

Translational control: mTOR signalling  
and the use of next-generation  
sequencing methods

Thesis submitted to the University of Nottingham  
for the degree of Doctor of Philosophy

February 2013

Thomas J. Jackson

For Mum,

## Abstract

Translation is a multi-stage process comprising initiation, elongation and termination. It has been suggested that the initiation phase is the rate limiting step of this process. In this thesis the contributions of how changes in initiation and elongation rates lead to overall alterations in gene expression pathways were investigated in a number of different systems.

It has been shown previously that increased expression of  $\text{tRNA}_i^{\text{Met}}$  is associated with tumourigenesis however the precise role of  $\text{tRNA}_i^{\text{Met}}$  in this process was unclear. Data obtained from cells with increased  $\text{tRNA}_i^{\text{Met}}$  copy number show that the associated increase in proliferation is transient and unlikely to play a major role in cancer.

In collaboration with Dr Owen Samson's group, it was shown that early dysplastic changes in intestinal tumourigenesis are driven by increased translation elongation via mammalian target of rapamycin 1 (mTORC1). It was found that constitutively active mutant K-Ras confers resistance to mTORC1 inhibition, and combined mTORC1/2 inhibition but that these tumours are acutely sensitivity to loss of the mTORC2 component, Rictor.

Interestingly, changes in translation elongation rate were also identified in cells cooled to  $32^\circ\text{C}$  and this was associated with the reprogramming of gene expression under these conditions.

Finally, the use of the next-generation sequencing technique ribosome profiling illustrated some potential challenges of using this approach to infer biologically relevant conclusions. These include: biases in fragment and library generation, limited read depth and statistical inference from low-count data. An alternate library generation method reduced bias, but reads predicted to be highly structured were still over-represented. The use of a newly developed thermostable ligase did not remedy this problem, but this may be due to additional biases associated with the particular ligase used.

## **Acknowledgements**

When I first told my parents I wanted to be a clinical academic they were keen to impress upon me how frustrating research can be. The past four years have, at times, proved them right, but it has always been bearable and for the most part rewarding and fun. That wouldn't have been the case without many people to whom I would like to give my thanks.

To Anne Willis for choosing to take me on as a student and for the support and guidance she has given me since. Despite her busy schedule she always made time for me.

To Owen Samson, Alessandro Scopelitti, and especially Liam Faller. I feel genuinely privileged to have been involved in such a fantastic collaboration.

To members of the Willis and Bushell groups. They have all supported and humoured me, but particular thanks must go to Amandine Bastide, John Knight, Xavier Pichon, Sam Johnston, Thomas Sbaratto, Mike Screen, Tuija Poyry, Ruth Spriggs, Carolyn Jones, Nick Burgoyne and Lindsay Wilson.

To Terence Stephenson for allowing me to defer my progression on the medical course and to Jim Lowe for allowing me to extend this for an additional year to get the most out of my research.

Finally, to friends and family who have been kind enough to at least feign interest and understanding as I talked excitedly at them about my research. They will be granted what I hope will be a temporary reprieve whilst I return to being a medical student.

## Contents

Abstract	1
Acknowledgements .....	2
Contents	3
List of Figures	10
Abbreviations	13
Chapter 1. Introduction.....	18
1.1 Mechanism and regulation of translation initiation.....	19
1.1.a Recruitment of tRNA <sub>i</sub> <sup>Met</sup> to the 40S ribosomal subunit .....	21
1.1.b Start site recognition .....	24
1.1.c Recruitment of the pre-initiation complex to mRNA .....	24
1.1.d Large ribosomal subunit joining .....	27
1.2 Mechanism and regulation of translation elongation.....	27
1.2.a Recruitment of aminoacyl-tRNAs .....	27
1.2.b Peptide bond formation .....	28
1.2.c Ribosomal translocation .....	28
1.3 Mechanism of translation termination and ribosomal recycling .....	31
1.4 mTOR signalling .....	32
1.4.a mTORC1 signalling .....	33
1.4.b mTORC2 signalling .....	38
1.5 Translational control and cancer .....	39

1.5.a	The eIF4F complex and cancer .....	39
1.5.b	Ternary complex and cancer .....	41
1.5.c	Translation elongation and cancer .....	42
1.6	Project aims .....	42
Chapter 2.	Materials and methods.....	44
2.1	Solutions common to multiple methods.....	44
2.2	Cell culture.....	44
2.2.a	Monolayer cultures .....	44
	i) Routine cell culture.....	44
	ii) DNA/RNA transfection.....	45
2.2.b	Matrigel cultures .....	45
	i) WT crypt culture .....	46
	ii) Culture of adenomas .....	46
2.3	<i>in vivo</i> methods .....	47
2.3.a	Details of Cre Recombinase inducible mice .....	47
2.3.b	Rapamycin treatment.....	48
2.3.c	Low dose cycloheximide treatment. ....	48
2.3.d	Survival studies .....	48
2.3.e	Intestinal regeneration assay .....	48
2.3.f	BrdU incorporation in vivo .....	49
2.3.g	Murine intestinal crypt extraction for polysome profiling .....	49
	i) Standard method.....	49

ii) Rapid extraction method.....	49
2.4 Functional assays for cultured cells.....	49
2.4.a Cell proliferation assay .....	49
2.4.b Protein synthesis assay.....	50
i) Adherent cells.....	50
ii) Matrigel cultures .....	51
2.5 DNA methods .....	51
2.5.a Phenol-chloroform extraction .....	51
2.5.b Reverse transcription.....	52
2.5.c Polymerase chain reaction (PCR).....	52
2.5.d Agarose gel separation and extraction.....	53
2.5.e Restriction endonuclease digestion of plasmid DNA .....	53
2.5.f DNA cloning .....	54
i) DNA ligation.....	54
ii) Transformation of competent E.Coli .....	54
iii) Miniprep extraction of plasmid DNA.....	55
2.6 RNA methods.....	55
2.6.a Acid phenol RNA extraction.....	55
2.6.b Trizol RNA extraction .....	55
2.6.c Denaturing agarose gel electrophoresis.....	55
2.6.d Passive transfer northern blotting for agarose gels .....	56
2.6.e Denaturing Polyacrylamide gel electrophoresis (PAGE).....	57

i) Acid Urea PAGE.....	57
ii) TBE-Urea PAGE .....	58
2.6.f Northern blotting of PAGE gels .....	59
i) Wet Electrophoretic transfer northern blotting for Acid Urea PAGE gels .....	59
ii) Semi-dry transfer northern blotting for TBE Urea PAGE gels.....	59
2.6.g Probing northern blots with radiolabelled probes.....	59
i) Probing northern blots with [ <sup>32</sup> P]-α-dCTP radiolabelled DNA.....	59
ii) Probing northern blots with [ <sup>32</sup> P]-γ-ATP radiolabelled oligonucleotides .....	60
2.6.h Polysome profiling.....	61
i) Preparation of sucrose gradients .....	61
ii) Preparation of adherent cells for polysome profiling, scrape method .....	61
iii) Harringtonine run-off treatment.....	62
iv) Preparation of adherent cells for polysome profiling, direct lysis method.....	62
v) Lysis of pelleted colonic crypt epithelia.....	62
vi) Fractionation and RNA extraction from sucrose gradients .....	62
2.6.i Ribosome profiling.....	63
i) Preparation of sucrose gradients .....	63
ii) Preparation of cultured cells for ribosome profiling .....	63
iii) Fractionation and RNA extraction from sucrose gradients .....	63
iv) RPF sequencing library generation .....	64
v) Whole transcriptome sequencing library generation .....	66
2.6.j Bioinformatic analysis of ribosome profiling data.....	66

i) Mappings .....	66
ii) Alignment of fragments relative to CDS .....	66
iii) Determination of minimum read thresholds for inclusion of gene in hypothesis testing.....	67
iv) Calculation of translational efficiency .....	67
2.6.k    Generating RNA dimers .....	67
2.6.l    Generation of capped, radiolabelled, $\beta$ -actin mRNA .....	68
i) Template preparation.....	68
ii) <i>in vitro</i> transcription of moderate specific activity capped RNA.....	68
2.6.m    RNase protection assay .....	69
i) Generation of high specific activity probe .....	69
ii) Generation of radiolabelled RNA markers .....	70
iii) RNase protection .....	71
iv) Visualisation.....	71
2.6.n    Alternate ligation methods for deep sequencing.....	72
i) Selection of partially degenerate 20 nt RNA pool .....	72
ii) Ligation of Universal Cloning linker .....	72
iii) Reverse transcription with Ion Torrent compatible primer .....	73
iv) Circularisation of cDNA.....	74
2.6.o    Sequencing using the Ion Torrent .....	74
2.7    Protein methods.....	74
2.7.a    Immunoblotting.....	74

i) Solutions .....	74
ii) Lysate preparation.....	75
iii) SDS-Polyacrylamide gel electrophoresis.....	75
iv) Transfer of protein to PVDF membrane .....	76
v) Protein detection.....	76
2.7.b Immunohistochemistry .....	77
i) Slide preparation .....	77
Chapter 3. The contribution of tRNA <sub>i</sub> <sup>Met</sup> to oncogenic transformation .....	79
3.1 Introduction.....	79
3.2 Establishment of new high tRNA <sub>i</sub> <sup>Met</sup> MEF cell lines .....	80
3.3 tRNA <sub>i</sub> <sup>Met</sup> and translational control .....	85
3.4 Using ribosome profiling to identify differentially translated mRNAs .....	89
3.5 RNA sequencing using SOLiD technology .....	92
3.5.a SOLiD sequencing reaction overview .....	92
3.5.b Conversion of colour calls to nucleotide reads .....	93
3.6 Analysis of ribosome profiling read data.....	93
3.7 Elevated tRNA <sub>i</sub> <sup>Met</sup> is not always associated with the oncogenic phenotype.....	106
Chapter 4. mTOR signalling in colorectal cancer.....	107
4.1 Introduction.....	107
4.2 Inhibition of mTORC1 signalling significantly reduces the oncogenic effects of APC deletion in mice.....	109

4.3	The effects of upregulated mTORC1 signalling in APC deficient mice are primarily due to changes in translation elongation.....	113
4.4	Improvements in intestinal crypt isolation.....	124
4.5	K-Ras mutation results in adenomas dependent on mTORC2 signalling rather than mTORC1.....	126
4.6	Discussion .....	130
Chapter 5.	Translation elongation as a component of the cold shock response.....	133
5.1	Introduction.....	133
5.2	Ribosome profiling of cold shock treated cells.....	136
5.3	Investigating translatoe reads in the 3'UTR .....	140
5.4	Discussion .....	148
Chapter 6.	The identification and remediation of biases in next generation sequencing protocols.	150
6.1	Introduction.....	150
6.2	Ribosome protected fragments may not fully reflect ribosome occupancy.....	151
6.3	Ligation biases in deep sequencing libraries .....	163
6.4	Use of an alternate library preparation method reduces bias in Ion Torrent sequencing.....	167
6.5	Discussion .....	177
Chapter 7.	Summary.....	182
	Bibliography.....	187

## List of Figures

Figure 1-1 Schematic of the mechanism of translation initiation .....	20
Figure 1-2 uORFs and translation .....	23
Figure 1-3 Schematic of the mechanism of translation elongation .....	30
Figure 1-4 An overview of mTOR signalling.....	37
Figure 3-1 Construction of a selectable tRNA <sub>i</sub> <sup>Met</sup> expression vector .....	82
Figure 3-2 Elevated tRNA <sub>i</sub> <sup>Met</sup> is associated with increased protein synthesis and cell proliferation.....	84
Figure 3-3 Elevated tRNA <sub>i</sub> <sup>Met</sup> expression is associated with increased polysomal association	86
Figure 3-4 Expression and phosphorylation status of eIFs and oncogenic proteins in response to increased tRNA <sub>i</sub> <sup>Met</sup> .....	88
Figure 3-5 Ribosome profiling schematic .....	91
Figure 3-6 Mapping statistics for SOLiD sequencing runs .....	95
Figure 3-7 Ribosome profiling gives sub-codon resolution of ribosome density.....	98
Figure 3-8 Applying minimum read thresholds to sequencing data .....	101
Figure 3-9 Consequences of applying minimum read thresholds to sequencing data .....	103
Figure 3-10 Inter-replicate variation confounds attempts to identify differentially translated messages. ....	105
Figure 4-1 Inhibition of mTORC1 signalling prolongs life of APC deficient mice.....	111
Figure 4-2 Reduction of mTORC1 signalling by Raptor deletion reverses the hyperproliferative phenotype of APC deficiency.....	112
Figure 4-3 APC <sup>fl/fl</sup> and APC <sup>fl/fl</sup> Raptor <sup>fl/fl</sup> intestinal epithelia have altered polysomal association compared to WT.....	114
Figure 4-4 Low dose cycloheximide mimics the effect of rapamycin treatment in APC <sup>fl/fl</sup> intestines .....	115

Figure 4-5 Protein synthesis and translation elongation rates are elevated in cultured APC <sup>fl/fl</sup> crypts .....	118
Figure 4-6 Intestinal regeneration is sensitive to mTORC1 inhibition and requires the inhibition of eEF2k.....	120
Figure 4-7 Intestinal regeneration in Drosophila also requires mTORC1 to increase the activity of downstream eEF2.....	123
Figure 4-8 Improved method of intestinal epithelial extraction .....	125
Figure 4-9 Addition of K-Ras mutation to APC deficiency removes sensitivity to mTORC1 inhibition.....	127
Figure 4-10 K-Ras mutation creates tumours that are dependent on mTORC2 signalling. ...	129
Figure 5-1 The cold shock response in cultured mammalian cells.....	135
Figure 5-2 Translation elongation is reduced during cold shock.....	137
Figure 5-3 Changes in ribosome profiling reads in response to cold shock. ....	139
Figure 5-4 RNase protection assay for gapdh CDS .....	142
Figure 5-5 RNase protection assay on fragments within the monosome peak .....	145
Figure 5-6 RNase protection assay does not show evidence of 3' UTR ribosome footprints	147
Figure 6-1 RNase I digestion does not reduce all ribosome footprints to monosomes.....	154
Figure 6-2 The effect of buffer composition on the bi-ribosome peak.....	157
Figure 6-3 Increased RNase I digestion increases bi-ribosome peak in MEFs but not HeLa..	159
Figure 6-4 Transfected <i>in vitro</i> transcribed actin mRNA does not share the same polysomal distribution as the native transcript .....	162
Figure 6-5 Uneven read distributions within transcriptome libraries.....	164
Figure 6-6 Schematic of Life Technologies and alternate library generation methods .....	166
Figure 6-7 Alternative library preparation method compatible with Ion Torrent sequencing .....	169
Figure 6-8 Trimming of sequencing reads.....	171

Figure 6-9 Sequence specific bias in library constructed with Mth K97A .....	172
Figure 6-10 Alternate library generation methods reduce, but do not eliminate, bias.....	174
Figure 6-11 Over-represented fragments are predicted to have lower folding energies.....	176

## **Abbreviations**

A: Adenine

AMP: Adenosine monophosphate

AMPK: AMP activated protein kinase

ATP: Adenosine triphosphate

dATP: deoxyadenosine triphosphate

APC: Adenomatous polyposis coli

APS: Ammonium persulphate

BCP: 1-bromo-3-chloropropane

BrdU: 5-bromo-2'-deoxyuridine

BSA: Bovine Serum Albumin

CRC: Colorectal cancer

C: Cytosine

CTP: Cytidine triphosphate

dCTP: deoxycytidine triphosphate

DMEM: Dulbecco's Modified Eagle Medium

DNA: Deoxyribonucleic acid

DT: Diphtheria toxin

DTT: Dithiothreitol

EDTA: Ethylenediaminetetraacetic acid

eEF: eukaryotic elongation factor

eEF2k: eukaryotic elongation factor 2 kinase

eIF: eukaryotic initiation factor

eRF: eukaryotic release factor

ERK: Extracellular signal-regulated kinase

FAP: Familial adenomatous polyposis

FCS: Fetal Calf Serum

G: Guanine

GDP: Guanosine diphosphate

GEF: Guanine nucleotide exchange factor

GTP: Guanosine triphosphate

dGTP: deoxyguanosine triphosphate

H&E: Haematoxylin and eosin

HBSS: Hank's Balanced Salt Solution

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HSB: High salt buffer

ISC: Intestinal stem cell

LSB: Low salt buffer

MEFs: Mouse embryo fibroblasts

MEK: Mitogen-activated protein kinase kinase

MES: 2-(N-morpholino)ethanesulfonic acid

Mth: *Methanobacterium thermoautotrophicum* ligase

ORF: Open reading frame

PABP: Poly(A) binding protein

PAGE: Polyacrylamide gel electrophoresis

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

PDCD4: Programmed cell death 4

PDK1: Phosphoinositide dependent kinase 1

PEG: Polyethylene glycol

pHH3: Phospho-histone H<sub>3</sub>

PIC: Pre-initiation complex

PKC: Protein kinase C

PNK: Polynucleotide kinase

p90RSK: p90 Ribosomal S6 kinase

RNA: Ribonucleic acid

dsRNA: Double-stranded RNA

Rnl1: T4 RNA ligase 1

Rnl2: T4 RNA ligase 2

RPS6: Ribosomal protein S6

S6K: p70 Ribosomal protein S6 kinase(s)

SDS: Sodium dodecyl sulphate

SGK1: Serum and glucocorticoid induced kinase 1

SOLiD: Sequencing by oligonucleotide ligation and detection

T: Thymidine

TA: Transit amplifying

TC: Termination complex

TCA: Trichloroacetic acid

TE: Translational efficiency

TEMED: Tetramethylethylenediamine

TOP: 5' Terminal oligopyrimidine tract

TOR: Target of rapamycin

tRNA: Transfer RNA

TSC: Tuberous sclerosis complex

TTP: Thymidine triphosphate

dTTP: deoxythymidine triphosphate

uORF: upstream open reading frame

UTP: Uridine triphosphate

UTR: Untranslated region

VHL: von Hippel Lindau

## Chapter 1. Introduction

The regulation of the abundance and activity of cellular proteins is essential to the response to intracellular and extracellular conditions. Unsurprisingly, the control of the expression of protein coding genes is central to this. Gene expression is a multi-stage process involving the nuclear transcription of a pre-mRNA, splicing of this to form a mature mRNA, export to (and localisation within) the cytoplasm, translation of the coding sequence of the mRNA to yield the protein product, which may be subject to post translational modifications that affect its activity and/or degradation. Regulation can occur at any of these stages but, at least in cultured cells, the single biggest determinant of gene expression is the rate of translation of the mRNA message (Schwanhausser et al., 2011).

A complete round of translation in a cell culture system typically takes 2 minutes (Ruvinsky et al., 2005), and thus regulation at the level of translation allows for a rapid response to stimuli. Where components common to the translation of all mRNAs are regulated the response is global. Specific /differential regulation is determined by sequences and structures within the mRNA in addition to the role of RNA-binding proteins. Furthermore, these two forms of regulation are not mutually exclusive, for example allowing the global downregulation of translation rates but specific upregulation of the translation of mRNAs required to respond to a particular stimulus.

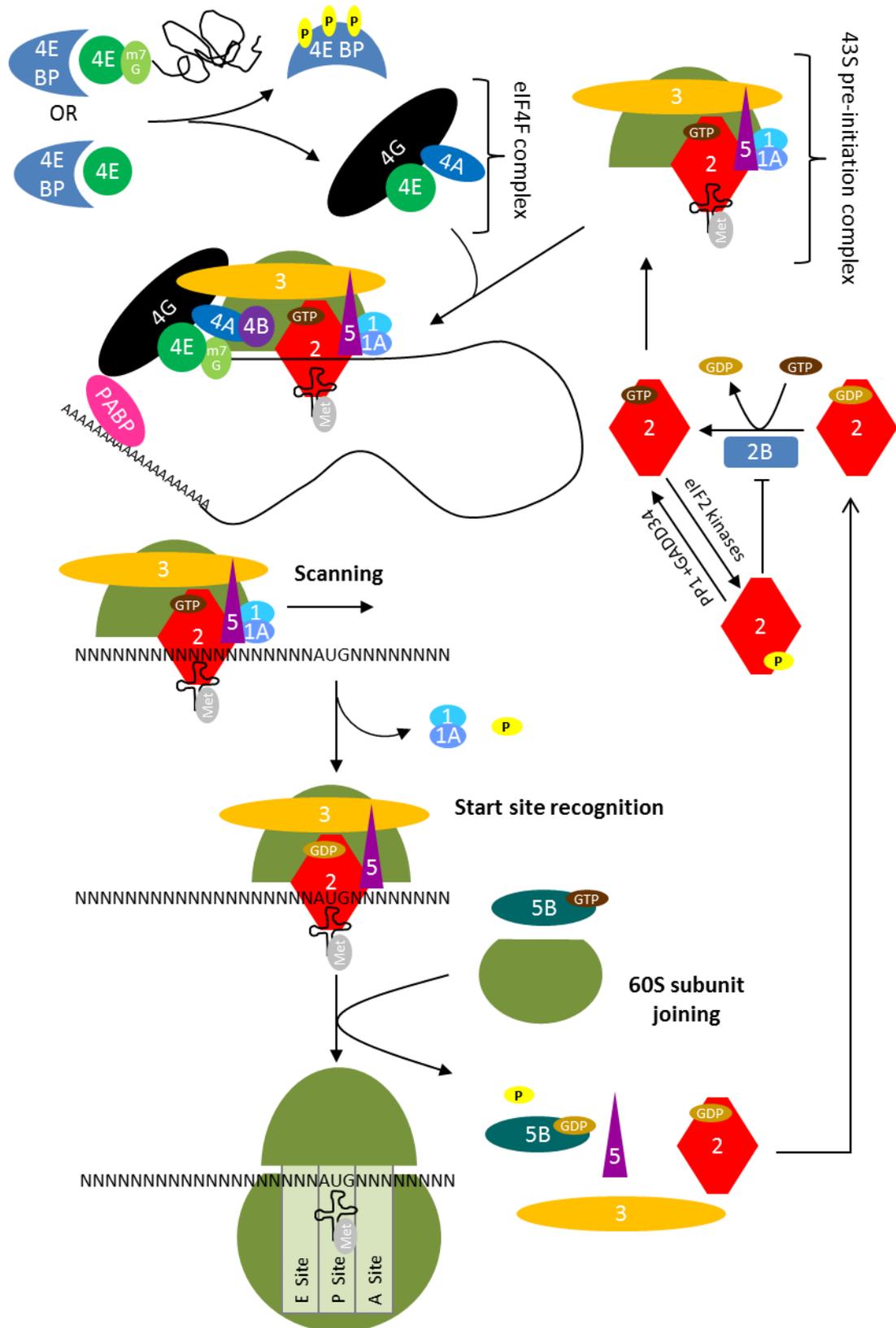
Translation can be divided into three primary stages: initiation, elongation and termination and ribosome recycling. Initiation is typically reported as being the rate limiting step in translation (Mathews et al., 2007) and modulation of initiation factors is observed in response to a range of stimuli. However, systematic modulation of different components of the translational machinery in exponentially growth yeast suggest that elongation may be

rate limiting under those conditions (Firczuk et al., 2013). Furthermore, this thesis will present evidence that modulation of elongation rate is central in other contexts.

This Introduction will outline the required molecular components of all three stages of translation, interweaving the effects of their regulation.

### **1.1 Mechanism and regulation of translation initiation**

Translation initiation is the process by which a translation competent ribosome is recruited to the start codon of an mRNA with the aid of initiation factors (eIFs: eukaryotic initiation factors). The current model for this involves the scanning of a 43S pre-initiation complex (PIC) – minimally consisting of the 40S ribosomal subunit, eIF1, eIF1A, eIF2-GTP, eIF3, eIF5 and complex of initiator tRNA methionine (Met-tRNA<sub>i</sub><sup>Met</sup>) – along the 5' untranslated region (UTR) of a message until a start site is recognised, upon which a 60S subunit is recruited and translation elongation begins (reviewed in (Aitken and Lorsch, 2012) and represented Figure 1.1).



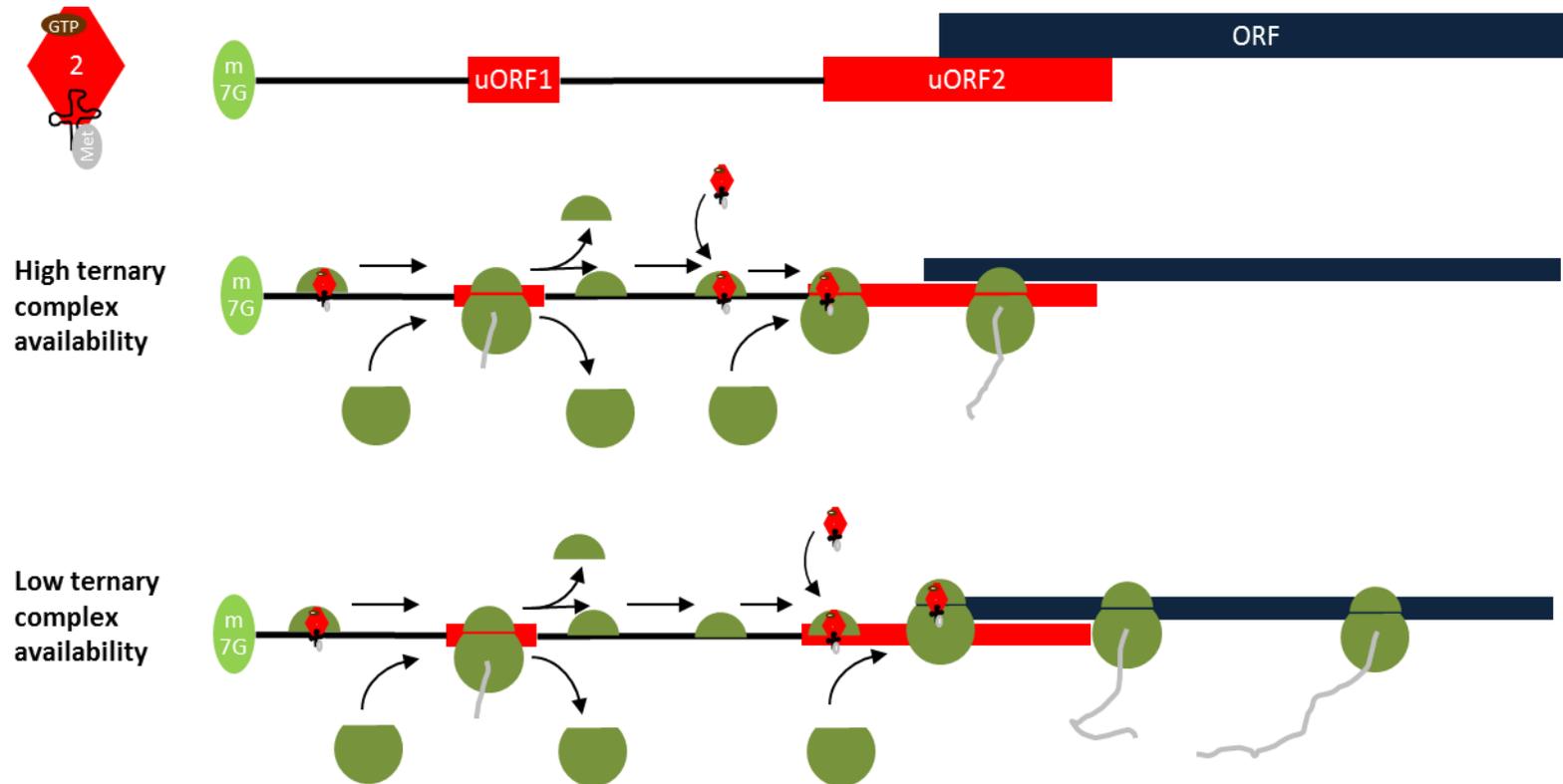
**Figure 1-1 Schematic of the mechanism of translation initiation**

Translation initiation is mediated by the numbered eukaryotic initiation factors (eIFs) depicted and described within the text. Global initiation rates are governed by the availability of either eIF4F complex or Met-tRNA<sub>i</sub><sup>Met</sup>-eIF2-GTP.

### 1.1.a Recruitment of tRNA<sub>i</sub><sup>Met</sup> to the 40S ribosomal subunit

Typically, start codons are decoded by tRNA<sub>i</sub><sup>Met</sup> (methionine codons in the body of a reading frame are decoded by elongator tRNA methionine, tRNA<sub>e</sub><sup>Met</sup>) and thus delivery of Met-tRNA<sub>i</sub><sup>Met</sup> may be a limiting factor for translation initiation. Met-tRNA<sub>i</sub><sup>Met</sup> is found in the PIC bound to eIF2-GTP. Human cell lysates contain eIF2 in a free form and as part of multifactor complex with eIF1, eIF2, eIF3 and eIF5, with recruitment of Met-tRNA<sub>i</sub><sup>Met</sup> possible in both forms (Sokabe et al., 2012). During initiation, eIF2 bound GTP is hydrolysed yielding eIF2-GDP. Met-tRNA<sub>i</sub><sup>Met</sup> binds eIF2-GTP with greater affinity than eIF2-GDP, therefore efficient recycling of GDP to GTP following a single round of initiation is required to recruit a new tRNA molecule (forming an active ternary complex of Met-tRNA<sub>i</sub><sup>Met</sup>-eIF2-GTP). eIF2B, a guanine nucleotide exchange factor (GEF) that facilitates the release of GDP from eIF2 (Williams et al., 2001), is sequestered by eIF2 phosphorylated on the  $\alpha$ -subunit (eIF2 $\alpha$ -P). Because eIF2 is vastly more abundant than eIF2B, modest changes in eIF2 $\alpha$ -P and/or eIF2B abundance can dramatically affect ternary complex availability. There are four eIF2 $\alpha$  kinases in mammals, each of which respond to different stresses including, but not limited to, amino acid starvation and DNA damage (reviewed in (Proud, 2005)). Equally there are eIF2 $\alpha$  phosphatases, most notably protein phosphatase 1, which is activated by GADD34, a protein induced when ternary complex availability is low forming a negative-feedback loop to allow recovery from translation inhibition following these stress responses (Novoa et al., 2001). Importantly, whereas eIF2 $\alpha$  phosphorylation may be associated with a reduction in global translation levels, the translation of some transcripts is upregulated. A mechanism underlying this is the presence of multiple upstream open reading frames (uORFs) within such transcripts. A paradigmatic example is the mammalian ATF4 mRNA which contains two uORFs, the second overlapping the coding region thus precluding translation of both the ATF protein and uORF-2 by the same ribosome (Figure 1-2). Following termination at the first uORF, the 40S subunit continues to scan and if active ternary complex is acquired, will

reinitiate at uORF-2. When ternary complex availability is reduced it becomes more likely that ribosomes acquire ternary complex after scanning past the start site of uORF-2 and thus re-initiate at the coding sequence (Vattem and Wek, 2004). ATF4 is a transcription factor that targets genes involved in amino acid synthesis and import, creating a closed-loop whereby phosphorylation of eIF2 $\alpha$  by GCN2 in response to amino acid depletion increases translation of ATF4 until amino acid levels are raised and eIF2 $\alpha$ -P levels fall (Harding et al., 2003). Up to 35 % of mRNAs are predicted to have uORFs (Iacono et al., 2005), however this does not mean all will be capable of supporting re-initiation, for example long uORFs do not result in reinitiation (Kozak, 2001, Luukkonen et al., 1995). Even where reinitiation does occur not all ribosomes do so and thus uORFs are generally inhibitory to translation.



**Figure 1-2 uORFs and translation**

Schematic showing the translation of an mRNA, such as ATF4, with two upstream open reading frames (uORFs), one overlapping the main ORF. Decreased ternary complex availability increases scanning past the second uORF increasing translation of the coding sequence. This is also the case even where the uORF does not overlap the main ORF as not all terminated ribosomes reinitiate.

### **1.1.b Start site recognition**

The scanning model requires a mechanism by which the PIC recognises a start codon and recruits a 60S subunit. The sequence of the nucleotides surrounding the start site - the Kozak consensus sequence, GCC(A/G)CCAUGG, (Kozak, 1987) – and the presence of eIF1A and eIF1 contribute to this. Initiation does occur from non-AUG codons, most commonly CUG, and from AUG codons in weaker context. However experiments performed *in vitro* show that mutation of transcripts away from the Kozak consensus reduces initiation and increases “leaky” scanning, whereby the PIC scans past an AUG and initiates at a downstream start codon (Kozak, 1991). Initiation does not occur in the absence of eIF1, and although eIF1A is somewhat dispensable *in vitro*, leaking scanning and recognition of codons in weak context is reduced when both factors are present (Pestova and Kolupaeva, 2002). Cryo-electron microscopy has shown that eIF1 and eIF1A bind to the yeast 40S subunit and collectively maintain a structure that is more permissive to scanning – an open conformation (Passmore et al., 2007). The binding sites of eIF1 and eIF1A on the yeast ribosome are analogous to those reported on the mammalian ribosome (Lomakin et al., 2003) and thus it is assumed the same conformation is found in all eukaryotes. Upon recognition of a start site, eIF2 bound GTP is hydrolysed by eIF5 and eIF1 is ejected resulting in a closed structure that can no longer scan (Passmore et al., 2007).

### **1.1.c Recruitment of the pre-initiation complex to mRNA**

Whilst the PIC can independently scan and initiate translation from the start site of mRNAs containing unstructured 5'UTRs such as  $\beta$ -globin (Pestova and Kolupaeva, 2002), it is typically recruited to mRNAs by additional factors and requires further factors to remove secondary structures inhibitory to scanning.

Eukaryotic mRNAs contain a 5' cap structure that is added concurrent with the transcription of the pre-mRNA (Shatkin, 1976). eIF4E binds to this cap complex and can recruit either of

two eIF4G isoforms (Gradi et al., 1998), both large (~220kDa) scaffold proteins. Mammalian eIF4G binds to eIF3 via a site in the central region of the protein (Korneeva et al., 2000), thus facilitating recruitment of the PIC to capped mRNA. eIF4G also contains two binding sites for the RNA helicase eIF4A (Imataka and Sonenberg, 1997) and PolyA-binding protein (PABP) (Imataka et al., 1998). The helicase activity of eIF4A is dramatically stimulated by the presence of eIF4G and eIF4B or eIF4H, with the eIF4B more potent than eIF4H *in vitro* (Ozes et al., 2011). However even these proteins are insufficient to permit scanning through highly structured mRNA and other helicases such as DHX29 are required in these cases (Pisareva et al., 2008). Programmed cell death 4 (PDCD4) inhibits eIF4A helicase activity and can be regulated at the transcriptional level or by phosphorylation which targets the protein for degradation (Lankat-Buttgereit and Goke, 2009). Two of the three isoforms of eIF4A are thought to be functionally interchangeable as helicases (Li et al., 1999), however recent work has shown that eIF4AII is specifically required for microRNA mediated repression of translation by a mechanism enhanced by structure within the 5' UTR (Meijer H, Lu W.T., Kong Y.W., Wilczynska A and Bushell M, in press). As its name suggests, PABP binds to the PolyA tail, and so the interaction between eIF4G and PABP results in the circularisation of mRNA. Circularisation increases translational efficiency (Niepel et al., 1999) and this is probably due to reducing the distance required for terminated ribosomes to travel before reinitiating rather than permitting scanning along the 3' UTR (Rajkowitsch et al., 2004).

eIF4E, eIF4G and eIF4A are collectively known as the eIF4F complex and the availability of this complex may affect the rate of initiation by reducing recruitment of the PIC to capped mRNAs. The binding of 4E-binding proteins (4E-BPs) to eIF4E prevents association with eIF4G and therefore are central to regulation of eIF4F availability. There are at least six phosphorylation sites on the 4E-BPs and an increase in phosphorylation decreases affinity for eIF4E. This process forms part of the mTORC1 signalling pathway which shall be introduced in greater detail in a later section. In addition, an increase in eIF4E levels will counter the

negative regulation by 4E-BPs. This will also be discussed later as it is frequently observed in tumorigenesis.

Whilst recruitment of the PIC is typically cap-dependent, this is impaired in a number of stress conditions, for example during picornavirus infection and apoptosis. Under both of these conditions 4E-BP1 phosphorylation is reduced (Bushell et al., 2006, Gingras et al., 1996) and eIF4G is cleaved between the eIF4E binding site and the eIF3 binding site (Clemens et al., 1998, Lamphear et al., 1995). The ability of the C-terminal (eIF3 binding) fragment of eIF4G to facilitate translation of a capped mRNA *in vitro* is drastically reduced, but not abolished (Ali et al., 2001). Nonetheless the drastic reduction of eIF4F availability caused by the above changes favours the translation of mRNAs that can initiate in a cap independent manner, allowing a global downregulation of protein synthesis but a relative upregulation of such transcripts. Many viral transcripts, including picornaviral ones, contain internal ribosome entry sites (IRESs) upstream of their coding sequences, able to initiate translation in the absence of a cap (the proof of this is rather elegantly demonstrated in (Chen and Sarnow, 1995)). Most IRESs are highly structured, but there is little commonality in these structures and the factor requirements of different viral IRESs varies greatly (Kieft, 2009). Whilst the identification of some putative IRES containing cellular transcripts may have been an artefact of the reporter system used (Kozak, 2008, Lemp et al., 2012), these concerns do not account for all IRES containing transcripts (Komar and Hatzoglou, 2011). Furthermore, despite a global decrease in ribosomal association and protein synthesis during apoptosis, the translation of a subset of mRNAs is either unchanged or upregulated (Bushell et al., 2006). Notably, 70 % of the 5'UTRs tested from this subset supported translation in classical bicistronic pRF construct without evidence of cryptic promoter/splicing activity suggesting they contain an IRES.

#### **1.1.d Large ribosomal subunit joining**

Following recognition of a start site, eIF5 and eIF2 are displaced from the 48S complex by the GTPase eIF5B, with complete dissociation occurring upon 60S subunit joining (Pisareva et al., 2006). eIF5 can remain associated with eIF2-GDP and inhibit GTP exchange. This association is increased during amino acid starvation further antagonising ternary complex recycling (Jennings and Pavitt, 2010).

### **1.2 Mechanism and regulation of translation elongation**

The process of translation elongation involves three sequential steps: recruitment of an aminoacyl-tRNA (aa-tRNA) complementary to the mRNA codon in the ribosomal A-site, peptide bond formation and ribosomal translocation to position the next (downstream) codon into the A-site. In eukaryotes the first of these steps is performed by eEF1, the second by the 60S ribosomal subunit, the third by eEF2 (Figure 1-3).

The comparative difficulty in generating crystal structures for eukaryotic, rather than prokaryotic, ribosomes means that much of the work on biophysical investigation of elongation has been performed with prokaryotic ribosomes and factors. However, the core of the ribosome is essentially conserved between all domains of life (Melnikov et al., 2012) as are elongation factors (Taylor et al., 2007a), making extrapolation from findings in prokaryotes appropriate.

#### **1.2.a Recruitment of aminoacyl-tRNAs**

To ensure the fidelity of translation it is crucial that aa-tRNAs with non-complementary anticodons do not take part in peptide bond formation. A degree of mismatch in the 3<sup>rd</sup> nucleotide (nt) of the codon can occur, wobble base-pairing, but the degree of fidelity, nor wobble-base pairing, can be accounted for by base-pairing energies alone. The ribosome itself and elongation factors therefore play a crucial role in accurate decoding (reviewed in (Ogle et al., 2003))

eEF1 consists of eIF1A and eIF1B, with the latter consisting of three subunits –  $\alpha, \beta, \gamma$ . The GTP-bound form of eEF1A can bind to aa-tRNA. Base-pairing can occur between the tRNA anticodon and A-site mRNA codon, but the amino acid end of the tRNA can only fully enter the A-site, and thus take part in peptide bond formation, upon release from eEF1. The hydrolysis of eEF1A-GTP to eEF1A-GDP drastically reduces the affinity for aa-tRNA. This process is mediated by the ribosome in a manner that recognises stem loop structures of cognate tRNAs but not mispaired ones (Connell et al., 2007, Pape et al., 1998). Antibiotics and/or rRNA mutations that affect the structure of the ribosome can therefore dramatically affect the fidelity of translation (Ogle et al., 2003). Much like the  $\text{tRNA}_i^{\text{Met}}$  ternary complex, GTP exchange must occur rapidly on eEF1A to allow recruitment of a new aa-tRNA. eEF1B acts as a GEF, with both the  $\alpha$  and  $\beta$  subunits of eEF1B containing GEF domains (reviewed in (Browne and Proud, 2002)).

### **1.2.b Peptide bond formation**

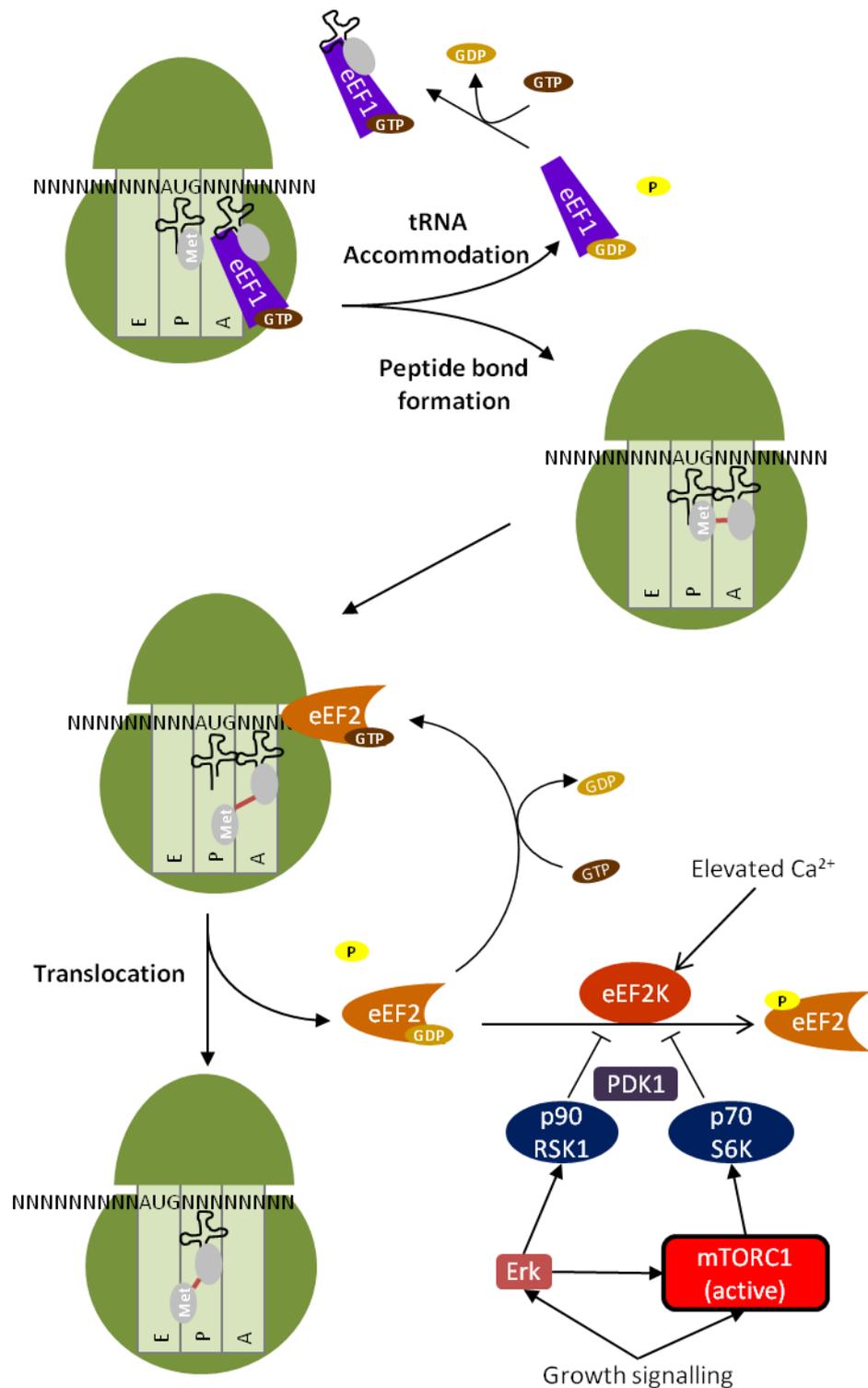
Following accommodation of the aminoacyl stem of the tRNA into the A site, peptide bond formation is catalysed by the 60S subunit. The catalytic region, the peptidyl transferase center (PTC), is within the ribosomal core, conserved in all domains, and is exclusively rRNA – the ribosome is a ribozyme (Cech, 2000).

### **1.2.c Ribosomal translocation**

The final step of elongation is the translocation of the ribosome to allow recognition of the next codon. Ribosomal translocation occurs via a ratchet like mechanism of the ribosomal subunits triggered by the binding and hydrolysis of eEF2-GTP (Spahn et al., 2004, Taylor et al., 2007b). eEF2 leaves the ribosome as eEF2-GDP, but guanosine exchange is sufficiently rapid as to not require an additional factor (unlike eEF1A). The eEF2 protein has two key residues that are sites of regulation. The first, H714 in human eEF2, is modified to diphthamide forming a unique site for ADP-ribosylation, and thus translational inhibition, by

diphtheria toxin (DT). Whilst mutation of this site to prevent diphthamide modification confers resistance to DT without affecting protein synthesis rates in cultured mammalian cells (Ivankovic et al., 2006), this is associated with developmental defects in mice (Liu et al., 2012). Interestingly, expression of eEF2 that does not contain diphthamide is associated with an increase in -1 frameshifting in yeast (Ortiz et al., 2006) and human cells (Liu et al., 2012) suggesting this modification is required for ribosomal translocation across all 3 nucleotides of a codon. The second site of modification is within the GTP-binding domain (T56 in human eEF2). Phosphorylation of this site, by eEF2 kinase (eEF2k) prevents binding of eEF2 to the ribosome (Carlberg et al., 1990). eEF2k can be phosphorylated on a number of serine and threonine residues, resulting in either the inhibition or activation of kinase activity. Under basal  $Ca^{2+}$  conditions (and contingent on the activity of phosphoinositide dependent protein kinase-1 (PDK1)) eEF2k is inhibited by phosphorylation on serine 366 by p70 ribosomal protein S6 kinase (S6K) or p90RSK1 downstream of mTORC1 and Ras-MEK-ERK pathways respectively (Wang et al., 2001). However, autophosphorylation of additional residues occurs in a calmodulin mediated response to elevated  $Ca^{2+}$  (Redpath and Proud, 1993) which can overcome S6K/p90RSK mediated inhibition (Wang et al., 2001). The relative importance of these two opposing regulatory pressures *in vivo* is not yet clear.

There are up to 9 residues of eEF2k that can be phosphorylated (Pyr Dit Ruys et al., 2012). Intriguingly a mutation that abolishes phosphorylation at serine 366 reduces the kinase activity of autophosphorylated eEF2k towards the synthetic substrate MH-1 peptide. This suggests there are important differences between the action of eEF2k on this peptide and the endogenous target eEF2 meaning interpretation of existing literature based on MH-1 as a read out should be done with caution.



**Figure 1-3 Schematic of the mechanism of translation elongation**

Aminoacyl tRNAs are recruited to the ribosome by elongation factor 1 (eEF1)-GTP. Hydrolysis of GTP upon recognition of cognate tRNA allows full entry into the A site. Peptide bond hydrolysis is catalysed by the peptidyl transferase centre within the 60S subunit. Elongation factor 2 is required for the translocation of the ribosome but can be inhibited by phosphorylation by elongation factor 2-kinase via mTORC1 or Erk mediated phosphorylation of PDK1 targets: ribosomal S6 kinases p70 and p90, under normocalcaemic conditions

### 1.3 Mechanism of translation termination and ribosomal recycling

In all domains, termination of translation requires the recognition of a stop codon in the A site by a class I release factor (RF) that mimics a tRNA and catalyses the hydrolysis of the tRNA in the P site releasing the peptide chain. This process also involves a class II RF, but there are significant differences in the actions of class II RFs between domains. Furthermore, whilst prokaryotes that have two class I RFs (RF1 and RF2) that each recognise two of the three stop codons, eRF1 is capable of recognising all 3 stop codons, thus explaining the slightly innumerate nomenclature of the eukaryotic class II RF, eRF3!

Monomeric eRF1 does not have a tRNA-like conformation, but binding to eRF3 induces a conformational change that permits mimicry of a tRNA. This additionally increases the affinity of eRF3 for GTP, meaning that termination is achieved by an eRF1-eRF3-GTP complex somewhat analogous to tRNA-eEF1-GTP (Cheng et al., 2009). A conformation change in the pre-termination ribosomal complex (pre-TC) occurs upon binding of eRF1, as part the eRF1-eRF3-GTP. Following ribosome-mediated GTP hydrolysis of eRF3, it is proposed the eRF1 is positioned closer to PTC resulting in rapid peptide release (Alkalaeva et al., 2006, Cheng et al., 2009).

Peptide release does not result in spontaneous ribosome dissociation, and eRF1 remains associated with the post-termination complex. Dissociation of the post-TC is required for ribosomes to participate in translation initiation (ribosome recycling) and is thus essential for the maintenance of translation. The full dissociation of both the 40S and the 60S subunits requires eIF3, eIF1 and eIF1A, however this will only occurs *in vitro* under low, likely sub-physiological,  $Mg^{2+}$  concentrations (Pisarev et al., 2010). Where  $Mg^{2+}$  concentration is not permissive, ABCE1 is required for ribosomal dissociation, coupling termination to ribosome recycling (Pisarev et al., 2010). ABCE1 is an ATPase highly conserved in eukaryotes that binds to both ribosomal subunits and eRF1 and following ATP hydrolysis has been modelled to

prise open the ribosome (Becker et al., 2012).

The importance of ABCE1, and ribosome recycling more generally, is not limited to ribosomes arrested on stop codons. Numerous surveillance mechanisms operate during the pioneer round of translation on a transcript, targeting it for degradation if there have been errors in transcription or processing (Maquat et al., 2010). In two of these mechanisms (the no-go and non-stop decay pathways), ribosomes are stalled on transcripts and translation termination and ribosome recycling must occur. This process involves a complex analogous to that required in termination on stop codons - a eRF1 homolog (Dom34 in yeast, Pelota in mammals) associated with GTP-bound-Hbs1 (a eRF3 homolog). Whilst ABCE1 is not strictly required for dissociation of post-TCs in yeast (Shoemaker et al., 2010), it is essential for this process in mammals (Pisareva et al., 2011).

#### **1.4 mTOR signalling**

Target of rapamycin (TOR) proteins are serine/threonine kinases that are present in two complexes, TOR complex 1 (TORC1) and TOR complex 2 (TORC2), conserved from yeast to human. In yeast the two TOR proteins are functionally redundant, whilst only one TOR protein is found in higher eukaryotes. TORCs form part of interconnected signalling networks that mediate both transcriptional and translation response to intracellular and extracellular growth signals. Aberrant TOR signalling has been reported to play a causal role in cancer, ageing and metabolic diseases such as type-II diabetes (Laplante and Sabatini, 2012).

That TOR proteins are inhibited by rapamycin was first identified in budding yeast (Heitman et al., 1991), however it took a further 11 years to recognise that they exist in both the rapamycin-sensitive TORC1 and the largely rapamycin insensitive TORC2 (Loewith et al., 2002). This goes some way to explaining why the details of TORC1 signalling are somewhat more elucidated than for TORC2. Whilst TOR signalling is broadly conserved, there are some differences between kingdoms which fall outside the scope of this thesis, which focuses

predominantly on mammalian systems. Both mammalian target of rapamycin complexes (mTORCs) are comprised of mTOR, mLST8, DEPTOR, Tti1 and Tel2 with mTORC1 additionally containing Raptor and PRAS 40, and mTORC2 containing Rictor, mSim1 and protor1/2 (Laplante and Sabatini, 2012). A complex of rapamycin and the intracellular FK506-binding protein (FKBP12) can allosterically inhibit some of the kinase activity of mTORC1 (Brown et al., 1995). mTOR signalling constitutes a complex and inter-connected network, depicted in Figure 1-4 and described below.

#### **1.4.a mTORC1 signalling**

The targets of mTORC1 are involved multiple different processes including, but not limited to, protein synthesis, lipid metabolism and autophagy (reviewed in (Laplante and Sabatini, 2012)). In the context of translational control, mTORC1 signalling increases protein synthesis via two primary targets: 4E-BPs and S6Ks, the latter of which must also be subsequently phosphorylated by PDK1 (Pullen et al., 1998). mTORC1 also inhibits Maf1, a negative regulator of RNA Polymerase III, such that increased mTORC1 activity is associated with increased transcription of Pol III genes such as tRNAs (Kantidakis et al., 2010). As described previously, phosphorylation of 4E-BPs reduces their affinity for eIF4E thereby increasing cap dependent translation. Of the S6K targets identified, eEF2k, PDCD4 and eIF4B may all play a role in the regulation of translation as described in earlier sections. Significantly, many transcripts encoding pro-proliferative proteins contain highly structured 5'UTRs meaning mTORC1 mediated increases in helicase activity of eIF4F is a mechanism by which the expression of such proteins is specifically up-regulated in response to growth signals (Ruggero, 2012).

The contribution of ribosomal protein S6 (RPS6) to translation is somewhat less well understood. A reduction in the phosphorylation of RPS6 correlates with a reduction in protein synthesis in response to rapamycin, but MEFs expressing a mutant RPS6 that cannot be phosphorylated at any of the 5 reported serine residues (RPS6<sup>P-/-</sup>) show an increase in

global protein synthesis rate relative to WT cells with phosphorylated RPS6 (Ruvinsky et al., 2005). An explanation for this is that despite the above correlation, phosphorylation of RPS6 does not increase protein synthesis but does compete with eEF2k phosphorylation for S6K such that when RPS6 cannot be phosphorylated there is increased phosphorylation of eEF2k. Consistent with this an increase in elongation rate was observed in RPS6<sup>P-/-</sup> MEFs but this was not dramatic enough to fully account for the increase in translation rate (Ruvinsky et al., 2005).

Inhibition of mTORC1 by rapamycin decreases the protein synthesis rate, but a subset of mRNAs show greater sensitivity to rapamycin inhibition – those containing 5'-terminal oligopyrimidine tracts (TOPs) (Jefferies et al., 1994). Rapamycin does not inhibit the S6K and 4E-BP arms of the mTORC1 pathways equally (Choo et al., 2008), which has driven the development of other active site inhibitors of mTORC1 (asTORi) such as Torin (Thoreen et al., 2009). These active site inhibitors affect the kinase activity of both mTORCs. mRNAs sensitive to both rapamycin and the asTORi pp242, contain a TOP or a stringent pyrimidine rich translation element (PRTE), and a shorter 5'UTR which, surprisingly given that eIF4B is a target of mTORC1, are predicted to be less structured (Hsieh et al., 2012). The mechanism(s) regulating TOP containing mRNAs remain unclear, as illustrated in the next paragraph.

