Human orthologues of *Saccharomyces cerevisiae* Trf4/5 non-canonical poly(A) polymerases

Kumudika Gibson

University of Oxford

Sir William Dunn School of Pathology

The Queen's College

Thesis submitted for the degree of Master of Science

at the University of Oxford

Trinity Term 2011

To

Nigel, Alexander, Sebastian and my parents

Role of the Author

The research presented in this thesis has not been submitted for any other degree. The experimental work presented was performed wholly by me, except where indicated to the contrary.

Kumudika Gibson
The Queen's College

A thesis submitted for the degree of Master of Science
Trinity Term 2011

Human orthologues of *Saccharomyces cerevisiae* Trf4/5 non-canonical poly(A) polymerases

RNA waste in the budding yeast nucleus is degraded 3'-5' *via* exosome-mediated pathways associated with the non-canonical poly(A) polymerases Trf4 and Trf5. Multiple sequence alignments identify POLS and PAPD5 as the human orthologues of Trf4 and Trf5. This research has investigated the degree of conservation and extent of divergence of POLS and PAPD5 from their orthologues. Mass spectrometric analysis identified interacting protein partners of POLS that are mostly nuclear and are associated with an intricate network of cellular processes, spanning/encompassing RNA metabolism including pre-mRNA splicing, DNA damage responses, chromosome biology, nuclear pore transport and structure. The human orthologue of the Trf4/5-associated helicase Mtr4 was identified as a POLS interactor along with DEAD box RNA helicases and zinc finger proteins. Surprisingly in the light of the identities of these interacting proteins, POLS was found to localize mainly to the cytoplasm with a small proportion visible in the nucleus. Co-immunoprecipitation experiments further indicated POLS interactions with CPSF-73 and BRCA1. Human pre-rRNA transcripts that were polyadenylated within the 3'ETS sequence were identified. Non-A residues incorporated within the A-rich tails of the human 3'ETS rRNA and β-actin mRNA were found to be PCR-induced artefacts. Individual down-regulation of POLS or PAPD5 resulted in a slow growth phenotype. A lethal phenotype was detected following the simultaneous down-regulation of both POLS and PAPD5 (DKD). Low concentrations of the anti cancer drug 5-Fluorouracil (5-FU) increased the accumulation of POLS in the nucleolus and 5-FU concentrations >25 μM caused POLS to move from the cytoplasm to the nucleus. DKD increased the sensitivity of U2-OS cells to 5-FU. I conclude that, although some aspects of the functions of Trf4 and Trf5 have been conserved, in human cells POLS has diverged considerably from its yeast counterparts, opening up many further lines of enquiry.

Table of contents

Abstract i

Contents ii

List of Figures iv

List of Tables vi

Glossary of Abbreviations vii

1 Introduction 1

1.1 RNA surveillance 1

1.1.1 RNA decay 1

1.1.2 Exosome 2

1.1.3 3’ and 5’ end modifications 11

1.2 TRF4 and TRF5 17

1.2.1 TRAMP complex 19

1.2.2 Partial structure of the TRAMP complex 23

1.2.3 Trf4/5 targets and mutant phenotypes 25

1.2.4 Ribosomal RNA 30

1.2.5 Orthologues of Trf4 and Trf5 39

1.3 A Comparison of canonical and non-canonical poly(A) polymerases 41

1.3.1 The family of non-canonical poly(A) polymerases 43

1.3.2 The human non-canonical poly(A) polymerases 46

1.3.3 Sequence comparison of POLS and PAPD5 50

1.3.4 Sequence comparison of POLS, PAPD5, Trf4 and Trf5 51

1.4 Aims of this study 54

2 Materials and Methods 55

2.1 Cell culture methods 55

2.2 RNA methods 59

2.3 DNA methods 64

2.4 Protein methods 74

3 The cellular functions of POLS and PAPD5 **85**

3.1 Introduction 85

3.2 Aims 88

3.3 POLS localization 88

3.4 Individual knockdown of *POLS* or *PAPD5* 96

3.5 Down-regulation of *POLS* and/or *PAPD5* results in a slow growth/lethal phenotype 98

3.6 Discussion 103

4 Biochemical functions of POLS and PAPD5 **106**

4.1 Introduction 106

4.2 Aims 110

4.3 Is/Are POLS or/and PAPD5 involved in the poly/oligo- adenylation of human rRNA tails? 110

4.4 Is the activity of 5-FU linked to POLS and/or PAPD5? 126

4.5 Discussion 130

5 Interacting protein partners of POLS **133**

5.1 Introduction 133

5.2 Aim 134

5.3 Identification of interacting protein partners of POLS by tandem affinity purification 134

5.4 Is there a POLS-associated human TRAMP-like complex? 138

5.5 Protein interactions with POLS 138

5.6 Discussion 148

6 Novel splice variants identified in an expression screen of human non-canonical poly(A) polymerases **152**

6.1 Introduction 152

6.2 Aim 153

6.3 Expression screen of the human non-canonical poly(A) polymerases 153

6.4 Discussion 157

7 Conclusions **158**

7.1 Future directions 163

References **164**

Appendices **173**

List of Figures

Figure 1.1	A schematic representation of the human exosome subunits and associated co-factors	3
Figure 1.2	Structure and subunits of the human exosome	3
Figure 1.3	Schematic representation of the pre-mRNA 3' end processing complex in mammals	13
Figure 1.4	Schematic representation of the *S. cerevisiae* TRAMP4 complex	20
Figure 1.5	Structure and homology of TRAMP components in humans and budding yeast.	21
Figure 1.6	Structure of a Trf4/Air2 subcomplex	24
Figure 1.7	Processing steps of the 47S pre-rRNA transcript in Hela cells	31
Figure 1.8	Chemical structure of 5-FU	35
Figure 1.9	Sequence alignment of POLS and PAPD5	51
Figure 1.10	Sequence alignment of Trf4, Trf5 POLS and PAPD5	53
Figure 2.1	pUHD10-3 plasmid map	71
Figure 2.2	pcDNA3.1 plasmid map	72
Figure 2.3	pGEX-6P-1 plasmid map	73
Figure 3.1	Subcellular localization of HA-tagged POLS in transiently transfected U2-OS cells	90
Figure 3.2	Subcellular localization of POLS-TAP in U2-OS cells	90
Figure 3.3	Cytoplasmic and nuclear fractionation of HA epitope-tagged POLS in U2-OS cells	92
Figure 3.4	POLS-HA expression in stable transfectants	93
Figure 3.5	Subcellular localization of POLS-HA in U2-OS cell stable transfectant K10	94
Figure 3.6	Subcellular localization of POLS-HA in U2-OS cell stable transfectant K8	95
Figure 3.7	POLS or PAPD5 knockdown	97
Figure 3.8	Depletion of POLS or/and PAPD5 results in a slow growth phenotype	99
Figure 3.9	Cell cycle analysis of POLS down-regulated *vs* scrambled control cells	101
Figure 3.10	Flow cytometric analysis of cell cycle profile of simultaneously down-regulated POLS and PAPD5 cells *vs* scrambled control	102
Figure 4.1	The human 47S pre-rRNA transcript.	108
Figure 4.2	Schematic representation of the 3'RACE assay used for the analysis of 3'ETS tails	111
Figure 4.3	Schematic representation of the 3'ETS tails detected by 3' RACE	112
Figure 4.4	β-actin products amplified from Taq or Expand high-fidelity polymerases after serial dilutions of cDNA	114
Figure 4.5	Sequence profiles showing non-A residues detected within the A tails	117
Figure 4.6	Comparison of poly(A) and residual 3'ETS tail lengths.	120
Figure 4.7	Northern hybridization of total RNA and poly(A) selected RNA with the h3'ETS probe	121

Figure 4.8	Schematic representation of the primers used to clone across the 28S-3'ETS boundary	122
Figure 4.9	Cloning across the human 28S and 3'ETS boundary	123
Figure 4.10	The 28S-3'ETS boundary products are RNA-specific	124
Figure 4.11	Schematic representation of the primers used for the reverse transcription and analysis of the product oligo-dT specificity	125
Figure 4.12	The 28S-3'ETS boundary products are not oligo(dT)-specific	125
Figure 4.13	Stable transfectant clone K10 expressing POLS-HA treated with varying concentrations of 5-FU for 48 hr	127
Figure 4.14	Stable transfectant clone K10 expressing POLS-HA treated with varying concentrations of 5-FU for 48 hr	129
Figure 5.1	An overview of the TAP strategy	135
Figure 5.2	SDS gel analysis of proteins co-purified with POLS-TAP and TAP	136
Figure 5.3	Analysis of POLS and CPSF-73 interaction	143
Figure 5.4	Analysis of POLS and BRCA1 interaction	145
Figure 5.5	Analysis of POLS-TAP and Ku80 interaction	145
Figure 5.6	Detection of immunoprecipitated POLS-HA	146
Figure 5.7	Analysis of POLS-HA and RAD50 interaction	146
Figure 5.8	Speculative model of the roles of POLS	150
Figure 6.1	The domain organization of the human non-canonical poly(A) polymerases	152
Figure 6.2	Splice variant of PAPD5 identified in this study	154
Figure 6.3	The first ZCCHC6 splice variant identified in this study	155
Figure 6.4	The second splice variant of ZCCHC6 identified in this study	156

List of Tables

Table 1.1	Yeast and human exosome components and exosome-associated factors	9
Table 2.1	siRNA duplex sequences used in this study	79
Table 2.2	Oligonucleotides used in the expression screen of the human non-canonical poly(A) polymerases	80
Table 2.3	Q-PCR oligonucleotides used in this study	81
Table 2.4	Oligonucleotides used in the 3' RACE screen	81
Table 2.5	Oligonucleotides used to clone across the human 28S and 3'ETS boundary	82
Table 2.6	Oligonucleotides used in northern analysis	82
Table 2.7	Oligonucleotides used to generate GST fusion proteins	82
Table 2.8	Antibodies used for western blot in this study	83
Table 2.9	Antibodies used for co-immunoprecipitation in this study	83
Table 2.10	Recipes for common buffers, solutions and media used in this study	84
Table 3.1	Non-canonical poly(A) polymerases in various species, their localization and functions	87
Table 4.1	One cDNA can give rise to many different PCR products	113
Table 4.2	A residues and non-A residues recorded in the actin tail screen with Taq or Expand high fidelity polymerases.	115
Table 4.3	The frequency of erroneous nucleotide incorporation for β-actin and 3'ETS rRNA with Taq or Expand highfidelity polymerases	118
Table 4.4	A residues and non-A residues recorded in the Slomovic *et al*., 2010 study for β-actin and 28S rRNA tail screens	118
Table 5.1	Proteins identified as interacting partners of POLS by mass spectrometry	137
Table 5.2	Relatedness of POLS-interacting zinc finger proteins with yeast Air1 and Air2	140
Table 5.3	RNA helicases and zinc finger proteins identified as interacting proteins with POLS by MS analysis	141
Table 5.4	Transcription-related proteins identified as interacting proteins with POLS by MS analysis	142
Table 5.5	DNA repair-related proteins identified in this study as interacting proteins of POLS	144
Table 5.6	Cohesion, nuclear structure and nuclear pore related proteins identified as interacting proteins with POLS by MS analysis	147
Appendix **A**	Human 3'ETS tails identified in the DKD and control cells	173
Appendix **B**	Human β-actin tails identified with Taq and Expand High Fidelity polymerases	174
Appendix **C**	Generation of POLS polyclonal antibody.	179
Appendix **D**	Interacting protein partners of POLS	180

Glossary of Abbreviations

aa	amino acid
bp	base pair
C.elegans	*Caenorhabditis elegans*
cDNA	complementary DNA
co-IP	co-immunoprecipitation
dATP	deoxy adenosine triphosphate
dCTP	deoxy cytosine triphosphate
dGTP	deoxy guanosine triphosphate
DIC	Differential interference contrast
dTTP	deoxy thymidine triphosphate
dNTP	deoxy ribonucleotide triphosphate
dH_2O	distilled water
DKD	double knockdown
D.melanogaster	*Drosophila melanogaster*
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ds	double-stranded
EDTA	ethylene diaminetetra-acetic acid
g	gram
g	relative centrifugal force
hr	hour
HA	haemagglutinin (YPYDVPDYA-aa)
KD	knock-down
KDa	Kilo Dalton
l	litre
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani
M	Molar (moles per litre)
min	minute
MOPS	3-(N-morpholino) propanesulfonic acid
mRNA	messenger RNA
miRNA	microRNA
nt	nucleotide

NTP	Nucleotidyltransferase
NuRD complex	Nucleosome remodeling and deacetylation
OD	optical density
ORF	open reading frame
PAP	poly(A) polymerase
PAPD5	poly(A) polymerase-associated domain containing protein 5
PCR	polymerase chain reaction
Pol	polymerase
POLS	DNA polymerase sigma
PtdIns4,5P2	Phosphatidylinositol 4,5-bisphosphate
RACE	rapid amplification of cDNA ends
RHAU	RNA Helicase associated with AU-rich element
RNA	ribonucleic acid
RRM	RNA recognition motif
rRNA	ribosomal RNA
Rrp6	ribosomal RNA-processing protein-6
RT	room temperature
RT-PCR	reverse transcription-PCR
S	Svedberg
S. cerevisiae	*Saccharomyces cerevisiae*
S.pombe	*Schizosaccharomyces pombe*
SDS	sodium dodecyl sulphate
sec	second
TAP-tag	tandem affinity purification tag
TBE	Tris/Borate/EDTA
TBS	Tris-buffered saline
TEMED	N,N,N’ N’-tetramethylethylenediamine
Tris	Tris(hydroxymethyl) aminomethane
tRNA	transfer RNA
ts	temperature sensitive
UTP	uracil triphosphate
UTR	untranslated region
W	Watt
X-gal	5-bromo-4-chloro-3-inodyl-β-D-galactopyranoside

1 Introduction

1.1 RNA surveillance

Most RNA is found in the cell within ribonucleoprotein complexes that include many RNA binding proteins. The formation of mature RNPs necessitates that all coding and non-coding RNAs undergo various processing and modification events, which need to be orchestrated with great accuracy. On some occasions errors arise in the events leading to the formation of mature RNPs. Defective RNP production has the capacity to impair the function of almost every biological system. Cells have developed surveillance mechanisms, which enable aberrant RNAs to be degraded to maintain the proper functioning of regulatory pathways.

1.1.1 RNA decay

RNA degradation is achieved *via* three major classes of enzymes: internal cleavage by endonucleases, 5' hydrolysis by 5' exonucleases and 3' degradation *via* 3' exonucleases. A major component of the RNA quality control machinery is the exosome. The exosome degrades many types of RNA that have been targeted by various surveillance pathways in the cytoplasm and nucleus *via* 3'-5' exonucleolysis (Houseley *et al.*, 2006). The exosome is also involved in trimming nuclear precursor transcripts in a 3'-5' orientation with high accuracy. The 5' human exonucleases XRN2 and XRN1 are involved in degrading RNAs from the 5' terminus. XRN2 is found both in the nucleus and cytoplasm while XRN1 localizes only to the cytoplasm. Budding yeast contains 5' exonucleases Rat1 and Xrn1. Rat1 is found in the nucleus while Xrn1 is cytoplasmic.

The exosome ensures normal RNA processing and turnover in the cell.

1.1.2 Exosome

The exosome is a multi-protein 400 kDa complex with 3' exonuclease and endonuclease activities. It was originally discovered through its role in 5.8S rRNA processing in the nucleus (Mitchell *et al.*, 2003). It plays a crucial role in many stages of RNA metabolism including quality control of mRNA (including histone mRNA), ribosomal RNA, tRNA, snRNA, snoRNA, processing of pre-mRNA and rRNA, and gene silencing. Hence, the exosome is associated with RNA processing and quality control of transcripts produced by all three polymerases (I, II and III). Most RNAs encounter the exosome before reaching their mature form. The exosome core forms a ring-shaped structure (Figures 1.1, 1.2). The ring consists of three distinct heterodimers made up of six different proteins: Rrp41-Rrp45, Rrp46-Rrp43 and Mtr3-Rrp42. The six-protein ring is stabilized by a cap consisting three different proteins: Rrp4, Csl4 and Rrp40. The cap proteins contain either a K-homology domain (KH) or ribosomal protein S1 (S1) type RNA binding domains (Anderson and Wang, 2009). Many structural and functional features are conserved among bacteria, archaea and eukaryotes. All components of the yeast exosome are essential for viability with the exception of Rrp6 (Allmang *et al.*, 1999a). Loss of Rrp6 results in reduced growth (Briggs *et al.*, 1998).

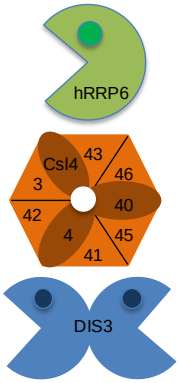


Figure 1.1 A schematic representation of the human exosome subunits and associated co-factors.
The human exosome constitutes a ring of three distinct heterodimers containing six RNase PH-like proteins: RRP41-RRP45, RRP46-RRP43 and MTR3-RRP42. Three additional proteins with a K-homology domain (KH) or S1 RNA-binding domains RRP4, Csl4 and RRP40 form a cap for stability. Associated co-factors include hRRP6 which harbours endonuclease activity and DIS3 which has both endo- and exo-nuclease activities.

The exosome is present both in the nucleus and cytoplasm. The nuclear exosome is present in much larger complexes (60-80S) compared to the approximately 10S cytoplasmic exosome (van Dijk *et al.*, 2007).

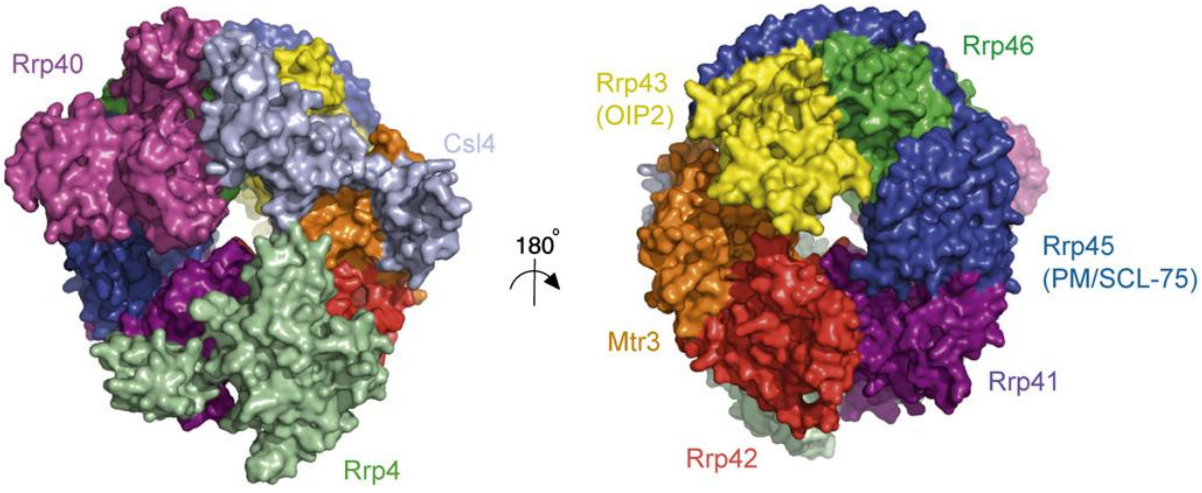


Figure 1.2 Structure and subunits of the human exosome.
Colour coded subunits are indicated. Images from Liu *et al.* (2006).

Reprinted from Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell 127, 1223-1237 with permission from Elsevier. doi:10.1016/j.cell.2006.10.037

In eukaryotes the RNAse PH units have lost their catalytic activity making the core inactive. The core exosome provides a structural scaffold while the catalytic activity is attributable to the co-factors that associate with the exosome.

1.1.2.1 Catalytic subunits of the exosome

Co-factors associated with the exosome include Dis3 and Rrp6. A+U rich elements, Puf proteins and Nrd1-Nab3 complex bind to specific sequences and recruit the exosome.

Dis3

Dis3, also known as Rrp44, and PM-Scl100 (yeast-Rrp6) associate stably with the exosome core and are responsible for the catalytic activity. Yeast Rrp6 is found only in the nucleus while the human orthologue PM-Scl100 localizes both to the nucleus and cytoplasm. Dis3 is found both in the nucleus and the cytoplasm in yeast (Synowsky *et al.*, 2009). Dis3 orthologues, hDis3 and hDis3L1, have been reported in humans (Staals *et al.,* 2010). Localized in the nucleus and cytoplasm, hDis3 contrasts with hDis3L1, which has a strictly cytoplasmic localization. Dis3 is a hydrolase which processively degrades both structured and unstructured RNA. In yeast, Rrp4, Csl4 and Rrr40 proteins are found at the opposite end to the Dis3 binding site. RNA is threaded through the central channel before being degraded by Dis3 (Malet *et al*., 2010). It is possible that a similar sequence of events take place in other eukaryotes. However, Dis3 co-purifies with the rest of the core in *S.cerevisiae* and in *D. melanogaster*, but not in *H. sapiens* or *T. brucei,* indicating possible functional differences (Malet *et al*., 2010).

Dis3 has been shown to possess both exonucleolytic and endonucleolytic activities. The *E.coli* degradosome, the main complex involved in RNA degradation, also contains both endonucleolytic and exonucleolytic activities (Dziembowski *et al.*, 2007). Degradosome functions are performed by proteins endoribonuclease E and polynucleotide phosphorylase, which harbour no similarity to Dis3. Dis3 has two cold shock domains (CSDs), an

exonuclease domain and a C-terminal S1 RNA binding domain. It also contains a pilT N-terminal (PIN) domain which harbours the endonuclease activity and tethers Dis3 to the core exosome. The nucleolytic activity of Dis3 requires a relatively high concentration (1-3 mM) of divalent cation (Mn^{2+}, Mg^{2+} or Zn^{2+}). Manganese was found to be the most efficient while magnesium least. However 3mM is much greater than the physiological concentration of Mn^{2+} in the yeast cell. Hence, the mechanism by which PIN domain achieves the required concentration of divalent cation to exert endonucleolytic activity is currently unknown. RNA-binding domain (RNB) exoribonucleases are most efficient with low magnesium concentrations 40-100μM (Dziembowski *et al.*, 2007). The requirement for a high concentration of Mn^{2+} is possibly a regulatory mechanism enabling only aberrant RNAs to be degraded by providing suboptimal conditions. If the environment was optimal for Dis3 function perhaps excessive transcript degradation may prevail.

In humans the endonucleolytic activity of the exosome is mainly a nuclear function attributed to hDIS3 (Lebreton *et al.*, 2008). Loss of one of the three highly conserved aspartic acids in hDIS3L1, results in poor conservation to the yeast Dis3 PIN domain. These aspartic acids have been shown to be essential for the endonucleolytic activity of PIN domains. The yeast PIN and RNB catalytic domain double mutants show an increased accumulation of exosome targets compared to the single mutants. For instance, enhanced accumulation of 5'ETS ribosomal RNA processing by-products and precursor of 5.8S rRNA were observed. This indicates that these domain activities are redundant. Point mutations in the Dis3 PIN domain have little impact on RNA processing and cell growth. Hence, in the absence of endonucleolytic cleavage, RNA is still degraded but perhaps inefficiently (Lebreton *et al.*, 2008). The RNB domain of hDis3L1 is associated with the exonucleolytic degradation of 28S rRNA in the cytoplasm (Staals *et al.*, 2010).

Rrp6

In the yeast nucleus, RNAs destined for degradation and processing intermediates are polyadenylated by the TRAMP (Trf4/5 non- canonical poly(A) polymerase, Air1/2 zinc finger protein and Mtr4 RNA helicase Polyadenylation) complex. Polyadenylated transcripts apart from poly(A)$^+$ mRNAs are channeled into Rrp6-exosome mediated RNA processing and RNA surveillance pathways. Rrp6 contains a core riboexonuclease domain, a domain involved in regulating its 3' end processing function and nuclear localization domain. Rrp6 works in a distributive manner and prefers unstructured substrates (Schmid and Jensen, 2008). Absence of Rrp6 is not lethal in *S.cerevisiae* but impairs growth and leads to temperature sensitivity. Cleavage intermediates destined to be 3' processed or degraded can be detected in the absence of Rrp6 (Egecioglu *et al.*, 2006).

Strains lacking Rrp6 show RNA processing defects that are distinct from those seen in core exosome mutants. The activities of the core exosome and Rrp6 are separable in RNA processing but not in RNA degradation. Rrp6 acts downstream from the core exosome, as exosome mutants accumulate pre–RNAs with 3' extensions longer than those found in *rrp6* mutants (Allmang *et al.*, 1999a). Rrp47, a core exosome component, promotes the catalytic activity of Rrp6 (Stead *et al.*, 2007). After the exosome produces the pre-5.8S rRNA, Rrp6 removes the last 30 nt from the 3' end of the precursor (Allmang *et al.*, 1999b). Rrp6 is also responsible for removing approximately 3 nt from some snoRNAs (van Hoof *et al.*, 2000).

The yeast nuclear exosome associates with the TRAMP complex. The cytoplasmic exosome interacts with the Ski complex. The nuclear exosome requires the RNA helicase Mtr4 (also known as Dob1), while the cytoplasmic exosome requires Ski2. These helicases are thought to help unwind secondary structures, enabling the TRAMP and Ski complexes to recruit substrate RNAs to the exosome.

In the cytoplasm, the exosome interacts with the GTP-binding protein Ski7, which recruits the exosome to ribosomes stalled at mRNA 3' ends that lack stop codons, promoting their degradation (Houseley *et al.*, 2006). In humans Ski7 is found in the nucleus too, although its function in the nucleus is unknown. Ski2, Ski3 and Ski8 proteins form a complex. Ski2 is an RNA helicase homologous to Mtr4, but Ski3 and Ski8 shows no homology to TRAMP components. Ski2-Ski3-Ski8 and Ski7 exosome complexes participate in the Nonsense-Mediated Decay **(**NMD) and non-stop decay mRNA surveillance pathways. No poly(A) polymerase has been linked to the cytoplasmic exosome. Cytoplasmic exosome functions in mRNA deadenylation and mRNA turnover. Exosome-mediated degradation of mRNAs follows the different pathways identified below:

A+U rich sequence elements-Human (AREs)

AREs found in the 3' UTRs of many regulated mRNAs which are degraded *via* the exosome. The exosome is recruited *via* ARE-binding proteins (AUF1, KSPR and TTP) or the RNA helicase RHAU. AREs are able to bind directly to the PM-Scl175 subunit and exacerbate decay (Hartung and Hopfner, 2007). AREs have been linked to influencing the 5' decapping and deadenylation steps of mRNA turnover (Gherzi *et al.*, 2004).

Puf proteins

Messenger RNA decay is thought to be synchronized by binding to PUF proteins. These PUF proteins are homologous to the *Drosophila* Pumilio RNA binding protein which binds the 3' UTR of developmentally regulated genes and represses translation (Wharton *et al.*, 1998). Puf proteins regulate gene expression post-transcriptionally in metazoan development. Puf3 binds *COX17* mRNA, which is rapidly deadenylated by the CCR4-NOT complex. The founding member of the Puf family of proteins, *Drosophila* Pumilio, acts with Nanos to

promote deadenylation of hunchback mRNA at the posterior end of the *Drosophila* embryo. Deadenylated hunchback mRNA then undergoes exosome-mediated degradation.

Nrd1-Nab3 complex

Nrd1, with its interacting partners, the RNA binding protein Nab3, RNA helicase Sen1, and cap binding complex (CBC), interacts with the nuclear exosome. The Nrd1 complex recruits the exosome to the 3' ends of RNA substrates to promote degradation and processing (Houseley and Tollervey, 2009). Nrd1 promotes degradation both *in vivo* and *in vitro*. However it is also reported that at some Nrd1 binding sites Nrd1 can also block 3' to 5' degradation by the exosome. Like exosome mutants, Nrd1 mutants show increased read-through transcription of some mRNA and sn/snoRNA genes. In the *nrd1-2* strain a dramatic increase in *CYC1* and *NAB2* mRNA was recorded. However no accumulation of extended *NAB3* and *SEN1* mRNAs was found in exosome mutants or *nrd1* mutants. Hence, the exosome might not be associated with the degradation of all read-through mRNAs. It is possible that some genes do not produce read-through transcripts. The RNA sequence 5'-CUU-3' is recognized by Nab3 and 5'-GUAAA-3' recognized by Nrd1. Nrd1 is capable of down-regulating its own mRNA levels, leading to exosome-mediated depletion of Nrd1. Cryptic unstable transcripts (CUTs) are rapidly degraded by the Nrd1-coupled exosome. Deletions or mutations in the Nrd1 complex or Rrp6 and Rrp4 leads to the accumulation of heterochromatic transcripts that would normally be degraded by this pathway (Vasiljeva and Buratowski, 2006).

The Nrd1/exosome complex is capable of terminating and degrading RNAP II transcripts (Vasiljeva *et al.*, 2008b) **.** Nrd1, Nab3 and Sen1 complex is also involved in Pol II transcription termination. The transcripts involved in this termination pathway include snRNA, snoRNA and CUTs which are all short transcripts. *S.cerevisiae* contains three C-terminal domain (CTD) interaction domain (CID) proteins, which bind the CTD of the Pol II largest

	yeast	y.loc.	human	h.loc.	Function in humans
Exosome core	Rrp4	N+C	hRrp4 (EXCSC2)	N+C	
	Rrp40	N+C	hRrp40 (EXCSC3)	N+C	
	Rrp41(Ski6)	N+C	hRrp41 (EXCSC4)	N+C	
	Rrp42	N+C	hRrp42 (EXCSC7)	N+C	structural scaffold
	Rrp43	N+C	hRrp43 (EXCSC8)	N+C	essential for viability
	Rrp45	N+C	hRrp45 (EXCSC9)	N+C	
	Rrp46	N+C	Rrp46 (EXCSC5)	N+C	
	Mtr3	N+C	hMtr3 (EXCSC6)	N+C	
	Csl4	N+C	hCsl4 (EXCSC1)	N+C	mRNA decay (Y)
Exonuclease	Rrp6	N	PMScl100(EXCSC10)	N+C	processing and decay
			Dis3L1	C	degradation of 28S rRNA
Exonuclease + Endonuclease	Rrp44 (Dis3)	N+C	Rrp44 (hDis3)	N+C	rRNA processing and decay
RNA helicase-TRAMP	Mtr4 (Dob1)	N	SKIV2L2 (hMtr4)	N+C	recruits the nuclear exosome (Y)
RNA helicase Ski complex	Ski2	C	SKIV2L2	N+C	recruits the cyto. exosome (Y)
	Ski3	C	hSki3	N+C	essential for viability
	Ski8	C	hSki8	N+C	essential for viability
Ski7	Ski7	C	?	?	mRNA decay
Co-factors	Rrp47	N	hRrp47	N	Interacts with SKIV2L2
	Mpp6	N	Mpp6	N	aids interaction of SKIV2L2 with Rrp6 and Rrp47
Nrd1-Nab3 complex	Nrd1	N	?	?	sn/snoRNA transcription
	Nab3	N	?	?	decay of some mRNA & cryptic pol
	Sen1	N	Senataxin	?	ll transcripts
NEXT complex	?	?	RBM7	N	degradation of promoter-upstream transcripts.
	?	?	ZCCHC8	N	
	Mtr4	N	hMTR4	N	

Table 1.1 Yeast and human exosome components and exosome-associated factors.

loc.: localization, N:nucleus, C: cytoplasm. Functions for the human components are shown. In the absence of a known function for human the yeast function is shown, indicated by (Y). Unknowns are indicated by: ?

subunit. Nrd1 binds to CTD phosphorylated at Ser5 found proximal to promoters, consistent with the termination of short transcripts (Vasiljeva and Buratowski, 2006; Vasiljeva *et al*., 2008a).

Messenger RNAs (mRNAs) are deadenylated and some mRNA transcripts are destroyed while others are kept for later release, readenylation and translation. mRNAs to be destroyed are sent to processing bodies (PB). Found in the cytoplasm, PB contain components of the 5'-3' decay machinery, nonsense-mediated decay pathway and in mammalian cells, RNA-induced silencing complex. Thus, PB are involved in mRNA decay and silencing processes. Stress granules found in the cytoplasm decide the fate of stress associated mRNAs. Stress granules are thought to act as intermediates between polysomes and PBs (Anderson and Wang, 2009; Kedersha *et al.*, 2005). It was thought that prior to 5' degradation in PBs mRNAs were released from polysomes; however in *S.cerevisiae* 5' degradation can occur on polysomes (Hu *et al.*, 2009). mRNA decay occurring 3'-5' *via* the exosome is described in section 1.1.2.

Quality control of cytoplasmic mRNAs

NMD occurs in aberrant mRNAs in which translation has been prematurely terminated owing to the presence of a nonsense codon within the coding region, also known as a premature termination codon (PTC), preventing the expression of truncated or erroneous proteins. Such transcripts are decapped and degraded 5'-3' or deadenylated and degraded 3'-5' *via* the exosome. The termination codon is sensed by a complex involving Upf1, Upf2 and Upf3. The NMD factors and PTC-containing transcripts are found in PBs in the cytoplasm where rapid destruction of aberrant mRNAs takes place (Sheth and Parker, 2003, 2006). The **Non Stop Decay (NSD)** pathway involves the recruitment of the exosome *via* Ski7 to ribosomes that have reached the end of mRNAs lacking translation termination codons. It is thought that Ski7 is able to recognize an empty A site produced when a ribosome reaches the 3' end of

an mRNA (Anderson and Wang, 2009). Exosome mutations prevent non-stop decay and lead to stabilization of a fully polyadenylated RNA. Proteins Hbs1 and Dom34 promote **No Go Decay (NGD)** by interacting with stalled ribosomes. Target substrates are cleaved endonucleolytically. **Ribosome Extension Mediated Decay (REMD)** involves destabilization of mRNAs when ribosomes inappropriately translate beyond the normal stop codon into the 3' UTR.

The best studied nuclear RNA waste destruction pathway is the yeast TRAMP (Trf4/5 non-canonical poly(A) polymerase, Air1/2 zinc finger protein and Mtr4 RNA helicase Polyadenylation) complex involved in the 3' to 5' exosome-mediated decay of non-coding transcripts. Non-coding transcripts may also enter the RNAi pathway although this is absent from *S. cerevisiae*.

1.1.3 3' and 5'end modifications

Transcripts undergo modifications at the 3' and 5' end which are important in deciding their fates in all eukaryotes.

1.1.3.1 3' end polyadenylation

Addition of a string of non-templated adenosine residues to the 3' end of a transcript is common to all organisms, except halophilic and some methanogenic archaea and *Mycoplasma*. In eukaryotes, canonical polyadenylation of pre-mRNA is initiated by endonucleolytic cleavage which is followed by the addition of poly(A) or poly(A)-rich sequences to the proximal cleavage products. In bacteria poly(A) destabilizes mRNA. In eukaryotes mature mRNA polyadenylation confers mRNA stability, proper translation initiation and nuclear export.

