Pierron, J. Mozziconacci, M. Kress and D. Weil (2017). "P-Body Purification Reveals the Condensation of Repressed mRNA Regulons." Mol Cell **68**(1): 144-157.e145.
Hug, N., D. Longman and J. F. Cáceres (2016). "Mechanism and regulation of the nonsense-mediated decay pathway." Nucleic Acids Res **44**(4): 1483-1495.
Hunter, R. W., Y. Liu, H. Manjunath, A. Acharya, B. T. Jones, H. Zhang, B. Chen, H. Ramalingam, R. E. Hammer, Y. Xie, J. A. Richardson, D. Rakheja, T. J. Carroll and J. T. Mendell (2018). "Loss of Dis3l2 partially phenocopies Perlman syndrome in mice and results in up-regulation of Igf2 in nephron progenitor cells." Genes Dev **32**(13-14): 903-908.
Huppert, J. L., A. Bugaut, S. Kumari and S. Balasubramanian (2008). "G-quadruplexes: the beginning and end of UTRs." Nucleic Acids Res **36**(19): 6260-6268.
Imashimizu, M., N. Shimamoto, T. Oshima and M. Kashlev (2014). "Transcription elongation. Heterogeneous tracking of RNA polymerase and its biological implications." Transcription **5**(1): e28285.
Ingelfinger, D., D. J. Arndt-Jovin, R. Luhrmann and T. Achsel (2002). "The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci." Rna **8**(12): 1489-1501.
Isakoff, M. S., S. S. Bielack, P. Meltzer and R. Gorlick (2015). "Osteosarcoma: Current Treatment and a Collaborative Pathway to Success." J Clin Oncol **33**(27): 3029-3035.
Ivanov, P., N. Kedersha and P. Anderson (2018). "Stress Granules and Processing Bodies in Translational Control." Cold Spring Harb Perspect Biol.
Januszyk, K., Q. Liu and C. D. Lima (2011). "Activities of human RRP6 and structure of the human RRP6 catalytic domain." Rna **17**(8): 1566-1577.
Jarrett, R. F. (2006). "Viruses and lymphoma/leukaemia." J Pathol **208**(2): 176-186.
Jeon, I. S., J. N. Davis, B. S. Braun, J. E. Sublett, M. F. Roussel, C. T. Denny and D. N. Shapiro (1995). "A variant Ewing's sarcoma translocation (7;22) fuses the EWS gene to the ETS gene ETV1." Oncogene **10**(6): 1229-1234.
Jinek, M., S. M. Coyle and J. A. Doudna (2011). "Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay." Mol Cell **41**(5): 600-608.
Jo, M. H., S. Shin, S. R. Jung, E. Kim, J. J. Song and S. Hohng (2015). "Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs." Mol Cell **59**(1): 117-124.
Johnson, J. M., J. Castle, P. Garrett-Engele, Z. Kan, P. M. Loerch, C. D. Armour, R. Santos, E. E. Schadt, R. Stoughton and D. D. Shoemaker (2003). "Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays." Science **302**(5653): 2141-2144.
Jones, C. I., D. P. Grima, J. A. Waldron, S. Jones, H. N. Parker and S. F. Newbury (2013). "The 5'-3' exoribonuclease Pacman (Xrn1) regulates expression of the heat shock protein Hsp67Bc and the microRNA miR-277-3p in Drosophila wing imaginal discs." RNA Biol **10**(8): 1345-1355.
Jones, C. I., A. L. Pashler, B. P. Towler, S. R. Robinson and S. F. Newbury (2016). "RNA-seq reveals post-transcriptional regulation of Drosophila insulin-like peptide dilp8 and the neuropeptide-like precursor Nplp2 by the exoribonuclease Pacman/XRN1." Nucleic Acids Res **44**(1): 267-280.
Jones, C. I., M. V. Zabolotskaya and S. F. Newbury (2012). "The 5' --> 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development." Wiley Interdiscip Rev RNA **3**(4): 455-468.
Kahvejian, A., Y. V. Svitkin, R. Sukarieh, M. N. M'Boutchou and N. Sonenberg (2005). "Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms." Genes Dev **19**(1): 104-113.
Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss and S. Ohno (2006). "Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay." Genes Dev **20**(3): 355-367.

Kedersha, N., G. Stoecklin, M. Ayodele, P. Yacono, J. Lykke-Andersen, M. J. Fritzler, D. Scheuner, R. J. Kaufman, D. E. Golan and P. Anderson (2005). "Stress granules and processing bodies are dynamically linked sites of mRNP remodeling." J Cell Biol **169**(6): 871-884.
Khabar, K. S. (2017). "Hallmarks of cancer and AU-rich elements." Wiley Interdiscip Rev RNA **8**(1).
Khan, A. A., D. Betel, M. L. Miller, C. Sander, C. S. Leslie and D. S. Marks (2009). "Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs." Nature Biotechnology **27**: 549.
Khemici, V. and P. Linder (2018). "RNA helicases in RNA decay." Biochem Soc Trans **46**(1): 163-172.
Koba, M. and J. Konopa (2005). "[Actinomycin D and its mechanisms of action]." Postepy Hig Med Dosw (Online) **59**: 290-298.
Kolb, S. J., S. Sutton and D. R. Schoenberg (2010). "RNA processing defects associated with diseases of the motor neuron." Muscle Nerve **41**(1): 5-17.
Komori, T. (2010). "Regulation of osteoblast differentiation by Runx2." Adv Exp Med Biol **658**: 43-49.
Kovac, M., C. Blattmann, S. Ribi, J. Smida, N. S. Mueller, F. Engert, F. Castro-Giner, J. Weischenfeldt, M. Kovacova, A. Krieg, D. Andreou, P. U. Tunn, H. R. Durr, H. Rechl, K. D. Schaser, I. Melcher, S. Burdach, A. Kulozik, K. Specht, K. Heinimann, S. Fulda, S. Bielack, G. Jundt, I. Tomlinson, J. O. Korbel, M. Nathrath and D. Baumhoer (2015). "Exome sequencing of osteosarcoma reveals mutation signatures reminiscent of BRCA deficiency." Nat Commun **6**: 8940.
Kowalinski, E., A. Kögel, J. Ebert, P. Reichelt, E. Stegmann, B. Habermann and E. Conti (2016). "Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex." Mol Cell **63**(1): 125-134.
Kridis, W. B., N. Toumi, H. Chaari, A. Khanfir, K. Ayadi, H. Keskes, T. Boudawara, J. Daoud and M. Frikha (2017). "A Review of Ewing Sarcoma Treatment: Is it Still a Subject of Debate?" Rev Recent Clin Trials **12**(1): 19-23.
Kurosaki, T., J. R. Myers and L. E. Maquat (2019). "Defining nonsense-mediated mRNA decay intermediates in human cells." Methods **155**: 68-76.
Kushner, S. R. (2004). "mRNA decay in prokaryotes and eukaryotes: different approaches to a similar problem." IUBMB Life **56**(10): 585-594.
Łabno, A., R. Tomecki and A. Dziembowski (2016). "Cytoplasmic RNA decay pathways - Enzymes and mechanisms." Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1863**(12): 3125-3147.
LaCava, J., J. Houseley, C. Saveanu, E. Petfalski, E. Thompson, A. Jacquier and D. Tollervey (2005). "RNA degradation by the exosome is promoted by a nuclear polyadenylation complex." Cell **121**(5): 713-724.
Lauvrak, S. U., E. Munthe, S. H. Kresse, E. W. Stratford, H. M. Namløs, L. A. Meza-Zepeda and O. Myklebost (2013). "Functional characterisation of osteosarcoma cell lines and identification of mRNAs and miRNAs associated with aggressive cancer phenotypes." Br J Cancer **109**(8): 2228-2236.
Lawrence, M. S., P. Stojanov, C. H. Mermel, J. T. Robinson, L. A. Garraway, T. R. Golub, M. Meyerson, S. B. Gabriel, E. S. Lander and G. Getz (2014). "Discovery and saturation analysis of cancer genes across 21 tumour types." Nature **505**(7484): 495-501.
Lee, D., D. Park, J. H. Park, J. H. Kim and C. Shin (2019). "Poly(A)-specific ribonuclease sculpts the 3' ends of microRNAs." Rna **25**(3): 388-405.
Lee, J. E., J. Y. Lee, J. Trembly, J. Wilusz, B. Tian and C. J. Wilusz (2012). "The PARN deadenylase targets a discrete set of mRNAs for decay and regulates cell motility in mouse myoblasts." PLoS Genet **8**(8): e1002901.
Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim and V. N. Kim (2003). "The nuclear RNase III Drosha initiates microRNA processing." Nature **425**(6956): 415-419.