Early work suggested that the regulation of TOP mRNAs by mTOR is mediated by S6K (Jefferies et al., 1997). However, knockout of both S6K isoforms in mouse embryo fibroblasts (MEFs) is not associated with the resistance of a TOP mRNA to rapamycin mediated inhibition suggesting TOP mRNAs are not regulated by S6K (Pende et al., 2004). Deletion of the two 4E-BPs expressed in MEFs removes the Torin 1 mediated inhibition of protein synthesis and significantly reduces, but does not abolish, the reduction in polysomal association of TOP/PRTE mRNAs under these conditions (Thoreen et al., 2012). That TOPs are still sensitive to Torin in the 4E-BP<sup>DKO</sup> cells suggests additional non-mTORC1/2 mediated

regulation of TOPs. This could explain why rapamycin, despite a limited ability to inhibit 4E-BP phosphorylation, can still affect TOPs mRNAs in S6K knockout mice. Furthermore, knockdown of raptor (and presumably mTORC1 activity) in HEK 293 cells has a minimal effect on the polysomal association of a TOP mRNA in response to insulin stimulation, unlike mTOR knockdown (Patursky-Polischuk et al., 2009).

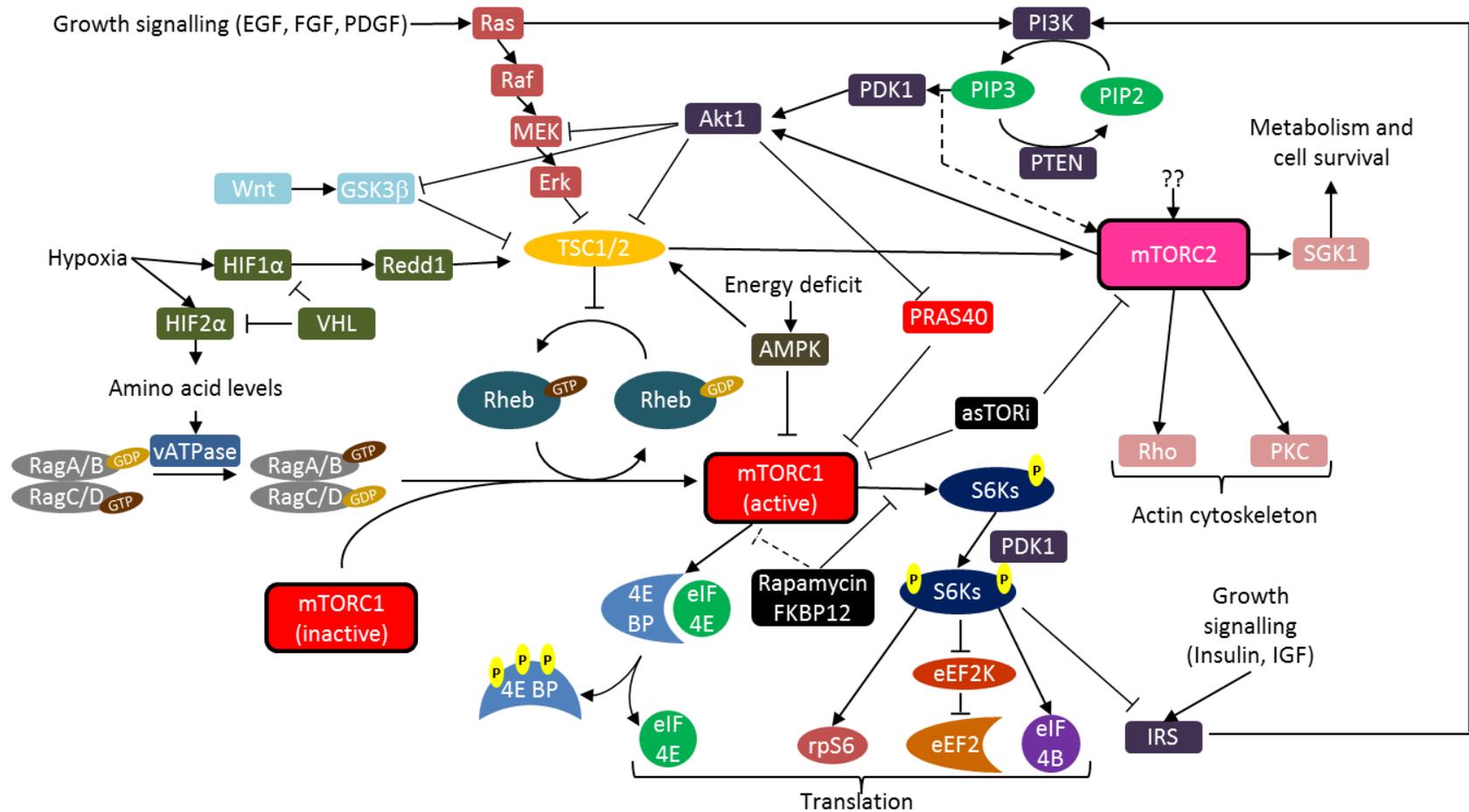
Four molecular mechanisms that regulate the kinase activity of mTORC1 in a physiological context have been identified. The activity of mTORC1 is contingent on interaction with GTP-bound Ras homolog enriched in brain (Rheb), which despite its name is widely expressed, and two mechanisms affect this.

The first involves the down regulation of translation in response to amino-acid starvation via a complex of the multi-protein GEF, Ragulator, and Rag-GTPases, bound to the lysosomal membrane (Bar-Peled et al., 2012). The Rag proteins exist as obligate heterodimers of either RagA or RagB with RagC or RagD. The GTP loading of RagA/B is always the opposite of RagC/D such that GDP-bound RagA or RagB will be found in a complex with GTP bound RagC or RagD. Upon activation in the presence of amino acids, Ragulator catalyses the GTP exchange creating GTP-bound RagA or RagB (Zoncu et al., 2011). This recruits mTORC1 to the lysosome allowing interaction with Rheb which is necessary for mTORC1 activity. Consistent with this model, expression of mutant forms of RagA or RagB that are constitutively bound by GTP results are insensitive to amino acid starvation. The second mechanism of regulation is via a dimer of the tuberous sclerosis complex proteins (TSC1/2). TSC1/2 catalyses the hydrolysis of Rheb-GTP to Rheb-GDP and thus is inhibitory to mTORC1 (Inoki et al., 2003a). Many different signalling pathways converge on TSC1/2. The Ras-MEK-ERK, PI3K and Wnt signalling pathways inhibit TSC1/2, thus coupling pro-proliferative signalling with increased mTORC1 activity and associated increases in translation. *Redd1*, a target of the hypoxia inducible transcriptional factor HIF1 $\alpha$ , activates TSC1/2, restricting proliferation when oxygen is limiting (Brugarolas et al., 2004). Interestingly, expression of HIF2 $\alpha$  causes an

increase in mTORC1 activity, but this transcription factor is suppressed under apathological conditions by von-Hippel Lindau tumour suppressor (VHL) and is typically expressed at much lower levels than HIF1 $\alpha$  (Elorza et al., 2012). AMP-activated protein kinase (AMPK) activates TSC1/2, coupling energy deficit with reduced protein synthesis via mTORC1 (Inoki et al., 2003b).

The other two mechanisms affect the composition of mTORC1. In addition to activating TSC1/2, AMPK also phosphorylates the mTORC1 essential Raptor, targeting the protein for degradation (Gwinn et al., 2008). Conversely, Akt1, a downstream effector of mTORC2, inhibits TSC1/2 and additionally dissociates PRAS40, an inhibitor of mTORC1, with Raptor, increasing mTORC1 activity (Vander Haar et al., 2007).

Finally, interaction between downstream mTORC1 effectors and upstream signalling pathways may occur. For example, S6K inhibits insulin receptor substrate, which is an activator of phosphatidylinositide-3-kinases (PI3K). PI3K activates PDK1 which activates Akt1, which in turn inhibits TSC1/2, meaning mTORC1 activation via PI3K sets off a negative feedback loop (Manning, 2004).



**Figure 1-4 An overview of mTOR signalling**

IGF = Insulin-like growth factor, EGF = Epidermal growth factor, FGF = Fibroblast growth factor, PDGF = Platelet-derived growth factor. For full details of other abbreviations see accompanying text.

#### 1.4.b mTORC2 signalling

Both the upstream regulation and downstream effectors of mTORC2 are less well characterised than for mTORC1. mTORC2 appears to be regulated by growth factor signalling, but, unlike mTORC1, the evidence for activation due to nutrient depletion is unclear (Oh and Jacinto, 2011). *In vitro*, the presence of PIP3 increases mTORC2 activity suggesting mTORC2 may be regulated by PI3K signalling (Gan et al., 2011). Furthermore, in another contrast to mTORC1, stimulation of mTORC2 requires TSC1/2 (Huang et al., 2008).

The identified targets of mTORC2 kinase activity include Akt1, serum and glucocorticoid induced kinase 1 (SGK1), protein kinase C (PKC) and Rho GTPase.

Akt1 can be phosphorylated by mTORC2 at two residues – T450 and S473 – that are distinct from the T308 site phosphorylated by PDK1. Phosphorylation at T450 occurs in a co-translational manner on the nascent peptide (Oh et al., 2010) and is believed to be stable, whereas S473 is phosphorylated in a temporal manner in response to growth signals (Oh and Jacinto, 2011). Knockdown of mTORC2 essential component, Rictor, in a number of human cell lines reduces the phosphorylation of not only S473, but also T308 to the same extent (Sarbasov et al., 2005). However this is not observed in MEFs derived from embryos with homozygous deletion of Rictor (Guertin et al., 2006) suggesting the interactions of PDK1 and mTORC2 mediated phosphorylation of Akt1 are context-dependent. Evidence from these Rictor<sup>-/-</sup> MEFs, suggests S473 phosphorylation is required for Akt1 activity against FoxO proteins but not TSC1/2 (and thus mTORC1).

SGK1 has a diverse range of targets involved in such processes as ion transport, cell survival, and growth (Lang et al., 2006). To give just one example of the contribution of SGK1 to cell survival, the pro-apoptotic transcription factor FoxO3a is targeted for degradation upon SGK1 mediated phosphorylation (Brunet et al., 2001). This contribution to the evasion of apoptosis and the role of other SGK1 targets in oncogenesis provide a link between mTORC2

and cancer that is independent of Akt1 (Bruhn et al., 2010). Furthermore PKC and Rho proteins are involved in the regulation of the actin-cytoskeleton and therefore may explain the contribution of mTORC2 to cell migration and metastasis (Gulhati et al., 2011).

## **1.5 Translational control and cancer**

In the United Kingdom, the estimated lifetime risk of developing a cancer is 1 in 3, with over a quarter of all deaths caused by it (ONS, 2009).

As illustrated in the previous section, translational control is central to the regulation of cellular growth in response to internal and external stimuli, for example via mTORC1. It should therefore come as little surprise that dysregulated translational control can result in persistent growth in the absence of proliferative stimuli – a neoplasm. The acquisition of other characteristic hallmarks of a malignant neoplasm (cancer), such as invasive potential, resistance to cell death and angiogenesis (Hanahan and Weinberg 2011) are mediated by subsets of proteins that may also be regulated at the level of translation.

Aberrant translation in cancer has been reported to be mediated by increased ribosome abundance, changes in the expression or activity of initiation and elongation factors, altered expression of RNA-binding proteins, and microRNA mediated regulation of transcripts encoding oncogenes and tumour suppressors (Blagden and Willis, 2011, Kong et al., 2012, Ruggero, 2012). This section will focus primarily on the modes by which initiation and elongation factors are dysregulated.

### **1.5.a The eIF4F complex and cancer**

Where eIF4F is limiting for translation initiation, increasing availability will cause increases in global synthesis and differential upregulation of mRNAs with long structured 5' UTRs (Koromilas et al., 1992) such as vascular endothelial growth factor (VEGF) (Scott et al., 1998). The latter is an example of the link between translational control and angiogenesis required to function in the tumour microenvironment. eIF4E is typically the least abundant of eIF4F

factors. Increased expression of eIF4E correlates with worse outcomes in cancers from at least 8 different organs (Silvera et al., 2010) and overexpression of eIF4E in a mouse models is sufficient to induce transformation (Ruggero et al., 2004). However, eIF4E availability is actually governed by ratio of eIF4E to hypophosphorylated 4E-BPs and whilst decreased expression and/or increased phosphorylation of 4E-BP has been observed in cancers, prognostic value is contingent on the 4E-BP:eIF4E axis (Coleman et al., 2009). Numerous different molecular changes have been reported to drive increased 4E-BP phosphorylation via mTORC1. Tuberous sclerosis, an autosomal dominant disorder is characterised by loss of function mutations in, or deletion of, *tsc1* and *tsc2* resulting in multi-organ neoplasia. Whilst these neoplasms are non-malignant, they still are still associated with severe pathology as they compete for space with healthy tissue, such as in the brain. In the context of sporadic cancers, components of the multiple signalling pathways upstream of TSC1/2 are frequently mutated. The three pathways most pertinent to this thesis are the Wnt signalling, PI3K/Akt1 and Ras-MEK-ERK pathways. The latter exhibit cross-talk that is positive – Ras activation increases PI3K activity – and negative – Akt1 inhibits MEK – with the balance between them affected by upstream growth factor signalling which may be altered in cancers (Aksamitiene et al., 2012, Castellano and Downward, 2011). Upregulation of Wnt signalling, observed in 93 % sporadic colorectal cancers (Cancer Genome Atlas, 2012), drives cell proliferation, at least in part, by increased expression of GSK3 $\beta$ , an inhibitor of TSC1/2. The tumour-suppressor gene *pten* is deleted or epigenetic silenced in many different cancers (Goel et al., 2004, Kang et al., 2002), thus relieving repression of PI3K signalling to both mTORCs. Mutations in the Ras and Raf proteins that result in constitutive activation of the ERK pathway have been reported in numerous different tumours of epithelial origin (Dhillon et al., 2007). Underscoring the importance of this pathway, approximately 90 % of pancreatic cancers have oncogenic mutant K-Ras (Thomas et al., 2007). Other components of the eIF4F complex may also play a role in oncogenesis. Evidence for

the involvement of eIF4A *per se* (rather than as a consequence increased eIF4F availability) are the observations that decreased PDCD4 expression drives proliferation in renal, brain and lung cancers cell lines (Jansen et al., 2004) and that increases in eIF4B expression in non-Hodgkin's lymphoma correlates with poor clinical outcomes (Dr E. Horvilleur, unpublished observations). eIF4G-1 is vastly over-expressed in ~70 % of squamous cell lung cancers relative to normal tissue (Bauer et al., 2002) and overexpression of eIF4G is sufficient to transform MEFs, independent of eIF4E expression (Fukuchi-Shimogori et al., 1997).

### **1.5.b Ternary complex and cancer**

Both decreases and increases in ternary complex availability may play a role in cancer. Reduced availability favours the translation of number of proto-oncogenes with multiple uORFs, such as *bcl-2* (Willis, 1999) and *Her2* (Child et al., 1999), but increased availability may increase global translation which can drive proliferation. Whilst overexpression of eIF2 $\alpha$  is sufficient to drive oncogenic transformation in MEFs (Donze et al., 1995) and is associated with benign and malignant tumours colonic neoplasms (Rosenwald et al., 2003), inhibition of the eIF2 $\alpha$  kinase PERK vastly reduced the size of tumours in a Ras-transformed mouse xenograph model (Fels and Koumenis, 2006). A potential explanation of these seemingly paradoxical findings is that reduced ternary complex availability causes the preferential translation of mRNAs necessary for a protective response to cellular stresses such a DNA damage (Powley et al., 2009) and hypoxia which are typical of the tumour microenvironment. Increased ternary complex availability is therefore disadvantageous for a tumour in the early stages of tumourigenesis, but the increased protein synthesis rates may be beneficial following the acquisition of mechanisms to evade stress-induced pro-apoptotic signalling (Adams and Cory, 2007) and/or increased vascularisation via a PERK mediated mechanism (Fels and Koumenis, 2006). Consistent with this increased expression and decreased phosphorylation of eIF2 $\alpha$  is observed in a variety of metastatic cancers (Silvera et al., 2010). The corollary of this is that the efficacy of chemotherapeutics that induce cell

stress, such as DNA damaging agents etoposide and cisplatin, may be enhanced when eIF2 $\alpha$  kinases are also inhibited (Pataer et al., 2009), but that this approach may be counter-productive for more advanced tumours.

### **1.5.c Translation elongation and cancer**

Although the majority of work on the contribution of translational dysregulation in cancer has focussed on initiation, three strands of evidence implicate translation elongation.

First, as has already been described, eEF2 is negatively regulated by eEF2k. Inhibition of eEF2k under normocalaemic conditions is achieved by mTORC1 dependent (S6K) and mTORC1 independent (p90RSK) pathways that are frequently upregulated in cancer (see 1.2.c and 1.5.a). Second, eEF2 is overexpressed in gastric and colorectal cancers and increased expression in a gastric carcinoma cell line increases tumour volumes when injected into immunocompromised (nude) mice (Nakamura et al., 2009). The third line of evidence relates to the two isoforms of eEF1A (the tRNA binding subunit of eEF1). eEF1A1 is the predominant form in normal tissue, but eEF1A2 expression is highly increased in estrogen receptor positive breast cancers (Tomlinson et al., 2007) and overexpression in a pancreatic cancer cell line increases the growth of nude mice xenographs (Cao et al., 2009). However, overexpression of eEF1A2 in NIH 3T3 MEFs is not associated with a statistically significant increase in translation rates, suggesting that the oncogenic capacity of eEF1A2 is due to moonlighting functions (Janikiewicz, 2011).

### **1.6 Project aims**

As described within this chapter, translation is a highly complex process that is regulated (and consequently can be dysregulated) at many stages. The work in this thesis focuses on a number of these stages in a variety of different contexts. Chapters 3 and 4 explore the role of translation and cancer, focussing on the role of tRNA<sub>i</sub><sup>Met</sup> in the transformation of cultured cells and mTOR signalling in colorectal cancer respectively. Chapter 5 investigates the

contribution of translation elongation to the multi-faceted cold shock response. Finally, drawing on experiences of using the next-generation sequencing techniques within this thesis, Chapter 6 attempts to characterise the strengths and limitations of these approaches and propose improvements to sequencing protocols.

## **Chapter 2. Materials and methods**

### **2.1 Solutions common to multiple methods**

1X TBE: 89 mM Tris, 89 mM boric acid, 2.5 mM Na<sub>2</sub>EDTA, pH 8.3

1X TAE: 40 mM Tris-HCl, 2 mM Na<sub>2</sub>EDTA, pH 8.0

20X SSC: 3 M NaCl, 300 mM sodium citrate, pH 7.4

50X Denhardt's Solution: 1 % (w/v) Ficoll, 1 % (w/v) polyvinyl pyrrolidone, 1 % (w/v) BSA

Formamide loading dye: 96 % Formamide, 10 mM EDTA, 0.01 % bromophenol blue, 0.01 % xylene cyanol)

LB Broth: 1 % w/v Bactotryptone, 0.5 % w/v bacto yeast extract, 1 % w/v NaCl.

### **2.2 Cell culture**

#### **2.2.a Monolayer cultures**

##### **i) Routine cell culture**

NIH 3T3 mouse embryonic fibroblasts (MEFs), described in Marshall et al., 2008, were kindly provided by Prof Robert White. HeLa and HEK293 cells were from Willis Laboratory stocks, originally purchased from American Type Culture Collection (ATCC). Cells were cultured in gamma sterilized plasticware (TPP c/o Helena Biosciences) under a humidified atmosphere containing 5 % CO<sub>2</sub> at 37°C.

MEFs were cultured in Dulbecco's modified Eagle's medium (DMEM) without sodium pyruvate (Invitrogen) supplemented with 10 % v/v fetal calf serum (FCS; Helena Biosciences), 2mM L-Glutamine (Sigma), 100U/ml Penicillin, 100U/ml Streptomycin and, following transfection with pPUR vectors, 2.5 µg/ml Puromycin. HeLa and HEK293 cells were cultured in DMEM supplemented with 10 % FCS and 2mM L-glutamine.

For routine culture, cells grown to full confluence were washed with phosphate buffered saline (PBS) and treated with 1X Trypsin-0.5mM EDTA (PAA) for 5 min at 37°C. Once cells had detached, fresh medium was added to neutralise the trypsin and 10-40 % seeded into to a new flask.

Where a specific seeding density was required, a haemocytometer was used to count cell density from a neutralised sample stained with 0.4 % w/v trypan blue (Gibco) to exclude dead cells.

For plating cells 48 hours prior to assaying

Cell type	6cm dish	10cm dish	15cm dish
NIH 3T3 MEFs	n/a	$5 \times 10^5$	$1.2 \times 10^6$
HeLa	$3.2 \times 10^5$	n/a	$2 \times 10^6$

## ii) DNA/RNA transfection

The 2D MEF clone was transfected with 2 µg of puromycin resistance plasmid (pPUR) with or without the tRNA<sup>Met</sup> gene, in a 6-well plate format using Lipofectamine2000, following the manufacturer's instructions.

For transfection of [<sup>33</sup>P] labelled capped RNA (2.6.i), 4µg RNA was transfected into a ~90 % confluent 6cm dish of HeLa cells using Lipofectamine 2000. Cells were re-fed 90 min prior to transfection. 1 or 2 hrs after transfection, cells were treated as described in 2.6.hiv).

### 2.2.b Matrigel cultures

Matrigel cultures were derived and maintained by Dr Liam Faller, (Beatson Institute, Glasgow).

### **i) WT crypt culture**

Intestines were isolated from mice sacrificed under UK Home Office guidelines. The intestinal villi were removed by scraping and intestines chopped into small pieces prior to repeated washing with cold PBS (typically around 10 times). Crypts were detached by incubation in either 2mM (duodenum) or 25mM (colon) EDTA in PBS at 4°C for 30min. The intestine was broken by repeated pipetting. This was repeated 3 times giving 4 fractions. The first fraction was discarded as it was rarely enriched for crypts. Crypt fractions were combined with advanced DMEM/F12 (ADF) and pelleted. Pellets were resuspended in ADF and repeatedly passed through a cell strainer to separate crypts. Crypts were resuspended in 10ml ADF and the number of crypts in 20 µl of suspension determined. As much supernatant as possible was removed from the suspension and Matrigel added to give a final concentration of 2000 crypts/ml. 50 µl of Matrigel mix was added to each well of a pre warmed 24-well plate and incubated at 37°C, 5-10 min to solidify the gel. Crypts were cultured in 500 µl ADF in the presence of 100ng/ml EGF, 500ng/ml R-spondin and 100ng/ml Noggin. Growth factors were replenished every two days with media changed twice per week. Organoids were passaged by mechanically pipetting the Matrigel, diluted with ADF, and pelleted (600 rpm, 2 min) to separate crypts from Matrigel plug. The supernatant (with the Matrigel) was carefully removed and the washing process repeated to dilute Matrigel. The number of crypts was quantified, fresh Matrigel added and seeded as described above.

### **ii) Culture of adenomas**

Adenomas were removed from the small intestine and cut into small pieces. These were washed repeatedly then incubated in 5mM EDTA-PBS for 10 min, room temperature. The pieces were then washed 2-3 times with PBS and incubated at 37°C, 30 min, in 5 ml 10X trypsin, 1X DNase buffer and 100-200U DNaseI. Trypsin was neutralised with 5ml ADF and the tube vigorously shaken to dissociate cells. This was repeated until 40-45 ml ADF had been added. The cells were pelleted and counted and seeded as described in 2.2.bi) but at a

density of 100-200 cells per well. Routine culturing is as described in 2.2.bi) but without R-Spondin and growth factors supplemented once per week.

### 2.3 *in vivo* methods

All mouse work was performed by Dr Liam Faller (Beatson Institute, Glasgow) under UK Home Office guidelines.

#### 2.3.a Details of Cre Recombinase inducible mice

Outbred male mice from 6 to 12 weeks of age were used.

The following mice strains were used:

*S6K1/2*<sup>-/-</sup>: systemic deletion of both ribosomal S6 protein kinases (Pende et al., 2004)

*4EBP DKO*<sup>-/-</sup>: systemic knockout of both 4E-BP1 and 4E-BP2 (Le Bacquer et al., 2007)

*rpS6*<sup>p/-</sup>: systemic mutation of all 5 serine residues that are phosphorylated by S6K (Ruvinsky et al., 2005)

*eEF2k*<sup>-/-</sup>: systemic deletion of eEF2k (Dorovkov et al., 2002)

*Apc*<sup>580S</sup>: Exon 14 of APC flanked by LoxP sites for creation of inducible truncation mutation that removes APC function (Shibata et al., 1997)

*Raptor*<sup>fl/fl</sup>: Raptor gene flanked by LoxP sites for inducible deletion (Polak et al., 2008)

*Rictor*<sup>fl/fl</sup>: Rictor gene flanked by LoxP sites for inducible deletion (Cybulski et al., 2009)

*KRas*<sup>G12D/+</sup>: K-Ras G12D mutant allele preceded by transcription termination stop element flanked by LoxP sites allowing for inducible expression of the mutant allele (Jackson et al., 2001)

For conditional knockouts, mice were either crossed with a VillinCre<sup>ER</sup> strain (Andreu et al., 2005) or an Lgr5Cre strain (Sato et al., 2009). Villin is expressed along the crypt axis, Lgr5 is specific to intestinal stem cells and Paneth cells.

To induce recombination in VillinCre<sup>ER</sup> mice, animals were given two injections (IP) of tamoxifen (80 mg/kg) on consecutive days. For LGR5Cre<sup>ER</sup> mice, a single IP injection of 3mg tamoxifen was given. At the appropriate time point, mice were sacrificed and the small intestines were removed and flushed with water (at least three mice were used for each experiment), and epithelial cells prepared according to the applicable protocol.

### **2.3.b Rapamycin treatment**

Rapamycin was made up as a 50 mg/ml stock solution in ethanol and diluted in PBS + 5 % Tween 80, 5 % PEG 400. Mice were IP injected with 10 mg/kg rapamycin daily.

### **2.3.c Low dose cycloheximide treatment.**

Mice were treated with 50 mg/kg cycloheximide on days 2, 3, 4, and sacrificed 6 hrs after last dosing

### **2.3.d Survival studies**

Lgr5Cre strains were used for conditional deletion. Mice were induced and two cohorts assembled. The first started rapamycin treatment at day 10, for 30 days, to assess the ability of rapamycin to prevent tumour growth. The second began treatment when showing signs of illness (pale paws, weight loss, hunching), and were treated for 60 days to assess the ability of the drug to regress established tumours.

### **2.3.e Intestinal regeneration assay**

VillinCre strains were used for conditional deletion. Mice were exposed to gamma-irradiation 137Cs sources. This delivered gamma-irradiation at 0.423 Gy/min, to a total of 14 Gy per mouse. 72 hr after irradiation, mice were sacrificed, and intestines isolated and scored for

crypt regeneration. This scoring was carried out by counting the number of regenerating crypts in up to 10 cross-sections of the small intestine and taking an average.

#### **2.3.f BrdU incorporation in vivo**

250 µl BrdU was injected (IP) into each mouse 2 hours prior to sacrificing. IHC was used to visualise BrdU incorporation and proliferation was scored manually.

#### **2.3.g Murine intestinal crypt extraction for polysome profiling**

VillinCre strains were used for conditional deletion.

##### **i) Standard method**

Mice were sacrificed, intestines removed and flushed with PBS containing 100 µg/ml cycloheximide (PBS-CHX). Intestines were scraped briefly to remove villi then incubated with agitation for 15 min at 37°C in HBSS containing 100 µg/ml cycloheximide, 10 mM EDTA to dislodge the crypts. Cells were pelleted in 50 ml falcon tubes at 4000 rpm, 4°C, for 5 min, resuspended in 1 ml ice cold PBS-CHX, transferred to 1.5 ml eppendorf and pelleted again. The supernatant was aspirated and pellets snap frozen in liquid nitrogen prior to storage at -80°C.

##### **ii) Rapid extraction method**

The intestines were extracted as described above, however the time spent in warm media was reduced. Specifically: following incubation of intestine in warm media containing cycloheximide (100µg/ml) for 5 min the intestine was placed in ice cold PBS-CHX and vigorously shaken to dislodge the crypts. All other steps were as described in 2.3.gi)

### **2.4 Functional assays for cultured cells**

#### **2.4.a Cell proliferation assay**

$2 \times 10^4$  cells were seeded in 4 12-well plates, in 1 ml media. After 3 hours, one plate was processed as follows. The media was aspirated, cells washed with warm PBS, then fixed with

500  $\mu$ l 100 % methanol for 30 min. The methanol was aspirated, and the well covered with 500  $\mu$ l 1 % w/v methylene blue in borate buffer (10 mM  $\text{Na}_2\text{B}_4\text{O}_7$ , pH 8.5) for 30 min. Excess stain was removed by at least 4 washes of borate buffer, until no further blue colour could be eluted. Residual buffer was carefully aspirated and the dye eluted with 1 ml acidified ethanol (1:1 v/v 0.1M HCl and ethanol) for 1 hr with gentle rocking. The absorbance at 650nm of 200  $\mu$ l of eluate was read using a spectrophotometric plate reader.

This process was repeated for the other plates at 24, 48 and 72 hr. A growth curve was plotted using the mean of readings from 6 wells.

#### **2.4.b Protein synthesis assay**

##### **i) Adherent cells**

$1 \times 10^5$  cells were seeded in a 6-well format in 2 ml media. At roughly 80 % confluency (typically 2 days), media was removed and replaced with 500  $\mu$ l of methionine-free DMEM supplemented with 10 % dialysed FCS, 2 mM L-Glutamine and 3  $\mu$ l/ml media [ $^{35}\text{S}$ ]-Met label (PerkinElmer or Hartmann Analytic), and returned to the incubator. After 30 min the cells were placed on ice, washed with ice cold PBS containing 0.3 % (w/v) L-Methionine (PBS-Met), and lysed with 400  $\mu$ l 1X Passive lysis buffer (Promega).

Lysates were centrifuged for 2 min at 13,000 rpm, 4°C. 300  $\mu$ l of the supernatant was added to 300  $\mu$ l 25 % (w/v) TCA and left on ice for 30 min to precipitate. The remaining supernatant was transferred to a fresh eppendorf and quantified using Bradford reagent (Bio-Rad) following the manufacturer's instructions.

Vacuum filtration was used to apply precipitated protein to glass microfibre filters (Whatman), pre-soaked in 25 % TCA. Filters were washed twice with 70 % industrial methylated spirits, twice with acetone and left to dry.

Radiolabelled protein was liberated using Ecoscint scintillation fluid (National Diagnostics) and quantified in a Wallac 1409 scintillation counter (Pegasus Scientific). Counts were normalised to protein mass taken from the mean of three Bradford reactions.

## **ii) Matrigel cultures**

Matrigel cultures were washed with warm PBS then incubated with 1ml of ADF supplemented with 100ng/ml EGF, 500ng/ml R-spondin and 100ng/ml Noggin, and 3 µl/ml [<sup>35</sup>S]-Met label and returned to the incubator for 30 min. The cultures were placed on ice and washed with ice cold PBS-Met. Cultures were scraped into 1 ml of PBS-met and dispersed by 30 pipette strokes using a P200 tip. Cells were pelleted at 600 rpm and the supernatant and Matrigel “plug” carefully removed. This process was repeated 3 times to obtain a Matrigel-free cell pellet. Cells were lysed in 150 µl RIPA buffer (50 mM Tris-HCl pH 7.5, 150mM NaCl, 0.5 % sodium deoxycholate, 1 % Triton X-100, 0.1 % SDS) and cleared by centrifugation at 13,000 rpm, 4°C, 5 min. Subsequent steps were essentially as described in 2.4.bi) with quantification by BCA Assay (Pierce).

## **2.5 DNA methods**

### **2.5.a Phenol-chloroform extraction**

An equal volume of 1:1 Phenol:chloroform was added to DNA-containing solution and shaken vigorously. Following 2 min centrifugation at 13,000 rpm, 4°C, the upper aqueous layer was transferred to a new eppendorf tube and the above was repeated but with chloroform only. 1/10<sup>th</sup> volume of 3 M NaOAc pH 5.2 was added to the aqueous phase and mixed thoroughly. If the expected yield was less than 1 µg, 1 µl GlycoBlue™ (Ambion) was also added. Either 1 volume of isopropanol or 2.5 volumes of ethanol was added, mixed thoroughly and left to precipitate for a minimum of 30 min at -20°C.

DNA was pelleted by centrifugation at 13,000 rpm, 4°C, for 20 min. The pellet was washed with 75 % ethanol, pelleted, and the supernatant carefully removed. Following a brief (5-10min) air dry the pellet was resuspended in nuclease free water and, where necessary, quantified on a Nanodrop2000.

### 2.5.b Reverse transcription

5 µg of total RNA, extracted as described in 2.6.b was brought to 10 µl with nuclease free water. 1 µl of 300ng/µl random primers (Invitrogen), 1 µl of 5 µg/µl oligodT<sub>20</sub> and 1 µl 10mM dNTPs (Invitrogen) were added. The mixture was incubated at 65°C for 5 min then on ice.

4 µl 5X First strand buffer, 1 µl 0.1M DTT, 1 µl Supersaseln (Ambion) and 1 µl Superscript III (Invitrogen) were added, mixed and incubated at 50°C for 30 min. Reactions were inactivated at 70°C for 15 min. RNA was degraded either by using 2 units of RNaseH and incubating at 37°C for 20 min; or 0.1M NaOH, 0.05M EDTA and incubating at 90°C for 15 min, followed by neutralisation with 1M HEPES pH 7.5.

### 2.5.c Polymerase chain reaction (PCR)

Reactions were prepared as described in the table below. To reduce pipetting error and allow for optimisation of annealing temperatures, reactions were typically prepared as a 100 µl master mix, split between 4 or 5 tubes and run on a gradient thermocycler.

Component	20 µl reaction	100 µl mix
10X PCR buffer with Mg <sup>2+</sup> (Roche)	2	10.0
10mM dNTPs	0.4	2.0
10mM Forward primer	0.4	2.0
10mM Reverse primer	0.4	2.0
RT reaction mix	0.4	2.0
Nuclease free water	16.2	81.2

Taq DNA polymerase	0.2	0.8
--------------------	-----	-----

For fragments that were difficult to amplify, 5M betaine was added to a final concentration of 1M with the water adjusted accordingly to maintain the same final volume.

#### **2.5.d Agarose gel separation and extraction**

Agarose was dissolved in 1X TAE, with heating, to give an appropriate percentage solution to resolve DNA fragments of the size of interest (typically 1 % w/v). After allowing to cool a little, SybrSafe (Invitrogen) was added to 1X (10,000X stock). The molten agarose was poured into taped trays, combs inserted and left to set.

4 volumes of sample were mixed with 1 volume of 5X DNA gel loading dye (30 % glycerol, 0.1 % xylene cyanol, 0.1 % bromophenol blue). Tape was removed from the trays and the gel submerged in 1X TAE. The samples were loaded and the gel run at 120 V for 1-2 hrs, or until sufficient separation had been achieved.

Gels were visualised on a Safelmager and the gel fragment corresponding to the correct size excised. DNA was recovered either using the QIAquick gel extraction kit (Qiagen) following the manufacturer's instructions or using the glass wool method detailed below.

A large gauge needle was used to pierce a 0.5 ml eppendorf. A small quantity of glass wool was placed in the bottom of the tube and this was placed in a 1.5 ml eppendorf. The gel slice was placed in the smaller tube and centrifuged 13000 rpm for 10min (for larger gel slices it was necessary to centrifuge for longer and/or split the slice between multiple tubes). The eluate was phenol chloroform extracted as described in 2.5.a.

#### **2.5.e Restriction endonuclease digestion of plasmid DNA**

Plasmid DNA was digested by 10-20 units of restriction endonucleases from NEB in the appropriate buffer from the supplier. Reactions were typically in 20-50 µl with not more than

5 % glycerol (the enzymes are stored in 50 % glycerol). Reactions were incubated at 37°C for 1-2 h and heat inactivated 65°C for 10 min.

## **2.5.f DNA cloning**

### **i) DNA ligation**

For ligation into pGEM-T Easy vector (Promega), PCR products were used without further modification as *Taq* DNA polymerase leaves 3' A overhangs. For ligation of tRNA<sub>i</sub><sup>Met</sup> fragments into pPur vector, fragments were digested with *EcoRI*. The linearised pPur vector was dephosphorylated with 1U calf intestinal phosphatase (Roche) at 37°C for 1 hr prior to ligation.

For ligation into pGEM-T Easy vector, a 3:1 molar ratio of insert to vector was used. Typically 50 ng of vector was used in a 10 µl volume containing 1X Rapid ligation buffer (Promega) and 3U T4 DNA polymerase (Promega). Reactions were incubated at room temperature for 1 hr or 4°C overnight.

For ligation of tRNA<sub>i</sub><sup>Met</sup> fragments the ratio was 10:1 to encourage multiple inserts per vector.

### **ii) Transformation of competent E.Coli**

5 µl of ligation mix was added to 50 µl of competent *E.Coli* suitable for blue white screening (DH5α or TOP10) and kept on ice for 10-15 min. The mix was heat shocked at 43°C for exactly 1 min then placed immediately on ice for 10 min.

1 ml of warm LB was added to each mix and incubated at 37°C, 1 hr, with shaking. Cells were then pelleted, resuspended in 50 µl warm media and spread on LB-Agar plates with 100 µg/ml ampicillin and left to grow at 37°C overnight. For blue-white selection of T-Easy transformants the agar plates also contained 100 µM Isopropyl β-D-1-thiogalactopyranoside and 40 µg/ml X-Gal.

Colonies were picked with a sterile pipette tip and placed in 5-10 ml LB with 100 µg/ml ampicillin and incubated at 37°C with shaking, overnight.

### **iii) Miniprep extraction of plasmid DNA**

Cultures were pelleted at 4000 rpm for 10 min, and the DNA extracted with the Miniprep SV Wizard Kit (Promega) according to manufacturer's instructions.

## **2.6 RNA methods**

### **2.6.a Acid phenol RNA extraction**

An equal volume of acid-phenol pH 4.5 was added to RNA-containing solution and shaken vigorously. Following 2 min centrifugation at 13,000 rpm, 4°C, the upper aqueous layer was transferred to a new eppendorf tube and the above was repeated but with chloroform in place of acid-phenol. 1/10<sup>th</sup> volume of 3 M NaOAc pH 5.2 was added to the aqueous phase and mixed thoroughly. If the expected yield was less than 1 µg, 1 µl GlycoBlue™ was also added. 1 volume of isopropanol was added, mixed thoroughly and left to precipitate for a minimum of 30 min at -20°C.

RNA was pelleted by centrifugation at 13,000 rpm, 4°C, for 20 min. The pellet was washed with 75 % ethanol, pelleted, and the supernatant carefully removed. Following a brief (5-10min) air dry the pellet was resuspended in nuclease free water.

### **2.6.b Trizol RNA extraction**

1ml of Trizol Reagent (Sigma) was added to each well of a 6-well plate of PBS washed cells. RNA was extracted following the manufacturer's instructions. Pellets were typically resuspended in 10-20 µl nuclease free water and quantified on a nanodrop2000.

### **2.6.c Denaturing agarose gel electrophoresis**

1 % RNA gels were made by heating 1 g of agarose in 72.5 ml RNase free water and once dissolved, adding 10 ml of 10X MOPS (200 mM MOPS, 50 mM NaOAc, 10 mM EDTA, pH 7.0)

and 17.5 ml of formaldehyde. Gel plates and combs were cleaned with RNase ZAP (Sigma) prior to use.

Up to 5  $\mu$ l sample was mixed with 2  $\mu$ l of 10X MOPS, 3.5  $\mu$ l of formaldehyde, 10  $\mu$ l of deionised formamide, and 2  $\mu$ l of formamide loading dye, heated to 65°C for 15 minutes, then snap cooled on ice. Samples were loaded into the gel and run at 100 V for approximately 1 hour in 1X MOPS.

#### **2.6.d Passive transfer northern blotting for agarose gels**

Prior to transfer, RNA gels were incubated for 20 min at room temperature in 1X denaturing buffer (150 mM NaCl, 50 mM NaOH) followed by a further 20min in 1X neutralising buffer (150 mM NaCl, 100 mM Tris-HCl, pH 7.4).

RNA was transferred to Zeta-Probe nucleic acid blotting membrane (Bio-Rad) from the gel by capillary transfer for at least 16 hours. Briefly, a salt-bridge was created by making a wick from 3MM blotting paper soaked in 20X SSC as a transfer buffer. The gel was placed on top followed by the membrane, which had been washed first in nuclease-free water then in 20X SSC. 3 pieces of blotting paper were soaked in nuclease-free water and placed on top of the membrane. The stack was completed by covering with absorbent tissues and compressed with a 1 kg weight. RNA was fixed to the membrane by UV cross-linking. The cross-linked RNA was stained for 30 sec with methylene blue stain (0.02 % methylene blue, 0.3 M NaOAc pH 5.2), followed by repeated washing in water, to visualise the ribosomal RNA bands which were marked with a pencil. The stain was removed by agitating the blot in 1 % SDS/1X SSC for approximately 15 minutes.

## 2.6.e Denaturing Polyacrylamide gel electrophoresis (PAGE)

### i) Acid Urea PAGE

A 6.5 % Polyacrylamide 8 M Urea 0.1 M NaAc pH 5 gel mix was prepared as in the table below, with rotation but without heating.

Component	Final concentration	20cm gel	30cm gel
Urea (99 % purity; Sigma aldrich)	8 M	14.4 g	24 g
Accugel 40 % polyacrylamide (19:1 acrylamide/bisacrylamide; BioRad)	6.5 %	4.875 ml	8.125 ml
1 M NaOAc, filtered	0.1 M	3 ml	5 ml
TEMED	-	45 µl	75 µl
Nuclease-free water	-	10 ml	15 ml

Mixes were brought to pH5 with Glacial acetic acid, (approx 600 µl for 30 ml, 1 ml for 50 ml) and adjusted to 29 ml (20 cm gel) or 48.5 ml (30 cm gel) with nuclease free water.

Prior to assembly, gel plates were treated with RNase ZAP (Sigma) and coated with dimethyldichlorosilane. Plates were clamped between 0.75 mm spacers. 1 ml (20 cm gel) or 1.5 ml (30 cm gel) of 25 % APS was added to the gel mix and immediately cast and left to polymerise for 90 min.

2 µg total RNA (<5 µl) was combined with an equal volume of acid-urea loading buffer (8 M Urea, 0.1 M NaOAc pH 5, 0.05 % Xylene Cyanol, 0.05 % Bromophenol Blue) and kept on ice. Deacylated controls were obtained by combining 2 µg total RNA with an equal volume of 0.4 M Tris-HCl pH 9.5 and incubating at 37°C for 30 min then combining with an equal volume of loading buffer.

Electrophoresis was performed at 4°C, at 120 V/cm in 0.1 M NaOAc pH 5. Following a 30 min pre-electrophoresis step each well was flushed with running buffer to remove residual urea. Samples were immediately loaded and run until the bromophenol blue had reached the end of the gel (typically 18 hr for 30 cm gel).

## ii) TBE-Urea PAGE

TBE-Urea gel mixes were prepared as in the table below, with rotation but without heating.

Component	Final concentration	20cm gel
Urea (99 % purity; Sigma aldrich)	7 M	12.6 g
Accugel 40 % polyacrylamide (19:1 acrylamide/bisacrylamide; BioRad)	15 %	11.25 ml
	10 %	7.5 ml
10X TBE, filtered	0.1 M	3 ml
TEMED	-	15 µl
Nuclease-free water	-	5.5 ml

Mixes were adjusted to 30 ml with nuclease –free water and 150 µl 25 % APS added. The mix was cast into 20cm long plates prepared as described in 2.6.ei) and left to polymerise for approximately 1 hr.

RNA samples (<5 µl) were combined with an equal volume of TBE-Urea loading buffer (7 M Urea, 1x TBE, 12 % Ficoll 400, 0.05 % Xylene Cyanol, 0.05 % Bromophenol Blue) and kept on ice. Just before loading to the gel, samples were incubated at 75°C for 2 min then placed on ice for at least 2 min.

Electrophoresis was performed at room temperature, at 21W in 1X TBE. Following a 30 min pre-electrophoresis step each well was flushed with running buffer to remove residual urea. Samples were loaded immediately and run until the bromophenol blue had reached the end of the gel (typically 60-90 min for 20 cm gel).

## **2.6.f Northern blotting of PAGE gels**

### **i) Wet Electrophoretic transfer northern blotting for Acid Urea PAGE gels**

The region of gel between the two dyes was cut and placed on chromatography paper, wetted in 1X TAE. N+ Hybond Membrane (Amersham) was washed first in nuclease free water, then 1X TAE and placed over the gel. The membrane was covered with a final sheet of chromatography paper to complete the sandwich and air bubbles removed with a serological pipette. The sandwich was loaded into TransBlot® Cell (BioRad) containing 1X TAE.

Electrophoresis was performed at 25 V for 90 min at 4°C. RNA was fixed to the membrane by baking for 1-2 hr at 80°C.

### **ii) Semi-dry transfer northern blotting for TBE Urea PAGE gels**

Transfer was performed as described in 2.6.fi) but with 1X TBE in place of 1X TAE, Trans-Blot® SD Semi-Dry Transfer Cell (BioRad) in place of the wet transfer apparatus and electrophoresis performed at 300 mA for 90 min at 4°C.

## **2.6.g Probing northern blots with radiolabelled probes**

### **i) Probing northern blots with [<sup>32</sup>P]-α-dCTP radiolabelled DNA**

~60 ng of template DNA in 15µl of nuclease-free water was denatured by heating to 95°C for 2min and then snap cooled on ice. 5 µl of 5X labelling buffer (Promega), 1 µl 10 mg/ml bovine serum albumin (BSA), 0.5 µl 25 mM dNTP (containing only dATP, dTTP and dGTP), 1 µl Klenow polymerase and 2.5 µl of [<sup>32</sup>P]-α-dCTP (400 Ci/mol, 10 mCi/ml; Hartmann Analytic) were added to the template DNA and the mixture was incubated at 37°C for 1 hr. The product was then filtered through a Sephadex G-50 size exclusion column (GE life Sciences) and the flow-through, containing the purified probe was denatured at 95°C for 3 min and then snap cooled on ice.

Northern blots were pre-hybridised by rotation in a rotisserie oven at 65°C in 10 ml of Church Gilbert solution for at least 30 min. The cooled denatured probe was then added directly to the Church Gilbert solution (140 mM Na<sub>2</sub>HPO<sub>4</sub>, 70 mM NaH<sub>2</sub>PO<sub>4</sub>, 7 % SDS) and the blots were hybridised by overnight rotation at 65°C. Northern blots were then washed by rotation for 15 min at room temperature, twice in each of the 3 wash solutions (northern wash 1: 0.1 % SDS, 2X SSC; northern wash 2: 0.1 % SDS, 0.5X SSC; northern wash 3: 0.1 % SDS, 0.1X SSC). Blots were wrapped in cling film and exposed to an imaging plate (Fujifilm) overnight before being visualised on a molecular imager (Bio-Rad).

**ii) Probing northern blots with [<sup>32</sup>P]-γ-ATP radiolabelled oligonucleotides**

Probe target	Sequence
tRNA <sub>i</sub> <sup>Met</sup>	5' – cgcgtcagtctcataatctgaag – 3'
tRNA <sub>e</sub> <sup>Met</sup>	5' – cccataaccagaggtc gatgg – 3'
eEF1a1	5' – agtctctgtgtcctggggcatc – 3'

50-100 pmol oligonucleotide was incubated at 37°C for 30 min in a 10 µl reaction containing 1 µl 10X PNK buffer, 2.5 µl [<sup>32</sup>P]-γ-ATP (400 Ci/mol, 10 mCi/ml; Hartmann Analytic), 1 µl PNK (NEB). Excess radiolabel was removed using Sephadex G-25 size exclusion column (GE life Sciences).

Membranes were incubated at 42°C for a minimum of 3 hours in prehybridisation solution (6X SSC, 10X Denhardt's Solution, 0.5 % SDS). Prehybridisation solution was replaced with 15ml hybridisation solution (6X SSC, 0.1 % SDS), probe added and incubated overnight at 42°C.

Blots were washed sequentially 6X SSC, 4X SSC and 2X SSC. Each wash was performed twice for 10 min, rotating at room temperature. Washed membranes were wrapped in cling film and exposed to an imaging plate (Fujifilm) overnight.

## **2.6.h Polysome profiling**

### **i) Preparation of sucrose gradients**

Sucrose solutions were made by dissolving the following concentrations of sucrose – 50 %, 40 %, 30 %, 20 %, 10 % – in 1X sucrose gradient buffer (300 mM NaCl, 15 mM MgCl<sub>2</sub>, 15 mM Tris-Cl pH 7.5, 100 µg/ml cycloheximide, 1 mg/ml heparin). 2ml of each concentration was carefully pipetted into 12ml Sorvall polyallomer centrifuge tubes, with the gradients frozen at -80°C between the addition of each layer. Completed gradients were stored at -20°C and equilibrated at 4°C for at least 8 hrs prior to use.

### **ii) Preparation of adherent cells for polysome profiling, scrape method**

Cells were seeded as outlined in 2.2.a. The growth medium was replenished 90 min prior to harvest. 5 min prior to harvest cells were incubated with 100 µg/ml cycloheximide (CHX). Cells were placed immediately on ice, washed with ice-cold PBS-CHX, then scraped into 10 ml of PBS-CHX. Cell suspensions were gently pelleted (1400 rpm, 5 min, 4°C) and resuspended in 1 ml PBS-CHX for transfer to 1.5 ml eppendorf tubes. 100 µl was transferred to separate tubes. All tubes were centrifuged at 3000 rpm for 5 min at 4°C, and the supernatant removed. Tubes containing the small (100 µl) pellet were stored at -20°C for subsequent western blot analysis, the larger pellets were lysed in 1 ml 1X polysome lysis buffer (1X sucrose gradient buffer, 1 % Triton X100) and centrifuged at 13,000 rpm, 4°C, for 1 min. The supernatants were carefully layered onto pre-prepared sucrose density gradients. Centrifugation was at 38,000 rpm for 2 hr at 4°C in a Beckman Coulter ultracentrifuge.

### **iii) Harringtonine run-off treatment**

A stock of 1000X harringtonine (2 mg/ml in DMSO) was added to cells or matrigel cultures to a 1X concentration and the plates returned to the incubator. At set time periods after addition of harringtonine (typically, 60 s, 120 s, 180 s etc.) cycloheximide was added to a 1X concentration and the cells were prepared as described in 2.6.hii). Crypt cultures were scraped into ice cold PBS-CHX and dispersed by 30 pipette strokes using a P200 tip. Cells were pelleted at 600 rpm and the supernatant and matrigel “plug” carefully removed. This process was repeated twice to obtain a matrigel-free cell pellet. This was lysed in the standard polysome lysis buffer.

### **iv) Preparation of adherent cells for polysome profiling, direct lysis method**

Cells were cultured as described above, but residual liquid was vacuum aspirated after PBS-CHX wash and cells scraped into an appropriate volume of ice cold 1X lysis buffer (1 ml for 15 cm dish; 400 µl for 10 cm dish; 200 µl for 6 cm dish). Lysates were cleared by centrifugation at 13,000 rpm at 4°C for 3 min and centrifuged through sucrose gradient as described above.

### **v) Lysis of pelleted colonic crypt epithelia**

1ml ice cold lysis buffer was added to crypt cell pellets obtained as described in 2.3.g. The pellets were disrupted by gentle pipetting and the lysate cleared by centrifugation at 13,000 rpm at 4°C for 3 min. In later experiments using the rapid extraction method, the lysis buffer was supplemented with 500U/ml RNasin (Promega).

### **vi) Fractionation and RNA extraction from sucrose gradients**

The centrifuged sucrose density gradients were analysed by continuous flow through a type 12 optical unit, which read the OD at 254 nm. 65 % blue sucrose solution was used to force the gradient through the reader and the absorbance spectrum was detected and recorded by an UA-6 UV/VIS detector (Presearch Ltd., Hitchin, Herts). 1 ml gradient fractions were

collected and dispensed directly into 3 ml of 7.7 M guanidine hydrochloride. 4 ml of absolute ethanol was added to each fraction, mixed by inversion and precipitated overnight at -20°C.

The gradient fractions were centrifuged at 4,000 rpm for 45 min at 4°C. The supernatant was aspirated and the RNA pellets re-suspended in 360 µl of RNase free water. 40 µl of 3 M NaOAc, pH 5.0 and 1 ml of absolute ethanol were added to the re-suspended RNA. This was incubated at -20°C for at least 2 hr and centrifuged at 13,000 rpm for 30 minutes at 4°C. The supernatant was removed and the RNA pellets were washed in 75 % ethanol by vortexing followed by 5 min on a shaker. The samples were then centrifuged for 10 min at maximum speed at 4°C. Ethanol was removed and the pellets were air dried before being re-suspended in either 25 or 50 µl nuclease-free water and stored at -80°C.

## **2.6.i Ribosome profiling**

### **i) Preparation of sucrose gradients**

Sucrose solutions were made as described in 2.6.hi) with 1X sucrose gradient buffer replaced with 1X RPF buffer (300mM NaCl, 15 mM MgCl<sub>2</sub>, 15mM Tris-Cl pH 7.5, 100 µg/ml cycloheximide, 2 mM DTT).

### **ii) Preparation of cultured cells for ribosome profiling**

Cells were prepared as described in 2.6.hii) and 2.6.hiv) with 1X lysis buffer replaced with 1X RPF lysis buffer (1X RPF buffer, 1 % TritonX 100, 500 U/ml RNasin). RNaseI (Ambion) was added to the lysates at a final concentration of 1 U/µl and incubated at room temperature, with rotation for 30 min. Lysates were placed briefly on ice then loaded onto prepared sucrose gradients and centrifuged as described in 2.6.hii).

### **iii) Fractionation and RNA extraction from sucrose gradients**

The centrifuged sucrose density gradients were analysed by continuous flow through a type 12 optical unit, which read the OD at 254 nm as previously described (2.6.hvi). 500 µl

gradient fractions were collected in 2 ml DNA/RNA LoBind tubes (Eppendorf) and stored on ice until all gradients were fractionated. 1.4 ml Trizol LS (Sigma) was added to each gradient fraction, mixed and left at room temperature for 5 min to dissociate ribosomes from mRNA and any other nucleoprotein complexes. 200 µl BCP was added to each sample and mixed by vigorous shaking for 15 sec. Samples were left to stand for 5-10 min at room temperature before being centrifuged at 13,000 rpm at 4°C for 15 min. The upper, aqueous, layer was carefully removed, mixed with 1 µl GlycoBlue and 1 ml isopropanol and left at room temperature for 15min. Precipitated RNA was pelleted by centrifugation at 13,000 rpm at 4°C for 20 min. Pellets were washed with 75 % ethanol and centrifuged at 13,000 rpm at 4°C for 10 min. Following careful removal of the supernatant, pellets were air dried and resuspended in 10 µl nuclease free water.

#### **iv) RPF sequencing library generation**

The fractions corresponding to the monosome peak were size selected using either a 10 % or 15 % precast Novex® TBE-Urea gel (Invitrogen) gel. 5 µl of each fraction were mixed with 5 µl TBE-Urea loading dye and incubated at 75°C for 3 min then placed on ice and loaded within 10 min of denaturing. 5 µl of marker mix (either 27mer and 33mer or 28mer and 34mer, 500 nM) was prepared in the same way.