Polyadenylation can also be a transient process involved in processing and decay. A wide variety of transiently polyadenylated RNAs have been recorded in cells. These poly(A) tails are not added to the mature 3' end of transcript but to the 3' end of degradation or processing intermediates. Transient poly(A) tails are not abundant as they undergo rapid degradation or processing. Polyadenylated RNA fragments are found in human and trypanosomal mitochondria. In yeast short A tail transcripts (LaCava *et al.*, 2005) and in humans short poly(A)/oligo(A) tail transcripts (West *et al.*, 2006) were shown to accumulate in the absence of the exosome. Human 18S and 28S rRNAs also containing poly(A)/oligo(A) tails corresponding to degradation or processing intermediates have been reported (Slomovic *et al.*, 2006). Aberrant RNAs, non functional RNAs and processing intermediates are poly/oligo-adenylated by the TRAMP or TRAMP-like complexes which undergo endo-/exo-nucleolytic processing or exonucleolytic degradation.

i. Nuclear polyadenylation and deadenylation

The canonical hexanucleotide (AAUAAA) polyadenylation signal found on most eukaryotic pre-mRNAs initiates cleavage and subsequent polyadenylation by the nuclear poly(A) polymerase (PAP), which is considered the major polyadenylation pathway in eukaryotes (Edmonds, 2002). The AAUAAA sequence is found 15-30 nucleotides upstream of the cleavage site. Located downstream from the cleavage site is the GU-rich downstream element (DSE). The upstream element (USE) is U-rich. The hexanucleotide sequence is associated with the Cleavage and Polyadenylation Specificity Factor (CPSF) which is required for cleavage. CPSF is comprised of a complex of proteins CPSF-160, CPSF-100, CPSF-73 and CPSF-30. The polyadenylation cleavage stimulation factor (CstF) is bound to the DSE and is also required for cleavage. CF I_m, CF II_m and Symplekin are also associated with 3' end processing of pre-mRNA in mammalian cells (Figure 1.3).

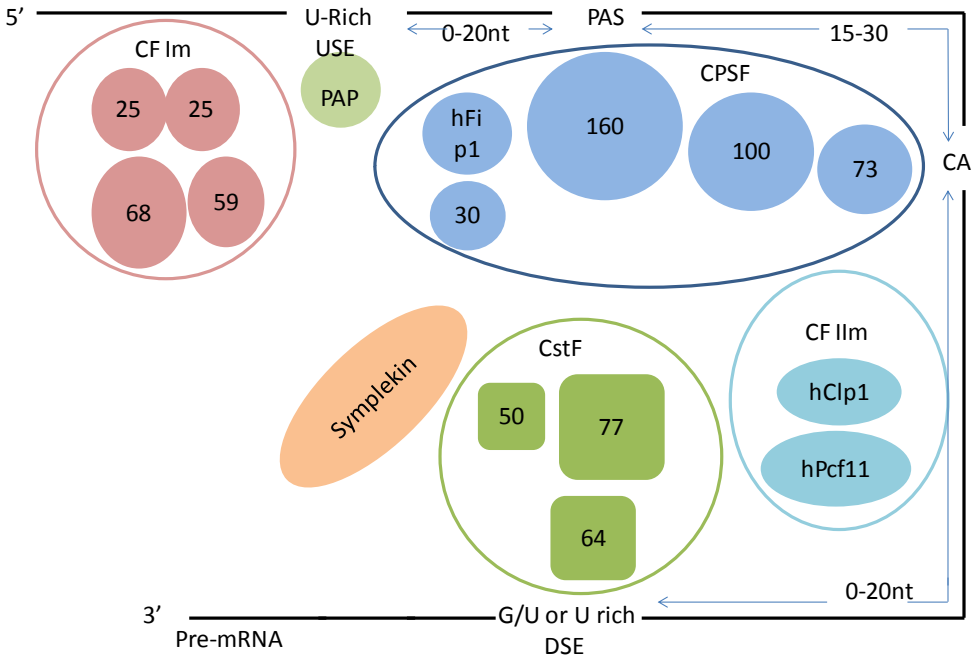


Figure 1.3 Schematic representation of the pre-mRNA 3' end processing complex in mammals.

CPSF-160 recognizes the polyadenylation signal (PAS) while CstF-64 recognizers the DSE. Together these complexes promote RNA cleavage between these *cis* elements. Cleavage factors such as CF I_m and CF II_m are also required for this process. The USE which is U-rich is found 0-20nt upstream from the PAS. The PAS is located 15-30nt upstream from the cleavage site. The cleavage site is found 0-20nt upstream from the DSE.

The nuclear poly(A) binding protein (PABPN1) found in mammalian cells binds the poly(A) tail. CPSF and PABPN1 are believed to enhance the activity of PAP. CPSF recruits PAP to RNAs containing the AAUAAA or a close varient of the sequence at the 3' end and PABPN1 is attracted to terminal oligo(A) tracts. With either factor alone (CPSF or PABPN1) PAP adds two to five nucleotides before dissociating. In the presence of both CPSF and PABPN1 incorporation of the poly(A) tail is processive. In human cells tails are elongated to up to 200-250nt in this processive manner by PAP before termination.

Nearly all genes have additional polyadenylation signals in their 3' UTR (Tian *et al.*, 2005). Cancer cells use alternative cleavage and polyadenylation (APA) signals to generate mRNAs with short 3' UTRs. As a consequence they are no longer regulated by the stronger regulation of long 3'UTR repressive elements, leading to increased protein production. Mayr and Bartel report ten-fold more protein production, mainly due to loss of microRNA-mediated repression (Mayr and Bartel, 2009).

Most data relating to nuclear deadenylation come from studies of the *S. cerevisiae* TRAMP complex (discussed in section 1.2.1).

ii. Cytoplasmic polyadenylation and deadenylation

Cytoplasmic polyadenylation of mRNA leads to stability while deadenylation causes translational repression regulating gene expression. In many animal species, during early development cytoplasmic addition of poly(A) controls translation of specific maternal mRNAs (Barnard *et al.*, 2004). Messenger RNAs receive poly(A) of different lengths and at different times during development. In *Xenopus* oocytes, cycling somatic cells and in neurons the cytoplasmic polyadenylation element binding (CPEB) protein associates with cytoplasmic polyadenylation element (CPE), which specifies RNAs to be polyadenylated. As in the nucleus, CPSF binds the AAUAAA hexanucleotide. Symplekin binds to both CPEB and CPSF and keeps these factors in close contact until CPEB gets phosphorylated and activates polyadenylation by xGLD-2.

The Schuster group report the accumulation of 28S rRNAs and β-actin mRNA fragments containing poly(A)/oligo(A) tails in both the nucleus and cytoplasm of human cells as a result of RNAi-mediated silencing of PM-Scl100, hXrn1, hDis3 or hDis3L (Slomovic *et al.*, 2010). This is the first report of poly(A)/oligo(A) tail addition in the cytoplasm to aberrant transcripts associated with degradation in human cells.

Deadenylation in the cytoplasm is carried out by Mg^{2+}-dependent exoribonucleases that hydrolyse RNA in the 3'-5' direction releasing 5' AMP. The three main complexes involved in deadenylation include CCR4-CAF-NOT, PAN2-3 and PARN. Ccr4 is the major deadenylase in yeast and humans (Bilger *et al.*, 1994). In mammalian cells Pan2-Pan3 reduces the tail to 50nt, the rest of the tail is degraded by Ccr4-Not complex as in *S.cerevisiae* or by PARN. Recruitment of Pan2-Pan3 to mRNA 3'ends is carried out by PABP, which also inhibits exonucleolytic activity of the Ccr4-Not complex. PABP is unable to protect a tail of 10-15nt resulting from deadenylation. This short tail is recognized by Lsm1-7 followed by subsequent decapping (Chicoine *et al.*, 2007). Regulatory elements in the 3' UTR are known to mediate the recruitment of Ccr4-Not.

None of the deadenylases is essential for yeast cell viability. However, *A.thaliana* and *Xenopus laevis* PARN and *D.melanogaster* CCR4 and *C.elegans* CCF-1 are essential during early development (Goldstrohm and Wickens, 2008). Reduction of *D.melanogaster* CCR4 and *C.elegans* CCF-1 results in sterility. Deadenylation can occur in granules or as cells enter meiosis. For instance maternal mRNAs are deadenylated *via* PARN by its release to the cytoplasm in *Xenopus* meiosis due to nuclear breakdown. Deadenylases can be found in germline and neuronal cells. They also localize to stress granules and processing bodies. The mRNA 5' cap can influence mRNA deadenylation by stimulating PARN. This is due to the direct interaction with PARN and the cap structure which does not occur with other deadenylases.

iii. Nuclear uridylation

In the nucleus, uridine residues are added to the small nuclear RNA U6, enabling splicing and stabilization of the transcript (Trippe *et al.*, 2006).

iv. Cytoplasmic uridylation

Some non-canonical poly(A) polymerases can incorporate residues other than AMP (Read *et al.*, 2002, Kwak *et al.*, 2004; Kwak and Wickens, 2007; Rissland *et al.*, 2007). Non-canonical poly(A) polymerases that incorporate UTP include *S.pombe* Cid1, *H.sapiens*: ZCCHC11, ZCCHC6, *C.elegans* PUPs: PUP-1(K10D2.3), PUP-2(K10D2.2), PUP-3(F59A3.9) and Arabidopsis At1. Uridylation and deadenylation are thought to act redundantly to stimulate decapping (Rissland and Norbury, 2009). *S. pombe* Cid1-dependent uridylation of polyadenylated mRNA and uridylation of human ZCCHC11-mediated histone mRNAs directs decapping *via* the Lsm1-7 complex and promotes degradation (Rissland *et al.*, 2007; Schmidt *et al.*, 2011).

By analogy with nuclear oligo(A) tails, oligo(U) tails might act as anchors for the cytoplasmic exosome . *In vitro* Rrp44 the cytoplasmic exosome nuclease prefers AU-rich sequences over oligo(A) supporting this observation (Liu *et al.*, 2006). Uridylation of pre-miRNA prevents maturation by Dicer leading to degradation of these pre-miRNAs (Heo *et al.*, 2009). In *Aspergillus nidulans* decapping is associated with a novel CUCU sequence modification added by CutA, belonging to the nucleotidyltransferase superfamily. Mutations of CutA lead to reduced decapping and subsequent transcript degradation (Morozov *et al.*, 2010).

1.1.3.2 5’ end modifications

The 7-methylguanosine cap structure from the 5’ end of mRNA is removed before degradation. Dcp1 and Dcp2 are involved in transcript decapping. The cytoplasmic Lsm1-7 complex involved in mRNA decapping shows strong binding preference for oligoadenylated RNA over polyadenylated and non-polyadenylated RNA (Chowdhury *et al.*, 2007). In fission yeast and mammalian cells decapping *via* the Lsm complex is stimulated by 3’-oligouridylation (Rissland and Norbury, 2009, Schmidt *et al*., 2011). The cap structure is broken down *via* hydrolysis by the scavenger enzyme DcpS. This is followed by the 3’-5’

exonucleolytic decay by the exosome or 5'-3' exonucleolytic decay Xrn1 (Yeast Rat1). Yeast Rat1 acts in a complex with cofactors Rai 1 and Rtt103 (Kim *et al.*, 2004). Capped RNA is resistant to Xrn1 and RNA carrying a 5' monophosphate is strongly preferred over RNA with a 5' hydroxyl end.

1.2 TRF4 and TRF5

S.cerevisiae Trf4 (YOL113W) and Trf5 (YNL299W) are two widely studied proteins that are involved in nuclear RNA surveillance. In 1995 the Christman group first identified *TRF4*, a non-essential gene, in a screen for mutations that are non-viable in combination with a DNA topoisomerase I (*top1*) null mutation (Sadoff *et al.*, 1995). In 1996, the same group observed that the hypersensitivity of *trf4* mutants to the topoisomerase I inhibitor and antitumor drug Camptothecin could be suppressed by overexpression of *TRF4*. The same approach led to the identification of *TRF5,* a *TRF4*-related gene (Castano *et al.*, 1996).

Originally, on the basis of an apparent template-dependent DNA polymerase activity, *TRF4* was assigned to represent the fourth class of essential nuclear DNA polymerases: DNA polymerase κ (Wang *et al*., 2000b). In this study the Trf4 was expressed in *E.coli* which has DNA pol I activity. A neutralizing monoclonal antibody was used to inhibit the *E.coli* DNA pol I activity to ensure the DNA polymerase activity measured was specific to Trf4. The Trf4 DNA polymerase activity shown was dependent on Mg^{2+}. Edward *et al.* (2003) supported this view by showing Trf4 to interact with the C-terminal domain of DNA polymerase ε (Polε, Pol2), a functional and genetic interaction stimulating the polymerase activity of DNA pol ε. This is based on defective sister chromatid cohesion observed in *pol* ε mutants, as in Trf4/5 mutants. Pol ε and Trf4 have been shown to interact with *SMC1,* a subunit of the cohesin complex. Pol ε also interacts with *ECO1/CTF7*, required for establishing sister chromatid cohesion.

Contradicting these observations, (Haracska *et al.*, 2005)) reported Trf4 and Trf5 to be RNA polymerases. Trf4 shows strong affinity for Mn^{2+} and incorporates AMP into the free 3'-hydroxyl end of an RNA primer. No poly(A) polymerase activity was detected when Mg^{2+} was present. Recent years have seen Trf4 and Trf5 assigned to the non-canonical poly(A) polymerase family.

Both Trf4 and Trf5 proteins are documented as localizing to the nucleus (Houseley and Tollervey, 2006; Walowsky *et al.*, 1999). The Trf4 amino acid sequence is 47.1% identical and 59.4% similar to Trf5, with much of the similarity concentrated in the central region. Deletion of both Trf4 and Trf5 is lethal (Castano *et al.*, 1996), indicating that they have overlapping functions that are essential for cell viability. These polymerases lack an RNA binding domain. They exert their functions by associating with other proteins, forming the TRAMP complex (LaCava *et al.*, 2005; Vanacova *et al.*, 2005). Aberrant RNAs produced by different biogenesis pathways by all three nuclear RNA polymerases are substrates for this pathway. Known substrates to date include aberrantly modified tRNAs, precursors of rRNA, CUTs, snRNA and snoRNA (Houseley and Tollervey, 2006; Kadaba *et al.*, 2006; Vanacova *et al.*, 2005; Wyers *et al.*, 2005). TRAMP complexes polyadenylate target RNAs and activate the exosome. Aberrant RNPs produced by different biogenesis pathways are subjected to 3' to 5' exosome-mediated degradation.

No poly(A) polymerase activity is detected with the Trf4-236 and Trf5-D233A, Trf5-D235A (DADA) catalytic activity point mutations (Houseley and Tollervey, 2006). In these mutants the aspartate residues have been replace by alanines. Interestingly Trf4 and not Trf5 is associated with telomere maintenance and antisence RNA-mediated regulatory pathways of gene expression (San Paolo *et al.*, 2009). Trf5 over-expression was shown to hyperadenylate pre-rRNAs while over-expression of Trf4 led to hyperadenylation of an aberrant pre-tRNA (Houseley and Tollervey, 2006). The above indicates that these polymerases are substrate-specific for at least some targets.

1.2.1 TRAMP complex

The TRAMP complex was first identified in *S.cerevisiae* using Mtr4 as bait in a two–hybrid screen (LaCava *et al.*, 2005). (Vanacova et al., 2005)) used affinity tagged Trf4 to copurify the TRAMP complex from *S.cerevisiae* whole cell extracts. Two different TRAMP complexes, TRAMP4 and TRAMP5 have been identified in *S. cerevisiae*. TRAMP4 complex is composed of Mtr4 helicase, the poly(A) polymerase Trf4 and the RNA binding Air2 zinc finger protein (LaCava *et al.*, 2005; Vanacova *et al.*, 2005). The poly(A) polymerase Trf5 is associated with the TRAMP5 complex along with the Mtr4 helicase and zinc finger proteins Air1/Air2 (Houseley and Tollervey, 2006). The TRAMP complex is recruited to its target RNAs *via* the RNA-binding proteins Air1/Air2. Trf4/Trf5 is/are responsible for polyadenylating the substrate RNAs. The poly(A) tract allows recruitment and activation of the exosome, which degrades the RNA in a 3'-5' direction, facilitated by the helicase Mtr4 as depicted in Figure.1.3 (Houseley and Tollervey, 2006; LaCava *et al.*, 2005).

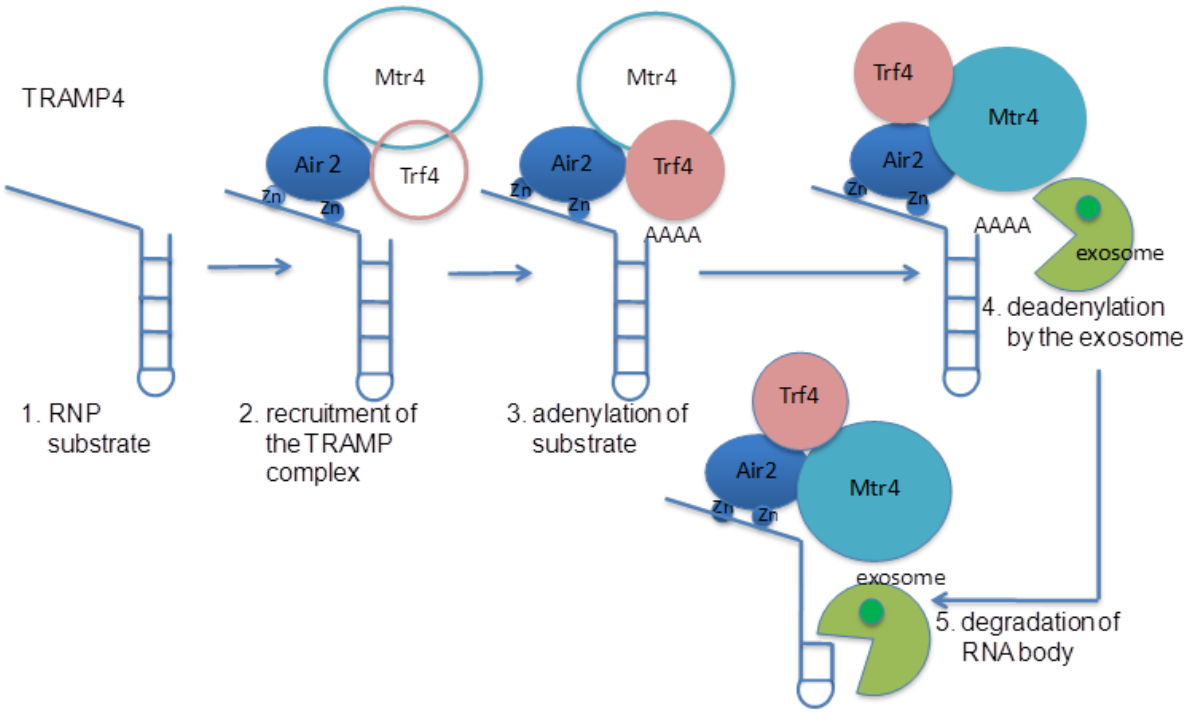


Figure 1.4 Schematic representation of the *S. cerevisiae* TRAMP4 complex. Components of the TRAMP4 complex include Trf4 non-canonical polymerase, Air2 zinc finger proteins and Mtr4 RNA helicase. The TRAMP4 complex is recruited to its target RNAs *via* the RNA binding protein Air2. The substrate RNA is polyadenylated by Trf4. Aided by the Mtr4 the exosome is recruited and activated by the poly(A) tail (derived from LaCava *et al.*, 2005). Adapted with permission from Elsevier.

The TRAMP complex has two main functions. First, the addition of poly(A) tails to transcripts gives a docking station for the exosome. Second, it makes the exosome function more efficiently independently of the catalytic activity of Trf4 and Trf5. Poly(A) tails of 1-7nt added by Trf5 have been observed in *vitro*. The poly(A) tails seen in *rrp6Δ* strain lacking Trf4 are longer. Hence the short A tail detected for Trf5 could be due to TAP-tag interfering with the reaction, non physiological conditions or the poly(A) tails could be added by another enzyme (Houseley and Tollervey, 2006). A recent study reports Trf4 to add 3-4 adenosine residues where the tail length is controlled by Mtr4 (Jia *et al.*, 2011).

Air2 shows 41% amino acid identity with Air1. These proteins contain five tandemly arranged zinc fingers of the CCHC type.

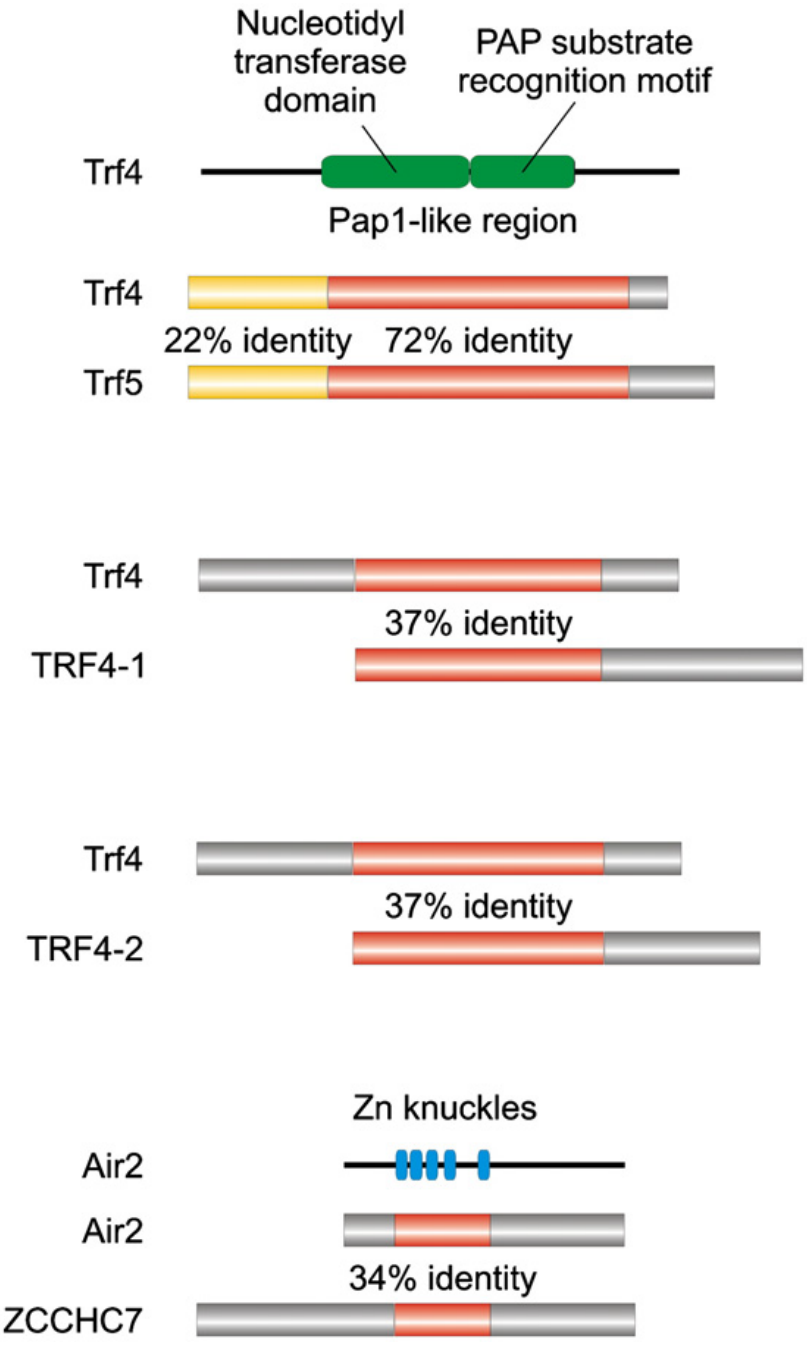


Figure 1.5 Structure and homology of TRAMP components in humans and budding yeast.

Red/orange represent homologous regions. Grey areas show no homology to other budding yeast proteins. Figure from Houseley and Tollervey, (2008). To be noted POLS and PAPD5 are also known as TRF4-1 and TRF4-2. ZCCHC7 is thought to be the human equivalent of Air2.

Yeast Mtr4 binds ATP and translocates along substrate, unwinding secondary structure or displacing bound proteins and/or RNA. Mtr4 localizes in the nucleoplasm and nucleolus. It is involved in degradation and maturation of many nuclear RNA substrates *via* the exosome. Polyadenylation is increased in strains lacking Mtr4; hence it is needed for polyadenylation-associated degradation. The Trf4/5-exosome pathway is conserved in the distantly related fission yeast *S. pombe.* Cid14 forms a complex with Mtr4 and Air1 [known as Cid14-Air1 complex (CAC)], similar to TRAMP complex found in *S.cerevisiae* (Keller *et al.*, 2010). Expression profiling data obtained using Affymetrix tiling arrays for *cid14Δ* and *air1Δ* were found to overlap only partially, suggesting that Air1 and Cid14 may have other independent functions.

The *T. brucei* Mtr4 (*Tb*MTR4) protein shares 41.5% identity with yeast Mtr4. TAP-tagged *Tb*MTR4 localized to the nucleus, possibly in the nucleolus, similar to the yeast situation (Cristodero and Clayton, 2007). Also in the absence of *Tb*MTR4, as with yeast, polyadenylated rRNA precursors accumulated. For instance, defects in 5.8S rRNA processing and growth arrest were caused by the depletion of *Tb*MTR4. Growth arrest was followed by death after the depletion of *Tb*Mtr4. Proteomic analysis of TbncPAP1 detected interactions with an Mtr4 RNA helicase orthologue and several RNA binding proteins, indicating the presence of a trypanosome TRAMP-like complex. Existence of a *T. brucei* TRAMP complex suggests the conservation of the TRAMP complex in eukaryotes, as does the presence of the CAC complex in *S.pombe*.

Lubas *et al.* (2011) report a human NEXT complex, which includes hMTR4, the zinc finger protein ZCCHC8 and the putative RNA-binding protein RBM7. This complex found in the nucleus, is associated with degradation of promoter–upstream transcripts (PROMPTs). The authors also identify putative homologues of TRAMP subunits PAPD5, hMTR4 and ZCCHC7 (hAIR2) in their SILAC/MS data. An exclusively nucleolar localization was reported for ZCCHC7.

1.2.2 Partial structure of the TRAMP complex

To date only a partial structure of the complex is available. Hamill *et al*. (2010) used Trf4/Air2 sub-complexes lacking proteolytically sensitive N and C termini from both proteins to avoid degradation problems. However both truncated forms were active in the polyadenylation assay and were able to recognize aberrant forms of tRNA. Of the five zinc fingers of Air2, the Trf4 central domain interacts with two. This interacting surface is highly conserved across eukaryotes suggesting that the function is conserved. Using a short RNA oligonucleotide (A_5) substrate, the first zinc finger was shown to enhance the polyadenylation activity towards tRNA-Ala. The second and third zinc fingers were dispensable for this polyadenylation activity. The second and third zinc fingers are probably involved in binding other RNA substrates, or might interact with Trf4 or Mtr4.

It is envisaged that mutations in the fourth and fifth zinc fingers would disrupt complex formation based on disabling RNA binding. N- and C-terminal portions of the Trf4/Air2 and one or more of the Air2 zinc fingers are involved in substrate binding. The zinc finger near the N terminus is thought to play a role in RNA binding and polyadenylation activity of Trf4, as other studies involving the full-length complex report greater activity compared to the mutant forms used in this study. The three most N-terminal zinc fingers of Air2 are reported to enhance the activity of Trf4 towards large RNA substrates.

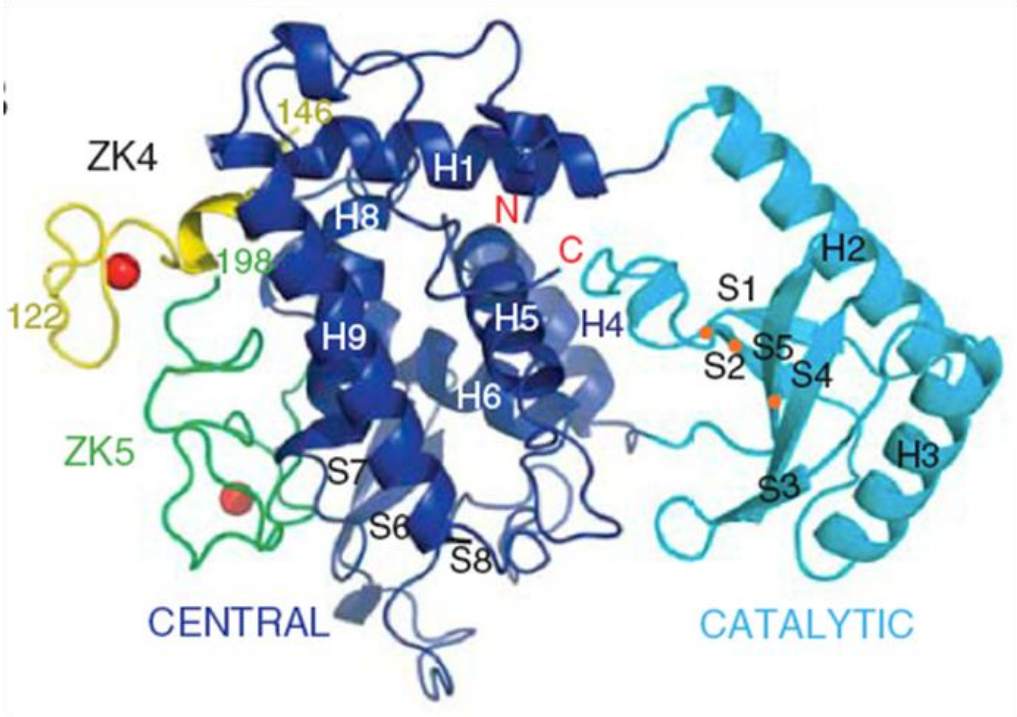


Figure 1.6 Structure of a Trf4/Air2 subcomplex
Ribbon diagram of Trf4, in yellow the catalytic domain and in blue the central domain. Orange dots represent positions of the aspartate residues in the catalytic triad. The fourth and fifth zinc knuckles of Air2 and linker regions are yellow and green respectively, and zincs are red. Figure from (Hamill *et al*., 2010).

Trf4 and Trf5 target many substrates. These include tRNA, small nucleolar RNA, small nuclear RNA, cryptic unstable transcripts, rRNA and histone mRNA. A detailed discussion of these substrates is to follow. Ribosomal RNA will be discussed in a special section as analysis of biochemical function of the human orthologues were performed on rRNA processing. Trf4 and Trf5 are also known to be associated with maintenance of telomere length, chromosome condensation and sister chromatid cohesion, phosphate metabolism and removal of spliced out introns. A brief discussion of these is to follow.

1.2.3 Trf4/5 targets and mutant phenotypes

i. tRNA

TRM6 methylates adenine 58 (A58) of tRNAi ^{Met}. A mutation in the gene *TRM6* stops the addition of a methyl group to A58 resulting in unstable $tRNA_i^{Met}$ (Anderson *et al.*, 1998). This disrupts the D-and T-loop interactions essential for $tRNA_i^{Met}$ to fold in its native L-shaped conformation, leading to tRNA degradation, although the exact mechanism of differentiating hypormodified tRNA from normal tRNA is still not known. Accumulation of hypomodified $tRNA_i^{Met}$ in *trf4* and *rrp44* mutants has been observed. The overexpression of Trf4 results in the accumulation of polyadenylated hypomodified $tRNA_i^{Met}$. Therefore, a tRNA surveillance pathway exists involving Trf4 and the exosome for polyadenylation and degradation of hypomodified $tRNA_i^{Met}$. Trf4 is also capable of polyadenylating newly transcribed precursor $tRNA_i^{Met}$. As a catalytic mutant is unable to stimulate the degradation process, continuous readenylation is required for this surveillance process (Kadaba *et al.*, 2004; Vanacova *et al.*, 2005).

ii Small nucelolar RNA (snoRNA) and small nuclear RNA (snRNA)

snRNAs are associated with pre-mRNA splicing and snoRNA with rRNA processing and modifications. Polyadenylated precursor snoRNA and snRNA have been recorded in strains lacking Rrp6 (Allmang *et al*., 2000; Mitchell and Tollervey, 2003). Studies have implicated Trf4 or Trf5 being responsible for this polyadenylation (Egecioglu *et al.*, 2006; Houseley *et al.*, 2006).

Other studies have shown that Pap1 is responsible for the poly(A) tail addition to snRNA and snoRNA in *rrp6Δ*. These transcripts are polyadenylated only because their normal processing is blocked. For instance, after Rnt1 cleavage, U4 RNA would normally be processed by Rrp6

and the exosome. However in the *rrp6Δ* strain, 3' end processing is inhibited, leading to polyadenylation by Pap1 (van Hoof *et al*., 2000).

Termination of snoRNA transcripts occurs at two sites; a Nrd1/Nab3 binding region (site I) and a sequence resembling the cleavage/polyadenylation signal (site II). Grzechnik and Kufel (2008) show that lack of Trf4 inhibits polyadenylation at one site and shifts termination to the other end as a consequence of the delayed 3' end processing. A Pap1 mutation has a more severe effect, with no poly(A) tail added at both terminators, resulting in no synthesis of precursor or mature snoRNAs. Hence the authors of this study speculate that Trf4/5 might initiate the process by adding short oligo(A) tails, which are further extended by Pap1. This is followed by recruitment of exonucleases including Rrp6/exosome. The Nab/Nrd major processing pathway for sn/snoRNA has no proven links with Pap1 to date. Although the exact mechanism is unknown, snoRNA transcriptional termination achieved *via* the Nrd1/Nab complex is lost in the absence of Trf4 or Mtr4. Hence, Trf4 and Mtr4 contributions are still of importance in snoRNA processing. Trf4 is involved in adenylating pre-snoRNAs to facilitate their processing *via* Nrd1 (Grzechnik and Kufel, 2008).

iii **C**ryptic **u**nstable **t**ranscript**s** (CUTs)

Transcribed from intergenic spacer and heterochromatic regions in yeast, CUTs are expressed at a very high level but are very short-lived (Wyers *et al.*, 2005). CUTs are easily detectable in exosome, TRAMP and Nrd1 mutants. These 200-800nt long non coding RNAs are capped and polyadenylated. Termination of CUTs depends on the Nrd1-Nab3 pathway. Nrd1 and Nab3 recruit the TRAMP complex to the CUT (Thiebaut *et al.*, 2006). Stimulation or targeting of the core exosome independent of its polyadenylation activity appears to be the role of Trf4 in CUT degradation. The *trf4-236* mutant, which lacks polyadenylation activity *in vitro*, accumulated fewer CUT transcripts that were not degraded compared to the *trf4* mutant.

Trf5 shows no significant effect on CUTs, however, when no Trf4 is present, Trf5 seems to play a role in CUT degradation (Wyers *et al.*, 2005). A polyadenylation-defective *trf4* mutant is able to recruit Trf5, which adds poly(A) tail to facilitate exosome-mediated degradation. This is supported by the finding that on deletion of *TRF5* in the *trf4-236* strain, the defect in CUT degradation was enhanced (Wyers *et al.*, 2005).

Heterochromatin-derived CUTs play a role in silencing and copy number control of rDNA in yeast, maintaining genome integrity (Vasiljeva *et al.*, 2008a; Vasiljeva *et al.*, 2008b). CUTs observed in Rrp6 mutants are degradation intermediates which are polyadenylated by Trf4. The longer poly(A)-containing species that accumulate in the absence of Trf4 and Rrp6 are thought to be polyadenylated by canonical PAP (Wyers *et al.*, 2005).