Li, D., T. Wei, C. M. Abbott and D. Harrich (2013). "The Unexpected Roles of Eukaryotic Translation Elongation Factors in RNA Virus Replication and Pathogenesis." Microbiology and Molecular Biology Reviews **77**(2): 253-266.
Li, F., L. Zhang, W. Li, J. Deng, J. Zheng, M. An, J. Lu and Y. Zhou (2015). "Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/beta-catenin pathway." Oncotarget **6**(8): 6001-6013.
Li, W. M., T. Barnes and C. H. Lee (2010). "Endoribonucleases--enzymes gaining spotlight in mRNA metabolism." Febs j **277**(3): 627-641.
Liberti, M. V. and J. W. Locasale (2016). "The Warburg Effect: How Does it Benefit Cancer Cells?" Trends Biochem Sci **41**(3): 211-218.
Libri, V., P. Miesen, R. P. van Rij and A. H. Buck (2013). "Regulation of microRNA biogenesis and turnover by animals and their viruses." Cell Mol Life Sci **70**(19): 3525-3544.
Lin, B. C., D. Huang, C. Q. Yu, Y. Mou, Y. H. Liu, D. W. Zhang and F. J. Shi (2016). "MicroRNA-184 Modulates Doxorubicin Resistance in Osteosarcoma Cells by Targeting BCL2L1." Med Sci Monit **22**: 1761-1765.
Lin, C. J., J. Wen, F. Bejarano, F. Hu, D. Bortolamiol-Becet, L. Kan, P. Sanfilippo, S. Kondo and E. C. Lai (2017). "Characterization of a TUTase/RNase complex required for Drosophila gametogenesis." Rna **23**(3): 284-296.
Lionetti, M., M. Barbieri, K. Todoerti, L. Agnelli, S. Fabris, G. Tonon, S. Segalla, I. Cifola, E. Pinatel, P. Tassone, P. Musto, L. Baldini and A. Neri (2015). "A compendium of DIS3 mutations and associated transcriptional signatures in plasma cell dyscrasias." Oncotarget **6**(28): 26129-26141.
Liu, E. Y., C. P. Cali and E. B. Lee (2017). "RNA metabolism in neurodegenerative disease." Disease Models & Mechanisms **10**(5): 509.
Lloyd, J. P. B. (2018). "The evolution and diversity of the nonsense-mediated mRNA decay pathway." F1000Res **7**: 1299.
Long, R. M. and M. T. McNally (2003). "mRNA decay: x (XRN1) marks the spot." Mol Cell **11**(5): 1126-1128.
Lubas, M., M. S. Christensen, M. S. Kristiansen, M. Domanski, L. G. Falkenby, S. Lykke-Andersen, J. S. Andersen, A. Dziembowski and T. H. Jensen (2011). "Interaction profiling identifies the human nuclear exosome targeting complex." Mol Cell **43**(4): 624-637.
Lubas, M., C. K. Damgaard, R. Tomecki, D. Cysewski, T. H. Jensen and A. Dziembowski (2013). "Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA." Embo j **32**(13): 1855-1868.
Luhrmann, R., B. Kastner and M. Bach (1990). "Structure of spliceosomal snRNPs and their role in pre-mRNA splicing." Biochim Biophys Acta **1087**(3): 265-292.
Luke, B. and J. Lingner (2009). "TERRA: telomeric repeat-containing RNA." Embo j **28**(17): 2503-2510.
Luke, B., A. Panza, S. Redon, N. Iglesias, Z. Li and J. Lingner (2008). "The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae." Mol Cell **32**(4): 465-477.
Luo, Y., Z. Na and S. A. Slavoff (2018). "P-Bodies: Composition, Properties, and Functions." Biochemistry **57**(17): 2424-2431.
Lykke-Andersen, S. and T. H. Jensen (2007). "Overlapping pathways dictate termination of RNA polymerase II transcription." Biochimie **89**(10): 1177-1182.
Mack, G. S. (2007). "MicroRNA gets down to business." Nature Biotechnology **25**: 631.
Mahdavi, M., M. Nassiri, M. M. Kooshyar, M. Vakili-Azghandi, A. Avan, R. Sandry, S. Pillai, A. K. Lam and V. Gopalan (2019). "Hereditary breast cancer; Genetic penetrance and current status with BRCA." J Cell Physiol **234**(5): 5741-5750.
Malecki, M., S. C. Viegas, T. Carneiro, P. Golik, C. Dressaire, M. G. Ferreira and C. M. Arraiano (2013). "The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway." Embo j **32**(13): 1842-1854.
Mao, J., G. Zhuang and Z. Chen (2017). "Genetic Polymorphisms of Insulin-Like Growth Factor 1 Are Associated with Osteosarcoma Risk and Prognosis." Med Sci Monit **23**: 5892-5898.

Martin, J. W., J. A. Squire and M. Zielenska (2012). "The genetics of osteosarcoma." Sarcoma **2012**: 627254.
Martino, E., G. Casamassima, S. Castiglione, E. Cellupica, S. Pantalone, F. Papagni, M. Rui, A. M. Siciliano and S. Collina (2018). "Vinca alkaloids and analogues as anti-cancer agents: Looking back, peering ahead." Bioorg Med Chem Lett **28**(17): 2816-2826.
Mayr, C. (2016). "Evolution and Biological Roles of Alternative 3'UTRs." Trends Cell Biol **26**(3): 227-237.
Mayya, V. K. and T. F. Duchaine (2019). "Ciphers and Executioners: How 3'-Untranslated Regions Determine the Fate of Messenger RNAs." Front Genet **10**: 6.
Medina, D. A., A. Jordán-Pla, G. Millán-Zambrano, S. Chávez, M. Choder and J. E. Pérez-Ortín (2014). "Cytoplasmic 5'-3' exonuclease Xrn1p is also a genome-wide transcription factor in yeast." Front Genet **5**: 1.
Memczak, S., M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D. Mackowiak, L. H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M. Landthaler, C. Kocks, F. le Noble and N. Rajewsky (2013). "Circular RNAs are a large class of animal RNAs with regulatory potency." Nature **495**(7441): 333-338.
Mendell, J. T., N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo and H. C. Dietz (2004). "Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise." Nat Genet **36**(10): 1073-1078.
Menezes, M. R., J. Balzeau and J. P. Hagan (2018). "3' RNA Uridylation in Epitranscriptomics, Gene Regulation, and Disease." Front Mol Biosci **5**: 61.
Meola, N., M. Domanski, E. Karadoulama, Y. Chen, C. Gentil, D. Pultz, K. Vitting-Seerup, S. Lykke-Andersen, Jens S. Andersen, A. Sandelin and Torben H. Jensen (2016). "Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts." Molecular Cell **64**(3): 520-533.
Mercer, T. R. and J. S. Mattick (2013). "Structure and function of long noncoding RNAs in epigenetic regulation." Nature Structural &Amp; Molecular Biology **20**: 300.
Meyer, S., C. Temme and E. Wahle (2004). "Messenger RNA turnover in eukaryotes: pathways and enzymes." Crit Rev Biochem Mol Biol **39**(4): 197-216.
Miao, L., H. Yao, C. Li, M. Pu, X. Yao, H. Yang, X. Qi, J. Ren and Y. Wang (2016). "A dual inhibition: microRNA-552 suppresses both transcription and translation of cytochrome P450 2E1." Biochim Biophys Acta **1859**(4): 650-662.
Milac, A. L., E. Bojarska and A. Wypijewska del Nogal (2014). "Decapping Scavenger (DcpS) enzyme: Advances in its structure, activity and roles in the cap-dependent mRNA metabolism." Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1839**(6): 452-462.
Mittelman, D. and J. H. Wilson (2013). "The fractured genome of HeLa cells." Genome Biol **14**(4): 111.
Moe, K., H. M. Thu, W. M. Oo, K. M. Aye, T. T. Shwe, W. Mar and C. D. Kirkwood (2009). "Genotyping of rotavirus isolates collected from children less than 5 years of age admitted for diarrhoea at the Yangon Children's Hospital, Myanmar." Vaccine **27 Suppl 5**: F89-92.
Monahan, Z., F. Shewmaker and U. B. Pandey (2016). "Stress granules at the intersection of autophagy and ALS." Brain Res **1649**(Pt B): 189-200.
Moncla, A., C. Missirian, P. Cacciagli, E. Balzamo, L. Legeai-Mallet, J. L. Jouve, B. Chabrol, M. Le Merrer, G. Plessis, L. Villard and N. Philip (2007). "A cluster of translocation breakpoints in 2q37 is associated with overexpression of NPPC in patients with a similar overgrowth phenotype." Hum Mutat **28**(12): 1183-1188.
Montecucco, A., F. Zanetta and G. Biamonti (2015). "Molecular mechanisms of etoposide." Excli j **14**: 95-108.
Moon, S. L., J. R. Anderson, Y. Kumagai, C. J. Wilusz, S. Akira, A. A. Khromykh and J. Wilusz (2012). "A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular exoribonuclease XRN1 and alters host mRNA stability." RNA **18**(11): 2029-2040.
Moon, S. L., J. G. Blackinton, J. R. Anderson, M. K. Dozier, B. J. Dodd, J. D. Keene, C. J. Wilusz, S. S. Bradrick and J. Wilusz (2015). "XRN1 stalling in the 5' UTR of Hepatitis C virus and Bovine Viral