Marker name	Sequence
27mer	AAGCGUGUACUCCGAAGAGGAUCCAAA
33mer	AGGCAUUAACGCGAACUCGGCCUACAAUAGUGA
28mer	AGCGUGUACUCCGAAGAGGAUCCAACGU
34mer	GGCAUUAACGCGAACUCGGCCUACAAUAGUGACGU

Wells were flushed with loading buffer, samples loaded and electrophoresis performed at room temperature, 180 V, until the bromophenol blue marker ran off the bottom of the gel (typically 45-60 min dependent on gel percentage).

The gel was incubated at room temperature for 5-10 min in 1X TBE 1X SybrGold (Invitrogen) then visualised using the Safelmager (Invitrogen). A smear was observed in the lanes corresponding to fractions extracted from the gradient as not all RNA is digested to ~30nt fragments. The region between the markers was excised and eluted from the gel slices by incubating in 400  $\mu$ l 300 mM NaCl, 0.1 % SDS, 1 mM EDTA overnight at 4°C, with rotation. The eluate was transferred to a new tube, 1  $\mu$ l GlycoBlue added and the RNA was precipitated by addition of 1 ml of ethanol and incubation for at least 2 hr at -20°C. RNA was pelleted, washed with 75 % ethanol and resuspended in 10  $\mu$ l nuclease-free water.

To modify the ends of the rpf fragments PNK was used in two separate buffers: the first to dephosphorylate the 3' end, the second to phosphorylate the 5' end. Briefly, 9  $\mu$ l sample was mixed with 10  $\mu$ l 2X MES buffer (200 mM MES-NaOH pH 5.5, 600 mM NaCl, 20 mM  $\beta$ -mercaptoethanol, 20 mM MgCl<sub>2</sub>), 1  $\mu$ l PNK and incubated at 37°C for 6 hr. RNA was recovered using Purelink RNA micro kit (Invitrogen) following the manufacturer's instructions with elution with 12  $\mu$ l nuclease-free water. 1  $\mu$ l eluate was stored for analysis with the remaining 10-11  $\mu$ l 5'-phosphorylated by addition of 1.5  $\mu$ l 10X PNK buffer, 1.5  $\mu$ l 10mM ATP, 1  $\mu$ l PNK and incubation at 37°C for 30 min. RNA was recovered using Purelink RNA micro kit and eluted with 10  $\mu$ l nuclease-free water.

Unmodified, 3'-modified and fully modified RNA was analysed on 2100 Bioanalyzer (Agilent) using a small RNA chip to confirm fragments had not been partially degraded during the phosphorylation process and to provide an estimate of concentration.

10 ng rpf fragments were subjected to SOLiD™ Small RNA library (Applied Biosystems) construction according to manufacturer's instructions, with the maximum recommended 18 cycles of in-gel PCR.

#### **v) Whole transcriptome sequencing library generation**

Total RNA was extracted from pelleted cells as described in 2.6.b. 50 µg total RNA was subjected to two rounds of polyA+ selection using Dynabeads (Invitrogen) according to manufacturer's instructions. 500 ng polyA+ RNA was subjected to SOLiD™ whole transcriptome library (Applied Biosystems) construction according to manufacturer's instructions.

#### **2.6.j Bioinformatic analysis of ribosome profiling data**

Analysis was performed by Dr Ruth V. Spriggs, using pipelines supplied by Applied Biosystems and custom scripts as described below.

##### **i) Mappings**

Colour space reads were mapped using BioScope (Applied Biosystems) with default parameters to the mouse reference genome (mm9, NCBI Build 37) or human reference genome (hg19 GRCh37) as appropriate. Reads were also mapped to a filter reference file supplied by Applied Biosystems for human and created using rRNA sequences from the NCBI Nucleotide database and tRNA sequences from the genomic tRNA database (<http://gtrnadb.ucsc.edu/>) for mouse and human genomes. Where reads mapped to multiple genomic locations, the location determined "best" by the BioScope software was used.

##### **ii) Alignment of fragments relative to CDS**

The nucleotide mappings were converted into SAM format using SAMtools ([samtools.sourceforge.net](http://samtools.sourceforge.net)) and mapped to refGene annotations to determine the position of the 5' end of each mapped read with relation to the start of the coding sequence. The same

was performed relative to the end of the coding sequence codon. The number of counts at each position was obtained and expressed as a ratio relative to the total number of aligned reads (reads per million, rpm).

### **iii) Determination of minimum read thresholds for inclusion of gene in hypothesis testing**

Genes were ranked based on the total number of reads mapping to that gene across both replicates of a given library type. This list was divided into bins of 1000 genes. For each gene the fraction of reads that came from replicate1 (ie rep1/total) was calculated and the standard deviation of these fractions within each bin calculated. The threshold was the start of the bin at which the standard deviation began to plateau.

### **iv) Calculation of translational efficiency**

For each gene that passed the minimum reads threshold in all 4 conditions (whole transcriptome and RPF libraries for both conditions), the translational efficiency value (TE) was calculated as follows: rpm in translatome library/rpm in transcriptome library

### **2.6.k Generating RNA dimers**

300 pmoles of 28mer and 34mer RNA oligonucleotides were 5' phosphorylated with 10 U PNK in a 50 µl reaction containing 1X PNK buffer and 1 mM ATP incubated at 37°C for 30 min. RNA was recovered as described in 2.6.a. To minimise the circularisation of the 5' phosphorylated RNA by the ligase, a solution with a 20:1 ratio of 5' hydroxylated:5' phosphorylated forms of each oligonucleotides was made. This was subjected to ligation with 5U T4 RNA ligase 1 in 1X T4 Ligase buffer (NEB), 1 mM ATP, in a final reaction volume of 10 µl. After incubation at 37°C for 15 min, 10 µl RNA loading dye was added and the RNA denatured at 80°C for 3 min. The RNA was separated on a 15 % TBE Urea gel and the dimer fragment recovered as described in 2.6.iiv).

## 2.6.1 Generation of capped, radiolabelled, $\beta$ -actin mRNA

### i) Template preparation

Full length  $\beta$ -actin cDNA was amplified as described in 2.5.b and 2.5.c using the following primers and cycling conditions:

Forward primer: 5' – accgccgagaccgcgt – 3'

Reverse primer: 5' – tttattcaactggtctcaagtcagtgtac – 3'

$98^{\circ}\text{C}^{2\text{min}} + 30*(98^{\circ}\text{C}^{20\text{sec}} + 54.7^{\circ}\text{C}^{20\text{sec}} + 70^{\circ}\text{C}^{30\text{sec}}) + 70^{\circ}\text{C}^{10\text{min}}$

The PCR product was ligated into T-Easy vector, transformed and 7 white colonies selected to be grown on. Each colony was digested with XmaI and XmnI to determine the orientation of the fragment relative to the T7 promoter. A plasmid with the insert in the sense orientation relative to T7 was transformed and grown in 100 ml LB with the plasmid extracted using the Pure Link Maxiprep kit (Invitrogen).

20  $\mu\text{g}$  plasmid was digested with SpeI (i.e. just downstream of the end of the 3'UTR). A small aliquot was removed and separated on an agarose gel as described in 2.5.d to confirm digestion was complete. The digested sample was treated with 100  $\mu\text{g}/\text{ml}$  Proteinase K and 0.1 % SDS at  $50^{\circ}\text{C}$  for 30 min to digest SpeI and residual RNase A from the maxiprep. The sample was then phenol-chloroform extracted, ethanol precipitated (2.5.a) and resuspended in 10  $\mu\text{l}$  nuclease free water.

### ii) *in vitro* transcription of moderate specific activity capped RNA

The following reaction was assembled at room temperature in the following order: 4  $\mu\text{l}$  5X transcription buffer (Fermentas), 2  $\mu\text{l}$  mix of NTPs each at 5 mM, 8  $\mu\text{l}$  [ $^{33}\text{P}$ ]- $\alpha$ -UTP (3000 Ci/mmol, 30 mCi/ml; Hartmann Analytic), 1  $\mu\text{l}$  40 mM Anti-reverse cap analogue (Ambion), 0.5  $\mu\text{l}$  0.1U/ $\mu\text{l}$  Inorganic pyrophosphatase (Fermentas) and 1 $\mu\text{l}$  Supersasin. 1  $\mu\text{l}$  of linearised plasmid (1  $\mu\text{g}/\mu\text{l}$ ) and 1.5  $\mu\text{l}$  of T7 RNA polymerase were added and the reaction incubated at

37°C for 2 hr. Template was degraded with DNaseI at 37°C, 15 min.

To this reaction was added 43 µl nuclease free water, 25 µl 4X PAP buffer (200mM Tris pH 7.9, 1M NaCl, 40mM MgCl<sub>2</sub>) and 10 µl 10mM ATP. 2.5 µl was removed to a new eppendorf to serve as a no enzyme control, 4 µl E-PAP (Ambion) was added to the main reaction and incubated at 37°C for 45min. 2.5 µl was removed and the two 2.5µl aliquots were denatured with 7.5 µl formaldehyde loading dye (Ambion) at 75°C for 10 min and separated on a denaturing agarose gel as described in 2.6.c. RNA was visualised by soaking the gel in 1X MOPS 1X SybrGold for 15 min in the dark.

The remainder of the polyadenylated message was recovered by acid phenol chloroform extraction and isopropanol precipitation.

## 2.6.m RNase protection assay

### i) Generation of high specific activity probe

Fragments corresponding to the CDS and 3'UTR of homo sapiens *gapdh* were generated by PCR amplification of cDNA from HEK293 cells using *Taq* polymerase as outlined in 2.5.c, using primers and cycling conditions detailed below.

Target	Forward Primer 5'-3'	Reverse Primer 5'-3'	Product size bp
Gapdh CDS	gtcaaggctgagaacgggaa	tgatgacccttttggtccc	184
Cycling conditions : 98°C <sup>2min</sup> + 25*(98°C <sup>20sec</sup> + 55.3°C <sup>30sec</sup> + 70°C <sup>30sec</sup> ) +70°C <sup>10min</sup>			
Gapdh 3' UTR	Ccccagcaagagcacaaga	tggttgagcacagggtacttt	184
Cycling conditions : 98°C <sup>2min</sup> + 25*(98°C <sup>20sec</sup> + 55.3°C <sup>30sec</sup> + 70°C <sup>30sec</sup> ) +70°C <sup>10min</sup>			

Fragments were gel purified as described in 2.5.d and ligated in to pGEM-T Easy vector (Promega) as described in 2.5.f.

The vector was linearised as described in 2.6.li) with *SpeI* and recovered by phenol chloroform extraction and ethanol precipitation.

The linearised plasmid was used as a template for *in vitro* transcription. The following reaction was assembled at room temperature in the following order: 4 µl 5X transcription buffer (Fermentas), 2 µl mix of ATP GTP UTP each at 5mM, 2.4 µl 100µM CTP, 2.5 µl [<sup>32</sup>P]-α-CTP (400 Ci/mmol, 10 mCi/ml; Hartmann Analytic), 1µl RNasin and 6.1 µl nuclease free water. 1 µl of linearised plasmid (1 µg/µl) and 1 µl of T7 RNA polymerase were added and the reaction incubated at 37°C for 2 hr. Template was degraded with DNaseI at 37°C for 15 min and the probe recovered by acid phenol chloroform extraction and isopropanol precipitation.

The probe was resuspended in 5 µl nuclease free water, 5 µl RNA loading dye added and the sample denatured at 80°C for 3 min before being run on a 6 % TBE-Urea gel at 180 V until the bromophenol blue dye front ran off the bottom. The gel was wrapped in saran wrap and exposed to X-Ray film for approximately 10 min. The film was developed and then overlaid onto the saran wrap to guide the excision of the correct sized band from the gel, thus excluding undersized transcription products. The probe was recovered from the gel slice as described in 2.6.iv) and resuspended in 20 µl nuclease free water. 1 µl was added to 2ml EcoScint to quantify the specific activity of the probe.

## **ii) Generation of radiolabelled RNA markers**

The small RNA fragments (28mer, 34mer and their respective dimers) were labelled with [<sup>32</sup>P]-γ-ATP as described in 2.6.gii). To label the 0.1-2kb RNA ladder, 4 µg of ladder was dephosphorylated with 2U Calf Intestinal Phosphatase (Roche) in a 20 µl volume containing 1X dephosphorylation buffer (ThermoScientific) at 37°C for 1 hr. RNA was recovered by acid phenol extraction and isopropanol precipitation (2.6.a). The pellet was resuspended in 8 µl nuclease free water and labelled with [<sup>32</sup>P]-γ-ATP.

### **iii) RNase protection**

5  $\mu$ l of RNA containing solutions (typically  $\frac{1}{2}$  of the RNA content of a fraction from an rpf gradient or 2  $\mu$ g yeast tRNA to act as a negative and positive controls) were added to a 15  $\mu$ l mixture containing 10  $\mu$ l formamide, 2  $\mu$ l 10X hybridisation salts (4 M NaCl, 400 mM PIPES pH 6.8, 20 mM EDTA) and  $2.5 \times 10^4$  dpm of probe. To ensure consistency the mixture was made up as a master mix, with 5 % excess to account for pipetting error.

The 20  $\mu$ l reactions were incubated in a thermocycler at 98°C for 3min then 42°C, overnight to permit annealing of the probe.

200  $\mu$ l RNase digestion buffer (10 mM Tris, 300 mM NaCl, 5 mM EDTA, pH 7.5) and 4.4  $\mu$ l of 50X RNaseA/T1 stock (250  $\mu$ g/ml RNase A, 5000U/ $\mu$ l RNase T1 in storage buffer of 50 mM Tris pH 7.5 50 % glycerol) were added to all samples other than the undigested control, and incubated 37°C, 45 min.

RNA was precipitated by addition of 1  $\mu$ l glycoblue and 660  $\mu$ l precipitation buffer (2M Guanidine-HCl [in later experiments this was replaced with 1M Guanidium isothiocyanate], 66.6 % isopropanol, 0.134 % SDS, 13.4 mM DTT), mixed thoroughly and incubated for at least 45 min at -20°C. RNA was recovered by centrifugation (13,000 rpm at 4°C for 30 min) washed twice with 75 % ethanol and resuspended in 5  $\mu$ l nuclease free water and 5 $\mu$ l RNA loading dye.

### **iv) Visualisation**

Samples were denatured at 80°C for 3 min, placed on ice, then separated on a 10 % TBE Urea gel as described in 2.6.eii). Only  $\frac{1}{20^{\text{th}}}$  of the volume of the undigested probe samples were loaded.  $1 \times 10^4$  dpm of 34 and 28mer markers and 0.1-2kb ladders were used to determine the sizes of bands.. The gels were fixed in 40 % v/v Methanol 5 % v/v Glacial acetic acid for

20 min then vacuum dried, 65°C for 1 hr then 80°C for 3 hr. Dried gels were exposed to a phosphorimager screen for 24 hr and visualised on a phosphorimager.

## **2.6.n Alternate ligation methods for deep sequencing**

### **i) Selection of partially degenerate 20 nt RNA pool**

A random number generator (<http://www.random.org>) was programmed to output values between 1 and 4 inclusive. Each base was assigned a number: 1 = A, 2 = C, 3 = G, 4 = U. The program was run 10 times and the output converted giving the following sequence:

GCAGUUGCCA. An RNA pool with this sequence followed by 10 degenerate nucleotides (5' – GCAGUUGCCANNNNNNNNNN – 3') was synthesised (ThermoScientific Molecular Biology).

This RNA pool has both 5' phosphate and 3'OH moieties necessary for ligation.

### **ii) Ligation of Universal Cloning linker**

40 ng RNA pool (~6 picomoles) was added to 1µl of 100 ng/µl Universal Cloning Linker (~18 picomoles; NEB) and brought to 10 µl with nuclease-free water. RNA was denatured at 80°C for 2 min then placed immediately on ice. For ligation with T4 RNA Ligase 2 truncated K227Q mutant (Rnl2tr K227Q) the following were added: 2 µl RNA ligase buffer (NEB), 1 µl SuperaseIn, 6 µl 50 % PEG8000 (NEB) and 1 µl Rnl2tr K227Q (NEB), mixed thoroughly and incubated at 25°C for 3 hr.

For ligation with *Methanobacterium thermoautotrophicum* RNA ligase K97A mutant (Mth K97A) the following were added: 2 µl NEB1 buffer, 1µl SuperaseIn, 6µl nuclease free water and 1 µl Mth K97A (NEB), mixed thoroughly and incubated at 65°C for 1 hr.

Reactions were terminated by heat inactivation at 90°C for 2 min, then placed on ice. RNA was precipitated by adding 205 µl nuclease free water, 25µl 3M NaOAc pH 5.2, 1 µl GlycoBlue and 250µl isopropanol, mixing thoroughly and incubating at -20°C for at least 30

min. RNA was pelleted at 13,000 rpm 30min, 4°C, washed twice with 75 % ethanol and left to air dry before being resuspended in 5 µl of nuclease free water.

5 µl of formamide loading dye was added to each sample, denatured at 80°C for 3 min then snap cooled on ice. 250 ng small RNA marker (Abnova) was prepared in the same way. Samples were run on a precast 15 % TBE-Urea gel (Invitrogen) in 1X TBE, at 180 V until the bromophenol blue dye front was at the bottom of the gel.

RNA was visualised with SybrGold and a band corresponding to 37 nt was excised and recovered as described in 2.6.iv).

### **iii) Reverse transcription with Ion Torrent compatible primer**

2 µl of 1.25 µM Ion Torrent compatible primer

(5'

CGCCTTG GCC/Sp/CACTCA/Sp/CCTCTCTATGGGCAGTCGGTGATATCTATTGATGGTGCCTACAG –

3' where Sp is an 18-atom hexa-ethyleneglycol spacer) was added to each sample, denatured at 80°C for 3min then snap cooled on ice. To this was added: 4 µl 5X First strand buffer, 1 µl 0.1M DTT, 1 µl SuperaseIn, 1 µl 10mM dNTPs and 1 µl Superscript III, mixed thoroughly and incubated at 48°C for 30 min, with heated lid. 2.5 µl each of 1N NaOH and 0.5 M EDTA were added, mixed thorough, and incubated at 98°C for 20 min to hydrolyse RNA. cDNA was precipitated as in 2.6.nii), denatured in formamide loading buffer at 98°C for 3 min, and run on a 6 % TBE-Urea gel, 180V, until the bromophenol blue had run off the bottom. Upon visualising with SybrGold, three bands were seen. The upper band (~80bp) was excised, being careful to avoid the lower two bands which were unincorporated RT primer and RT primer dimers. cDNA was recovered as described in 2.6.iv) though without SDS in the elution buffer. cDNA was resuspended in 15 µl nuclease free water.

#### **iv) Circularisation of cDNA**

2 µl Circligase buffer, 1 µl 50mM MnCl<sub>2</sub>, 1 µl 10mM ATP and 1 µl Circligase (Epicentre) were added to cDNA, mixed thoroughly and incubated at 60°C for 1 hr, with heated lid. Reactions were terminated by heat inactivation, 80°C, 10 min then precipitated as described in 2.6.nii), before being resuspended in 12 µl nuclease free water.

#### **2.6.o Sequencing using the Ion Torrent**

cDNA libraries were amplified and subject to emulsion PCR as described in the Ion Total RNA-Seq Kit v2 (Life Technologies). Templated beads were deposited on the 318 Chips and sequenced.

### **2.7 Protein methods**

#### **2.7.a Immunoblotting**

##### **i) Solutions**

Whole cell lysis Buffer: 10 mM Tris pH 7.5, 50 mM NaCl, 0.5 % NP40, 0.5 % Sodium Deoxycholate, 0.5 % SDS, 10 mM iodoacetamide.

Phospho-preserving lysis buffer: 10 mM Tris pH 7.5, 50 mM NaCl, 0.5 % NP40, 0.5 % Sodium Deoxycholate, 0.5 % SDS, 10 mM iodoacetamide, 0.5M NaF

2X SDS sample buffer: 100 mM Tris pH 6.8, 20 % glycerol, 8 % SDS, 0.2 % bromophenol blue, 20 % β-mercaptoethanol, 2 mM EDTA

4X SDS-PAGE resolving buffer: 1.5 M Tris, 1 % SDS pH 8.8.

4X SDS-PAGE stacking buffer: 0.25 M Tris, 0.2 % SDS pH 6.8.

10X SDS running buffer: 250 mM Tris, 1.92 M glycine, 1 % SDS pH 8.3.

Transfer buffer: 50 mM Tris, 192 mM glycine, 20 % methanol.

Ponceau Stain: 0.5 % w/v Ponceau S in 5 % w/v TCA.

TBST (Tris buffered saline, Tween): 10 mM Tris pH 8.0, 0.9 % NaCl, 0.1 % Tween

Luminol solution: 50mg Luminol (5-amino-2,3-dihydro-1,4-phyhalazinedione) in 0.1 M Tris-HCl pH 8.6.

Enhancer: 11 mg para-coumaric acid in 10 ml DMSO.

### ii) Lysate preparation

Whole cell lysis buffer or phospho-preserving lysis buffer was added directly to cell pellets obtained from polysome profiling experiments. Buffers were supplemented with protease inhibitors (complete mini, EDTA free protease inhibitor cocktail, Roche). Lysates were homogenised by 30 passages through a 25G needle. Insoluble material was pelleted by centrifugation at 13,000rpm for 10min at 4°C. The cleared lysates were stored at -20°C.

### iii) SDS-Polyacrylamide gel electrophoresis

SDS-PAGE gels were cast and run in a Bio-Rad Protean system. Depending on resolution required, the resolving gel mixtures were prepared as described below.

% Gel	4X Resolving Buffer (ml)	30 % Acrylamide (ml)	Distilled water (ml)
10	2.5	3.33	4.17
12.5	2.5	4.17	3.33
15	2.5	5	2.5

Gel mixes were polymerised by addition of ammonium persulphate and TEMED (N,N,N',N'-Tetramethylethylenediamine) to concentrations of 0.125 % w/v and 0.001 % v/v respectively. The stacking gel was prepared by mixing 1.8 ml of water, 2.5 ml of stacking

buffer, 0.67 ml of 30 % acrylamide/ 0.8 % bisacrylamide, 25 µl of 25 % APS and 5 µl of TEMED. An equal volume of 2X loading buffer was added to each lysate and denatured at 95°C for 5 min prior to loading. Gels were run in 1X SDS running buffer at a constant voltage of 120 V for approximately 2 hr, until the bromophenol blue dye front reached the bottom of the gel.

#### **iv) Transfer of protein to PVDF membrane**

Separated proteins were transferred to a PVDF membrane (Biorad) by semi-dry blotting. The membrane, pre-soaked in methanol, and gel were washed in transfer buffer and then sandwiched between layers of 3MM paper, also soaked in transfer buffer. Transfer was carried out at 10 V for 90 min. To confirm the efficacy of transfer, membranes were incubated for 5min in ponceau stain then washed repeatedly with water to reveal the bands. The location of each lane was marked with pencil before further washes to remove the stain. Non-specific binding was blocked by incubation with 5 % non-fat dried milk in TBST for 1 hr at room temperature.

#### **v) Protein detection**

Membranes were incubated overnight at 4°C in a solution of 5 % milk-TBST containing an appropriate concentration of primary antibody. Excess antibody was removed by 3x 10 min washes with TBST. Membranes were then incubated for 1 hr at room temperature with a horseradish peroxidase conjugated secondary antibody raised against mouse or rabbit immunoglobulins, diluted 1:5,000 in 5 % milk-TBST. Excess antibody was removed with 3 TBST washes, 5-10 min each. 1 ml luminol, 10µl enhancer and 3.1µl 3 % (v/v) hydrogen peroxide were mixed to create an enhanced chemiluminescent (ECL) mix. The ECL mix was applied to the membrane for 5 min and the luminescence visualised by exposure to X-ray film.

Antibodies

Target	Supplier (Clone)	Dilution	Source
β-Actin	Sigma	1:5000	Mouse
Cyclin D1	Santa Cruz	1:1000	Rabbit
Cyclin D1	Cell Signalling	1:1000	Mouse
c-Myc	Cell Signalling (D84C12)	1:1000	Rabbit
eIF4E	Willis Lab	1:1000	Rabbit
eIF2α	Abcam	1:500	Mouse
eIF2α-phospho	Cell Signalling	1:1000	Rabbit
eIF2B-ε	Cell Signalling	1:1000	Rabbit

### 2.7.b Immunohistochemistry

All immunohistochemical methods were performed by Dr Liam Faller at the Beatson Institute, Glasgow.

#### i) Slide preparation

Slides were deparafinised using standard methods, incubated in citrate buffer (10mM Sodium Citrate pH 6.0) for 30 min at 95°C, then left to cool. After washing with distilled water slides were incubated in 3 % H<sub>2</sub>O<sub>2</sub> for 10 min then washed for 5 min in TBST. Slides were blocked in 5 % normal goat serum-TBST (TBST-5 % NGS) for 30 min at RT then and incubated at 4°C overnight with primary antibodies diluted in TBST-5 % NGS as follows:

Target	Supplier (Clone)	Dilution	Source
Phospho S6K	Abcam (#ab32359)	1:200	Rabbit

Phospho rps6	Cell signalling (D57.2.2E)	1:400	Rabbit
Phospho 4E-BP1	Cell Signalling (#2855)	1:500	Rabbit

Slides were washed twice for 5 min each with TBST, then incubated with biotinylated anti-Rabbit 2° antibody (ABC kit; Vector Laboratories) for 30min at RT. The slides were then washed again twice for 5 min each with TBST then incubated with pre-mixed ABC complex for 30 min at RT. Excess ABC was removed in the same manner as the antibodies and the slide visualised with DAB reagent according to the manufacturer's instructions.

## Chapter 3. The contribution of tRNA<sub>i</sub><sup>Met</sup> to oncogenic transformation

### 3.1 Introduction

As described in the introduction, many components of the translational machinery may be dysregulated in oncogenesis. Much work has focussed on the role of initiation factors and microRNAs, however another class of non-coding RNAs – tRNAs – may also play a role in the formation of some tumours.

Microarray studies have revealed that the total expression levels of tRNAs are higher in cell lines derived from breast cancers compared to non-transformed breast epithelial cell lines (Pavon-Eternod et al., 2009) and in samples from lymphoma patients (A. Lund, unpublished observations). This is perhaps not surprising, as the activity of RNA Polymerase III, the polymerase responsible for transcription of tRNA, is regulated by well characterised oncoproteins and tumour suppressors such as c-myc and p53 respectively (White, 2004).

Notably, this upregulation is not uniform; tRNA transcripts encoding the same amino acid are differentially upregulated within each cancer cell line (Pavon-Eternod et al., 2009). However, some phenomena are correlative but not causative (Payne et al., 2008). The extent to which changes in tRNA abundance affect the translational efficiency of specific transcripts and thus are causative of transformation is an unresolved question.

There are two strands of evidence in support of a causative role, even if the exact mechanism underlying this is contested (Qian et al., 2012, Stadler and Fire, 2011). First, in cases of degeneracy (i.e. where an amino acid is encoded by multiple codons), the abundance of each anticodon is correlated with the abundance of the codon within the transcriptome of budding yeast (Ikemura, 1985) and mammalian reticulocytes (Hatfield and Rice, 1986). That these correlations differ between organisms, suggests that selective pressure operates to determine codon bias – i.e. codon bias can provide a selective

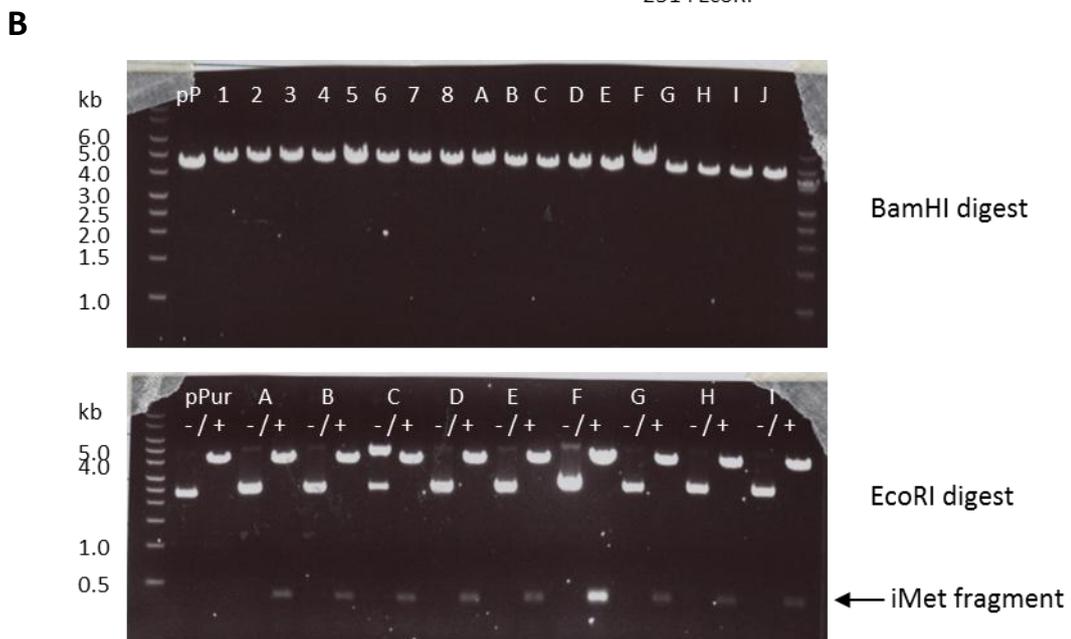
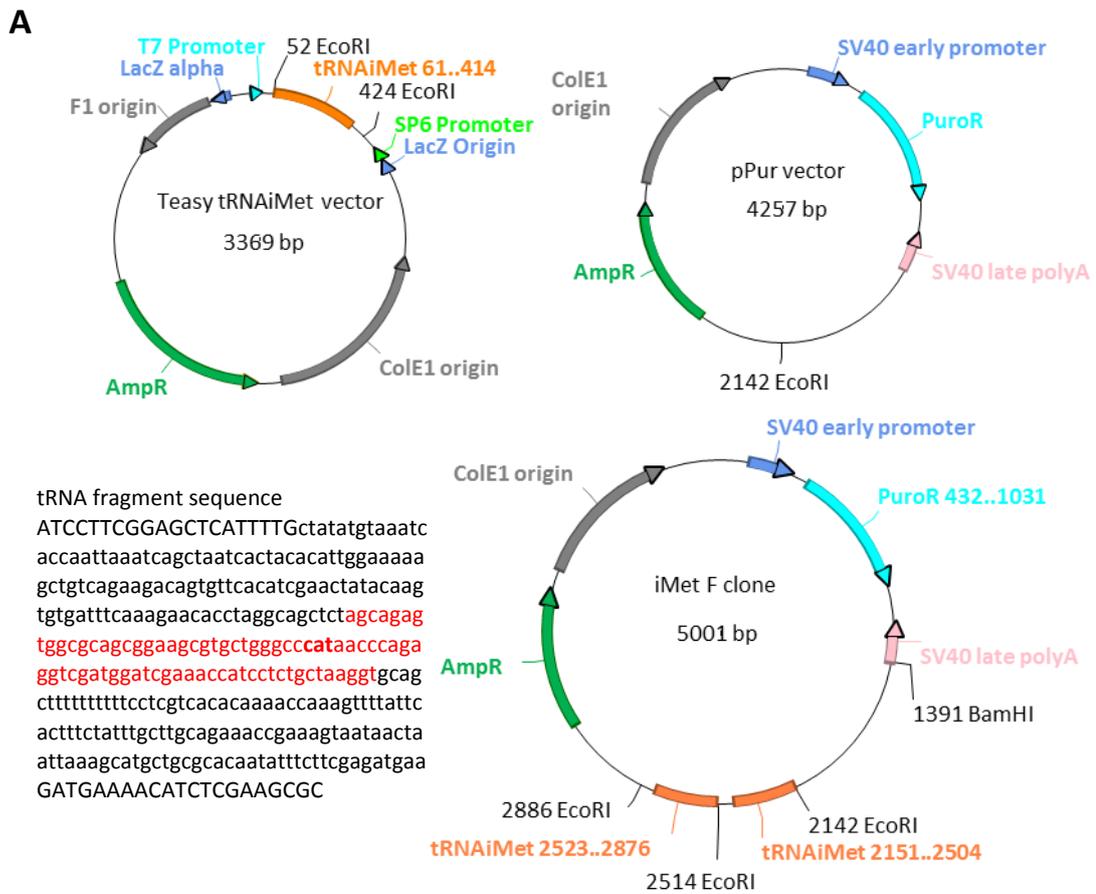
advantage for an organism/tumour. Furthermore the translational efficiency of endogenous proteins can be increased by mutating rarer codons to synonymous commoner ones, as has been shown for nPTB (Robinson et al., 2008) and EPO (Kim et al., 1997). Second, elevated expression of tRNA<sub>i</sub><sup>Met</sup> in mouse embryo fibroblasts has been associated with increased proliferation and oncogenic capacity (Marshall et al., 2008). Significantly, overexpression caused an increase in both the global translation rate and the translational efficiency of two transcripts, c-myc and cyclin D1.

We planned to use a recently described deep sequencing technique, ribosome profiling, to identify other transcripts differentially translated under high tRNA<sub>i</sub><sup>Met</sup> conditions.

### **3.2 Establishment of new high tRNA<sub>i</sub><sup>Met</sup> MEF cell lines**

Initial work using the MEF cell lines 2D and 4C described in Marshall et al., (kindly donated by Professor Robert J White) proved unproductive because over multiple (>10) passages the overexpression of tRNA<sub>i</sub><sup>Met</sup> was lost. These cell lines were constructed by co-transfection of a hygromycin resistance plasmid and a T-Easy based vector with or without the tRNA<sub>i</sub><sup>Met</sup> gene. It is possible that genomic integration had not occurred such that over time hygromycin selection permitted growth of cells without the tRNA<sub>i</sub><sup>Met</sup> plasmid. To avoid this issue, the fragment containing the tRNA<sub>i</sub><sup>Met</sup> gene was cloned into a puromycin resistance vector (pPur) using the EcoRI site. There are estimated to be 9 copies of the tRNA<sub>i</sub><sup>Met</sup> gene within the mouse genome (full details available from <http://gtrnadb.ucsc.edu/Mmusc10/mm10-tRNAs.fa>; sequences starting AGCAGAT) so a vector with only one copy of the gene would be unlikely to result in a significant increase in expression. Thus, a vast excess of insert:vector was used to encourage ligation of multiple copies of the gene within a single vector. To identify if any clones did indeed contain multiple copies of the fragment, an equal mass of each vector clone was digested with either BamHI to linearise, or EcoRI to yield the original inserts. All linearised clones were larger than pPur vector alone, with the F clone (iMetF)

being larger than all others (Figure 3.1 B, upper gel). Upon digestion with EcoRI, iMetF also yielded a much more intense band at ~350 bp suggesting the increased size was due to extra copies of the tRNA<sup>Met</sup> gene. Subsequent cloning yielded a vector larger than iMetF, estimated to contain 3 copies of the gene (named 3Met).



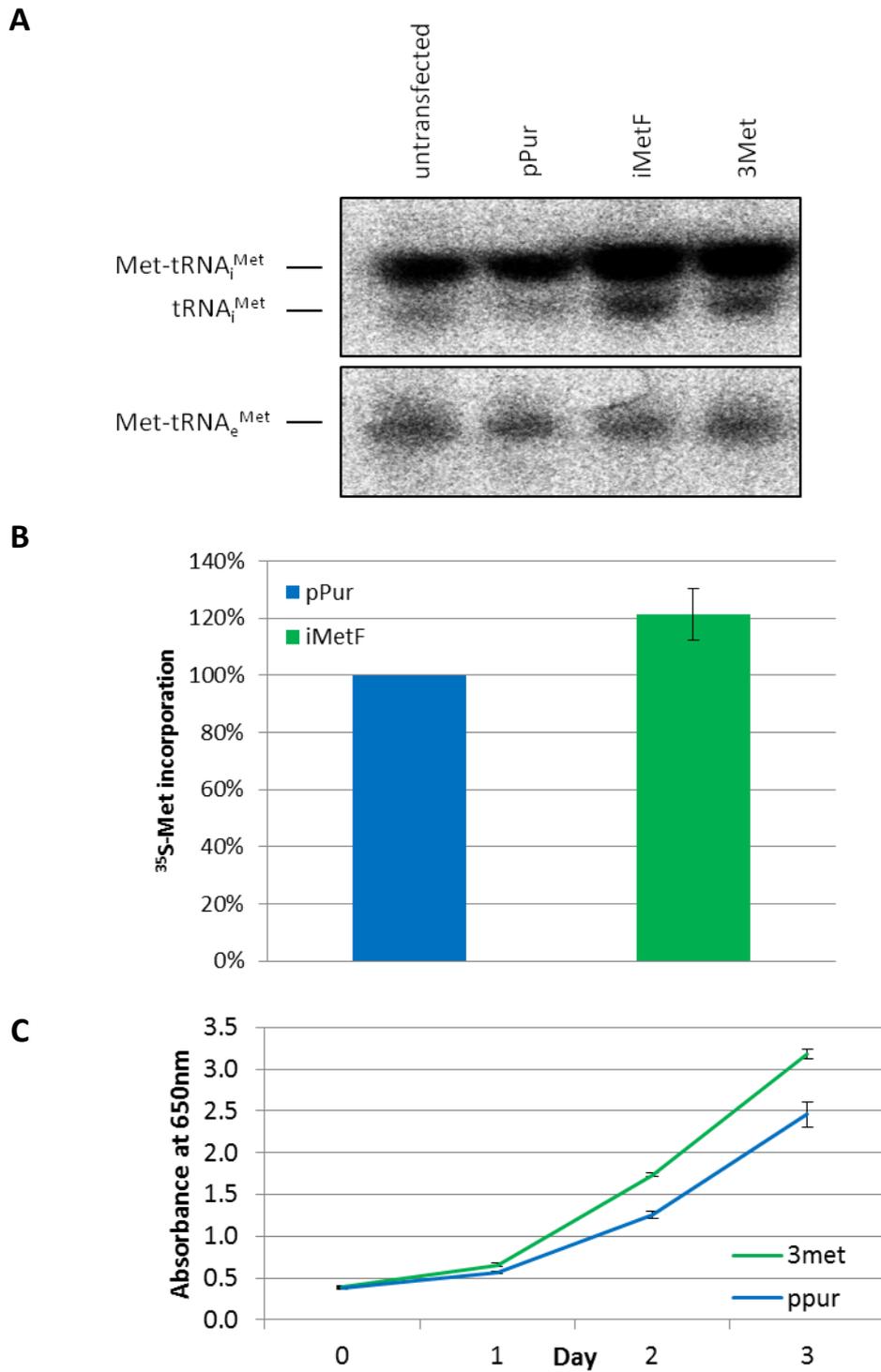
**Figure 3-1 Construction of a selectable tRNA<sub>i</sub><sup>Met</sup> expression vector**

**A** A genomic fragment containing the mammalian tRNA<sub>i</sub><sup>Met</sup> sequence (red text) was excised from a T-Easy vector backbone and ligated into puromycin resistance vector, pPur, using the EcoRI site. A 10:1 insert:vector ratio was used to encourage the ligation of multiple fragments within the same vector.

**B** Upon linearization, clone F was shown to be larger than other ligated vectors. Digestion with EcoRI revealed this clone contained at least two copies of the tRNA<sub>i</sub><sup>Met</sup> gene.

NIH 3T3 MEFs were transfected with either pPur, iMetF or 3Met vectors and selected using puromycin as described in Materials and Methods. Interestingly, total RNA from iMetF and 3Met transfected cells showed equally increased expression of tRNA<sub>i</sub><sup>Met</sup> compared to the pPur containing controls (Figure 3.2 A). This is somewhat surprising as the 3Met vector has more copies of the tRNA gene. That expression levels are similar may be due to an increased transfection efficiency of iMetF such that there are multiple copies of the vector within some cells creating a mixed population, or alternatively that there is a limit to Pol III activity upon a single vector. Importantly, by separating RNA under acidic conditions it was possible to confirm that the increased tRNA levels were of the aminoacylated form. Given that each amino acid has its own aminoacyl-tRNA-synthetase that specifically recognises tRNAs with the correct anticodon, it is most likely the upper band in the northern blots are methionyl-tRNA, capable of forming ternary complex and able to participate in translation initiation.

As was observed by (Marshall et al., 2008), overexpression of tRNA<sub>i</sub><sup>Met</sup> was associated with modest increases in protein synthesis (Figure 3.2 B) and cell proliferation (Figure 3.2 C)



**Figure 3-2 Elevated tRNA<sub>i</sub><sup>Met</sup> is associated with increased protein synthesis and cell proliferation.**

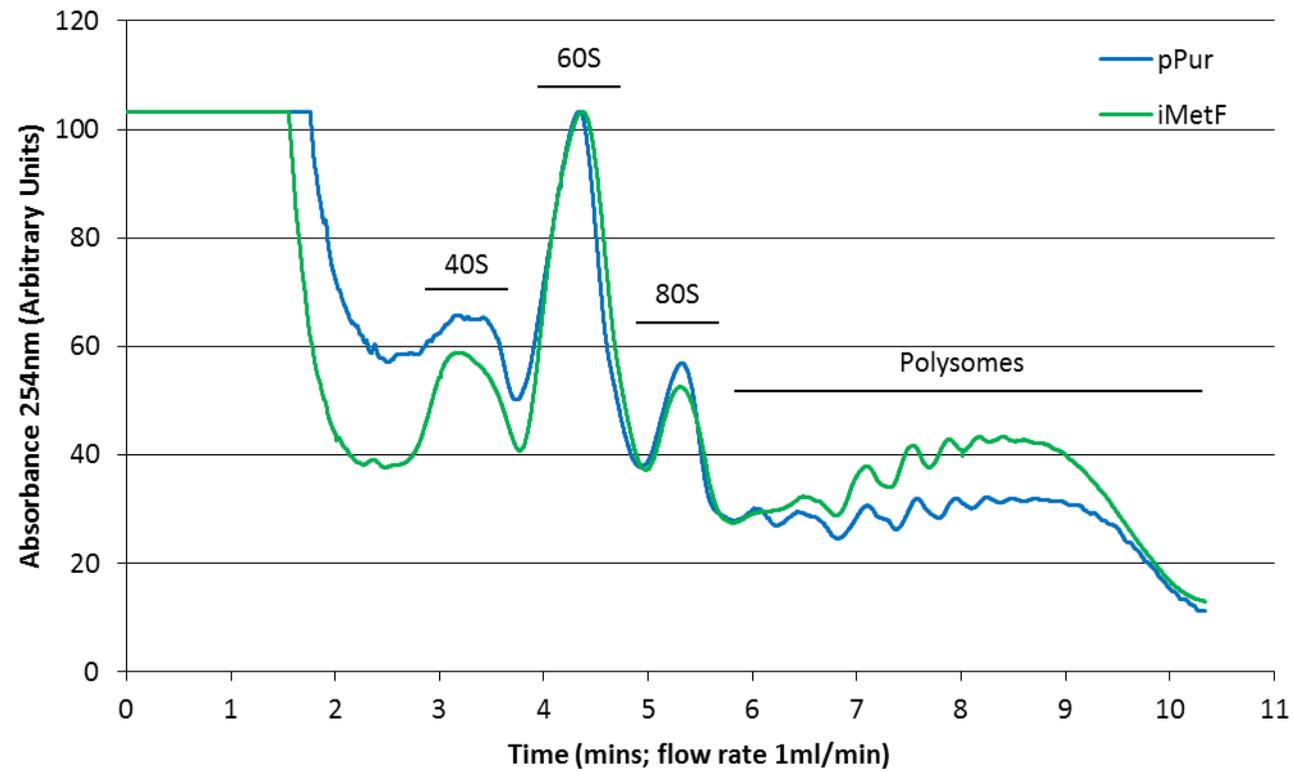
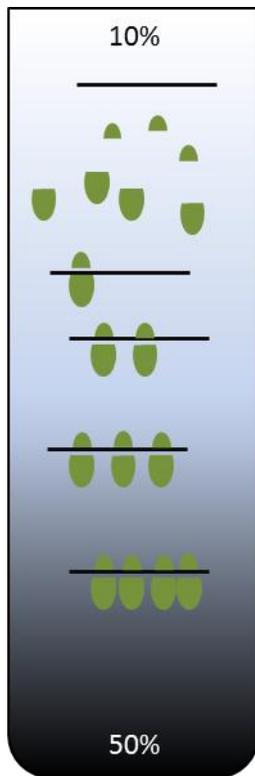
**A** Representative northern blot of total RNA separated by acid-urea PAGE. NIH 3T3 MEFs were stably transfected with the vector listed above the lane.

**B** Representative methionine incorporation assay data. Values are the mean of 3 independent experiments, each from 3 wells. In each experiment values were normalised to pPur, thus the standard deviation is only calculated for iMet.

**C** Representative methylene blue proliferation assay. Values are the mean of 6 wells +/- SD

### 3.3 $\text{tRNA}_i^{\text{Met}}$ and translational control

To investigate the change in polysomal association of mRNAs, cycloheximide treated post-nuclear lysates were subject to sucrose gradient ultracentrifugation. An increase in polysomal association of RNA was observed in both high  $\text{tRNA}_i^{\text{Met}}$  cell lines (a representative trace of iMetF is shown in Figure 3.3, 3Met was very similar). When taken with the increased rate of protein synthesis, the data suggest that translation initiation is the predominant stage of regulation within these cells (a reduction in polysomal association with an increase in protein synthesis rate would suggest elongation as a primary stage of regulation, a topic that shall be explored in later chapters).



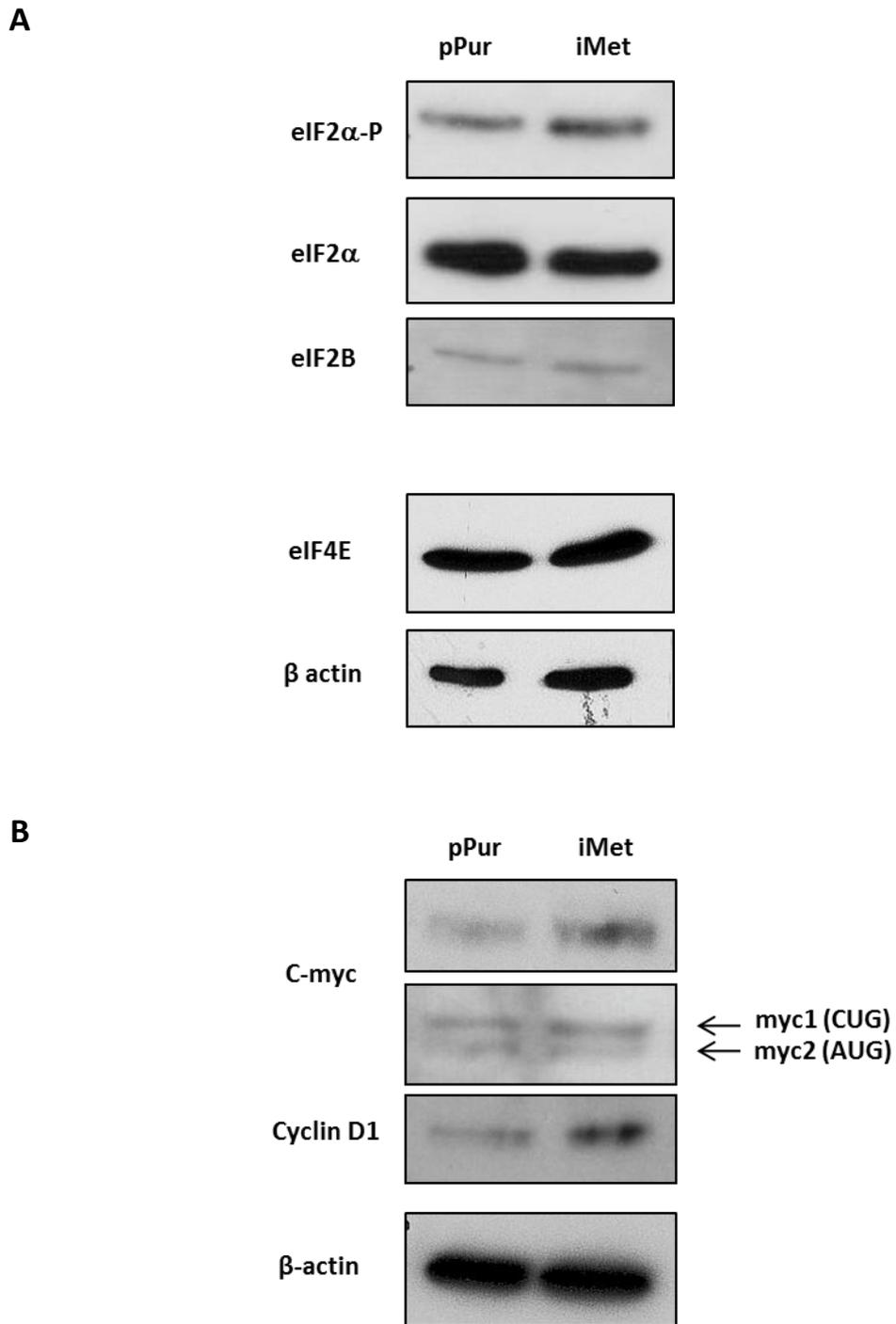
**Figure 3-3 Elevated  $tRNA_i^{Met}$  expression is associated with increased polysomal association**

Total cytoplasmic extracts were separated by sucrose gradient ultracentrifugation, represented in the image (left). Gradients were separated by upward displacement giving the representative absorbance trace shown (right).

As described in the Introduction, global translation initiation is typically limited either by the availability of the eIF4F complex or of ternary complex. Given the change in translation initiation in response to increased  $\text{tRNA}_i^{\text{Met}}$  it was assumed that the level of ternary complex was limiting in the cell line used. There was no dramatic decrease in the phosphorylation of eIF2 or an increase in eIF2B in iMet cells compared to the control (Figure 3-4 A), suggesting that  $\text{tRNA}_i^{\text{Met}}$  alone is limiting for ternary complex formation in these cells. There was no difference in eIF4E expression, however a definitive answer on the phosphorylation status of the 4E binding proteins could not be obtained. In the absence of these data, the possibility that increased  $\text{tRNA}_i^{\text{Met}}$  levels also increase eIF4F availability cannot be excluded, though the application of Occam's razor would suggest that this is unlikely.

Consistent with the earlier report of Marshall et al., increased expression of c-myc and cyclin D1 in response to increased  $\text{tRNA}_i^{\text{Met}}$  was observed (Figure 3-4 B). It was presumed that this was due to increased translational efficiency, but on reflection an assessment of mRNA abundance should have been performed to confirm this. Notably, the c-myc isoform (myc1) that is initiated from an alternate start codon (CUG) was expressed at higher levels in the iMet cells. Whilst it has been reported that CUG start codons can be translated by their cognate  $\text{tRNA}^{\text{Leu}}$  in immune cells (Starck et al., 2012), it was also the case that such CUG start codons nonetheless have higher occupancy of  $\text{tRNA}_i^{\text{Met}}$  than  $\text{tRNA}^{\text{Leu}}$ , consistent with the classical scanning model of translation initiation. The increased expression of the CUG initiated myc isoform led us to speculate that increased ternary complex availability may increase translation from less favoured start codons more generally and that this might partially explain the phenotypic changes observed.

To investigate this, and other potential mechanisms, attempts were made to identify other mRNAs that were also differentially translated in response to increased  $\text{tRNA}_i^{\text{Met}}$  and whether there were common features or motifs that contributed to this.



**Figure 3-4 Expression and phosphorylation status of eIFs and oncogenic proteins in response to increased  $tRNA_i^{Met}$**

A) Representative western blots for the initiation factors listed.  $\beta$ -actin was used as a loading control in addition to ponceau staining (not shown).

B) Representative western blots for c-myc and cyclin D1. By running a PAGE gel so that the 50kDa marker is about to run off the bottom it is possible to resolve two myc isoforms, one of which is initiated from a non-AUG start codon.

### **3.4 Using ribosome profiling to identify differentially translated mRNAs**

The previous approach of the Willis lab to identifying changes in translational efficiency between different conditions was to use translational profiling (Bushell et al., 2006, Powley et al., 2009). In this technique, RNA from the subpolysomal fractions of the sucrose density gradient are pooled, reverse transcribed and labelled with a fluorescent Cy dye. The remaining, polysomal, fractions are treated in the same way but with a different dye. These cDNA pools are then hybridised to microarrays containing probes to known mRNAs and the ratio of each dye intensity used as a measure of translational efficiency.

However, this approach has three key limitations. First, it only identifies changes in the distribution of mRNA across a single point in the density gradient. For example, an mRNA that is associated with >8 ribosomes (fractions 9-11 in Figure 3-3) in condition A and 4-5 ribosomes (fraction 7 in Figure 3-3) in condition B would not be identified as differentially translated by this method despite at least a 2 fold difference in translational efficiency.

Second, expression data from a cDNA microarray is limited to the subset of complementary probes on its surface; there may be tissue specific genes that show differential regulation but cannot be detected as there are no complementary probes present on the array. Third, the technical biases that plague microarrays are well characterised (Smyth and Speed, 2003), but the normalisation methods used to correct these make assumptions that are not necessarily applicable to translational profiling data (e.g. that the ratio of each dye should be 1:1).

Subsequent work within the Willis lab has sought to develop new normalisation approaches, the details of which are beyond the scope of this thesis.

For the above reasons it was decided to use the recently developed next-generation sequencing technique ribosome profiling (Ingolia et al., 2009) to identify differentially translated mRNAs.

Ribosome profiling involves the sequencing of two pools of RNA collected simultaneously – one derived from all polyadenylated RNAs (the polyA+ transcriptome), the other from fragments created by digesting ribosome bound mRNA (the translome). The former will give a value proportional<sup>1</sup> to the abundance of an mRNA, the latter the degree of ribosomal association. By dividing one by the other it is possible to estimate the translational efficiency of an mRNA. Hypothesis testing can then be performed to identify mRNAs that show significant differences in translational efficiency between conditions.