In human cells, transcripts similar to CUTs, Promoter upstream transcripts (PROMPTs), are found following the depletion of the nuclear exosome (Preker *et al.*, 2008). These are highly unstable. PROMPTs are bi-directionally transcribed from a 1.5-kb region upstream from the transcription start sites of active protein coding genes. They are controlled by the same promoter as the downstream mRNA. Non-coding RNAs could be involved in DNA methylation as PROMPTs are linked with promoters that contain many CpG dinucleotides.

iv Histone mRNA

In the *trf4-ts trf5Δ* mutant all four core histone mRNAs were found to be over-expressed in S phase, while deletion of *TRF4* alone had no such effect (Reis and Campbell, 2007). Histone over-expression was found to be toxic to *trf4Δ, trf5Δ and rrp6Δ* mutants. The reduction in histone levels rescued the synthetic lethality of *trf4Δ rad53Δ*, which is possibly explained by the synthetic genetic interactions observed between *TRF4* and genes that negatively regulate histone levels including *RAD53.*

Both *trf4Δ* and *trf4Δ-ts trf5Δ* exhibit a slow S-phase compared to wild type cells. It is thought unlikely that this is due to DNA damage, slow or aberrant DNA synthesis, as no DNA damage response was observed in *trf* mutants. Supporting this, no synthetic interactions were observed between *TRF4* and checkpoint kinases *MEC1* or *TEL1,* indicating Rad53 function to be specific. Loss of Rrp6 also was noted to lead to the same histone mRNA phenotype as *trf4Δ-ts trf5Δ*. Therefore, slow S-phase and related defects in chromosome maintenance in *trf* mutants in part at least are attributed to defects in regulation of histone mRNAs.

Synthetic interactions in *asf1⊿ trf4⊿* and *hir⊿ trf4⊿* strains were observed to increase histone mRNA levels. However, as Asf1 and the Hir1 complex have a vast spectrum of targets, the roles were assigned to transcriptional regulation or effects on the stability of histone mRNA. Inabilities to detect histone mRNA degradation intermediates also make it difficult assign histone degradation as a Trf4/5-related activity. Thus, global alterations in histone metabolism are thought to account for the over-expression of histones rather than reduced histone mRNA degradation (Reis and Campbell, 2007). The observed slow progression through S-phase of the *trf4-ts trf5Δ* mutant in response to HU treatment indicates the DNA damage checkpoint to be intact as, noted by Reis and Campbell.

Histone mRNAs are found in the cytoplasm while *S.cerevisiae* Trf4, Trf5 and Rrp6 are known to be localized in the nucleus making it difficult to envisage interactions. It was also reported that Air1 and Air2 were not required for the maintenance of normal histone mRNA levels. All other Trf4 and Trf5 targets are regulated *via* the TRAMP complex.

v Maintenance of telomere length

Telomere length varies from 100nt to 10kb. Disruption of Trf4 causes shortening of telomeres. Trf4 is thought to stimulate the degradation of the **t**elomerase **c**omponent **1**

(TLC1) in a polyadenylation-independent mechanism. TLC1 is part of a ribonucleoprotein required for telomere replication in a variety of organisms. *TRF4* deletion causes the accumulation of high levels of TLC1 which possibly interferes with the recruitment of telomerase at the chromosome ends and then leads to telomere shortening (San Paolo *et al.*, 2009).

vi Chromosome condensation and sister chromatid cohesion

Trf4 and Trf5 are known to affect mitotic events. Trf4 associates with both Smc1 involved in cohesion and Smc2, required for mitotic chromosome condensation. Smc1 mutants, and a *top1 trf4-ts* double mutant are defective in nuclear segregation at the restrictive temperature. *TRF4* and *TRF5* are reported to have important roles in sister chromatid cohesion (Wang *et al.*, 2000b). However, a study by the Nasmyth group concluded that *TRF4* deletion alone had no effect on sister chromatid cohesion in mitotic metaphase cells or on chromosome segregation during meiosis (Petronczki *et al.*, 2004). This observation was based on *TRF4* entirely being responsible for this activity. The possibility that *TRF5* could be compensating for *TRF4* was not addressed. Another possibility would be that chromosome condensation and sister chromatid cohesion are secondary effects detected due to Trf4 and Trf5 function on another substrate.

vii Phosphate metabolism

Genes involved in phosphate metabolism are regulated *via* Trf4/5 in an antisense RNA-mediated manner (San Paolo *et al.*, 2009).

viii Removal of spliced out introns

Removal of spliced out Introns is linked to Trf4 in a mechanism independent of its polyadenylation activity (San Paolo *et al.*, 2009).

Ribosomal RNA a target substrate of Trf4 and Trf5 will be discussed in detail as the analysis of the biochemical functions of the human orthologues was performed on this substrate.

1.2.4 Ribosomal RNA

A subset of pre-rRNAs are polyadenylated by TRAMP, processed or degraded by the exosome in the nucleus.

1.2.4.1 Ribosomal RNA degradation and processing

Ribosomes contain highly conserved rRNA species produced by a complex assembly pathway. Human acrocentric chromosomes 13, 14, 15, 21, and 22 contain 200-400 copies of rRNA genes. The number of copies of rDNA can vary according to cell type. The level of rRNA transcription is determined by initiation and not the number of transcribed genes in exponentially growing yeast. Active rRNA genes are replicated in early and the silent ones in late S phase (Lafontaine and Tollervey, 2001).

Structural rearrangement and nucleotide modifications followed by cleavage, give rise to mature ribosomal subunits. Pre-rRNA is separated by two internal transcribed spacers, ITS1 and ITS2 which contain regulatory elements such as gene promoters, spacer promoters, repetitive enhancer elements, and transcription terminators. ITS1 and ITS2 are flanked by two external transcribed spacers, the 5’ ETS and the 3’ ETS (Hadjiolova *et al.*, 1993).

Pre-rRNAs vary in size between species due to transcribed spacer differences and variable segments within the mature ribosomal RNAs. The yeast 35S nucleolar RNA (nRNA) is one of

the shortest while the human 47S nRNA transcript is one of the longest. The 13.3 kb human primary transcript is flanked by 30 kb of intergenic spacer (IGS). The core sequence elements within the mature RNA are highly conserved. Approximately half of the precursor is lost during ribosome biogenesis. The primary transcript, transcribed by RNA polymerase I is processed through a series of endo- and exonucleoytic processing steps and modifications to 18S, 5.8S and 28S rRNA. The 5S rRNA is synthesized by RNA Polymerase III.

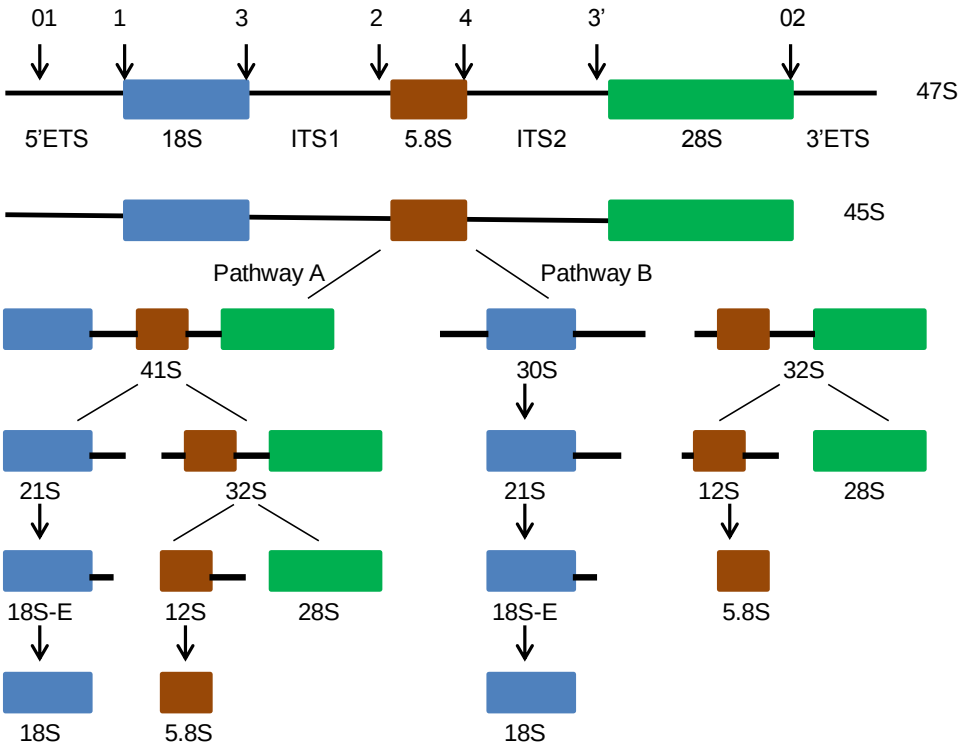


Figure 1.7 Processing steps of the 47S pre-rRNA transcript in Hela cells.

The regions annotated bellow are the rRNA sequences 18S, 5.8S and 28S flanked by two external transcribed spacers; 5'ETS, 3'ETS and separated by two internal transcribed spacers; ITS1, ITS2. Annotated above 0-4 indicate positions of the major processing sites. The main difference between pathway A and B depends on the temporal order the endonuclease cuts at sites 01 and 02. Schematic and nomenclature according to Hadjiolova *et al*., 1993.

The *S.cerevisiae* processing pathway is the most studied and best established. In addition endo- and exonucleolytic cleavages that remove the transcribed spacers vary according to species, cell type or physiological state (Rouquette *et al.*, 2005). Thus, comparisons need to be made with caution. In yeast the primary pre-rRNA transcript is cleaved by Rnt1 to produce

the 35S precursor. The next steps in processing follow a strong 5'-3' bias in order of cleavage. The A' cleavage was thought to be the primary cleavage site in mammals which requires U3 snoRNA. Down-regulation of the 5' exonuclease Xrn2 by Wang and Pestov (2011) has highlighted processing events undetected before. Cleavage events can occur prior to A' cleavage, with products rapidly destroyed by Xrn2 in mouse 3T3 cells. This perhaps acts as a checkpoint in pre-rRNA maturation to ensure docking of U3 snoRNA to the formation of stable 5' ends protected from exonuleolytic degradation. Hence, Xrn2 has acquired a proof-reading role determining if the pre-rRNA intermediates are to be processed or degraded. The final stages of yeast 5.8S rRNA processing occur in the cytoplasm (Thomson and Tollervey, 2010).

1.2.4.2 Polyadenylation of ribosomal RNA

Ribosomal RNA were not thought to be polyadenylated as they lack a canonical poly(A) signal and are not transcribed by Pol II. Fodor *et al.*, (2000) have shown that synthesis of RNA transcripts from pol II is not an absolute requirement for 3' end processing. RNA transcripts derived from a modified viral RNA (vRNA) template by the influenza virus RNA polymerase can be cleaved and polyadenylated by the cellular 3' end processing machinery *in vivo*. An influenza vRNA template coding for the neuraminidase protein was engineered by replacing the viral poly(A) site, the U6 sequence, with a negative-sense eukaryotic poly(A) signal derived from the rabbit β-globin gene to demonstrate this process.

A small proportion of polyadenylated ribosomal RNA transcripts have been detected. Fang *et al.*, 2004 reported that in yeast <0.1% of 25S related RNA is polyadenylated in the wild type strain and this population increases 100-fold in an *rrp6Δ* strain. The relative level of polyadenylated ribosomal RNA is only about 1/20th of poly(A)$^{+}$ mRNA. Pre-ribosomal RNAs are polyadenylated by TRAMP and degraded by the exosome in the nucleus. Absence of Trf4 suppressed polyadenylation and stabilized the pre-rRNA (27S) and rRNA (25S).

Hyperadenylation and stabilization of RNA was noted in the absence of Rrp6 (Dez *et al.*, 2006). The temperature sensitive mutant of *pap1-1* does not add a poly(A) tail to mRNA at the nonpermissive temperature. Polyadenylated rRNA is not detected at permissive temperature but is visible in an *rrp6Δ* strain. A similar result was noted for Pap1 by poly(dT) selection and northern blotting. Fang *et al.* (2004) assigned Pap1 to be responsible for polyadenylating rRNA. There are no records of Pap1 mutations stabilizing the transcripts in a wild type or *rrp6* mutant background. Polyadenylated rRNA does not accumulate in an *rrp6-13* strain. Hence, Rrp6p may play a role in removing the poly(A) tail or degrading RNAs that carry a poly(A) tail. While Trf4-dependent polyadenylation leads to degradation of the transcript the role of Pap1 polyadenylation of rRNAs is unclear.

The Schuster group using molecular and bioinformatics approaches have shown that human 28S and 18S rRNAs contain adenosine tails of homo- or heteropolymeric nature (Slomovic *et al.*, 2006). The heteropolymeric tails detected are thought to resemble the tails added post-transcriptionally by bacterial PNPase and the hyperthermophilic archaeal exosome complex, stimulating degradation. The homopolymeric tails observed are believed to be produced in the nucleus while the heteroploymeric tails in the cytoplasm. Another possibility is that the heteropolymeric tails are added by the human orthologues of *TRF4* and *TRF5* and the homopolymeric tails by the canonical PAP, however this is thought unlikely as the tails detected are dispersed along the rRNA.

1.2.4.3 Mutant phenotypes detected in relation to ribosomal RNA processing and degradation

In *air1*Δ, *air2*Δ, *trf4*Δ and *trf5*Δ strains 23S pre-rRNA accumulation was observed. A *trf4-236 trf5Δ* strain showed the most amount of accumulation of 23S pre-rRNA indicating that both Trf4 and Trf5 are involved in pre-rRNA degradation. RNA maturation defects observed in *trf4* are exacerbated by the depletion of Trf5. Polyadenylated RNA accumulated to higher levels in strains lacking Mtr4 or Rrp6 in comparison to strains lacking Rrp41 or Rrp44 core

exosome components, suggesting that Rrp6 is the main deadenylase (Dez *et al.*, 2006). Overexpression of Trf5 in an *rrp6Δ* strain resulted in greatly increased polyadenylation of 5.8S and 30S pre-rRNA (Houseley and Tollervey, 2006). Strains with defects in the exosome or Mtr4 showed an accumulation of 5' ETS and maturation defects in 7S pre-rRNA which were not observed in *trf4Δ* or *trf5Δ* strains. Depletion of Mtr4 results in the accumulation of polyadenylated pre-rRNA. Hence, Mtr4 links polyadenylation by Trf4 and Trf5 to exosome-mediated degradation.

1.2.4.4 Degradation of mature ribosomal RNA

Ribosomal RNAs containing individual point mutations in the 25S peptidyl transferase centre or 18S decoding site are down-regulated in *S.cerevisiae.* This down-regulation is due to decreased stability of the mature rRNAs and ribosomal subunits. Hence, there is a quality control mechanism in place to detect and eliminate aberrant rRNA components of mature ribosomes. This non-functional rRNA decay (NRD) pathway, occurs in the cytoplasm (LaRiviere *et al.*, 2006). The mechanism of this decay pathway is undefined to date.

The anti-cancer drug 5-FU is known to have many target substrates. Increased accumulation of polyadenylated transcripts have been documented in exosome and *cid14* mutants treated with 5-FU.

1.2.4.5 5-FU

5-FU is a widely used chemotherapeutic drug for the treatment of solid colorectal, breast, digestive tract and head/neck cancers. 5-FU is an analogue of uracil with a fluorine atom at the c-5 position in place of a hydrogen atom (Figure 1.8).

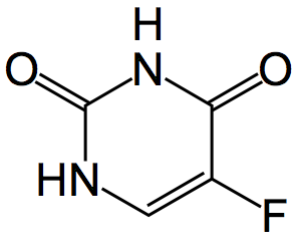


Figure 1.8 Chemical structure of 5-FU.
An analogue of uracil with a fluorine atom at the c-5 position.

1.2.4.6 Possible mechanisms of 5-FU

A vast amount of research performed over decades has assigned many mechanisms of action for 5-FU; the principal mechanism is controversial. DNA or RNA could be the target. RNA targets include rRNA, tRNA, snRNA and mRNA. The evidence for these target substrates will be detailed bellow.

5-FU effects on DNA

The drug was originally considered to act as an inhibitor of DNA synthesis by interfering with the activity of thymidylate synthetase (TS). 5-FU follows the same metabolic pathway as uracil. Inhibition of TS by the 5-FU metabolite, FdUMP results in loss of dTMP and dTTP, resulting in cell cycle arrest in S-phase. TS inhibition also leads to the accumulation of dUMP resulting in an increase in dUTP which could be misincorporated into DNA. Deoxyuridine incorporation results in DNA damage, replication fork collapse and cell cycle arrest. A DNA repair mutation *apn1Δ* in budding yeast increases sensitivity to 5-FU-induced DNA damage, as noted by a genetic screen (Hoskins and Butler, 2008).

5-FU effects on RNA

The cyotoxic and apoptotic effects of 5-FU are relieved by uridine but not thymidine, which could be interpreted as evidence that RNA is the target (Fang *et al.*, 2004). Inhibition of pre-rRNA processing, alterations in mRNA expression, inhibition of mRNA splicing and interference with tRNA modifications have been reported as a result of 5-FU incorporation into RNA (Ghoshal and Jacob, 1994).

5-FU effects on rRNA

5-FU addition has no effect on 47S precursor synthesis. However the formation of 28S and 18S is inhibited by the drug. Increase in steady state levels of 47S rRNA indicates unaltered Pol I transcription. 5-FU is thought to inhibit cleavage of pre-rRNA at the primary site (Wilkinson and Pitot, 1973).

Incorporation of 5-FU into mouse sarcoma cells resulted in decreased synthesis of 18S and 28S RNA, consistent with the inhibition of rRNA maturation by 5-FU (Takimoto *et al.*, 1987) though not proving that inhibition of rRNA processing by 5-FU is responsible for cytotoxicity in these cells. The 5-FU concentration used by Takimoto *et al.* (1987) did not affect cell growth. Perhaps 5-FU has a transient effect on rRNA maturation (Parker and Cheng, 1990). Wild-type cells exposed to 5-FU accumulate polyadenylated noncoding RNA in an exosome-dependent manner. 5-FU treatment resulted in accumulation of polyadenylated fragments of 27S rRNA and a defective Rrp6 enhanced this effect in yeast (Hoskins and Butler, 2008). The defective ribosomal RNA degradation observed in *S.cerevisiae rrp6* mutants and *S.pombe cid14* mutants are enhanced by the addition of 5-FU. An accumulation of increased polyadenylated rRNA intermediates was detected in both instances. Therefore, a general pathway for degrading rRNA precursors may be inhibited by 5-FU (Wang *et al*., 2008).

5-FU treatment of heterozygous deletion strains of *RRP6, RRP44, RRP41* and *RRP46* produced 2- to 10-fold more A2-C2 rRNA intermediate product than seen in a isogenic wildtype strain treated with 5-FU. This suggests that 5-FU inhibits cell growth *via* a mechanism that disrupts rRNA processing by the exosome complex. In these 5-FU treated cells, G2/M arrest was recorded. Of all the exosome mutants tested, the greatest sensitivity to 5-FU was detected in *rrp6*. Microarray and northern analysis have shown that 5-FU treatment causes an accumulation of rRNA in wild type yeast which is enhanced in *rrp6* mutants. *RRP6* was also identified as a potential target of 5-FU in a genome-wide screen.

Down-regulation of the human homologue of Rrp6, PM-Scl100, in Hela cells greatly impaired growth of cells treated with 5-FU compared to control cells. Selective stabilization of PM-Scl100 RNA substrates indicates that 5-FU directly targets PM-Scl100 and not core exosome components. The specificity of targeting Rrp6 is limited to some tumors as differential sensitivities to 5-FU were recorded for different cell lines (Kammler *et al.*, 2008). However, in this study drug sensitivity was only recorded at two time points. To make a solid claim of Rrp6 being the target the screening should have been extended to include at least one more time point. It is possible that the screen was stopped at two time points as no down-regulation of PM-Scl100 was detected after 3 days as siRNA duplexes are known to be effective only for a short time. If this was the case a plasmid construct could have been employed. Plasmid constructs are known to be effective in down-regulation of genes for longer periods.

In 5-FU treated *Drosophila* S2 cells, by gel filtration chromatography, a change in relative molecular mass of the complex containing Rrp6 and the core exosome was detected. This implies alterations in the complex due to 5-FU incorporation. This results in aberrant transcripts synthesized in 5-FU treated cells are not turned over efficiently (Silverstein *et al*., 2011).

5-FU effects on tRNA

Incorporation of 5-FU into tRNA was noted to contribute to defective protein synthesis (Dolnick and Pink, 1985).

5-FU effects on mRNA

In 1982, Glazer and Lloyd reported the cytotoxic effects of 5-FU were due to its incorporation into nuclear RNA. They observed an increase in dihydrofolate reductase (DHFR) mRNA by three-fold compared to total cellular RNA when cells were treated with 5-FU. This enhancement was observed in the 3.5 kb mRNA but not in the 1.0 kb mRNA of DHFR. Another study by (Armstrong *et al.*, 1986) reported a three-fold enhancement of in DHFR mRNA in the nucleus but no increase in the cytoplasm of 5-FU treated cells. This is perhaps due to a reduction in the amount of DHFR pre-mRNA processed into mature mRNA. It is also possible that 5-FU decreases the degradation of nuclear DHFR mRNA. A number of studies have documented a change in the ratio of pre-mRNA to mature mRNA in cells exposed to 5-FU, suggesting that intron splicing has been affected by 5-FU.

5-FU effects on snRNA

Incorporation of 5-FU into U2 snRNA greatly reduced the splicing activity but incorporation of 5-FU into U6 snRNA functioned less efficiently though it was catalytically active. This may be due to structural differences or role played by the different snRNA in splicing (Lenz *et al.*, 1994). U2 snRNA base pairs with the pre-mRNA at the branch point. U4/U6/U5 form a tri-snRNP complex before associating with the splicing complex. Armstrong *et al.* (1986) have shown that in nondenaturing gel electrophoresis 5-FU alters the migration of U4 and U6 snRNAs. This is possibly reflective of secondary structure alterations in the snRNAs due to the incorporation of 5-FU (Doong and Dolnick, 1988).

The above gives an insight to the many pathways by which 5-FU may be functional. It is possible that 5-FU has many targets. Understanding the action of 5-FU could increase its efficacy and usage. It is possible that 5-FU may affect component/s of the TRAMP complex.

To date some Trf4/Trf5 orthologues and their functions have been elucidated. A brief summary of these are discussed in the next section. These include *S.pombe* Cid14, *C.elegans GLD-4, D. melanogaster DmTRF4-1* and *DmTRF4-2 and T. brucei TbncPAP1 and TbncPAP2.*

1.2.5 Orthologues of Trf4/Trf5

S.pombe Cid14

Cid14 is localized in the nucleolus and is required for maintaining nucleolar structure. A minor population of Cid14-dependent, fully processed and intermediate polyadenylated rRNA transcripts accumulated in a *dis3-54* exosome mutant. This indicates that Cid14 is involved in rRNA degradation. Cells lacking Cid14 are viable although they are prone to elevated levels of chromosome missegregation; they are also hypersensitive to 5-FU (Win *et al.*, 2006).

A silencing pathway within the centromeric DNA repeats requiring Cid14 has been described by the Moazed laboratory (Buhler *et al.*, 2007). Double stranded RNA cleavaged by Dicer generated short interfering RNAs (siRNA) which mediate post-transcriptional gene silencing (PTGS) or RNA induced transcriptional silencing (RITS) in a TRAMP and RNA-dependent RNA polymerase (RDRP) dependent manner. Deletion of *cid14*$^{+}$ led to the loss of *ura4*$^{+}$ silencing at the silent mating type locus as assessed by the growth on 5-fluoro-orotic acid (5-FOA)-containing medium. The degradation of transcripts arising from the centromeric siRNA was dramatically reduced in *cid14Δ* cells. The structure of heterochromatin was preserved in *cid14Δ* cells as H3K9 methylation and Swi6 localization was unaffected. Deletion of *cid14*$^{+}$

was found to lead to a dramatic decrease in siRNA levels. Transgene silencing at centromeric repeats and silent mating type loci required Cid14. Insertion of a *ura4*$^{+}$ gene into centromeric heterochromatin led to the generation of siRNA. These siRNAs mediated silencing by direct slicing of *ura4*$^{+}$ mRNA by Ago1. Centromeric repeat silencing requires RNAi but silencing at the silent mating loci does not. Therefore Cid14 is involved in efficient heterochromatin gene silencing *via* the exosome in the presence or absence of the RNAi machinery. Polyadenylation by Cid14 is disrupted by mutations in the catalytic core and disrupts the silencing function *in vivo.* This gene silencing pathway triggered by double-stranded RNA, is not conserved in *S. cerevisiae* but conserved in a close relative *Saccharomyces castellii* (Drinnenberg *et al.*, 2009). It will be of interest to see if this silencing pathway is conserved in humans.

C.elegans GLD-4

GLD-4 appears to be the *C.elegans* orthologue of Trf4/5, with very similar nucleotide transferase (NTP) domain and poly(A) polymerase (PAP)-associated domain. GLD-4 is found in cytoplasmic P granules. GLS-1 a protein associated with the RNA binding protein GLD-3, enhances the enzymatic activity of GLD-4. Predominantly expressed in germ cells, GLD-4 activity is essential in the absence of GLD-2 for early meiotic progression (Schmid *et al., 2009)*.

D. melanogaster DmTRF4-1* and *DmTRF4-2

DmTRF4-1 and DmTRF4-*2*, two *Drosophila melanogaster* orthologues of Trf4/5, have been identified. DmTRF4-1 displays poly(A) polymerase activity *in vitro* but DmTRF4-2 does not. Disruption of DmTRF4-1 results in a lethal phenotype, while knockdown of DmTRF4-2 does not affect the viability of the flies. DmTRF4-1, which localizes to the nucleolus, causes an increase in poly(A) tail lengths of snRNAs when over-expressed (Nakamura *et al.*, 2008).

T. brucei TbncPAP1 and TbncPAP2

Two *T. brucei* Trf4/5 orthologues, TbncPAP1 and TbncPAP2, implicate the presence of nuclear PAPs involved in targeting aberrant RNAs for exosome mediated degradation in protozoans (Etheridge *et al*., 2009).

1.3 A Comparison of canonical and non-canonical poly(A) polymerases

Trf4 and Trf5 belong to a family of non-canonical poly(A) polymerases conserved from yeast to humans. These proteins belong to the ribonucleotidyl transferases (rNTrs) within the DNA polymerase β (Pol β) superfamily. The well known canonical poly(A) polymerase first identified in 1960 from extracts of calf thymus also belongs to this same family. A characteristic feature of ribonucleotidyl transferases is the addition of nucleotides to the 3' end of an RNA in a template-independent manner.

Structural features

Canonical and non-canonical PAPs contain a nucleotide transferase (NTP) domain and a poly(A) polymerase (PAP)-associated domain. All three canonical PAPs (α, β and γ) contain an RNA-recognition motif (RRM). The RRM is found near the C-terminus in canonical PAPs, contrasting with the situation in non-canonical PAPs where, when present, the RRM domain is found located near N-terminus of the protein. In eukaryotic canonical PAPs, the active site is located between the NTP and RRM domains. A glycine-serine (GS) motif and a conserved aspartic acid triad form the catalytic site. The incoming ATP is embedded in the central domain and interacts with the 3' end of the single-stranded RNA primer (Martin and Keller, 1996). Data relating to structural features of non-canonical PAP's are beginning to appear. A partial structure of the TRAMP complex has been reported. The *S.pombe* Cid1 crystal structure (Luke Yates, Robert Gilbert and Chris Norbury, unpublished) has revealed that

basic amino acid residues on the surface, confer RNA-binding capacity to Cid1 (Sophie Fleurdepine and Chris Norbury, unpublished). Some non-canonical polymerases contain zinc fingers. Zinc fingers are thought to aid RNA binding based on mutational and structural studies (Amarasinghe *et al.*, 2000). It is envisaged that different regions will be of importance to accommodate the diverse biological functions of non-canonical polymerases**.**

Functional features

The canonical PAPs, PAP-α and PAP-γ , add poly(A) tails during the processing of the 3' ends of mRNA precursors in the nucleus. PAPβ is involved in polyadenylation of testis-specific mRNA (Kashiwabara *et al.*, 2000). Polyadenylation by canonical PAP is fast and processive. The long (approximately 200nt in humans) poly(A) tails added by PAP stabilize mRNA by binding to poly(A) binding proteins (PABPs) protecting the 3' end from deadenylation and subsequent degradation by the exosome.

Non-canonical poly(A) polymerases are found localized to the nucleus, cytoplasm and mitochondrion, indicating diverse biological functions. RNA turnover, DNA integrity checkpoint responses and RNAi-dependent heterochromatin formation are roles assigned to these polymerases. The mechanisms by which these functions are achieved are not well defined. Adenosine and/or uridine residue addition to the 3' end of various substrates seem to be an important process in reaching these objectives.

The conservation of the catalytic mechanism among nucleotidyltransferases is believed to impose rigid constraints on the evolution of the catalytic domain, making the more divergent central domain a likely candidate to determine substrate specificity. Another possibility is that there are other sequences possibly in the 3' UTR that determine substrate specificity.

The *T. brucei* RET2, an RNA editing TUTase, is a member of the DNA polymerase β superfamily that binds UTP. The RET2 catalytic domain structure is very similar to yeast nuclear PAP, however the overall shape differs. Purines are excluded from the nucleotide binding site while a bound water molecule is positioned to interact with UTP (Aphasizhev, 2005). This divergence could be attributed to functional specificity of RET1 and related TUTases (RET2 and TUT 3-7).

1.3.1 The family of non-canonical poly(A) polymerases

The *S.pombe Cid1* (caffeine-induced death suppressor 1) was identified in 2000 forming a novel class of non-canonical poly(A) polymerases (Wang *et al.*, 2000a). Cid1 was identified in a screen as conferring resistance to the combination of hydroxyurea (HU) and caffeine. HU is a ribonucleotide reductase inhibitor and caffeine inhibits the cell cycle checkpoint signaling kinase Rad3 (Wang *et al.*, 1999; Wang *et al.*, 2000a).

Non-canonical poly(A) polymerases found in different species are described below;

S. pombe

The *S. pombe* genome encodes six non-canonical poly(A) polymerases. **Cid1** is a cytoplasmic polymerase. When expressed in bacteria it was shown to have both poly(A) and poly(U) activity *in vitro* (Read *et al.*, 2002; Rissland *et al.*, 2007). However, Cid1 purified from *S. pombe* possesses strong poly(U) polymerase activity (Rissland *et al.,* 2007). Cid1 has been shown to target mRNA for degradation by the addition of short U-tails in the cytoplasm (Rissland and Norbury 2009). Not much is known to date about **Cid11** apart from its nuclear and cytoplasmic localization, and meiosis specific expression. **Cid12**, is a nuclear protein recruited by the RITS complex during RNAi- dependent heterochromatin formation, gene silencing and faithful chromosome segregation. **Cid13**, when over expressed was able to partially override the hydroxyurea sensitivity of *rad*3 mutants but was not able to suppress the replication checkpoint defect of these mutants. Cid13, found in the cytoplasm, was

thought to maintain normal levels of dNTPs by stabilizing *suc22* mRNA, which encodes a subunit of ribonucleotide reductase (Saitoh *et al.*, 2002). Contrastingly, another study demonstrated that total Suc22 protein levels are not altered by the function of *cid13* either in the presence or absence of HU. It is possible that other targets of Cid13 are responsible for dNTP pool regulation. Perhaps Cid13-mediated polyadenylation in the cytoplasm controls *suc22* mRNA translation that is not reflected at Suc22 protein level. This study concluded that HU sensitivity of *cid13Δ* was not due to the failure to up-regulate Suc22 protein levels (Read *et al.*, 2002). **Cid14** is a functional homologue of *S.cerevisiae* Trf4/5. **Cid16** may be a mitochondrial PAP on the basis of the localization of a yellow fluorescent protein (YFP)-tagged protein (Wood *et al.*, 2002).

***Caenorhabditis elegans* germ*l*ine *d*evelopment defective-2 (*GLD-2*)**

GLD-2 is a cytoplasmic poly(A) polymerase required to complete both spermatogenesis and oogenesis. On its own GLD-2 shows very little poly(A) polymerase activity. The RNA binding protein GLD-3 assists GLD-2 to bind specific mRNAs stimulating the polymerase activity of GLD-2. ***GLD-3*** is associated with the siRNA gene silencing pathway. It is needed for fertility and viability at high temperatures. (Chen *et al.*, 2005). ***GLD-4*** appears to be the *C.elegans* orthologue of Trf4/5.

Xenopus GLD-2

This polymerase was first identified as a regulator of poly(A) tail length and consequently a determinant of translational efficiency (Kwak *et al.*, 2004). The presence of poly(A) nuclease PARN in the CPEB/CPSF complex along with symplekin and *GLD-2* ensures the addition of a short poly(A) tail. Upon maturation of oocytes, PARN is released and the poly(A) tail is elongated by *GLD-2*. The RNA binding proteins maskin, Pumilio and Nanos are capable of

suppressing this activity. Although XGLD-2 is found both in the nucleus and cytoplasm the role played in the nucleus is unclear.

It is unlikely that the nuclear enzyme is required for cytoplasmic polyadenylation events during oocyte maturation, as the maturation process was shown to be unaffected in enucleated oocytes (Fox *et al.*, 1989).

Mus musculus GLD-2

Murine *GLD-2* is a cytoplasmic polymerase. During oogenesis translation of specific dormant mRNAs is induced by the addition of poly(A) tails by murine *GLD-2* which acts as a regulator of metaphase progression (Kashiwabara *et al.*, 2008). The catalytic and central domain sequences are highly similar to *Xenopus* and human *GLD-2* (74% and 94%). Gld-2-deficient mice are normal in behavior and health. *GLD-2* is not essential for polyadenylation, perhaps due to another PAP. In murine liver cells *GLD-2* adds a single adenosine residue which stabilizes mature miR-122 (Katoh *et al.*, 2009).

Drosophila GLD-2 (Wisp)

In late stages of oogenesis and early stages of embryogenesis Wisp is required for poly(A) tail lengthening of mRNA. Wisp interacts with Orb which is a homologue of CPEB (Benoit *et al.*, 2008). *Drosophila* BicC, a homologue of *C. elegans GLD-3,* is found in an RNP complex with Wisp. BicC has been reported to elongate poly(A) tails during oogenesis as well as recruit the deadenylase complex CCR4-NOT (Chicoine *et al.*, 2007). **DmTrf4-1** and **DmTrf4-2** orthologues of Trf4 and Trf5 are discussed in section 1.2.4.7.

1.3.2 The human non-canonical poly(A) polymerases

The human genome encodes seven non-canonical poly(A) polymerases. Data are only just beginning to emerge regarding the biological functions of these genes.

PAPD4 (Hs1, HsGLD-2, TUTase 2)

PAPD4 is a cytoplasmic poly(A) polymerase regulating poly A tail length of neuronal mRNA at synapses (Rouhana *et al*., 2005). Katoh *et al*. (2009) demonstrate the achievement of cellular stability of miR-122 *via* 3' adenylation by *PAPD4*.

ZCCHC6 (Hs2, PAPD6, cid1, TUTase 7)

In vitro studies involving *Xenopus* oocytes and from human cells have shown this enzyme to be a poly(U) polymerase (Kwak *et al.*, 2004; Rissland *et al.*, 2007). It has been identified as a possible human orthologue of *S.pombe* Cid1. *ZCCHC6* contains three zinc finger motifs. Arsenite treatment of human cancer cells localizes this protein in stress granules pointing to a role in stress-induced RNA metabolism (Andrea Mikulasova and C.J.N., unpublished data). In this study two splice variants coding for proteins of 1498 and 1541 were identified (Figures 6.3 and 6.4).