Diarrhea virus is associated with dysregulated host mRNA stability." PLoS Pathog **11**(3): e1004708.
Moon, S. L., B. J. Dodd, D. E. Brackney, C. J. Wilusz, G. D. Ebel and J. Wilusz (2015). "Flavivirus sfRNA suppresses antiviral RNA interference in cultured cells and mosquitoes and directly interacts with the RNAi machinery." Virology **485**: 322-329.
Mori, F., K. Tanji, Y. Miki, Y. Toyoshima, H. Sasaki, M. Yoshida, A. Kakita, H. Takahashi and K. Wakabayashi (2018). "Immunohistochemical localization of exoribonucleases (DIS3L2 and XRN1) in intranuclear inclusion body disease." Neurosci Lett **662**: 389-394.
Morris, M. R., D. Astuti and E. R. Maher (2013). "Perlman syndrome: overgrowth, Wilms tumor predisposition and DIS3L2." Am J Med Genet C Semin Med Genet **163c**(2): 106-113.
Mozzetta, C., E. Boyarchuk, J. Pontis and S. Ait-Si-Ali (2015). "Sound of silence: the properties and functions of repressive Lys methyltransferases." Nat Rev Mol Cell Biol **16**(8): 499-513.
Mukherjee, D., M. Gao, J. P. O'Connor, R. Raijmakers, G. Pruijn, C. S. Lutz and J. Wilusz (2002). "The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements." Embo j **21**(1-2): 165-174.
Mukhopadhyay, S., R. J. Kuhn and M. G. Rossmann (2005). "A structural perspective of the flavivirus life cycle." Nat Rev Microbiol **3**(1): 13-22.
Mullen, T. E. and W. F. Marzluff (2008). "Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'." Genes Dev **22**(1): 50-65.
Murn, J. and Y. Shi (2017). "The winding path of protein methylation research: milestones and new frontiers." Nat Rev Mol Cell Biol **18**(8): 517-527.
Nagarajan, V. K., C. I. Jones, S. F. Newbury and P. J. Green (2013). "XRN 5'→3' exoribonucleases: structure, mechanisms and functions." Biochim Biophys Acta **1829**(6-7): 590-603.
Neugebauer, K. M. (2002). "On the importance of being co-transcriptional." J Cell Sci **115**(Pt 20): 3865-3871.
Newbury, S. and A. Woollard (2004). "The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis." Rna **10**(1): 59-65.
Nguyen, H. A., E. D. Hoffer and C. M. Dunham (2019). "Importance of tRNA anticodon loop modification and a conserved, noncanonical anticodon stem pairing in tRNA(Pro) CGG for decoding." J Biol Chem.
Nishimura, T., Z. Padamsi, H. Fakim, S. Milette, W. H. Dunham, A. C. Gingras and M. R. Fabian (2015). "The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs." Cell Rep **11**(9): 1425-1436.
Nowotny, M. (2009). "Retroviral integrase superfamily: the structural perspective." EMBO Rep **10**(2): 144-151.
O'Brien, J., H. Hayder, Y. Zayed and C. Peng (2018). "Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation." Frontiers in Endocrinology **9**(402).
Ochiai, H., S. Okada, A. Saito, K. Hoshi, H. Yamashita, T. Takato and T. Azuma (2012). "Inhibition of insulin-like growth factor-1 (IGF-1) expression by prolonged transforming growth factor-beta1 (TGF-beta1) administration suppresses osteoblast differentiation." J Biol Chem **287**(27): 22654-22661.
Olivera-Couto, A., V. Salzman, M. Mailhos, M. A. Digman, E. Gratton and P. S. Aguilar (2015). "Eisosomes are dynamic plasma membrane domains showing pil1-lsp1 heteroligomer binding equilibrium." Biophys J **108**(7): 1633-1644.
Ozaki, T. (2015). "Diagnosis and treatment of Ewing sarcoma of the bone: a review article." J Orthop Sci **20**(2): 250-263.
Pak, D., R. Root-Bernstein and Z. F. Burton (2017). "tRNA structure and evolution and standardization to the three nucleotide genetic code." Transcription **8**(4): 205-219.
Palmini, G., F. Marini and M. L. Brandi (2017). "What Is New in the miRNA World Regarding Osteosarcoma and Chondrosarcoma?" Molecules **22**(3).
Patel, A. G. and S. H. Kaufmann (2012). "How does doxorubicin work?" Elife **1**: e00387.

Pavlova, N. N. and C. B. Thompson (2016). "The Emerging Hallmarks of Cancer Metabolism." Cell Metab **23**(1): 27-47.
Payne, A. M., J. Zorman, M. Horton, S. Dubey, J. ter Meulen and K. A. Vora (2013). "Caspase activation as a versatile assay platform for detection of cytotoxic bacterial toxins." J Clin Microbiol **51**(9): 2970-2976.
Peccarelli, M. and B. W. Kebaara (2014). "Regulation of natural mRNAs by the nonsense-mediated mRNA decay pathway." Eukaryot Cell **13**(9): 1126-1135.
Pelechano, V., W. Wei and L. M. Steinmetz (2015). "Widespread Co-translational RNA Decay Reveals Ribosome Dynamics." Cell **161**(6): 1400-1412.
Perry, J. A., A. Kiezun, P. Tonzi, E. M. Van Allen, S. L. Carter, S. C. Baca, G. S. Cowley, A. S. Bhatt, E. Rheinbay, C. S. Pedamallu, E. Helman, A. Taylor-Weiner, A. McKenna, D. S. DeLuca, M. S. Lawrence, L. Ambrogio, C. Sougnez, A. Sivachenko, L. D. Walensky, N. Wagle, J. Mora, C. de Torres, C. Lavarino, S. Dos Santos Aguiar, J. A. Yunes, S. R. Brandalise, G. E. Mercado-Celis, J. Melendez-Zajgla, R. Cardenas-Cardos, L. Velasco-Hidalgo, C. W. Roberts, L. A. Garraway, C. Rodriguez-Galindo, S. B. Gabriel, E. S. Lander, T. R. Golub, S. H. Orkin, G. Getz and K. A. Janeway (2014). "Complementary genomic approaches highlight the PI3K/mTOR pathway as a common vulnerability in osteosarcoma." Proc Natl Acad Sci U S A **111**(51): E5564-5573.
Peter, M., J. Couturier, H. Pacquement, J. Michon, G. Thomas, H. Magdelenat and O. Delattre (1997). "A new member of the ETS family fused to EWS in Ewing tumors." Oncogene **14**(10): 1159-1164.
Pillai, R. S., S. N. Bhattacharyya, C. G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand and W. Filipowicz (2005). "Inhibition of translational initiation by Let-7 MicroRNA in human cells." Science **309**(5740): 1573-1576.
Pitchiaya, S., L. A. Heinicke, J. I. Park, E. L. Cameron and N. G. Walter (2017). "Resolving Subcellular miRNA Trafficking and Turnover at Single-Molecule Resolution." Cell Reports **19**(3): 630-642.
Pitchiaya, S., L. A. Heinicke, J. I. Park, E. L. Cameron and N. G. Walter (2017). "Resolving Subcellular miRNA Trafficking and Turnover at Single-Molecule Resolution." Cell Rep **19**(3): 630-642.
Popp, M. W. and L. E. Maquat (2013). "Organizing principles of mammalian nonsense-mediated mRNA decay." Annu Rev Genet **47**: 139-165.
Porrua, O. and D. Libri (2015). "Transcription termination and the control of the transcriptome: why, where and how to stop." Nat Rev Mol Cell Biol **16**(3): 190-202.
Protter, D. S. W. and R. Parker (2016). "Principles and Properties of Stress Granules." Trends in Cell Biology **26**(9): 668-679.
Proudfoot, N. J. (2011). "Ending the message: poly(A) signals then and now." Genes Dev **25**(17): 1770-1782.
Qu, S., Z. Liu, X. Yang, J. Zhou, H. Yu, R. Zhang and H. Li (2018). "The emerging functions and roles of circular RNAs in cancer." Cancer Lett **414**: 301-309.
Rajagopalan, P. T., Z. Zhang, L. McCourt, M. Dwyer, S. J. Benkovic and G. G. Hammes (2002). "Interaction of dihydrofolate reductase with methotrexate: ensemble and single-molecule kinetics." Proc Natl Acad Sci U S A **99**(21): 13481-13486.
Rajkowitsch, L., D. Chen, S. Stampfl, K. Semrad, C. Waldsich, O. Mayer, M. F. Jantsch, R. Konrat, U. Blasi and R. Schroeder (2007). "RNA chaperones, RNA annealers and RNA helicases." RNA Biol **4**(3): 118-130.
Ramskold, D., E. T. Wang, C. B. Burge and R. Sandberg (2009). "An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data." PLoS Comput Biol **5**(12): e1000598.
Ray, R. B., L. M. Lagging, K. Meyer, R. Steele and R. Ray (1995). "Transcriptional regulation of cellular and viral promoters by the hepatitis C virus core protein." Virus Res **37**(3): 209-220.
Rehwinkel, J., I. Behm-Ansmant, D. Gatfield and E. Izaurralde (2005). "A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing." Rna **11**(11): 1640-1647.