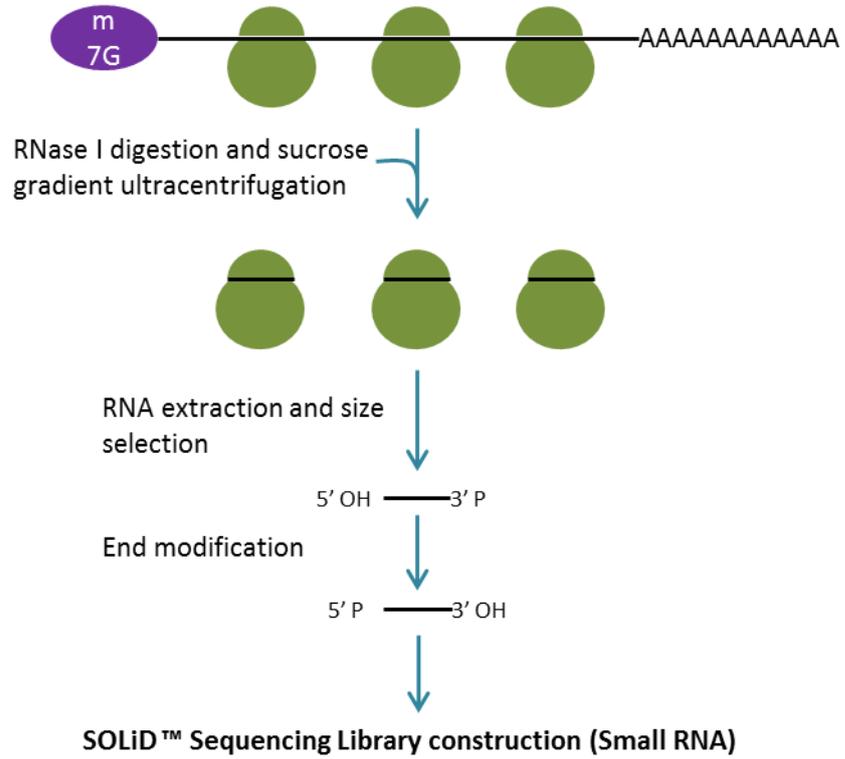
The process of fragment generation for ribosome profiling is represented (Figure 3-5) and will be described below. To assess the whole transcriptome, polyA+ RNA was selected from total RNA extracts from an aliquot of cells pelleted prior to lysis. PolyA+ RNA was fragmented with RNase III and used to generate a sequencing library compatible with the SOLiD sequencing platform. On reflection, it would have been more appropriate to extract the RNA from the post nuclear lysate as there may be mRNAs that show differential nuclear:cytoplasmic localisation between conditions. In addition, fragmentation with RNase III results in patchy coverage, an issue which will be addressed in a later chapter.

The isolation of ribosome protected mRNA fragments exploits the fact that mammalian ribosomes arrested on mRNAs protect a region of approximately 30 nt within the ribosomal cleft from digestion by RNases (Wolin and Walter, 1988). When a post nuclear lysate is digested with RNaseI and separated by sucrose gradient ultracentrifugation these fragments are present within the 80S peak. RNA can be extracted from this region of the gradient using the acid phenol method and size selected on a denaturing acrylamide gel. However, digestion with RNaseI leaves fragments with 5' hydroxyl and 3' phosphate groups which cannot be successfully ligated to adaptors to form the templates for cDNA sequencing libraries. These ends are therefore modified with polynucleotide kinase.

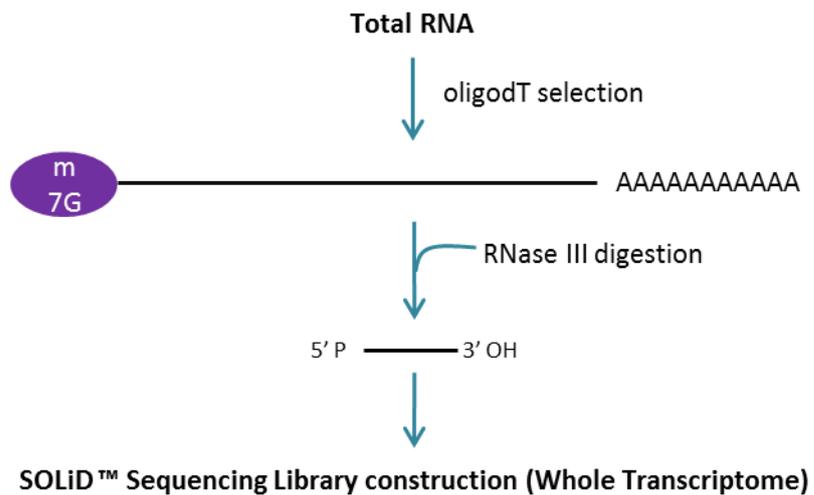
---

<sup>1</sup> A full explanation as to why these values are proportional rather than absolute is described in a later section.

**A**



**B**



**Figure 3-5 Ribosome profiling schematic**

Ribosome protected and RNase III digested fragments are generated in parallel and subjected to next generation sequencing. See Materials and Methods for full details

### 3.5 RNA sequencing using SOLiD technology

Fragments were ligated to proprietary duplexed DNA adaptors with 6n degenerate overhangs. The degenerate overhangs should permit hybridisation of all RNA fragments, with the common adaptor sequences acting as primer binding sites for subsequent reverse transcription and PCR amplification. These cDNA libraries were then amplified onto magnetic beads by PCR within an oil emulsion. The result was a mix of beads with no DNA (where either no cDNA read or polymerase was present within the emulsion vesicle), monoclonal beads (only one cDNA read was within the vesicle) and polyclonal beads (multiple cDNA reads within the vesicle). Successfully amplified beads were isolated. A fixed number of beads were deposited onto the surface of the slide, left to adhere and any unbound beads washed off. The slide was loaded into the sequencer and processed using either a 50bp run (whole transcriptome) or 35 bp run (translatome) using SOLiD technology. The details of this technology are available on the Life Technologies website<sup>2</sup> but it is also described briefly below with a view to helping the reader understand the sources of variance within sequencing data. Polyclonal beads were excluded from downstream analysis.

#### 3.5.a SOLiD sequencing reaction overview

In each cycle of the sequencing reaction, degenerate 8mer oligonucleotides are annealed to the bead bound reads. To align the 8mers to the start of the read, a primer antisense to the common RNA adapter is first annealed to the reads. Each oligonucleotide is 3' conjugated to one of 4 fluorophores dependent on the two 5' most nucleotides (e.g. blue = AA, CC, GG or TT; green = AC, CA, TG or GT). Following ligation, the oligonucleotide is cleaved after the 5<sup>th</sup> nucleotide, releasing the fluorophore. An image of the slide is captured at this point such that a colour is associated with each bead. This ligation-cleavage-imaging cycle is repeated such that the 1<sup>st</sup> and 2<sup>nd</sup>, 6<sup>th</sup> and 7<sup>th</sup>, 11<sup>th</sup> and 12<sup>th</sup>, etc. nucleotides of the read are

---

<sup>2</sup> E.g.:

[http://www3.appliedbiosystems.com/cms/groups/mcb\\_marketing/documents/generaldocuments/cms\\_057810.pdf](http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_057810.pdf)

interrogated. The oligonucleotides are then denatured and removed from the slide and the process is repeated with a primer that is offset by one nucleotide such that the 2<sup>nd</sup> and 3<sup>rd</sup>, 7<sup>th</sup> and 8<sup>th</sup>, 12<sup>th</sup> and 13<sup>th</sup> nucleotides are interrogated. This continues until all positions have been interrogated twice.

The result is a series of colour calls for each position of each bead/read. For example, assuming the adaptor ends with a T, then the read TGGGGTG... would have the colour call: blue (TT)-green(GT) and green(TG) and green (TG) when the primer with a +1 offset is used.

### **3.5.b Conversion of colour calls to nucleotide reads**

Colour calls are then mapped to both a reference genome and to a filter set containing tRNAs and rRNAs. The use of the reference genome resolves situations where two colour calls overlapping the same nucleotide are discordant (eg calling yellow-green for +1 offset string in the above example). This can occur if an oligonucleotide hybridised with a sequence that is not 100 % complementary such that the colour called does not correspond to the true dinucleotide sequence. In some cases the image cannot be correctly read (because the reaction was unsuccessful and/or the reads were too close together) so no genomic mapping can be made. By merging the reads that successfully map to a genomic location and are not filtered a list of (predominantly) mRNA reads and their nucleotide sequence is obtained.

### **3.6 Analysis of ribosome profiling read data**

Due to differences in the number of adherent beads and the quality values from the sequencing reactions the total number of reads differs between libraries (see Figure 3-6). In order to make comparisons between libraries, expression values are normalised relative to the total number of reads – expressing the abundance of a sequence/mRNA in reads per million (rpm). To illustrate why this is necessary, imagine a scenario where the same library was sequenced twice, the first run yielding twice as many reads as the second (this would be described as having twice the read depth). For gene x there may be 100 reads in run one but

only 50 in run two. Looking at the raw values alone would suggest there was differential expression, which is not the case (the libraries were identical).

Differences in read depth, along with the fact that libraries are amplified prior to sequencing, explains why sequencing data are only ever proportional to the abundance or ribosomal association of an mRNA.

When comparing the mapping statistics of translome and transcriptome libraries two differences are striking. First, the number of reads that map uniquely to genomic location is much higher in the transcriptome libraries (Figure 3-6, bottom row). The likely cause is the whole transcriptome libraries contained longer fragments of which 50 bp were sequenced, whereas ribosome protected fragments are ~ 30 nt. The second difference is that the whole transcriptome libraries are less contaminated by tRNA and rRNA sequences. rRNA contamination of the translome libraries occurs because RNaseI will also digest the rRNA components of the ribosome itself as well as mRNA that is not protected by the ribosome. So long as a partially digested ribosome retains an 80S structure these rRNA fragments, some of which are ~30 nt in length, will be found within the monosome peak and ultimately will be incorporated within the sequencing library. Ultimately this means that even for similar numbers of total reads there are considerably fewer useful reads generated by sequencing translome libraries – the effective read depth is lower. How this affects attempts to accurately identify differentially translated mRNAs genome-wide will be discussed later in this section.

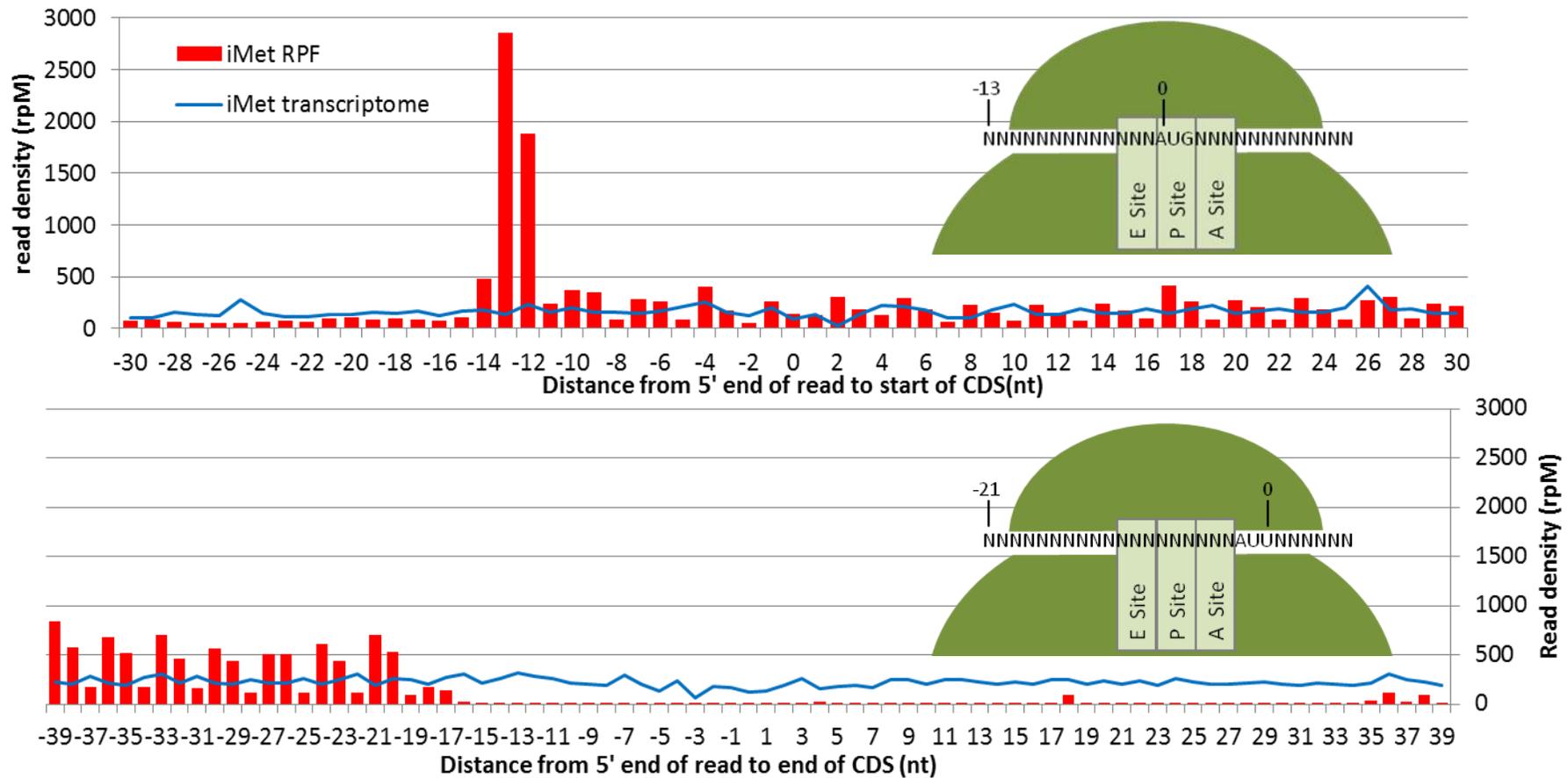
	Ppur Whole transcriptome				Ppur RPF				iMet Whole transcriptome				iMet RPF			
	Replicate 1		Replicate 2		Replicate 1		Replicate 2		iMetF		3Met		iMetF		3Met	
<b>Total reads</b>	73,284,040	100%	92,688,211	100%	55,039,740	100%	95,795,673	100%	41,004,828	100%	84,806,960	100%	70,707,146	100%	105,952,909	100%
<b>Reads mapped to genome</b>	50,944,225	70%	67,907,508	73%	21,060,761	38%	41,303,251	43%	13,833,758	34%	60,815,214	72%	36,234,043	51%	50,438,232	48%
<b>Reads mapped to filter set</b>	395,576	1%	762,920	1%	14,833,841	27%	42,412,968	44%	403,038	1%	763,316	1%	34,008,741	48%	44,134,842	42%
<b>Reads mapped, not filtered</b>	50,612,362	69%	67,284,384	73%	10,060,845	18%	9,621,477	10%	13,446,462	33%	60,193,065	71%	10,312,738	15%	18,160,457	17%
<b><i>Reads with too many mappings (N &gt;= 10)</i></b>	4,271,846	8%	4,759,162	7%	3,487,836	35%	2,415,710	25%	1,434,928	11%	4,428,550	7%	1,941,379	19%	5,037,955	28%
<b><i>Reads with number of mappings in proper range (N &lt; 10)</i></b>	46,340,516	92%	62,525,222	93%	6,573,009	65%	7,205,767	75%	12,011,534	89%	55,764,515	93%	8,371,359	81%	13,122,502	72%
<b><i>Reads uniquely aligned</i></b>	36,322,456	72%	49,468,003	74%	1,658,788	16%	2,513,242	26%	7,566,758	56%	44,300,874	74%	4,429,271	43%	5,049,273	28%

**Figure 3-6 Mapping statistics for SOLiD sequencing runs**

Mapping statistics for all 8 libraries. Reads were mapped to the mouse reference genome mm9 and, independently, to a filter reference set containing rRNA and tRNA sequences. Percentages in the first four rows are expressed relative to the total. Unfiltered mapped reads are further divided based on the number of genomic locations they map to (bottom three rows) with percentages relative to the total unfiltered mapped reads in italics

Before testing for differential translation of mRNAs it was necessary to confirm that the unfiltered reads from the translome libraries mapped to positions consistent with ribosome protection. Using a custom Perl script (written and executed by Dr Ruth V. Spriggs), the 5' end of each read was aligned relative to both the start and the end of the coding sequence of the RefSeq that the read best mapped to. The number of reads aligning at each nucleotide position was determined, summed for all RefSeqs and expressed as a read density (rpm). Reassuringly, the translome reads mapped almost exclusively to the 5'UTR and CDS of transcripts (Figure 3-7). There is a striking 3n periodicity within the CDS-aligned reads, consistent with capture of ribosomes on in frame codons. That there are reads at +1 and +2 positions relative to the start of the codon is unlikely to be due to the capture of ribosomes during translocation because the treatment with cycloheximide, which blocks the E site preventing further translocation of the ribosome (Pestova and Hellen, 2003, Schneider-Poetsch et al., 2010), is for considerably longer than the time required for translocation. A more likely explanation is that the RNase I digestion does not generate ribosome protected fragments of identical length, perhaps due to slight conformational differences in the ribosome at different codons, so the 5' most end of the reads are offset by 1 or 2 nucleotides. The read density on the 5'UTR is lower than in the CDS and lacks periodicity. This is consistent with the fact that only ~40 % of messages are reported to have uORFs and that these may be in a different frame to the CDS. Finally, the highest read density is clustered around the start site. This may, in part, be because the initiation of the full competent 80S ribosome at the start site involves multiple biochemical changes which may be comparatively slow compared to elongation. However it is also likely that part of this density is artefactual: because cycloheximide will only inhibit elongating ribosomes; newly initiated ribosomes can form during the treatment period on start sites unoccupied at the start of treatment.

Reads from the transcriptome libraries mapped to all regions of mRNAs, as would be expected. However, the density was not even across all positions as would be expected if the RNA was randomly fragmented and sequenced without bias. The source and potential remediation of biases in the sequencing protocols are addressed in a later chapter, although it is worth noting at this stage that RNase III does have a preference for double stranded RNA and was subsequently reported by Applied Biosystems to give more uneven coverage than chemical fragmentation, but that any biases should broadly apply equally between libraries unless there are substantial changes in the transcriptome.



**Figure 3-7 Ribosome profiling gives sub-codon resolution of ribosome density**

The 5' end of each mappable read was aligned relative to the CDS start (top) and CDS stop (bottom) of the RefSeq it mapped to. The number of counts at each position across all RefSeqs was obtained and expressed as a ratio relative to the total number of aligned reads – reads per Million.

The next step was to identify differentially translated mRNAs. A naïve (and ultimately inappropriate approach) would be to generate translational efficiency values (TE;  $\text{translatome rpm} / \text{transcriptome rpm}$ ) for each transcript in each replicate and do hypothesis testing (say with a familiar parametric test such as the student's t test), with a multiple testing correction to account for the fact that when testing genome wide the number of comparisons being performed means there will necessarily be some false positives.

To appreciate why such an approach would be incorrect it may be helpful to first outline a mathematical understanding of the problem.

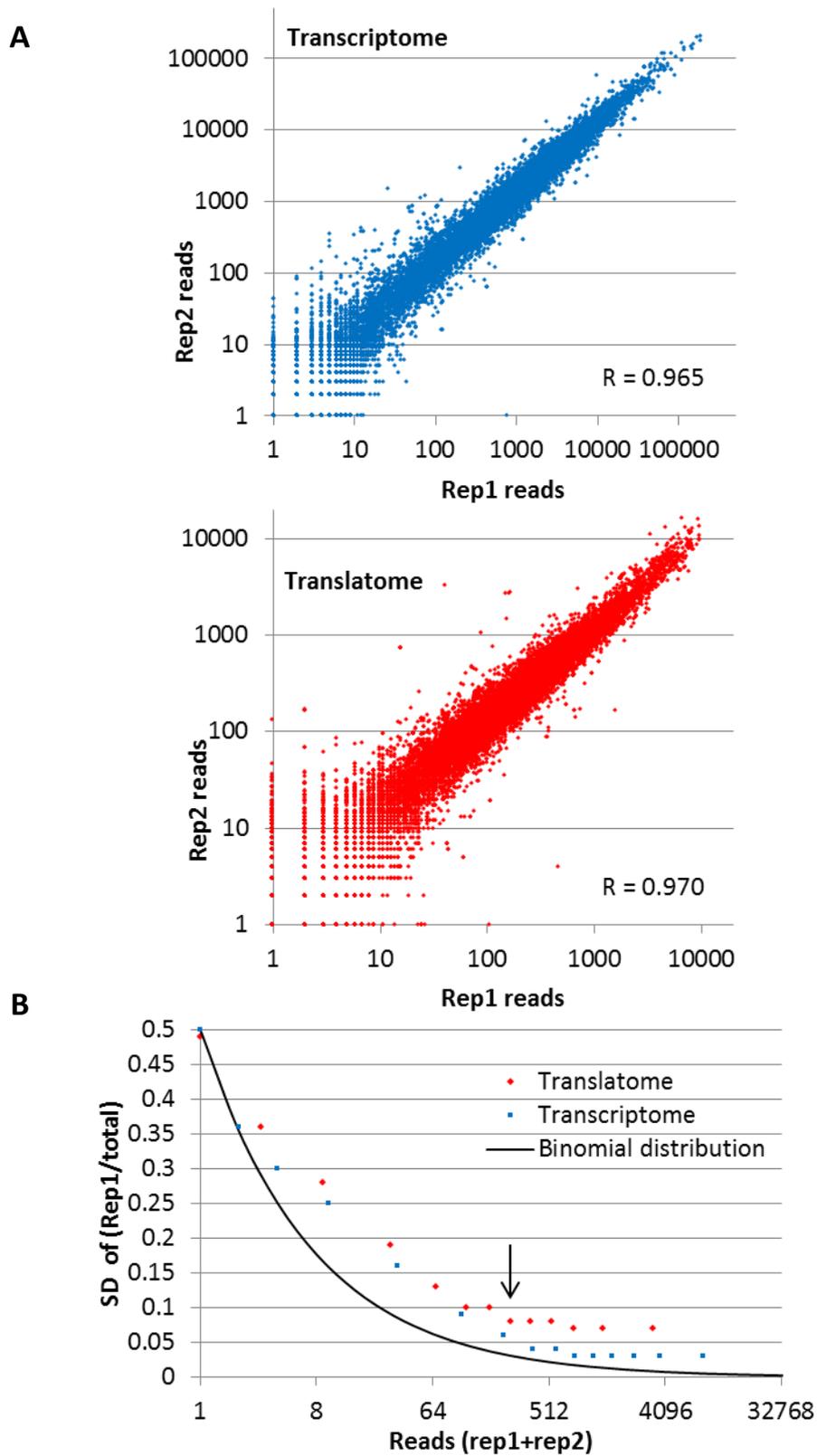
As with any quantitative property of interest, the expression (or ribosome occupancy) of any transcript ( $t$ ) can be considered to be a random variable with two parameters: mean ( $\mu_t$ ), and variance ( $\sigma_t^2$ ), and a probability distribution (e.g.: Gaussian/Normal). Experimental data are simply  $n$  samples of this random variable under different conditions resulting in a sample mean ( $\bar{x}$ ) and sample variance ( $s^2$ ) for each condition. Using a parametric test to identify genes differentially expressed in two conditions (A, B) involves making an assumption as to the distribution of this variable, using the sample means ( $\bar{x}_{tA}, \bar{x}_{tB}$ ) and variances ( $s_{tA}^2, s_{tB}^2$ ) to estimate probability that the two conditions share the same mean ( $\mu_{tA} = \mu_{tB}$ ).

The t-test assumes both that the data are continuous and normally distributed, neither of which are true for sequencing. Sequencing data are essentially the result of a counting process with discrete integer counts (reads), and random count based observations (e.g. radioactive decay) follow a Poisson distribution. However early RNA-seq data showed sequencing data has more variability than this and thus is not well modelled by such a distribution; a negative binomial distribution is felt to better approximate sequencing data (Anders and Huber, 2010, Robinson et al., 2010). This allows the variance to be modelled as two components – random shot noise and raw variance. The raw variance (which would be expected to be predominantly biological variance as standardised protocols should minimise

technical biases) will differ between datasets; the shot noise is, by definition, a Poisson distributed component. Information is pooled across genes with similar total read counts because two data points are insufficient to accurately estimate the per-gene variances.

Ultimately, it was not possible to use edgeR and DESeq to test for differential translation, because the hypothesis test for these packages is for the simple case comparing two transcriptome libraries. In that case there are two random variables (mRNA abundance in condition A, mRNA abundance in condition B) each negatively binomially distributed so the probability of the observed data occurring under the assumption of the null hypothesis can be estimated using an exact test. Translational efficiency is a compound variable:  $\mu_{\text{ribosome occupancy}} / \mu_{\text{mRNA abundance}}$ . This will not be negatively binomially distributed nor is it safe to assume that they will have similar distributions under the null hypothesis.

However, what the approaches of edgeR and DESeq highlight is that there are multiple components of variance. Importantly, because shot noise is Poisson distributed (i.e. it is equal to the mean) it has a greater effect on low count data than high count data. For example if  $\bar{x}_{tA} = 4$  then  $s_{tA} = \sqrt{4 + \text{biological variance}}$  and if  $\bar{x}_{tB} = 400$  then  $s_{tB} = \sqrt{400 + \text{biological variance}}$ . The degree to which the sample mean is influenced by this shot noise can be compared by computing the signal to noise ratio ( $\bar{x}/s$ ) for each condition. For simplicity, assume there is no biological variance (e.g. technical replicates where there should be no differential expression), then the signal to noise ratio is much smaller in A (2) than B (20). This means that, all other things being equal, translational efficiency values computed from low count data are less accurate than those from high count data. This partially explains the fanning out of the low count data when biological replicates are plotted against each other (Figure 3-8 A).



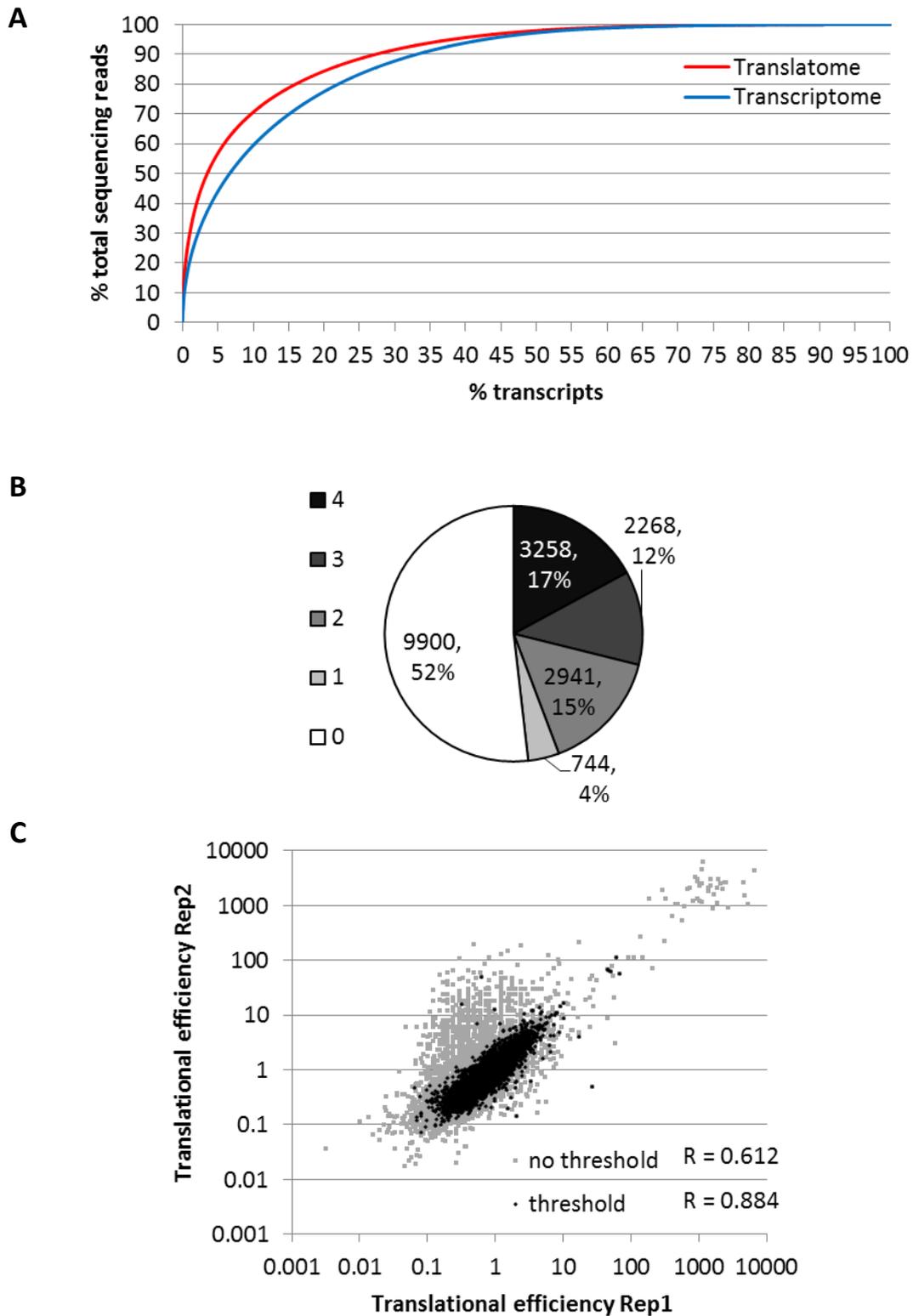
**Figure 3-8 Applying minimum read thresholds to sequencing data**

**A** Correlation of raw read counts for each RefSeq between iMet replicates

**B** Scatter plot of a measure of inter replicate variance within equally sized bins. For equally sized libraries governed purely by random noise a binominal distribution would be expected. The start of the plateau (>260 reads, indicated with arrow) was chosen as the threshold.

Being unable to use parametric hypothesis tests necessitated the use of non-parametric tests that use ranks derived from the raw means. However, if the signal to noise ratio is much lower for low count data, then using these methods would likely yield false positives of transcripts with very few counts because their rank would be more of a reflection of noise. To avoid this, Ingolia has suggested determining a threshold number of counts beyond which transcripts have broadly similar variances (Ingolia et al., 2009). To do this, genes are binned based on the total number of transcripts between two biological replicates and the standard deviation of the fraction of reads coming from a particular replicate computed within that bin. A plot of these per-bin variances is shown in Figure 3-8 B.

If transcripts were distributed on a purely random basis between replicates of equal depth (i.e. pure shot noise, analogous to flipping a coin) the variances would follow a binominal distribution. That they are over dispersed suggests there are additional sources of variance within all bins – biological and technical variance. Beyond 260 counts the variance plateaus. This is somewhat higher than the 128 counts in (Ingolia et al., 2009) and 200 in (Ingolia et al., 2011), but the biological and technical variability will differ between experiments and must be estimated from the data rather than simply taking a numerical value as appropriate in all cases as appeared to be done in (Thoreen et al., 2012).



**Figure 3-9 Consequences of applying minimum read thresholds to sequencing data**

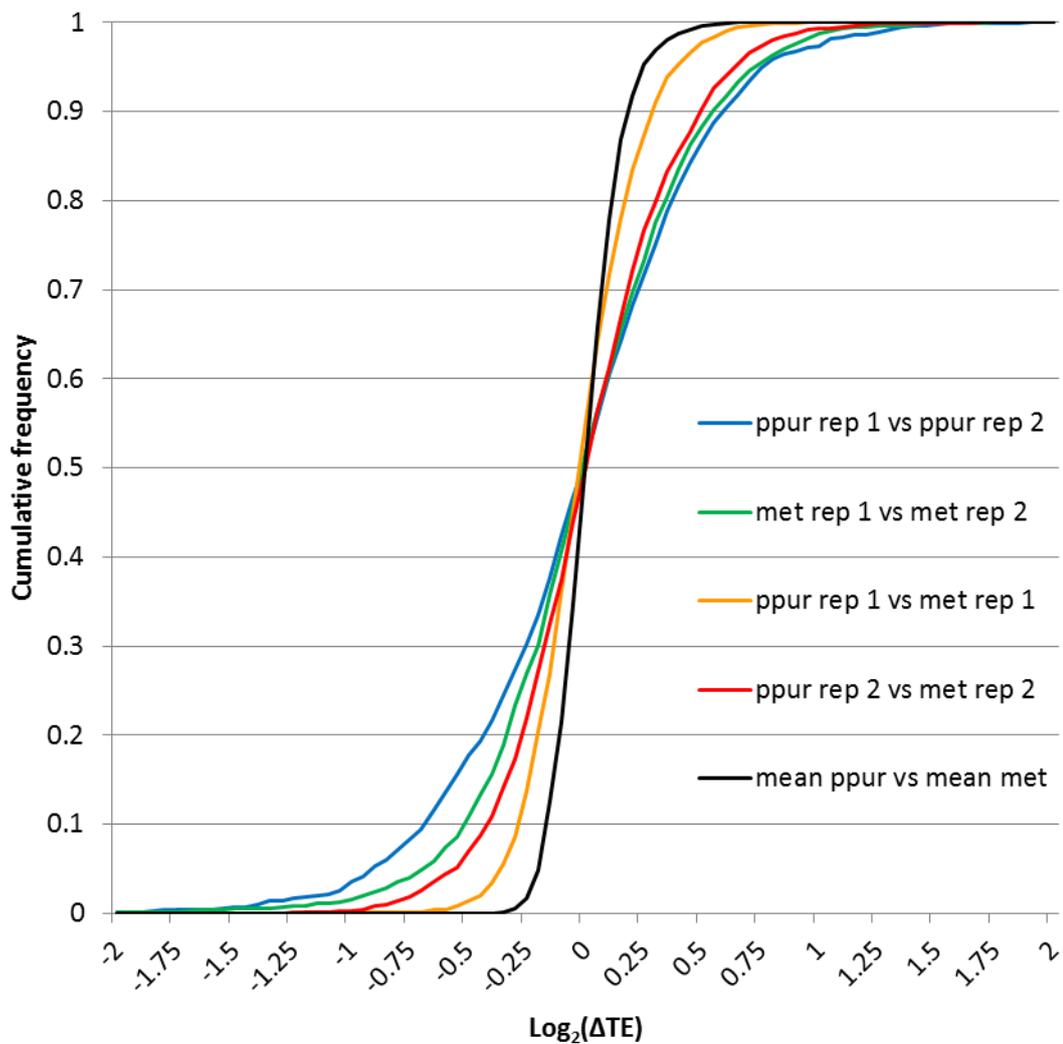
**A** Cumulative frequency plot of the % of total non-filtered reads that map to transcripts

**B** Pie chart of genes that meet the chosen threshold (260 counts) in all four library types (pPur/iMet, transcriptome/translatome).

**C** Scatter plot of translational efficiency values for the two iMet replicates with and without application of the threshold.

This use of thresholds effectively limits identification of differentially expressed mRNAs to those that are highly expressed. This is because the abundance and translation of genes follow a Pareto-like distribution – the 20 % most abundant/translated messages account for approximately 80 % of all sequencing reads (Figure 3-9 A). Consequently only 3258 genes crossed the threshold in all four library types (Figure 3-9 B). However, employing this cut off approach significantly improved the correlation of TE values between the replicates (Figure 3-9 C). This validates the threshold approach, but also highlights that when using very few replicates the technique does not give an accurate genome-wide picture of the translational landscape. For example, neither c-myc nor cyclin D1 met the threshold.

To test for differential translation between these ~3000 genes the TE values for each replicate were calculated. These values were analysed with RankProd, a nonparametric test that has been successfully used on microarray DNA data where parametric testing would have been inappropriate (Breitling et al., 2004). No differentially translated genes were identified. This may have been because 2 replicates was too few for such a conservative test so another approach of Ingolia was used (Hsieh et al., 2012). The biological replicates were used to estimate the range of  $\Delta$ TE values that occur simply through biological variance and determine the  $\Delta$ TE values beyond which a comparison between conditions is significantly distinct from biological variance. Unfortunately a greater range of  $\Delta$ TE values was obtained when comparing biological replicates than when comparing different conditions isolated on the same day and sequenced on the same sequencing run (Figure 3-10, blue and green vs yellow and red lines). Ultimately it was concluded the signal-to-noise ratio for the means was too low to identify differentially translated mRNAs from these two replicates.



**Figure 3-10 Inter-replicate variation confounds attempts to identify differentially translated messages.**

For the comparisons given in the figure legend, the logfold change difference in translational efficiency (TE; see materials and methods for more details) was determined for each RefSeq with over 260 reads in all libraries. This data is represented in a cumulative frequency plot, with the median fold change approximately 0.

### **3.7 Elevated $\text{tRNA}_i^{\text{Met}}$ is not always associated with the oncogenic phenotype.**

By generating more biological replicates of the ribosome profiling data it was hoped the signal-to-noise ratio of the  $\Delta\text{TE}$  values could be improved. Unfortunately when working to generate these replicates it was noticed that pPur and iMet cells were showing similar translation rates, despite the high  $\text{tRNA}_i^{\text{Met}}$  levels being maintained in the iMet cells. After some investigation, the cause of the adaptive response in these cells could not be determined and it was decided to pursue other projects. Soon after, the paper from Marshall et al., was retracted due to evidence that elements of figures were duplicated.

## Chapter 4. mTOR signalling in colorectal cancer

### 4.1 Introduction

The epithelia of both the small intestine and colon are organised into finger-like villi projecting from crypts anchored to the intestinal mucosa. The existence of villi massively increases the surface area of the intestine, aiding absorption from the lumen; loss of villi is associated with a number of pathologies. Intestinal stem cells (ISCs) are located at the base of the crypts and are the origin of all cell types within the crypt and villus (Scheepers et al., 2012). These stem cells can either self-renew or differentiate into one of five epithelial lineages (Gerbe et al., 2011). The first of these is the Paneth cell, that resides adjacent to ISCs and regulate their activity (Sato et al., 2009). The other four lineages are derived from intermediary cells termed “transit amplifying (TA) cells”. TA cells are initially found within the crypts but divide up to 6 times pushing older TA cells up toward the villi (Marshman et al., 2002). Cells become more differentiated as they progress up the crypt-villi axis. Differentiated cells at the top of the villi undergo apoptosis and are lost into the lumen – intestinal epithelial turnover time ranges from 2-5 days depending on the location in the intestine (Creamer et al., 1961).

Colorectal cancer (CRC) is the third most common cancer in the world with 1,234,000 new cases in 2008 (Ferlay et al., 2010). The disease has a characteristic sequelae, involving the accumulation of driver mutations leading to a gradually more aggressive phenotype (Fearon and Vogelstein, 1990). Mutations accumulate due to defects in chromosomal stability (CIN) or mismatch repair (MMR) pathways. In the latter case, loss of MMR proteins results in tumours with extensive microsatellite instability (MSI+) and mutation rates 10-100 fold higher than in CIN+ tumours (Cancer Genome Atlas, 2012). Approximately 3 % of CRC patients have Lynch Syndrome, autosomal dominant mutations in MMR proteins (Lynch et

al., 2009), and in total ~15 % tumours have a MSI+ hypermutated phenotype (Aaltonen et al., 1998).

Irrespective of mutational status, the initial event in colorectal tumourigenesis is typically dysregulated Wnt-signalling (Fearon and Vogelstein, 1990). Wnts are a class of extra-cellular ligand, which upon binding to cell membrane-bound Frizzled receptors trigger the disassembly of a degradation complex that restricts the activity of the pro-proliferative transcription factor  $\beta$ -catenin (Ilyas, 2005). Hyperactive Wnt signalling thus leads to formation of adenomas that contain both ISCs and Paneth cells (Scheper et al., 2012).

Adenomatous polyposis coli (APC) is an essential component of the degradation complex. In 70 % of sporadic tumours the activity of APC is lost due to mutation or deletion (Luchtenborg et al., 2004) increasing the availability of  $\beta$ -catenin. Familial adenomatous polyposis (FAP), an autosomal dominant syndrome with an incidence of 1:10,000 (Jarvinen, 1992), is characterised by loss of function mutations in APC. Loss or mutation(s) in the second allele of APC result in the development of adenomas, and in the absence of colectomy, almost 100 % progression to CRC by the age of 40 (Merg et al., 2005).

Progression from adenomas to invasive adenocarcinomas (cancers) is associated with loss of function mutations in p53 (60 % of MSI-), and gain of function mutations in K-Ras (30 % of MSI- tumour) and PI3-K (18 % of MSI-) (all statistics are taken from (Cancer Genome Atlas, 2012)).

As highlighted in the Introduction, mTOR signalling is frequently upregulated in cancer and CRC is no exception. Evidence from cell lines and human biopsies suggested that mTORC1 signalling was upregulated in CRC (Zhang et al., 2009). Initial work from Dr Owen Sansom's group showed that treatment with rapamycin was sufficient to prevent hyperproliferation in APC null mice. Whilst long term rapamycin inhibition has been reported to reduce mTORC2 formation in some cell lines by sequestering free mTOR (Sarbasov et al., 2006), no decrease

in Akt1 phosphorylation was observed suggesting that mTORC1 but not mTORC2 signalling is crucial in the early stages of CRC development. However, the results of human clinical trials using analogues of rapamycin (rapalogs) have been disappointing (Markman et al., 2010), suggesting a more comprehensive understanding of the interplay between mTOR signalling and other common mutations acquired during CRC development is necessary to guide novel therapies.

Using both mouse and *Drosophila* models of intestinal tumourigenesis we have attempted to characterise the mechanisms by which mTORC1 activity drives intestinal proliferation and to explain the poor performance of rapalogs in human trials.

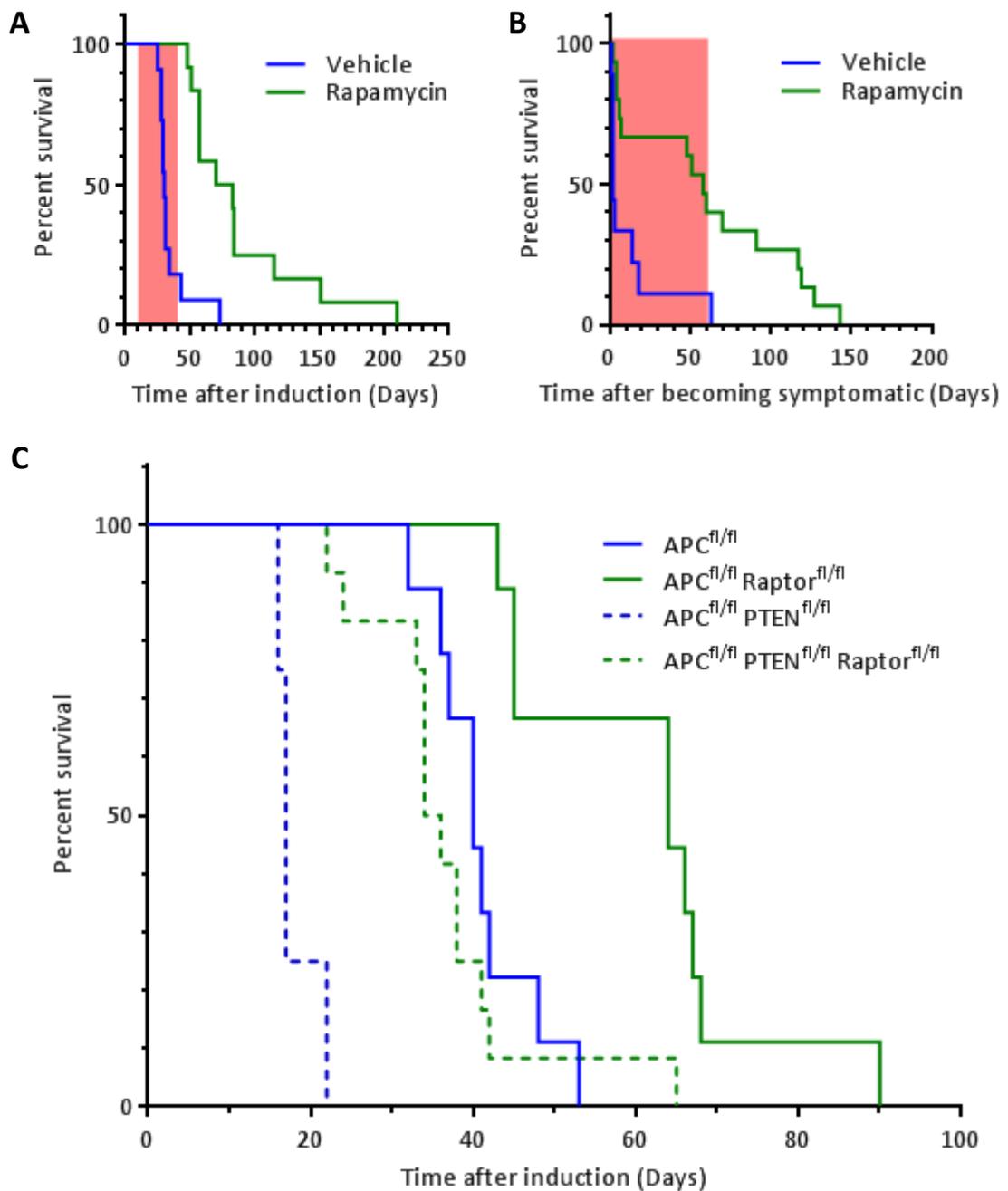
This work was performed in collaboration with Dr William (Liam) Faller and Dr Alessandro Scopelitti, both of Owen Sansom's group at the Beatson Institute, Glasgow. Both have kindly given permission for their work to be presented within this thesis.

#### **4.2 Inhibition of mTORC1 signalling significantly reduces the oncogenic effects of APC deletion in mice.**

The Sansom group have an extensive range of mice strains with alleles either deleted systemically or under the control of a tamoxifen inducible Cre recombinase system. The inducible mice system allows specific knockdown of the gene in either the whole intestinal epithelia (villin) or *Igr5+* cells (ISCs), full details of this can be found in the Materials and Methods (2.3.a). The similarities in cell differentiation and proliferation between the small intestine and the colon make it an appropriate model for the study of intestinal tumours in general.

Homozygous loss of function mutation of APC ( $APC^{fl/fl}$ ) was lethal to the mice with 100 % penetrance. When rapamycin treatment was administered 10 days after induction of APC loss, mice were protected against these lethal effects for the period of treatment (Figure 4-

1A). Extending the treatment for 180 days resulted in comparable survival and phenotype to wild type mice (data not shown). Rapamycin also prolonged life in mice treated once they had developed clinical symptoms (Figure 4-1 B) suggesting that this approach may be clinically relevant. Rapamycin treatment may not inhibit all mTORC1 activity (Thoreen et al., 2009) so a genetic approach to studying the role of mTORC1 was also used. Intestine specific knockout of the mTORC1 essential component Raptor also prolonged life in both APC<sup>fl/fl</sup> and APC<sup>fl/fl</sup> PTEN<sup>fl/fl</sup> genetic backgrounds (Figure 4-1C). That Raptor knockout was not fully protective for these mice (unlike rapamycin treatment) could be because recombination is not 100 % efficient so some residual mTORC1 signalling was present. Consistent with the survival data and the role of Raptor in mTORC1 signalling, crypt size and phosphorylation of mTORC1 targets 4E-BP1 and S6K was markedly reduced in APC<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> mice compared to APC<sup>fl/fl</sup> (Figure 4-2).



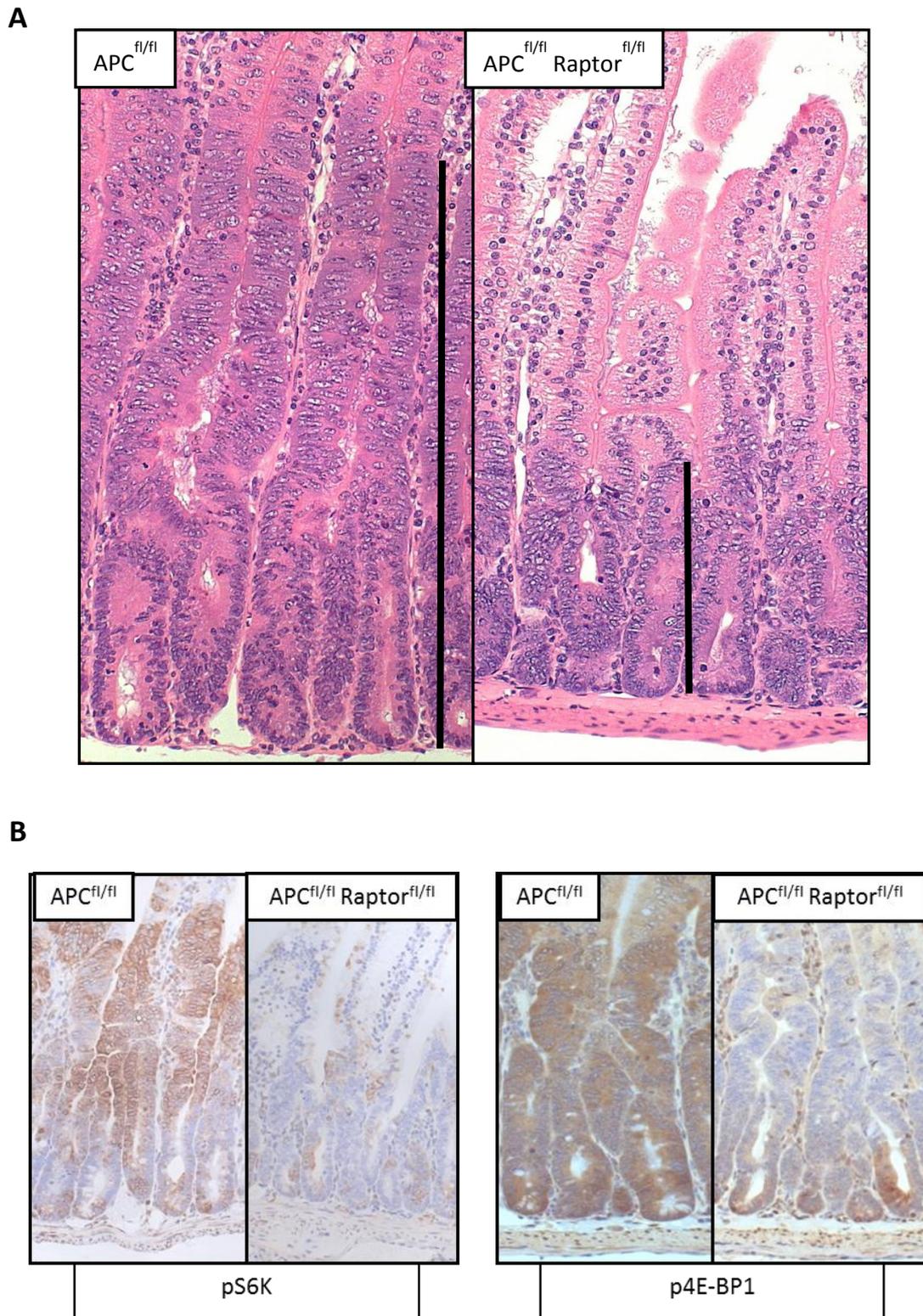
**Figure 4-1 Inhibition of mTORC1 signalling prolongs life of APC deficient mice**

On day 0 mice were administered with tamoxifen to induce Cre recombination to flox out the alleles described. Recombination occurs within the small intestine.

**A** Survival curve for mice with homozygous deletion of APC. 10 days after APC deletion, mice were treated for 30 days with rapamycin or a vehicle control (red region). N = 13 in treatment arm, 14 in vehicle control arm.

**B** As per A, but rapamycin treatment commenced when animals showed signs of illness and continued for 60 days. N = 15 in treatment arm, 9 in vehicle control

**C** Survival curves for mice with or without loss of mTORC1 by deletion of Raptor as described in the legends. N=10 in APC<sup>fl/fl</sup> arm and 9 in APC<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> arm (left). N = 4 in APC<sup>fl/fl</sup> PTEN<sup>fl/fl</sup> arm and 12 in APC<sup>fl/fl</sup> PTEN<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> arm. All experiments in this Figure performed by Dr W. Faller (Beatson Institute, Glasgow). Differences in survival in response to mTORC1 inhibition was statistically significant (p<0.001) in all cases.



**Figure 4-2 Reduction of mTORC1 signalling by Raptor deletion reverses the hyperproliferative phenotype of APC deficiency**

**A** Representative images of haematoxylin and eosin (H&E) stained slices from murine intestines. The black bars indicate the intestinal crypts.