ZCCHC11 (Hs3, PAPD3, TUTase 4)

ZCCHC11 has been shown to have poly(U) polymerase activity (Kwak *et al.*, 2004). It is highly expressed in the thymus, spleen, testes and lung. A recent study has shown *ZCCHC11* to be a terminal U-transferase targeting human histone mRNAs for degradation following inhibition or completion of DNA replication (Schmidt *et al.*, 2011).

Two other independent studies have shown *ZCCHC11* to regulate miRNA biogenesis. Narry Kim's laboratory has shown Lin28 a pluripotency factor to interact with pre-let-7 (*let-7* is a tumour suppressor miRNA) and recruits *ZCCHC11*, which adds an oligo uridine tail of 10-30nt to the 3' terminus of the pre-miRNA, blocking Dicer processing and leading to the down-regulation of *let-7* miRNA (Heo *et al.*, 2009).

(Jones *et al.*, 2009) showed that *ZCCHC11* regulates cytokine expression by uridylating cytokine-targeting miRNA. The *miR-26* family of miRNAs target IL-6. In control cells 78% of the miRNA26a sequences were found to contain 1-3 uridines but when *ZCCHC11* is disrupted the proportion of 1-3 uridines found falls to less than 0.1%. Thus, *ZCCHC11* is involved in regulating gene expression through uridylation of miRNAs. It was also observed that cytokine mRNAs needing *ZCCHC11* for stability and maximal protein expression are regulated by the poly(A) tail length at a post-transcriptional level. Down-regulation of *ZCCHC11* results in short poly(A) tails on these cytokine mRNAs. Hence, *ZCCHC11* directly or indirectly regulates cytokine expression *via* regulation of poly(A) tail length. *ZCCHC11* also has been shown to be a negative regulator of Toll-like receptor (TLR) signaling pathway. TLRs are responsible for sensing microbial pathogens in immune responses (Minoda *et al*., 2006).

***PAPD1* (*Hs4, mtPAP, TUTase 1*)**

PAPD1 localizes constitutively to mitochondria and is the only human enzyme in this family to contain a mitochondrial targeting sequence at the N terminus. It adds a poly(A) tail to mitochondrial mRNAs. (Nagaike *et al.*, 2005). Mutagenesis and biochemical studies show that dimerization of PAPD1 is essential for its catalytic activity although the purpose of dimerization is currently unclear. PAPD1 has a high substrate affinity for ATP and UTP. However this enzyme is reported to be able to use GTP and CTP as substrates (Bai *et al.*, 2010). The poly(A) tails added by PAPD1 stabilize human mitochondrial mRNA while in other

organisms mitochondrial mRNA decay is initiated by polyadenylation. Some mitochondrial mRNAs lack a complete stop codon but terminate in uridine. Hence polyadenylation generates a UAA stop codon. Inhibition of PAPD1 leads to decrease in the level of polyadenylated fraction, while the oligoadenylated fraction is reported to increase accordingly.

PAPD1 was suggested as an enzyme that adds uridine residues to the 3' end of histone mRNAs initiating histone mRNA degradation in the cytoplasm (Mullen and Marzluff, 2008). A recent study found *PAPD1* not to be involved in histone mRNA degradation (Schmidt *et al.*, 2011). These contradicting results require further investigation, although it is difficult to envisage a role for a mitochondrial poly(A) polymerase in terminal U-transferase activity in the cytoplasm.

***PAPD2* (U6-tutase, *TUT1, Hs5,* star *PAP*)**

U6-tutase was identified by its terminal uridylyl transferase activity, with a very high specificity to its substrate U6 snRNA, being the only substrate accepted from total cellular RNA (Trippe *et al.,* 2006). A component of the spliceosome, U6 snRNA is known to play a major role in pre-mRNA splicing. The addition of up to 3 uridine residues to the 3' end during U6 snRNA splicing by U6-tutase is a vital process in snRNA maturation. RNAi knockdown experiments have identified U6-tutase as an essential protein for cell proliferation which localizes in nucleoli with a transient expression in Hela cells.

The amino acid sequence of U6-tutase was found to be highly divergent from poly(A) polymerases and from trypanosome TUTases. U6-tutase has a split PAP domain linked by a proline-rich sequence, a C2H2 zinc finger at the amino terminus, a PAP-associated domain, an R/S repeat, a nuclear localization signal and an RRM (Mellman *et al.*, 2008). *PAPD2* is the only human non-canonical polymerase to contain a RRM. It is found near the N-terminus

of the protein. Perhaps the C2H2 zinc finger motif found in PAPD2 aids RNA binding as the case with the distantly related trypanosomal RNA editing TUTase, RET-1.

Another study (Mellman *et al.*, 2008) assigned the localization of this protein to nuclear speckles and named it Star-PAP. This polymerase interacts with phosphatidylinositol 4-phosphate 5-kinases (PIPKIs). The product of PIPKIα, PtdIns4,5P2 stimulates PAP activity in vitro. Star-PAP was demonstrated to have both poly(A) polymerase as well as poly(U) polymerase activity. In normal physiological conditions the ATP concentration is greater than the UTP concentration and thus *in vivo* ATP is considered the preferred substrate for Star-PAP. The expression of some mRNAs involved in detoxification and/or oxidative stress responses by adding poly(A) tails were shown to be regulated by Star-PAP based on RNAi knock-down experiments. However, if as stated above this is an essential polymerase affecting cell viability the levels of mRNA would be lower compared to the control experiments due to cell death. Star PAP binds directly to both pre-mRNA and CPSF. A cleavage complex is formed between HO-1 RNA a target of star PAP and CPSF. This differs from the situation with canonical PAP in that CPSF interaction with PAP recruits PAP to the cleavage site. (Murthy and Manley, 1992, 1995).

POLS (Hs6, TRF4-1, TUT5, PAPD7)

Sequence analogy indicates *POLS* to be a human orthologue of *TRF4/5*.

PAPD5 (Hs7, TRF4-2, TUT3)

PAPD5, the other human orthologue of *TRF4/5,* was recently reported to be associated with the degradation of PROMPTs *via* the NEXT complex. The authors also found PAPD5 to interact with ZCCHC7 (hAIR2) and with hRRP6. In ActD treated Hela cells PAPD5 was shown to adenylate 5.8S 5'ETS rRNA transcripts (Lubas *et al*., 2011). *PAPD5* has also been

implicated in uridylation of histone mRNA 3' ends leading to histone mRNA turnover in the cytoplasm. Another study disagrees with *PAPD5* association with histone mRNA turnover (Schmidt *et al.*, 2011). As *PAPD5* is an orthologue of *TRF4/5* and the involvement of mouse PAPD5 in adenylating rRNA intermediate transcripts, a poly(A) polymerase activity is more likely to be the function of this polymerase.

1.3.3 Sequence comparison of POLS and PAPD5

Sequence comparison of POLS and PAPD5 indicate a very similar middle region. POLS and PAPD5 show 42.6% identity and 50.5% similarity (Figure 1.9).

```
PAPD5           MYRSGERLLGSHALPAEQRDFLPLETTNNNNNHHQPGAWARRAGSSASSPPSASSSPHPS
POLS            ------------------------------------------------------------


PAPD5           AAVPAADPADSASGSSNKRKRDNKPSGGRAAGGGRADGGGVVYSGTPWKRRNYNQGVVGL
POLS            ------------------------------------------------------------


PAPD5           HEEISDFYEYMSPRPEEEKMRMEVVNRIESVIKELWPSADVQIFGSFKTGLYLPTSDIDL
POLS            ----------MSPCPEEAAMRREVVKRIETVVKDLWPTADVQIFGSFSTGLYLPTSDIDL
                          *** ***  ** ***:***:*:*:***:********* ***********


PAPD5           VVFGKWENLPLWTLEEALRKHKVADEDSVKVLDKATVPIIKLTDSFTEVKVDISFNVQNG
POLS            VVFGKWERPPLQLLEQALRKHNVAEPCSIKVLDKATVPIIKLTDQETEVKVDISFNMETG
                ******* **  **:*****:**:  *:***************  **********:: *


PAPD5           VRAADLIKDFTKKYPVLPYLVLVLKQFLLQRDLNEVFTGGIGSYSLFLMAVSFLQLHPRE
POLS            VRAAEFIKNYMKKYSLLPYLILVLKQFLLQRDLNEVFTGGISSYSLILMAISFLQLHPRI
                ****::**:: *** :****:******************** ****:***:********


PAPD5           DACIPNTNYGVLLIEFFELYGRHFNYLKTGIRIKDGGSYVAKDEVQKNMLDGYRPSMLYI
POLS            DARRADENLGMLLVEFFELYGRNFNYLKTGIRIKEGGAYIAKEEIMKAMTSGYRPSMLCI
                **   : * *:**:********:***********:**:*:**:*: * *  ******* *


PAPD5           EDPLQPGNDVGRSSYGAMQVKQAFDYAYVVLSHAVSPIAKYYPNNETESILGRIIRVTDE
POLS            EDPLLPGNDVGRSSYGAMQVKQVFDYAYIVLSHAVSPLARSYPNRDAESTLGRIIKVTQE
                **** ***************** *****:*********:*: *** ::** *****:**:*


PAPD5           VATYRDWISKQWGLKNRPEPSCNGNGVTLIVDTQQLDKCNNNLSEENEALGKCRSKTSES
POLS            VIDYRRWIKEKWGSKAHPSPGMDSR----IKIKERIATCNGEQTQNREPESPYGQRLTLS
                *  ** ** ::** * :* *  :      *   :::  ** : ::: *       : : *
```

```
PAPD5           LSKHSSNSSSGPVSSSSATQSSSSDVDSDATPCKTP-----------------------
POLS             LSSPQLLSSGS--SASSVSSLSGSDVDSDTPPCTTPSVYQFSLQAPAPLMAGLPTALPMP
                 **     **    *:** :  * ******: ** **

PAPD5           -------------------------------KQLLCRPSTGNRVGSQDVSLESSQAVGKM
POLS            SGKPQPTTSRTLIMTTNNQTRFTIPPPTLGVAPVPCRQAGVEGTASLKAVHHMSSPAIPS
                                              : ** :  :   *       *

PAPD5           QSTQTTNTSNSTNKSQHGSARLFRSSSKGFQGTTQTSHGSLMTNKQHQGKSNNQYYHG--
POLS            ASPNPLSSPHLYHKQHNGMKLSMKGSHGHTQGGGYSSVGSGGVRPPVGNRGHHQYNRTGW
                 * :  : :   :* ::*    :: *    **   :* **         : ::** :

PAPD5           --KKRKHKRDAPLSDLCR
POLS            RRKKHTHTRDSLPVSLSR
                  **: * **:    * *
```

Figure 1.9 Sequence alignment of POLS and PAPD5.
Represented in red is the NTP transferase domain and in blue the PAP-associated domain. Asterisks indicate identical residues and double dots indicate similar residues. Highligted in yellow are the amino acids corresponding to the primers used to generate the POLS-GST-fusion protein. Note at the time of design of the POLS-GST fusion protein the N-terminus of PAPD5 was unknown. Thus, with the information available today it is possible that the antibody is capable of detecting endogenous POLS as well as PAPD5.

1.3.4 Sequence comparison of POLS, PAPD5, Trf4 and Trf5

Using sequence alignments alone it is difficult to predict if POLS is orthologous to Trf4 and PAPD5 is orthologous to Trf5 and *vice versa,* as the alignments only very marginally differ. Trf4 and Trf5 show 51% similarity, POLS and PAPD5 show 50% similarity, POLS and Trf4 show 22% similarity, POLS and Trf5 show 21% similarity, PAPD5 and Trf4 show 22% similarity and PAPD5 and Trf5 show 17% similarity (Figure 1.10). Clustalw2 was used to obtain these alignments and scores.

```
Trf4            MGAKSVTASSSKKIKNRHNGKVKKSKKIKKVRKPQKSISLNDENEVEILPSRNEQETNKL 60
Trf5            --MTRLKAKYSPTKGKRKEDKHTKRMRKSSFTRTQKMLEVFNDN----RSHFNKYESLAI 54
POLS            ------------------------------------------------------------
PAPD5           ----------MYRSGERLLGSHALPAEQRDFLPLETTNNNNNHHQPGAWARRAGSSASSP 50

Trf4            PKDHVTADGILVLEHKSDDDEGFDVYDGHFDNPTDIPSTTEESKTPSLAVHGDEKDLANN 120
Trf5            DVEDDDTFGNLVLMENDKSDVDIPVIE-------EVTSSEDEQRAESSKRN---NSLEDN 104
POLS            ------------------------------------------------------------
PAPD5           PSASSSPHPSAAVPAADPADS-------------ASGSSNKRKRDNKPSGGRAAGGGRAD 97

Trf4            DDFISLSASSEDEQAEQEEEREKQ-------ELEIKKEKQKEILNTD------YPWILNH 167
Trf5            QDFIAFSDSSEDETEQIKEDDDERSSFLLTDEHEVSKLTSQQSLNTESACNVEYPWIRNH 164
POLS            ------------------------------------------------------------
PAPD5           GGGVVYSGT-------------------------------------------PWKRRN 112

Trf4            DHSKQKEISDWLTFEIKDFVAYISPSREEIEIRNQTISTIREAVKQLWPDADLHVFGSYS 227
Trf5            CHSKQRRIADWLTSEIKDFVHYISPSKNEIKCRNRTIDKLRRAVKELWSDADLHVFGSFA 224
POLS            ----------------------MSPCPEEAAMRREVVKRIETVVKDLWPTADVQIFGSFS 38
PAPD5           YNQG----VVGLHEEISDFYEYMSPRPEEEKMRMEVVNRIESVIKELWPSADVQIFGSFK 168
                                     :**  :*   * ..:. :. .:*:**. **:::***:

Trf4            TDLYLPGSDIDCVVTSELGGKESRNNLYSLASHLKKKNLATE--VEVVAKARVPIIKFVE 285
Trf5            TDLYLPGSDIDCVVNSRNRDKEDRNYIYELARHLKNKGLAIR--MEVIVKTRVPIIKFIE 282
POLS            TGLYLPTSDIDLVVFG----KWERPPLQLLEQALRKHNVAEPCSIKVLDKATVPIIKLTD 94
PAPD5           TGLYLPTSDIDLVVFG----KWENLPLWTLEEALRKHKVADEDSVKVLDKATVPIIKLTD 224
                *.**** **** ** .    * ..  :  *   *::: :*    ::*: *: *****: :

Trf4            PHSGIHIDVSFERTNGIEAAKLIREWLDDTPGLRELVLIVKQFLHARRLNNVHTGGLGGF 345
Trf5            PQSQLHIDVSFERTNGLEAAKLIREWLRDSPGLRELVLIIKQFLHSRRLNNVHTGGLGGF 342
POLS            QETEVKVDISFNMETGVRAAEFIKNYMKKYSLLPYLILVLKQFLLQRDLNEVFTGGISSY 154
PAPD5           SFTEVKVDISFNVQNGVRAADLIKDFTKKYPVLPYLVLVLKQFLLQRDLNEVFTGGIGSY 284
                  : :::*:**:  .*:.**.:*:::  . . *  *:*::****  * **:*.***:..:

Trf4            SIICLVFSFLHMHPRIITNEIDPKDNLGVLLIEFFELYGKNFGYDDVALGSSDGYPVYFP 405
Trf5            TVICLVYSFLNMHPRIKSNDIDVLDNLGVLLIDFFELYGKNFGYDDVAISISDGYASYIP 402
POLS            SLILMAISFLQLHPRI--DARRADENLGMLLVEFFELYGRNFNYLKTGIRIKEGGAYIAK 212
PAPD5           SLFLMAVSFLQLHPRE--DACIPNTNYGVLLIEFFELYGRHFNYLKTGIRIKDGGSYVAK 342
                ::: :. ***::***   :      * *:**::******::*.* ...:  .:* .
```

```
Trf4            KSTWSAIQPIKNPFSLAIQDPGDESNNISRGSFNIRDIKKAFAGAFDLLTNRCFELHSAT 465
Trf5            KSCWRTLEPSRSKFSLAIQDPGDPNNNISRGSFNMKDIKKAFAGAFELLVNKCWELNSAT 462
POLS            EEIMKAMTSGYRPSMLCIEDPLLPGNDVGRSSYGAMQVKQVFDYAYIVLS-HAVSPLARS 271
PAPD5           DEVQKNMLDGYRPSMLYIEDPLQPGNDVGRSSYGAMQVKQAFDYAYVVLS-HAVSPIAKY 401
                ..    :        * *:**   .*::.*.*:.  ::*:.*  *: :*  :. .  :

Trf4            FKDRLGKSILGNVIKYRGKARDFK-------------------------DERGLVLNKA 499
Trf5            FKDRVGKSILGNVIKYRGQKRDFN-------------------------DERDLVQNKA 496
POLS            YPNRDAESTLGRIIKVTQEVIDYRRWIKEKWGSKAHPSPGMDSR----IKIKERIATCNG 327
PAPD5           YPNNETESILGRIIRVTDEVATYRDWISKQWGLKNRPEPSCNGNGVTLIVDTQQLDKCNN 461
                : :.  :* **.:*:   :   :.                           . :   :

Trf4            IIENENYHKKRSRIIHDEDFAEDTVTSTATATTTDDDYEITNPPAKKAKIE--------- 550
Trf5            IIENERYHKRRTRIVQEDLFINDTEDLPVEEIYKLDEPAKKKQKAKKDKREGEIK----- 551
POLS            EQTQNREPESPYGQRLTLSLSSPQLLSSGS--SASSVSSLSGSDVDSDTPPCTTPSVYQF 385
PAPD5           NLSEENEALGKCRSKTSESLSKHSSNSSSGPVSSSSATQSSSSDVDSDATPCKTP----- 516
                   ::.            : .     .       .    .   ...

Trf4            ------------------------------------------------------------
Trf5            ----------------------------------------------------------KSAI 555
POLS            SLQAPAPLMAGLPTALPMPSGKPQPTTSRTLIMTTNNQTRFTIPPPTLGVAPVPCRQAGV 445
PAPD5           ----------------------------KQLLCRPSTGNR------------VGSQDVSL 536

Trf4            -----------EKPESEPAKRNSG----ETYITVS-----SEDDDE-------------- 576
Trf5            PSPPPDFGVSRSKLKRKVKKTDQGSLLHQNNLSIDDLMGLSENDQESDQDQKGRDTPSGQ 615
POLS            EGTASLKAVHHMSSPAIPSASPNPLSSPHLYHKQHNGMKLSMKGSHGHTQGGGYSSVGSG 505
PAPD5           E---SSQAVGKMQSTQTTNTSN---STNKSQHGSARLFRSSSKGFQGTTQ----TSHGSL 586
                            .              .           * .. .

Trf4            DGYNPYTL------------------------------ 584
Trf5            DEKSPLETKTVDAQTRRDYWLSKGQA------------ 641
POLS            GVRPPVGNRGHHQYNRTGWRRKKHTHTRDSLPVSLSR 542
PAPD5           MTNKQHQGKSNNQY----YHGKKRKHKRDAPLSDLCR 619
```

Figure 1.10. Sequence alignment of Trf4, Trf5 POLS and PAPD5

Represented In red is the NTP transferase domain and in blue the PAP associated domain. Asterisks indicate identical residues and double dots indicate similar residues.

1.4 Aims of this study:

Multiple sequence alignments assign *POLS* and *PAPD5* as the human orthologues of *S.cerevisiae TRF4/5.* Although reports suggest that the TRAMP/exosome pathway is likely to be conserved in human cells, the poly(A) polymerase(s) responsible has/have yet to be identified. A recent study assigns PAPD5 involvement in PROMPTs degradation and processing of 5'ETS rRNA (Lubas *et al*., 2011).

This thesis has investigated the degree of conservation and extent of divergence of POLS and PAPD5 from their orthologues in relation to cellular and biochemical functions. Attempts were also made to find interacting protein partners of POLS.

In brief, in chapter 3 as a first step the localization of POLS was sought in order to follow up its functions. The effects on cell growth as a result of individual down-regulation of *POLS* or *PAPD5* and the simultaneous down-regulation of *POLS* and *PAPD5* were investigated. Chapter 4 analyses the role of POLS and PAPD5 in adenylation of the human 28S rRNA 3'ETS. In particular, the possibility that POLS or/and PAPD5 might be responsible for adding non-A residues to the oligo A tails were assayed. The possibility that these genes might be controlling the processivity was also investigated. Attempts made to identify interacting protein partners of POLS are described in chapter 5. Novel splice variants identified in an expression screen not reported in public databases are documented in Chapter 6.

2 Materials and Methods

All chemicals used were supplied by Sigma, enzymes by Roche and radioisotopes by PerkinElmer LAS (UK) Ltd., unless otherwise stated. RNA and DNA concentrations were determined using a NanoPhotometer (IMPLEN).

2.1 Cell culture methods

2.1.1 Growth and maintenance of human cells

Cells were grown at 37ºC with 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM, PAA laboratories) containing 5% foetal calf serum (FCS, Biosera), 100 U/ml penicillin and 100 µg/ml streptomycin (PAA laboratories). U2-OS cells were obtained from the American Type Culture Collection (ATCC). HEK-293T cells were acquired from lab stocks.

2.1.2 Freezing cells for storage

Adherent cells were dissociated by adding phosphate buffered saline (PBS) containing 0.15% trypsin (PAA laboratories) and 1 mM ethylene diaminetetra-acetic acid (EDTA) and collected by centrifugation at 234 x *g* at room temperature (RT) for 4 min. Cells were washed with 1x (PBS, PAA laboratories). Non-adherent cells were harvested by gently tapping plates and aspirating cells with a 10 ml pipette followed by centrifugation at 234 x *g* at room temperature (RT) for 4 min. Cells were resuspended in freezing medium [growth medium as above containing 10% dimethyl sulfoxide (DMSO)] and stored at -80ºC in cryo-vials before transfer to liquid nitrogen for long term storage.

2.1.3 siRNA transfection

Two gene-specific siRNA and two scrambled control duplexes (Table 2.1) were transfected into U2-OS or HEK-293T cells using Lipofectamine 2000 (Invitrogen), according to the manufacturer's instructions. Briefly, cells were grown in antibiotic-free medium and plated at approximately 25% confluency the day before the transfection in 9 cm plates. The next day, growth medium was removed and cells were washed twice in PBS. Cells were then incubated at 37ºC with 3.5 ml OPTI-MEM (Invitrogen). Transfections were set up in 1.5 ml Eppendorf tubes at room temperature (RT). In one tube siRNA duplexes were combined with OPTI-MEM to a total volume of 750 µl and in another tube Lipofectamine was combined with OPTI-MEM to a total volume of 750 µl and incubated for 5 min. A 100 nM final concentration of duplexes was employed throughout, apart from when gene knockdown efficiencies were optimised. The duplexes and Lipofectamine were mixed and left for 20-25 min at RT before addition to the cells. After 4-6 hrs cells were supplemented with 5 ml of growth medium and incubated at 37ºC. Cells were harvested after two days or as needed. For transfections performed in 6-well plates, quantities of siRNA duplexes, Lipofectamine and OPTI-MEM were scaled down accordingly.

2.1.4 Plasmid transfection

Plasmids pcDNA3-*POLS*-HA and pcDNA3-*POLS*-TAP (constructed as described in section 2.3.15) were transfected into U2-OS or HEK-293T cells as above (2.1.3) with 12 µg of the respective plasmid DNA replacing the siRNA duplexes.

2.1.5 Generation of stable transfectants

The plasmid pcDNA3-*POLS*-HA, which contains the neomycin resistance cassette, was transfected as above into U2-OS cells. Cells expressing *POLS*-HA were selected using 400 µg/ml G418. After approximately two weeks, single colonies appeared on 9 cm plates. Each

colony was carefully lifted with a 1 ml Gilson tip into a single well in a 96-well plate. When this well was confluent the clone was then transferred into a 24-well plate, and then onto a 6-well plate, and finally to a 9 cm plate maintaining the clone in 400 µg/ml G418 at all times.

2.1.6 Immunofluorescence microscopy

Cells were grown on 0.15 mm cover slips and were transiently transfected as described above. After 48 hr, growth medium was removed and cells were washed twice with PBS. Cells were fixed with 4% paraformaldehyde (v/v) in 500 mM HEPES pH 7.5 (Gibco) for 30 min at 4ºC. After three washes in PBS, cells were permeabilized with 0.5% Triton X-100 (v/v) in PBS for 1 hr at 4ºC, followed by blocking with 10% FCS (v/v) in PBS for 1 hr at 37ºC. Anti-HA [1:1000 in PBS containing 0.1% Tween 20 (v/v)] or anti–TAP [1:1000 in PBS containing 0.1% Tween 20 (v/v), Open Biosystems] antibodies were added and incubated for 2 hr (RT)/overnight (4ºC) along with a control with no primary antibody. Cells were washed three times in PBS containing 0.1% Tween 20 (v/v) for 10 min each at RT with gentle agitation. Cy3-conjugated anti mouse IgG secondary antibody [1:1000 in PBS containing 0.1% Tween 20 (v/v)] was added to all cells. After 30 min incubation, three washes in PBS containing 0.1% Tween 20 (v/v) were performed. Vectashield mounting medium (Vector Laboratories) containing 4', 6-diamidino-2-phenylindole (DAPI) was used to mount cover slips onto slides. Images were collected with a CCD camera mounted on a Zeiss inverted microscope with a 63x objective lens. A fixed exposure time was used for each set of images. MetaMorph software (Universal Imaging, Downingtown, Pennsylvania, United States) was used to view the images. Some cells were visualized by confocal microscopy (Olympus FluoView FV1000) using a 60x objective lens.

2.1.7 Cell growth assay

Control and siRNA-treated cells were subjected to a MTT (3-(4, 5-Dimethyl thiazol-2-yl)-2,5-diphenyltetrazolium bromide)-based assay (EZ4U, Biomedica). The reduction of yellow coloured tetrazolium compounds by mitochondrial reductase into red formazan derivatives was recorded by measuring absorbance at 450 nm. Transfections were performed as described in section 2.1.3 and after 4 hr, cells were trypsinised and split into 24-well plates at a density of $2.5x10^4$/well. At each time point substrate was added to the wells and left for colour development for 3 hr. The A_{450} at this time point was measured using a micro-plate reader (ELx808, BioTex). Each time point was performed in triplicate and individual measurements were recorded by taking the mean of triplicate readings.

2.1.8 Trypan blue assay

Trypan blue assays were performed in parallel with cell growth assays. Transfections and sub-culturing were performed as above. The medium in which cells were growing was saved and adherent cells were trypsinised and resuspended in the saved growth medium. After the addition of 0.4% trypan blue, viable and dead cells were counted using a haemacytometer at each time point. Each time point was performed in duplicate and individual measurements were recorded by taking the mean of triplicate readings.

2.1.9 Flow cytometry

Cells were trypsinised and washed once with PBS. The cell pellet was fixed by resuspension in 2 ml of 70% ethanol per 9 cm plate for at least 30 min at 4ºC. Fixed cells were collected by centrifugation at 234 x *g* for 4 min at 4ºC, and were rehydrated by two 10 ml PBS washes in PBS. Cells were re-suspended in PBS containing 40 µg/ml propidium iodide and 100 µg/ml RNase A. Prior to flow cytometric analysis (FACScan, Becton Dickinson) cells were incubated in the dark for 15 min at RT.

2.1.10 Cycloheximide treatment

To 75% confluent U2-OS cells in 9 cm plates, cycloheximide was added (from a 100 mg/ml stock in DMSO) to give a final concentration of 10^5, 10^4, 10^3, 10^2, 10^1 or 10^{-1} nM and cells were incubated at 37ºC. After 30 min cells were subjected to flow cytometric analysis as described in section 2.1.9.

2.1.11 5-FU treatment and colony formation assay

U2-OS cells were transfected with POLS, PAPD5, (POLS + PAPD5) or scrambled control siRNA duplexes. Six hours later, cells were plated at 150 cells per 9 cm plate and incubated at 37ºC. This cell number enabled counting of individual colonies after 10 days. Different concentrations of 5-FU were added the day after plating and cells were incubated for two days. The drug was then washed out with 1x PBS. DMEM containing 5% FCS,100 U/ml penicillin and 100 µg/ml streptomycin was added to the plates which were incubated for 8 days. After removing the growth medium, colonies were stained with crystal violet [0.05% crystal violet (w/v) in 25% methanol (v/v] and were counted using a colony counter (Stuart Scientific). Each treatment was performed in triplicate.

2.2 RNA methods

2.2.1 RNA extraction

Total RNA was isolated using Trizol (Applied Biosystems) according to the manufacturer's instructions. Briefly, cells were harvested from a 9 cm culture dish by trypsinisation and collected in a 1.5 ml screw-cap centrifuge tube. Cells were resuspended in 1 ml of Trizol and lysates were incubated at RT for 10-15 min. To this lysate 200 µl of chloroform was added and tubes were tightly capped and shaken vigorously for 15 sec. After 2 min incubation at

RT, and centrifugation at 16,000 x *g* for 20 min at 4ºC, the aqueous phase was transferred to a fresh tube containing 500 µl of isopropanol. This was mixed and centrifuged at 16,000 x *g* for 20 min at RT. RNA pellets were washed with 1 ml of 75% ethanol per 1 ml Trizol, and recovered by centrifugation at 16,000 x *g* at RT for 10 min. The resulting pellet was air dried for 5 min. RNA was then dissolved in 50 µl of de-ionized water and stored at 4ºC for immediate use or -80ºC for long-term storage.

2.2.2 DNase treatment

To eliminate any contaminating DNA, 20 µg of RNA was incubated at 37ºC for 1 hr with 1 unit of RQ-DNase I (Promega), 2 µl of 10x DNase buffer (Promega) in a total volume of 20 µl made up with deionised water. A further 0.5 units of RQ-DNase was added and incubated for a further 30 min. The reaction was stopped by incubation at 70ºC for 10 min.

2.2.3 cDNA synthesis

SuperScript II (Invitrogen) was used according to the manufacturer's instructions. Briefly, 0.5-1 µg of total RNA was incubated at 65ºC for 5 min with 1 µl of 10 mM dNTP mix (10 mM each of dATP, dCTP, dGTP, dTTP) and 1 µl of 50 µM oligonucleotide primer. This was cooled for 1 min before the addition of the following: 4 µl of 5x RT-buffer (Invitrogen), 4 µl of 25 mM $MgCl_2$, 2 µl of 0.1 M dithiothreitol (DTT), and 1 µl of RNAseOUT recombinant RNase inhibitor. Following this an incubation for 2 min at 42ºC and then the addition of 25 units of SuperScript II and a further incubation of 58 min at 42ºC was carried out. The reaction was stopped by incubation at 70ºC for 10 min.

2.2.4 Ethanol precipitation

To concentrate RNA or DNA, 2.5 volumes of 100% ethanol and 0.1 volumes of 3 M sodium acetate pH 5.2 were added, mixed gently and incubated at -20ºC overnight. The precipitate

was centrifuged at 16,000 x *g* for 20 min at RT. Pellets were air dried and re-suspended in de-ionized water (normally 30 µl).

2.2.5 Semi-quantitative RT-PCR screen of human non-canonical poly(A) polymerases

cDNA was synthesised (as described above in section 2.2.3), and using the primer pairs listed in Table 2.2, segments of each cDNA were amplified by PCR (described section 2.3.8) to obtain products in the region of 500 base pairs. To provide a measure of relative quantitation, a range of cycle numbers were used for each gene.

2.2.6 Quantitative PCR (Q-PCR)

Q-PCR was performed using SYBR green PCR Master Mix (Stratagene/Bioline) on a Rotagene 6000/3000 real time PCR system. Primers used are listed in Table 2.3. Q-PCR reactions were set up as follows: 5 µl of SYBR mix, 0.5 µl of 10 µM sense oligonucleotide primer, 0.5 µl of 10 µM antisense oligonucleotide primer, 1 µl cDNA and deionised water to make up to 10 µl. For each sample, a control reaction was performed using template prepared in the absence of reverse transcriptase. Thermal cycling was performed as follows: 95^{o}C for 15 min, followed by 40 cycles at 94^{o}C for 15 sec, 60^{o}C for 15 sec and 72^{o}C for 30 sec. All RNA levels were normalized to mature 28S rRNA.

2.2.7 3' RACE

1 µg of total RNA was used as the starting material for reverse transcription with an oligo-dT_{16}-adaptor primer using SuperScript II (Invitrogen) as described above in section 2.2.3. Fivefold serial dilutions of the resulting cDNA were used for the actin screen. PCR was performed with a gene-specific primer (actin F or 3'ETSFF) and oligo-dT_{16} adaptor for 25

cycles. A second nested PCR of 30 cycles (for actin 20 cycles) was performed using another gene-specific primer (actin FF or 3'ETSF) and oligo-dT_{16} adaptor. For the actin screen the above procedure was carried out with two polymerases; Taq polymerase (Fisher scientific) and Expand high-fidelity (Roche). 3' ETS tails were screened using Taq polymerase only.

2.2.8 Poly(A) selection of RNA

Total RNA was extracted as described in section 2.2.1 and poly(A) selection was performed using PolyATtract mRNA isolation system (Promega) according to the manufacturer's protocol.

2.2.9 Northern analysis

3-10 µg of RNA was resuspended in 3x gel loading buffer [50% formamide (v/v), 6% formaldehyde (w/v), 0.025% bromophenol blue (w/v), 20 mM MOPS pH 7.0, 8 mM sodium acetate pH 5.2, 1 mM EDTA pH 8.0, 10% glycerol (v/v), 20 µg/ml ethidium bromide], denatured at 65°C for 10 min and then cooled on ice prior to loading. A 2% agarose (w/v) gel containing 0.7 M formaldehyde was run at 5V/cm in 1x MOPS. RNA was visualized on a UV transilluminator. For the transfer, the gel was rinsed in water for 10 min, followed by two washes in 10x SSC for 10 min. RNA was transferred onto a positively charged nylon transfer membrane (Hybond N^+, Amersham) via capillary transfer set up with Whatman 3 MM paper overnight. The next day the RNA was UV cross-linked to the membrane using a UVC 500 crosslinker (Stratagene, full power, 30 sec). Membranes were stored at RT.

Actin probe (amplified using primers NAcF1 and NAcR1, Table 2.6) was radiolabelled with [α^{32}P] dCTP by random priming, using a Rediprime II Kit (GE Healthcare) according to the manufacturer's instructions. Briefly, 25 ng of gel- purified DNA was diluted in 45 µL 1X Tris-EDTA (TE) and denatured at 95°C for 5 min. Tubes were placed on ice for 5 min. DNA and 5 µl [α^{32}P] dCTP (2.4 MBq) was added to the reaction tube, mixed and incubated for 30 min at

37°C. The reaction was stopped by adding 5 µl of 0.2 M EDTA. Unincorporated nucleotides were removed with a size exclusion column (MicroSpin S-300 HR column, GE Healthcare) which was prepared by centrifugation at 1 min at 1000 x *g*. This was followed by a 2 min centrifugation at 1000 x *g* with the sample. The membrane was pre-hybridized by incubating for 1-2 hr at 42°C in 10-20 ml of hybridization buffer [5 X SSC, 50% formamide (v/v), 1% SDS (w/v), 1 X Denhardt's solution, 100 µg/ml salmon sperm DNA, 10% dextran sulphate (w/v)]. The actin probe was denatured by heating for 5 min at 95°C and chilling on ice before adding to the hybridization buffer and incubating overnight at 42°C. The membrane probed with actin was washed in 1X SSC containing 0.1% SDS (w/v) for 15 min at 42ºC then 15 min at 65°C, followed by two 10 min washes at 65°C in 0.1X SSC containing 0.1% SDS (w/v). The membrane was wrapped in plastic film and exposed to a phosphor screen (Molecular Dynamics). Screens were scanned using a FLA5000 Fuji Phosphoimager. Images were analyzed with AIDA analysis software (Raytest GmbH).