Reimann, E., S. Koks, X. D. Ho, K. Maasalu and A. Martson (2014). "Whole exome sequencing of a single osteosarcoma case--integrative analysis with whole transcriptome RNA-seq data." Hum Genomics **8**: 20.
Reimão-Pinto, M. M., R. A. Manzenreither, T. R. Burkard, P. Sledz, M. Jinek, K. Mechtler and S. L. Ameres (2016). "Molecular basis for cytoplasmic RNA surveillance by uridylation-triggered decay in Drosophila." Embo j **35**(22): 2417-2434.
Rickel, K., F. Fang and J. Tao (2017). "Molecular genetics of osteosarcoma." Bone **102**: 69-79.
Roberts, K. G. and C. G. Mullighan (2015). "Genomics in acute lymphoblastic leukaemia: insights and treatment implications." Nat Rev Clin Oncol **12**(6): 344-357.
Robinson, P. J., M. A. Pringle, C. A. Woolhead and N. J. Bulleid (2017). "Folding of a single domain protein entering the endoplasmic reticulum precedes disulfide formation." J Biol Chem **292**(17): 6978-6986.
Robinson, S., A. Oliver, T. Chevassut and S. Newbury (2015). "The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease." Biomolecules **5**(3): 1515.
Robinson, S. R., S. C. Viegas, R. G. Matos, S. Domingues, M. Bedir, H. J. S. Stewart, T. J. Chevassut, A. W. Oliver, C. M. Arraiano and S. F. Newbury (2018). "DIS3 isoforms vary in their endoribonuclease activity and are differentially expressed within haematological cancers." Biochem J **475**(12): 2091-2105.
Rogers, K. M. and R. M. Conran (2019). "Educational Case: Pediatric Osteosarcoma." Acad Pathol **6**: 2374289519833902.
Romero-Cordoba, S. L., I. Salido-Guadarrama, M. Rodriguez-Dorantes and A. Hidalgo-Miranda (2014). "miRNA biogenesis: biological impact in the development of cancer." Cancer Biol Ther **15**(11): 1444-1455.
Rose, A. E., L. Poliseno, J. Wang, M. Clark, A. Pearlman, G. Wang, Y. S. d. M. E. C. Vega, R. Medicherla, P. J. Christos, R. Shapiro, A. Pavlick, F. Darvishian, J. Zavadil, D. Polsky, E. Hernando, H. Ostrer and I. Osman (2011). "Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression." Cancer Res **71**(7): 2561-2571.
Rytelewski, M., S. Maleki Vareki, L. S. Mangala, L. Romanow, D. Jiang, S. Pradeep, C. Rodriguez-Aguayo, G. Lopez-Berestein, R. Figueredo, P. J. Ferguson, M. Vincent, A. K. Sood and J. D. Koropatnick (2016). "Reciprocal positive selection for weakness - preventing olaparib resistance by inhibiting BRCA2." Oncotarget **7**(15): 20825-20839.
Sampath, D., C. Liu, K. Vasan, M. Sulda, V. K. Puduvalli, W. G. Wierda and M. J. Keating (2012). "Histone deacetylases mediate the silencing of miR-15a, miR-16, and miR-29b in chronic lymphocytic leukemia." Blood **119**(5): 1162-1172.
Schaeffer, D., J. A. Somarelli, G. Hanna, G. M. Palmer and M. A. Garcia-Blanco (2014). "Cellular migration and invasion uncoupled: increased migration is not an inexorable consequence of epithelial-to-mesenchymal transition." Mol Cell Biol **34**(18): 3486-3499.
Schilke, K., F. Schaefer, R. Waldherr, W. Rohrschneider, C. John, U. Himbert, E. Mayatepek and G. Tariverdian (2000). "A case of Perlman syndrome: fetal gigantism, renal dysplasia, and severe neurological deficits." Am J Med Genet **91**(1): 29-33.
Schmidt, E. K., G. Clavarino, M. Ceppi and P. Pierre (2009). "SUnSET, a nonradioactive method to monitor protein synthesis." Nat Methods **6**(4): 275-277.
Schmidt, M. J., S. West and C. J. Norbury (2011). "The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation." Rna **17**(1): 39-44.
Schneider, C., E. Leung, J. Brown and D. Tollervey (2009). "The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome." Nucleic Acids Res **37**(4): 1127-1140.
Schoenberg, D. R. and L. E. Maquat (2012). "Regulation of cytoplasmic mRNA decay." Nat Rev Genet **13**(4): 246-259.
Segalla, S., S. Pivetti, K. Todoerti, M. A. Chudzik, E. C. Giuliani, F. Lazzaro, V. Volta, D. Lazarevic, G. Musco, M. Muzi-Falconi, A. Neri, S. Biffo and G. Tonon (2015). "The ribonuclease DIS3

promotes let-7 miRNA maturation by degrading the pluripotency factor LIN28B mRNA." Nucleic Acids Res **43**(10): 5182-5193.
Senger, D. R., C. A. Perruzzi, M. Streit, V. E. Koteliansky, A. R. de Fougerolles and M. Detmar (2002). "The alpha(1)beta(1) and alpha(2)beta(1) integrins provide critical support for vascular endothelial growth factor signaling, endothelial cell migration, and tumor angiogenesis." Am J Pathol **160**(1): 195-204.
She, M., C. J. Decker, D. I. Svergun, A. Round, N. Chen, D. Muhlrad, R. Parker and H. Song (2008). "Structural basis of dcp2 recognition and activation by dcp1." Mol Cell **29**(3): 337-349.
Sheth, U. and R. Parker (2003). "Decapping and decay of messenger RNA occur in cytoplasmic processing bodies." Science **300**(5620): 805-808.
Shimakami, T., D. Yamane, R. K. Jangra, B. J. Kempf, C. Spaniel, D. J. Barton and S. M. Lemon (2012). "Stabilization of hepatitis C virus RNA by an Ago2-miR-122 complex." Proc Natl Acad Sci U S A **109**(3): 941-946.
Shukla, S., G. A. Bjerke, D. Muhlrad, R. Yi and R. Parker (2019). "The RNase PARN Controls the Levels of Specific miRNAs that Contribute to p53 Regulation." Mol Cell **73**(6): 1204-1216.e1204.
Shukla, S., G. A. Bjerke, D. Muhlrad, R. Yi and R. Parker (2019). "The RNase PARN Controls the Levels of Specific miRNAs that Contribute to p53 Regulation." Mol Cell.
Shwetha, S., A. Kumar, R. Mullick, D. Vasudevan, N. Mukherjee and S. Das (2015). "HuR Displaces Polypyrimidine Tract Binding Protein To Facilitate La Binding to the 3' Untranslated Region and Enhances Hepatitis C Virus Replication." J Virol **89**(22): 11356-11371.
Sierecki, E. (2018). "The Mediator complex and the role of protein-protein interactions in the gene regulation machinery." Semin Cell Dev Biol.
Silva, P. A., C. F. Pereira, T. J. Dalebout, W. J. Spaan and P. J. Bredenbeek (2010). "An RNA pseudoknot is required for production of yellow fever virus subgenomic RNA by the host nuclease XRN1." J Virol **84**(21): 11395-11406.
Singh, A., P. Trivedi and N. K. Jain (2018). "Advances in siRNA delivery in cancer therapy." Artif Cells Nanomed Biotechnol **46**(2): 274-283.
Siomi, M. C., K. Sato, D. Pezic and A. A. Aravin (2011). "PIWI-interacting small RNAs: the vanguard of genome defence." Nat Rev Mol Cell Biol **12**(4): 246-258.
Skourti-Stathaki, K., N. J. Proudfoot and N. Gromak (2011). "Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination." Mol Cell **42**(6): 794-805.
Slomovic, S. and G. Schuster (2011). "Exonucleases and endonucleases involved in polyadenylation-assisted RNA decay." Wiley Interdiscip Rev RNA **2**(1): 106-123.
Snee, M. J., W. C. Wilson, Y. Zhu, S. Y. Chen, B. A. Wilson, C. Kseib, J. O'Neal, N. Mahajan, M. H. Tomasson, S. Arur and J. B. Skeath (2016). "Collaborative Control of Cell Cycle Progression by the RNA Exonuclease Dis3 and Ras Is Conserved Across Species." Genetics **203**(2): 749-762.
Sobell, H. M. (1985). "Actinomycin and DNA transcription." Proc Natl Acad Sci U S A **82**(16): 5328-5331.
Song, M. G., S. Bail and M. Kiledjian (2013). "Multiple Nudix family proteins possess mRNA decapping activity." Rna **19**(3): 390-399.
Song, M. G. and M. Kiledjian (2007). "3' Terminal oligo U-tract-mediated stimulation of decapping." Rna **13**(12): 2356-2365.
Sorensen, P. H., S. L. Lessnick, D. Lopez-Terrada, X. F. Liu, T. J. Triche and C. T. Denny (1994). "A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG." Nat Genet **6**(2): 146-151.
Staals, R. H., A. W. Bronkhorst, G. Schilders, S. Slomovic, G. Schuster, A. J. Heck, R. Raijmakers and G. J. Pruijn (2010). "Dis3-like 1: a novel exoribonuclease associated with the human exosome." Embo j **29**(14): 2358-2367.
Stewart, M. (2019). "Polyadenylation and nuclear export of mRNAs." J Biol Chem **294**(9): 2977-2987.
Stoecklin, G. and N. Kedersha (2013). "Relationship of GW/P-bodies with stress granules." Adv Exp Med Biol **768**: 197-211.