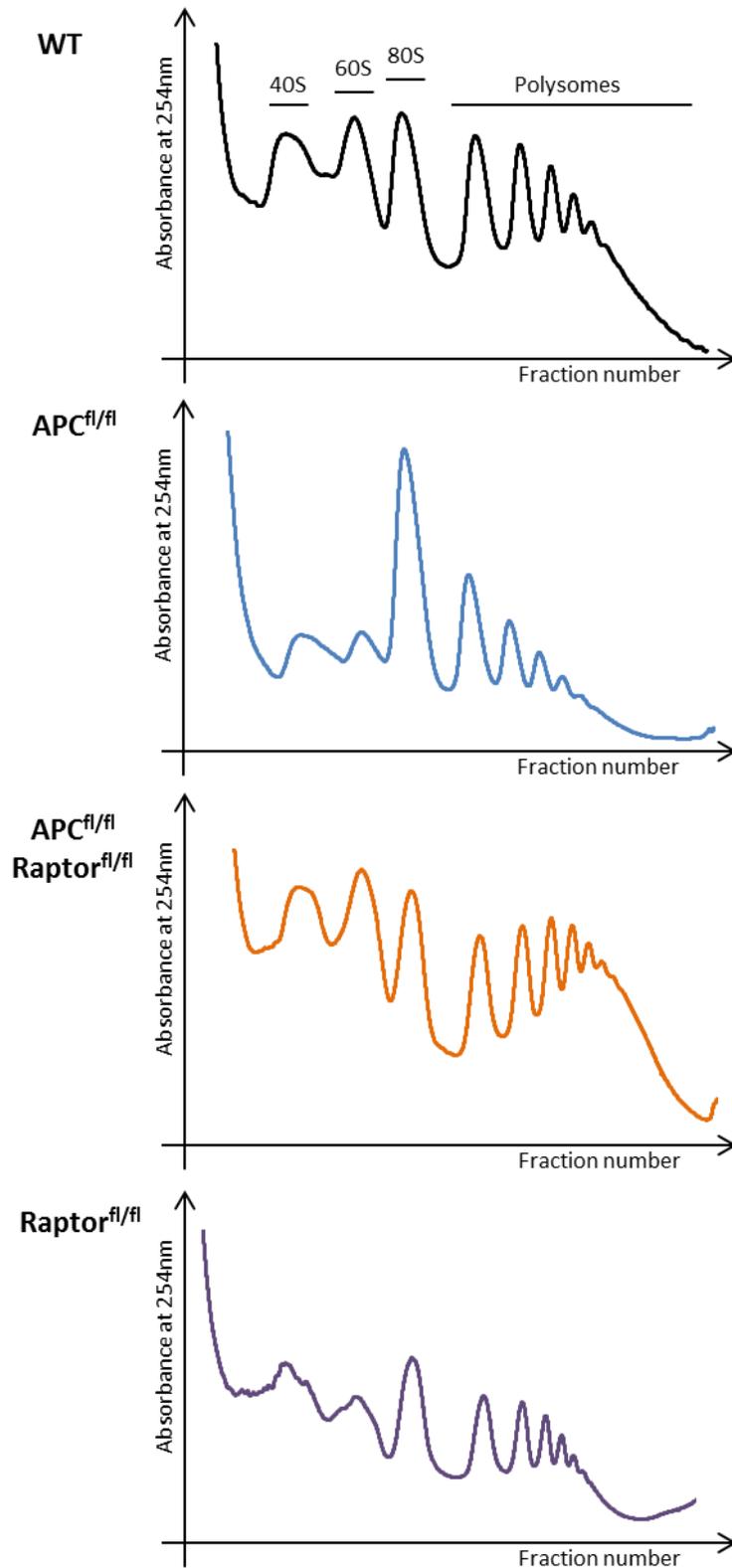
**B** Immunohistochemistry for targets phosphorylated by mTORC1: S6K and 4E-BP1

All experiments in this Figure performed by Dr W. Faller (Beatson Institute, Glasgow)

### **4.3 The effects of upregulated mTORC1 signalling in APC deficient mice are primarily due to changes in translation elongation**

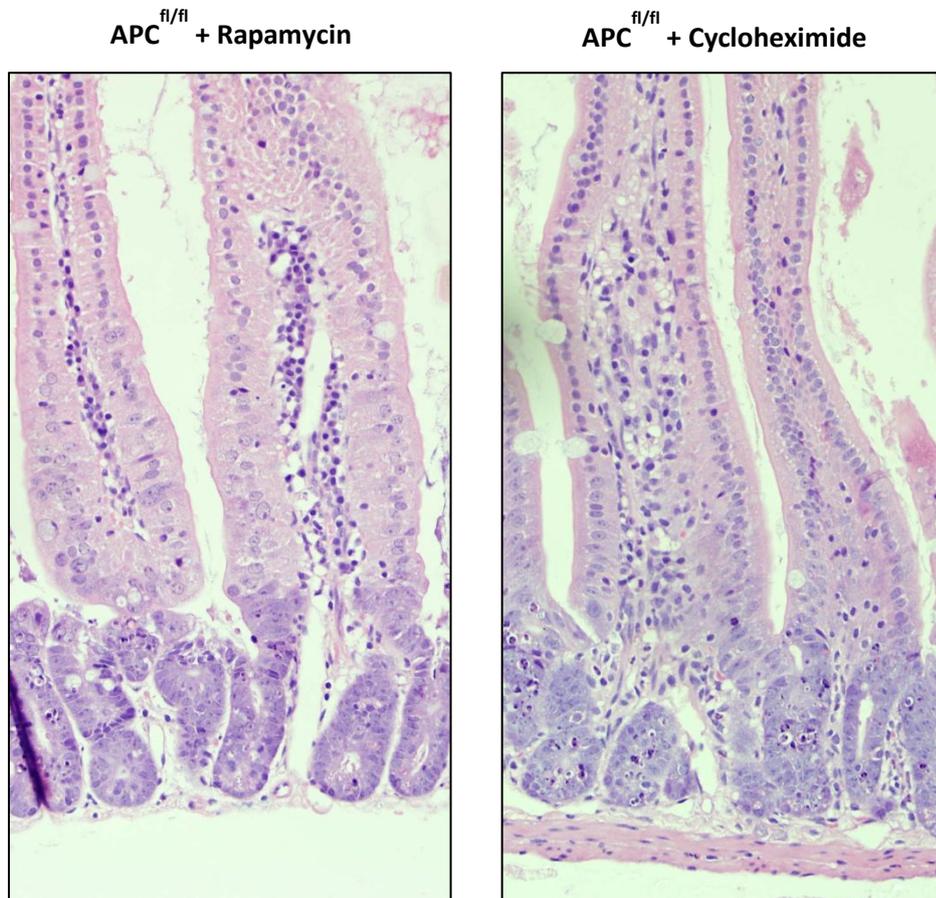
To further investigate the mTORC1 driven changes in translational control in response to APC loss, the crypt epithelial cells were extracted from induced mice and subjected to sucrose gradient analysis. 4-5 mice of each genotype were profiled with a characteristic trace obtained for each (see Figure 4-3). Surprisingly crypt epithelia from APC<sup>fl/fl</sup> mice showed decreased polysomal association relative to the WT mice. Cellular proliferation requires the accumulation of cell mass so an increase in proliferation typically requires an increase in protein synthesis rate (Zetterberg and Killander, 1965). A change in either the distribution of cell types within the crypts or of cell size could uncouple protein synthesis rates from proliferation (e.g. an increase in secretory cells that have lower proliferation rate relative to protein synthesis rates or the replication of smaller cells); neither was observed. In the vast majority of studied cases of translational control, an increase in protein synthesis is due to an increase in initiation and thus causes an increase in polysomal association. The most plausible explanation consistent with a decrease in polysomal association and an increase in protein synthesis is a reduction in ribosome transit time: i.e. APC deficiency is associated with an increase in translation elongation rate. Relative to the WT, there is a marked increase in polysomal association in the APC<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> crypts but not Raptor<sup>fl/fl</sup> suggesting that translation elongation is the locus point of regulation for mTORC1 signalling in the context of APC loss but not in general.

In support of the hypothesis that upregulation of translation elongation is central to Wnt driven proliferation, treatment with low doses of cycloheximide (to lower translation elongation rates) reduces crypt proliferation to the same extent as rapamycin treatment in the APC<sup>fl/fl</sup> mice (Figure 4-4).



**Figure 4-3  $APC^{fl/fl}$  and  $APC^{fl/fl} Raptor^{fl/fl}$  intestinal epithelia have altered polysomal association compared to WT**

Representative polysome traces from epithelial extracts from mice with the genotypes listed. Due to differences in the mass of material obtained from different mice the traces have been scaled to be comparable to each other.



**Figure 4-4 Low dose cycloheximide mimics the effect of rapamycin treatment in APC<sup>fl/fl</sup> intestines**

H&E stained intestinal sections from APC<sup>fl/fl</sup> mice treated with either rapamycin or low doses of cycloheximide as described in the Materials and Methods.

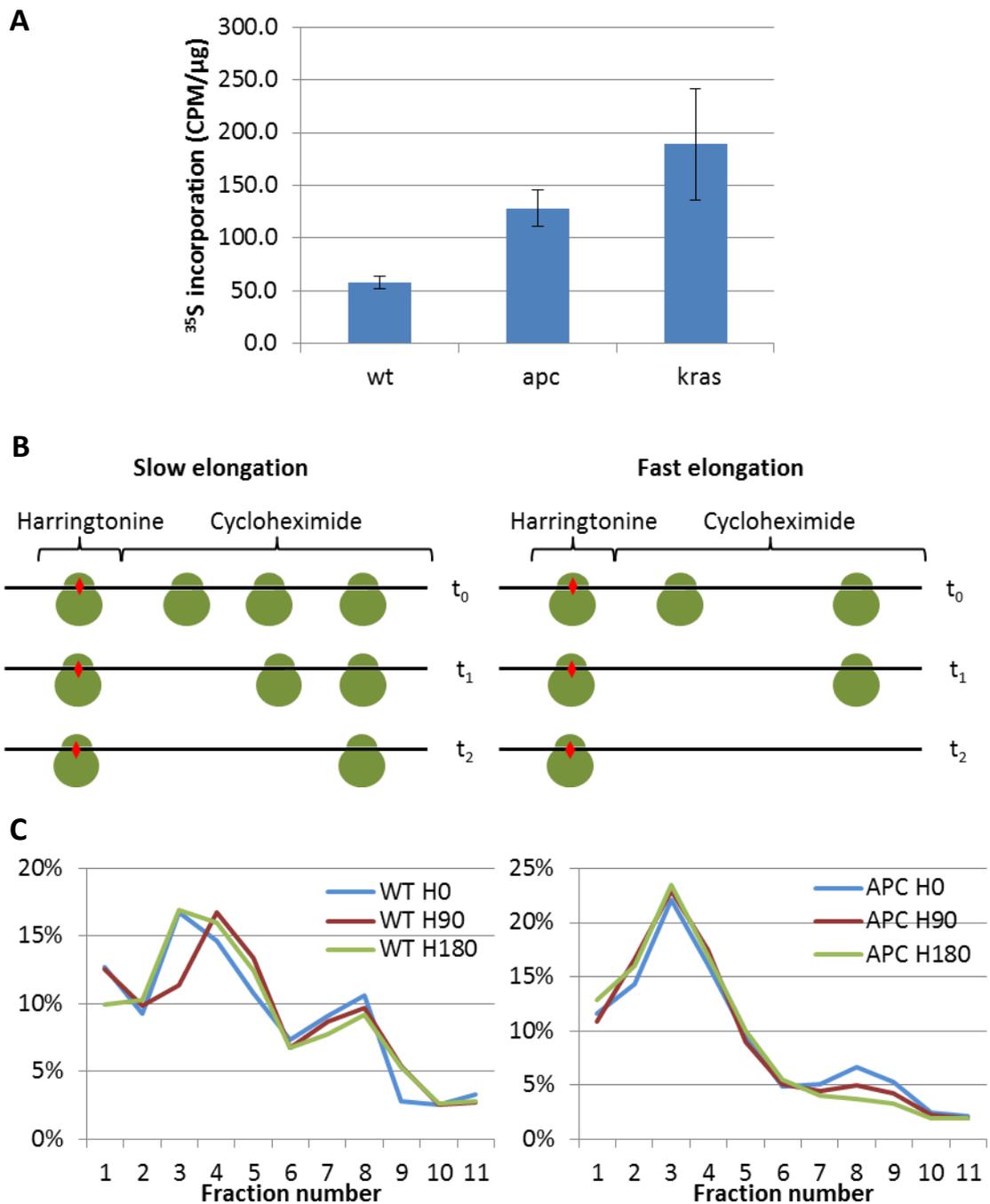
All experiments in this Figure performed by Dr W. Faller (Beatson Institute, Glasgow)

However, the effect of cycloheximide treatment on crypt proliferation may simply be due to a decrease in protein synthesis rather than directly implicating translation elongation, therefore it was necessary to determine the translation elongation rates within the intestine. Performing such investigations within an *in vivo* model was too technically challenging and costly, so an *in vitro* Matrigel culture model was used. A full description of this system can be found within the Materials and Methods (2.2.b) but essentially crypt epithelia (WT) were extracted from mice and dispersed within a layer of Matrigel. When cultured in media containing a range of growth factors crypt stem cells differentiate to form organoids – proliferating crypts that invade upwards into the Matrigel to form crypts with the same morphology as is seen *in vivo*. Cre recombinase was induced in some WT cultures to give matched APC<sup>fl/fl</sup> crypts. Unfortunately cells derived from the APC<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> mice did not culture under these conditions, but this system did allow investigation of the changes in response to APC deficiency.

Protein synthesis rates are higher in the APC<sup>fl/fl</sup> crypts compared to wild type (Figure 4-5A), consistent with the assumption that increased protein synthesis facilitates cell proliferation.

Assessment of elongation rates is somewhat more challenging, in part because ribosome transit times are so short (half ribosome transit times are typically in region of 2 min (Ruvinsky et al., 2005)). The classical radioincorporation based assay requires large numbers of replicates and time points to generate accurate regression lines of the rate of incorporation into the total and released pools of protein (Fan and Penman, 1970). Slight differences in gradients of these lines have dramatic impacts on the x-axis intercept and subsequent transmit time estimates. Instead we chose to perform a harringtonine run-off assay. This exploits the ability of harringtonine to inhibit the elongation of newly initiated ribosomes but not those already elongating (Fresno et al., 1977). Treatment with cycloheximide (which inhibits all elongating ribosomes) at fixed time points after

harringtonine is added gives an indication of the rate of ribosomal transit (see Figure 4-5B for illustration). The mass of material obtained from crypt cultures was considerably smaller than from the whole mouse so a well resolved polysome profile could not be obtained, but the RNA extracted from each fraction of the gradient could be reliably quantified. Similar to the *in vivo* observations, RNA was more polysomally associated in WT crypts than the APC<sup>fl/fl</sup> ones (Figure 4-5C). Furthermore the reduction in polysomal association during the harringtonine run-off was much greater in the APC<sup>fl/fl</sup> crypts implying the translation elongation rate is higher in these crypts.



**Figure 4-5 Protein synthesis and translation elongation rates are elevated in cultured APC<sup>fl/fl</sup> crypts**

**A** [<sup>35</sup>S]-Met label incorporation from WT, APC<sup>fl/fl</sup> and APC<sup>fl/fl</sup> KRas<sup>G12D/+</sup> crypts. 30min pulse. Data are mean of 2 independent wells +/- SD.

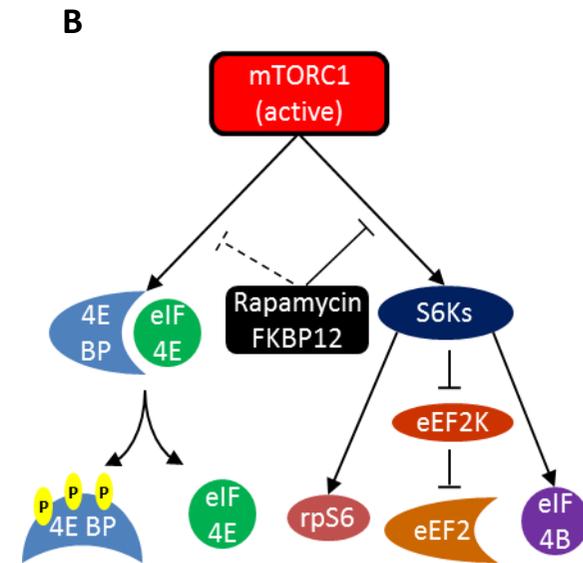
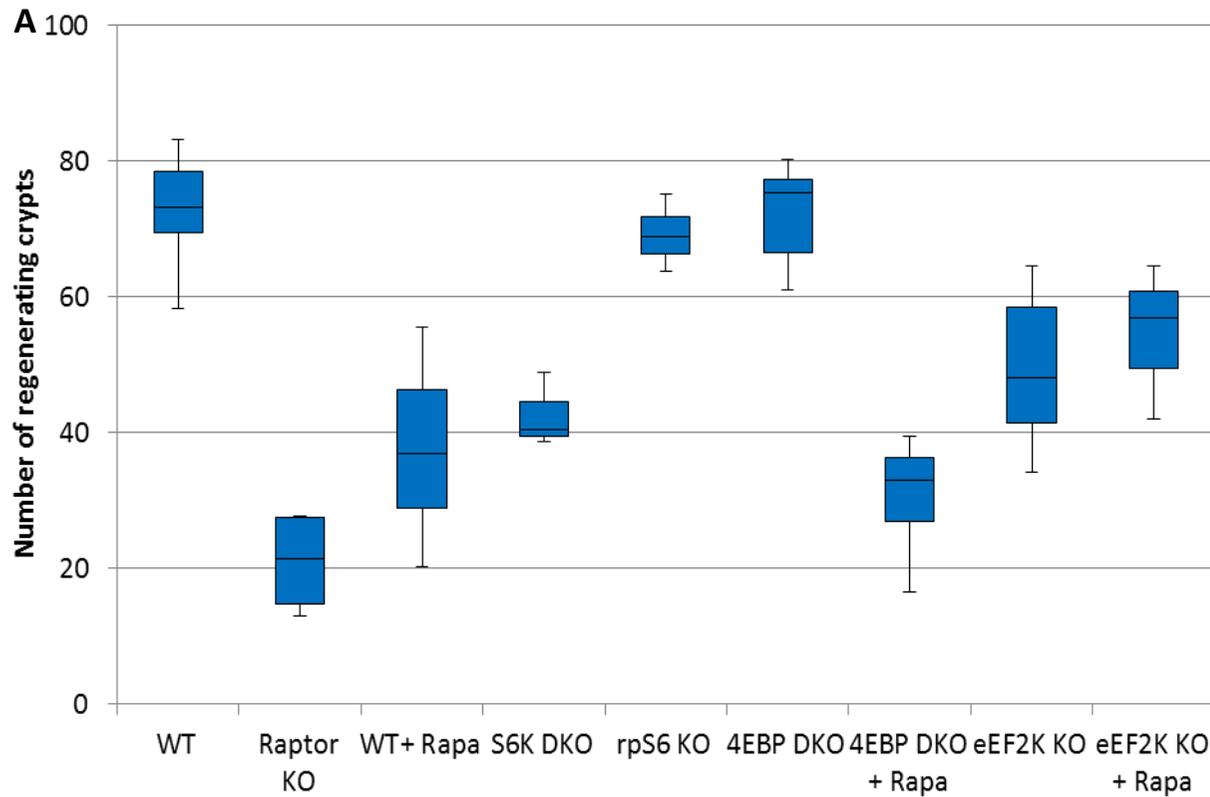
**B** A schematic representing the ribosome association of a theoretical mRNA after treatment with harringtonine and then cycloheximide at set time points afterwards. The AUG start site is indicated with a diamond.

**C** Polysomal association of RNA represented as the % of the total RNA isolated from each fraction of the sucrose density gradient. H0 traces were from crypts treated with harringtonine and cycloheximide simultaneously; H90: 90s between addition of harringtonine and cycloheximide; H180: 180s gap.

Having obtained evidence suggesting APC deficiency drives proliferation due to an increase in translation elongation it was important to confirm that this was mediated through the well described mTORC1 signalling pathways, using the *in vivo* models of the Sansom laboratory.

Survival studies are exceptionally time and resource intensive so an alternative was used: intestinal regeneration after IR irradiation. Regeneration under these conditions is driven by Wnt signalling and thus analogous to APC deficiency (Ashton et al., 2010).

Treatment of WT mice with rapamycin caused a significant reduction in intestinal regeneration compared to the vehicle treated controls (Figure 4-6 A). A similar result was obtained in Raptor<sup>fl/fl</sup> mice, suggesting that mTORC1 is required for intestinal regeneration. As described previously mTORC1 has two well characterised targets, 4E-BPs and S6K, each with their own downstream targets (Figure 4-6B). Mice with systemic homozygous deletion of both 4E-BP1 and 4E-BP2 (4E-BP3 is reported to not be expressed in MEFs (Dowling et al., 2010a) and we assume the same is true in the murine intestine) showed no difference in the capacity to regenerate compared with WT mice. 4E-BP DKO mice retain the same degree of sensitivity to rapamycin as WT. This suggests rapamycin inhibition of Wnt-dependent proliferation is mediated by activation of S6K and not inhibition of 4E-BPs. Consistent with this, mice with systemic homozygous deletion of both S6K isoforms have significant reduced intestinal regeneration, comparable to that observed with rapamycin treatment. Rapamycin treatment does not reduce regeneration as much as Raptor<sup>fl/fl</sup>, so it is possible that rapamycin does not fully inhibit mTORC1 signalling to 4E-BPs. However, the comparative difference between Raptor<sup>fl/fl</sup> and rapamycin treated mice suggested that proliferation is more dependent on the S6K arm of the mTORC1 signalling pathway than the 4EBP arm.



**Figure 4-6 Intestinal regeneration is sensitive to mTORC1 inhibition and requires the inhibition of eEF2k**

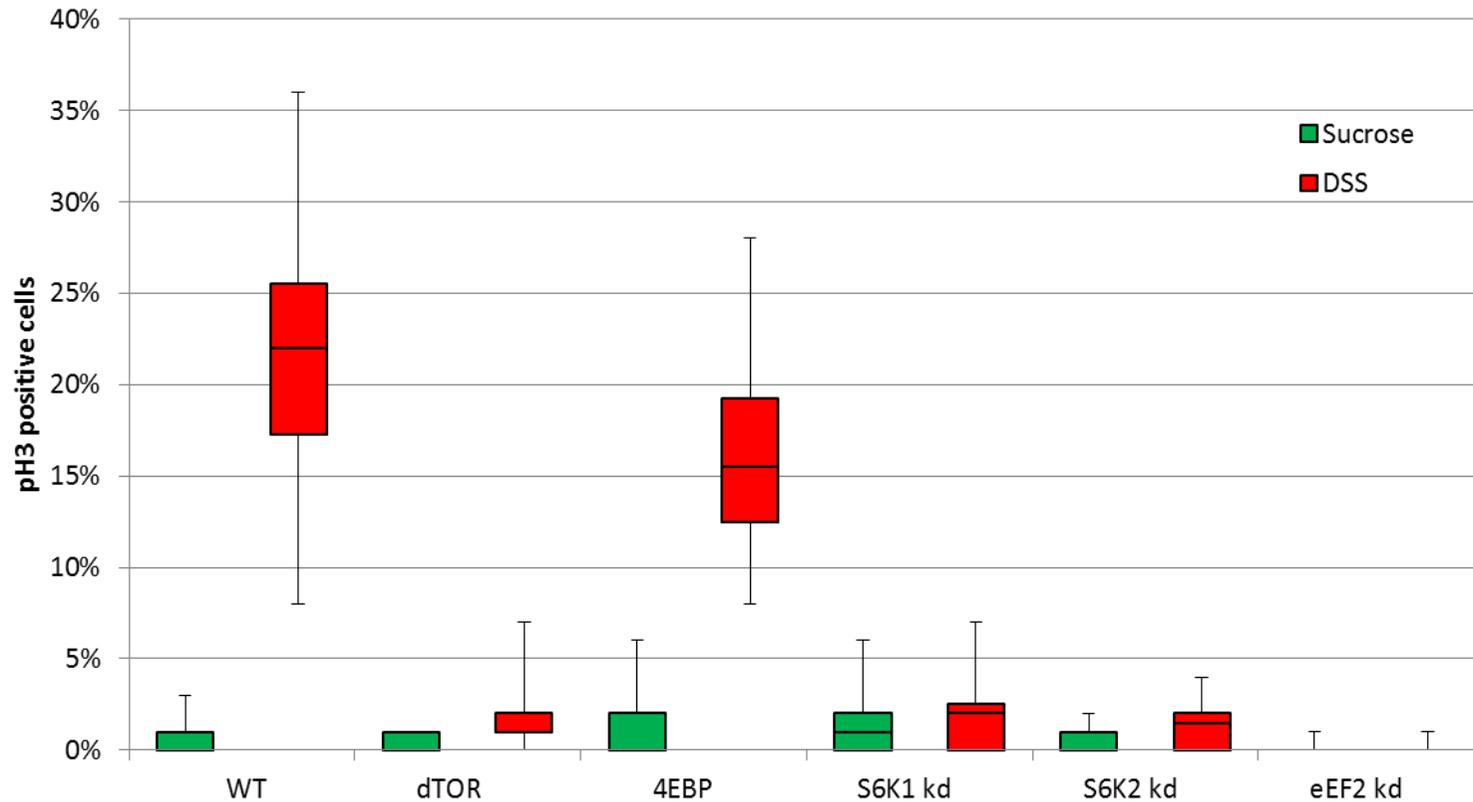
**A** Box and whisker blots of number of regenerating crypts in a murine intestine 72 hours after IR irradiation. Numbers are derived from 10 fields of view, n= 16, 4, 6, 3, 4,5,4,11,3 (left to right) with the box representing the interquartile range and the whiskers the full range of the data. All groups (other than WT) describe floxed gene(s) in the mouse as detailed in the Materials and Methods. All experiments in this Figure performed by Dr W. Faller (Beatson Institute, Glasgow)

**B** Representation of some of the downstream effectors of mTORC1 signalling and how they mediate their effects on translation

S6K has numerous downstream targets, which affect different cellular processes. Mice expressing a mutant form of RPS6 that cannot be activated by S6K-mediated phosphorylation were able to regenerate as efficiently as WT. Despite being used as a marker of mTORC1 activation, it would appear that RPS6 does not mediate intestinal regeneration. Interestingly a whole mouse knockout of eEF2k did show a slight reduction in regeneration. It is unclear why this occurs; mTORC1 negatively regulates eEF2k so loss of eEF2k might be expected to give a similar result to the WT (where mTORC1 signalling is active). Irrespective, the intestinal regeneration of eEF2k<sup>-/-</sup> mice is not reduced by inhibition of mTORC1 by rapamycin. Other targets of S6K were not investigated so may play some role. However the complete rapamycin insensitivity of the eEF2k<sup>-/-</sup> mice in addition to the observations from the Matrigel crypt culture system strongly suggests that Wnt-dependent proliferation requires increased translational elongation mediated through mTORC1.

To further confirm this, similar experiments were performed by Dr Alessandro Scopelitti in a *Drosophila* system. The crypt morphology and signalling pathways in the *Drosophila* intestine are very similar to those in mammals (Apidianakis and Rahme, 2011). Furthermore, the Sansom group have shown that intestinal damage with dextran sodium sulphate (DSS) results in the activation of Wingless (Wg), the *drosophila* Wnt homolog (Cordero et al., 2012). The results observed in *Drosophila* are fully consistent with those seen in the mouse (Figure 4-7). Expression of a dominant negative form of dTOR prevents intestinal regeneration, suggesting that TOR signalling is required for Wnt/Wg mediated proliferation. Unlike mammals, *Drosophila* have only one 4E-BP. If repression of 4E-BP by dTORC1 is required to mediate intestinal regeneration then overexpression of 4E-BP should significantly reduce proliferation, but this was not observed. By contrast, RNAi knockdown of either of the S6Ks prevented intestinal regeneration. Finally, and crucially, RNAi knockdown of the downstream target of eEF2k, eEF2, prevented intestinal regeneration. Collectively the data

presented in this section strongly suggest that mTORC1 mediates Wnt-driven proliferation in the intestine via the upregulation of translation elongation.



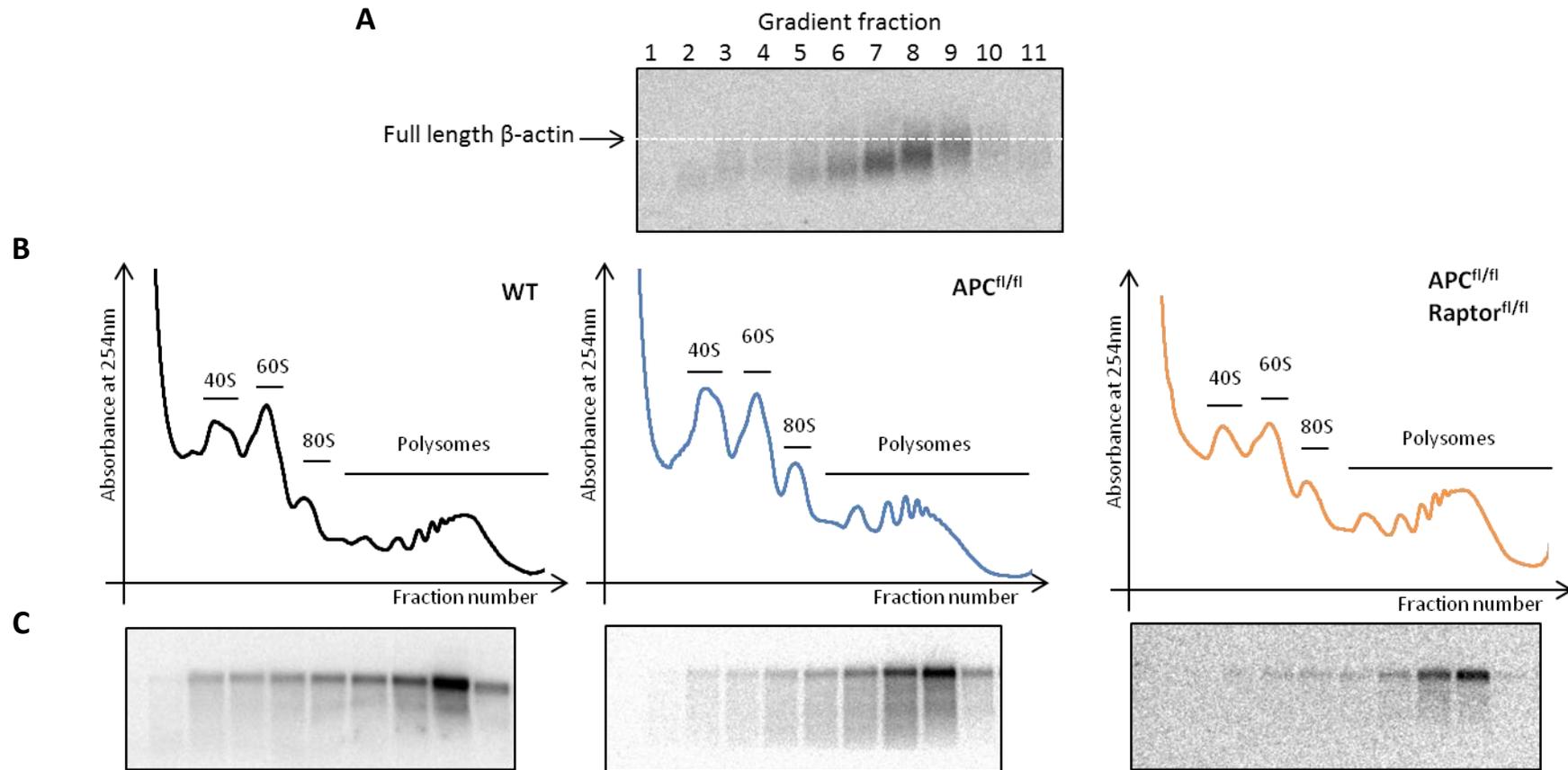
**Figure 4-7 Intestinal regeneration in Drosophila also requires mTORC1 to increase the activity of downstream eEF2**

Box and whisker plot of the number of phospho-histone3+ cells in a drosophila intestine after damage with DSS. pHH3 is a marker of regeneration. dTOR = overexpression of dominant negative dTOR; 4EBP = over expression of 4EBP; S6K2 = RNAi to S6K2; S6K1 = RNAi to S6K1 eEF2 = RNAi to eEF2. n= 20, 14, 9, 9, 9, 8, 13, 15, 9, 10, 9, 6 (left to right). Experiments performed by Dr A. Scopeletti (Beatson Institute)

#### 4.4 Improvements in intestinal crypt isolation

To further investigate the role of mTORC1 in APC deficiency it would have been informative to identify if any mRNAs were differentially translated under these conditions using RNA from the polysome profiles of the murine intestinal epithelia. However, northern blotting with a probe complementary to a region of the highly abundant transcript  $\beta$ -actin showed evidence of RNA degradation within these samples (Figure 4-8A). The probe hybridised to RNA of the correct size in heavy polysomes fractions, but progressively shorter species in lighter fractions. This suggests that there was degradation of mRNA between adjacent ribosomes (somewhat analogous to the process for generating ribosome protected fragments). This was observed in samples from all mice genotypes strongly suggesting the source was the method of crypt isolation and/or lysis conditions.

The original protocol involved approximately 15 minutes of incubation in warmed media to detach epithelial cells after the intestines were flushed with PBS-CHX. Endogenous RNases will function optimally under these conditions so are a plausible cause of the degradation. To address this, the extractions were repeated with a method that involved brief treatment in warmed media containing cycloheximide to facilitate uptake of the drug and rapid translational inhibition, with subsequent epithelial extraction performed by brief, but vigorous, shaking in ice cold PBS-CHX as described in 2.3.gii). Cell pellets were lysed in a buffer containing additional RNase inhibitors to supplement the action of heparin in the standard buffer. The polysome profiles obtained under these conditions differ from those from the original method, showing a greater degree of polysomal association (Figure 4-8B). Importantly, the RNA of WT and APC<sup>fl/fl</sup> Raptor<sup>fl/fl</sup> epithelia were associated with heavier polysomes than in a APC<sup>fl/fl</sup> background, consistent with the initial observations. Reassuringly, northern blots for  $\beta$ -actin showed no signs of the degradation seen in the original samples (Figure 4-8C).



**Figure 4-8 Improved method of intestinal epithelial extraction**

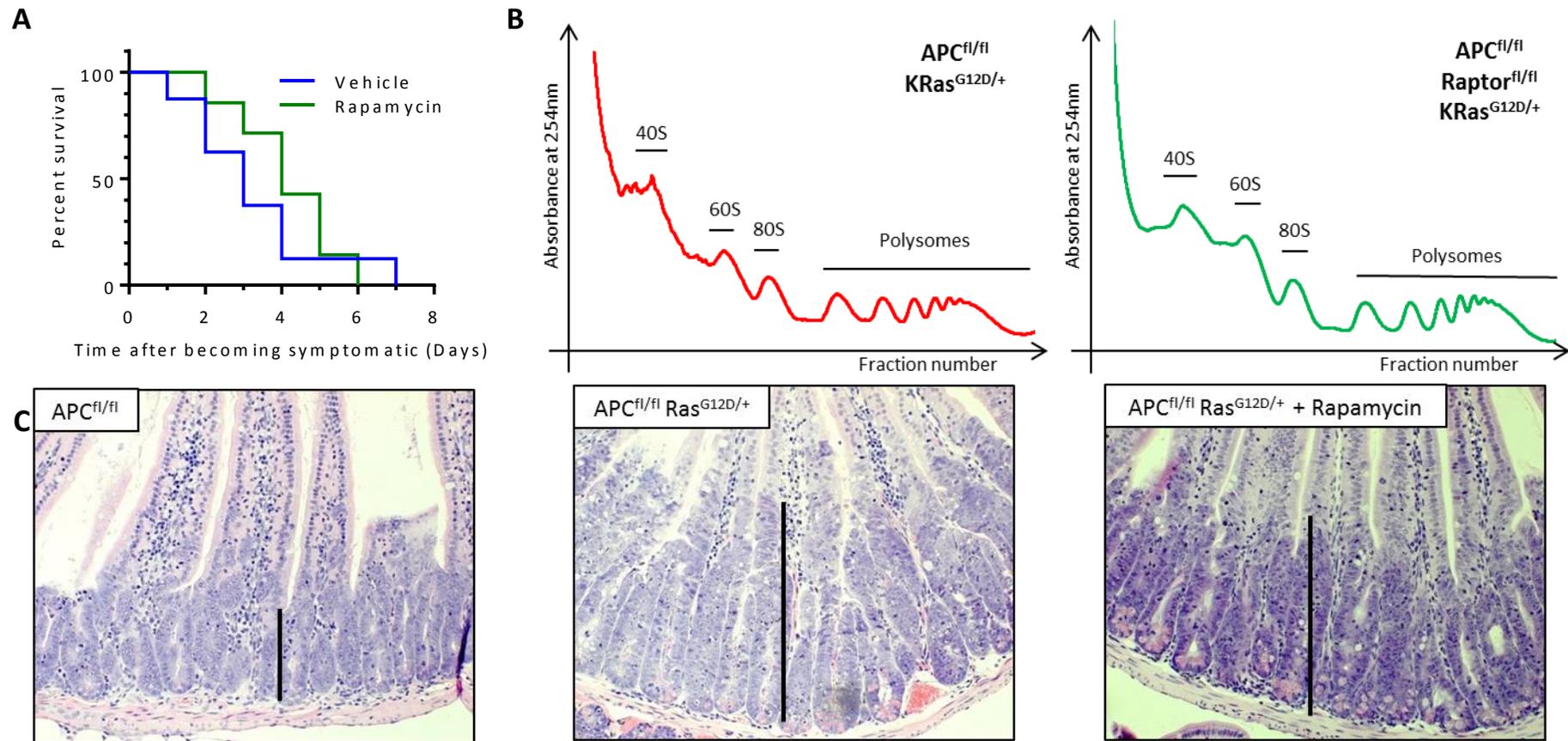
A Northern blot for  $\beta$ -actin on gradient fractions taken from colonic epithelia of a WT mouse. Similar blots were obtained for all other genotypes  
 B Polysome profiles of epithelial cells obtained using a method that minimises the potential for RNase mediated degradation  
 C Northern blots for  $\beta$ -actin on the gradient fractions corresponding to the traces above.

#### 4.5 K-Ras mutation results in adenomas dependent on mTORC2 signalling rather than mTORC1

As mentioned in the introduction to this chapter, in spite of the promising data from *in vivo* mouse models the use of rapalogs in the treatment of colorectal cancers has been disappointing. It has been suggested the loss of the inhibitory feedback loop of S6K on PI3K signalling may be the cause of this (Carracedo and Pandolfi, 2008). However, mice which have homozygous PTEN deletion are also responsive to rapamycin therapy (Figure 4-1B) despite PTEN being a negative regulator of Akt1 (a key effector of PI3K signalling). An alternative hypothesis is that other genetic changes acquired during tumourigenesis may be responsible for resistance to mTORC1 inhibition. This is also a plausible explanation of the poor performance of clinical trials as the patients recruited are often those resistant to existing therapies.

A common change associated with the progression of early adenomas to more proliferative late adenomas and adenocarcinomas is mutation of K-Ras (Fearon and Vogelstein, 1990). Mutation results in constitutive activation which may explain why therapeutic inhibition of upstream EGF signalling lacks efficacy in patients with Ras-mutant tumours (Preneen et al., 2010). In addition K-Ras activates PI3K, driving signalling independent of the S6K feedback loop.

Rapamycin treatment of mice with the heterozygous G12D mutation of K-Ras in addition to APC deficiency does not prolong survival (Figure 4-9A). These mice have enlarged crypts relative to APC<sup>fl/fl</sup> mice which were not decreased by rapamycin treatment (Figure 4-9C). Finally, no difference in polysomal association between APC<sup>fl/fl</sup> KRas<sup>G12D/+</sup> and APC<sup>fl/fl</sup> KRas<sup>G12D/+</sup> Raptor<sup>fl/fl</sup> mice was observed (Figure 4-9B)



**Figure 4-9 Addition of K-Ras mutation to APC deficiency removes sensitivity to mTORC1 inhibition**

**A** Survival curve for mice with  $APC^{fl/fl}$   $K-Ras^{G12D/+}$  intestines with or without rapamycin treatment. Dr W. Faller (Beatson Institute, Glasgow)

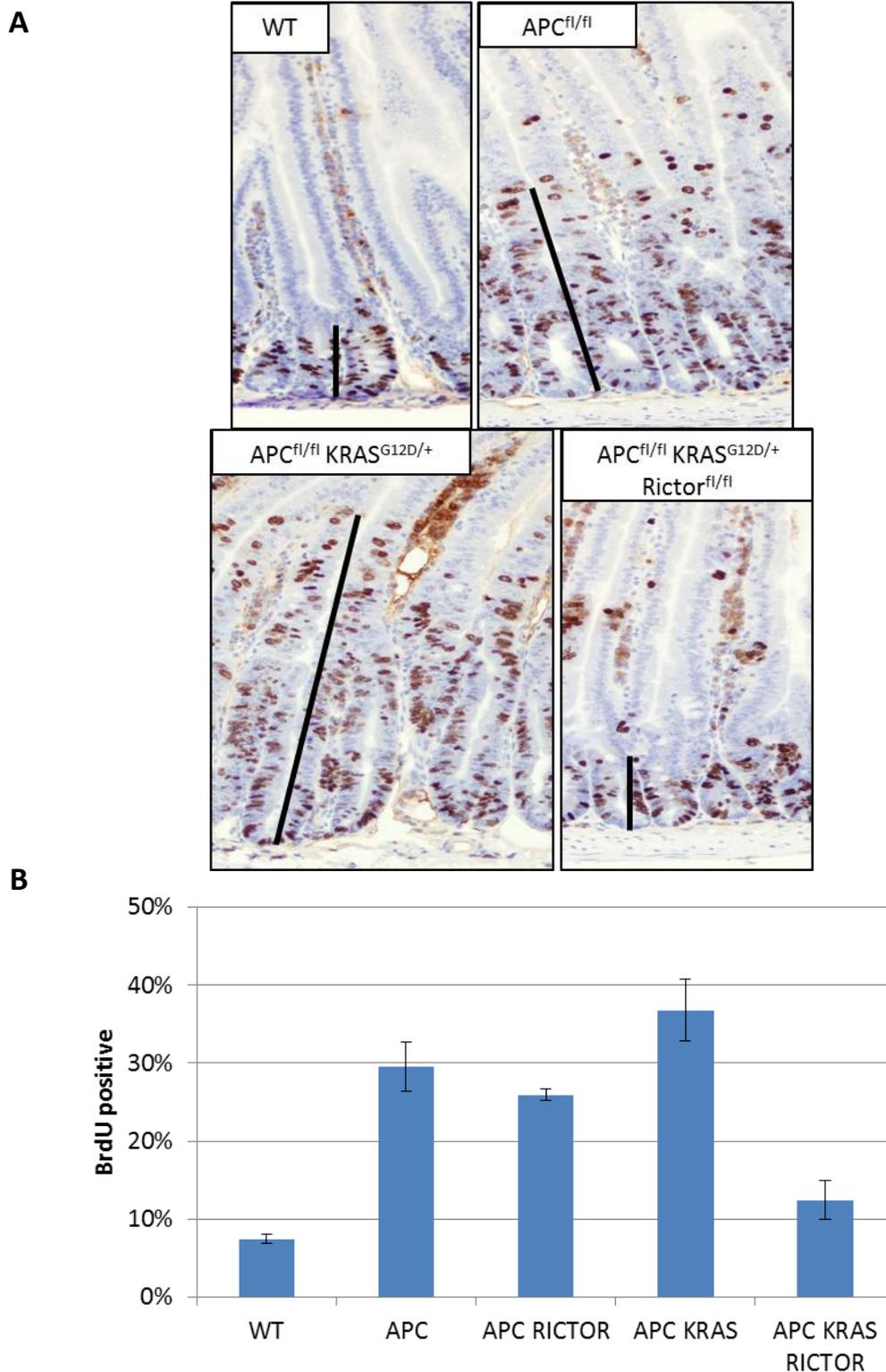
**B** Polysome profiles of intestinal epithelia with indicated genotypes

**C** H&E stained intestinal slices from the indicated genotypes, right panel shows a mouse treated with rapamycin. Dr W. Faller (Beatson Institute, Glasgow)

Treatment of K-Ras mutant mice with pp242, to inhibit both mTORC1 and mTORC2, did result in a modest reduction in crypt size but not even to APC<sup>fl/fl</sup> levels (data not shown). This suggests that the combination of APC loss and K-Ras mutation makes the tumour dependent on mTORC2 signalling, but that either the pp242 dose was not sufficient to fully reverse the phenotype or that targeting both mTORC1 and mTORC2 simultaneously is counterproductive. Intriguingly, treatment of APC<sup>fl/fl</sup> or APC<sup>fl/fl</sup> PTEN<sup>fl/fl</sup> mice with pp242 did not improve survival and is under investigation within the Sansom group.

To investigate the importance of mTORC2 signalling alone, rather than both complexes as occurs when using pp242, mice with inducible homozygous deletion of Rictor were bred.

The percentage of proliferating cells with crypts of APC<sup>fl/fl</sup> mice was unaffected by Rictor deletion (Figure 4-10). As expected, there was increased proliferative activity in the crypts of APC<sup>fl/fl</sup> KRas<sup>G12D/+</sup> mice compared to the APC<sup>fl/fl</sup> mice. APC<sup>fl/fl</sup> KRas<sup>G12D/+</sup> Rictor<sup>fl/fl</sup> mice showed significant reduction in proliferative activity to levels comparable to WT crypts. The morphology of crypts from these mice was exceptionally similar to those from WT mice suggesting that normal intestinal function was present. Survival studies for these mice are on-going, but it appears that whilst mTORC1 signalling is required in early adenoma formation, K Ras mutation results in late adenomas that are dependent almost exclusively on Rictor. Rictor has been described to act outside of mTORC2, most notably in complex with integrin-linked kinase (ILK) that can phosphorylate Akt1 on S473 (i.e. the same site as mTORC2) (McDonald et al., 2008). However, increased expression of ILK is an event observed in benign polyps from FAP patients (Marotta et al., 2001) as well as carcinomas (Marotta et al., 2003). That Rictor deletion does not reduce proliferation in the APC<sup>fl/fl</sup> mice suggests it unlikely that Rictor is predominantly acting via ILK in the intestine. In the absence of confounding data, it was presumed that Rictor-mediated effects in KRas mutant tumours were via mTORC2.



**Figure 4-10 K-Ras mutation creates tumours that are dependent on mTORC2 signalling.**

**A** Immunohistochemistry for BrdU incorporation over 2 hr within the intestines of the genotypes indicated.

**B** 10 fields of view were quantified to give the % of BrdU positive cells in each of the indicated genotypes. Values are mean +/- SD n=3, except for APC<sup>fl/fl</sup> KRAS<sup>G12D/+</sup> where n=5. All experiments in this Figure by Dr W. Faller (Beatson Institute, Glasgow)

## 4.6 Discussion

In this chapter it was shown the mTORC1 signalling is necessary for the hyperproliferative phenotype associated with APC deficiency.

Given that many recent reviews highlight the role of 4E-BPs and translational initiation in mTORC1 signalling (Dowling et al., 2010b, Laplante and Sabatini, 2012, Populo et al., 2012) a primary role for translation elongation in APC deficiency may appear rather surprising.

Perhaps it should not be. In addition to eEF2k being a target of S6K, two recent papers that have used ribosome profiling in cultured cells find the eEF2 is in the top 3 most differentially translated mRNAs in response to modulations of mTORC1 activity (Hsieh et al., 2012, Thoreen et al., 2012). That Thoreen et al., report that inhibition of mTOR signalling with Torin has no effect on the translation rate of 4E-BP DKO MEFs suggests that the contribution of the 4E-BP and S6K arms of mTORC1 signalling may differ between cellular contexts.

Work in cultured cell lines suggests that resistance to active-site mTOR inhibitors (asTORi) can be acquired through downregulation of 4E-BP expression thus increasing eIF4E availability (Alain et al., 2012). Similarly it is possible that treatment with rapamycin could drive selection of tumour cells with upregulation of eEF2, but APC<sup>fl/fl</sup> mice given long term rapamycin maintenance therapy are phenotypically normal so this seems unlikely.

Furthermore, when mice with heterozygous mutation of APC (APC<sup>Min/+</sup>: analogous to the genetic status of FAP patients) were bred until they acquired the second-hit mutation and became sick, rapamycin treatment caused regression of these adenomas (W. Faller, unpublished observations). These observations may offer a potential therapy for FAP kindreds. Given that the existing management of people with FAP is regular (invasive) colonoscopies and eventual bowel resection, a pharmacological therapy would dramatically improve the quality of life for these families.

However, when patients present with sporadic CRC their tumour will typically have acquired more genetic changes than just APC mutation. In the *in vivo* mouse model, the combination of APC deficiency and K-Ras mutation conferred resistance to mTORC1 inhibition consistent with similar findings in human colorectal cancer cell lines (Di Nicolantonio et al., 2010). K-Ras mutated tumours are resistant to other standard chemotherapeutics (Preneen et al., 2010) and genetic testing for K-Ras mutation has recently been approved for the management of colorectal cancer<sup>3</sup> so these patients can be advised to take part in trials of novel therapies tailored to their molecular phenotype. Combination therapies targeting both MEK (a downstream target of Ras) and the PI3K/mTOR pathway have resulted in partial responses for a subset of Ras-mutated human advanced cancers (Shimizu et al., 2012). A larger trial treating patients with K-Ras mutated tumours with drugs targeting these pathways is ongoing<sup>4</sup>.

However another centre found that matching colorectal cancer patients resistant to conventional therapies to clinical trials on the basis of their molecular profile provided no benefit (Dienstmann et al., 2012). There are three non-mutually exclusive explanations for this. The first is that the doses used in these trials were not biologically effective (this in part may be due to overlapping toxicity profiles limiting the use of full doses of combination therapies). The second is that patients were incorrectly matched due to deficiencies in the methods used to characterise the molecular phenotype. Third, patients were incorrectly matched because of an insufficient understanding of how different mutations affect interactions between different signalling pathways.

Our data address the third of these concerns. That inhibition of mTORC1 does not protect mice with K-Ras mutated tumours and that treatment with dual mTORC1/2 inhibitors is less effective than genetic deletion of an mTORC2 specific component suggest selective inhibition

---

<sup>3</sup> [http://www.accessdata.fda.gov/cdrh\\_docs/pdf11/p110030a.pdf](http://www.accessdata.fda.gov/cdrh_docs/pdf11/p110030a.pdf)

<sup>4</sup> <http://clinicaltrials.gov/ct2/show/NCT01347866?term=mTOR+K-ras&rank=2>

of mTORC2 may be a more effective approach. Additionally, a monotherapy may reduce the toxicity profile of the treatment thus increasing efficacy. To date no selective pharmacological inhibitor of mTORC2 has been developed, but delivery of a siRNA to Rictor, targeted to CRC cells may have promise (see (Kong et al., 2012) for a review of this approach).

Finally, determining whether other genetic changes found in some K-Ras mutated colorectal tumours (eg loss of p53 or PTEN) modulate the dependency on mTOR signalling would be invaluable in determining the scope for inhibition of mTORC2 as a treatment strategy.

## Chapter 5. Translation elongation as a component of the cold shock response

### 5.1 Introduction

Culturing cells at hypothermic temperatures has appeal for the industrial production of recombinant proteins. Superficially, lower temperatures are cheaper to maintain, but more significantly the yield of some recombinant proteins is increased in cells cultured in mild hypothermic conditions ( $\sim 30^{\circ}\text{C}$ ), despite a reduction of total protein synthesis rates (Al-Fageeh et al., 2006). This increased yield is not observed for all recombinant proteins. The characterisation of the mechanisms associated with cold stress (the cold shock response), would therefore be helpful in designing protein expression platforms that will maximise the yield of any given protein.

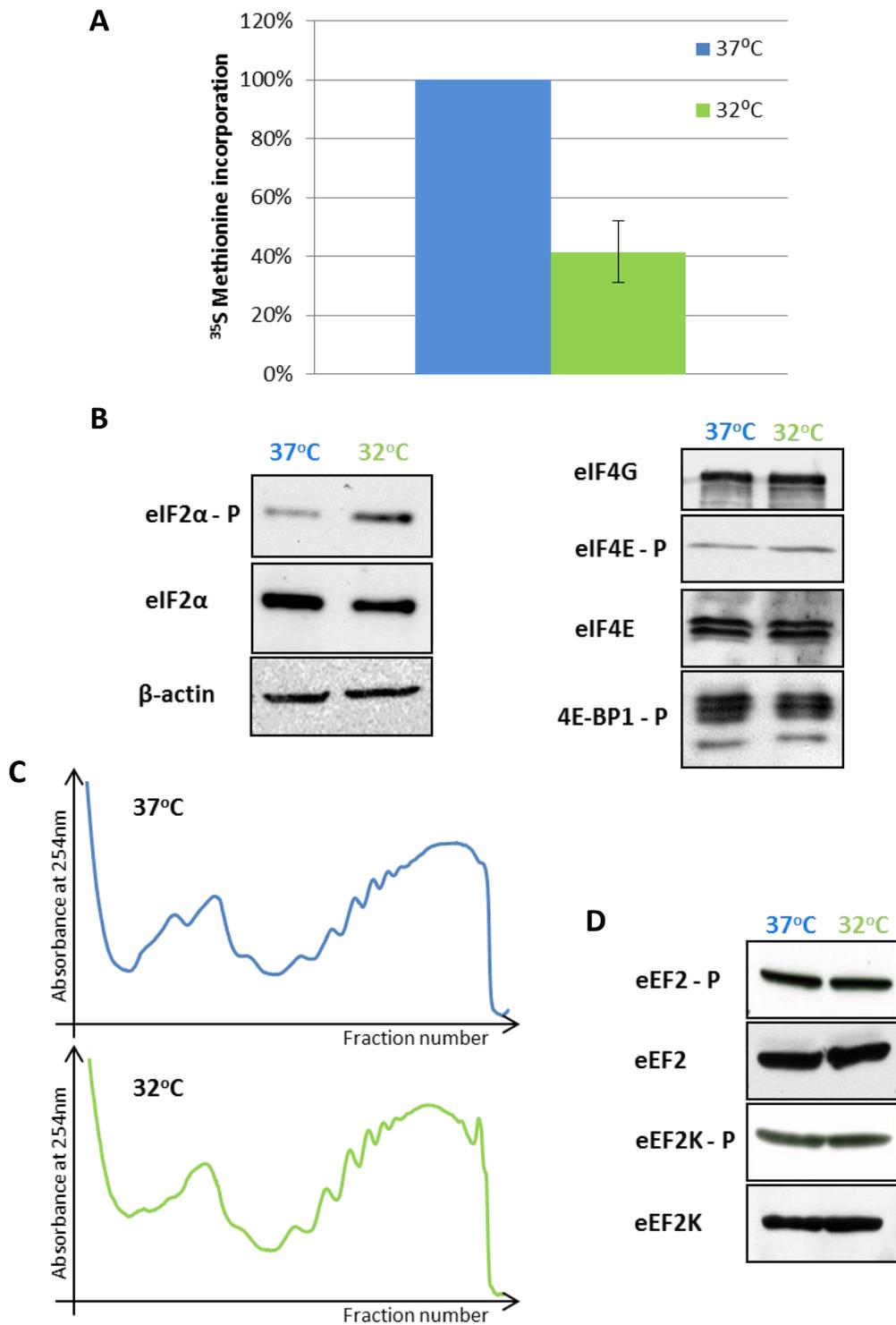
The term “cold shock response” has been used to define a range of temperature treatments from near zero Celsius to moderate hypothermia. This is unfortunate as the responses differ between these conditions. Near zero conditions are associated with severe growth arrest (Al-Fageeh et al., 2006) and a 95 % reduction in protein synthesis (Hofmann et al., 2012).

Characterisation of these conditions may be of use to improving preservation of transplant organs but of little use to industry. The cold shock response described in this chapter is to that of mild hypothermia.

Initial work by Dr Amandine Bastide (Willis group) showed that culturing the kidney epithelial cell line, HEK293, at  $32^{\circ}\text{C}$  for 24 hr resulted in a  $\sim 60\%$  reduction in protein synthesis (Figure 5-1A) and an increase in fraction of eIF2 that is phosphorylated (Figure 5-1B). There were no significant changes in the abundance of eIF4G or eIF4E or the phosphorylation of 4E-BP1. A slight increase in phosphorylation of eIF4E was observed in the cold shocked cells.

Phosphorylation of eIF4E may be associated with an increase in protein synthesis (Bianchini

et al., 2008), but given the reduction in global protein synthesis observed in response to cold shock this is either not the case in HEK293 cells or else the effect is masked by reductions via other mechanisms. The increase in eIF2 phosphorylation would be expected to reduce the translation initiation rate and thus reduce the polysomal association of mRNA. However there was little difference in polysomal association of cells treated at 37°C and 32°C (Figure 5-1 C). An explanation consistent with this observation is that a reduction in translation elongation is also a component of the cold shock response. This change was not mediated by changes in phosphorylation of either of the canonical regulators of elongation: eEF2k and eEF2 (Figure 5-1 D). To further investigate the role of elongation in cold shock I collaborated with Dr Bastide and Dr John Knight (both of the Willis Group).



**Figure 5-1 The cold shock response in cultured mammalian cells**

HEK293s were cultured for 24 hr at 37°C or 32°C, data presented are all from Dr A. Bastide.