The 3' ETS (N3'ETSF, Table 2.6) and 47S (N47S1, Table 2.6) oligonucleotide probes were 5' radiolabelled as follows: reactions were set up to include 10 pmol primer, 1.5 µl 10X polynucleotide kinase (PNK) buffer, 5 µl [γ^{32}P] ATP (2.4 MBq), 1.5 µL PNK and de-ionized water to make a volume of 15 µl. The reaction was allowed to proceed at 37°C for 30-45 min. Unincorporated nucleotides were removed using Roche Mini-Quickspin oligo columns. The column was prepared by vortexing for 5 sec and centrifugation at 1,000 x *g* for 1 min. The labelling reaction was made up to a volume of 50 µl and was loaded on to the column and centrifuged for 4 min at 1,000 x *g*. For the 3'ETS probe pre-hybridization was carried out as described above but in hybridization buffer (1% SDS (w/v), 6X SSPE, 100 mg/ml salmon sperm DNA) at 60°C. The membrane was washed three times with 6X SSPE, 0.1% SDS (w/v) at 60ºC, wrapped in plastic film and exposed to a phosphor screen (Molecular Dynamics). Screens were scanned using a FLA5000 Fuji Phosphoimager. Images were analyzed with AIDA analysis software (Raytest GmbH).

2.2.10 Stripping northern blots

The membrane was immersed in boiling water containing 0.1% SDS (w/v), and allowed to cool. The blot was rinsed briefly in 2X SSC and monitored for removal of the probe. Stripped membranes were stored at -20°C.

2.3 DNA methods

2.3.1 Determination of *E.coli* cell concentration

An Eppendorf Biophotometer was used to measure the absorbance at 600 nm (A_{600}). An A_{600} of 1 was taken to be equivalent to $2x10^7$ cells per ml.

2.3.2 Growth and maintenance of *E.coli*

Transformed TOP 10 [F- *mcrA* (*mrr-hsd*RMS-*mcrBC*) 80lacZM15 *lacX74 recA1 ara139* (ara-leu)7697 *galU galK rpsL* (Str^R) *endA1 nupG*] or BL21 (F–, *omp*T, *hsd*S_B (r_B–, m_B–), *dcm*, *gal*, λ(DE3), pLysS, Cm^r) cells were grown in Luria Bertani (LB) agar/medium containing ampicillin (100 µg/ml). For BL21 cells chloramphenicol (50µg/ml) was also added to maintain the pLysS plasmid carrying a T7 lysozyme gene. T7 lysozyme lowers the background expression level of target genes under the control of the T7 promoter without affecting the level of expression achieved following induction with isopropyl β-D-1-thiogalactopyranoside (IPTG).

2.3.3 Generation of chemically competent *E.coli* cells

For blue-white colony screening, competent TOP 10 cells were generated by the rubidium chloride method. The calcium chloride method was used to generate chemically competent BL21 cells.

Rubidium chloride method

A single TOP 10 colony was resuspended in 5 ml LB medium containing ampicillin (100 µg/ml) and grown overnight at 37ºC. The following day 250 ml LB medium containing ampicillin (100 µg/ml) was inoculated with 2.5 ml of this starter culture and grown until cells were in mid-exponentional phase. Cells were harvested in chilled 50 ml Falcon tubes by centrifugation at 4000 x *g*, 15 min at 4ºC. The cell pellet was resuspended in 200 ml cold filter-sterilized TFBI (30 mM KAc, 100 mM $RbCl_2$, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 15% glycerol (v/v) pH 5.8). Cells were incubated on ice for 20 min and harvested by centrifugation (4000 x *g*, 15 min at 4ºC). Cells were resuspended in cold filter-sterilized TFBII (10 mM MOPS, 75 mM $CaCl_2$, 10 mM $RbCl_2$, 15% glycerol (v/v) pH 6.5) and incubated on ice for 20 min. 200 µl aliquots were snap-frozen in ethanol-dry ice and stored at -80ºC.

Calcium chloride method

A 5 ml starter culture of BL21 cells was grown in LB medium containing ampicillin (100 µg/ml). The following day 250 ml LB medium containing ampicillin (100 µg/ml) was inoculated with 2.5 ml of starter culture. When A_{600} was between 0.4-0.6, cells were collected by centrifugation at 4000 x *g* for 15 min at 4ºC. The pellet was resuspended in 20 ml of 50 mM $CaCl_2$ and chilled on ice for 30 min. Cells were collected by centrifugation at 4000 x *g* for 15 min at 4ºC and resuspended in 4 ml of 50 mM $CaCl_2$ for 30 min. 200 µl aliquots were snap-frozen in ethanol-dry ice and stored at -80ºC.

2.3.4 Chemical transformation of *E.coli* TOP10 and BL21 cells

An aliquot of competent cells (prepared as above in section 2.3.3) was thawed gently on ice and 100 ng of plasmid DNA or 10 µl of ligation reaction was added. The cell suspensions were incubated on ice for 30-45 min, heat shocked for 90 sec at 42ºC and returned to ice. To each sample 600 µl of pre- warmed LB (37ºC) was added and plated on LB agar plates containing ampicillin (100 µg/ml) and incubated at 37ºC overnight. Transformants were grown up in 3ml of LB medium containing ampicillin (100 µg/ml) overnight. Colony PCR was performed to identify positive transformants. In some cases a small scale plasmid preparation was performed and the presence of an insert was verified by restriction enzyme digestion. The DNA sequence was analysed to confirm plasmid identity.

2.3.5 TA cloning

TA cloning (Invitrogen) was carried out according to the manufacturer's instructions. Briefly, 10-100 ng of PCR product was ligated at RT, overnight. The ligation reaction mixtures were set up as follows: 1 µl of 10X ligation buffer, 0.5 µl of pCR2.1 vector, 10-100 ng of fresh PCR product, 1 µl T4 DNA ligase made up to a final volume of 10 µl with de-ionized water. Reaction mixes were chemically transformed into *E.coli* TOP10 cells. Blue-white colony screening was carried out using X-Gal (40µg/ml) on LB ampicillin (100 µg/ml) plates.

2.3.6 Plasmid preparations (small scale)

Cell pellets were obtained by centrifugation (16,000 x *g* for 1 min at RT) of overnight saturated liquid cultures of *E.coli* transformants. The following buffers were added in sequence to the cell pellet; 200 µl of resuspension buffer-P1 (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RNase A), 200 µl of lysis buffer-P2 [200 mM NaOH, 1% SDS (w/v)], and 350 µl of neutralization buffer-P3 (3 M potassium acetate pH 5.5). The suspension was

mixed by inversion and centrifuged for 10 min at 16,000 x *g* at RT. Avoiding the white precipitate, the supernatant was added to a fresh tube containing 750 µl of 100% ethanol, mixed and centrifuged for 20 min at 16,000 x *g* at RT. The supernatant was removed and the pellet was resuspended in 30-50 µl of de-ionized water.

2.3.7 Plasmid preparations (large scale)

A Qiagen maxi prep kit was used according to the manufacturer's instructions. Briefly, cells were harvested by centrifugation at 4000 x *g* for 10 min at 4ºC. The pellet was resuspended in 10 ml of cold buffer P1 (as above in section 2.3.6). Buffer P2, 10 ml (as above in section 2.3.6) was added, mixed gently and left at RT for 5 min. This was followed by the addition of 10 ml of chilled neutralization buffer P3 (as above in section 2.3.6), and an incubation on ice for 20 min. The suspension was centrifuged for 30 min at 20,000 x *g* at RT. The supernatant was added to a Qiagen-tip 500 column, which was equilibrated using 10 ml buffer-QBT [750 mM NaCl, 50 mM MOPS pH 7.0, 15% isopropanol (v/v) 0.15% Triton X-100 (v/v)] prior to the addition of supernatant. The column was washed twice with 30 ml buffer-QC [1.0 M NaCl, 50 mM MOPS pH 7.0, 15% isopropanol (v/v)]. DNA was eluted with 15 ml of buffer-QF [1.25 M NaCl, 50 mM Tris-HCl pH 8.5, 15% isopropanol (v/v)]. DNA was precipitated with 0.7 volumes of isopropanol, at RT. The precipitate was collected by centrifugation at 15,000 x *g* for 30 min at 4ºC. The DNA pellet was washed in 5 ml of 70% ethanol at RT and centrifuged at 15,000 x *g* for 10 min. Supernatant was removed and the pellet was air-dried for 5-10 min and dissolved in 1 ml de-ionized water.

2.3.8 PCR

For the template 1 µl of cDNA made as described in section 2.2.3, 100 ng of genomic DNA or 10 ng plasmid DNA was used in a reaction volume of 50 µl. The PCR master mix was made up as follows: 5 µl of 10X PCR buffer (Fisher scientific), 1 µl dNTP mix 10 mM

(Biorad), 3 µl 25 mM $MgCl_2$, 1 µl of 10 µM sense oligonucleotide primer, 1 µl of 10 µM antisense oligonucleotide primer, 0.2 units of Taq polymerase (Fisher scientific), made up to a total volume of 49 µl with de-ionized water. Thermal cycling was performed in a GeneAMP PCR system 9700 as follows: denaturation at 95ºC for 5 min, 25-36 cycles of [denaturation at 95ºC for 10 sec, annealing at 50-60ºC for 1 min, extension at 72ºC for 1 min], final elongation at 72ºC for 10 min and chilled at 4ºC until further use. For templates with high GC content, a 0.375 mM $MgCl_2$ concentration was employed compared to the 1.5 mM $MgCl_2$ concentration used above. Addition of 5% DMSO aided the amplification of regions with stable secondary structures.

2.3.9 Colony PCR from bacterial cells

A single colony of bacteria was resuspended in the PCR mix as above. Thermal cycling was carried out as above for 25-30 cycles.

2.3.10 Agarose gel electrophoresis

DNA loading dye [6X DNA loading dye: 0.01% bromophenol blue (w/v), 0.1 M EDTA pH 8.0 and 30% sucrose (w/v)] 0.2 volumes was added to 20 µl samples prior to loading 1-2% agarose gels containing 0.5 µg/ml ethidium bromide prepared with 1X Tris-Borate-EDTA (TBE) buffer. Gel electrophoresis was carried out in 1X TBE buffer at 2-10 V/cm. DNA was visualized using a UV transilluminator (UVP). Hyperladder I or IV (Bioline) was used as a size marker.

2.3.11 Restriction enzyme digestion

Each digest was set up as follows: 0.2-1 µg DNA, 3 µl appropriate restriction buffer, 1 µl of restriction enzyme(s). The total volume was adjusted to 30 µl with de-ionized water.

Reactions were incubated for 1-2 hr at the optimum temperature as specified by the manufacturer, and digestion verified by agarose gel electrophoresis.

2.3.12 DNA sequencing

Dideoxy-terminator DNA sequencing was carried out using the BigDye Terminator Cycle Sequencing Kit, Version 3.1 (Applied Biosystems). 1 µl of BigDye reagent was mixed with 200 ng of DNA, 3.2 pmol oligonucleotide primer, 1.5 µl of 5X reaction buffer (Applied Biosystems) and made up to a final volume of 10 µl with de-ionized water. Reactions were thermally cycled in a GeneAmp PCR system 9700 (Applied Biosystems) as follows: 96ºC for 4 min, 25 cycles of [96ºC for 10 sec, 50ºC for 5 sec, and 60ºC for 4 min] and chilled at 4ºC. Samples were ethanol precipitated and DNA pellets were stored at 4ºC. Sequencing was performed using a 3730x/ DNA analyzer by GeneService, Department of Biochemistry, University of Oxford. Sequence data were analyzed using ApE (universal) software.

2.3.13 Gel Purification

DNA was purified from agarose gel slices using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's instructions. Briefly, the DNA fragment was excised from the agarose gel and the gel slice was weighed in a tube. Three volumes of buffer QG (composition unknown- proprietary) was added to one volume of gel (for 100mg, 100µl of QG). The gel slice was dissolved by incubation at 50ºC for 10 min. To the suspension one gel volume of isopropanol was added and mixed. To bind DNA the sample was added to a QIAquick column and centrifuged at 12000 x *g* for 1 min at RT. To wash 0.75 ml of buffer PE (composition unknown- proprietary) was added to the column and centrifuged at 12000 x *g* for 1 min at RT. The flow-through was discarded and the column was centrifuged at 12000 x *g* for 1 min RT to remove any residual buffer. The column was placed in a clean 1.5ml

Eppendorf tube and 30µl of de-ionized water was added to the column which was centrifuged at 12000 x *g* for 1 min to recover the DNA.

2.3.14 Ligation

Ligation reactions contained 25-100 ng vector and insert DNA, 10 units of T4 DNA ligase, 1 µl of 10X ligase buffer and were adjusted to a final volume of 10 µl with de-ionized water. Reactions were incubated at room temperature for 4 hr or overnight.

2.3.15 Construction of *POLS/PAPD5* expression plasmids

pcDNA-*POLS*-HA

A pUHD10-3 vector (gift from Soyoung Min) containing a sequence encoding a HA2His6 tag and the full length *POLS* cDNA (ENST00000230859) from the LAK-1 clone (gift from Yasuhito Abe) were digested with restriction enzymes *Sac* II and *Not* I at the 5′ and 3′ ends, respectively. *POLS* was ligated into the pUHD10 vector as described in section 2.3.14. For expression in mammalian cells, this cDNA was subsequently re-cloned with the HA-tag sequence into pcDNA3.1+ (Invitrogen) by digesting both plasmids with *Hind* III (5’ end) and *BamH* I (3’ end). Ligation was performed as in section 2.3.14.

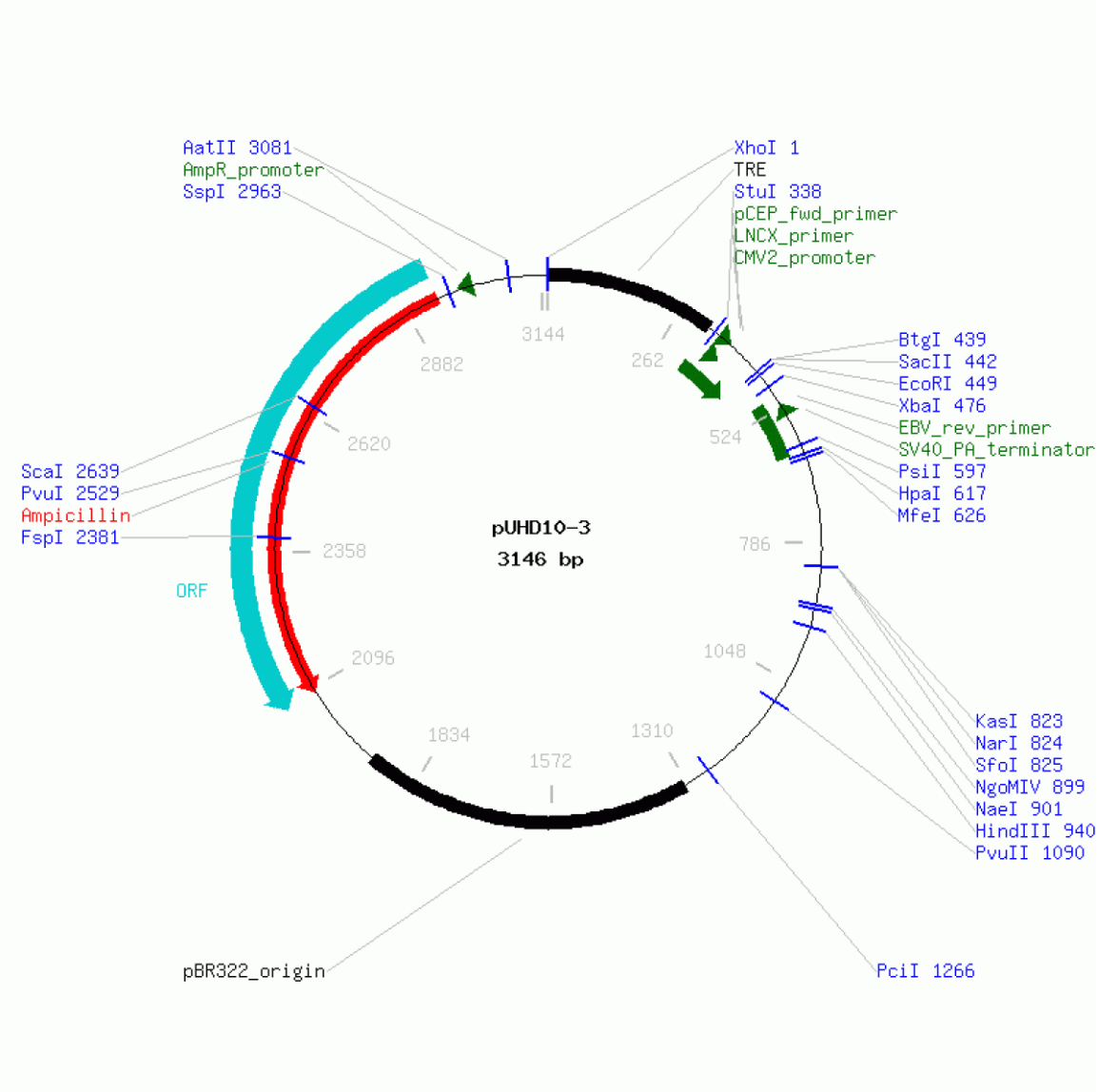
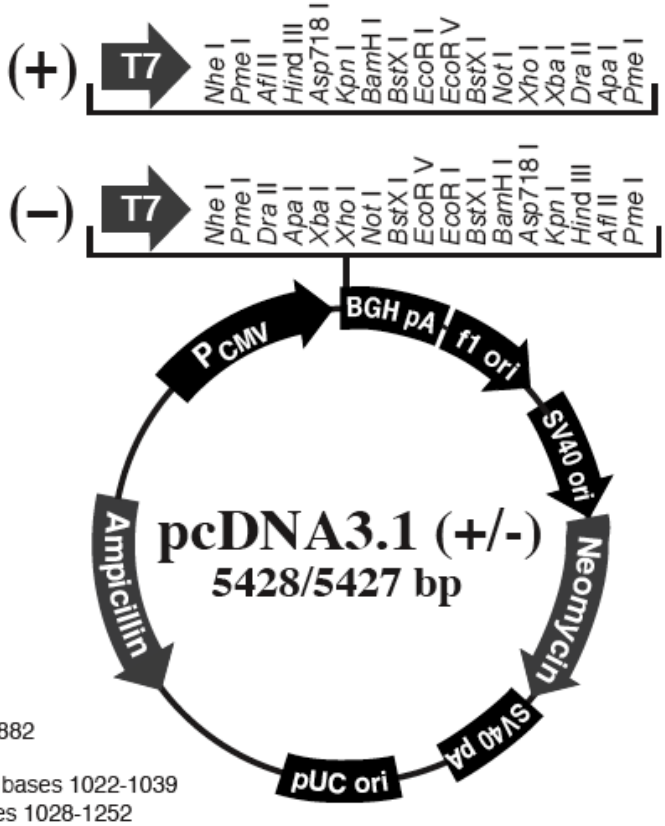


Figure 2.1 pUHD10-3 plasmid map (from Lablife web pages).

POLS cDNA and pUHD10 vector were digested with the restriction enzymes *Sac* II and *Not* I before ligation. Note the pUHD10 vector containing the HA2His6 (gift from Soyoung Min) contained the *Not* I and *BamH* I restriction enzyme sites. pUHD10-POLS-HA was inserted into the pcDNA3.1+ by digesting both plasmids with *Hind* III and *BamH* I followed by ligation.



Comments for pcDNA3.1 (+)
5428 nucleotides

CMV promoter: bases 232-819
T7 promoter/priming site: bases 863-882
Multiple cloning site: bases 895-1010
pcDNA3.1/BGH reverse priming site: bases 1022-1039
BGH polyadenylation sequence: bases 1028-1252
f1 origin: bases 1298-1726
SV40 early promoter and origin: bases 1731-2074
Neomycin resistance gene (ORF): bases 2136-2930
SV40 early polyadenylation signal: bases 3104-3234
pUC origin: bases 3617-4287 (complementary strand)
Ampicillin resistance gene (*bla*): bases 4432-5428 (complementary strand)
ORF: bases 4432-5292 (complementary strand)
Ribosome binding site: bases 5300-5304 (complementary strand)
bla promoter (P3): bases 5327-5333 (complementary strand)

Figure 2.2 pcDNA3.1 plasmid map (from Invitrogen life technologies).

To express proteins in mammalian cells pcDNA3.1+ plasmid was used. To insert HA-tag both pUHD10 and pcDNA3.1+ plasmids were digested with *Hind* III and *BamH* I. For the insertion of TAP-tag restriction enzymes *Not* I and *Xba* I were used to exercise the TAP-tag from the pcDNA3.1+ plasmid containing the TAP-tag. pcDNA-POLS-HA was digested with *Not* I and *Xba* I before ligation of the TAP-tag.

pcDNA-*POLS*-TAP

Restriction enzymes *Not* I and *Xba* I were used to excise the TAP-tag coding sequence from a pcDNA3.1+ plasmid containing the TAP-tag (gift from Andrea Mikulasova). The HA-tag sequence was excised from pcDNA- *POLS*-HA using the same restriction enzymes. The TAP-tag sequence was ligated as described in section 2.3.14 into a pcDNA3.1+ derivative encoding *POLS*.

PGEX-6P-I-*POLS*-GST

Hela cell cDNA was used to amplify the appropriate regions listed in section 2.4.8 using the primers listed in Table 2.6. Restriction enzymes *Bam*H I and *Not* I were used to digest the inserts and PGEX-6P-I vector encoding the glutathione S-transferase (GST) tag. Each gene fragment was ligated as described in section 2.3.14 into the PGEX-6P-I vector.

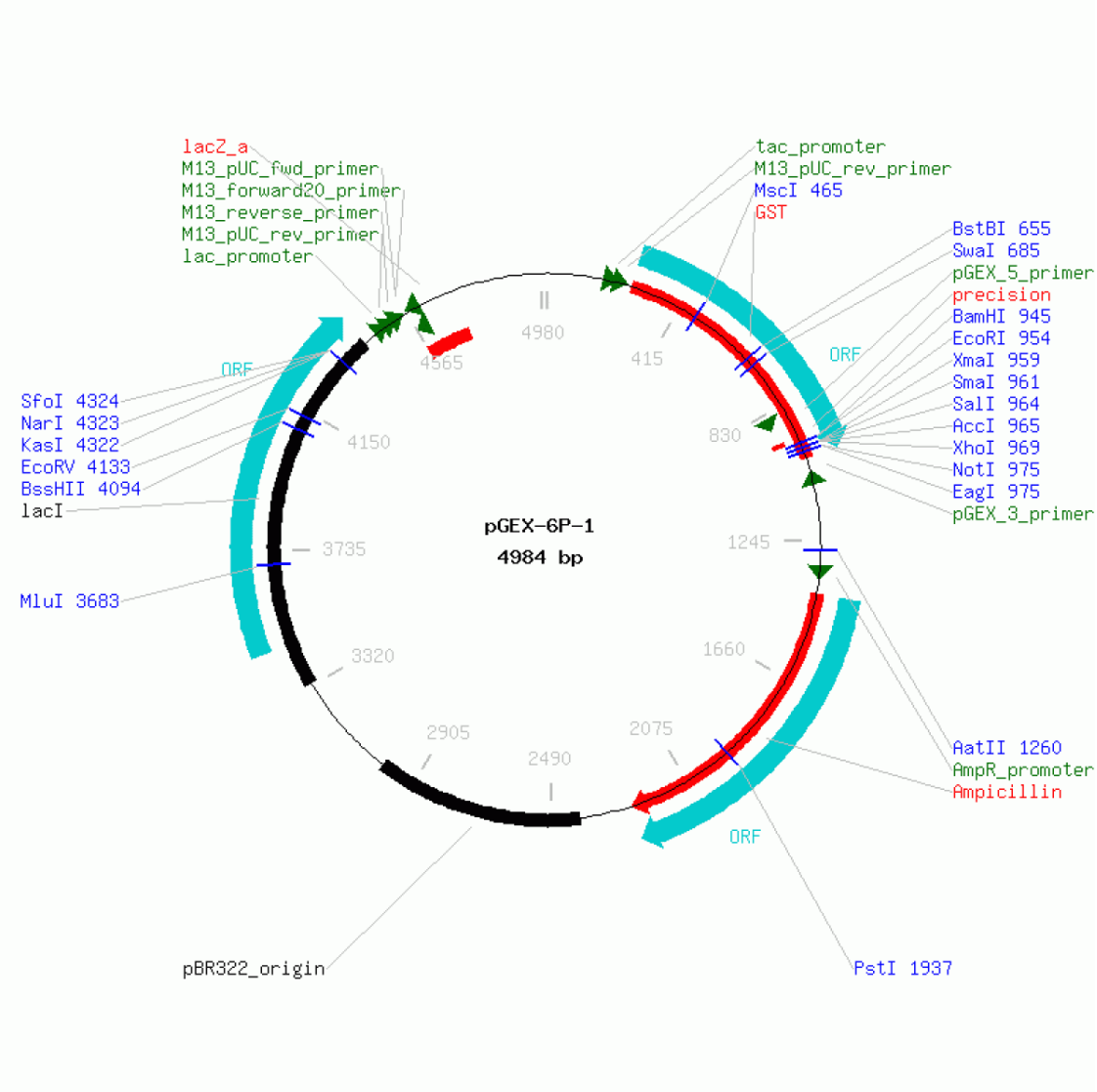


Figure 2.3 pGEX-6P-1 plasmid map (from Lablife web pages).
The PCR product and pGEX-6P-1 vector were digested with *Bam*H I and *Not* I before ligation.

2.4 Protein methods

2.4.1 Protein extraction from cultured human cells

Growth medium was removed and cells were washed with PBS. To 5-10x 10^6 cells, 60-80µl lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.6, 10 mM EDTA, 1% Triton X-100 (v/v), 1 mM PMSF, EDTA-free protease inhibitor tablet (Roche)/10 ml of solution) was added and cells were harvested immediately. The lysate was centrifuged at 16,000 x *g* for 25 min at 4ºC. The supernatant fraction was saved and the protein concentration was measured using the Bradford assay. BSA protein standards between 1-25 µg/ml were used to generate a standard curve. 1 ml 20% Bradford solution (Biorad) was mixed with each sample and incubated at RT for 2 min. After incubation the A_{595} was measured using an Eppendorf Biophotometer. A standard curve was plotted with A_{595} against concentration and sample concentration was determined from this curve.

2.4.2 Preparation of denaturing protein gels (SDS-PAGE)

A 10% resolving gel [375 mM Tris-Cl pH 8.8, 0.1% SDS (w/v), 10% acrylamide: bisacryamide (37.5:1, Biorad) and a 6% stacking gel (125 mM Tris-Cl pH6.8, 0.1% SDS (w/v), 6% acrylamide:bisacryamide (37.5:1, Biorad)], were polymerized by adding a final concentration of 0.03% APS and 0.2% TEMED. To each protein sample 2x SDS sample buffer was added and boiled briefly. This was followed by centrifugation at 16,000 x *g* for 1 min and loaded on a 10% acrylamide gel, prepared as stated above. A colorPlus prestained protein ladder (New England Biolabs) or rainbow molecular weight marker (Amersham Biosciences) was used as a size marker. Electrophoresis was performed at 80 V for 30 min and then at 150 V for 2-2.5 hr in 1X SDS running buffer [25 mM Tris, 190 mM glycine, 0.1% SDS (w/v)].

2.4.3 Western blotting

Transfer of protein onto Hybond^{TM} nitrocellulose membranes (Amersham) was carried out using a semi-dry blotter (Biorad) used at constant current of 0.04 A for 90 min. The membrane and blotting paper were soaked in semi-dry transfer buffer [48 mM Tris pH8.3, 39 mM glycine, 1.3 mM SDS and 10% methanol (v/v)] prior to transfer. Following the transfer the membrane was blocked with 5-10% low fat dried milk (w/v) in Tris-buffered saline (TBS), containing 0.1% Tween 20 (v/v) for 1 hr, and incubated with primary antibody diluted in TBS containing 0.1% Tween 20 (v/v) and 3% low fat dried milk (w/v) for 1 hr/overnight. The membrane was washed three times with TBS containing 0.1% Tween 20 (v/v) and 3% low fat dried milk (w/v). Secondary antibody conjugated to horseradish peroxidise diluted in TBS containing 0.1% Tween 20 (v/v) and 3% low fat dried milk (w/v) was added and incubated for 30 min. The membrane was washed twice with TBS containing 0.1% Tween 20 (v/v) and once in TBS. ECL (Santa Cruz Biotechnology) developing agent was added for 1 min before exposure to Kodak XAR-5 films (GE Healthcare). Antibodies used in this study are tabulated in Table 2.8. To quantify the amount of protein on blots AIDA software (Raytest GmbH) was used.

2.4.4 Staining total protein

Coomassie stain [0.1% Coomassie brilliant blue (w/v), 10% acetic acid (v/v), 20% methanol (v/v)] was used to visualize total protein in SDS-PAGE gels (1 hr, RT). The excess stain was removed by soaking gels for as long as necessary in destaining solution [10% acetic acid (v/v), 40% methanol (v/v)] at RT.

2.4.5 Stripping western blots

Membranes were soaked at 50ºC for 45 min in stripping buffer [2% SDS (w/v), 60mM Tris-HCl pH 6.8, 0.8% beta-mercaptoethanol (v/v)] with some agitation; following this membranes were rinsed in de-ionized water for 1-2 hr.

2.4.6 Cytoplasmic and nuclear fractionation

A 9 cm plate of U2-OS cells was washed in PBS solution, resuspended in 0.5 ml of lysis buffer I [10 mM Tris pH 7.5, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40 (v/v), EDTA-free protease inhibitor tablet (Roche)/10 ml of solution], and incubated for 5 min on ice. The suspension was under-layered with 0.5 ml lysis buffer I containing 24% sucrose (w/v). Lysates were then centrifuged at 16,000 x *g* for 20 min at 4ºC. The supernatant (cytoplasmic fraction) was transferred to a new tube. The pellet (nuclei) was re-suspended in RIPA buffer [15mM NaCl, 1% Triton X-100 (v/v), 0.5% sodium deoxycholate (w/v), 0.1% SDS (w/v), and 10% glycerol (v/v), EDTA-free protease inhibitor tablet (Roche)/10 ml of solution], vortexed and incubated on ice for 5 min. The suspension was centrifuged at 12,000 x *g*, at 4ºC for 10 min to obtain the nuclear fraction.

2.4.7 Tandem affinity purification of POLS interacting proteins

HEK-293T cells were transfected with pcDNA-*POLS*-TAP as in section 2.1.3., Cells were harvested 32 hr post-transfection, and were lysed in Tris lysis buffer [50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 0.5% Igepal CA-630 (v/v), 1 mM DTT, 1 mM PMSF, and one Complete Mini, EDTA-free protease inhibitor tablet (Roche)/10 ml of solution] on ice. *POLS*-TAP was purified by incubating 350 µl of cell lysate with 20 µl IgG FastFlow Sepharose slurry (Amersham) at 4ºC for 2 hr. The Sepharose beads were washed three times with wash buffer [10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5% Igepal CA-630 (v/v), 1 mM PMSF]. The

eluate from the beads was subjected to electrophoresis on a 4-12% bis-tris gel (Invitrogen). Gel pieces were digested with trypsin which was resuspended in the buffer provided (1µg/µl, Promega) and desalted on a C18 packed pipette tip. Samples were injected onto an Ultimate 3000 nano HPLC (DIonex) system coupled to an Orbitrap mass spectrometer (Thermo Electron). Peak lists were generated using DTAsupercharge and subjected to analysis by Mascot (Matrixscience). Data were searched against the IPI_human database. Lists of proteins present in pcDNA-TAP (control), *POLS*-pcDNA-TAP and proteins detected in both pcDNA-TAP and *POLS*-pcDNA-TAP were generated using Mascot software.

2.4.8 Polyclonal antibody production

A polyclonal antibody was raised against recombinant peptides of *POLS* encoded by nucleotides 4 to 244 of the open reading frame respectively. The GST-fusion plasmids were constructed as stated in section 2.3.15. These fusion proteins were expressed in *E.coli* BL21. Glycerol stocks of each transformant were frozen at -80ºC.

Protein induction (small scale)

A 15 ml single colony starter culture of each GST-fusion transformant was grown in LB containing ampicillin (100 µg/ml) until A_{600} reached 0.4-0.6. At this point the culture was divided into two tubes containing 5 ml each. Protein expression of one culture was induced with a final concentration of 100 µM IPTG. The cultures were grown overnight at 30ºC. The following day the cells were harvested by centrifugation at 16,000 x *g* for 1 min at 4ºC in 1.5 ml Eppendorf tubes. For each tube which contained 1.5 ml of culture 200 µl of Bugbuster (Novagen) containing 0.2 µl Benzonase (Novagen) was added and the pellet was resuspended. The suspension was rotated at 4ºC for 15 min. Samples were centrifuged at 16,000 x *g*, 4°C for 25 min. The supernatant and pellet fractions were saved and subjected to 10% SDS (w/v) PAGE as described in section 2.4.3. The gel was stained with coomassie

stain and the protein induction was monitored. The POLS-GST fusion protein was 40 kDa and PAPD5-GST fusion protein was 34.5 kDa. Both were present in both the supernatant and pellet fractions.

Protein induction (large scale)

From a 5 ml overnight pre-culture 2.5 ml was used to inoculate a 250 ml LB containing ampicillin (100 µg/ml). When the culture reached an A_{600} of 0.5, protein expression was induced by the addition of IPTG to a final concentration of 100 µM. The following day the cells were harvested by centrifugation at 4000 x *g* for 10 min at 4ºC in 50 ml Falcon tubes. For each tube which contained 50 ml of culture 2 ml of Bugbuster (Novagen) containing 2 µl Benzonase (Novagen) was added and the pellet was resuspended. The suspension was rotated at 4ºC for 15 min. Samples were centrifuged at 16,000 x *g*, 4°C for 25 min. The supernatant fraction was separated on a large 10% SDS (w/v) PAGE gel. Total protein was stained with Coomassie Brilliant Blue; after destaining, the gel was washed in de-ionized water and the band containing POLS-GST or PAPD5-GST was cut out and placed in 1.5 ml Eppendorf tubes with 10 µl of PBS to prevent the gel slices drying out in transit. To be noted the GST segment was excised before the gel slice was used to immunise female New Zealand White rabbits (CovalAb, France). Crude serum was provided 90 days after immunization and was tested for reactivity with the endogenous POLS/PAPD5 proteins by western blotting in comparison with the corresponding pre-immune serum.