Strong, M. J. (2010). "The evidence for altered RNA metabolism in amyotrophic lateral sclerosis (ALS)." J Neurol Sci **288**(1-2): 1-12.
Sun, L., C. Yang, J. Xu, Y. Feng, L. Wang and T. Cui (2016). "Long Noncoding RNA EWSAT1 Promotes Osteosarcoma Cell Growth and Metastasis Through Suppression of MEG3 Expression." DNA Cell Biol **35**(12): 812-818.
Symmons, M. F. and B. F. Luisi (2009). "Through Ancient Rings Thread Programming Strings." Structure **17**(11): 1429-1431.
Szczepińska, T., K. Kalisiak, R. Tomecki, A. Labno, L. S. Borowski, T. M. Kulinski, D. Adamska, J. Kosinska and A. Dziembowski (2015). "DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts." Genome Res **25**(11): 1622-1633.
Tassano, E., J. Buttgereit, M. Bader, M. Lerone, M. T. Divizia, R. Bocciardi, F. Napoli, G. Pala, F. Sloan-Bena, S. Gimelli and G. Gimelli (2013). "Genotype-Phenotype Correlation of 2q37 Deletions Including Gene Associated with Skeletal Malformations." PLoS One **8**(6): e66048.
Tazi, J., N. Bakkour and S. Stamm (2009). "Alternative splicing and disease." Biochim Biophys Acta **1792**(1): 14-26.
Tesina, P., E. Heckel, J. Cheng, M. Fromont-Racine, R. Buschauer, L. Kater, B. Beatrix, O. Berninghausen, A. Jacquier, T. Becker and R. Beckmann (2019). "Structure of the 80S ribosome-Xrn1 nuclease complex." Nat Struct Mol Biol.
Tharun, S., D. Muhlrad, A. Chowdhury and R. Parker (2005). "Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3' end protection." Genetics **170**(1): 33-46.
Thibault, P. A., A. Huys, Y. Amador-Canizares, J. E. Gailius, D. E. Pinel and J. A. Wilson (2015). "Regulation of Hepatitis C Virus Genome Replication by Xrn1 and MicroRNA-122 Binding to Individual Sites in the 5' Untranslated Region." J Virol **89**(12): 6294-6311.
Tian, Z. Z., X. J. Guo, Y. M. Zhao and Y. Fang (2015). "Decreased expression of long non-coding RNA MEG3 acts as a potential predictor biomarker in progression and poor prognosis of osteosarcoma." Int J Clin Exp Pathol **8**(11): 15138-15142.
Tollervey, J. R., T. Curk, B. Rogelj, M. Briese, M. Cereda, M. Kayikci, J. Konig, T. Hortobagyi, A. L. Nishimura, V. Zupunski, R. Patani, S. Chandran, G. Rot, B. Zupan, C. E. Shaw and J. Ule (2011). "Characterizing the RNA targets and position-dependent splicing regulation by TDP-43." Nat Neurosci **14**(4): 452-458.
Tomecki, R., K. Drazkowska, I. Kucinski, K. Stodus, R. J. Szczesny, J. Gruchota, E. P. Owczarek, K. Kalisiak and A. Dziembowski (2014). "Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target." Nucleic Acids Res **42**(2): 1270-1290.
Tomecki, R. and A. Dziembowski (2010). "Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism." Rna **16**(9): 1692-1724.
Tomecki, R., M. S. Kristiansen, S. Lykke-Andersen, A. Chlebowski, K. M. Larsen, R. J. Szczesny, K. Drazkowska, A. Pastula, J. S. Andersen, P. P. Stepien, A. Dziembowski and T. H. Jensen (2010). "The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L." Embo j **29**(14): 2342-2357.
Towler, B. P., C. I. Jones, K. L. Harper, J. A. Waldron and S. F. Newbury (2016). "A novel role for the 3'-5' exoribonuclease Dis3L2 in controlling cell proliferation and tissue growth." RNA Biol **13**(12): 1286-1299.
Towler, B. P., C. I. Jones, S. C. Viegas, P. Apura, J. A. Waldron, S. K. Smalley, C. M. Arraiano and S. F. Newbury (2015). "The 3'-5' exoribonuclease Dis3 regulates the expression of specific microRNAs in Drosophila wing imaginal discs." RNA Biol **12**(7): 728-741.
Towler, B. P. and S. F. Newbury (2018). "Regulation of cytoplasmic RNA stability: Lessons from Drosophila." Wiley Interdiscip Rev RNA **9**(6): e1499.
Treiber, T., N. Treiber and G. Meister (2019). "Regulation of microRNA biogenesis and its crosstalk with other cellular pathways." Nat Rev Mol Cell Biol **20**(1): 5-20.

Treiber, T., N. Treiber, U. Plessmann, S. Harlander, J. L. Daiss, N. Eichner, G. Lehmann, K. Schall, H. Urlaub and G. Meister (2017). "A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis." Mol Cell **66**(2): 270-284.e213.
Truitt, M. L. and D. Ruggero (2016). "New frontiers in translational control of the cancer genome." Nat Rev Cancer **16**(5): 288-304.
Tu, J., Z. Huo, J. Gingold, R. Zhao, J. Shen and D. F. Lee (2017). "The Histogenesis of Ewing Sarcoma." Cancer Rep Rev **1**(2).
Turc-Carel, C., A. Aurias, F. Mugneret, S. Lizard, I. Sidaner, C. Volk, J. P. Thiery, S. Olschwang, I. Philip, M. P. Berger and et al. (1988). "Chromosomes in Ewing's sarcoma. I. An evaluation of 85 cases of remarkable consistency of t(11;22)(q24;q12)." Cancer Genet Cytogenet **32**(2): 229-238.
Unterholzner, L. and E. Izaurralde (2004). "SMG7 acts as a molecular link between mRNA surveillance and mRNA decay." Mol Cell **16**(4): 587-596.
Ustianenko, D., D. Hrossova, D. Potesil, K. Chalupnikova, K. Hrazdilova, J. Pachernik, K. Cetkovska, S. Uldrijan, Z. Zdrahal and S. Vanacova (2013). "Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs." Rna **19**(12): 1632-1638.
van Hoof, A., R. R. Staples, R. E. Baker and R. Parker (2000). "Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA." Mol Cell Biol **20**(21): 8230-8243.
Vanas, V., B. Haigl, V. Stockhammer and H. Sutterluty-Fall (2016). "MicroRNA-21 Increases Proliferation and Cisplatin Sensitivity of Osteosarcoma-Derived Cells." PLoS One **11**(8): e0161023.
Vaskovicova, K., T. Awadova, P. Vesela, M. Balazova, M. Opekarova and J. Malinsky (2017). "mRNA decay is regulated via sequestration of the conserved 5'-3' exoribonuclease Xrn1 at eisosome in yeast." Eur J Cell Biol **96**(6): 591-599.
Vishnoi, A. and S. Rani (2017). "MiRNA Biogenesis and Regulation of Diseases: An Overview." Methods Mol Biol **1509**: 1-10.
Wahid, F., A. Shehzad, T. Khan and Y. Y. Kim (2010). "MicroRNAs: synthesis, mechanism, function, and recent clinical trials." Biochim Biophys Acta **1803**(11): 1231-1243.
Waldron, J. A. (2014). The Role of the Exoribonuclease Pacman/Xrn1 in Wing Development in *Drosophila*. Doctor of Philosophy, Brighton & Sussex Medical School.
Waldron, J. A., C. I. Jones, B. P. Towler, A. L. Pashler, D. P. Grima, S. Hebbes, S. H. Crossman, M. V. Zabolotskaya and S. F. Newbury (2015). "Xrn1/Pacman affects apoptosis and regulates expression of hid and reaper." Biol Open **4**(5): 649-660.
Wang, C., F. Schmich, S. Srivatsa, J. Weidner, N. Beerenwinkel and A. Spang (2018). "Context-dependent deposition and regulation of mRNAs in P-bodies." Elife **7**.
Wang, M. and P. J. Casey (2016). "Protein prenylation: unique fats make their mark on biology." Nat Rev Mol Cell Biol **17**(2): 110-122.
Wang, S., Z. Han, D. Libri, O. Porrua and T. R. Strick (2019). "Single-molecule characterization of extrinsic transcription termination by Sen1 helicase." Nat Commun **10**(1): 1545.
Wang, Z., Y. Wang, T. Liu, Y. Wang and W. Zhang (2019). "Effects of PIWI/MID domain of Argonaute protein on the association of miRNAi's seed base with the target." Rna.
Watson (2013). Molecular biology of the gene. San Francisco, Pearson.
Wegert, J., N. Ishaque, R. Vardapour, C. Georg, Z. Gu, M. Bieg, B. Ziegler, S. Bausenwein, N. Nourkami, N. Ludwig, A. Keller, C. Grimm, S. Kneitz, R. D. Williams, T. Chagtai, K. Pritchard-Jones, P. van Sluis, R. Volckmann, J. Koster, R. Versteeg, T. Acha, M. J. O'Sullivan, P. K. Bode, F. Niggli, G. A. Tytgat, H. van Tinteren, M. M. van den Heuvel-Eibrink, E. Meese, C. Vokuhl, I. Leuschner, N. Graf, R. Eils, S. M. Pfister, M. Kool and M. Gessler (2015). "Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors." Cancer Cell **27**(2): 298-311.
Wei, R., G. Cao, Z. Deng, J. Su and L. Cai (2016). "miR-140-5p attenuates chemotherapeutic drug-induced cell death by regulating autophagy through inositol 1,4,5-trisphosphate kinase 2 (IP3k2) in human osteosarcoma cells." Biosci Rep **36**(5).
Will, C. L. and R. Lührmann (2011). "Spliceosome structure and function." Cold Spring Harb Perspect Biol **3**(7).