**A** [<sup>35</sup>S] Met-label incorporation over 30 min. Data are the mean of 8 independent experiments +/- SD. Experiments were the average of 3 wells normalised to the 37°C value

**B** Representative western blots for the abundance and phosphorylation of initiation factors

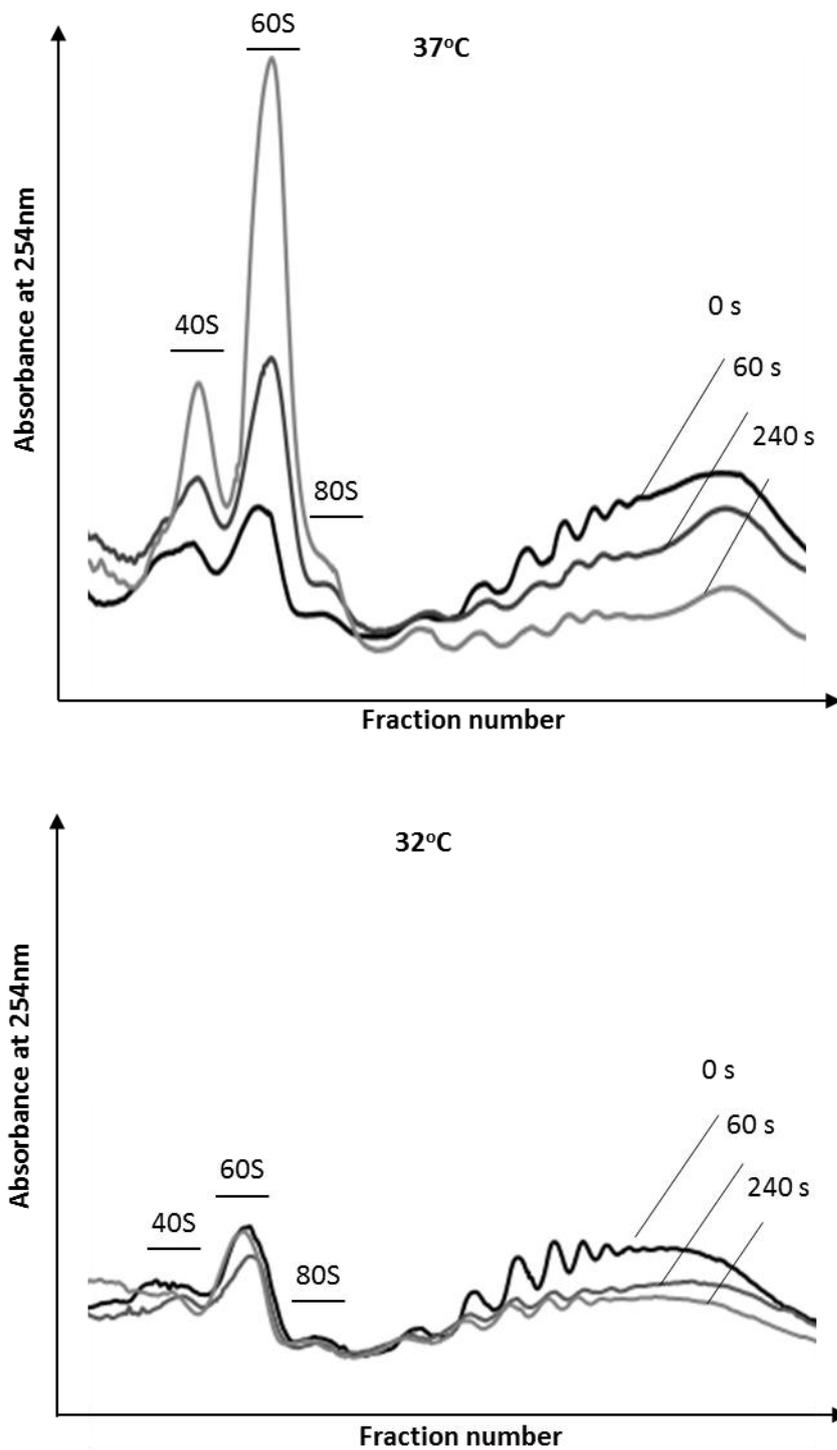
**C** Representative polysome traces

**D** Representative western blots for the abundance and phosphorylation of key elongation factors.

## 5.2 Ribosome profiling of cold shock treated cells.

To test the assumption that elongation rates were decreased in response to cold shock the harringtonine run-off approach was used (described in 2.6.hiii) and 4.3). The reduction in polysomal association occurred more rapidly in cells cultured at 37°C than those cultured at 32°C (Figure 5-2). The possibility that harringtonine was less efficiently absorbed into the cell at 32°C, allowing for new cycles of initiation and elongation during the run-off period, cannot be excluded but the data from the harringtonine run-off assay are nonetheless consistent with a reduction in elongation rate being partially responsible for the reduction in protein synthesis rates in cold shock.

The harringtonine run-off approach can be coupled with ribosome profiling to estimate global elongation rates and, for transcripts sequenced to sufficient depth, gene specific elongation rates. The report that the global elongation rate in cultured MEFs is 5 codons/second (Ingolia et al., 2011) created concern that using the same time intervals used for the initial run-off assays (Figure 5-2) would result in ribosome occupancy estimates only for long transcripts. To attempt to estimate elongation rates across a range of transcript lengths, run-off times of 60, 90 and 120 seconds were used. Translatome libraries were obtained as described in 2.6.i and analysed as per 2.6.j. There was no time dependent reduction in read density downstream from the start site in the translatome libraries, so it was concluded that the harringtonine treatment had not been successful in this instance. Therefore, further work focussed on the H0 libraries, i.e. those from cells treated with cycloheximide alone.

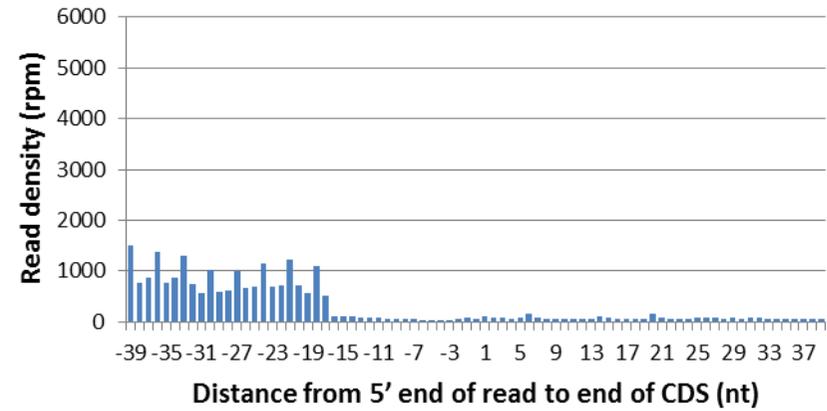
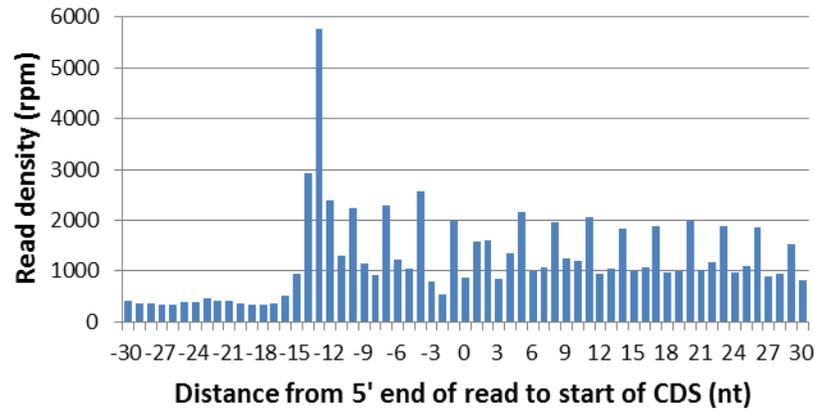


**Figure 5-2 Translation elongation is reduced during cold shock**

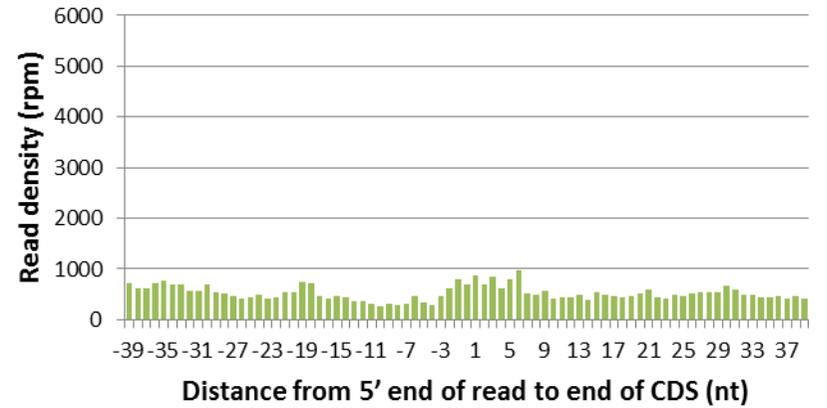
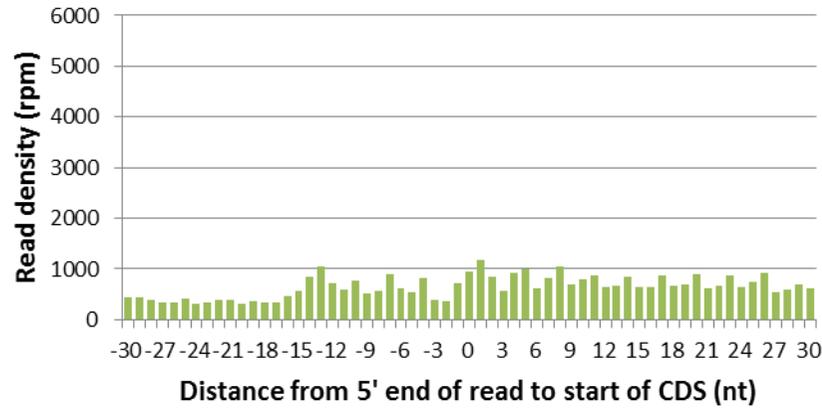
Representative sucrose gradient profiles from HEK 293 cells subject to harringtonine run-off after 24 hr at 37°C or 32°C. Harringtonine was added to the media, the cells returned to the incubator and cycloheximide added at the indicated time points.

The translatome library from cells cultured at 37°C showed a prominent start site peak, periodicity within the CDS and the absence of reads mapping to the 3' UTR (Figure 5-3A). The cold shock translatome library showed no prominent start site peak, a partial reduction of periodicity within the CDS and extensive read density on 3'UTRs (Figure 5-3B). The read density on 3'UTRs, and Western blots that showed no difference in the size of proteins in response to cold shock, suggested there was no failure of translation termination, but a failure of ribosomal release. This was surprising as translation termination is normally tightly coupled with ribosome release (see 1.2). Whilst there have been reports of termination codon run-through in yeast in response to stress (Gerashchenko et al., 2012) this is typically only the first few nucleotides downstream and not throughout the 3' UTR as observed. To resolve whether the observed 3' UTR occupancy was a result of a technical flaw in the ribosome profiling method or a genuine component of the cold shock response, an RNase protection assay was used to compare the monosome fragments obtained at 37°C and 32°C. This approach removed the need to size select fragments and generate sequencing libraries, two potential sources of technical error.

**A**



**B**



**Figure 5-3 Changes in ribosome profiling reads in response to cold shock.**

**A** Alignment of reads from translatoome libraries derived from HEK293 cells cultured at 37°C, relative to start site (left) and stop site (right)

**B** Same as A but for HEK293 cells cultured at 32°C for 24 hr.

### 5.3 Investigating translome reads in the 3'UTR

RNase protection assays involve the hybridisation of a radiolabelled probe complementary to a portion of the transcript of interest to an RNA pool. The hybridised region is protected from digestion by RNase A and RNase T1. If the hybridisation conditions used are correctly optimised, and assuming the probe is in excess of the target, the activity of the probe fragments that remain after RNase digestion will be proportional to the abundance of the specific target. Performing the assay in the presence of a total RNA sample will give a reading of the abundance of the targeted transcript; performing it on RNA extracted from the monosome peak of an RNaseI digested lysate from cycloheximide treated cells (i.e. the location of RPFs) will give a reading of ribosome density.

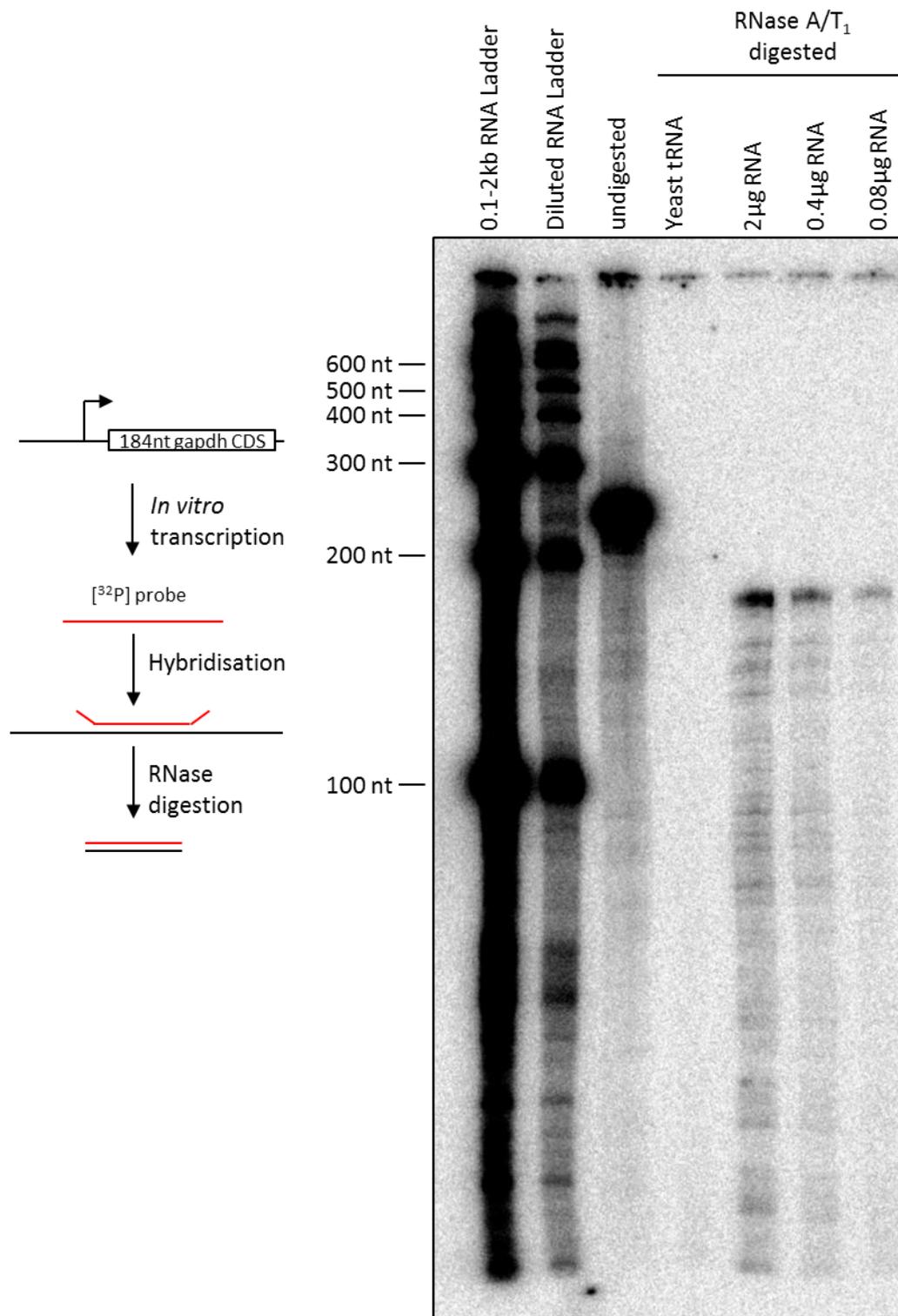
*Gapdh* mRNA was well expressed at both temperatures and showed extensive 3'UTR reads in the translome library from 32°C but not 37°C. The 3'UTR of *gapdh* is 184 nt, considerably shorter than the CDS of either short or long isoforms (881 nt and 1007 nt respectively). Using a longer probe for the CDS region might have required a higher hybridisation temperature and would give a stronger radioactive signal, so for convenience both probes contained 184 nt complementary to their respective targets. Probes were larger than 184 nt as they also contained flanking regions of the vector.

In the absence of RNase there was little degradation of the probe (Figure 5-4). Hybridisation with yeast tRNA did not protect the probe from digestion (no input control), however the total RNA sample resulted in protection of a ~184 nt band. A five-fold dilution of total RNA was adequately detected, demonstrating the assay is quantitative. It is possible that there was hybridisation to transcripts other than *gapdh* that share a similar sequence. This is unlikely; RNase A/T1 is able to cleave probes between as few as two adjacent mismatches

(Lau et al., 1993)<sup>5</sup> and whilst there are bands smaller than 184 nt it is more likely that these are due to RNA degradation prior to loading as this is was also observed in the lanes containing the ladders.

---

<sup>5</sup> Because RNase A/T1 collectively cleave after C, G and U residues, if the region of mismatch on the probe is 2 more A's flanked by an A on either side then this won't be cleaved, but this will be comparatively rare.



#### Figure 5-4 RNase protection assay for gapdh CDS

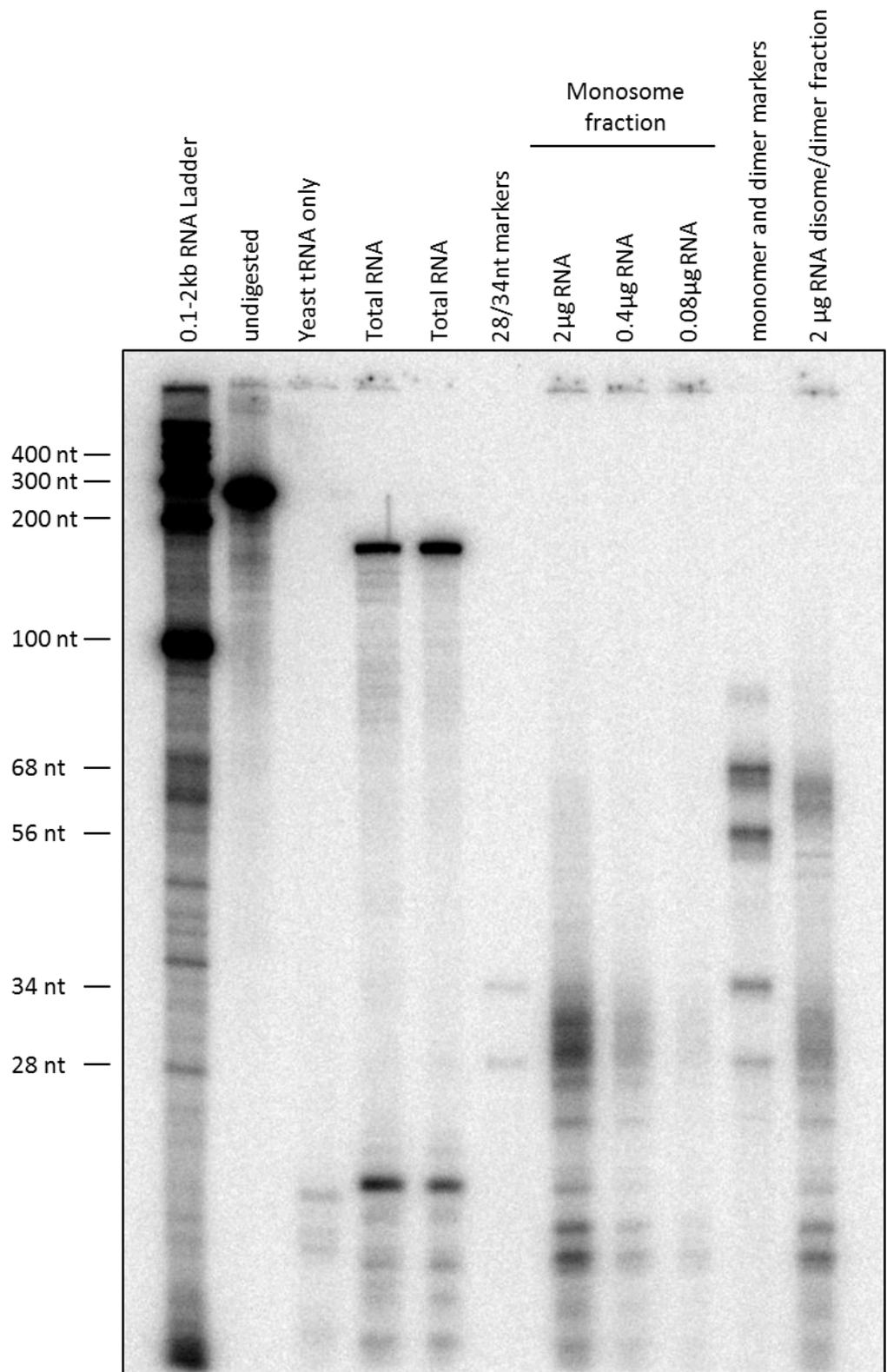
An autoradiograph from an RNase protection assay. The indicated masses of total RNA extracted from HEK293 cells were hybridised with a  $[^{32}\text{P}]$  labelled probe to gapdh and digested as described in 2.6.m. A negative control with yeast tRNA and an undigested control were also performed. The resulting products were recovered, denatured and separated on a 6 % denaturing acrylamide gel. The probe contains a 184 nt region antisense to gapdh CDS.

Ribosome protected fragments (RPFs) are ~ 30 nt; much shorter than the probes. All other things being equal, this means the predicted  $T_m$  for the probe:full length transcript hybrid is higher than for the RPF:probe hybrid. An often quoted prediction formula (Bodkin and Knudson, 1985) is recommended for hybrids longer than 100 bp (Sambrook, 1989) so may not have been appropriate for RPFs. Nevertheless this formula for a 30 nt fragment with 25 % GC under the conditions used for the total RNA gave a  $T_m$  of 52.9°C so it was decided to initially test to see if RPFs would hybridise under those conditions.

In lanes from reactions containing the monosome fraction, a smear of 28-34 nt in length was observed, consistent with hybridisation to RPFs. As with the total RNA samples, a 5-fold dilution of the monosome sample could also be detected. Interestingly, when RNA from the fraction corresponding to two ribosomes was used for the assay bands were observed between 28 and 34 nt and between 56 and 68 nt (far right lane, Figure 5-4). The ~30 nt smear may be due to two individual ribosomes bound to each other after RNaseI digestion (ribosome dimers). The ~60 nt smear maybe due to due adjacent ribosomes being so close on an mRNA as to prevent RNase digestion between them (disomes). Further evidence for these phenomena and their implications for ribosome profiling are presented in the next chapter.

In order to resolve the RPFs a higher percentage denaturing acrylamide gel was used than had been the case for the RPA on total RNA. This revealed some very short (<25 nt) bands in the yeast tRNA control lane, which were also present in all lanes from reactions that were treated with RNase. This result suggests that some of the probe must have hybridised to itself, to an adjacent probe or the yeast tRNA (in the no input lane). Given that the shorter the hybrid the higher the GC % necessary to retain the same  $T_m$ , at least one of two things are likely to be true: 1) There were very GC-rich sections of the probe, 2) the hybridisation temperature was lower than it needed to be to detect RPFs. Consistent with the latter, the

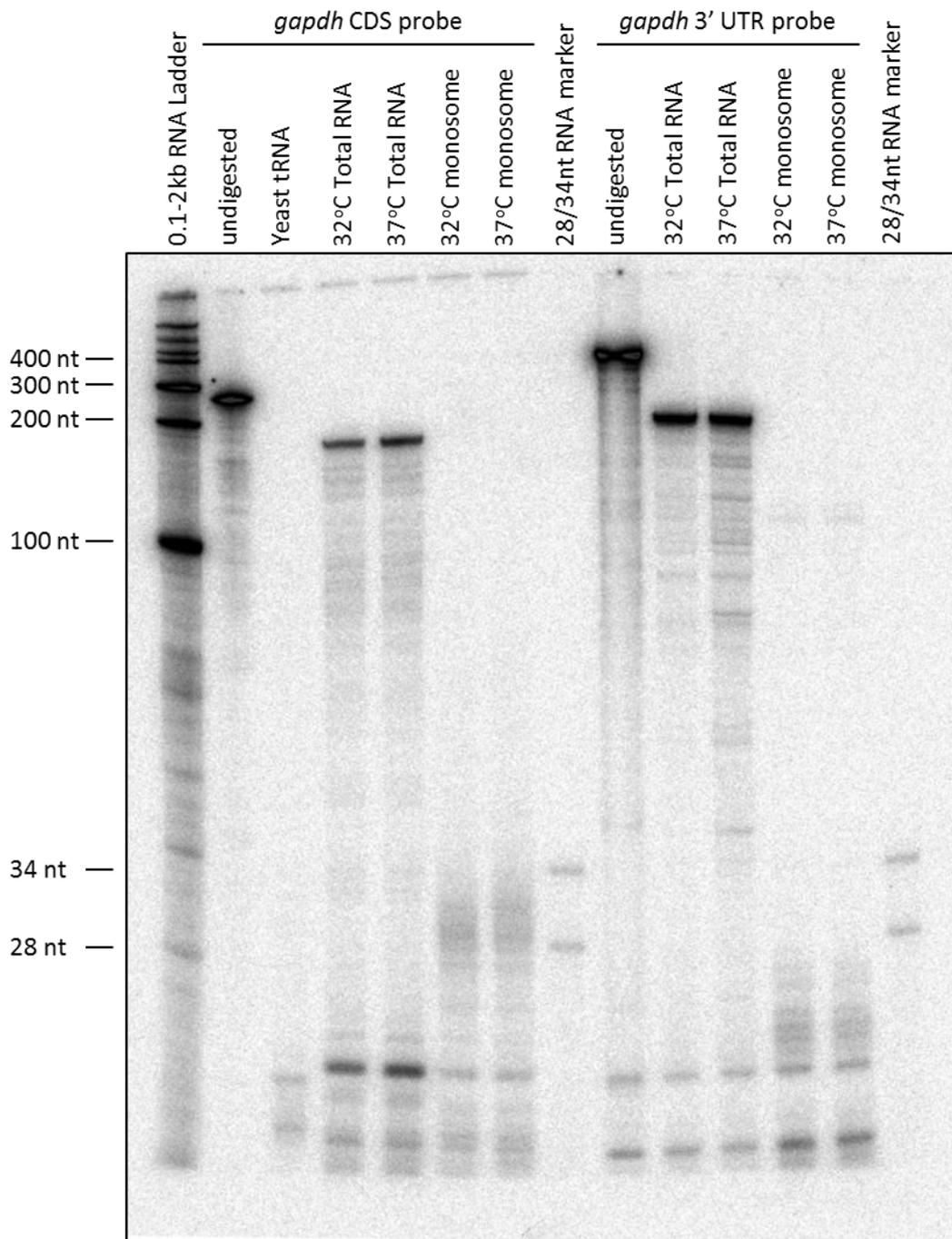
total RNA lanes contained a prominent band below 28 nt which could be microRNAs (~21 nt), however bands of the same size were observed in the no input lane suggesting self-hybridisation was a more likely cause.



**Figure 5-5 RNase protection assay on fragments within the monosome peak**

An autoradiograph from an RNase protection assay performed with the gapdh CDS probe on the indicated masses of RNA obtained from the monosome peak of RNaseI digested HeLa lysate. The outermost lane is from the RNA obtained from the peak corresponding to two ribosomes. The undigested control is 1/10<sup>th</sup> of total input.

If the temperature was too low then it is possible that ~30 nt fragments of rRNA or other mRNAs with 1 mismatch to the probe were detected in the monosome fractions. However if this was the case then it would be unlikely that none of the signal would be from the hybridisation of gapdh-derived RPFs. Therefore the absence of any smear at ~30 nt would be indicative of no ribosome occupancy within region complementary to the probe. When the monosome fractions from the samples that were ribosome profiled were assayed, a 30 nt smear was observed with the CDS probe but not the 3' UTR probe (Figure 5-6). From this it was concluded the translome reads that mapped to the 3'UTRs were likely the result of a technical problem when the ribosome profiling was performed.



**Figure 5-6 RNase protection assay does not show evidence of 3' UTR ribosome footprints**  
 An autoradiograph from an RNase protection assay performed with either *gapdh* CDS probe or 3' UTR probe on the H0 samples that were sequenced and represented in Figure 5-3. Both probes contain a 184 nt region antisense to *gapdh* (CDS or 3'UTR as appropriate)

#### 5.4 Discussion

The cold shock response involves the modulation of multiple different cellular processes. That culturing HEK293s in hypothermic conditions does not reduce polysomal association despite a reduction in global protein synthesis rates suggests elongation rates are reduced. Consistent with this, ribosome run-off after harringtonine treatment is much slower at 32°C. This change in elongation rate is not mediated by eEF2 or eEF2k, so perhaps the cause is thermodynamic phenomenon (i.e. at lower energy levels recruitment of tRNAs and subsequent peptide bond catalysis occurs more slowly). The failure of the harringtonine treatment in the ribosome profiling experiment meant it was not possible to address whether some transcripts are differentially regulated at the level of elongation.

That reads from translome libraries mapped extensively to 3'UTRs was surprising, but subsequent RNase protection assays suggest this was due to a technical error. It is unclear precisely at what stage of the process this occurred. There were two important differences between the gel electrophoresis used to resolve RPFs in the ribosome profiling compared to the RPAs. First, the gels used for ribosome profiling were considerably shorter than for RPAs so the resolution would have been less crisp (i.e. under or oversized, non RPF fragments could have been selected). Second, the markers used in the RPA were 28 and 34 nt rather than the 27 and 33 nt used in the profiling work. The reason for the difference was that the first ribosome profiling paper used these shorter markers in a yeast system (Ingolia et al., 2009) but a subsequent publication from the same group working in a mammalian system used the larger ones (Ingolia et al., 2011). The RNase protection assay with the 3'UTR probe does show bands just below the 28 nt marker so it is possible that these were extracted with the 27/33 nt markers. It should be noted that the 27/33 markers did give the expected distribution of RPFs in the iMet work and in the 37°C libraries, and a degree of periodicity at 32°C so they can be used to capture RPFs, but there is a risk of contaminating fragments

might result in spurious conclusions.

Ultimately these experiences served as a cautionary tale. In subsequent collaborative work, that is not the scope of this thesis, the 28/34 nt markers were used and gave sequencing data with the expected RPF periodicity.

## Chapter 6. The identification and remediation of biases in next generation sequencing protocols.

### 6.1 Introduction

Next generation sequencing is a powerful tool. It allows analysis of the whole genome, transcriptome or translome without any prior assumptions as to their composition.

However, as with all techniques there are limitations. As highlighted in Chapter 3, statistical approaches commonly applied to RNA-seq experiments are not applicable to ribosome profiling data such that differential translational efficiency can only be accurately tested for well translated, abundant transcripts.

More fundamentally, the assumption of all next-generation sequencing work is that the read data obtained is a good estimate of the relative abundance of features within a biological system (e.g. microRNA/mRNA abundance, ribosome occupancy, RNA-Protein interactome). Biases in the technique will reduce the extent to which this assumption is true and thus the possible strength of conclusions that can be inferred from sequencing data.

A simple reason why data may not fully reflect the biological complexity is if insufficient RNA is used to generate the sequencing library meaning fragments from lower abundance transcripts are excluded. *In silico* fragmentation of the human RefSeq library into 30 nt fragments yields  $4.2 \times 10^8$  fragments, 93 % of which are unique (Julian Wang, Willis lab). Assuming equal nucleotide content, the molecular weight of a 30 nt RNA will be 9759 g/mol. Typically no less than 10ng of RPF fragment pool is used in a ribosome profiling experiment:  $\sim 6.1 \times 10^{11}$  molecules. Even factoring in that 80 % of these fragments will come from 20 % of

transcripts it seems unlikely transcripts are under-sequenced because of a lack of input material<sup>6</sup>.

However, there are significant differences in the data obtained when sequencing an identical pool of RNA or DNA on different platforms (Quail et al., 2012, Toedling et al., 2012). This implies there are at least two sources of bias in sequencing data: library preparation and sequencing reaction chemistry (both of which differ between platforms). There is also a third source of bias that may be particularly significant in ribosome profiling: fragment generation. Ribosome profiling assumes that the ratios of ribosome protected fragments (RPFs) in the fragment pool are proportional to the steady state ribosome density across a message. If this is not the case then making conclusions on changes in the location of ribosomes should be made with caution – it could be that the ribosome occupancy is actually increased at a feature of interest but that biases in fragment generation result in a lower proportion of RPF fragments from that feature in the fragment pool.

To briefly summarise, biases at any of the three stages identified will lead to results that are not representative of the biological system being studied. Identifying how significantly these biases skew data and potential remedies to them is thus important. This chapter will focus on biases in fragment isolation in the context of ribosome profiling and in library preparation on Life Technologies platforms.

## **6.2 Ribosome protected fragments may not fully reflect ribosome occupancy**

Central to ribosome profiling is the nuclease digestion of lysates leaving fragments where ribosomes are bound: ribosome footprints. A single eukaryotic ribosome will protect ~30 nt from digestion, but protein-mRNA complexes may also do the same. For this reason

---

<sup>6</sup> Current sequencing platforms have a maximum yield of approximately  $2 \times 10^8$  reads per multiplexed library, not all of which are useable, so not every possible fragment is eventually sequenced

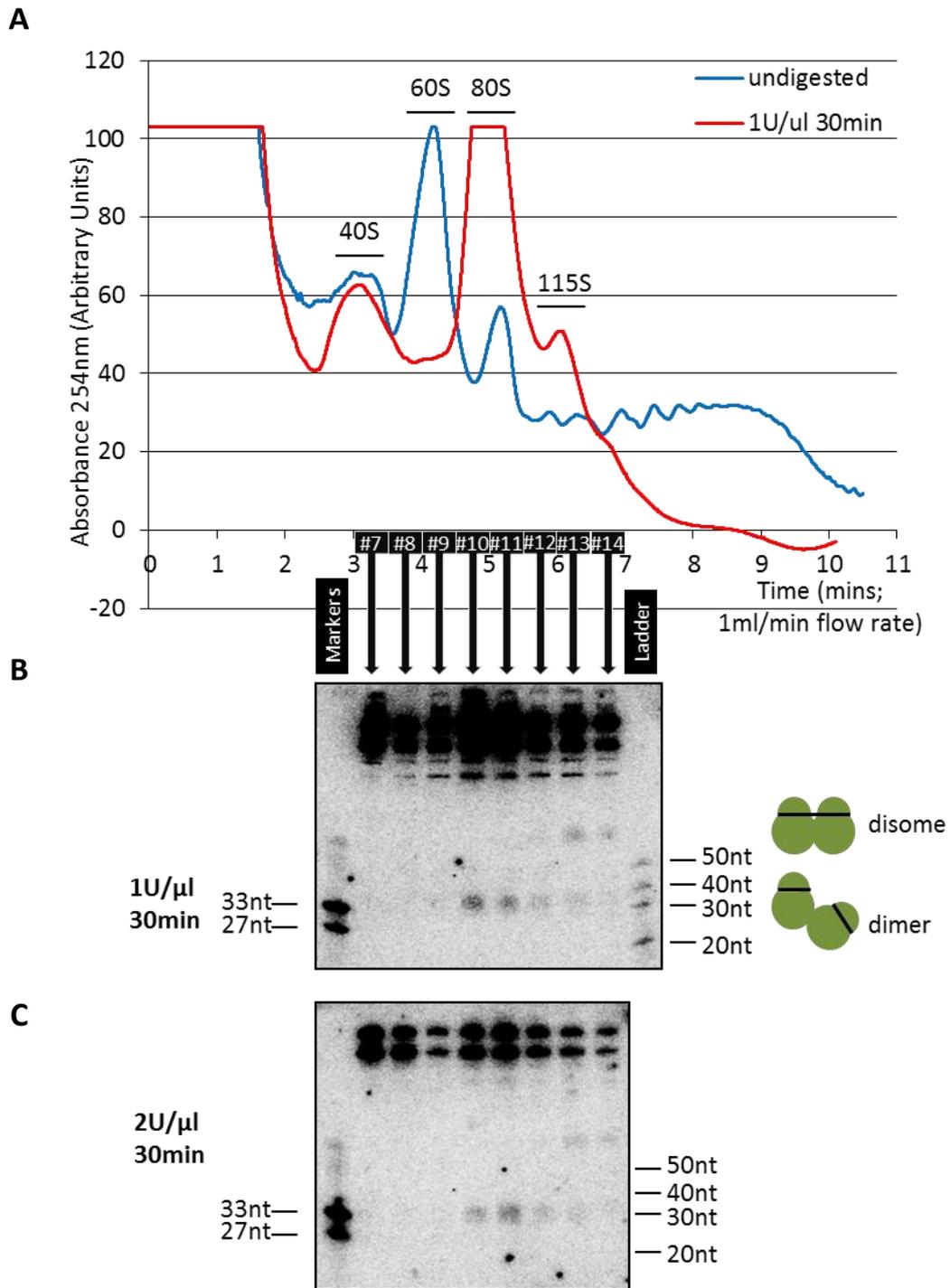
footprints are isolated from the monosome peak of a sucrose gradient (i.e. just one ribosome with associated with an RPF).

It is then assumed that all ribosome occupancy is reflected in the relative distributions of these 30 nt RPFs. For instance, given two equally abundant mRNAs where the first has twice the ribosome occupancy, fragments from the first will appear twice as frequently in the translational fragment pool.

However, the sucrose gradient traces from RNase-digested lysates showed peaks corresponding to multiple ribosomes (Figure 6-1 A). Notably, this is also seen in all published papers that display sucrose gradient traces from ribosome profiling in eukaryotic systems (Ingolia et al., 2009, Stadler and Fire, 2011). The existence of these peaks could be due to insufficient RNase digestion (i.e. concentration, time or temperature not optimised) or due to adjacent ribosomes being so close as to prevent the region between them being accessed by RNase: ribosome stacking (Wolin and Walter, 1988). A hybrid of these two explanations is also theoretically possible – the shorter the distance between adjacent ribosomes the smaller the region of mRNA that can act as an initial binding site for RNase I, leading to regions of high ribosome density being digested less efficiently than regions of low density. Northern blotting with a 20 nt fragment complementary to a region within the CDS of an abundant mRNA, revealed that whilst the monosome peak contains mRNA fragments of ~30 nt, the bi-ribosome peak (i.e. 2 ribosomes present) contains a mixture of ~30 nt and ~60 nt fragments (Figure 6-1 B). A similar observation was made using an RNase protection assay with a probe spanning 184 nt of the CDS of *gapdh* mRNA (Figure 5-5). Protection of a ~60 nt fragment is consistent with ribosome stacking. If the hybrid explanation were true, then species slightly larger than the equivalent of two monosomal RPFs would be expected. This was not the case in the RNase protection assay. Whilst there are larger molecular weight bands in the northern blots, these are much larger than 60 nt, are found in all fractions and are reduced in intensity after washing at a higher temperature (not shown). It was therefore

presumed that they were the result of non-specific binding to non-mRNA species. An increase in RNase concentration significantly reduced the intensity of these non-specific bands but did not result in the loss of the ~60 nt band (Figure 6-1 C).

The presence of a ~30 nt band in the bi-ribosome peak is somewhat perplexing; if they are RPFs then they are from monosomally associated mRNA. The 1 ml of lysate added atop the gradients represents ~10 % total volume so it is possible that the disomes at the top of the lysate sediment alongside monosomes at the bottom of the lysate. Another possibility, described in Chapter 5, is that the dimerisation of ribosomes after digestion results in the associated RPFs sedimenting in the bi-ribosome peak.



**Figure 6-1 RNase I digestion does not reduce all ribosome footprints to monosomes**

A) Representative sucrose gradient trace of lysates from cycloheximide treated MEFs. Lysates were prepared as described in 2.6.iii) and 2.6.hii)

B) Northern blot with a 20 nt DNA probe anti-sense to eEF1a. Each lane contains 1/5<sup>th</sup> of the RNA extracted from the fraction listed above. An illustration of ribosome stacking (giving disomes) and post digestion ribosome dimerisation is shown on the right.

C) Same as B but with digestion performed with twice the RNase I concentration

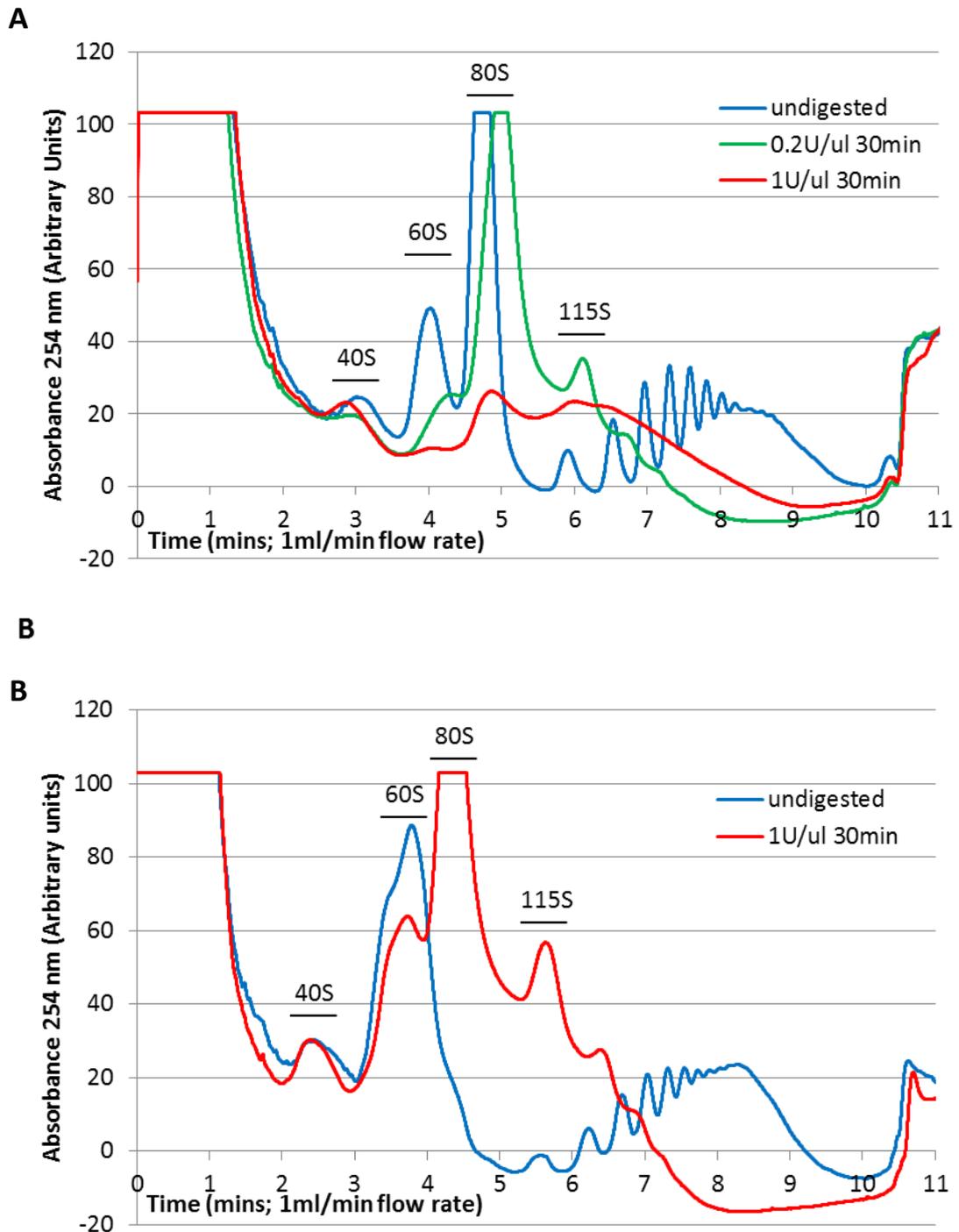
Ribosome dimers have been detected in undigested lysates from excised liver and cell lines of hamster and rat, but not from mice or human cell lines (Reader and Stanners, 1967).

Dialysis against RSB lysis buffer, which has near physiological levels of  $K^+$  and  $Mg^{2+}$  (Romani, 2007, Williams and Wacker, 1967), reduced, but did not abolish, dimerisation in that case. RNase digestion of the hamster and rat lysates did not remove the 115S (bi-ribosome) peak, but did result in the radioactivity from the  $^{14}C$ -Phe used to trace nascent protein synthesis being localised in the 80S peak and not the 115S peak. Therefore the authors concluded that the dimers were not associated with translating mRNA.

By contrast, the ribosome profiling work presented in this thesis shows what may be ~30 nt mRNA fragments in the 115S peak of RNase digested lysates from human (Figure 5-5) and mouse cells lines (Figure 6-1). A possible explanation for the difference between these data and the previous observations (Reader and Stanners, 1967) is that under harsh RNase I treatment there is extensive digestion of the ribosome itself (hence the contaminating rRNA fragments in translational libraries) which may expose new sites on the partially digested ribosome that can hybridise to another. High cation concentrations and that lysates are placed on ice after digestion and centrifuged at  $4^{\circ}C$  may further favour dimerisation.

To investigate the role of cation concentration in the presence of the 115S peak, HeLa cells were lysed in either a buffer used routinely within the Willis lab (HSB; see 2.6.iii) or a lower salt and magnesium buffer used by (Guo et al., 2010) (LSB). The sucrose gradients used for fractionation contained the same ion concentrations as the buffer. The polysomal distribution was different in the undigested samples. High salt conditions dissociate 80S ribosomes not bound to mRNA (vacant couples) (Falvey and Staehelin, 1970), hence the prominent 80S peak in Figure 6-2 A is not seen in Figure 6-2 B. It is unclear whether the relative increase in heavy polysomes in HSB is because polysomes are not as stable in LSB or because HSB results in aggregation of lighter polysomes. Interestingly the RNase I concentration used in the HSB resulted in destruction of the monosome peak in the LSB. This

was somewhat surprising as when RNase I was first characterised it was reported to have optimal activity in NaCl or KCl concentration of ~200mM with an approximately 10 % decrease in activity at both 300mM and 100mM (Spahr and Hollingworth, 1961). Increasing  $Mg^{2+}$  was reported to reduce activity, but not as dramatically is seen in Figure 6-2 so there is presumably an additive effect of high salt and high magnesium. Crucially, bi-ribosome peaks were present in both LSB and HSB, though less in the LSB. This could be because there were fewer heavy polysomes in the LSB and thus less ribosome stacking or because the ionic conditions affect dimer formation. Comparison of the relative abundance of dimers and disome fragments would distinguish between these two possibilities, but this was not performed. When the  $Mg^{2+}$  concentration in the lysis buffer was reduced to 1.5mM no polysomes could be recovered, possibly because intracellular magnesium levels are higher than has been previously reported.



**Figure 6-2 The effect of buffer composition on the bi-ribosome peak**

Identical plates of HeLa cells were lysed directly and digested under conditions described in the legends.

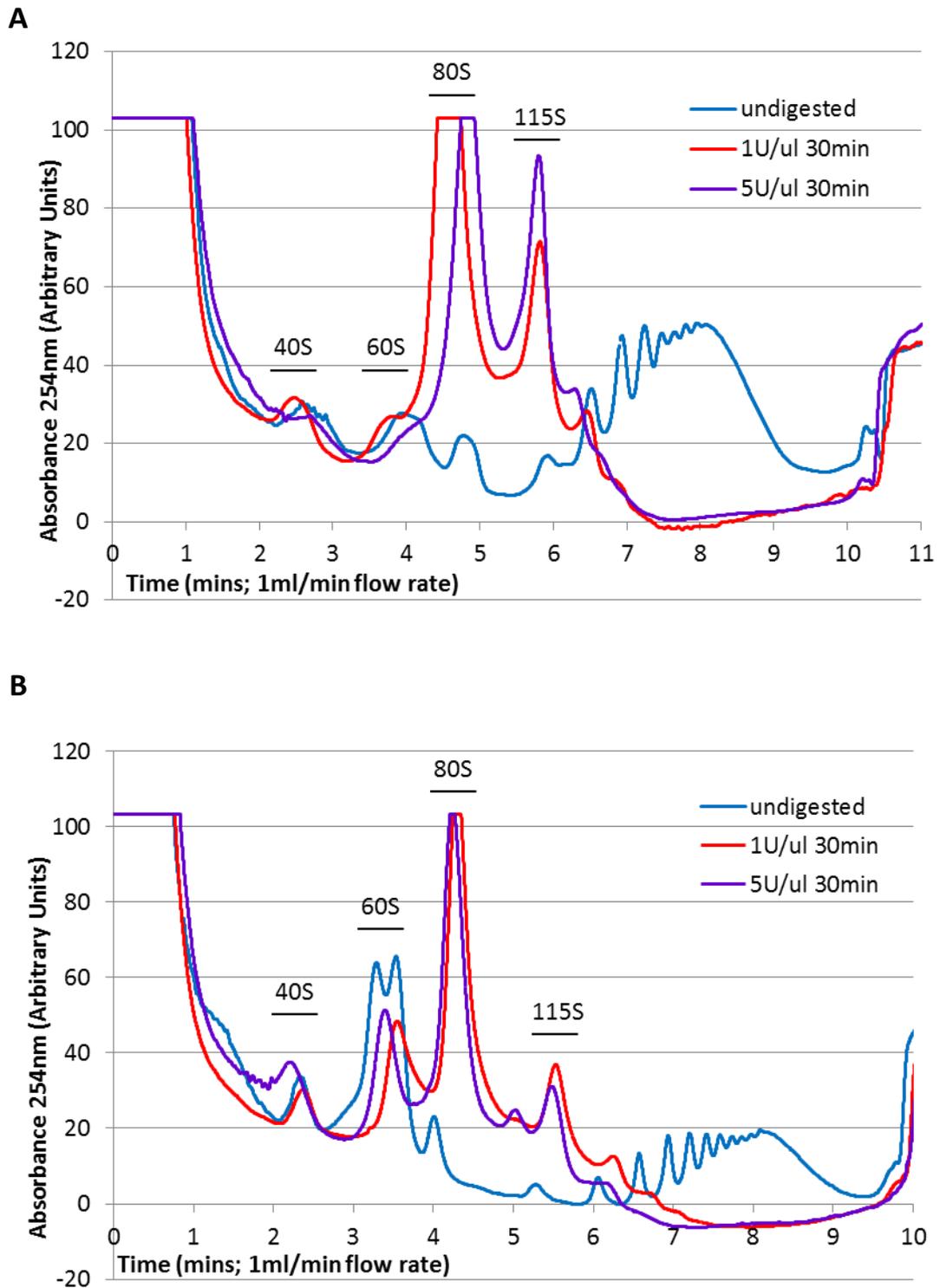
**A** Lysis and sucrose density ultracentrifugation in a low salt, low magnesium buffer (100mM KCl, 5mM MgCl<sub>2</sub>, 10mM Tris-HCl pH 7.4, 100 µg/ml cycloheximide, 2 mM DTT). The offset of the 80S peak in the digested samples compared with the undigested samples is likely due to pipetting error when the gradients were made.

**B** As A but in a high salt, high magnesium buffer (300mM NaCl, 15 mM MgCl<sub>2</sub>, 15mM Tris-HCl pH 7.5, 100 µg/ml cycloheximide, 2 mM DTT)

To address concerns that the 115S peak may be due to insufficient RNaseI activity in the HSB a range of RNaseI concentrations were used on lysates from NIH 3T3 MEFs<sup>7</sup> and HeLas. Surprisingly, in MEF lysates the 115S peak was increased when higher RNaseI concentrations were used (Figure 6-3 A). To gain comparable width peaks in sucrose gradient traces, it was necessary to maintain the same volume of lysate at each RNaseI condition so lysates with increased RNase concentration also had higher concentrations of glycerol (which inhibits RNase). The high RNase concentration condition had 2.5 % glycerol and there was no increase in the height of peaks larger than disomes so it seems unlikely the RNase activity was lower at the higher concentration. A possible explanation is with more digestion there is more opportunity for dimerisation. In HeLa extracts, increasing RNase concentration had no effect on the size of the 115S peak and only a modest decrease in heavier polysomes (Figure 6-3 B).

---

<sup>7</sup> The undigested lysates are more polysomal that was seen in Chapter 2 and Figure 6-1, this is due to the extraction method being altered to replace scraping of cells into PBS and subsequent pelleting with lysis of cells directly on the plate. The scrape-first method appears to reduce polysomal association by an unknown mechanism.



**Figure 6-3 Increased RNase I digestion increases bi-ribosome peak in MEFs but not HeLa**

**A** Identical plates of MEFs were lysed in high salt buffer and digested with the indicated concentrations of RNase I in the figure legends.

**B** As A, but with HeLas and with a smaller volume of lysate loaded onto the gradient to give greater separation between peaks

Collectively the above experiments suggest that RNase I digestion is not sufficient to yield 30 nt RPFs from all locations of ribosome occupancy. However both the RPA and northern blot can suffer the same limitation: non-specific hybridisation of the probe. A strategy to address this would be to specifically label mRNA and detect the location of labelled fragments in the sucrose gradient, thus removing any hybridisation steps. Labelling intracellularly will not result in specific labelling of mRNA as all RNAs use the same nucleotides. Treatment with 300nM CX-5461 can reduce rRNA synthesis by up to 80 % with only modest reduction in the abundance of select mRNAs and translation (Drygin et al., 2011). However rRNA accounts for ~70 % of total RNA so even an 80 % reduction in synthesis would likely result in more labelled rRNA than mRNA, and associated background fragments. Therefore transfection of *in vitro* transcribed mRNA is necessary for this strategy. Attempts to use 4-thiouridine as a labelling molecule were unsuccessful as conjugation to biotin and subsequent detection with streptavidin-HRP, which can be performed in solution, could not be achieved on labelled RNA immobilised on a nylon membrane (i.e. northern blotted). [<sup>33</sup>P] labelled mRNA was used instead as no secondary detection method was required. In order to maximise the specific activity within the RPF pool (aiding detection), increase the likelihood of observing ribosome stacking and reduce the possibility that overexpression of the transcript would substantially perturb the cells, an abundant mRNA that is associated with heavy polysomes was chosen:  $\beta$ -actin.