2.4.9 Co-immunoprecipitation

Protein A or G beads were washed four times with wash buffer (as described in section 2.4.7). Cells were lysed with Tris-lysis buffer (as described in section 2.4.7); of this lysate 10% was saved for later analysis by western blotting. To pre-clear the lysate 10 µl protein A or G beads were added and the mixture was rotated at 4ºC for 1 hr. This suspension was

centrifuged at 2,000 x *g* for 4 min at 4ºC and beads were discarded. The lysate was divided between two tubes; to one tube antibody was added, at a concentration of 2 µg/ml and to the other tube no antibody was added and both were incubated overnight at 4ºC. The next day, 10 µl of the appropriate beads were added, rotated for 1 hr at 4ºC and centrifuged at 2,000 x *g* for 4 min at 4ºC. Supernatants from all fractions were saved for analysis by western blotting. Beads were washed four times with wash buffer (section 2.4.7). To the beads 40 µl of 2X SDS sample buffer was added and boiled for 5 min. Samples were centrifuged at 16,000 x *g* for 3 min at RT before loading on a 10% SDS-PAGE gel. Immunoblotting was performed as described in section 2.4.3.

siRNA duplex	**sense strand sequence**	**anti-sense strand sequence**
POLS-250	5’ GGCUAGGUACCAAUAAUAtt 3’	5’ UAUUAUUGGUACCGUAGCCtt 3’
POLS-250C	5’ UCAUAACUAUGCAAGCGGAtt 3’	5’UCCGCUUGCAUAGUUAUGAtt 3’
POLS-888	5’ CUUCGAUUAUGCCUACAAtt 3’	5’ UAUGUAGGCAUAAUCGAAGtt 3’
POLS-888C	5’ AUCAAUCCGUAUUAGCUUCtt 3’	5’ GAAGCUAAUACGGAUUGAUtt 3’
PAPD5-697	5’ AGCAUACUAGGUAGAAUAAtt 3’	5’ UUAUUCUACCUAGUAUGCUtt 3’
PAPD5-697C	5’ AAUAGGAUAGAUCAGACUAtt 3’	5’ UAGUCUGAUCUAUCCUAUUtt 3’
PAPD5-1306	5’ CUGUUCAGACGUUGACUUAtt 3’	5’ UAAGUCAACGUCUGAACAGtt 3’
PAPD5-1306C	5’ AUUACGUUGCCAGAUUGUCtt 3’	5’ GACAAUCUGGCAACGUAAUtt 3’

Table 2.1 siRNA duplex sequences used in this study.

Name of oligo	Sequence	ENSEMBL transcript ID
PAPD5 F	5’ GCTAGATGGCTACAGGCCATCAATG 3’	ENST00000436909
PAPD5 R	5’ GTTGTACCTTGGAAGCCTTTGCTGG 3’	ENST00000436909
POLS F	5’ ATTCAGACACACCGCCCTGCACAAC 3’	ENST00000230859
POLS R	5’ AGCCAGGAGCCATTATCTGCTGAGG 3’	ENST00000230859
U6 TUTase F	5’ CTCGGAACTAGCAGAGACCCCAAAG 3’	NM_022830.1
U6 TUTase R	5’ CAGGGAGCCACGAAGATCCCAACAA 3’	NM_022830.1
ZCCHC11 F	5’ CCATCTTATTCTCCTCAGCCTCAGCC 3’	ENST00000257177
ZCCHC11 R	5’ GAATCGGGCATTTCCCTGATGCATG 3’	ENST00000257177
ZCCHC6 F	5’ CGGACACTTCATGAAGGACTGTCCT 3’	ENST00000375963
ZCCHC6 R	5’ GCCTGATCTACACTGCTGTGTCCTG 3’	ENST00000375963
hGLD-2 F	5’ CGAGTTCGTCCGTTAGTGCTGGTGA 3’	NM_173797
hGLD-2 R	5’ CAGGACAGCAGCTCTTACAGGTAGT 3’	NM_173797
mtPAP F	5’ ACCAGGCCTCCGGATTTCAGTGTGA 3’	ENST00000263063
mtPAP R	5’ TCCAGGCACTTTCTCGGGCCAAATC 3’	ENST00000263063

Table 2.2 Oligonucleotides used in the expression screen of the human non-canonical poly(A) polymerases.

Name Q-PCR oligo	Sequence
POLS-F	5’ ATTCAGACACACCGCCCTGCACAAC 3’
POLS-R	5’ CCGCTTGTCTGCAAGGAACAGGAGC 3’
PAPD5-F	5’ CTGCCACACAGTCCAGCTCTAGTGA 3’
PAPD5-R	5’ GTTGTACCTTGGAAGCCTTTGCTGG 3’
Actin-F	5’ GAATGATGAGCCTTCGTGC 3’
Actin-R	5’ CTGGTCTCAAGTCAGTGTACA 3’
GAPDH-F	5’ CGGAGTCAACGGATTTGGTC 3’
GAPDH-R	5’ ATTTGCCATGGGTGGAATCA 3’
28SMF	5’ ACGAGCGCACGTGTTAGGACC 3’
28SMR	5’ TACCCAGGTCGGACGACCGA 3’

Table 2.3 Q-PCR oligonucleotides used in this study.

Name of oligo	Sequence
3'ETSF	5’ CGGCCTCTCCCGCCGACCG 3’
3'ETSFF	5’ TCCGTCCTTCCGTTCGTCTT 3’
actinF	5’ GAATGATGAGCCTTCGTGC 3’
actinFF	5’ TCTTAAGGAGAATGGCCCAGTCCTCT 3’
dT-adaptor	5’ GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT 3’

Table 2.4 Oligonucleotides used in the 3’ RACE screen.

Name of oligo	Sequence	Position
28SM3'-a (F1)	5’ TCGTAGACGACCTGCTTCTG 3’	4927
28SM3'-b (F2)	5’ AAGTCAGCCCTCGACACAAG 3’	4997
3'ETSR (R1)	5’ GAAGACGAACGGAAGGACGGA 3’	51
3'ETSRR (R2)	5’ CGGTCGGCGGGAGAGGCCG 3’	79
dT-adaptor	5’ GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT 3’	-
adaptor	5’ GACTCGAGTCGACATCGA 3’	-

Table 2.5 Oligonucleotides used to clone across the human 28S and 3’ETS boundary. Site of oligonucleotides relative to 5’ end of the mature 28S and 5’ end of 3’ETS are indicated in the third column**.**

Name of oligo	Sequence
N3’ETSF	5’ CGTTCGTCTTCCTCCCTCCCGGCCTCTCCCG 3’

Table 2.6 Oligonucleotide used in northern analysis.

Name of oligo	Sequence
POLS-GSTF1	5’ AGCGCTGGATCCTCCCCTTGTCCTGAAGAAGCA 3’
POLS-GSTR1	5’ AGCTCGCGGCCGCTCATACCGTAGCCTTGTCAAG 3’

Table 2.7 Oligonucleotides used to generate POLS-GST fusion protein.

Name	Type	Dilution	Source
α-Tubulin (TAT-1)	Mouse monoclonal	1:10000	Keith Gull
α-TOPO2α	Rabbit polyclonal	1:2000	Ian Hickson
α –HA.11	Mouse monoclonal	1:1000	Cambridge Bioscience
α -mouse-HRP	Goat polyclonal	1:2000	Sigma
α -rabbit-HRP	Goat polyclonal	1:2000	Sigma
α -RAD50	Mouse monoclonal	1:250	Sigma
POLS-crude serum	Rabbit polyclonal	1:2000	This study

Table 2.8. Antibodies used for western blot in this study.

Name	Type	Dilution	Source
CPSF	Rabbit polyclonal	2 µg/ml	Abcam
BRCA1	Goat polyclonal	2 µg/ml	Santa Cruz Biotechnology
Ku 80	Mouse monoclonal	2 µg/ml	Fisher Scientific
α –HA.11	Mouse monoclonal	2 µg/ml	Cambridge Bioscience

Table 2.9 Antibodies used for co-immunoprecipitation in this study.

Common buffers, solutions and media	**Recipes**
20X SSC:	3 M NaCl**,** 0.3 M sodium citrate, pH 7.0
1X TE	10 mM Tris, 1 mM EDTA, pH 8.0
1X PBS	137 mM NaCl, 2.7 mM KCl, 100 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4
1X TBE	10 mM Tris, 10 mM H_3BO_3, 2 mM EDTA, pH 8.0
50X Denhardt's	1% Ficoll [type 400, (w/v)], 1% polyvinylpyrrolidone (w/v), 1% bovine serum albumin (w/v)
1X SSPE	150 mM NaCl**,** $NaH_2PO_4.H_2O$, 10mM EDTA, pH 7.4
1X MOPS	20 mM MOPS pH 7.0, 8 mM sodium acetate pH 5.2,1 mM EDTA pH 8.0
Luria-Bertani (LB) broth	1% Bacto-tryptone (w/v), 0.5% Bacto-yeast extract (w/v), 1% NaCl (w/v)
LB agar	LB (as above) with 1.5% Bacto-agar

Table 2.10 Recipes for common buffers, solutions and media used in this study.

3 The cellular functions of POLS and PAPD5

3.1 Introduction

The localization of a protein can be very informative about of its function. The ncPAPs studied to date are localized to the nucleus, cytoplasm or mitochondrion, indicating diverse biological functions. Of the seven human ncPAPs, the subcellular localization of the following are known (Table 3.1); U6TUTase in the nucleus (Trippe *et al.*, 2006; Mellman *et al.*, 2008), ZCCHC11 in the cytoplasm (Kwak *et al.*, 2008; Schmidt *et al.,* 2011), mtPAP in mitochondria (Tomecki *et al.*, 2004; Nagaike *et al.*, 2005), hGLD-2 (Richter, 2007) and ZCCHC6 (A. Mikulasova, personal communication; Schmidt and Norbury, 2010) in both the nucleus and cytoplasm.

The localization human PAPD5 remains controversial at present. The Ensembl database shows two transcripts containing open reading frames (ORF) of 619 aa and 698 aa for *PAPD5*. In preparative work for this thesis a further splice variant encoding a protein of 665 aa was identified (Figure 6.2). N-terminal myc-tagged (pcDNA4-myc-Flag) PAPD5 was reported to localize predominantly to the cytoplasm but was also present in the nucleus in some cells as visualized by confocal microscopy (Mullen and Marzluff, 2008). To be noted, a sequence encoding a 572 aa variant of PAPD5 (Q8NDF8) was cloned into the pcDNA4 in the latter study. A very recent study reports PAPD5 to be localized in the nucleus including the nucleoli (Lubas *et al.,* 2011). In the latter study PAPD5 was fused to the C-terminal EGFP tag in Hela cells or N-terminal EGFP tag in HEK-293T cells. PAPD5 was also found to co-purify with factors associated with nuclear function, consistent with the functions of other related orthologues. Thus, the authors claim their clone to be expressing physiologically

relevant PAPD5. UniProtKB/Swiss-Prot database lists four isoforms for Q8NDF8 produced by alternative splicing, coding for proteins of 572, 370, 465 and 585 aa. The implication from the alignment shown in the Lubas *et al.* (2011) study is that a PAPD5 sequence of 572 aa was used for localization purposes. The reason for these apparently discrepant findings is therefore currently unclear. There are no reports of the localization of POLS to date.

On the basis of multiple sequence comparisons *POLS* and *PAPD5* have been classified as human orthologues of *TRF4* and *TRF5*. Trf4 and Trf5 are documented as localizing to the nucleus, with Trf5 enriched in the nucleolus (Walowsky *et al*., 1999; Houseley and Tollervey, 2006). A slow growth phenotype is documented for *trf4* or *trf5* single mutants, while the double mutant of *trf4 trf5* has a lethal phenotype. The *S.pombe* Trf4/Trf5 homologue Cid14 is a constitutively nucleolar protein. Cells lacking Cid14, although viable, are prone to elevated levels of chromosome missegregation (Win *et al*., 2006). *D. melanogaster* DmTRF4-1 localizes to the nucleolus and disruption of DmTRF4-1 results in a lethal phenotype. Knock-down of DmTRF4-2 does not affect the viability of the flies (Nakamura *et al.,* 2008). The *C.elegans* Trf4/Trf5 orthologue GLD-4 is found in cytoplasmic P granules and plays an essential role in the absence of GLD-2 in early meiotic progression (Schmid *et al.*, 2009). Hence, these Trf4 and Trf5 orthologues are associated with diverse biological functions consistent with their diverse localizations.

Species	Name	Location	Function
S. cerevisiae	Trf4	N	tRNA, sn/snoRNA, CUT and rRNA degradation. rRNA processing.
	Trf5	N	rRNA processing and degradation
S. pombe	Cid1	C	mRNA degradation
	Cid11	N and C	unknown
	Cid12	N	gene silencing and chromosome segregation
	Cid13	C	cytoplasmic mRNA activation
	Cid14	N	rRNA polyadenylation and degradation, maintain nucleolar structure, gene silencing
	Cid16	unknown	unknown
C. elegans	GLD-2	C	required to complete spermatogenesis and oogenesis.
	GLD-3	C	gene silencing
	GLD-4	C	unknown
Xenopus	GLD-2	N and C	oocyte maturation
Mus Musculus	GLD-2	C	oocyte maturation
Drosophila	GLD-2 (Wisp)	C	oocyte maturation
	DmTRF4-1	Nu	Polyadenylation of snRNA
	DmTRF4-2	unknown	unknown
H.sapiens	PAPD5	N and Nu	degradation of PROMTs, polyadenylation of 5' ETS rRNA
	POLS	N and C	unknown
	U6 TUTase	N	U6 processing, mRNA cleavage
	ZCCHC11	C	pre-miRNA degradation
	ZCCHC6	N and C	unknown
	hGLD-2	N and C	cytoplasmic mRNA activation, miRNA stabilization
	mtPAP	M	mtRNA adenylation, histone mRNA degradation

Table 3.1 Non-canonical poly(A) polymerases in various species, their localization and functions.

N:nucleus, Nu:nucleolus, C:cytoplasm, M:mitochondria

3.2 Aims

The catalytic and PAP-associated domains of Trf4, Trf5, POLS and PAPD5 are well conserved (Figure 1.7). In this study, I set out to characterize the cellular functions of POLS and PAPD5. The degree of conservation and extent of divergence from their orthologues in other species were investigated.

3.3 POLS localization

As a first step towards characterization of the cellular functions of POLS, its subcellular localization was determined.

3.3.1 POLS expression in transient transfectants

In the absence of an antibody capable of detecting endogenous POLS, transiently transfected HA epitope-tagged POLS (Figure 3.1), and TAP-tagged POLS (Figure 3.2) were analysed in U2-OS cells by immunofluorescence microscopy. The constructions of these plasmids are described in the methods section 2.3.15. U2-OS cells were used because they are an osteosarcoma cell line with a flat morphology aiding clear visualization of the cytoplasm and nucleus. In the majority of cells predominantly cytoplasmic staining was observed with significant, but less intense, staining in the nucleus with both tagged proteins. Among cells transiently transfected with HA-tagged POLS, nuclear speckle staining was detected in 0.2% cells. About 2% cells were noted to contain only intense cytoplasmic staining although it is possible that the minor nuclear staining visible in the majority of cells was not seen in these cells due to their (flattened morphology. In transient transfection experiments some cells take up DNA while others do not; Figure 3.1 and Figure 3.2 are

representative of the localization detected under these circumstances. Cells that fail to take up DNA provide a convenient internal control for the specificity of anti-HA or anti-TAP immunofluorescence. To calculate the percentage of cells showing specific staining, 1000 cells were counted for each treatment. The TAP-tag is capable of binding to IgG as it contains protein A, resulting in a signal with any IgG antibody. In Figure 3.2 the immunofluorescene signal detected when TAP-specific antibody was used was above this level.

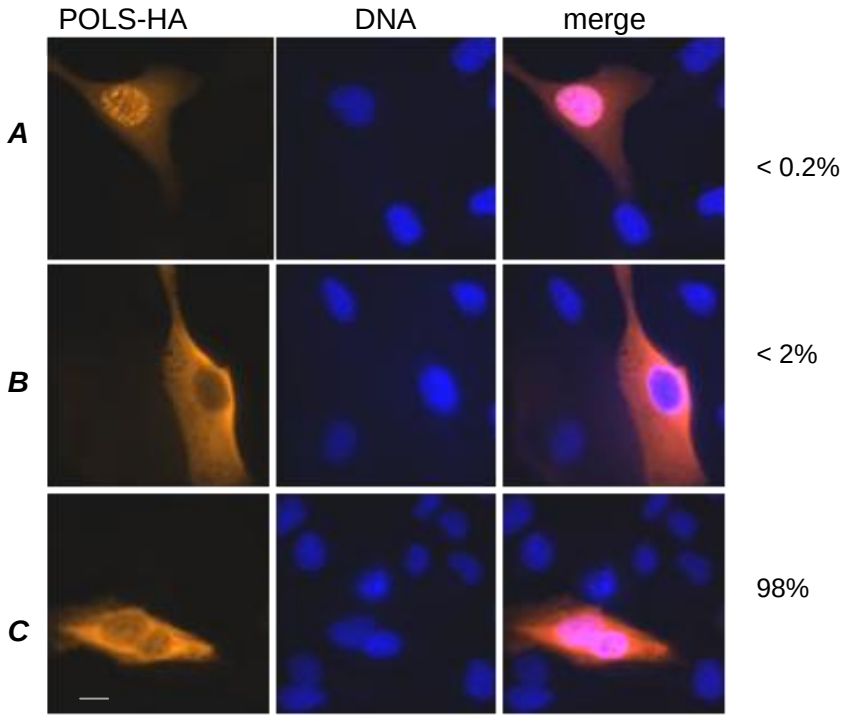


Figure 3.1 Subcellular localization of HA-tagged POLS in transiently transfected U2-OS cells.

A Cells expressing POLS-HA in nuclear speckles. ***B*** Cells expressing POLS-HA in the cytoplasm. ***C*** Cells expressing POLS-HA in the nucleus and the cytoplasm. Bar 20µm.

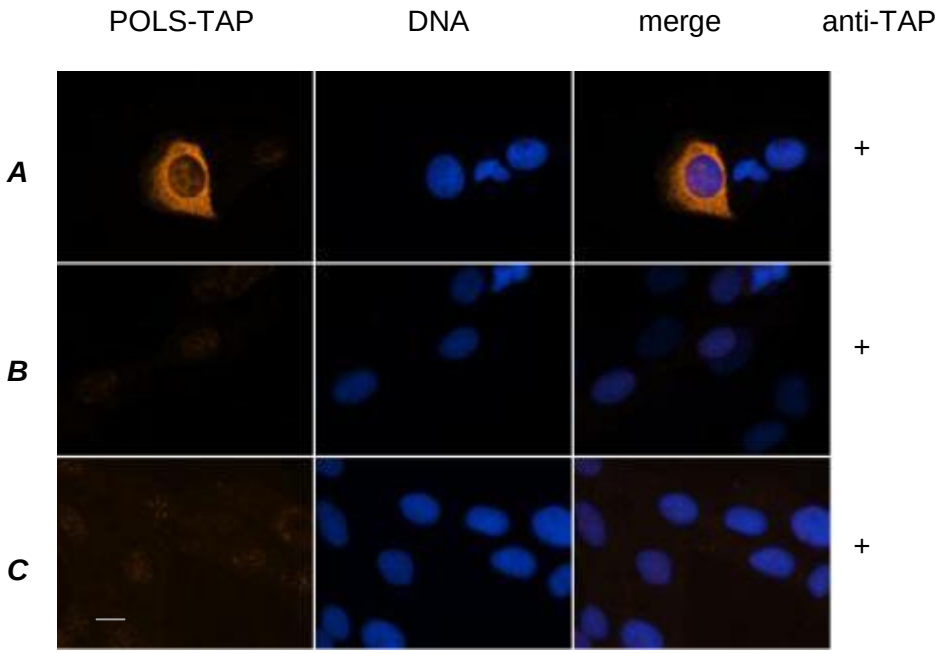


Figure 3.2 Subcellular localization of POLS-TAP in U2-OS cells.

A Transiently transfected POLS-TAP expression. ***B*** Cells transfected with pCDNA vector containing the TAP–tag stained with primary and secondary antibodies. ***C*** U2-OS cells stained with primary and secondary antibodies. Bar 20µm.

To quantify the nuclear and cytoplasmic distribution of POLS-HA by an independent method, biochemical fractionation was performed on U2-OS cells transiently transfected with HA epitope-tagged POLS. The HA tagged POLS was used to perform the fractionation instead of TAP-tagged POLS as the TAP-tag could give a non-specific signal by attaching to IgG. Equivalent proportions of nuclear and cytoplasmic fractions were compared by western blotting (Figure 3.3). α-Tubulin and topoisomerase IIα served as controls for cytoplasmic and nuclear fractions respectively. Note that the cytoplasmic fraction is contaminated with the nuclear fraction. This however does not negate the conclusion that POLS is present in both compartments, as cytoplasmic levels of POLS-HA were much higher than those of contaminating topo IIα in the cytoplasmic fraction. Additionally the nuclear fraction was devoid of highly expressed tubulin, suggesting that the nuclear presence of POLS was not a result of cytoplasmic contamination. This is reassuring as only 3.7% localized to the nuclear compartment as quantified by AIDA software (methods section 2.4.3). The cytoplasmic/nuclear fractionation substantiated the immunofluorescence data, further suggesting that POLS is predominantly cytoplasmic, with significant but less intense staining in the nucleus.

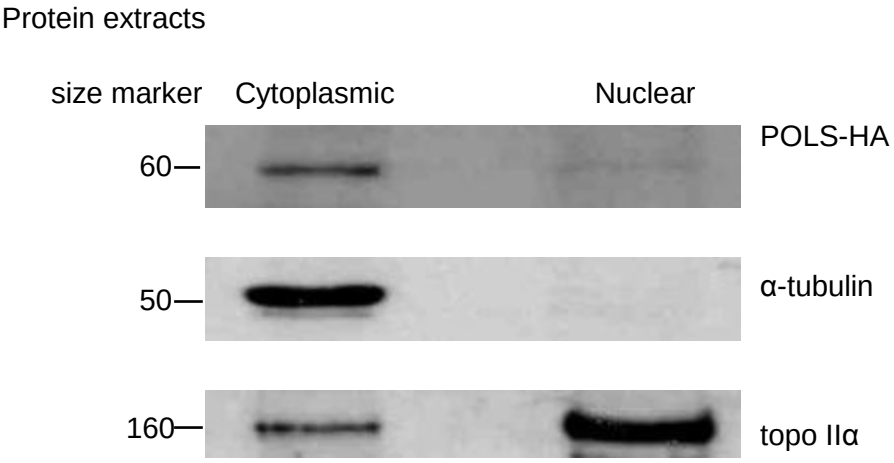


Figure 3.3 Cytoplasmic and nuclear fractionation of HA epitope-tagged POLS in U2-OS cells.

POLS-HA was analysed by western blotting with an HA-specific antibody. α-tubulin and topoisomerase IIα (topo IIα) served as controls for cytoplasmic and nuclear localization, respectively.

3.3.2 POLS expression in stable transfectants

Individual cells in a transiently transfected population take up different amounts of DNA, resulting in varied expression. In order to observe the expression of more physiologically relevant levels of POLS, stable U2-OS transfectants expressing HA epitope-tagged POLS were generated by G418 selection. Individual transfectants were found to be expressing different amounts of POLS in comparison to α-tubulin (Figure 3.4). The level of POLS-HA expression of these stable transfectants in comparison to endogenous POLS is unknown at present.

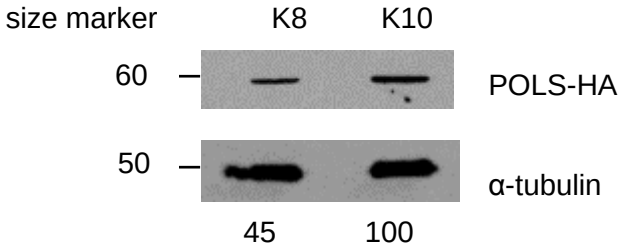


Figure 3.4 POLS-HA expression in stable transfectants.
Expression levels compared to α-tubulin: K10>K8. Amount of POLS-HA quantified as a percentage compared to α-tubulin is shown below the blot.

In the stable transfectant K10 with a high expression level of POLS (222% more than transfectant K8) the majority of cells contained cytoplasmic staining with significant, but less intense, staining in the nucleus (Figure 3.5), agreeing with the localization seen in the majority of transient transfectants. Cells in cytokinesis displayed intense accumulation of POLS in the densest region of the cytoplasm due to cell rounding (Figure 3.5***B***).

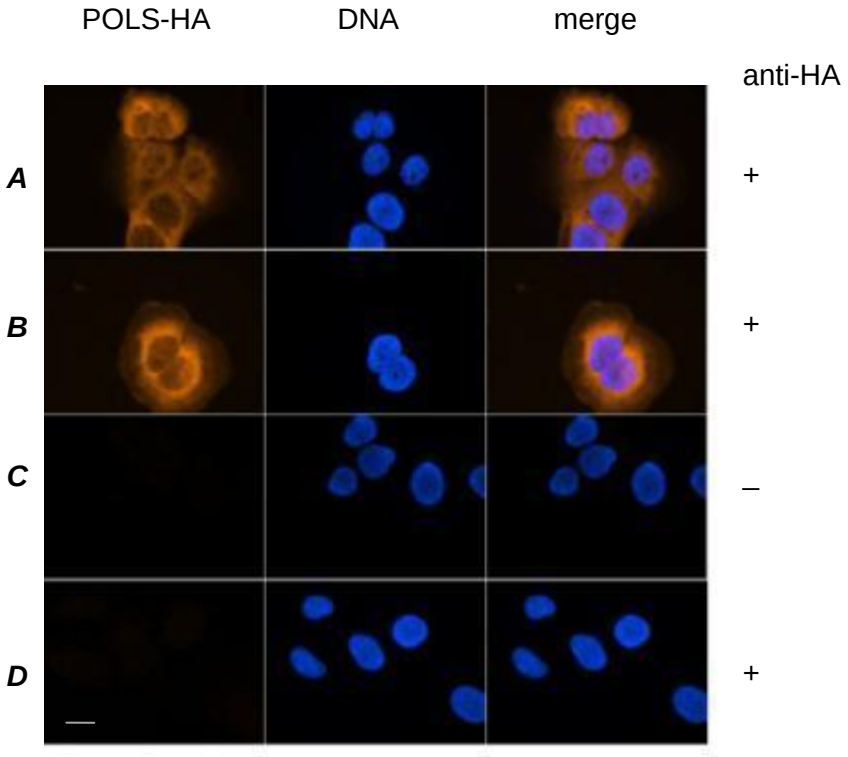


Figure 3.5 Subcellular localization of POLS-HA in U2-OS cell stable transfectant K10.

A, B Stable transfectant expressing POLS-HA. ***B*** Cell undergoing cytokinesis. ***C*** Stable transfectant with no primary antibody. ***D*** U2-OS parental cells with primary and secondary antibodies. Bar 20µm.

In stable transfectant K8, which expressed 45% of the level of POLS-HA seen in K10, nuclear speckles were visible (Figure 3.6 ***I-L***). In 50% cells these speckles were found to be concentrated in the nucleolus (Figure 3.6 ***A-D***). It remains to be confirmed by the use of a nucleolar marker that these structures are indeed nucleoli. They were assigned as nucleoli by the hollow distinct structure visible in DAPI stained nuclei. A low level of speckled staining was also noted in the cytoplasm.

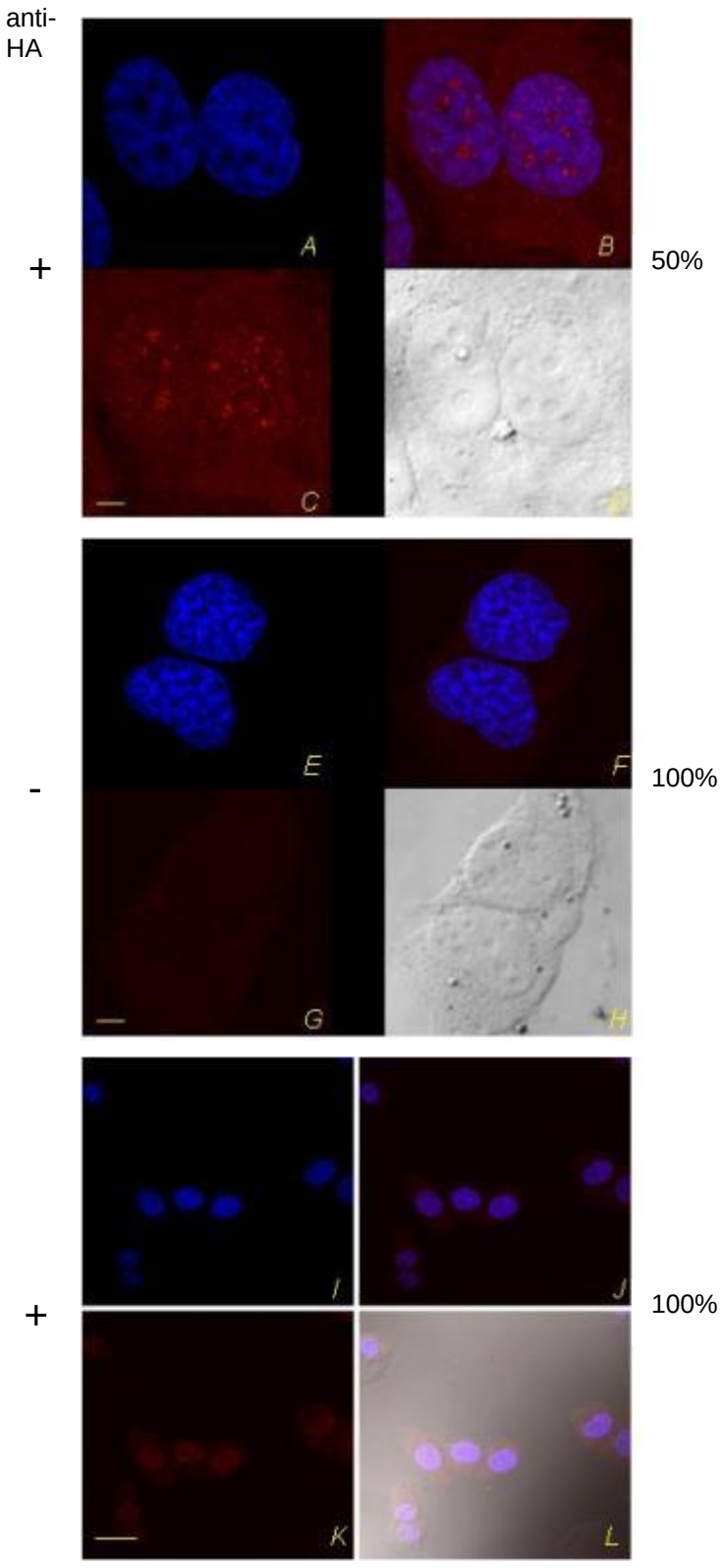


Figure 3.6 Subcellular localization of POLS-HA in U2-OS cell stable transfectant K8.

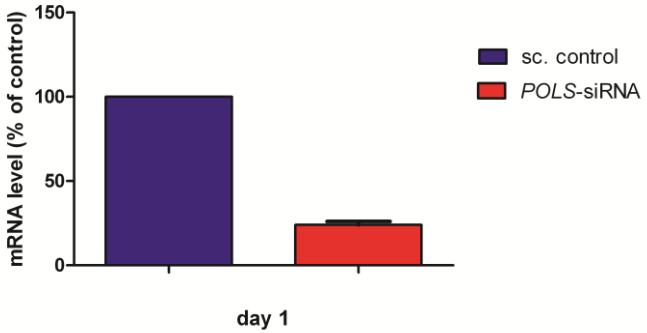
A-H Images visualized by a z-stack projection. Bar 5 µm. ***I-L*** confocal images Bar 20 µm. ***A,E,I*** nuclei stained with DAPI, ***B*** POLS-HA expressed in nuclear speckles accumulating in the nucleolus overlayed with DAPI stained nuclei. ***C*** POLS-HA expressed in nuclear speckles accumulating in the nucleolus, ***D,H*** DIC images, ***F*** POLS-HA stable transfectants with no primary antibody overlayed with DAPI stained nuclei, ***G*** POLS-HA stable transfectant with no primary antibody, ***J*** POLS-HA expressed in nuclear speckles overlayed with DAPI stained nuclei, ***K*** POLS-HA expressed in nuclear speckles ***L*** POLS-HA expression overlayed with DAPI stained nuclei and DIC images.

3.4 Individual knockdown of *POLS* or *PAPD5*

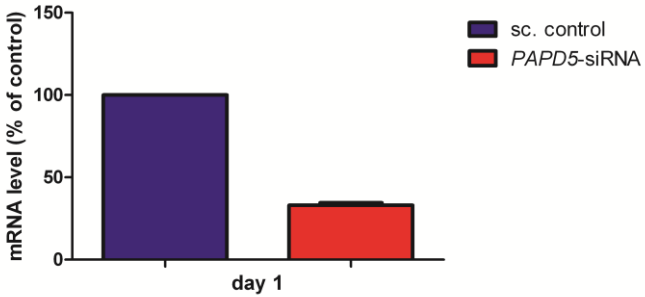
To investigate the cellular functions of POLS and PAPD5, the corresponding mRNAs were knocked down using siRNA duplex transfections (Figure 3.7) and the effects were monitored *via* cell viability assays. The siRNA duplexes used to knock-down POLS and PAPD5 are listed in Table 2.1. In U2-OS cells, 24 hr post-transfection 76% *POLS* (Figure 3.7***A***) and 67% PAPD5 (Figure 3.7***B***) mRNA knock-down was detected compared to the scrambled control by q-PCR. In a stable U2-OS transfectant (K8), POLS-HA protein was reduced to below the limit of detection by 11 hr post-siRNA knock-down, indicating that POLS-HA has a short half-life (Figure 3.7***C***). No data for PAPD5 protein is available due to the lack of an antibody capable of detecting endogenous PAPD5. My attempt to amplify the N-terminus of PAPD5 based on the mouse sequence was unsuccessful. Thus, no tagged construct of PAPD5 was available to study PAPD5 protein levels. The entry of the human N-terminus PAPD5 sequence in public databases is very recent.

.

A. **Quantitative PCR (Q-PCR) assay**



B. **Quantitative PCR (Q-PCR) assay**



C. **Western blot**

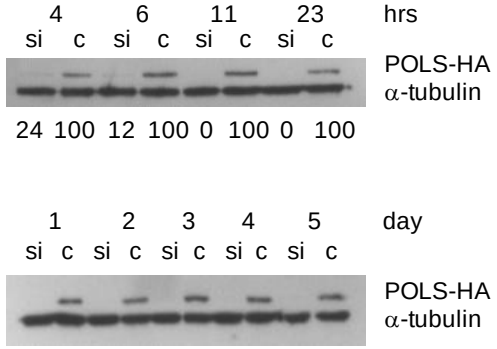


Figure 3.7 *POLS* or *PAPD5* knockdown.

A *POLS*, ***B*** *PAPD5* mRNA levels quantified by q-PCR in U2-OS cells. Data are indicative of three biological replicates performed in triplicate. ***C*** Western blot showing POLS-HA protein levels. si: total cell lysate of POLS siRNA treated cells, c: total cell lysate of control cells. The amount of POLS-HA is quantified as a percentage compared to the scrambled control at each time point. Data are indicative of two biological replicates. A stable U2-OS transfectant K8 was used to study the protein down-regulation.