Wirth, S. M. (2015). "Afatinib in Non-Small Cell Lung Cancer." J Adv Pract Oncol **6**(5): 448-455.
Wojtowicz, J. M. and N. Kee (2006). "BrdU assay for neurogenesis in rodents." Nat Protoc **1**(3): 1399-1405.
Wolf, J. and L. A. Passmore (2014). "mRNA deadenylation by Pan2-Pan3." Biochem Soc Trans **42**(1): 184-187.
Wu, C., X. Zhu, K. Tao, W. Liu, T. Ruan, W. Wan, C. Zhang and W. Zhang (2018). "MALAT1 promotes the colorectal cancer malignancy by increasing DCP1A expression and miR203 downregulation." Mol Carcinog **57**(10): 1421-1431.
Wu, D., D. Muhlrad, M. W. Bowler, S. Jiang, Z. Liu, R. Parker and H. Song (2014). "Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1 for decapping activation." Cell Res **24**(2): 233-246.
Xiao, S., F. Houser-Scott and D. R. Engelke (2001). "Eukaryotic ribonuclease P: increased complexity to cope with the nuclear pre-tRNA pathway." J Cell Physiol **187**(1): 11-20.
Xie, L., Z. Yao, Y. Zhang, D. Li, F. Hu, Y. Liao, L. Zhou, Y. Zhou, Z. Huang, Z. He, L. Han, Y. Yang and Z. Yang (2018). "Deep RNA sequencing reveals the dynamic regulation of miRNA, lncRNAs, and mRNAs in osteosarcoma tumorigenesis and pulmonary metastasis." Cell Death Dis **9**(7): 772.
Xu, M., H. Jin, C. X. Xu, B. Sun, Z. Mao, W. Z. Bi and Y. Wang (2014). "miR-382 inhibits tumor growth and enhance chemosensitivity in osteosarcoma." Oncotarget **5**(19): 9472-9483.
Xu, M., H. Jin, C. X. Xu, B. Sun, Z. G. Song, W. Z. Bi and Y. Wang (2015). "miR-382 inhibits osteosarcoma metastasis and relapse by targeting Y box-binding protein 1." Mol Ther **23**(1): 89-98.
Yamashita, A., N. Izumi, I. Kashima, T. Ohnishi, B. Saari, Y. Katsuhata, R. Muramatsu, T. Morita, A. Iwamatsu, T. Hachiya, R. Kurata, H. Hirano, P. Anderson and S. Ohno (2009). "SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay." Genes Dev **23**(9): 1091-1105.
Yang, X. and K. Qian (2017). "Protein O-GlcNAcylation: emerging mechanisms and functions." Nat Rev Mol Cell Biol **18**(7): 452-465.
Yang, Z. J., C. E. Chee, S. Huang and F. A. Sinicrope (2011). "The role of autophagy in cancer: therapeutic implications." Mol Cancer Ther **10**(9): 1533-1541.
Yin, L. M., Y. Wei, Y. Wang, Y. D. Xu and Y. Q. Yang (2013). "Long term and standard incubations of WST-1 reagent reflect the same inhibitory trend of cell viability in rat airway smooth muscle cells." Int J Med Sci **10**(1): 68-72.
Yonekawa, T. and A. Thorburn (2013). "Autophagy and cell death." Essays Biochem **55**: 105-117.
Zakrzewska-Placzek, M., F. F. Souret, G. J. Sobczyk, P. J. Green and J. Kufel (2010). "Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA." Nucleic Acids Res **38**(13): 4487-4502.
Zangari, J., M. Ilie, F. Rouaud, L. Signetti, M. Ohanna, R. Didier, B. Roméo, D. Goldoni, N. Nottet, C. Staedel, J. Gal, B. Mari, B. Mograbi, P. Hofman and P. Brest (2017). "Rapid decay of engulfed extracellular miRNA by XRN1 exonuclease promotes transient epithelial-mesenchymal transition." Nucleic Acids Res **45**(7): 4131-4141.
Zekri, L., E. Huntzinger, S. Heimstadt and E. Izaurralde (2009). "The silencing domain of GW182 interacts with PABPC1 to promote translational repression and degradation of microRNA targets and is required for target release." Mol Cell Biol **29**(23): 6220-6231.
Zekri, L., D. Kuzuoglu-Ozturk and E. Izaurralde (2013). "GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation." Embo j **32**(7): 1052-1065.
Zhang, G., M. Hubalewska and Z. Ignatova (2009). "Transient ribosomal attenuation coordinates protein synthesis and co-translational folding." Nat Struct Mol Biol **16**(3): 274-280.
Zhang, K., N. Dion, B. Fuchs, T. Damron, S. Gitelis, R. Irwin, M. O'Connor, H. Schwartz, S. P. Scully, M. G. Rock, M. E. Bolander and G. Sarkar (2002). "The human homolog of yeast SEP1 is a novel candidate tumor suppressor gene in osteogenic sarcoma." Gene **298**(2): 121-127.
Zhang, S. Z., L. Cai and B. Li (2017). "MEG3 long non-coding RNA prevents cell growth and metastasis of osteosarcoma." Bratisl Lek Listy **118**(10): 632-636.

Zhang, W., X. Shen, C. Wan, Q. Zhao, L. Zhang, Q. Zhou and L. Deng (2012). "Effects of insulin and insulin-like growth factor 1 on osteoblast proliferation and differentiation: differential signalling via Akt and ERK." Cell Biochem Funct **30**(4): 297-302.
Zhang, X. O., H. B. Wang, Y. Zhang, X. Lu, L. L. Chen and L. Yang (2014). "Complementary sequence-mediated exon circularization." Cell **159**(1): 134-147.
Zhang, Y., M. Hu, L. Liu, X. L. Cheng, J. Cai, J. Zhou and T. Wang (2018). "Anticancer effects of Rosmarinic acid in OVCAR-3 ovarian cancer cells are mediated via induction of apoptosis, suppression of cell migration and modulation of lncRNA MALAT-1 expression." J buon **23**(3): 763-768.
Zhang, Z., Y. W. Qin, G. Brewer and Q. Jing (2012). "MicroRNA degradation and turnover: regulating the regulators." Wiley Interdiscip Rev RNA **3**(4): 593-600.
Zheng, D., A. Frankish, R. Baertsch, P. Kapranov, A. Reymond, S. W. Choo, Y. Lu, F. Denoeud, S. E. Antonarakis, M. Snyder, Y. Ruan, C. L. Wei, T. R. Gingeras, R. Guigo, J. Harrow and M. B. Gerstein (2007). "Pseudogenes in the ENCODE regions: consensus annotation, analysis of transcription, and evolution." Genome Res **17**(6): 839-851.
Zhu, Z., J. Tang, J. Wang, G. Duan, L. Zhou and X. Zhou (2016). "MiR-138 Acts as a Tumor Suppressor by Targeting EZH2 and Enhances Cisplatin-Induced Apoptosis in Osteosarcoma Cells." PLoS One **11**(3): e0150026.
Zhuo, M., C. Yuan, T. Han, J. Cui, F. Jiao and L. Wang (2018). "A novel feedback loop between high MALAT-1 and low miR-200c-3p promotes cell migration and invasion in pancreatic ductal adenocarcinoma and is predictive of poor prognosis." BMC Cancer **18**(1): 1032.
Zinder, J. C. and C. D. Lima (2017). "Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors." Genes Dev **31**(2): 88-100.

Appendix

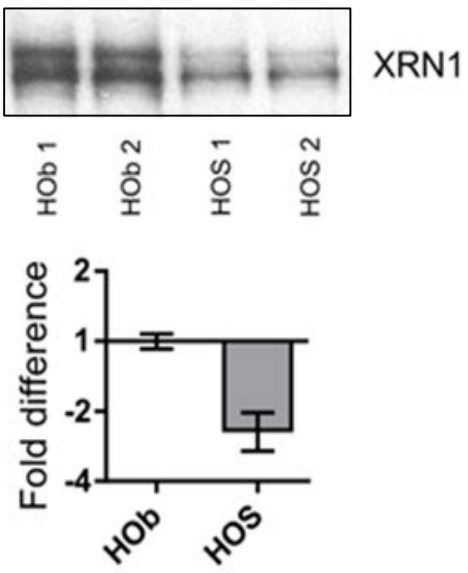


Figure A.1. Previous XRN1 protein expression analysis by Dr Chris Jones, Post-Doctoral researcher in the Newbury lab. The above Western gel image and graph show that XRN1 is down regulated in the HOS cell line during previous observational experiments which is not recapitulated by the author. XRN1 expression normalized to TUBULIN, statistical analysis unknown.