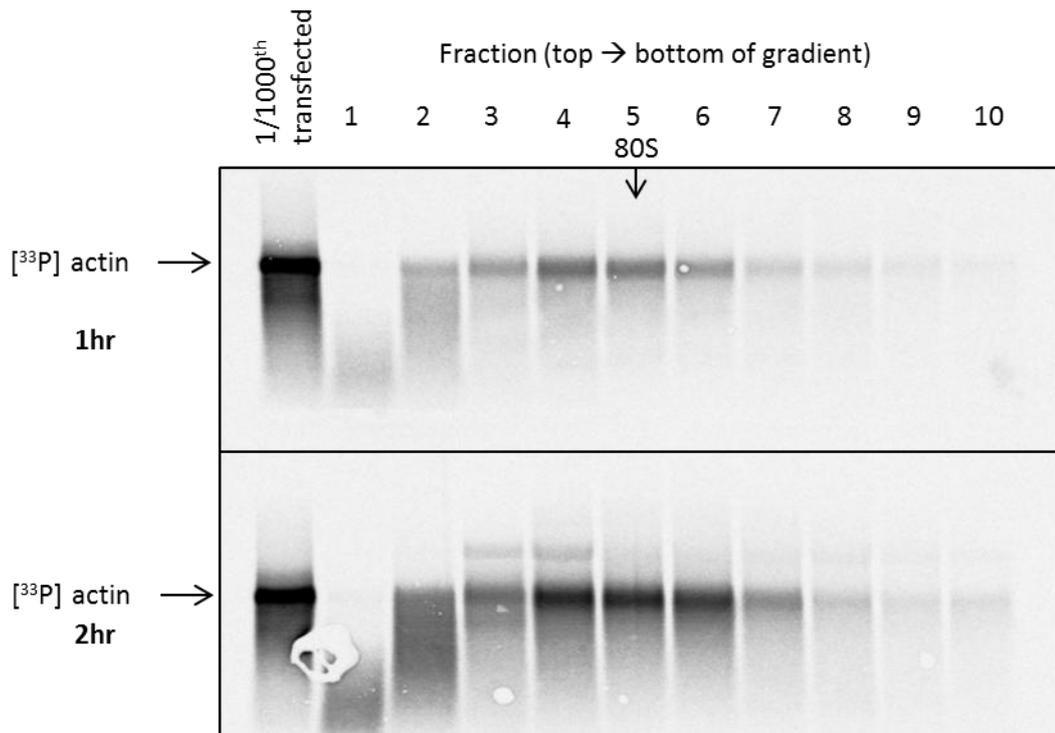
1-2 hr after lipofection of [<sup>33</sup>P]-actin mRNA into HeLa cells, the labelled mRNA was found in light polysomal fractions, unlike the native transcript that was in heavy polysomes (Figure 6-4). Lipid-based transfection has been reported to retain mRNAs in discrete, stabilised particles (Barreau et al., 2006). To investigate whether this accounted for the reduced polysomal association of the transfected mRNA, the experiment was repeated with nucleofection (i.e. electroporation) in place of lipofection. This did not address the issue of

polysomal association as lysates obtained 3 hr<sup>8</sup> after nucleofection showed extensive degradation and radioactive bands corresponding to the 18S and 28S rRNA (data not shown). Degradation was also observed in the lipofected samples (see fractions 1 and 2 of Figure 6-4) and lysates taken 4 hr after lipofection had no full length [<sup>33</sup>P] transcript (data not shown). By contrast the half-life of endogenous  $\beta$ -actin mRNA in HeLa cells has been reported to be 6-12 hr (Khalili and Weinmann, 1984). There are two possible explanations for the reduced polysomal association and rapid degradation of the [<sup>33</sup>P] labelled mRNA. First, during the transcription and processing of endogenous mRNAs, proteins are associated that protect the mRNA from degradation and aid translation – these are not present in the “naked” RNA transfected into cells. Second, regions of the T-Easy vector that flank the cloning site were present in the actin mRNA, effectively altering the 5'UTR and 3'UTR. The linearisation site (SpeI) for the template plasmid was only 4 nt downstream from the cloning site and created no additional microRNA sites so it is unlikely that the 3' UTR mediated any affect. However, after these experiments were performed it was noticed that the 50 nt between the cloning site and the T7 transcriptional start site of pGEM-T Easy does contain a start codon in good Kozak context: GGCCGCCATGG (ggcRccAUGG). If translated, this would create a 42 amino acid uORF that overlaps the CDS of the actin transcript and so might explain the reduced polysomal association of the *in vitro* transcribed mRNA. Using SP6 polymerase to transcribe the mRNA from a vector with the actin sequence in the opposite orientation would have produced a transcript without an AUG in such good context but this was not performed. The lack of success in this approach means that it may be possible that the bi-ribosome peak is not particularly significant and the data presented thus far are simply a consequence of methodological artefacts.

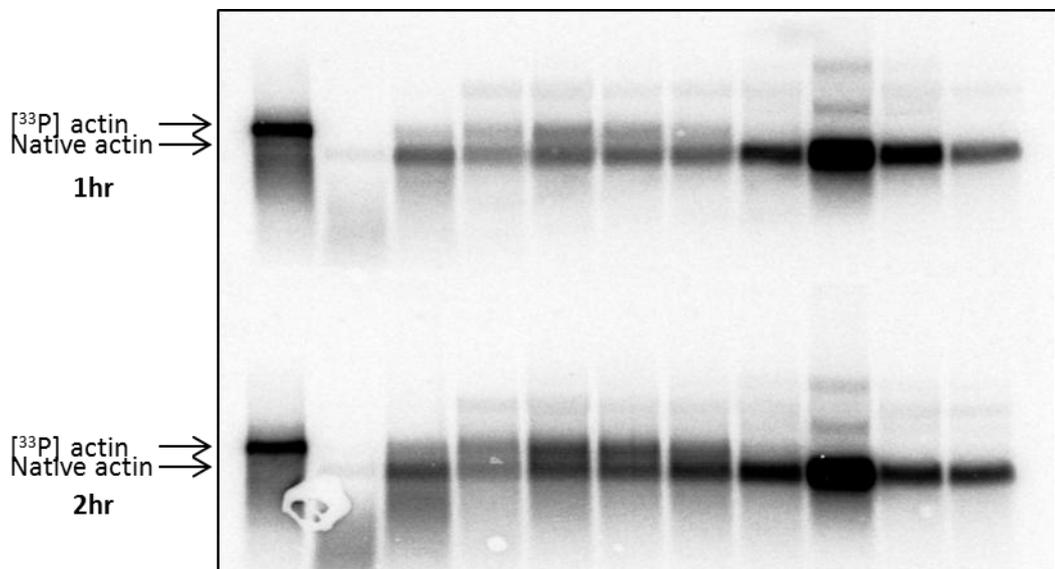
---

<sup>8</sup> Cells were left for 3 hr to allow attachment to the culture plates

**A**



**B**



**Figure 6-4 Transfected *in vitro* transcribed actin mRNA does not share the same polysomal distribution as the native transcript**

**A** Polysomal association of  $[^{33}\text{P}]$  labelled capped mRNA 1hr (top) and 2hr (bottom) after lipofection into HeLa cells.  $\frac{1}{4}$  of RNA was from each fraction was northern blotted. The monosome peak was within fraction 5.

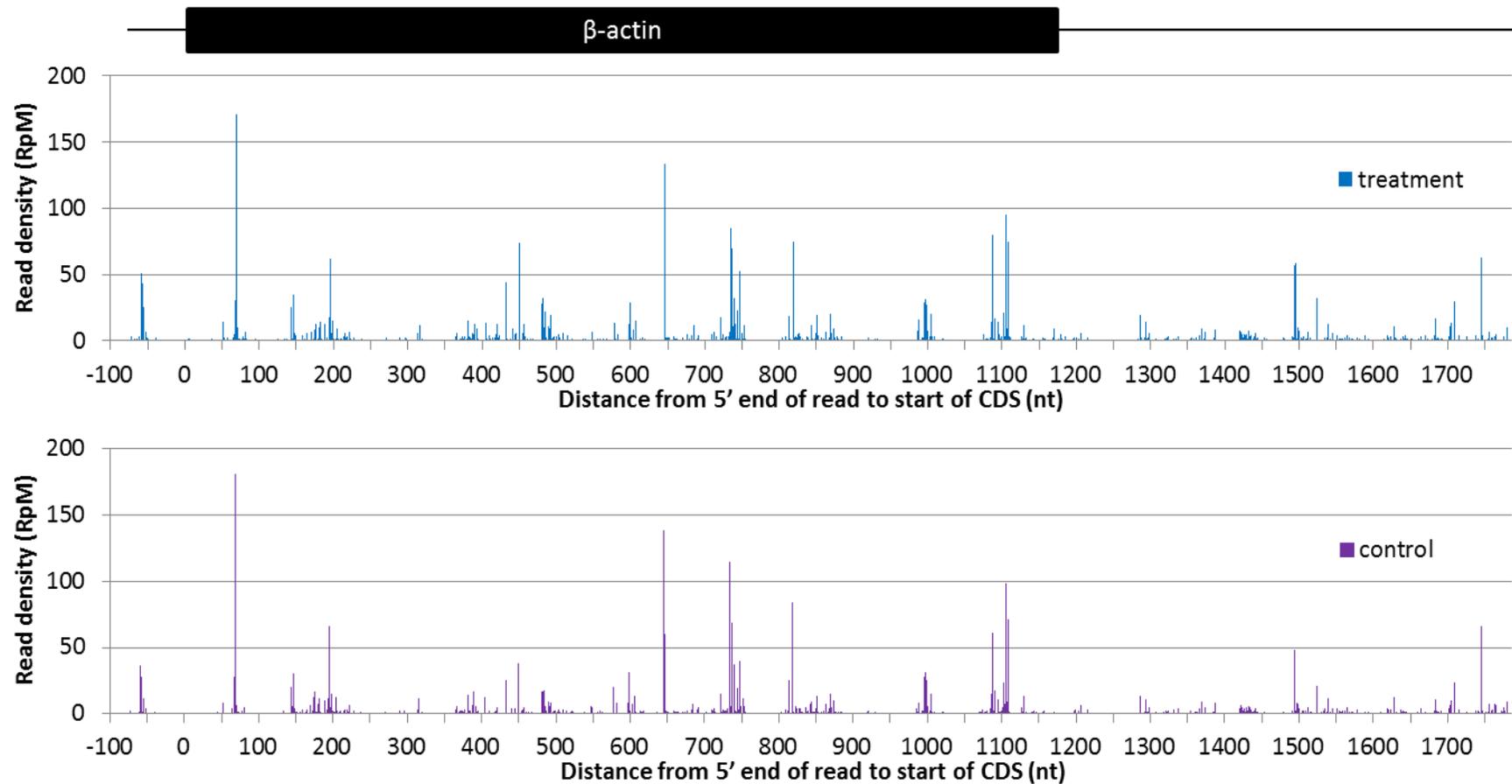
**B** The membrane in A was probed with a  $[^{32}\text{P}]$  labelled probe complementary to actin. The transfected mRNA had a longer polyA tail and so is visible above the native transcript.

### 6.3 Ligation biases in deep sequencing libraries

Even if no biases are introduced at the level of fragment generation, sequencing data may be skewed by biases in the library generation protocols. In order to sequence RNA fragments using current platforms, they must be ligated to adaptors that can be used to amplify all fragments onto beads (or in the case of Illumina directly and later versions of SOLiD technology directly onto the sequencing slide) to be sequenced.

SOLiD transcriptome libraries are generated from RNase III fragmented polyA+ RNA. It was assumed that the fragmentation protocol had been optimised by Life Technologies to be essentially random and thus that sequenced reads obtained would map evenly across the transcript. This is not the case (Figure 6-5). This is at least in part because digestion by RNase III is not unbiased (it has preference for dsRNA), a fact acknowledged in later editions of the SOLiD Total RNA-Seq kit manual. Alkaline (i.e. chemical) fragmentation methods give much more even coverage than enzymatic ones (Danny Nedialkova, personal communication, 26<sup>th</sup> September 2012).

Given that the uneven coverage in the transcriptome libraries is similar between the control and iMet cells (see Figure 6-5 and Chapter 3) it seems reasonable that comparison between conditions for a given transcript can be performed regardless of biases. However, if library generation biases are substantial they will prevent accurate quantification of the abundance of a fragment/transcript relative to the whole transcriptome. This would limit the ability to answer a question such as: what is the most abundant mRNA/microRNA in a particular cell? Furthermore it may limit attempts to make position specific inferences about ribosome profiling data – e.g.: was there a lack of translome reads mapping to a particular feature because there was genuinely no ribosomes there or because the fragments generated were selected against due to biased library preparation?



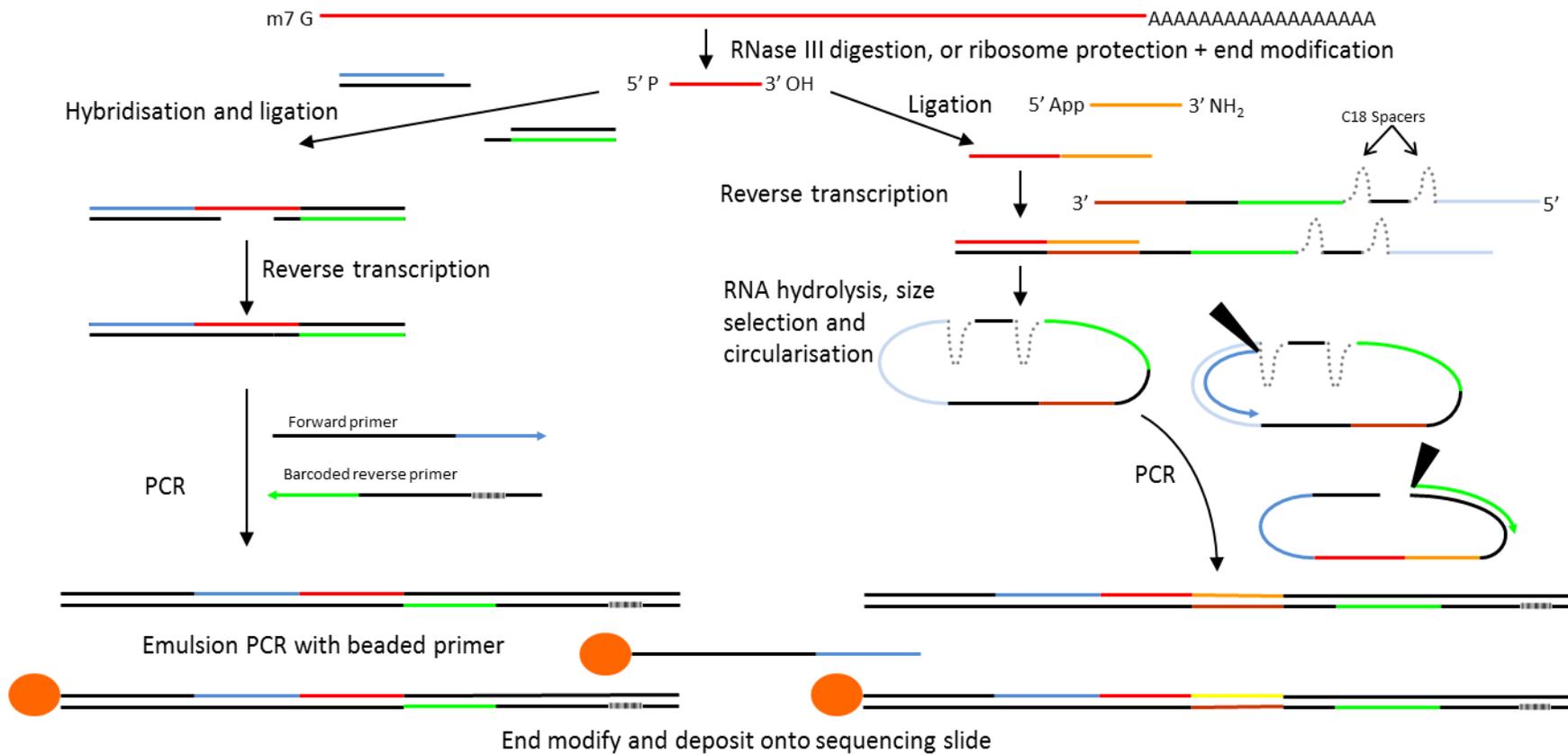
**Figure 6-5 Uneven read distributions within transcriptome libraries**

The distribution of reads from transcriptome libraries prepared with fragmentation by RNase III that map to β-actin mRNA. β-actin is the most abundant mRNA in these two sequencing libraries. Treatment is over expression of tRNA<sub>i</sub><sup>Met</sup> (see Chapter 3 for details of cell lines).

There are three steps in library generation: ligation of adaptors, reverse transcription to generate cDNA and PCR of the resulting cDNA. The use of random hexamers to prime reverse transcription does bias sequencing data (Hansen et al., 2010), but the use of a universal RT primer that binds to a sequence on the adaptor circumvents this. PCR amplification results in under-representation of both AT rich and GC rich regions (Aird et al., 2011, Dohm et al., 2008) but this can be minimised by the use of polymerases that have been generated through molecular evolution to reduce these biases (e.g. KAPA HiFi) or by using alternative protocols if the likely AT content of the library is known (Oyola et al., 2012). It may also be possible to correct for PCR biases using statistical methods (Benjamini and Speed, 2012), but it is unclear how appropriate these methods will be for transcriptome data where GC content may play a role in ribosome occupancy under different conditions.

The approaches for the ligation of fragments differ between Illumina and Life Technologies platforms. For completeness and to aid in understanding of referenced work that follows, both are briefly described below.

In the Illumina approach, RNA adaptors are sequentially ligated to the 3' and the 5' ends of fragments using truncated T4 RNA ligase 2 (trRnl2) and T4 RNA ligase 1 (Rnl1) respectively. Rnl1 shows sequence specific biases (Jayaprakash et al., 2011) and the Illumina strategy gives libraries with very patchy coverage (Ingolia et al., 2009). Life Technologies platforms (SOLiD and Ion Torrent) use duplexed adaptors with degenerate 6 nucleotide overhangs that can hybridise to the fragments and guide ligation with T4 RNA ligase 2 (Patent US 20/00279305 A1 and Figure 6-6). However, nucleotide composition is known to affect hybridisation efficiency (hence why random hexamer priming is biased) and in experiments ligating single fragments and overhang adaptors differential ligation efficiency has been observed (Zhuang et al., 2012). Furthermore the degree of predicted structure correlates with the representation of different variants of miR-1 in SOLiD libraries (Tian et al., 2010).



**Figure 6-6 Schematic of Life Technologies and alternate library generation methods**

The standard library construction for SOLiD and Ion Torrent libraries (left) involves the hybridisation of fragments to duplex adaptors with degenerate 6n ends, these can then act as primers for reverse transcription and PCR amplification (green and blue regions). The alternate method (right) uses ligation of an 5' adenylated, 3' blocked adaptor (yellow) to the 3' end of each fragment which is used to prime reverse transcription using a primer suitable for priming PCR when the fragment is circularised. Barcoding libraries allows multiplexing of sequencing reactions which improves yield and reduces hands on time.

#### 6.4 Use of an alternate library preparation method reduces bias in Ion Torrent sequencing

An alternative approach to the Illumina protocol (Ingolia et al., 2009) is represented in Figure 6-6. In this method, an adaptor is ligated to the 3' end of fragments with a K227Q mutant of trRnl2 that reduces 5' adenylation and thus circularisation of the RNA fragments (Violet et al., 2011). The adaptor is used to prime reverse transcription with a primer containing PCR sites to the Illumina specific sequencing primers. Different RT primers can be designed to allow amplification by PCR primer pairs used for other platforms. The use of this alternate approach does improve coverage relative to the standard Illumina protocol, but no published comparison has been made to the Life Technologies approach.

TrRnl2 K227Q is reported to show a preference for fragments with unstructured 3' ends and those predicted to co-fold with the adaptor (Zhuang et al., 2012). RNA structure is reduced at high temperature so the use of a thermostable ligase may result in more even coverage. A recently characterised K97A point mutant of an RNA ligase from *Methanobacterium thermoautotrophicum* (Mth K97A) is able to ligate only pre-adenylated DNA to the 3' end of ssDNA or RNA (Zhelkovsky and McReynolds, 2012) and so maybe expected to act like trRnl2 but function at 65°C.

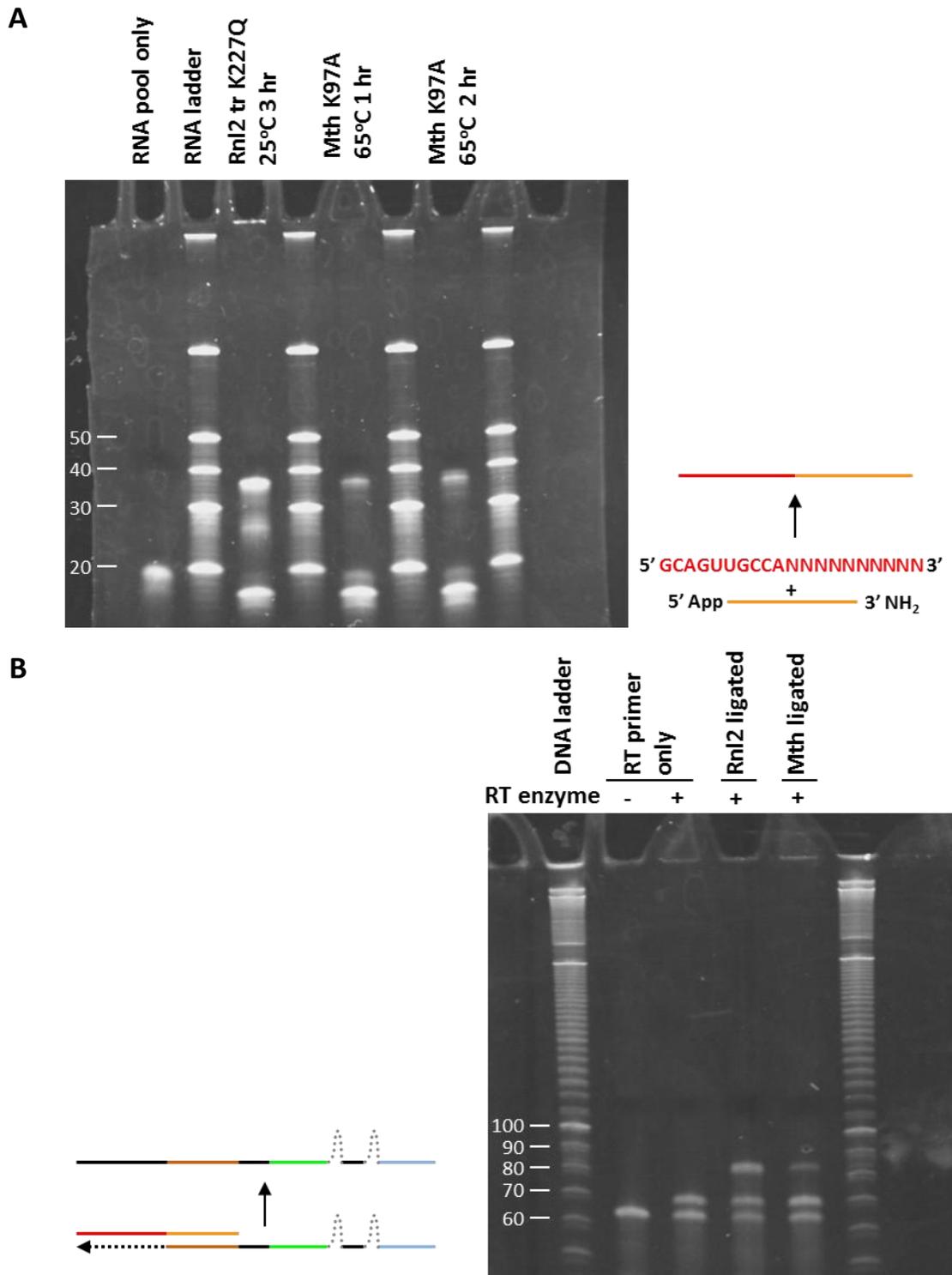
To compare the biases of the standard Life Technologies approach against the alternative approach and to investigate the potential of Mth K97A to replace trRnl2, libraries from an RNA pool of known composition were generated using each method and sequenced on an Ion Torrent. The RNA pool used contained partially degenerate 20 nt RNAs:

5'- GCAGUUGCCANNNNNNNNNN – 3'. Degeneracy at 10 positions yields 1,048,576 ( $4^{10}$ )

unique sequences and as the synthesis of the degenerate section was requested to be performed with equimolar ratios of each NTP at each position it was assumed that all fragments were at an equimolar concentration within the pool. The typical yield from the

largest Ion Torrent chip is ~3 million usable reads. Having more reads than unique sequences increases the likelihood that each sequence will be represented at least once. This allows for characterisation of all biases, rather than just the most prominent ones and accounts for the fact that PCR and/or sequencing reaction biases would be expected to reduce the representation of some fragments even in the absence of ligation biases. Crucially, the PCR amplification, bead preparation and deposition of all 3 libraries (standard, alternate with trRnl2 K227Q, alternate with Mth K97A) were identical and performed by the same researcher (Carolyn Jones, Willis Lab).

Rnl2tr K227Q was able to ligate almost all of the RNA pool to the adenylated linker giving fragments of 47 nt (Figure 6-7 A). By contrast Mth K97A did not ligate all of the RNA pool under the conditions recommended by (Zhelkovsky and McReynolds, 2012) and this was not significantly improved by increasing the reaction time. The enzyme was in excess of the RNA pool but not of the linker which may have been the cause of the incomplete ligation. An inability to ligate all fragments in the pool is not an issue *per se* unless the enzyme shows preference as to which fragments it does ligate. These ligated fragments were successfully reverse transcribed giving a product of ~80 nt as expected (Figure 6-7 B). It is worth noting that RT primer dimers form in the absence of ligated fragment (figure 6-7B) and that if the concentration of RT primer is too high no 80 nt product forms in the presence of ligated fragments (data not shown). The hydrolysis of the RNA following the RT reaction leaves single stranded ends which allow for the circularisation with the thermostable CirLigase I™. This results in a site complementary to the forward primer upstream of the fragment and a site identical to the reverse primer downstream of the linker, permitting amplification of the libraries. These sites are the 8 nt assumed, on the basis of the similarity to the SOLiD protocols, to be within the two proprietary Ion Torrent adaptors. Consistent with this assumption amplified libraries were of the expected size (data not shown).



**Figure 6-7 Alternative library preparation method compatible with Ion Torrent sequencing**

The partially degenerate RNA library shown on the top left was ligated with either truncated Rnl2 K227Q at 25°C or Mth K97A at 65°C.

**A** The products, visualised after separation on a denaturing gel. The 37 nt band was extracted from the gel and reverse transcribed with a custom RT primer then subjected to alkaline hydrolysis to degrade the RNA.

**B** The resulting RT reactions. The ligation products from 1hr Mth K97A treated were used.

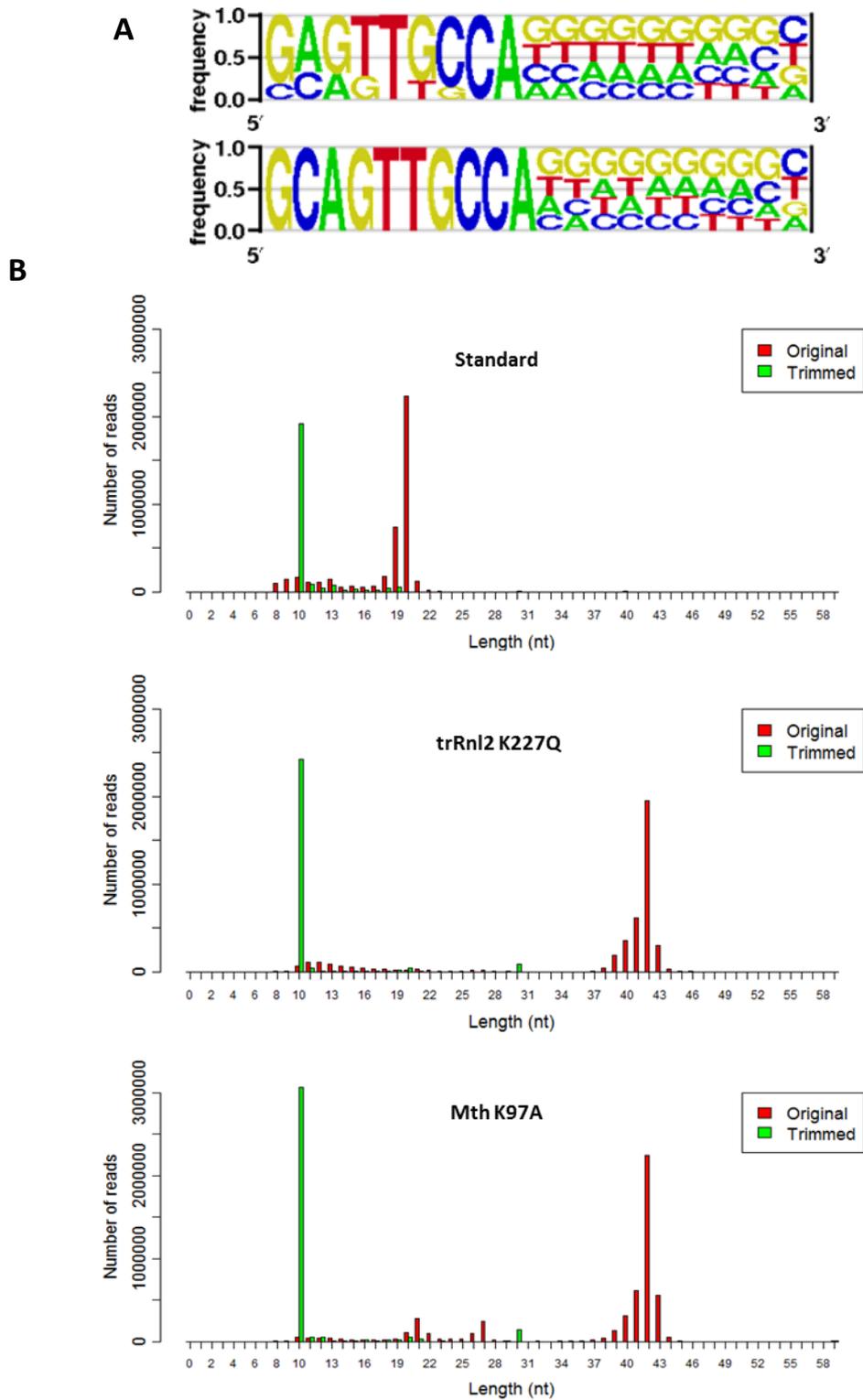
Sequencing data having been obtained, bioinformatic analysis was performed to compare the degree of bias in each method and attempt to identify potential causes of this. These analyses were done with the support of Dr Nicholas Burgoyne (MRC Toxicology Unit).

The standard Ion Torrent pipeline was used to trim the Ion Torrent adaptor sequences away from reads. This resulted in reads of ~20 nt from libraries generated with the standard protocol and ~42 nt for the alternate protocols (Figure 6-8, original reads). It might be expected the alternate libraries would have 41 nt reads: 20 nt fragments+17 nt linker +4 nt sequence between linker binding site and PCR site in the RT primer. A potential explanation for the oversized reads is tailing activity of the reverse transcriptase used, adding an extra base between the 5' end of the fragment and the beginning of the PCR primer site.

To identify possible sequence specific-biases, non-Ion Torrent adaptors and the conserved (i.e. non-degenerate) regions were further trimmed (Figure 6-8, trimmed reads).

There were a substantial number of 19 nt reads in the standard libraries. These truncated reads did not contain all 10 nt of the conserved sequence (Figure 6-8). The reported insertion/deletion rate on the IonTorrent is 1 %, but perhaps this particular sequence is more prone to being misread. However, the last few nucleotides of the conserved sequence were always present in the untrimmed reads, which allowed alignment to the conserved section to guide the trimming. This resulted in almost all trimmed reads being 10 nt therefore any sequence specific biases are unlikely to be due an artefact of trimming.

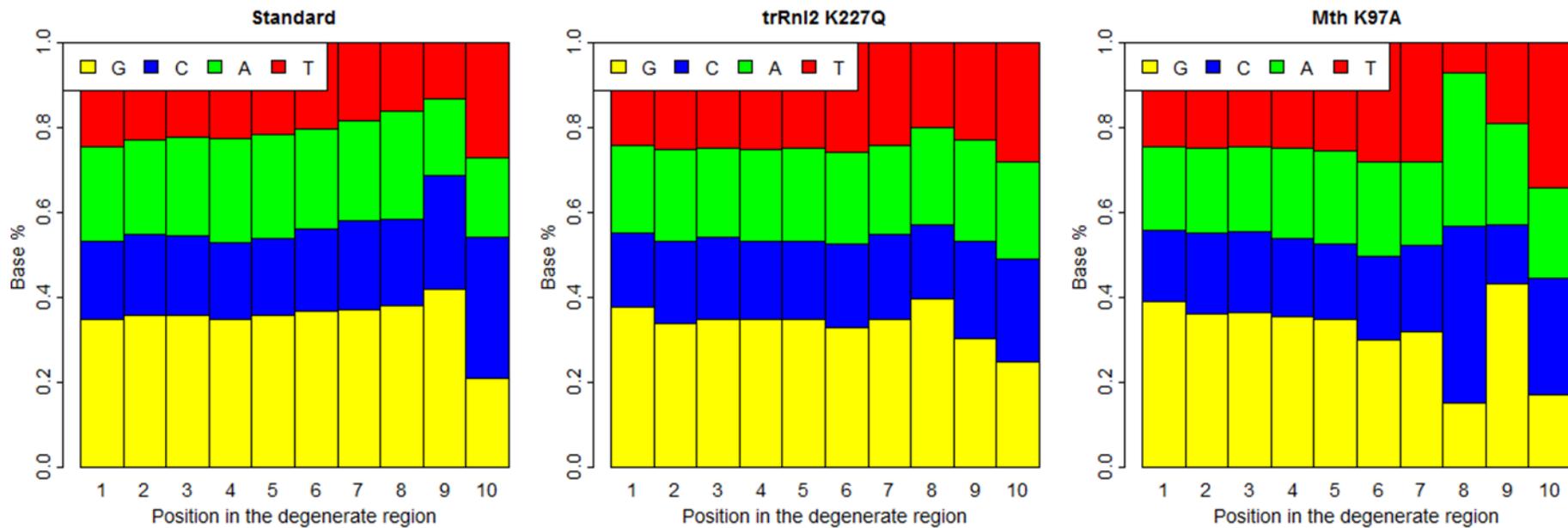
Whilst all libraries showed higher than expected representation of guanines, no position specific biases were observed in the standard protocol or alternate protocol with trRnI2 K277Q (Figure 6-9). However, the use of Mth97A resulted in a profound bias for adenine or cytosine at position 8 in the degenerate region. Notably, the reporter RNA used to initially characterise this enzyme contained an A in the equivalent position (i.e. 2 nt from the end of the read) (Zhelkovsky and McReynolds, 2012).



**Figure 6-8 Trimming of sequencing reads**

**A** Frequency of each nucleotide for untrimmed reads from standard library. Top is 19 nt and bottom is 20 nt. The 10 nt conserved sequence in the partially degenerate RNA pool is GCAGUUGCCA

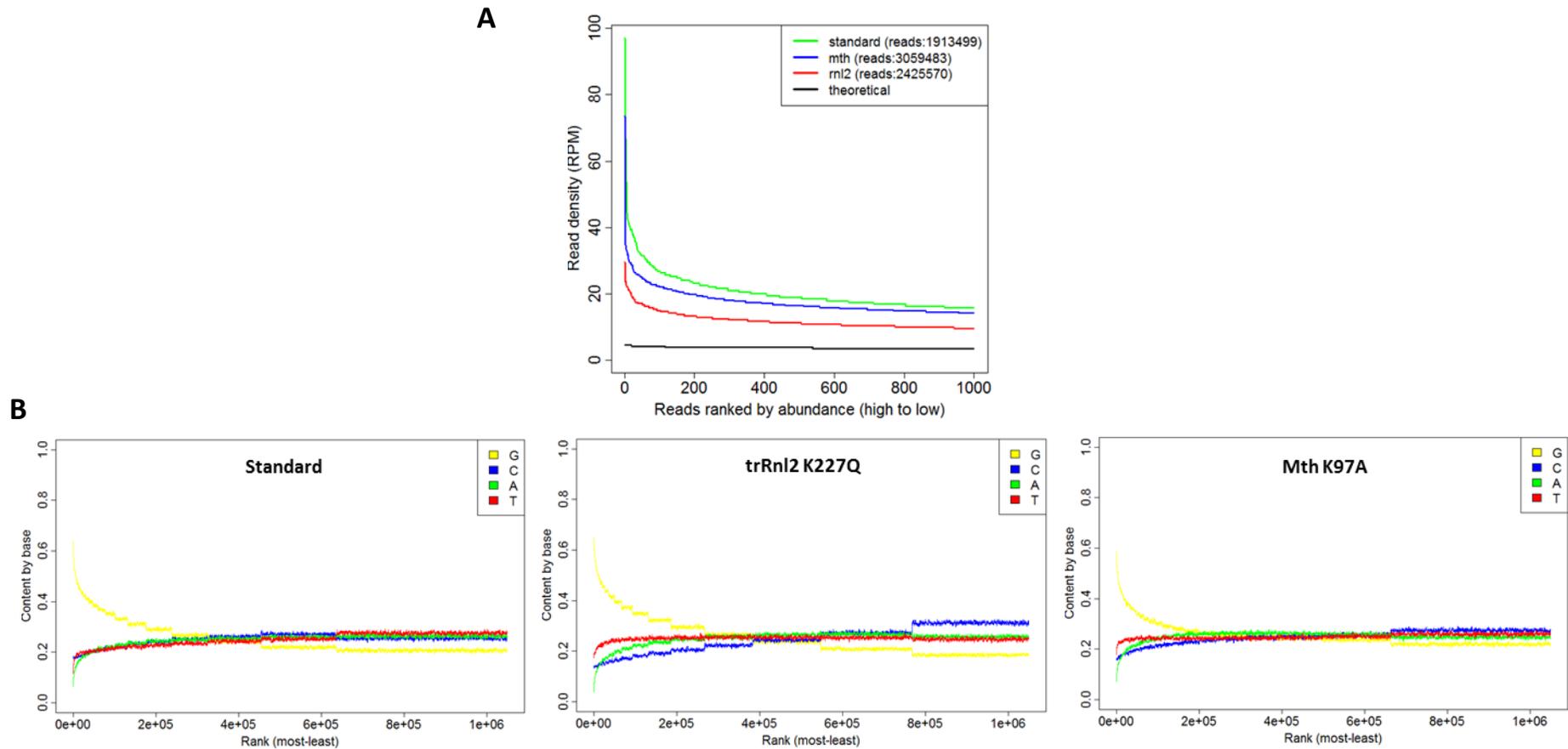
**B** Results of trimming conserved sequence from 5' end and universal adaptor from 3' end.



**Figure 6-9 Sequence specific bias in library constructed with Mth K97A**

Frequency of each base at each position across the degenerate region of all reads. Position 1 is the first nucleotide after the conserved 10 nt.

To compare the level of uneven coverage between the three libraries, the 10 nt degenerate regions of the reads were analysed. In the absence of any biases, the theoretical read distribution or a pool of  $4^{10}$  equimolar fragments should be binomially distributed with probability  $1/4^{10}$  and the number of trials (n) equal to the total number of sequencing reads. Each of the libraries has roughly 3 million reads so the theoretical was modelled as  $X \sim \text{Bin}(1/4^{10}, 3,000,000)$ . Any over-dispersion relative to this implies bias – this was observed for all 3 methods (Figure 6-10 A). However reads from trRnl2 K227Q alternate libraries show a distribution much closer to the theoretical than either of the other two libraries. This suggests that the alternate method will give more even (unbiased) coverage than the standard approach. Somewhat surprisingly, the use of Mth K97A did not reduce bias despite ligation occurring at temperatures less permissive for structure. Sequencing data from the standard and Mth libraries contained over 200 sequences at least 5X more abundant than would be expected without any biases. Initial examination suggested the most over-represented sequences in all libraries show a profound GC % biases, but strictly this is a purely G % bias (Figure 6-10 B). This G % preference may not be a reflection of ligation biases alone. It is possible that the Ion Torrent sequencer suffers from an A->G substitution bias that has not been previously reported, or alternatively that the oligo pool was not correctly formulated by the supplier.

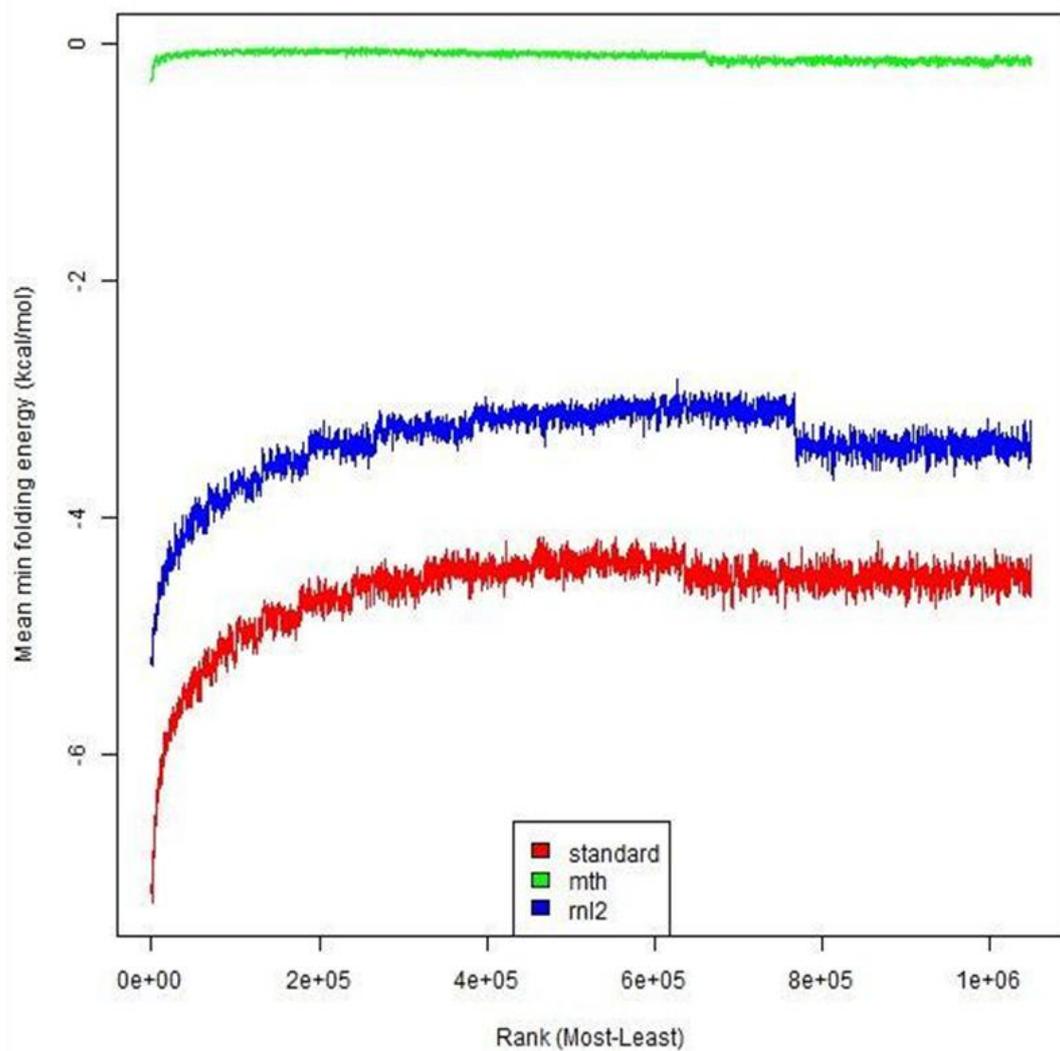


**Figure 6-10 Alternate library generation methods reduce, but do not eliminate, bias**

**A** Read density of 1000 most abundant fragments in each of the three libraries. The theoretical is  $X \sim \text{Binomial}(3 \times 10^6, 1/4^{10})$ .

**B** Percentage content for each nucleotide in last 10 nt of the fragments (ie the degenerate section) using a sliding window of 1000 across a ranked list of fragments.

Secondary structure of the fragment and cofolding with the adaptors have been suggested as possible causes of the ligation bias. It had been hoped that the use of a thermostable ligase to ligate the fragments at 65°C would therefore reduce bias. To understand why this was not the case, all 20 nt fragments that were represented in the sequencing libraries were folded *in silico* under the salt and temperature conditions used for library construction. Interestingly, some fragments were predicted to be structured even at 65°C and folding energy inversely correlated with over-representation in all three sequenced libraries (Figure 6-11).



**Figure 6-11 Over-represented fragments are predicted to have lower folding energies**

The predicted folding energies of the 10 nt trimmed read plus associated conserved region (i.e. a full length of 20 nt) were computed by Nick Burgoyne using ViennaFold. The plot shows the mean folding energy across sliding window of 1000 across the reads ranked by abundance (high to low).

It should be noted that fragments were denatured at 80°C in the alternate method and 65°C in the standard method prior to ligation/hybridisation so it may be that some fragments that were predicted to be structured were not at the time of ligation. In the case of the standard protocol, fragments were left for 5 min at 16°C to hybridise to adaptors prior to ligation overnight so it is theoretically possible all fragments hybridised perfectly to the adaptors without refolding. The likelihood of this would be affected by the relative abundance of the adaptors (the more abundant the adaptors the increased probability of inter-molecular hybridisation vs intra-molecular hybridisation). If hybridisation is extensive in the standard protocol then the cofolding of the adaptors and the fragments may play a role in ligation efficiency. It has been shown that Rnl2 can ligate fragments that create a 1-2 nt overhang when hybridised to a duplexed adaptor (Sorefan et al., 2012). If this occurs in Life Technologies protocols then the PCR products will contain mismatches as the upper strand (containing the fragment) will be longer than the lower one. Consistent with this, a considerable fraction of reads from the standard library were shorter than 20 nt and showed deletions in the conserved region were occurring in the 5' most nucleotides which hybridise with the adaptor overhangs (Figure 6-8 A).

A bioinformatics approach to assess how significant mishybridisation is in the biases observed in the standard library, could not be performed because the concentration of adaptors is a proprietary secret and existing folding programs cannot process duplexes.

*In silico* analysis of cofolding in the alternate protocol is possible and would be worthy of further investigation.

## **6.5 Discussion**

This chapter has focussed on the limitations of next generation sequencing. This is not to argue that the technique is without merit (nothing could be further from the truth!), but

rather to suggest where potential limitations may exist that could result in overinterpretation of sequencing data.

As described in the introduction of this chapter there are essentially three potential areas that could be considered: fragment generation, library construction and sequencing. Biases at any of these stages will lead to patchy/incomplete coverage of the transcriptome/translatome.

A fragment generation method that results in an RNA pool that reflects the biological system being studied is crucial, particularly when trying to infer ribosome density with ribosome profiling. Whilst the majority of polysomes are reduced to monosomes by RNase digestion, a bi-ribosome peak was observed in two different, commonly used buffers and was not reduced by increased RNase concentration. The ~30 nt mRNA fragments found in this peak are likely the result of either overlap between peaks or the dimerisation of ribosomes after digestion. Assuming ribosomes are not specialised (i.e. structurally different) at different locations then the dimers are likely formed in an unbiased fashion so are a curiosity rather than a problem.

The existence of ~60 nt fragments in this peak is more concerning as it suggests there might be disomes formed due to ribosome stacking at regions that are translated slowly/have pause sites.

Attempts have been made to investigate pause sites using ribosome profiling (Ingolia et al., 2011). In that report, a pause site was defined as a region with more than 25 times the median translatome density on the transcript. This was almost double the largest difference in regions assessed in the chemically fragmented transcriptome library so is unlikely to be purely a result of sequencing bias. They reported no secondary accumulation of ribosomes upstream of these sites, and so concluded that ribosomes do not stack in MEFs or if they do, upstream ribosomes displace downstream ones. Given that we find ~60 nt fragments in the bi-ribosome peak from MEFs digested with the same buffer used by (Ingolia et al., 2011) it at

least seems plausible that there is some stacking and that this results in fragments from these regions not sequenced. This would mean that regions of highest ribosome density within a transcript are paradoxically represented as areas of lower density! To investigate this further it will be necessary to sequence these disome fragments.

It may well be that in many experiments there is no significant difference in stalling and thus the disome phenomenon is not significant. However, recent work suggests the prokaryotic homologue of eIF5A (EF-P) is responsible for ribosome stacking at polyproline stretches (Doerfel et al., 2013, Ude et al., 2013). eIF5A contains an unusual amino acid, hypusine, that is absolutely required for the protein to promote translation elongation (Saini et al., 2009). Hypusination involves the transfer of 4-aminobutyl moiety from spermidine, in a process involving both deoxyhypusine synthase and deoxyhypusine hydroxylase (Park, 2006).

Ribosome profiling in conditions where any of these three factors may be down-regulated would thus be of interest.

Furthermore, there are other circumstances in which a reduction in elongation rates may occur (e.g. cold shock). This reduction in elongation rates may result in an increase in stacking at codons that are translated more slowly. As described in Chapter 3, the presence of codons decoded by less abundant tRNA – giving the transcripts a low tRNA adaption index (tAI) value (dos Reis et al., 2004) – is associated with reduced translational efficiency. It is assumed this is because such codons are translated more slowly. However, ribosome profiling of MEFs with the harringtonine run-off protocol showed no difference in translation elongation rates of transcripts with a lower tAI (Ingolia et al., 2011). It is worth noting that these calculations could only be reliably performed on abundant, relatively well translated mRNAs. Given selective pressures it may be that all these well translated mRNAs have comparatively high tAI relative to the whole transcriptome and so show no differential elongation.

Recent work has shown that ribosome profiling can be used to detect stalled ribosomes in response to proteotoxic stress (Liu et al., 2013). However, it remains possible that this increase in occupancy is actually an under-representation of the extent of ribosome stalling as the bi-ribosome peak was not examined. Ultimately it remains unclear that RNase digestion fully captures all locations of ribosome density and that this may have important implications when interpreting ribosome profiling data. To limit the possibility of misinterpretation of data it might be wise to always perform full sucrose gradient analysis of RNase digested lysates to compare the bi-ribosome peak between conditions rather than simply pellet ribosomes as recommended in more recent protocols (Ingolia et al., 2012).

The second area addressed by this chapter was library generation. This was initially motivated by the observation of patchy coverage in the transcriptome libraries. Whilst that was in part due to enzymatic fragmentation of the polyA<sup>+</sup> RNA, library generation is still a potential source of bias.

The use of an alternate library generation protocol reduced over-representation of fragments compared to the standard Life Technologies methods. Beyond a desire to identify which transcripts are genuinely highly expressed in a system, reducing over-representation may be important in the light of the current need to employ thresholds when detecting differential translation in ribosome profiling data (see Chapter 3). If certain fragments are over-represented they necessarily reduce the reads that map to other transcripts. This may limit the ability to accurately determine TE values for these transcripts. In addition, the use of sequencing to identify binding motifs or to assess which microRNAs are most abundant within a given cell may be skewed by these biases.

All libraries showed over-representation with some transcripts more than 10X abundant than theoretical. Libraries generated using the alternate method showed less over-representation than the standard IonTorrent library. Unfortunately the use of the thermostable ligase Mth

K97A did not improve upon the results obtained with trRnl2 K227Q. This is in part due to a position specific bias for adenine or guanine 3 nucleotides from the 3' end. In all libraries the most over-represented sequences were predicted to be structured under the conditions used for ligation. Both Rnl2 and Mth ligases are part of a conserved class of enzymes involved in repair of nicked dsRNA (Ho and Shuman, 2002) which may explain why the most over-represented fragments were those predicted to have secondary structure. Moreover, a role in repairing nicked dsRNA would favour those fragments that are able to co-fold more effectively with adaptor sequences and would make for an interesting further study using the sequencing data obtained, though this is much more challenging bioinformatically.

Replacing the ends of RNA linkers used in the standard Illumina protocol with degenerate sequences, so as to reduce the consistency of cofolding, reduces but does not eliminate over-dispersion (Sorefan et al., 2012). It may be of interest to see if the alternate library preparation method with trRnl2 can be improved by adopting this approach.

## Chapter 7. Summary

This thesis has described studies of translational control in a variety of different contexts.

There were several challenges faced working on the original thesis proposal: the role of  $\text{tRNA}_i^{\text{Met}}$  in oncogenic transformation. In support of the data shown in the original paper (Marshall et al., 2008), the data in Chapter 3, show an increase in both global translation rate and of selected messages following the re-establishment of high  $\text{tRNA}_i^{\text{Met}}$  levels in a NIH 3T3 cell line. However this phenotype was not stable, so despite time-consuming efforts to retain a consistent phenotype we must conclude  $\text{tRNA}_i^{\text{Met}}$  cannot drive oncogenic transformation in MEFs. Following an investigation into scientific misconduct within Professor Bob White's group, the original paper that claimed to show increased  $\text{tRNA}_i^{\text{Met}}$  levels could drive transformation was retracted. However, despite the lack of biological significance, this work allowed the development of the ribosome profiling method using the SOLiD sequencing platform. It was these experiences that motivated the final of three additional projects that were undertaken, all of which were productive.

The first, performed in collaboration with Dr Owen Sansom's group, used genetic manipulation localised to the intestinal epithelia of *in vivo* and *in vitro* mouse models to investigate mTOR signalling in colorectal cancer.

Inhibition of mTORC1 with rapamycin prevented the onset of benign (but pathological) intestinal tumours when administered shortly after APC loss, and improved survival when treatment commenced upon development of a sick phenotype. This suggests that rapamycin would be a promising therapy for FAP patients, which would be a significant improvement on the current recommended management: colon resection.

That deletion of Raptor provides little additional benefit beyond that conferred by rapamycin treatment suggests that proliferative mTORC1 signalling in the intestine is mediated predominantly via S6K and not 4E-BPs. S6K has multiple targets which affect different mechanistic processes; this thesis presented five distinct pieces of evidence to suggest that increased translational elongation is the key underlying mechanism by which Wnt-driven mTORC1 signalling leads to proliferation. First, the polysomal association in APC<sup>fl/fl</sup> mice was reduced relative to wild type, which in the absence of an uncoupling of protein synthesis from proliferation would imply a reduced ribosome transit time. Second, treatment of APC<sup>fl/fl</sup> mice with the elongation inhibitor cycloheximide resulted in a reduction in crypt proliferation comparable to rapamycin treatment. Third, APC<sup>fl/fl</sup> crypts cultured *in vitro* show an increase in protein synthesis rate and a decrease in ribosome run-off time when treated with harringtonine. Forth, systemic deletion of eEF2k, a negative regulator of elongation downstream of mTORC1, removes rapamycin induced inhibition of Wnt-driven proliferation. Finally, RNAi knockdown of eEF2, the downstream target of eEF2k, prevents intestinal regeneration in *Drosophila*. Collectively these data suggest that unlike in other organs, in the intestine, the increase in protein synthesis required to drive proliferation is due primarily to an increase in translation elongation and not initiation.

Approximately 40 % of CRC patients present with a K-Ras mutant tumour and respond poorly to existing therapies. The common G12D gain of function mutation conferred resistance to inhibition of mTORC1 (by rapamycin or Raptor knockout) and to both mTORC1 and mTORC2 (with pp242). Intriguingly, deletion of Rictor alone reduced proliferation to almost wild type levels in APC-null, K-Ras mutant epithelia, but not APC-null alone. In the absence of reports of other plausible Rictor targets that could mediate this effect, we hypothesise that selective inhibition of mTORC2 may be an effective therapy for Ras-mutant CRC. Investigating the effect of loss of other mTORC2 essential components, such as mSIN1, would be of value in

investigating this theory. Investigation using a selective inhibitor of mTORC2 would be ideal, but these have not been developed.

The second collaboration was within the Willis group, supported the research carried out by Dr Amandine Bastide to characterise the cold shock response. It was shown that translation elongation is vastly reduced under these conditions, which explains the maintenance of heavy polysomes at 32°C despite a 60 % reduction in protein synthesis rates. Notably these changes in elongation rate were not mediated via eEF2k. To investigate the possibility of differential regulation at the level of elongation ribosome profiling was performed.

Extensive read density was observed in the 3'UTRs of mRNAs in response to cold shock, although ultimately the use of RNase protection assays revealed that this was likely to be a technical artefact.

In the final project it was decided to better characterise the potential limitations of the ribosome profiling technique and of sequencing more generally. The great advantage of ribosome profiling is the ability to compare the ribosome occupancy at multiple different features such as coding sequences, uORFs or start sites. However this is best achieved when biases are minimised so that sequencing data is an accurate representation of the true biological situation.

A potentially troubling issue, consistently observed, was that, unlike in the schematics, RNase I digestion does not digest polysomes to monosomes. Through the use of northern blotting and RNase protection assays we have identified the formation of both ribosome dimers after digestion and digestion resistance disomes. Two pieces of additional work are necessary to confirm the significance of this. First, a more direct method of assessment of the dimer/disome RNA fragments as the data seen may be due to non-specific binding.