3.5 Down-regulation of *POLS* and/or *PAPD5* results in a slow growth/lethal phenotype

3.5.1 Individual down-regulation of *POLS* or *PAPD5*

Down-regulation of POLS or PAPD5 was achieved as described in section 2.1.3. Mitochondrial reductase activity, as a surrogate measure of biomass, was monitored using an MTT-based assay. The assay was performed at 24 hr intervals over a 4-day time course on scrambled control and *POLS* or *PAPD5* siRNA-treated U2-OS cells. *POLS* down-regulated cells accumulated a lower biomass compared to control cells (Figure 3.8***A***), as did *PAPD5* down-regulated cells (Figure 3.8***B***). The MTT-based assay measures total reductase activity rather than cell numbers. Hence, direct cell counting in the presence of trypan blue was used to monitor changes in cell number and distinguish between viable and dead cells. This assay showed that *POLS* down-regulated populations had fewer viable cells compared to the control (Figure 3.8***D***); The same was seen in the *PAPD5* down-regulated population (Figure 3.8***E***). However, no increase in dead cells was recorded in either case.

3.5.2 Simultaneous down-regulation of *POLS* and *PAPD5*

To investigate the potentially redundant cellular functions of *POLS* and *PAPD5*, the corresponding mRNAs were simultaneously knocked down using siRNA duplexes and the effects were monitored *via* cell viability assays. A 70% down-regulation of *POLS* and a 75% down-regulation of *PAPD5* mRNA compared to the scrambled control resulted in slower growth as measured by the MTT assay compared to the control population (Figure 3.8***C***). Trypan blue assays showed a decrease in viable cells in the gene down-regulated population compared to the control population (Figure 3.8***F***). The trypan blue assay recorded a significant difference in the number

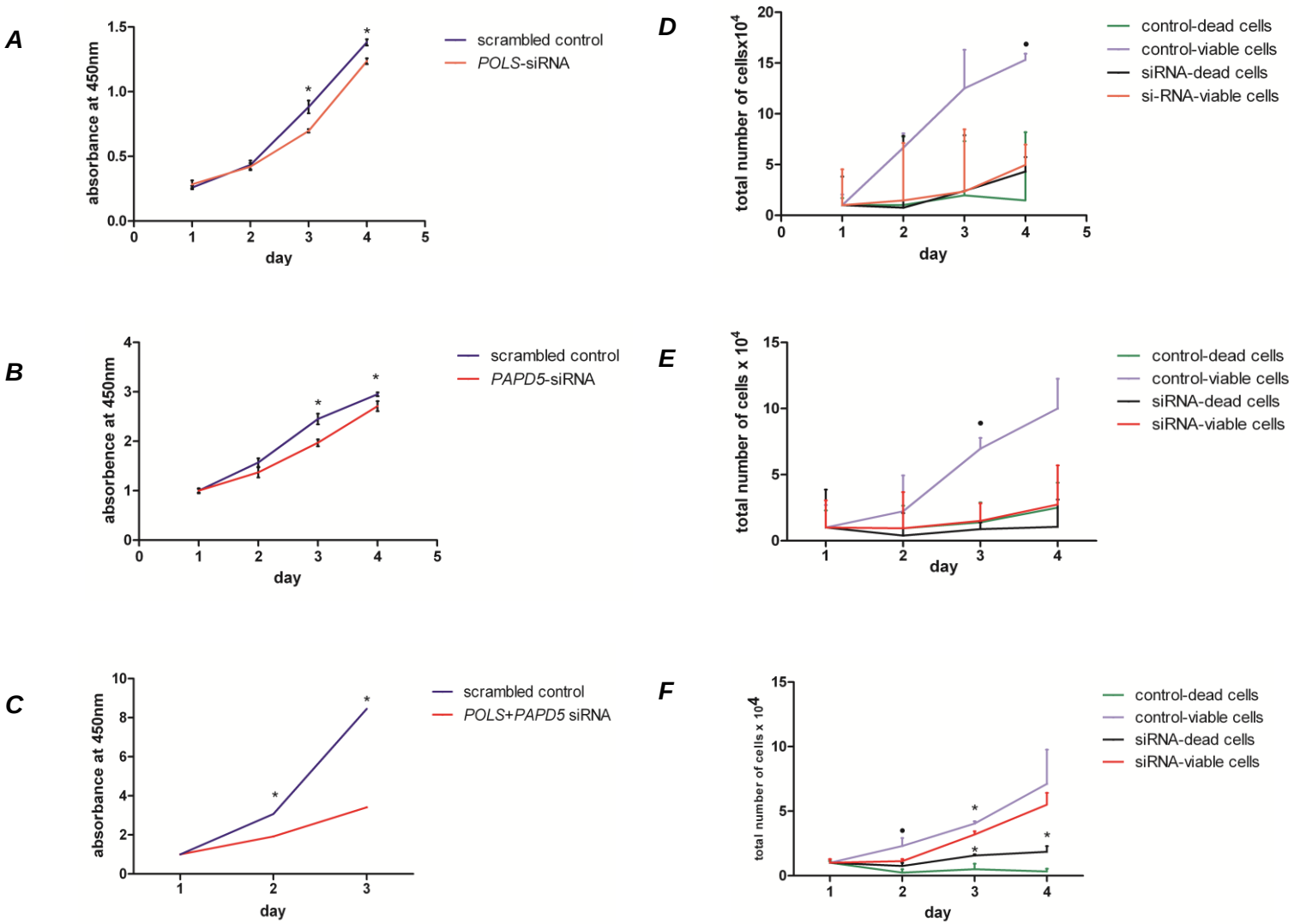


Figure 3.8 Depletion of POLS or/and PAPD5 results in a slow growth phenotype.

A,B,C. EZ4U assay was performed for 4 days on scrambled control and POLS or/and PAPD5 siRNA treated U2-OS cells. ***D,E,F.*** In parallel trypan blue assays were carried out to detect viable and dead cells. Results for the cell growth assay are of three independent experiments performed in duplicate. For the trypan blue assay two independent experiments were performed in duplicate. Results were normalized to 1 at day 1 to compare the increase in biomass or cell number. *p<0.01 ; •p<0.05

of dead cells between the DKD (POLS+PAPD5) and control populations (Figure 3.8***F***).

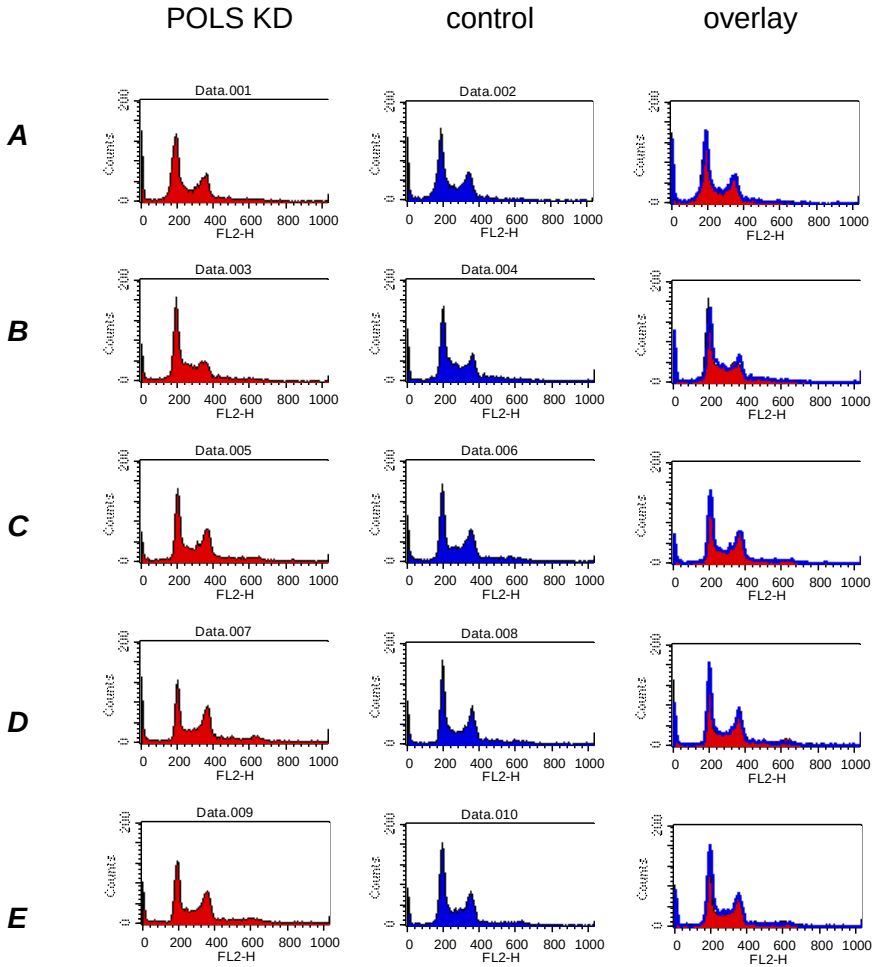
Flow cytometric analysis was performed to determine if changes in the cell cycle were responsible for the reduced proliferation rate observed in the gene down-regulated cells compared to the control cells.

3.5.3 Flow cytometric analysis of POLS knock-down

Cell cycle profiles of POLS down-regulated cells were compared to the scrambled control at 18, 24, 48, 72 and 96 hr (Figure 3.9). No significant difference was noted in the cell cycle profiles of POLS down-regulated *vs* control cell populations.

3.5.4 Flow cytometric analysis of POLS and PAPD5 double knock-down

Flow cytometric analysis was performed to observe the cell cycle profile when both genes were down-regulated at day 3, as the greatest growth difference was observed at this point (Figure 3.8***C***). Simultaneous down-regulation of *POLS* and *PAPD5* reduced the number of cells in G1 while an accumulation of cells in G2/M was noted (Figure 3.10). This experiment was performed only once.



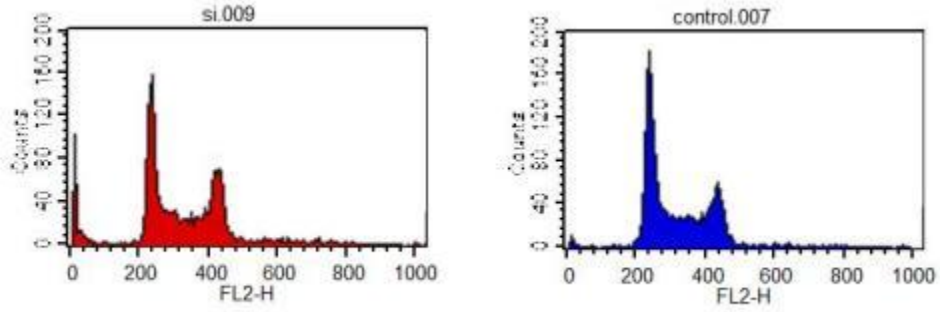
cell cycle phase	Percentage of cells	
	siRNA	control
G1-18hr	42	40
S-18hr	17	18
G2-18hr	24	24
G1-24hr	41	34
S-24hr	23	23
G2-24hr	22	25
G1-48hr	31	35
S-48hr	20	19
G2-48hr	31	29
G1-72hr	24	32
S-72hr	17	19
G2-72hr	32	29
G1-96hr	31	38
S-96hr	16	20
G2-96hr	28	30

Figure 3.9 Cell cycle analysis of POLS down-regulated *vs* scrambled control cells.

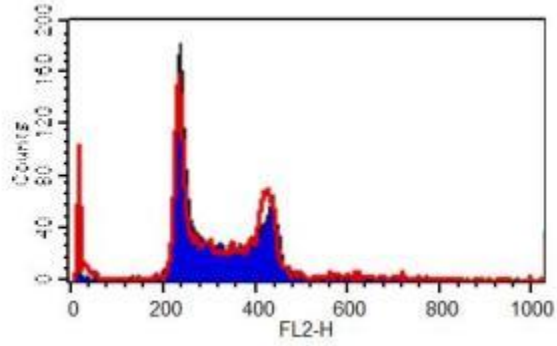
Red: *POLS* knocked down(KD), Blue: control**.** All histograms showing DNA content ***A***. 18 hr, ***B***. 24 hr ***C***. 48 hr, ***D***. 72 hr and ***E***. 96 hr. Table showing percentage of cells in the different phases of the cell cycle for three biological replicates.

Flow cytometric analysis of (POLS+PAPD5) DKD

A



B



C

phase	Percentage of cells	
	siRNA	control
<G1	6	2
G1	33	42
S	29	29
G2/M	22	20

Figure 3.10 Flow cytometric analysis of cell cycle profile of simultaneously down-regulated POLS and PAPD5 cells *vs* scrambled control.

Red: POLS+PAPD5 DKD, blue: control. Analysis was performed 72 hr post transfection. The greatest growth difference was observed at 72 hr as measured by the MTT-based assay (Figure 3.8***C***). ***A***. Plot showing DNA content of individual cells. ***B***. Overlay of POLS and PAPD5 down-regulated and control cells. ***C***. Table showing the percentage of cells in each phase of the cell cycle.

3.6 Discussion

The detection of POLS in the cytoplasm and nucleus in both transient transfectants and the stable transfectant K10 was unexpected, as with the exception of *C.elegans* GLD-4, Trf4/5 orthologues in other species are localized in the nucleus or/and nucleolus. It is conceivable that the tags used in this study could affect the localization. In the cancer cells used in this study the localization may differ from that in normal cells. The variation in localization detected in the POLS-HA transient transfection may be representative of cells in different stages of the cell cycle. The subcellular distribution could indicate that POLS has diverged from its orthologues, gaining functions in both the cytoplasm and nucleus during evolution.

The difference in the intracellular level of POLS expression or clonal variation may account for the difference in the localization of POLS-HA in stable transfectants K10 and K8. Under normal physiological conditions POLS could be localizing in the nucleus with an accumulation of speckles in the nucleolus in some cells as seen in the stable transfectant K8. When POLS is over-expressed, as possibly is the case with stable transfectant K10, it could be sequestered in the cytoplasm until required for functions in the nucleus. If on the other hand transfectant K10 is representative of expression of POLS under physiological conditions, POLS might have functions in both the nucleus and cytoplasm. For PAPD5 the clone used by Lubas *et al*. (2011) was claimed to be expressing physiologically relevant levels of PAPD5, as substantiated by their immunoprecipitation data, which showed PAPD5 to be associated with nuclear proteins. The PAPD5 clone used by Mullen *et al*. (2008) may have been expressing much greater quantities compared to endogenous PAPD5, resulting in cytoplasmic and nuclear localization. The variation in localization observed using transiently transfected POLS-HA may have been due to cells taking up different amounts of DNA, and hence expressing various levels of POLS-HA. The specificity of the polyclonal antibody raised to detect POLS has just been confirmed in this study, enabling comparison of levels of stable transfectant expression to endogenous POLS.

A slow growth phenotype was observed following knock-down of *POLS* or/and *PAPD5.* By flow cytometric analysis no differences in the cell cycle profiles were noted for POLS down-regulated compared to the control cells. The DKD (POLS+PAPD5) showed an increase in the proportion of cells in G2/M phases of the cell cycle compared to the control at 72 hr. There are many potential reasons for the altered cell cycle distribution.

TRF4 has been shown to be required for mitotic chromosome segregation and establishment of sister chromatid cohesion. Trf4 interacts with Smc1, a protein required for sister chromatid cohesion. As cohesion occurs at multiple sites along each chromosome, chromosomes with some defective cohesion sites segregate normally enabling the *trf4* single mutant to be viable. Slow proliferation observed with *POLS* and/or *PAPD5* down-regulated cells may be explained by cells completing the cell cycle more slowly due to defective chromosome segregation. If cells are taking longer to complete mitosis, accumulation of cells in G2/M would be visible.

Given the well-defined roles of Trf5 in rRNA metabolism and to a lesser extent Trf4, it is conceivable that the slowed growth is an indirect consequence of altered rRNA metabolism in POLS or/and PAPD5 down-regulated cells. The accumulation seen in G2/M when both *POLS* and *PAPD5* are down-regulated simultaneously could be due to cells progressing more slowly into mitosis compared to the control due to a reduced efficiency of translation. Impairment of rRNA turnover could lead to slow growth or death due to impaired ribosome function leading to eventual inhibition of translation. Human and murine PAPD5 have been reported to be involved in the degradation of incomplete RNA polymerase I transcripts (Lubas *et al.*, 2011; Shcherbik *et al.*, 2010), Ribosomal RNA processing may be impaired in PAPD5 knock-down cells. Accumulation of 47S rRNA was noted in the murine PAPD5 down-regulated cells. This would lead to decreased ribosome biogenesis, and hence translation. If human POLS is also associated with degradation of rRNA transcripts and both genes act in the same pathway, down-regulation of these genes could be detrimental to the cell.

The *trf4 trf5* double mutant has a lethal phenotype. Therefore, *TRF4* and *TRF5* are redundant for an essential function (Castano *et al*., 1996). Simultaneous down regulation of *POLS* and *PAPD5* led to reduced cell proliferation with increase in the number of dead cells at day 3 as measured by the trypan blue assay and flow cytometric analysis. Cohesin defects could be so severe that sister chromatid cohesion is disrupted in some cells, leading to eventual cell death. Thus, in human cells knock-down of both *POLS* and *PAPD5* is lethal making these essential genes too.

4 Biochemical functions of POLS and PAPD5

4.1 Introduction

S.cerevisiae Trf4 and Trf5 are responsible, as components of TRAMP complexes, for the transient polyadenylation of substrates of the 3'-5' exosome-mediated RNA decay and processing pathways in the nucleus. Trf4 and Trf5 have been shown to adenylate rRNA, tRNA, snRNA, snoRNA and cryptic pol II transcripts (Houseley and Tollervey, 2006; Kadaba *et al.*, 2006; Vanacova *et al.*, 2005; Wyers *et al.*, 2005). Recent reports indicate that these tails are 4-5nt in length (Jia *et al.,* 2011; Wlotzka *et al.*, 2011). The oligo(A) tails added by the TRAMP complex are too short to bind the canonical poly(A) binding protein (Pab1), which needs a minimum of around 12nt for binding (Wahle *et al.*, 1991). Thus, the oligo(A) and A-rich tails which are left unprotected provide an entry site for exonuclease-mediated degradation. The exact number of nucleotides that are needed for a "toe-hold" for the exosome is currently unknown. This degradation requires slow and distributive polyadenylation associated with frequent polymerase-RNA dissociation linked to degradation (LaCava *et al*., 2005). Among the other known orthologues of Trf4 and Trf5, Cid14 is associated with ribosomal RNA and DmTrf4-2 with snRNA polyadenylation.

Poly(A) tails containing non-A residues have been reported in bacteria, archaea, chloroplasts, human 18S and 28S ribosomal RNA, mouse 5'ETS ribosomal rRNA and β-actin mRNA (Shcherbik *et al.*, 2010; Slomovic *et al*., 2006; Slomovic *et al*., 2008; Slomovic and Schuster, 2008; West *et al.*, 2006). In *Escherichia coli* polyadenylation is carried out by nucleotidyltransferase-type poly(A) polymerase (Ntr-PAP) which produces homopolymeric tails. Polynucleotide phosphorylase (PNPase) adds heteropolymeric tails. The polymerase responsible for generating any given tail may be identifiable by the nucleotide composition of the tail.

In humans co-transcriptional cleavage (CoTC) is necessary for transcription termination of the β-globin gene (Dye and Proudfoot, 2001). Of the two naked ends produced by CoTC, the 3' product is degraded by the exonuclease Xrn2 while the 5' products are polyadenylated before degradation. Co-transcriptional cleavage products of human β-globin pre-mRNA accumulate when PM-Scl100 (hRrp6) or core exosome subunit PM-Scl175 is depleted. Thus, these transcripts are degraded by a mechanism involving the exosome and its co-factors. The short A tails detected on these transcripts were extended upon PM-Scl100 knock-down. It has been suggested that the human homologue/s of Trf4 and/or Trf5 might be responsible for generating these heteropolymeric tails on CoTC sequences as canonical PAP is adenosine specific (West *et al.*, 2006). The heteropolymeric tails detected on 18S and 28S rRNA were also tentatively attributed to *POLS* and/or *PAPD5* (Slomovic *et al.*, 2006).

In actinomycin D (ActD) treated murine 3T3 (Shcherbik *et al*., 2010) and Hela (Lubas *et al.*, 2011) cells, 5'ETS pre-rRNA transcripts generated by Pol I have been reported to be polyadenylated by PAPD5. ActD is a transcriptional inhibitor which shuts down Pol I transcription completely at a concentration of 2 µg/ml. To ensure ongoing transcription, both murine 3T3 and Hela cells were treated with 20 ng/ml of ActD. In murine 3T3 cells only a minor accumulation of polyadenylated 5'ETS transcripts was detected in the absence of ActD. Lubas *et al.* (2011) do not indicate if in Hela cells PAPD5 polyadenylated transcripts are found in the absence of ActD. Down-regulation of murine PAPD5 resulted in inhibition of Pol I transcription as quantified by the levels of 47S pre-rRNA. However in the presence of ActD, 5'ETS fragments still accumulated. An increase in steady state levels of polyadenylated 5'ETS RNAs after down-regulation of core exosome subunits Exosc6 and Exosc10 indicated the involvement of the exosome in degradation of these murine 5'ETS adenylated transcripts. By contrast, knock-down of murine POLS had no effect on pre-rRNA adenylation or degradation (Shcherbik *et al*., 2010).

In Hela cells depletion of either PAPD5 or the zinc finger protein ZCCHC7, the human orthologue of Air2 (Houseley and Tollervey, 2008), resulted in reduced levels of 5'ETS products, which were further reduced in hRRP40- or hMTR4-depleted backgrounds. ZCCHC7 has no polyadenylation activity on its own, but is thought to aid the activity of PAPD5. hMTR4 was dispensable for adenylation of 5'ETS fragments. A PAPD5 and ZCCHC7 dimer is thought sufficient for binding and targeting RNA, while hMTR4 aids the recruitment of the exosome for transcript degradation (Lubas *et al.*, 2011).

The primary rRNA precursor identifiable in humans is the 47S transcript (Figure 4.1). In this study, the 3' ETS was analyzed to address the biochemical functions of POLS and PAPD5.

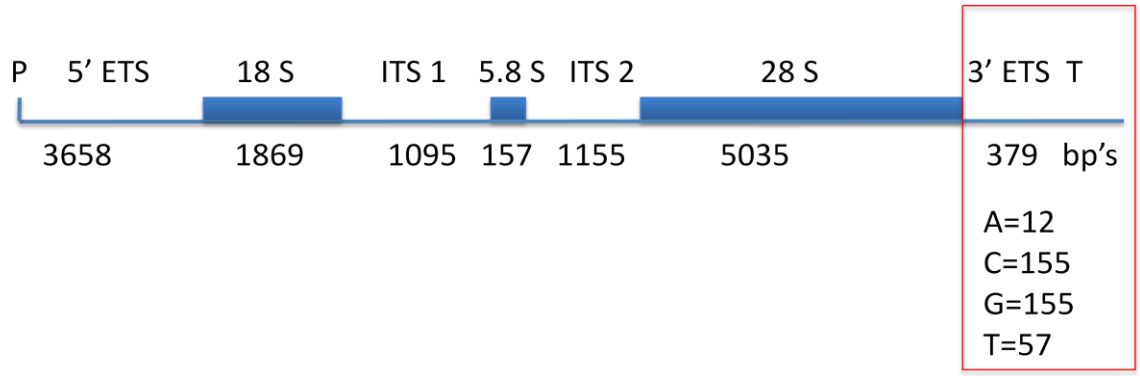


Figure 4.1 The human 47S pre-rRNA transcript.

The regions annotated above are the rRNA sequences 18S, 5.8S and 28S flanked by the two external transcribed spacers; 5'ETS, 3'ETS, and separated by the two internal transcribed spacers; ITS1, ITS2. P:promoter, T: termination site. The number of nucleotides of each region is indicated below the transcript. Indicated in the red box is the nucleotide composition of the 3'ETS.

Sequence and structural features of the human pre rRNA transcript

The transcribed spacer G+C contents range from 79-83%. Between bases 2887 and 3596 of the 28S variable region, the G+C content is 86%. The stability detected of the 28S region secondary structures can be attributed to this base composition. The 28S and spacer regions contain simple sequence repeats.

The 5'ETS is the most widely studied spacer region. Human 5'ETS bases 1800 to 2800 are well conserved and correspond to the murine 5'ETS bases 2070 to 3120. Hence, transcript processing in the two species could be similar in this region. ITS1 sequences also show a high level of conservation between humans and mice. In contrast, ITS2 sequences show poor conservation (Wellauer *et al.*, 1974).

Human and murine 3'ETS sequences are divergent at the primary sequence level; the human 3'ETS contains 379 bases in contrast to the 594 bases in the mouse. However, as in 3'ETS sequences in all organisms, it contains an extended hairpin structure just downstream from the 25-28S rRNA (Gonzalez *et al.*, 1990). In *S.pombe* it has been shown that this structure is required for the maturation of the 3' end of the 25S rRNA as well as for the processing of the ITS2 region. In budding yeast the first step in processing is carried out by Rnt1, which cleaves the stem of the 3'ETS. Both human and murine sequences share features such as pyrimidine tracts. Murine rRNA transcription terminates at 565 bases downstream from the 28S 3' end in a pyrimidine tract. Transcription termination in humans also takes place in a pyrimidine tract 352 (Gonzalez *et al.*, 1990) or 357 to 359 bases (Safrany *et al.*, 1989) downstream from the 28S 3' end.

As described in the introduction (Figure 1.4), the pre-rRNA transcript undergoes various processing and modification steps to generate the mature 28S, 18S and 5.8S RNAs. Polyadenylated ribosomal RNA transcripts are known to accumulate in exosome, *trf4, trf5*

and *cid14* mutants. This accumulation is due to the inhibition of the 3'-5' exosome mediated degradation/processing pathways.

Addition of 5-FU has been reported to increases the accumulation of polyadenylated transcripts in exosome and cid14 mutants.

4.2 Aims

In order to understand their biochemical functions, the involvement of human POLS and/or PAPD5 in the adenylation of fragments generated during rRNA processing was investigated in this study. Increased accumulation of polyadenylated transcripts in exosome and *cid14* mutants when exposed to the anti-cancer drug 5-FU have been reported. The possibility that POLS and/or PAPD5 might affect the sensitivity of cells to this drug was therefore also explored.

All experiments in this chapter were performed with HEK-293T cells due to the greater efficiency of transfection compared to U2-OS cells. A 70% down regulation of POLS and 55% down-regulation of PAPD5 was achieved in the following experiments.

4.3 Is/Are POLS or/and PAPD5 involved in the poly/oligo-adenylation of human 3'ETS rRNA tails?

The human 3'ETS was chosen to investigate if POLS and/or PAPD5 are/is responsible for adenylating transcripts. The human 3'ETS sequence is divergent from the mouse sequence. Most studies have been performed on the 5'ETS. Mouse and human PAPD5 is associated with the polyadenylation of 5'ETS rRNA. Thus, I choose to look at the 3'ETS to study the conservation/divergence of the function.

Using the 3' RACE technique (Figure 4.2) many pure A tails along with A tails containing occasional non-A residues were identified within the human 3'ETS region, which have not been documented to date (Figure 4.3). The tails identified are listed in Appendix **A**.

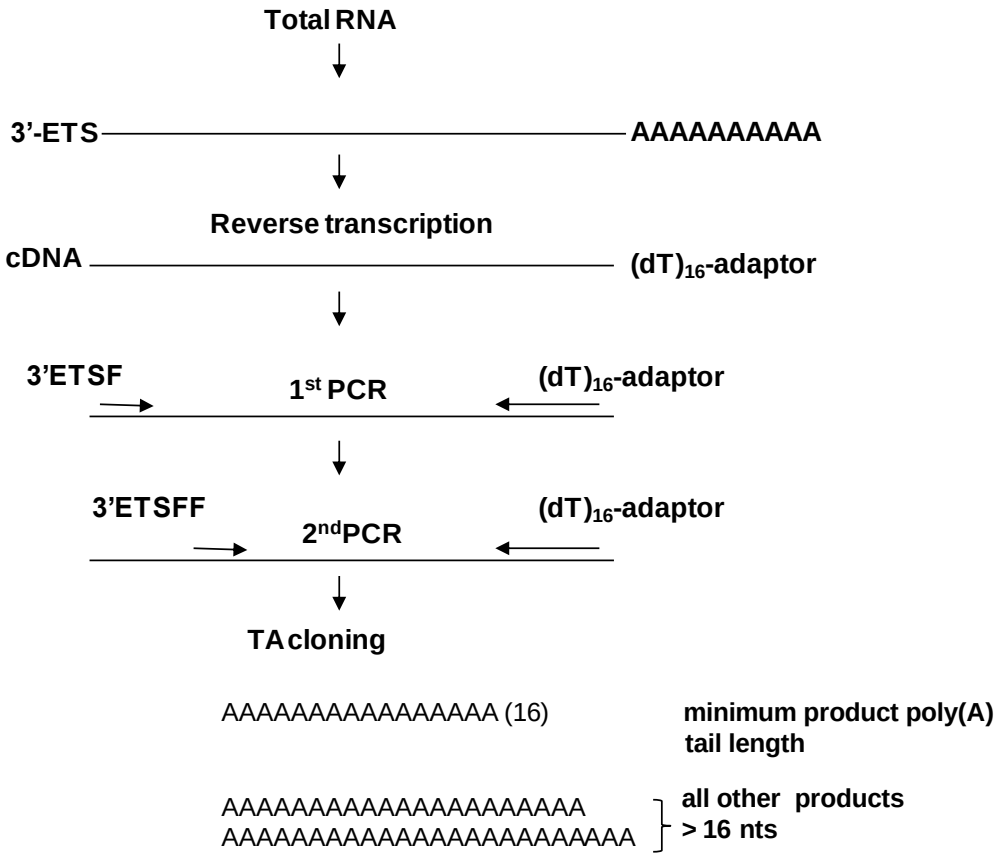


Figure 4.2 Schematic representation of the 3'RACE assay used for the analysis of 3'ETS tails.

Oligo/poly-adenylated 3'ETS tails were amplified from $(dT)_{16}$-adaptor primed cDNA with primers 3'ETSF and $(dT)_{16}$-adaptor followed by a second nested PCR with 3'ETSFF and $(dT)_{16}$-adaptor (Table 2.4) and cloned into pCR2.1(Invitrogen). Randomly selected clones were sequenced and aligned with the human 3'ETS sequence (Gonzalez *et al.*, 1990).

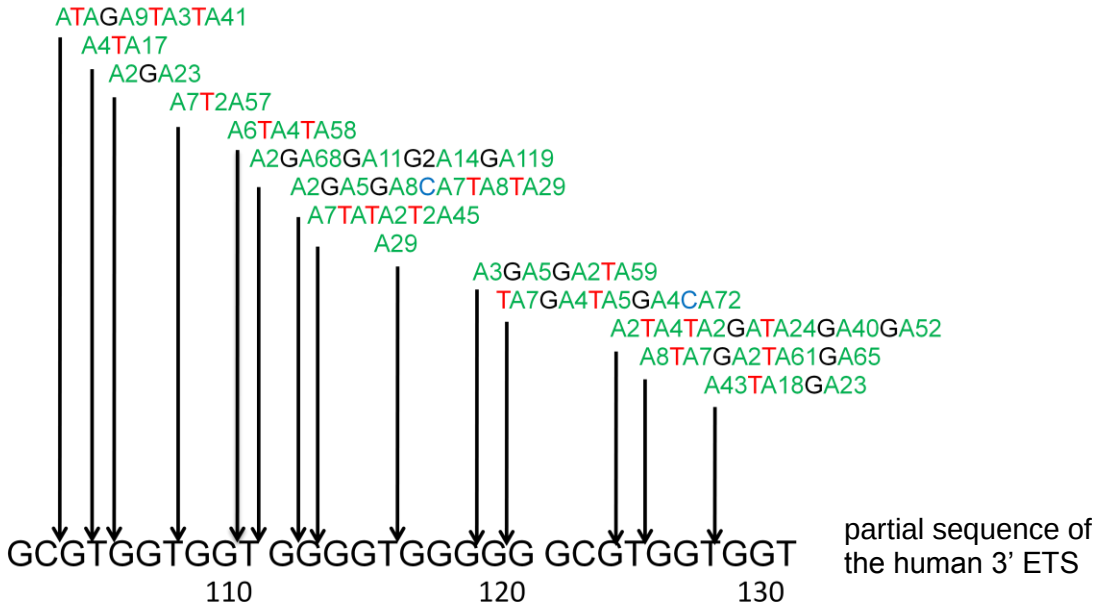


Figure 4.3 Schematic representation of the 3'ETS tails detected by 3' RACE.

Oligo/poly-adenylated 3'ETS tails were amplified by 3' RACE. Arrows indicate the last nucleotide in each clone before the start of an oligo/poly(A) sequence. 3' RACE products and northern analysis showed smears rather than a single product, consistent with the scattered positions of the 3' ends of the clones. Above is an illustrative sample of the tails detected within the 3'ETS. All the tails detected within the 3'ETS are tabulated in Appendix **A**.

During the course of the screen on several occasions individual clones derived from a single PCR reaction were found to have closely related, but distinct tail nucleotide compositions and a common poly(A) site. The tails shown in Table 4.1 were particularly striking.

Clone	Tail	Poly(A) site	Tail length
5_1	A5GA93	108	99
5_2	A100	108	100
5_3	A20GA15GA65	108	101
5_4	A5GA109	108	115
5_5	A25GA71	108	97
5_6	A63TA11TCA2GA20	108	100

Table 4.1 One cDNA can give rise to many different PCR products.
One PCR reaction gave rise to 6 different colonies each containing a different tail composition. The poly(A) site for each clone was the same. The total tail length varied marginally within the range of 97-115 nucleotides, and different clones contained different, but related patterns of non-A residues.

Thus, one cDNA was giving rise to many different PCR products. Here one tail contained 100 pure “A” residues; it was hypothesized that this was the parent molecule and the other nucleotides incorporated within the other tails are PCR-induced artefacts. This interpretation is supported by the presence of two clones with G residues at position 6. These were thought to result as the low abundance of these transcripts necessitated a high number of cycles (55) to be performed to amplify a product. Around the same time an article from the Schuster group reported the addition of poly(A) and poly(A)-rich tails during mRNA degradation in the cytoplasm of human cells (Slomovic *et al*., 2010). β-actin transcripts were shown to contain hetero- and homopolymeric tails including cytoplasmic RNAs with heteropolymeric tails added within the open reading frame, perhaps indicating that they were degradation intermediates. Could these heteropolymeric tails reported be pure A tails containing

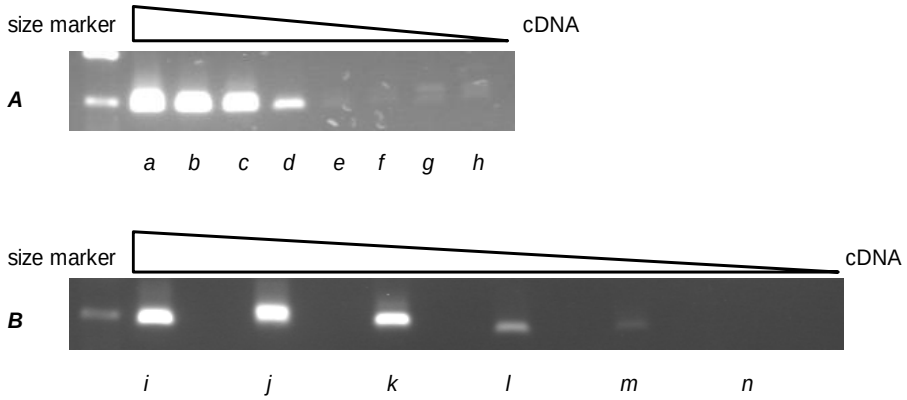
occasional non-A residues as a result of PCR artefacts? To investigate this further β-actin tails were screened using 3' RACE.