XRN1 overexpression studies

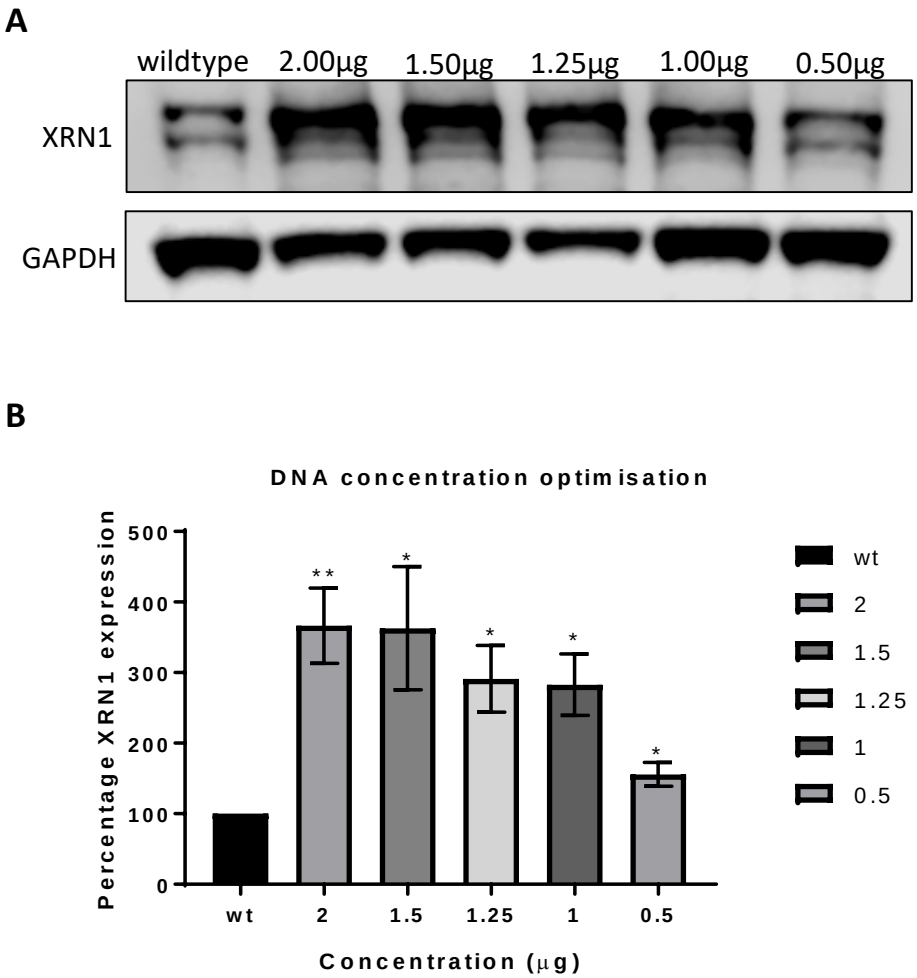


Figure A.2. Transfection of XRN1 DNA concentration optimisation. A) Gel visualization of XRN1 protein expression in SAOS-2 cells after being transfected with different concentrations of plasmid XRN1-containing DNA. B) Quantification of XRN1 expression compared to a wildtype control in SAOS-2 cells after 24hrs. Statistical analysis performed by unpaired t-test, where wildtype vs. 2µg DNA p= 0.0075, wildtype vs. 1.5µg DNA p=0.0396, wildtype vs. 1.25µg p=0.0156, wildtype vs. 1µg DNA p=0.0138 and wildtype vs. 0.5µg p=0.0299. Error bars represent SEM, based on n=3.

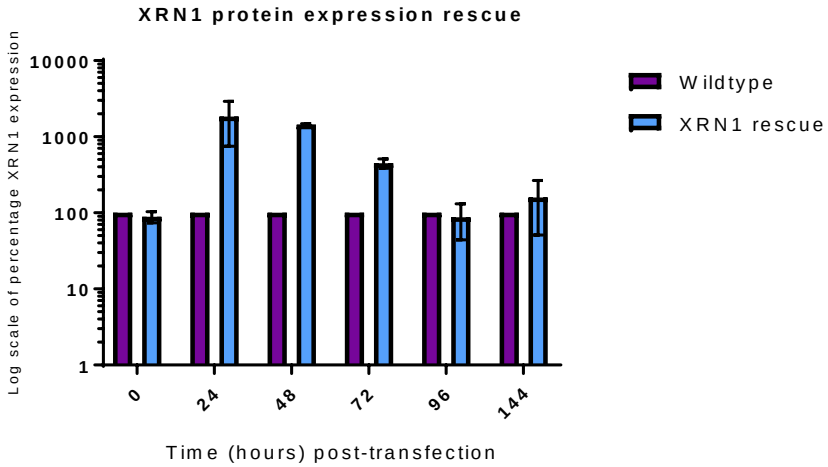


Figure A.3. Time course of XRN1 over expression in SAOS-2 cells when compared with wildtype cells. A time course of XRN1 expression was conducted over 144hrs, where cells were harvested at every 24hr time point (excluding 120hrs). This showed that XRN1 protein expression was initially up regulated at 24hrs, and expression remained up regulated until after 72hrs.

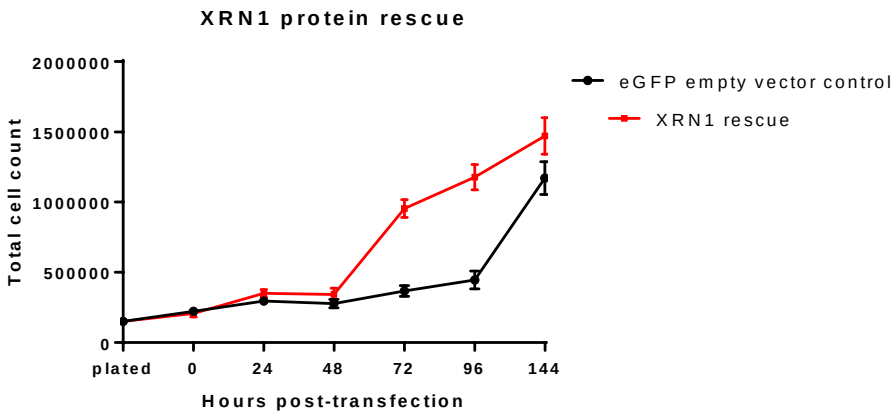
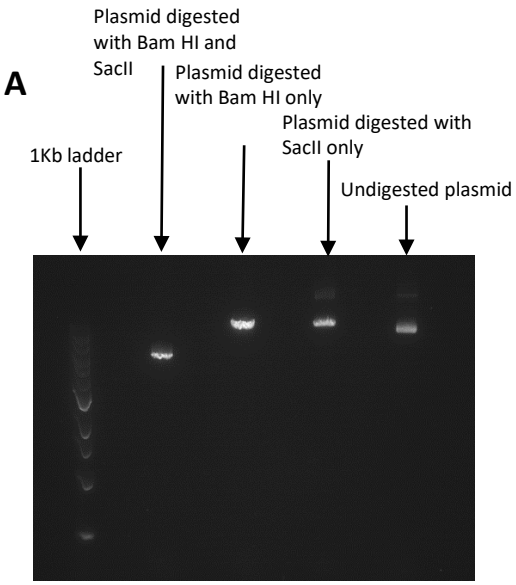
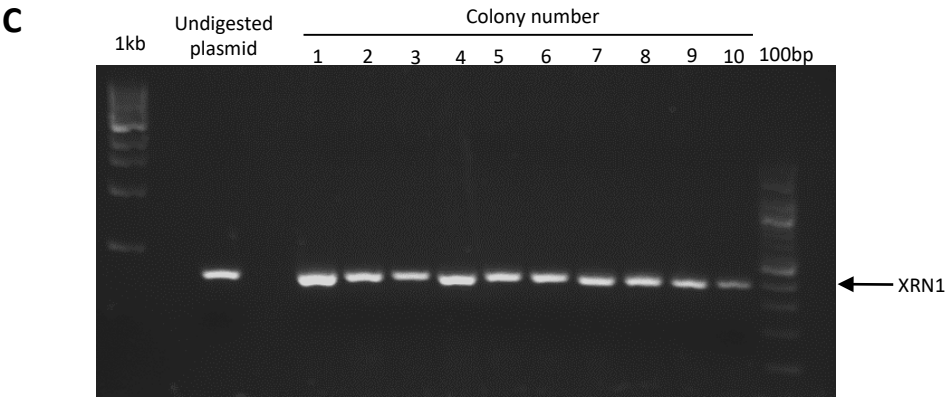


Figure A.4. Growth curve of cells when compared to eGFP empty plasmid control. This empty control plasmid was kindly donated by Dr. L. Mullen (Brighton & Sussex Medical School) to control for the XRN1-containing plasmid. A growth curve was performed over 144hrs and cell counts were conducted every 24hr timepoint (excluding 120hrs). Results at first showed that XRN1 might be rescuing cell line growth when transfected into SAOS-2 cells. However, this result was probably not because of XRN1 rescue, but rather that the empty plasmid (which contained a strong CMV promoter) was cytotoxic to the SAOS-2 cells, which made it seem as though XRN1 rescue was impacting positively on cell growth. In order to confirm this, it was necessary to try to create a control plasmid from the XRN1-containing plasmid (see further Appendix figures).

Creation of control plasmid without XRN1 sequence:



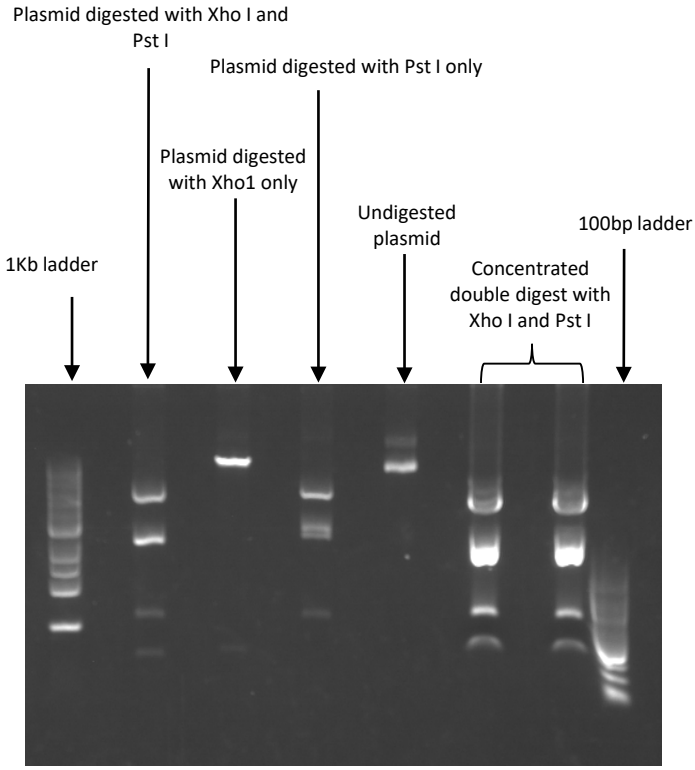
Gel to check digestion with appropriate restriction enzymes



After bacterial transformation, colonies were picked to test the plasmid. Colony formation PCR to check XRN1 has been removed from the plasmid. Removal unsuccessful.

Figure A.5. Cloning Attempt 1: Digestion with restriction enzymes Bam H1 and SacII.