Attempts to use transfection of radiolabelled RNA were unsuccessful both because of rapid RNA degradation and the possible presence of a functional uORF within the vector. However,

using a transcript with longer half-life and in a vector that doesn't have the uORF may address this issue. The second key experiment would be the sequencing of ~60 nt fragments within the bi-ribosome peak. If there is no difference in the representation of transcripts within this peak compared to the 30 nt fragments in the monosome peak then it would be safe to infer that no bias is introduced at the RNase digestion step.

Given the initial aim to characterise and reduce biases, it was pleasing that the use of an alternate library preparation method showed substantially reduced bias relative to the standard Life Technologies protocol. However, consistent with previous reports from other library protocols, reads with a high degree of secondary structure were over represented. Unfortunately the use of a thermostable enzyme (Mth K97A) did not rectify this problem. It is unclear whether this was due to additional biases of Mth K97A – for example a sequence specific bias was observed – or whether structure is correlative with, but not causative of bias. Irrespective, it would be interesting to use the dataset generated to investigate whether over represented fragments are also predicted to co-fold with adaptors, in agreement with the work of other groups. If this is the case then the use of partially degenerate adaptors would be expected to result in more representative libraries, though this may create difficulties in adaptor trimming.

A final area that would be useful to investigate would be increasing the effective read depth that can be achieved with ribosome profiling. Two potential approaches would be to develop robust statistical methods to reliably analyse low count data and to develop subtractive hybridisation methods against contaminating rRNA fragments. The latter has been described by (Ingolia et al., 2012), but it is unclear by how much this improves read depth or its applicability to different organisms.

In conclusion, my research has shown how global changes in post-transcriptional regulation of gene expression are associated with the development and progression of neoplasia, in

particular changes in elongation rates. Detailed analysis of ribosome profiling suggests that some further development of this technique is required, however this method is already beginning to revolutionise the analysis of the translome and will only become more powerful and more accessible as next-generation sequencing technologies mature.

## Bibliography

- AALTONEN, L. A., SALOVAARA, R., KRISTO, P., CANZIAN, F., HEMMINKI, A., PELTOMAKI, P., CHADWICK, R. B., KAARIAINEN, H., ESKELINEN, M., JARVINEN, H., MECKLIN, J. P. & DE LA CHAPELLE, A. 1998. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med*, 338, 1481-7.
- ADAMS, J. M. & CORY, S. 2007. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene*, 26, 1324-37.
- AIRD, D., ROSS, M. G., CHEN, W. S., DANIELSSON, M., FENNELL, T., RUSS, C., JAFFE, D. B., NUSBAUM, C. & GNIRKE, A. 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol*, 12, R18.
- AITKEN, C. E. & LORSCH, J. R. 2012. A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol*, 19, 568-76.
- AKSAMITIENE, E., KIYATKIN, A. & KHOLODENKO, B. N. 2012. Cross-talk between mitogenic Ras/MAPK and survival PI3K/Akt pathways: a fine balance. *Biochem Soc Trans*, 40, 139-46.
- AL-FAGEEH, M. B., MARCHANT, R. J., CARDEN, M. J. & SMALES, C. M. 2006. The cold-shock response in cultured mammalian cells: harnessing the response for the improvement of recombinant protein production. *Biotechnol Bioeng*, 93, 829-35.
- ALAIN, T., MORITA, M., FONSECA, B. D., YANAGIYA, A., SIDDIQUI, N., BHAT, M., ZAMMIT, D., MARCUS, V., METRAKOS, P., VOYER, L. A., GANDIN, V., LIU, Y., TOPISIROVIC, I. & SONENBERG, N. 2012. eIF4E/4E-BP Ratio Predicts the Efficacy of mTOR Targeted Therapies. *Cancer Res*, 72, 6468-6476.
- ALI, I. K., MCKENDRICK, L., MORLEY, S. J. & JACKSON, R. J. 2001. Truncated initiation factor eIF4G lacking an eIF4E binding site can support capped mRNA translation. *EMBO J*, 20, 4233-42.
- ALKALAEVA, E. Z., PISAREV, A. V., FROLOVA, L. Y., KISSELEV, L. L. & PESTOVA, T. V. 2006. In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. *Cell*, 125, 1125-36.
- ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol*, 11, R106.
- ANDREU, P., COLNOT, S., GODARD, C., GAD, S., CHAFEY, P., NIWA-KAWAKITA, M., LAURENT-PUIG, P., KAHN, A., ROBINE, S., PERRET, C. & ROMAGNOLO, B. 2005. Crypt-restricted proliferation and commitment to the Paneth cell lineage following Apc loss in the mouse intestine. *Development*, 132, 1443-51.
- APIDIANAKIS, Y. & RAHME, L. G. 2011. Drosophila melanogaster as a model for human intestinal infection and pathology. *Dis Model Mech*, 4, 21-30.
- ASHTON, G. H., MORTON, J. P., MYANT, K., PHESSÉ, T. J., RIDGWAY, R. A., MARSH, V., WILKINS, J. A., ATHINEOS, D., MUNCAN, V., KEMP, R., NEUFELD, K., CLEVERS, H., BRUNTON, V., WINTON, D. J., WANG, X., SEARS, R. C., CLARKE, A. R., FRAME, M. C. & SANSOM, O. J. 2010. Focal adhesion kinase is required for intestinal regeneration and tumorigenesis downstream of Wnt/c-Myc signaling. *Dev Cell*, 19, 259-69.
- BAR-PELED, L., SCHWEITZER, L. D., ZONCU, R. & SABATINI, D. M. 2012. Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell*, 150, 1196-208.
- BARREAU, C., DUTERTRE, S., PAILLARD, L. & OSBORNE, H. B. 2006. Liposome-mediated RNA transfection should be used with caution. *RNA*, 12, 1790-3.
- BAUER, C., BRASS, N., DIESINGER, I., KAYSER, K., GRASSER, F. A. & MEESE, E. 2002. Overexpression of the eukaryotic translation initiation factor 4G (eIF4G-1) in squamous cell lung carcinoma. *Int J Cancer*, 98, 181-5.
- BECKER, T., FRANCKENBERG, S., WICKLES, S., SHOEMAKER, C. J., ANGER, A. M., ARMACHE, J. P., SIEBER, H., UNGEWICKELL, C., BERNINGHAUSEN, O., DABERKOW, I., KARCHER, A.,

- THOMM, M., HOPFNER, K. P., GREEN, R. & BECKMANN, R. 2012. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature*, 482, 501-6.
- BENJAMINI, Y. & SPEED, T. P. 2012. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res*, 40, e72.
- BIANCHINI, A., LOIARRO, M., BIELLI, P., BUSA, R., PARONETTO, M. P., LORENI, F., GEREMIA, R. & SETTE, C. 2008. Phosphorylation of eIF4E by MNKs supports protein synthesis, cell cycle progression and proliferation in prostate cancer cells. *Carcinogenesis*, 29, 2279-88.
- BLAGDEN, S. P. & WILLIS, A. E. 2011. The biological and therapeutic relevance of mRNA translation in cancer. *Nat Rev Clin Oncol*, 8, 280-91.
- BODKIN, D. K. & KNUDSON, D. L. 1985. Assessment of sequence relatedness of double-stranded RNA genes by RNA-RNA blot hybridization. *J Virol Methods*, 10, 45-52.
- BREITLING, R., ARMENGAUD, P., AMTMANN, A. & HERZYK, P. 2004. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett*, 573, 83-92.
- BROWN, E. J., BEAL, P. A., KEITH, C. T., CHEN, J., SHIN, T. B. & SCHREIBER, S. L. 1995. Control of p70 s6 kinase by kinase activity of FRAP in vivo. *Nature*, 377, 441-6.
- BROWNE, G. J. & PROUD, C. G. 2002. Regulation of peptide-chain elongation in mammalian cells. *Eur J Biochem*, 269, 5360-8.
- BRUGAROLAS, J., LEI, K., HURLEY, R. L., MANNING, B. D., REILING, J. H., HAFEN, E., WITTERS, L. A., ELLISEN, L. W. & KAELIN, W. G., JR. 2004. Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes Dev*, 18, 2893-904.
- BRUHN, M. A., PEARSON, R. B., HANNAN, R. D. & SHEPPARD, K. E. 2010. Second AKT: The rise of SGK in cancer signalling. *Growth Factors*, 28, 394-408.
- BRUNET, A., PARK, J., TRAN, H., HU, L. S., HEMMING, B. A. & GREENBERG, M. E. 2001. Protein kinase SGK mediates survival signals by phosphorylating the forkhead transcription factor FKHRL1 (FOXO3a). *Mol Cell Biol*, 21, 952-65.
- BUSHELL, M., STONELEY, M., KONG, Y. W., HAMILTON, T. L., SPRIGGS, K. A., DOBBYN, H. C., QIN, X., SARNOW, P. & WILLIS, A. E. 2006. Polypyrimidine tract binding protein regulates IRES-mediated gene expression during apoptosis. *Mol Cell*, 23, 401-12.
- CANCER GENOME ATLAS, N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487, 330-7.
- CAO, H., ZHU, Q., HUANG, J., LI, B., ZHANG, S., YAO, W. & ZHANG, Y. 2009. Regulation and functional role of eEF1A2 in pancreatic carcinoma. *Biochem Biophys Res Commun*, 380, 11-6.
- CARLBERG, U., NILSSON, A. & NYGARD, O. 1990. Functional properties of phosphorylated elongation factor 2. *Eur J Biochem*, 191, 639-45.
- CARRACEDO, A. & PANDOLFI, P. P. 2008. The PTEN-PI3K pathway: of feedbacks and cross-talks. *Oncogene*, 27, 5527-41.
- CASTELLANO, E. & DOWNWARD, J. 2011. RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer*, 2, 261-74.
- CECH, T. R. 2000. Structural biology. The ribosome is a ribozyme. *Science*, 289, 878-9.
- CHEN, C. Y. & SARNOW, P. 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science*, 268, 415-7.
- CHENG, Z., SAITO, K., PISAREV, A. V., WADA, M., PISAREVA, V. P., PESTOVA, T. V., GAJDA, M., ROUND, A., KONG, C., LIM, M., NAKAMURA, Y., SVERGUN, D. I., ITO, K. & SONG, H. 2009. Structural insights into eRF3 and stop codon recognition by eRF1. *Genes Dev*, 23, 1106-18.
- CHILD, S. J., MILLER, M. K. & GEBALLE, A. P. 1999. Translational control by an upstream open reading frame in the HER-2/neu transcript. *J Biol Chem*, 274, 24335-41.

- CHOO, A. Y., YOON, S. O., KIM, S. G., ROUX, P. P. & BLENIS, J. 2008. Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc Natl Acad Sci U S A*, 105, 17414-9.
- CLEMENS, M. J., BUSHHELL, M. & MORLEY, S. J. 1998. Degradation of eukaryotic polypeptide chain initiation factor (eIF) 4G in response to induction of apoptosis in human lymphoma cell lines. *Oncogene*, 17, 2921-31.
- COLEMAN, L. J., PETER, M. B., TEALL, T. J., BRANNAN, R. A., HANBY, A. M., HONARPISHEH, H., SHAABAN, A. M., SMITH, L., SPEIRS, V., VERGHESE, E. T., MCELWAINE, J. N. & HUGHES, T. A. 2009. Combined analysis of eIF4E and 4E-binding protein expression predicts breast cancer survival and estimates eIF4E activity. *Br J Cancer*, 100, 1393-9.
- CONNELL, S. R., TAKEMOTO, C., WILSON, D. N., WANG, H., MURAYAMA, K., TERADA, T., SHIROUZU, M., ROST, M., SCHULER, M., GIESEBRECHT, J., DABROWSKI, M., MIELKE, T., FUCINI, P., YOKOYAMA, S. & SPAHN, C. M. 2007. Structural basis for interaction of the ribosome with the switch regions of GTP-bound elongation factors. *Mol Cell*, 25, 751-64.
- CORDERO, J. B., STEFANATOS, R. K., SCOPELLITI, A., VIDAL, M. & SANSOM, O. J. 2012. Inducible progenitor-derived Wingless regulates adult midgut regeneration in *Drosophila*. *EMBO J*, 31, 3901-17.
- CREAMER, B., SHORTER, R. G. & BAMFORTH, J. 1961. The turnover and shedding of epithelial cells. I. The turnover in the gastro-intestinal tract. *Gut*, 2, 110-8.
- CYBULSKI, N., POLAK, P., AUWERX, J., RUEGG, M. A. & HALL, M. N. 2009. mTOR complex 2 in adipose tissue negatively controls whole-body growth. *Proc Natl Acad Sci U S A*, 106, 9902-7.
- DHILLON, A. S., HAGAN, S., RATH, O. & KOLCH, W. 2007. MAP kinase signalling pathways in cancer. *Oncogene*, 26, 3279-90.
- DI NICOLANTONIO, F., ARENA, S., TABERNERO, J., GROSSO, S., MOLINARI, F., MACARULLA, T., RUSSO, M., CANCELLIERE, C., ZECCHIN, D., MAZZUCHELLI, L., SASAZUKI, T., SHIRASAWA, S., GEUNA, M., FRATTINI, M., BASELGA, J., GALLICCHIO, M., BIFFO, S. & BARDELLI, A. 2010. Deregulation of the PI3K and KRAS signaling pathways in human cancer cells determines their response to everolimus. *J Clin Invest*, 120, 2858-66.
- DIENSTMANN, R., SERPICO, D., RODON, J., SAURA, C., MACARULLA, T., ELEZ, E., ALSINA, M., CAPDEVILA, J., PEREZ-GARCIA, J., SANCHEZ-OLLE, G., AURA, C., PRUDKIN, L., LANDOLFI, S., HERNANDEZ-LOSA, J., VIVANCOS, A. & TABERNERO, J. 2012. Molecular profiling of patients with colorectal cancer and matched targeted therapy in phase I clinical trials. *Mol Cancer Ther*, 11, 2062-71.
- DOERFEL, L. K., WOHLGEMUTH, I., KOTHE, C., PESKE, F., URLAUB, H. & RODNINA, M. V. 2013. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science*, 339, 85-8.
- DOHM, J. C., LOTTAZ, C., BORODINA, T. & HIMMELBAUER, H. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res*, 36, e105.
- DONZE, O., JAGUS, R., KOROMILAS, A. E., HERSHEY, J. W. & SONENBERG, N. 1995. Abrogation of translation initiation factor eIF-2 phosphorylation causes malignant transformation of NIH 3T3 cells. *EMBO J*, 14, 3828-34.
- DOROVKOV, M. V., PAVUR, K. S., PETROV, A. N. & RYAZANOV, A. G. 2002. Regulation of elongation factor-2 kinase by pH. *Biochemistry*, 41, 13444-50.
- DOS REIS, M., SAVVA, R. & WERNISCH, L. 2004. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res*, 32, 5036-44.
- DOWLING, R. J., TOPISIROVIC, I., ALAIN, T., BIDINOSTI, M., FONSECA, B. D., PETROULAKIS, E., WANG, X., LARSSON, O., SELVARAJ, A., LIU, Y., KOZMA, S. C., THOMAS, G. &

- SONENBERG, N. 2010a. mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science*, 328, 1172-6.
- DOWLING, R. J., TOPISIROVIC, I., FONSECA, B. D. & SONENBERG, N. 2010b. Dissecting the role of mTOR: lessons from mTOR inhibitors. *Biochim Biophys Acta*, 1804, 433-9.
- DRYGIN, D., LIN, A., BLIESATH, J., HO, C. B., O'BRIEN, S. E., PROFFITT, C., OMORI, M., HADDACH, M., SCHWAEBE, M. K., SIDDIQUI-JAIN, A., STREINER, N., QUIN, J. E., SANIJ, E., BYWATER, M. J., HANNAN, R. D., RYCKMAN, D., ANDERES, K. & RICE, W. G. 2011. Targeting RNA Polymerase I with an Oral Small Molecule CX-5461 Inhibits Ribosomal RNA Synthesis and Solid Tumor Growth. *Cancer Research*, 71, 1418-1430.
- ELORZA, A., SORO-ARNAIZ, I., MELENDEZ-RODRIGUEZ, F., RODRIGUEZ-VAELLO, V., MARSBOOM, G., DE CARCER, G., ACOSTA-IBORRA, B., ALBACETE-ALBACETE, L., ORDONEZ, A., SERRANO-OVIEDO, L., GIMENEZ-BACHS, J. M., VARA-VEGA, A., SALINAS, A., SANCHEZ-PRIETO, R., MARTIN DEL RIO, R., SANCHEZ-MADRID, F., MALUMBRES, M., LANDAZURI, M. O. & ARAGONES, J. 2012. HIF2alpha acts as an mTORC1 activator through the amino acid carrier SLC7A5. *Mol Cell*, 48, 681-91.
- FALVEY, A. K. & STAEHELIN, T. 1970. Structure and function of mammalian ribosomes. I. Isolation and characterization of active liver ribosomal subunits. *J Mol Biol*, 53, 1-19.
- FAN, H. & PENMAN, S. 1970. Regulation of protein synthesis in mammalian cells. II. Inhibition of protein synthesis at the level of initiation during mitosis. *J Mol Biol*, 50, 655-70.
- FEARON, E. R. & VOGELSTEIN, B. 1990. A genetic model for colorectal tumorigenesis. *Cell*, 61, 759-67.
- FELS, D. R. & KOUMENIS, C. 2006. The PERK/eIF2alpha/ATF4 module of the UPR in hypoxia resistance and tumor growth. *Cancer Biol Ther*, 5, 723-8.
- FERLAY, J., SHIN, H. R., BRAY, F., FORMAN, D., MATHERS, C. & PARKIN, D. M. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*, 127, 2893-917.
- FIRCZUK, H., KANNAMBATH, S., PAHLE, J., CLAYDON, A., BEYNON, R., DUNCAN, J., WESTERHOFF, H., MENDES, P. & MCCARTHY, J. E. 2013. An in vivo control map for the eukaryotic mRNA translation machinery. *Mol Syst Biol*, 9, 635.
- FRESNO, M., JIMENEZ, A. & VAZQUEZ, D. 1977. Inhibition of translation in eukaryotic systems by harringtonine. *Eur J Biochem*, 72, 323-30.
- FUKUCHI-SHIMOGORI, T., ISHII, I., KASHIWAGI, K., MASHIBA, H., EKIMOTO, H. & IGARASHI, K. 1997. Malignant transformation by overproduction of translation initiation factor eIF4G. *Cancer Research*, 57, 5041-4.
- GAN, X., WANG, J., SU, B. & WU, D. 2011. Evidence for direct activation of mTORC2 kinase activity by phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem*, 286, 10998-1002.
- GERASHCHENKO, M. V., LOBANOV, A. V. & GLADYSHEV, V. N. 2012. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc Natl Acad Sci U S A*, 109, 17394-9.
- GERBE, F., VAN ES, J. H., MAKRINI, L., BRULIN, B., MELLITZER, G., ROBINE, S., ROMAGNOLO, B., SHROYER, N. F., BOURGAUX, J. F., PIGNODEL, C., CLEVERS, H. & JAY, P. 2011. Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *J Cell Biol*, 192, 767-80.
- GINGRAS, A. C., SVITKIN, Y., BELSHAM, G. J., PAUSE, A. & SONENBERG, N. 1996. Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. *Proc Natl Acad Sci U S A*, 93, 5578-83.
- GOEL, A., ARNOLD, C. N., NIEDZWIECKI, D., CARETHERS, J. M., DOWELL, J. M., WASSERMAN, L., COMPTON, C., MAYER, R. J., BERTAGNOLLI, M. M. & BOLAND, C. R. 2004. Frequent inactivation of PTEN by promoter hypermethylation in microsatellite instability-high sporadic colorectal cancers. *Cancer Research*, 64, 3014-21.

- GRADI, A., IMATAKA, H., SVITKIN, Y. V., ROM, E., RAUGHT, B., MORINO, S. & SONENBERG, N. 1998. A novel functional human eukaryotic translation initiation factor 4G. *Mol Cell Biol*, 18, 334-42.
- GUERTIN, D. A., STEVENS, D. M., THOREEN, C. C., BURDS, A. A., KALAANY, N. Y., MOFFAT, J., BROWN, M., FITZGERALD, K. J. & SABATINI, D. M. 2006. Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKCalpha, but not S6K1. *Dev Cell*, 11, 859-71.
- GULHATI, P., BOWEN, K. A., LIU, J., STEVENS, P. D., RYCHAHOU, P. G., CHEN, M., LEE, E. Y., WEISS, H. L., O'CONNOR, K. L., GAO, T. & EVERS, B. M. 2011. mTORC1 and mTORC2 regulate EMT, motility, and metastasis of colorectal cancer via RhoA and Rac1 signaling pathways. *Cancer Res*, 71, 3246-56.
- GUO, H., INGOLIA, N. T., WEISSMAN, J. S. & BARTEL, D. P. 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, 466, 835-40.
- GWINN, D. M., SHACKELFORD, D. B., EGAN, D. F., MIHAYLOVA, M. M., MERY, A., VASQUEZ, D. S., TURK, B. E. & SHAW, R. J. 2008. AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol Cell*, 30, 214-26.
- HANSEN, K. D., BRENNER, S. E. & DUDOIT, S. 2010. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res*, 38, e131.
- HARDING, H. P., ZHANG, Y., ZENG, H., NOVOA, I., LU, P. D., CALFON, M., SADRI, N., YUN, C., POPKO, B., PAULES, R., STOJDL, D. F., BELL, J. C., HETTMANN, T., LEIDEN, J. M. & RON, D. 2003. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell*, 11, 619-33.
- HATFIELD, D. & RICE, M. 1986. Aminoacyl-tRNA(anticodon): codon adaptation in human and rabbit reticulocytes. *Biochem Int*, 13, 835-42.
- HEITMAN, J., MOVVA, N. R. & HALL, M. N. 1991. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science*, 253, 905-9.
- HO, C. K. & SHUMAN, S. 2002. Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains (vol 99, pg 12709, 2002). *Proceedings of the National Academy of Sciences of the United States of America*, 99, 13961-13961.
- HOFMANN, S., CHERKASOVA, V., BANKHEAD, P., BUKAU, B. & STOECKLIN, G. 2012. Translation suppression promotes stress granule formation and cell survival in response to cold shock. *Mol Biol Cell*, 23, 3786-800.
- HSIEH, A. C., LIU, Y., EDLIND, M. P., INGOLIA, N. T., JANES, M. R., SHER, A., SHI, E. Y., STUMPF, C. R., CHRISTENSEN, C., BONHAM, M. J., WANG, S., REN, P., MARTIN, M., JESSEN, K., FELDMAN, M. E., WEISSMAN, J. S., SHOKAT, K. M., ROMMEL, C. & RUGGERO, D. 2012. The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature*, 485, 55-61.
- HUANG, J., DIBBLE, C. C., MATSUZAKI, M. & MANNING, B. D. 2008. The TSC1-TSC2 complex is required for proper activation of mTOR complex 2. *Mol Cell Biol*, 28, 4104-15.
- IACONO, M., MIGNONE, F. & PESOLE, G. 2005. uAUG and uORFs in human and rodent 5'untranslated mRNAs. *Gene*, 349, 97-105.
- IKEMURA, T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol*, 2, 13-34.
- ILYAS, M. 2005. Wnt signalling and the mechanistic basis of tumour development. *J Pathol*, 205, 130-44.
- IMATAKA, H., GRADI, A. & SONENBERG, N. 1998. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J*, 17, 7480-9.

- IMATAKA, H. & SONENBERG, N. 1997. Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol Cell Biol*, 17, 6940-7.
- INGOLIA, N. T., BRAR, G. A., ROUSKIN, S., MCGEACHY, A. M. & WEISSMAN, J. S. 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc*, 7, 1534-50.
- INGOLIA, N. T., GHAEMMAGHAMI, S., NEWMAN, J. R. & WEISSMAN, J. S. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, 324, 218-23.
- INGOLIA, N. T., LAREAU, L. F. & WEISSMAN, J. S. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, 147, 789-802.
- INOKI, K., LI, Y., XU, T. & GUAN, K. L. 2003a. Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes Dev*, 17, 1829-34.
- INOKI, K., ZHU, T. & GUAN, K. L. 2003b. TSC2 mediates cellular energy response to control cell growth and survival. *Cell*, 115, 577-90.
- IVANKOVIC, M., RUBELJ, I., MATULIC, M., REICH, E. & BRDAR, B. 2006. Site-specific mutagenesis of the histidine precursor of diphthamide in the human elongation factor-2 gene confers resistance to diphtheria toxin. *Mutat Res*, 609, 34-42.
- JACKSON, E. L., WILLIS, N., MERCER, K., BRONSON, R. T., CROWLEY, D., MONTOYA, R., JACKS, T. & TUVESON, D. A. 2001. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev*, 15, 3243-8.
- JANIKIEWICZ, J. 2011. *Tissue-specific variants of translation elongation factor eEF1A and their role in cancer*. PhD, University of Edinburgh.
- JANSEN, A. P., CAMALIER, C. E., STARK, C. & COLBURN, N. H. 2004. Characterization of programmed cell death 4 in multiple human cancers reveals a novel enhancer of drug sensitivity. *Mol Cancer Ther*, 3, 103-10.
- JARVINEN, H. J. 1992. Epidemiology of familial adenomatous polyposis in Finland: impact of family screening on the colorectal cancer rate and survival. *Gut*, 33, 357-60.
- JAYAPRAKASH, A. D., JABADO, O., BROWN, B. D. & SACHIDANANDAM, R. 2011. Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res*, 39, e141.
- JEFFERIES, H. B., FUMAGALLI, S., DENNIS, P. B., REINHARD, C., PEARSON, R. B. & THOMAS, G. 1997. Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k. *EMBO J*, 16, 3693-704.
- JEFFERIES, H. B., REINHARD, C., KOZMA, S. C. & THOMAS, G. 1994. Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. *Proc Natl Acad Sci U S A*, 91, 4441-5.
- JENNINGS, M. D. & PAVITT, G. D. 2010. eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. *Nature*, 465, 378-81.
- KANG, Y. H., LEE, H. S. & KIM, W. H. 2002. Promoter methylation and silencing of PTEN in gastric carcinoma. *Lab Invest*, 82, 285-91.
- KANTIDAKIS, T., RAMSBOTTOM, B. A., BIRCH, J. L., DOWDING, S. N. & WHITE, R. J. 2010. mTOR associates with TFIIC, is found at tRNA and 5S rRNA genes, and targets their repressor Maf1. *Proc Natl Acad Sci U S A*, 107, 11823-8.
- KHALILI, K. & WEINMANN, R. 1984. Actin mRNAs in HeLa cells. Stabilization after adenovirus infection. *J Mol Biol*, 180, 1007-21.
- KIEFT, J. S. 2009. Comparing the three-dimensional structures of Dicistroviridae IGR IRES RNAs with other viral RNA structures. *Virus Res*, 139, 148-56.
- KIM, C. H., OH, Y. & LEE, T. H. 1997. Codon optimization for high-level expression of human erythropoietin (EPO) in mammalian cells. *Gene*, 199, 293-301.

- KOMAR, A. A. & HATZOGLOU, M. 2011. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. *Cell Cycle*, 10, 229-40.
- KONG, Y. W., FERLAND-MCCOLLOUGH, D., JACKSON, T. J. & BUSHELL, M. 2012. microRNAs in cancer management. *Lancet Oncol*, 13, e249-58.
- KORNEEVA, N. L., LAMPHEAR, B. J., HENNIGAN, F. L. & RHOADS, R. E. 2000. Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J Biol Chem*, 275, 41369-76.
- KOROMILAS, A. E., LAZARIS-KARATZAS, A. & SONENBERG, N. 1992. mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J*, 11, 4153-8.
- KOZAK, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*, 15, 8125-48.
- KOZAK, M. 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem*, 266, 19867-70.
- KOZAK, M. 2001. Constraints on reinitiation of translation in mammals. *Nucleic Acids Res*, 29, 5226-32.
- KOZAK, M. 2008. Faulty old ideas about translational regulation paved the way for current confusion about how microRNAs function. *Gene*, 423, 108-15.
- LAMPHEAR, B. J., KIRCHWEGER, R., SKERN, T. & RHOADS, R. E. 1995. Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J Biol Chem*, 270, 21975-83.
- LANG, F., BOHMER, C., PALMADA, M., SEEBOHM, G., STRUTZ-SEEBOHM, N. & VALLON, V. 2006. (Patho)physiological significance of the serum- and glucocorticoid-inducible kinase isoforms. *Physiol Rev*, 86, 1151-78.
- LANKAT-BUTTGEREIT, B. & GOKE, R. 2009. The tumour suppressor Pcd4: recent advances in the elucidation of function and regulation. *Biol Cell*, 101, 309-17.
- LAPLANTE, M. & SABATINI, D. M. 2012. mTOR signaling in growth control and disease. *Cell*, 149, 274-93.
- LAU, E. T., KONG, R. Y. & CHEAH, K. S. 1993. A critical assessment of the RNase protection assay as a means of determining exon sizes. *Anal Biochem*, 209, 360-6.
- LE BACQUER, O., PETROULAKIS, E., PAGLIALUNGA, S., POULIN, F., RICHARD, D., CIANFLONE, K. & SONENBERG, N. 2007. Elevated sensitivity to diet-induced obesity and insulin resistance in mice lacking 4E-BP1 and 4E-BP2. *J Clin Invest*, 117, 387-96.
- LEMP, N. A., HIRAOKA, K., KASAHARA, N. & LOGG, C. R. 2012. Cryptic transcripts from a ubiquitous plasmid origin of replication confound tests for cis-regulatory function. *Nucleic Acids Res*, 40, 7280-90.
- LI, Q., IMATAKA, H., MORINO, S., ROGERS, G. W., JR., RICHTER-COOK, N. J., MERRICK, W. C. & SONENBERG, N. 1999. Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol Cell Biol*, 19, 7336-46.
- LIU, B., HAN, Y. & QIAN, S. B. 2013. Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. *Mol Cell*, 49, 453-63.
- LIU, S., BACHRAN, C., GUPTA, P., MILLER-RANDOLPH, S., WANG, H., CROWN, D., ZHANG, Y., WEIN, A. N., SINGH, R., FATTAH, R. & LEPLA, S. H. 2012. Diphthamide modification on eukaryotic elongation factor 2 is needed to assure fidelity of mRNA translation and mouse development. *Proc Natl Acad Sci U S A*, 109, 13817-22.
- LOEWITH, R., JACINTO, E., WULLSCHLEGER, S., LORBERG, A., CRESPO, J. L., BONENFANT, D., OPPLIGER, W., JENOE, P. & HALL, M. N. 2002. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Mol Cell*, 10, 457-68.

- LOMAKIN, I. B., KOLUPAEVA, V. G., MARINTCHEV, A., WAGNER, G. & PESTOVA, T. V. 2003. Position of eukaryotic initiation factor eIF1 on the 40S ribosomal subunit determined by directed hydroxyl radical probing. *Genes Dev*, 17, 2786-97.
- LUCHTENBORG, M., WEIJENBERG, M. P., ROEMEN, G. M., DE BRUINE, A. P., VAN DEN BRANDT, P. A., LENTJES, M. H., BRINK, M., VAN ENGELAND, M., GOLDBOHM, R. A. & DE GOEIJ, A. F. 2004. APC mutations in sporadic colorectal carcinomas from The Netherlands Cohort Study. *Carcinogenesis*, 25, 1219-26.
- LUUKKONEN, B. G., TAN, W. & SCHWARTZ, S. 1995. Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol*, 69, 4086-94.
- LYNCH, H. T., LYNCH, P. M., LANSPA, S. J., SNYDER, C. L., LYNCH, J. F. & BOLAND, C. R. 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet*, 76, 1-18.
- MANNING, B. D. 2004. Balancing Akt with S6K: implications for both metabolic diseases and tumorigenesis. *J Cell Biol*, 167, 399-403.
- MAQUAT, L. E., TARN, W. Y. & ISKEN, O. 2010. The pioneer round of translation: features and functions. *Cell*, 142, 368-74.
- MARKMAN, B., DIENSTMANN, R. & TABERNERO, J. 2010. Targeting the PI3K/Akt/mTOR pathway--beyond rapalogs. *Oncotarget*, 1, 530-43.
- MAROTTA, A., PARHAR, K., OWEN, D., DEDHAR, S. & SALH, B. 2003. Characterisation of integrin-linked kinase signalling in sporadic human colon cancer. *Br J Cancer*, 88, 1755-62.
- MAROTTA, A., TAN, C., GRAY, V., MALIK, S., GALLINGER, S., SANGHERA, J., DUPUIS, B., OWEN, D., DEDHAR, S. & SALH, B. 2001. Dysregulation of integrin-linked kinase (ILK) signaling in colonic polyposis. *Oncogene*, 20, 6250-7.
- MARSHALL, L., KENNETH, N. S. & WHITE, R. J. 2008. Elevated tRNA(iMet) synthesis can drive cell proliferation and oncogenic transformation. *Cell*, 133, 78-89.
- MARSHMAN, E., BOOTH, C. & POTTEN, C. S. 2002. The intestinal epithelial stem cell. *Bioessays*, 24, 91-8.
- MATHEWS, M. B., SONENBERG, N. & HERSHEY, J. W. B. 2007. Origins and Principles of Translational Control. In: MATHEWS, M. B., SONENBERG, N. & HERSHEY, J. W. B. (eds.) *Translational Control in Biology and Medicine* New York: Cold Spring Harbour Laboratory Press.
- MCDONALD, P. C., OLOUMI, A., MILLS, J., DOBREVA, I., MAIDAN, M., GRAY, V., WEDERELL, E. D., BALLY, M. B., FOSTER, L. J. & DEDHAR, S. 2008. Rictor and integrin-linked kinase interact and regulate Akt phosphorylation and cancer cell survival. *Cancer Res*, 68, 1618-24.
- MELNIKOV, S., BEN-SHEM, A., GARREAU DE LOUBRESSE, N., JENNER, L., YUSUPOVA, G. & YUSUPOV, M. 2012. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol*, 19, 560-7.
- MERG, A., LYNCH, H. T., LYNCH, J. F. & HOWE, J. R. 2005. Hereditary colon cancer--part I. *Curr Probl Surg*, 42, 195-256.
- NAKAMURA, J., AOYAGI, S., NANCHI, I., NAKATSUKA, S., HIRATA, E., SHIBATA, S., FUKUDA, M., YAMAMOTO, Y., FUKUDA, I., TATSUMI, N., UEDA, T., FUJIKI, F., NOMURA, M., NISHIDA, S., SHIRAKATA, T., HOSEN, N., TSUBOI, A., OKA, Y., NEZU, R., MORI, M., DOKI, Y., AOZASA, K., SUGIYAMA, H. & OJI, Y. 2009. Overexpression of eukaryotic elongation factor eEF2 in gastrointestinal cancers and its involvement in G2/M progression in the cell cycle. *Int J Oncol*, 34, 1181-9.
- NIEPEL, M., LING, J. & GALLIE, D. R. 1999. Secondary structure in the 5'-leader or 3'-untranslated region reduces protein yield but does not affect the functional interaction between the 5'-cap and the poly(A) tail. *FEBS Lett*, 462, 79-84.

- NOVOA, I., ZENG, H., HARDING, H. P. & RON, D. 2001. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2alpha. *Journal of Cell Biology*, 153, 1011-22.
- OGLE, J. M., CARTER, A. P. & RAMAKRISHNAN, V. 2003. Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem Sci*, 28, 259-66.
- OH, W. J. & JACINTO, E. 2011. mTOR complex 2 signaling and functions. *Cell Cycle*, 10, 2305-2316.
- OH, W. J., WU, C. C., KIM, S. J., FACCHINETTI, V., JULIEN, L. A., FINLAN, M., ROUX, P. P., SU, B. & JACINTO, E. 2010. mTORC2 can associate with ribosomes to promote cotranslational phosphorylation and stability of nascent Akt polypeptide. *Embo Journal*, 29, 3939-51.
- ONS. 2009. *Cancer incidence and mortality* [Online]. Available: <http://www.statistics.gov.uk/cci/nugget.asp?id=915> [Accessed].
- ORTIZ, P. A., ULLOQUE, R., KIHARA, G. K., ZHENG, H. & KINZY, T. G. 2006. Translation elongation factor 2 anticodon mimicry domain mutants affect fidelity and diphtheria toxin resistance. *J Biol Chem*, 281, 32639-48.
- OYOLA, S. O., OTTO, T. D., GU, Y., MASLEN, G., MANSKE, M., CAMPINO, S., TURNER, D. J., MACINNIS, B., KWIATKOWSKI, D. P., SWERDLOW, H. P. & QUAIL, M. A. 2012. Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes. *BMC Genomics*, 13, 1.
- OZES, A. R., FEOKTISTOVA, K., AVANZINO, B. C. & FRASER, C. S. 2011. Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J Mol Biol*, 412, 674-87.
- PAPE, T., WINTERMEYER, W. & RODNINA, M. V. 1998. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. *EMBO J*, 17, 7490-7.
- PARK, M. H. 2006. The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). *J Biochem*, 139, 161-9.
- PASSMORE, L. A., SCHMEING, T. M., MAAG, D., APPLEFIELD, D. J., ACKER, M. G., ALGIRE, M. A., LORSCH, J. R. & RAMAKRISHNAN, V. 2007. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell*, 26, 41-50.
- PATAER, A., SWISHER, S. G., ROTH, J. A., LOGOTHETIS, C. J. & CORN, P. G. 2009. Inhibition of RNA-dependent protein kinase (PKR) leads to cancer cell death and increases chemosensitivity. *Cancer Biol Ther*, 8, 245-52.
- PATURSKY-POLISCHUK, I., STOLOVICH-RAIN, M., HAUSNER-HANOCHI, M., KASIR, J., CYBULSKI, N., AVRUCH, J., RUEGG, M. A., HALL, M. N. & MEYUHAS, O. 2009. The TSC-mTOR pathway mediates translational activation of TOP mRNAs by insulin largely in a raptor- or rictor-independent manner. *Mol Cell Biol*, 29, 640-9.
- PAVON-ETERNOD, M., GOMES, S., GESLAIN, R., DAI, Q., ROSNER, M. R. & PAN, T. 2009. tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res*, 37, 7268-80.
- PAYNE, G. C., PAYNE, R. E. & FAREWELL, D. M. 2008. Rugby (the religion of Wales) and its influence on the Catholic church: should Pope Benedict XVI be worried? *BMJ*, 337, a2768.
- PENDE, M., UM, S. H., MIEULET, V., STICKER, M., GOSS, V. L., MESTAN, J., MUELLER, M., FUMAGALLI, S., KOZMA, S. C. & THOMAS, G. 2004. S6K1(-/-)/S6K2(-/-) mice exhibit perinatal lethality and rapamycin-sensitive 5'-terminal oligopyrimidine mRNA translation and reveal a mitogen-activated protein kinase-dependent S6 kinase pathway. *Mol Cell Biol*, 24, 3112-24.

- PESTOVA, T. V. & HELLEN, C. U. 2003. Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev*, 17, 181-6.
- PESTOVA, T. V. & KOLUPAEVA, V. G. 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev*, 16, 2906-22.
- PISAREV, A. V., SKABKIN, M. A., PISAREVA, V. P., SKABKINA, O. V., RAKOTONDRAFARA, A. M., HENTZE, M. W., HELLEN, C. U. & PESTOVA, T. V. 2010. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell*, 37, 196-210.
- PISAREVA, V. P., PISAREV, A. V., HELLEN, C. U., RODNINA, M. V. & PESTOVA, T. V. 2006. Kinetic analysis of interaction of eukaryotic release factor 3 with guanine nucleotides. *J Biol Chem*, 281, 40224-35.
- PISAREVA, V. P., PISAREV, A. V., KOMAR, A. A., HELLEN, C. U. & PESTOVA, T. V. 2008. Translation initiation on mammalian mRNAs with structured 5'UTRs requires DEXH-box protein DHX29. *Cell*, 135, 1237-50.
- PISAREVA, V. P., SKABKIN, M. A., HELLEN, C. U., PESTOVA, T. V. & PISAREV, A. V. 2011. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J*, 30, 1804-17.
- POLAK, P., CYBULSKI, N., FEIGE, J. N., AUWERX, J., RUEGG, M. A. & HALL, M. N. 2008. Adipose-specific knockout of raptor results in lean mice with enhanced mitochondrial respiration. *Cell Metab*, 8, 399-410.
- POPULO, H., LOPES, J. M. & SOARES, P. 2012. The mTOR Signalling Pathway in Human Cancer. *Int J Mol Sci*, 13, 1886-918.
- POWLEY, I. R., KONDRASHOV, A., YOUNG, L. A., DOBBYN, H. C., HILL, K., CANNELL, I. G., STONELEY, M., KONG, Y. W., COTES, J. A., SMITH, G. C., WEK, R., HAYES, C., GANT, T. W., SPRIGGS, K. A., BUSHELL, M. & WILLIS, A. E. 2009. Translational reprogramming following UVB irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes. *Genes Dev*, 23, 1207-20.
- PRENEN, H., TEJPAR, S. & VAN CUTSEM, E. 2010. New strategies for treatment of KRAS mutant metastatic colorectal cancer. *Clin Cancer Res*, 16, 2921-6.
- PROUD, C. G. 2005. eIF2 and the control of cell physiology. *Semin Cell Dev Biol*, 16, 3-12.
- PULLEN, N., DENNIS, P. B., ANDJELKOVIC, M., DUFNER, A., KOZMA, S. C., HEMMING, B. A. & THOMAS, G. 1998. Phosphorylation and activation of p70s6k by PDK1. *Science*, 279, 707-10.
- PYR DIT RUYS, S., WANG, X., SMITH, E. M., HERINCKX, G., HUSSAIN, N., RIDER, M. H., VERTOMMEN, D. & PROUD, C. G. 2012. Identification of autophosphorylation sites in eukaryotic elongation factor-2 kinase. *Biochem J*, 442, 681-92.
- QIAN, W., YANG, J. R., PEARSON, N. M., MACLEAN, C. & ZHANG, J. 2012. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet*, 8, e1002603.
- QUAIL, M. A., SMITH, M., COUPLAND, P., OTTO, T. D., HARRIS, S. R., CONNOR, T. R., BERTONI, A., SWERDLOW, H. P. & GU, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, 13, 341.
- RAJKOWITSCH, L., VILELA, C., BERTHELOT, K., RAMIREZ, C. V. & MCCARTHY, J. E. 2004. Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *J Mol Biol*, 335, 71-85.
- READER, R. W. & STANNERS, C. P. 1967. On the significance of ribosome dimers in extracts of animal cells. *J Mol Biol*, 28, 211-23.
- REDPATH, N. T. & PROUD, C. G. 1993. Purification and phosphorylation of elongation factor-2 kinase from rabbit reticulocytes. *Eur J Biochem*, 212, 511-20.

- ROBINSON, F., JACKSON, R. J. & SMITH, C. W. 2008. Expression of human nPTB is limited by extreme suboptimal codon content. *PLoS One*, 3, e1801.
- ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-40.
- ROMANI, A. 2007. Regulation of magnesium homeostasis and transport in mammalian cells. *Arch Biochem Biophys*, 458, 90-102.
- ROSENWALD, I. B., WANG, S., SAVAS, L., WODA, B. & PULLMAN, J. 2003. Expression of translation initiation factor eIF-2 $\alpha$  is increased in benign and malignant melanocytic and colonic epithelial neoplasms. *Cancer*, 98, 1080-8.
- RUGGERO, D. 2012. Translational Control in Cancer Etiology (vol 21, pg 474, 2011). *Cold Spring Harbor Perspectives in Biology*, 4.
- RUGGERO, D., MONTANARO, L., MA, L., XU, W., LONDEI, P., CORDON-CARDO, C. & PANDOLFI, P. P. 2004. The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nat Med*, 10, 484-6.
- RUVINSKY, I., SHARON, N., LERER, T., COHEN, H., STOLOVICH-RAIN, M., NIR, T., DOR, Y., ZISMAN, P. & MEYUHAS, O. 2005. Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes Dev*, 19, 2199-211.
- SAINI, P., EYLER, D. E., GREEN, R. & DEVER, T. E. 2009. Hypusine-containing protein eIF5A promotes translation elongation. *Nature*, 459, 118-21.
- SAMBROOK, J. F., F.E.
- MANIATIS, T. 1989. Hybridisation of Radiolabelled probes to immobilised nucleic acids *Molecular Cloning*. 2 ed.
- SARBASSOV, D. D., ALI, S. M., SENGUPTA, S., SHEEN, J. H., HSU, P. P., BAGLEY, A. F., MARKHARD, A. L. & SABATINI, D. M. 2006. Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. *Mol Cell*, 22, 159-68.
- SARBASSOV, D. D., GUERTIN, D. A., ALI, S. M. & SABATINI, D. M. 2005. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science*, 307, 1098-101.
- SATO, T., VRIES, R. G., SNIPPERT, H. J., VAN DE WETERING, M., BARKER, N., STANGE, D. E., VAN ES, J. H., ABO, A., KUJALA, P., PETERS, P. J. & CLEVERS, H. 2009. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 459, 262-5.
- SCHEPERS, A. G., SNIPPERT, H. J., STANGE, D. E., VAN DEN BORN, M., VAN ES, J. H., VAN DE WETERING, M. & CLEVERS, H. 2012. Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. *Science*, 337, 730-5.
- SCHNEIDER-POETSCH, T., JU, J., EYLER, D. E., DANG, Y., BHAT, S., MERRICK, W. C., GREEN, R., SHEN, B. & LIU, J. O. 2010. Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol*, 6, 209-217.
- SCHWANHAUSSER, B., BUSSE, D., LI, N., DITTMAR, G., SCHUCHHARDT, J., WOLF, J., CHEN, W. & SELBACH, M. 2011. Global quantification of mammalian gene expression control. *Nature*, 473, 337-42.
- SCOTT, P. A., SMITH, K., POULSOM, R., DE BENEDETTI, A., BICKNELL, R. & HARRIS, A. L. 1998. Differential expression of vascular endothelial growth factor mRNA vs protein isoform expression in human breast cancer and relationship to eIF-4E. *Br J Cancer*, 77, 2120-8.
- SHATKIN, A. J. 1976. Capping of eucaryotic mRNAs. *Cell*, 9, 645-53.
- SHIBATA, H., TOYAMA, K., SHIOYA, H., ITO, M., HIROTA, M., HASEGAWA, S., MATSUMOTO, H., TAKANO, H., AKIYAMA, T., TOYOSHIMA, K., KANAMARU, R., KANEGAE, Y., SAITO, I., NAKAMURA, Y., SHIBA, K. & NODA, T. 1997. Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science*, 278, 120-3.

- SHIMIZU, T., TOLCHER, A. W., PAPADOPOULOS, K. P., BEERAM, M., RASCO, D. W., SMITH, L. S., GUNN, S., SMETZER, L., MAYS, T. A., KAISER, B., WICK, M. J., ALVAREZ, C., CAVAZOS, A., MANGOLD, G. L. & PATNAIK, A. 2012. The clinical effect of the dual-targeting strategy involving PI3K/AKT/mTOR and RAS/MEK/ERK pathways in patients with advanced cancer. *Clin Cancer Res*, 18, 2316-25.
- SHOEMAKER, C. J., EYLER, D. E. & GREEN, R. 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science*, 330, 369-72.
- SILVERA, D., FORMENTI, S. C. & SCHNEIDER, R. J. 2010. Translational control in cancer. *Nature Reviews Cancer*, 10, 254-266.
- SMYTH, G. K. & SPEED, T. 2003. Normalization of cDNA microarray data. *Methods*, 31, 265-73.
- SOKABE, M., FRASER, C. S. & HERSHEY, J. W. 2012. The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. *Nucleic Acids Res*, 40, 905-13.
- SOREFAN, K., PAIS, H., HALL, A. E., KOZOMARA, A., GRIFFITHS-JONES, S., MOULTON, V. & DALMAY, T. 2012. Reducing sequencing bias of small RNAs. *Silence*, 3, 4.
- SPAHN, C. M., GOMEZ-LORENZO, M. G., GRASSUCCI, R. A., JORGENSEN, R., ANDERSEN, G. R., BECKMANN, R., PENCZEK, P. A., BALLESTA, J. P. & FRANK, J. 2004. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. *EMBO J*, 23, 1008-19.
- SPAHR, P. F. & HOLLINGWORTH, B. R. 1961. Purification and Mechanism of Action of Ribonuclease from Escherichia Coli Ribosomes. *Journal of Biological Chemistry*, 236, 823-&.
- STADLER, M. & FIRE, A. 2011. Wobble base-pairing slows in vivo translation elongation in metazoans. *RNA*, 17, 2063-73.
- STARCK, S. R., JIANG, V., PAVON-ETERNOD, M., PRASAD, S., MCCARTHY, B., PAN, T. & SHASTRI, N. 2012. Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science*, 336, 1719-23.
- TAYLOR, D. J., FRANK, J. & KINZY, T. G. 2007a. Structure and Function of the Eukaryotic Ribosome and Elongation Factors. In: MATHEWS, M. B., SONENBERG, N. & HERSHEY, J. W. (eds.) *Translational Control in Biology and Medicine*. New York: Cold Spring Harbour Laboratory Press.
- TAYLOR, D. J., NILSSON, J., MERRILL, A. R., ANDERSEN, G. R., NISSEN, P. & FRANK, J. 2007b. Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. *EMBO J*, 26, 2421-31.
- THOMAS, R. K., BAKER, A. C., DEBIASI, R. M., WINCKLER, W., LAFRAMBOISE, T., LIN, W. M., WANG, M., FENG, W., ZANDER, T., MACCONAILL, L., LEE, J. C., NICOLETTI, R., HATTON, C., GOYETTE, M., GIRARD, L., MAJMUDAR, K., ZIAUGRA, L., WONG, K. K., GABRIEL, S., BEROUKHIM, R., PEYTON, M., BARRETINA, J., DUTT, A., EMERY, C., GREULICH, H., SHAH, K., SASAKI, H., GAZDAR, A., MINNA, J., ARMSTRONG, S. A., MELLINGHOFF, I. K., HODI, F. S., DRANOFF, G., MISCHER, P. S., CLOUGHESY, T. F., NELSON, S. F., LIAU, L. M., MERTZ, K., RUBIN, M. A., MOCH, H., LODA, M., CATALONA, W., FLETCHER, J., SIGNORETTI, S., KAYE, F., ANDERSON, K. C., DEMETRI, G. D., DUMMER, R., WAGNER, S., HERLYN, M., SELLERS, W. R., MEYERSON, M. & GARRAWAY, L. A. 2007. High-throughput oncogene mutation profiling in human cancer. *Nat Genet*, 39, 347-51.
- THOREEN, C. C., CHANTRANUPONG, L., KEYS, H. R., WANG, T., GRAY, N. S. & SABATINI, D. M. 2012. A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature*, 485, 109-13.
- THOREEN, C. C., KANG, S. A., CHANG, J. W., LIU, Q., ZHANG, J., GAO, Y., REICHLING, L. J., SIM, T., SABATINI, D. M. & GRAY, N. S. 2009. An ATP-competitive mammalian target of

- rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J Biol Chem*, 284, 8023-32.
- TIAN, G., YIN, X., LUO, H., XU, X., BOLUND, L., ZHANG, X., GAN, S. Q. & LI, N. 2010. Sequencing bias: comparison of different protocols of microRNA library construction. *BMC Biotechnol*, 10, 64.
- TOEDLING, J., SERVANT, N., CIAUDO, C., FARINELLI, L., VOINNET, O., HEARD, E. & BARILLOT, E. 2012. Deep-sequencing protocols influence the results obtained in small-RNA sequencing. *PLoS One*, 7, e32724.
- TOMLINSON, V. A., NEWBERY, H. J., BERGMANN, J. H., BOYD, J., SCOTT, D., WRAY, N. R., SELLAR, G. C., GABRA, H., GRAHAM, A., WILLIAMS, A. R. & ABBOTT, C. M. 2007. Expression of eEF1A2 is associated with clear cell histology in ovarian carcinomas: overexpression of the gene is not dependent on modifications at the EEF1A2 locus. *Br J Cancer*, 96, 1613-20.
- UDE, S., LASSAK, J., STAROSTA, A. L., KRAXENBERGER, T., WILSON, D. N. & JUNG, K. 2013. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science*, 339, 82-5.
- VANDER HAAR, E., LEE, S. I., BANDHAKAVI, S., GRIFFIN, T. J. & KIM, D. H. 2007. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. *Nat Cell Biol*, 9, 316-23.
- VATTEM, K. M. & WEK, R. C. 2004. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A*, 101, 11269-74.
- VIOLETT, S., FUCHS, R. T., MUNAFO, D. B., ZHUANG, F. & ROBB, G. B. 2011. T4 RNA ligase 2 truncated active site mutants: improved tools for RNA analysis. *BMC Biotechnol*, 11, 72.
- WANG, X., LI, W., WILLIAMS, M., TERADA, N., ALESSI, D. R. & PROUD, C. G. 2001. Regulation of elongation factor 2 kinase by p90(RSK1) and p70 S6 kinase. *EMBO J*, 20, 4370-9.
- WHITE, R. J. 2004. RNA polymerase III transcription and cancer. *Oncogene*, 23, 3208-16.
- WILLIAMS, D. D., PRICE, N. T., LOUGHLIN, A. J. & PROUD, C. G. 2001. Characterization of the mammalian initiation factor eIF2B complex as a GDP dissociation stimulator protein. *J Biol Chem*, 276, 24697-703.
- WILLIAMS, R. J. & WACKER, E. C. 1967. Cation balance in biological systems. *JAMA*, 201, 96-100.
- WILLIS, A. E. 1999. Translational control of growth factor and proto-oncogene expression. *Int J Biochem Cell Biol*, 31, 73-86.
- WOLIN, S. L. & WALTER, P. 1988. Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J*, 7, 3559-69.
- ZETTERBERG, A. & KILLANDER, D. 1965. Quantitative cytophotometric and autoradiographic studies on the rate of protein synthesis during interphase in mouse fibroblasts in vitro. *Exp Cell Res*, 40, 1-11.
- ZHANG, Y. J., DAI, Q., SUN, D. F., XIONG, H., TIAN, X. Q., GAO, F. H., XU, M. H., CHEN, G. Q., HAN, Z. G. & FANG, J. Y. 2009. mTOR signaling pathway is a target for the treatment of colorectal cancer. *Ann Surg Oncol*, 16, 2617-28.
- ZHELKOVSKY, A. M. & MCREYNOLDS, L. A. 2012. Structure-function analysis of *Methanobacterium thermoautotrophicum* RNA ligase - engineering a thermostable ATP independent enzyme. *BMC Mol Biol*, 13, 24.
- ZHUANG, F., FUCHS, R. T., SUN, Z., ZHENG, Y. & ROBB, G. B. 2012. Structural bias in T4 RNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Res*, 40, e54.
- ZONCU, R., BAR-PELED, L., EFEYAN, A., WANG, S., SANCAK, Y. & SABATINI, D. M. 2011. mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+)-ATPase. *Science*, 334, 678-83.