1. Plasmid digestion



2. Ligation reaction - unsuccessful

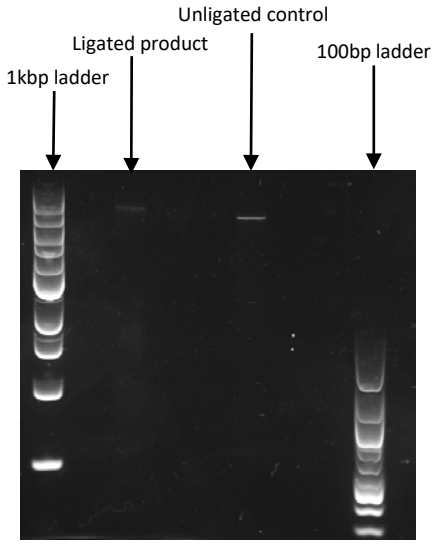


Figure A.6. Cloning Attempt 2: Conducted with 2 different restriction enzymes (Xho1 and Pst 1). Not1 oligos were annealed and ligated into the linearized empty vector to recircularise the plasmid in the ligation reaction. After bacterial transformation there was no growth during antibiotic selection, indicating that the ligation was not successful.

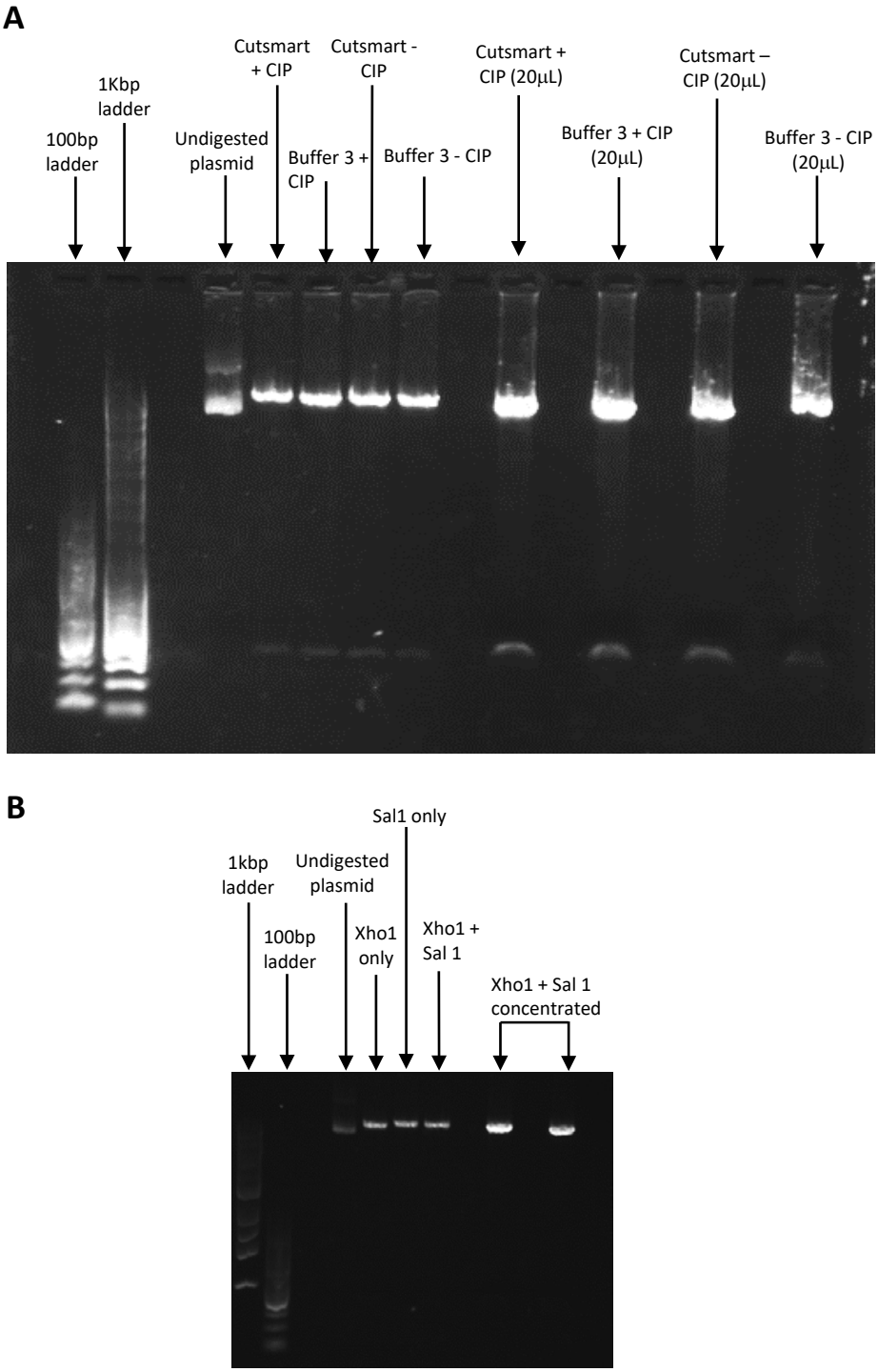
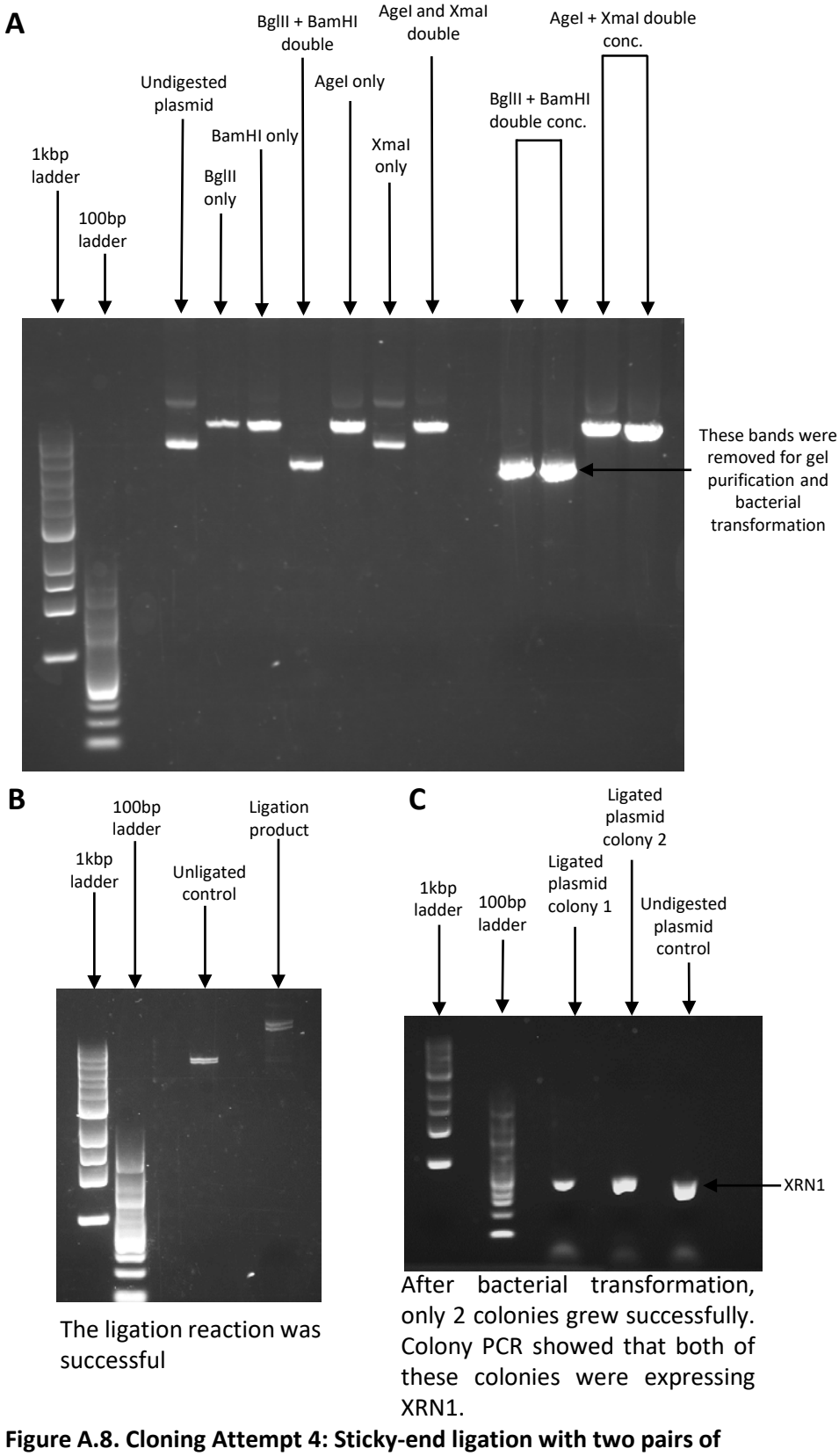


Figure A.7. Cloning Attempt 3: Digestion with Xho1 and Sal1 A) After digestion with Xho1 and SalI it was clear that only the Xho1 enzymes had successfully cut at the correct site. B) The SalI enzyme was retested using higher concentrations which again showed that it was not digesting the plasmid.



The ligation reaction was successful

After bacterial transformation, only 2 colonies grew successfully. Colony PCR showed that both of these colonies were expressing XRN1.

Figure A.8. Cloning Attempt 4: Sticky-end ligation with two pairs of compatible enzymes (BglII + BamHI) and (AgeI and Xma1).