4.3.1 Screening of tail composition using 3' RACE

Tail compositions were screened using 3' RACE as described in section 2.2.7 and illustrated in Figure 4.2.

4.3.1.1 Human beta actin tail composition

An $(dT)_{16}$-anchor primer was used to reverse transcribe cDNA from total RNA. A series of five-fold serial dilutions of cDNA was amplified with Taq or Expand high fidelity polymerases using primers listed in Table 2.5. Products obtained after 45 PCR cycles are shown in Figure 4.4.



.

Figure 4.4 β-actin products amplified from Taq or Expand high-fidelity polymerases after serial dilutions of cDNA.
Serial dilutions of the cDNA generated from $(dT)_{16}$-adaptor were amplified with actinF and $(dT)_{16}$-adaptor followed by a second nested PCR with actinFF and $(dT)_{16}$-adaptor. The highest and lowest dilution where a PCR product was visible was chosen to screen tails. ***A*** Taq polymerase. ***B*** Expand high-fidelity polymerase.

The highest (lanes *a* and *i*) and lowest (lanes *d* and *l*) concentration where a product was visible was used as the starting material to screen tail compositions. The abundance of template will be greatest at the highest concentration and very limited at the lowest concentration. A wide distribution of tails would be captured by this method. Thus, PCR

products of different sizes were visible as a smear after gel electrophoresis rather than a discrete band. These PCR products were TA cloned. Sequence analysis of these randomly cloned actin sequences showed a wide distribution of 3' ends. The nucleotide compositions of all tails detected were calculated after deducting the 16 adenosine residues corresponding to the $(dT)_{16}$ primer. The error occurrence was very significant with a low amount of template compared to a high amount of template for the PCR amplification with Taq polymerase. Substituting Taq polymerase with proof-reading Expand high fidelity no significant difference in error incorporation was detected between the low and high template concentrations (Table 4.2). All β-actin tail sequences determined are tabulated in Appendix **B.**

	A residues		Non-A residues		Chi-square 2x2
	cDNA concentration		cDNA concentration		
	Low	High	Low	High	
Taq polymerase	341	889	14	7	p<0.0001
High fidelity	295	228	3	0	p=0.0804

Table 4.2 A residues and non-A residues recorded in the actin tail screen with Taq or Expand high fidelity polymerases.

Oligo/poly-adenylated β-actin tails were amplified from $(dT)_{16}$-adaptor primed cDNA with primers actinFF and $(dT)_{16}$-adaptor followed by a second nested PCR with actinF and $(dT)_{16}$-adaptor (Table 2.4) and cloned into pCR2.1 (Invitrogen). Randomly selected clones were sequenced. After deducting 16 adenosine residues corresponding to the oligo $(dT)_{16}$ primer, A residues and non-A residues were calculated.

For β-actin mRNA tails the error rate was very significant for Taq polymerase. As expected this error rate was improved by the use of proof-reading Expand high fidelity polymerase,

where only pure A tails were detected at the highest concentration of cDNA. However, errors were detected at the lower concentration of cDNA.

The positions at which these errors occurred were apparently random (Figure 4.5). The frequency of erroneous nucleotide incorporation is tabulated in Table 4.3. The Schuster group observations are tabulated in Table 4.4.

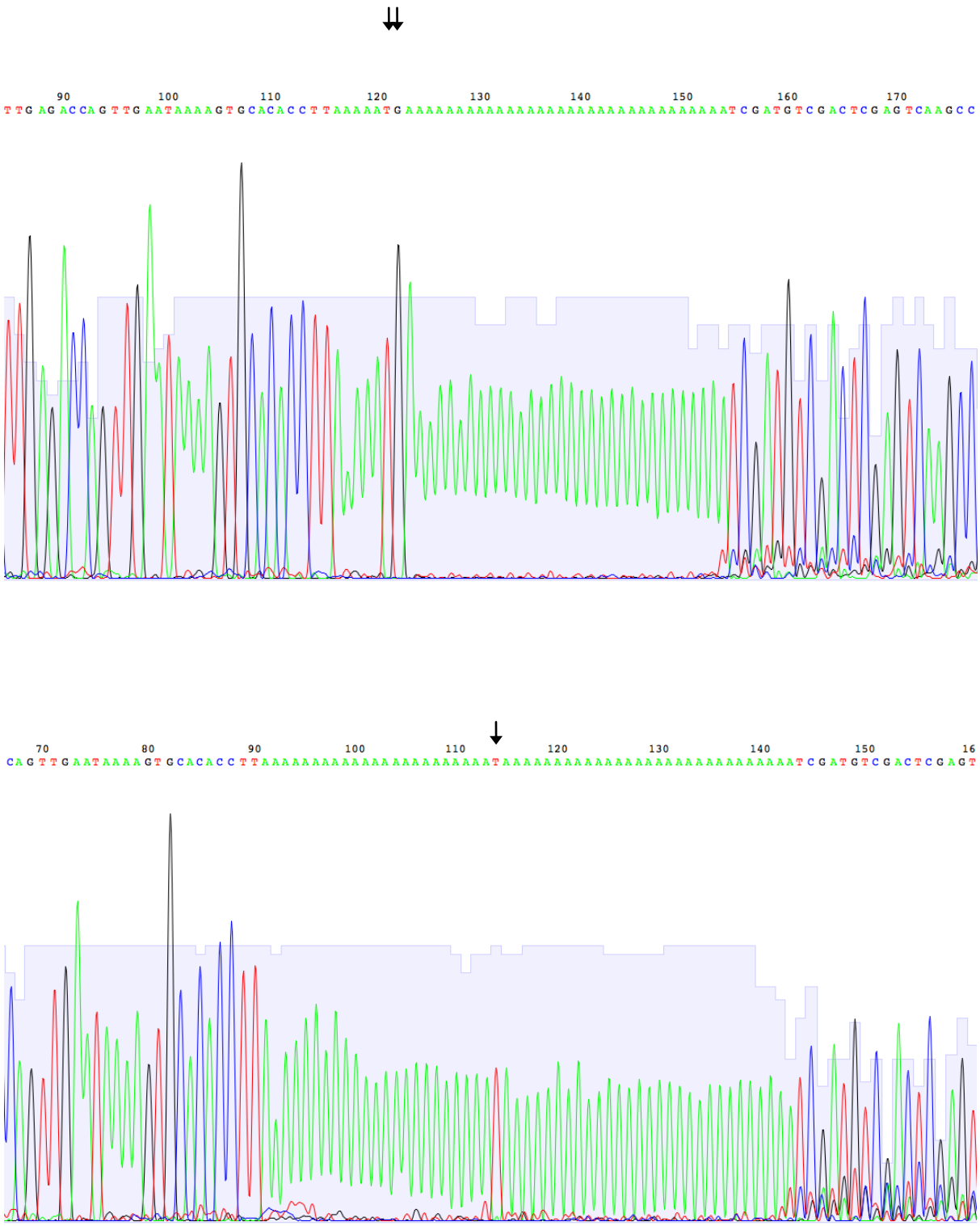


Figure 4.5 Sequence profiles showing non-A residues detected within the A tails.

Screening of β-actin tails using 3'RACE generated tails containing pure A residues and A-rich tails containing non-A residues. Black arrows indicate the positions at which non-A residues were incorporated to the A tails.

transcript	polymerase	cDNA concentration	nucleotide incorporation
total β-actin mRNA	Taq	high	G & U > C
total β-actin mRNA	HF	low	G & U > C
total β-actin mRNA	Taq	low	C > G & U
3' ETS	Taq	high	G & U > C

Table 4.3 The frequency of erroneous nucleotide incorporation for β-actin and 3'ETS rRNA with Taq or Expand highfidelity polymerases.

Non-A residue incorporation at low and high cDNA concentrations within the poly(A) tails.

	A residues	Non-A residues
total β-actin mRNA	914	42
β-actin nuclear fraction	211	7
β-actin cytoplasmic fraction	131	0
28S rRNA nuclear fraction	368	88
28S rRNA cytoplasmic fraction	637	63

Table 4.4 A residues and non-A residues recorded in the Slomovic *et al.,* 2010 study for β-actin and 28S rRNA tail screens.

The A residues and non-A residues were calculated from supporting information Slomovic *et al.*, 2010 Table S1. Oligo/polyadenylated 28S rRNA and β-actin mRNA sequences isolated from Hela cells using oligo(dT) RT-PCR.

The number of PCR cycles performed, the amount of cDNA or the polymerase used for this work is not documented. However it would be reasonable to assume that the polymerase used was Taq polymerase with a low abundance of cDNA and a relatively high number of

cycles were performed to obtain these tails based on the results obtained in this thesis (Table 4.2).

4.3.1.2 Human rRNA 3'ETS tail composition

An approach similar to that used for the analysis of β-actin mRNA tails was followed to characterize the composition of rRNA 3'ETS tails. Here 0.5 µg of RNA from siRNA-treated cells or control cells was used as the starting material for the reverse transcription. PCR primers listed in Table 2.4 were used to amplify the products with Taq polymerase using a nested PCR approach (Figure 4.2). Comparison of DKD (*POLS* + *PAPD5)* or PAPD5 knock-down cells to control cells yielded no difference in nucleotide composition of the tails. Both fractions contained pure A tails as well as A-rich tails containing non-A residues.

4.3.2 3' RACE to screen human 3'ETS rRNA tail length

3' RACE sequencing of the human 3'ETS RNA showed no statistically significant variation in tail length between the DKD of (POLS+PAPD5) compared to the control (Figure 4.6.***A***). This was also true for the individual down-regulation of PAPD5 (Figure 4.6 ***B***).

4.3.3 Does down-regulation of POLS or/and PAPD5 alter the processivity of 3'-5' degradation?

The 3'ETS tails showed no statistically significant difference between the double knock-down of POLS and PAPD5, or PAPD5 knock-down *vs* control for residual ETS length. Thus, the processivity in other words the poly(A) site of 3'-5' degradation did not appear to be altered by the down-regulation of these genes.

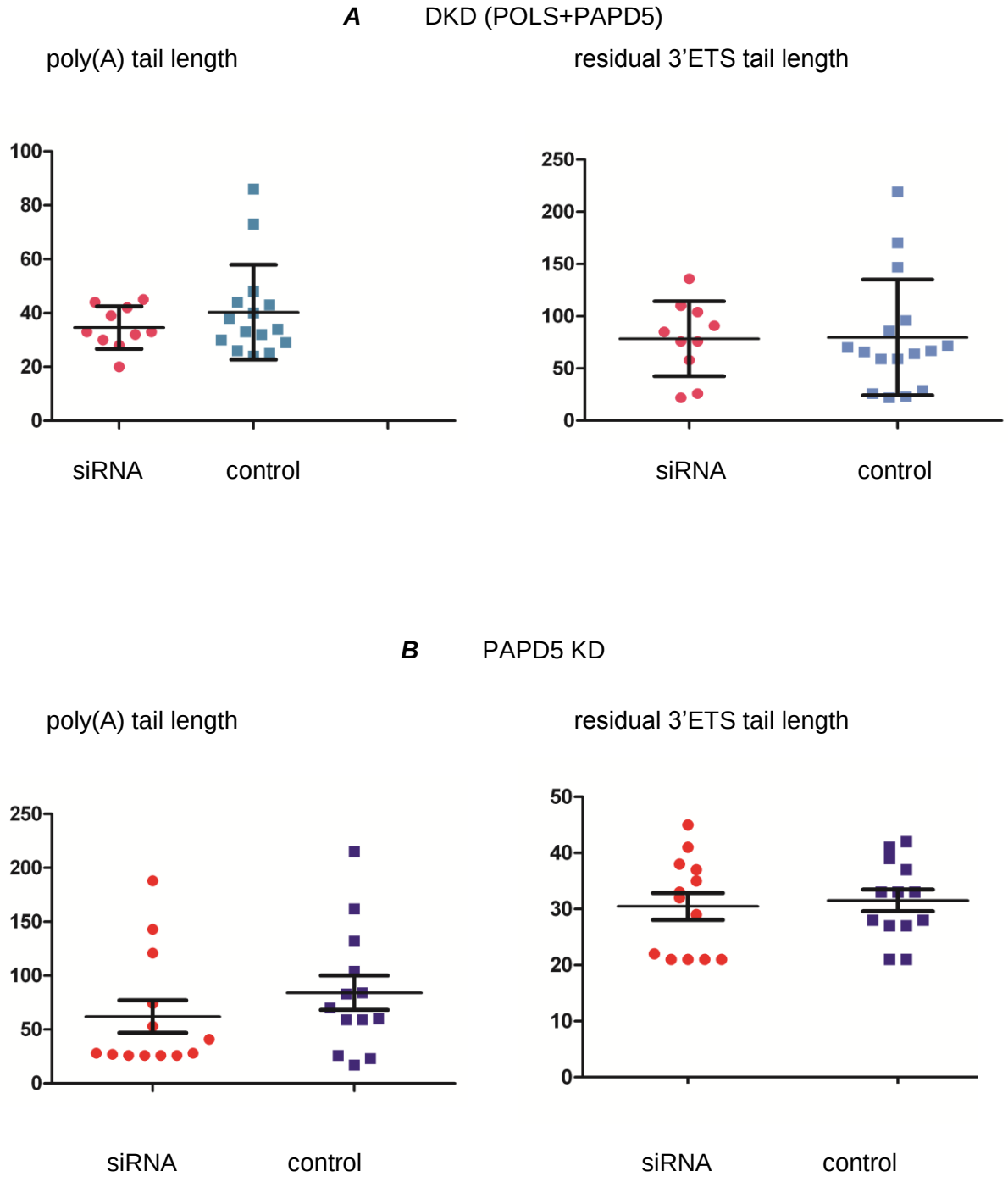


Figure 4.6 Comparison of poly(A) and residual 3'ETS tail lengths.

A DKD of (POLS+PAPD5); in pink the siRNA-treated and light blue scrambled control, ***B*** PAPD5 KD; in red the siRNA-treated and blue the scrambled control. Mean and standard deviation is shown for each population.

4.3.4 Are the 3'ETS rRNA tails detected added before or after cleavage of the 3'ETS?

To investigate if the polyadenylated 3'ETS rRNAs detected were generated before or after cleavage of the pre-rRNA, northern analysis was performed.

4.3.4.1 Northern analysis

A 3'ETS probe (Table 2.6) was used to analyse total RNA and polyadenylated RNA. Total RNA would contain all RNA species including the 47S pre-rRNA and all downstream processing products including the 3'ETS. This probe was able to detect species around 379 nt corresponding to the human 3'ETS as well as 47S precursor transcript in the total RNA. With the polyadenylated RNA only the high molecular weight transcripts were visible (Figure 4.7).

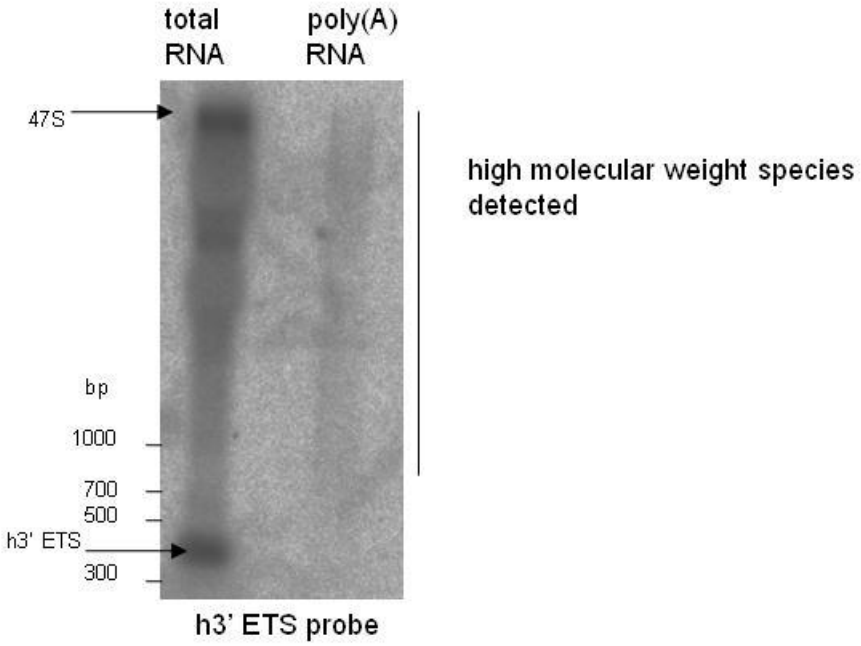


Figure 4.7 Northern hybridization of total RNA and poly(A) selected RNA with the h3'ETS probe.

Total RNA was extracted and poly(A)$^+$ RNA was prepared as described in material and methods section and analysed with the h3'ETS probe. A DNA ladder was used as a size marker.

Based on these results it can be concluded that the 3'ETS poly(A) tails detected by 3' RACE are still attached to the 28S transcript. In other words 3' adenylation precedes cleavage at the 3' end of the 28S sequence. In an attempt to capture the 3'ETS transcripts identified by northern analysis, cDNA cloning across the 28S-3'ETS boundary was undertaken.

4.3.4.2 Cloning of the human 28S-3'ETS boundary

Primers were designed spanning the 28S-3'ETS boundary as shown in Figure 4.8.

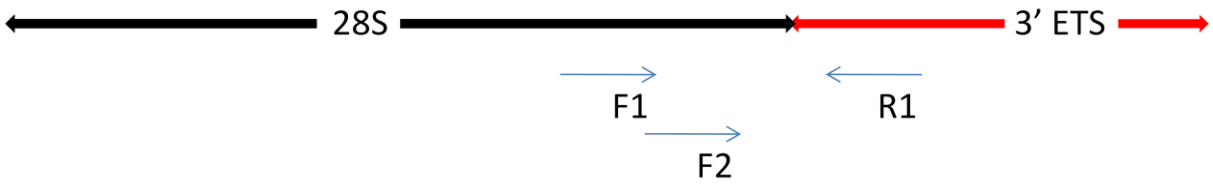


Figure 4.8 Schematic representation of the primers used to clone across the 28S-3'ETS boundary.
Forward primers F1 or F2 with reverse primer R1 was used to clone across the 28S-3'ETS boundary region.

Successful cloning across the boundary of the 28S ribosomal RNA and 3'ETS was achieved with the use of DMSO and low salt in the PCR reaction (Figure 4.9). DMSO disrupts the secondary structures found in rRNA. Low salt concentrations were found to aid amplifying regions with high GC nucleotide content.

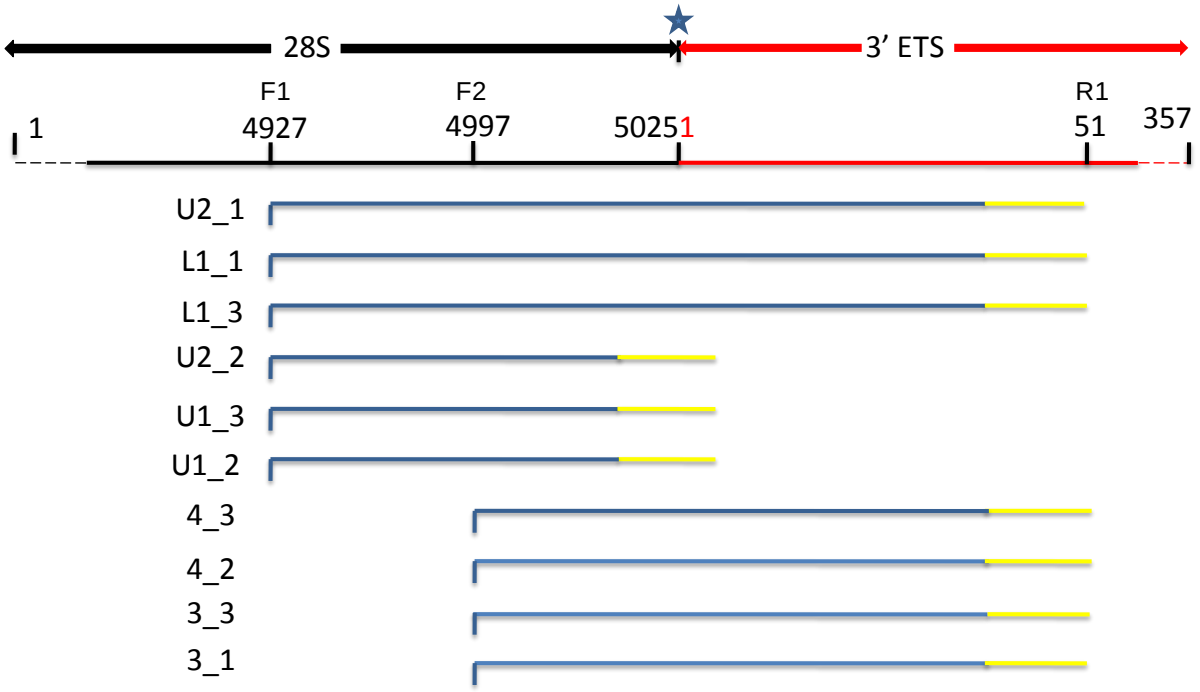


Figure 4.9 Cloning across the human 28S and 3'ETS boundary.

The end of the 28S and start of the 3'ETS is indicated by *. Positions of the primers used are indicated (F1, F2 and R1). The clones obtained are represented schematically. In yellow the reverse primer, in blue the forward primer and sequence detected through to the reverse primer.

4.3.5 Are the 28S-3'ETS boundary products RNA specific?

cDNA was made with 0.5µg of total RNA with Superscript II as described in section 2.2.3. Control tubes were set up eliminating reverse transcriptase (RT) from the reaction but with all other components. This control was required to demonstrate the RNA specificity of the product. If a product was visible in the no RT control, the product is derived from genomic DNA. Two different forward primers (F1 and F2) with two different reverse primers (R1 and R2) as shown in Figure 4.10 and listed in Table 2.5 were used to amplify the products.

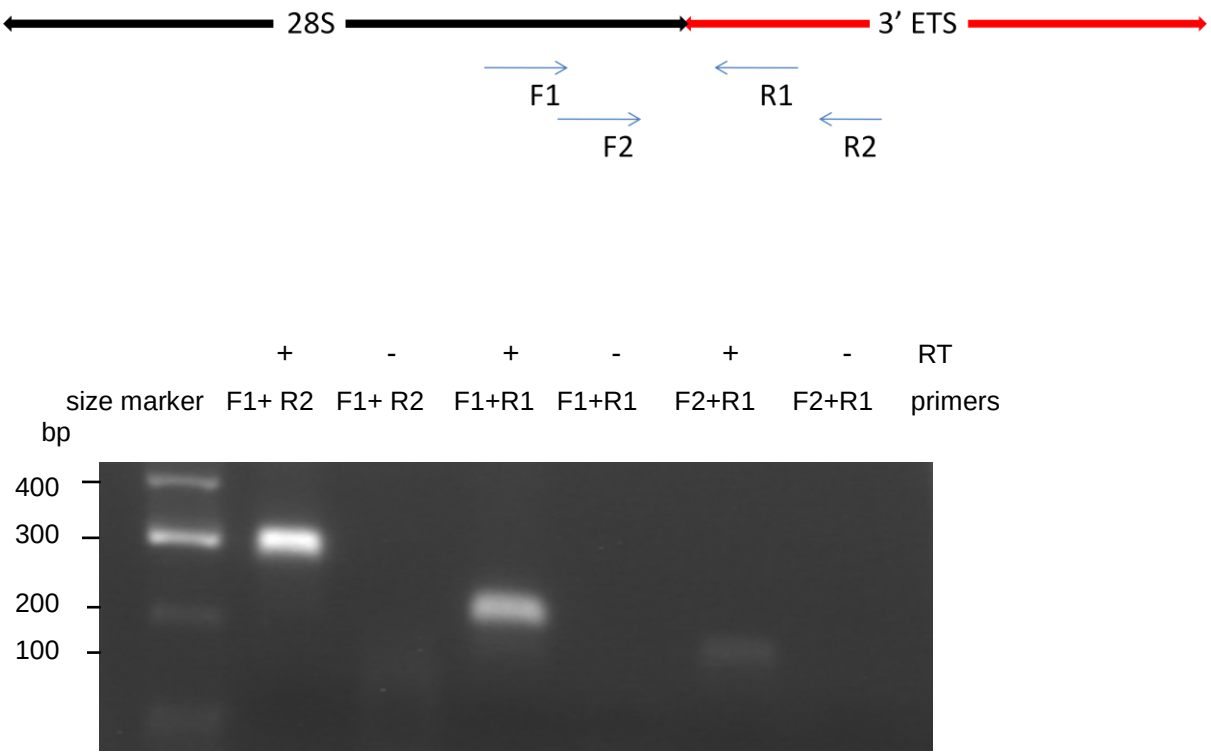


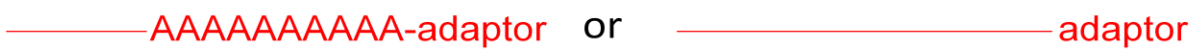
Figure 4.10 The 28S-3'ETS boundary products are RNA-specific.
cDNA synthesis was performed in the presence and absence of reverse transcriptase. Forward primers F1 or F2 was used in combination with reverse primers R1 or R2 to obtain products specified above the gel.

As can be seen in Figure 4.10 the 28S-3'ETS boundary products obtained were specific to the presence of RT, and by inference specific to the presence of RNA.

4.3.6 Are the 28S-3'ETS boundary products oligo(dT)-specific?

To find out if the products obtained were oligo(dT)-specific the primers shown in Figure 4.11, and listed in Table 2.5, was used to amplify the products.

cDNA synthesis



PCR

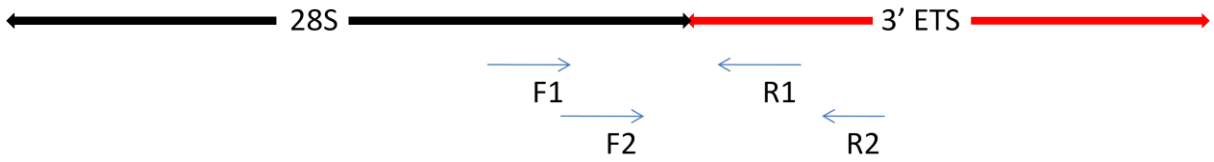


Figure 4.11 Schematic representation of the primers used for the reverse transcription and analysis of the product oligo(dT) specificity.
Oligo(dT)-adaptor or oligo(dT) was used to generate cDNA. The following primer combinations were used to generate the PCR products; F1+R1, F1+R2, F2+R2 and F2+R1.

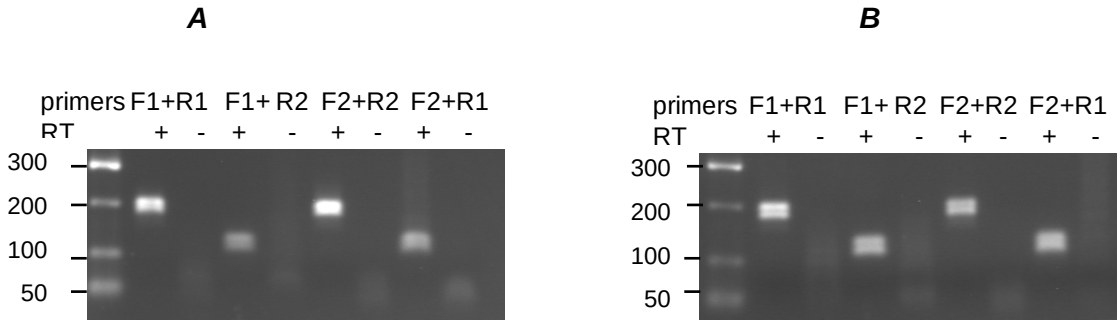


Figure 4.12 The 28S-3'ETS boundary products are not oligo(dT)-specific.
A,B PCR products were amplified using the primers shown in Figure 4.11, and listed in Table 2.5 ***A,*** Oligo(dT)-adaptor primer used for cDNA synthesis. ***B,*** Adaptor only primer used for cDNA synthesis. Indicated on the left are size markers in base pairs.

As could be seen in Figure 4.12, the products obtained by cloning across the 28S-3'ETS boundary were not oligo(dT)-specific. cDNA generated by adaptor primer produced the same

products as the cDNA generated using the oligo(dT)-adaptor primer. Limited sequencing did not reveal any poly(A) sequences but the product specificity was confirmed. This could be due to abundant products out competing relatively rare minor polyadenylated species. By sequencing 3' RACE products however, 'internal' poly(A) sequences occurring within the 3'ETS (Figure 4.3) were detected. Northern analysis (Figure 4.7) also indicated the presence of heterogenous polyadenylated transcripts much longer than the 3'ETS. Hence, it is very likely that the internally polyadenylated transcripts identified within the 3'ETS are derived from the 47S precursor rRNA.

4.4 Is the activity of 5-FU linked to POLS and/or PAPD5?

The defective ribosomal RNA degradation observed in *S.cerevisiae rrp6* mutants and *S.pombe cid14* mutants was enhanced by the addition of 5-FU (Fang *et al.*, 2004; Win *et al.*, 2006). An accumulation of polyadenylated rRNA intermediates was detected in both instances. Therefore, a pathway for degrading rRNA precursors may be inhibited by 5-FU. To investigate if POLS was associated with 5-FU activity, POLS localization was monitored in the presence of 5-FU. Colony formation assays were also performed in POLS and/or PAPD5 down-regulated cells to detect any change in cellular sensitivity to the drug.

4.4.1 5-FU treatment and immunofluorescence microscopy

In a U2-OS stable transfectant (K10) expressing POLS-HA, at 12.5µM and 25µM 5-FU concentrations POLS accumulated in the nucleoli with some POLS visible in the cytoplasm. In higher concentrations of 50 µM, 100 µM, 200µM and 400 µM most of the POLS was seen moving into the nucleus, leaving the nucleoli free of POLS (Figure 4.13).

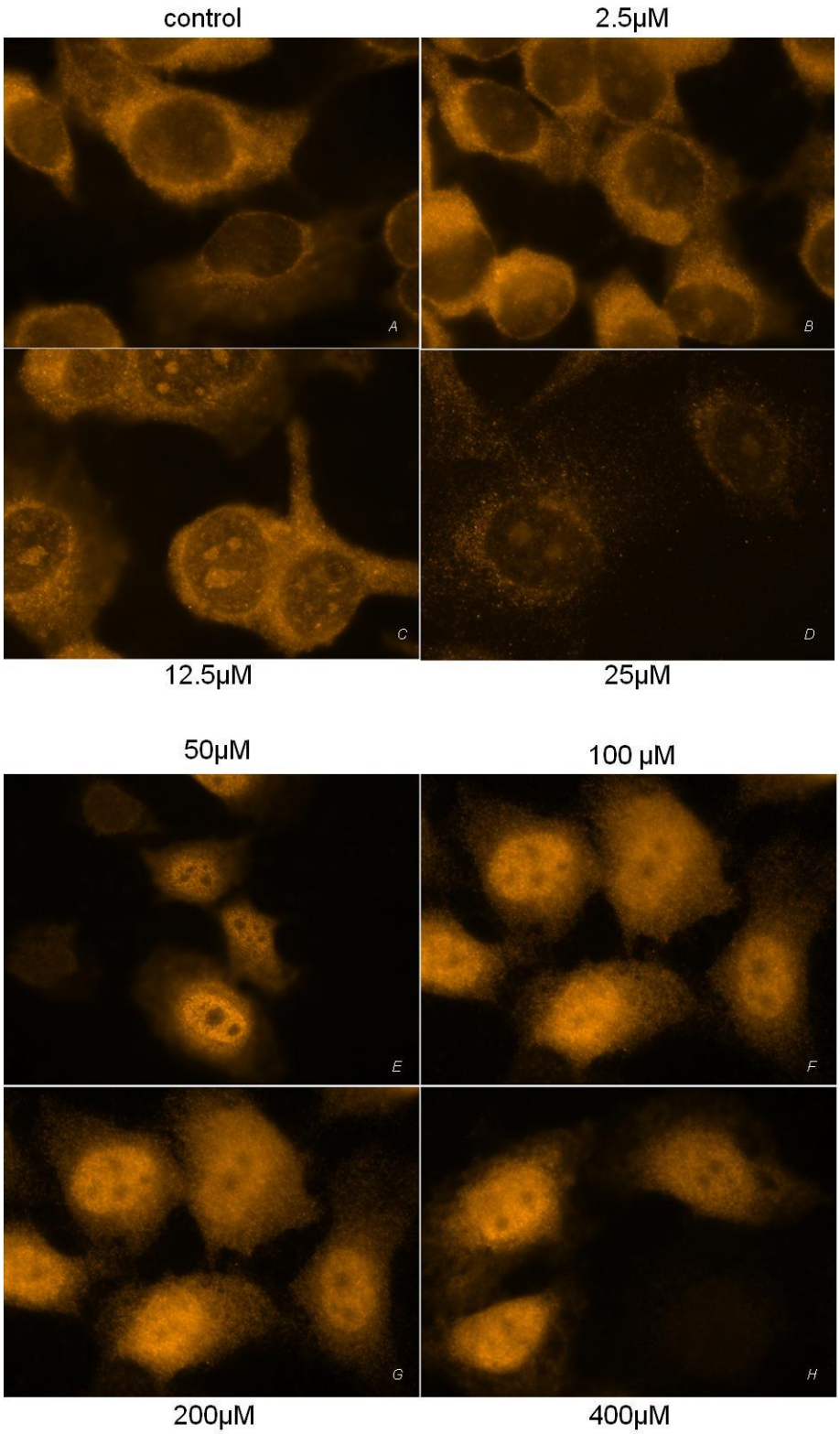


Figure 4.13 Stable transfectant clone K10 expressing POLS-HA treated with varying concentrations of 5-FU for 48 hr.

A, control, concentrations of 5-FU in μM; ***B***; 2.5, ***C***; 12.5, ***D***; 25, ***E***; 50, ***F***;100, ***G***;200, ***H***;400. Anti-HA antibody was used to detect expression of POLS by immunofluorescence**.**

4.4.2 Is the sensitivity of cells to 5-FU altered by the down-regulation of POLS and/or PAPD5?

Colony formation assay

POLS or PAPD5 individual knock-down and DKD (POLS + PAPD5) cells were treated with various concentrations of 5-FU two days after the transfection. Each treatment was carried out in triplicate (Figure 4.14). In the gene down-regulated cells the numbers of colonies formed were lower than the control. The knock-down of POLS and/or PAPD5 resulted in a slow growth phenotype (Figure 3.9). Thus, for POLS or PAPD5 gene down-regulated cells with no 5-FU added the number of foci formed would be expected to be the same but smaller foci than the control would be obtained. However with the (POLS+PAPD5) DKD cells, as some cell death occurred in the gene down-regulated population as measured by the trypan blue assay and flow cytometric analysis, a reduction in the number of colonies would be expected. In this experiment, the number of foci formed was lower in all gene down-regulated populations compared to the control. It is possible that in the individual knock-down cells the other orthologue was not able to support the function disrupted on its own for a prolonged period of time resulting in the loss of some colonies. The cells were incubated for a longer period in the colony formation assay compared to the cell growth and trypan blue assays.

There was a significant difference in the number of colonies formed at a concentration of 6.25 µM of 5-FU between the (POLS+PAPD5) DKD and control, indicating that DKD cells were significantly more sensitive at 6.25 µM 5-FU. At 50 µM, 100 µM and 200 µM concentrations cell survival was too low to allow an evaluation of any effect of gene down-regulation on the cellular sensitivity to 5-FU.

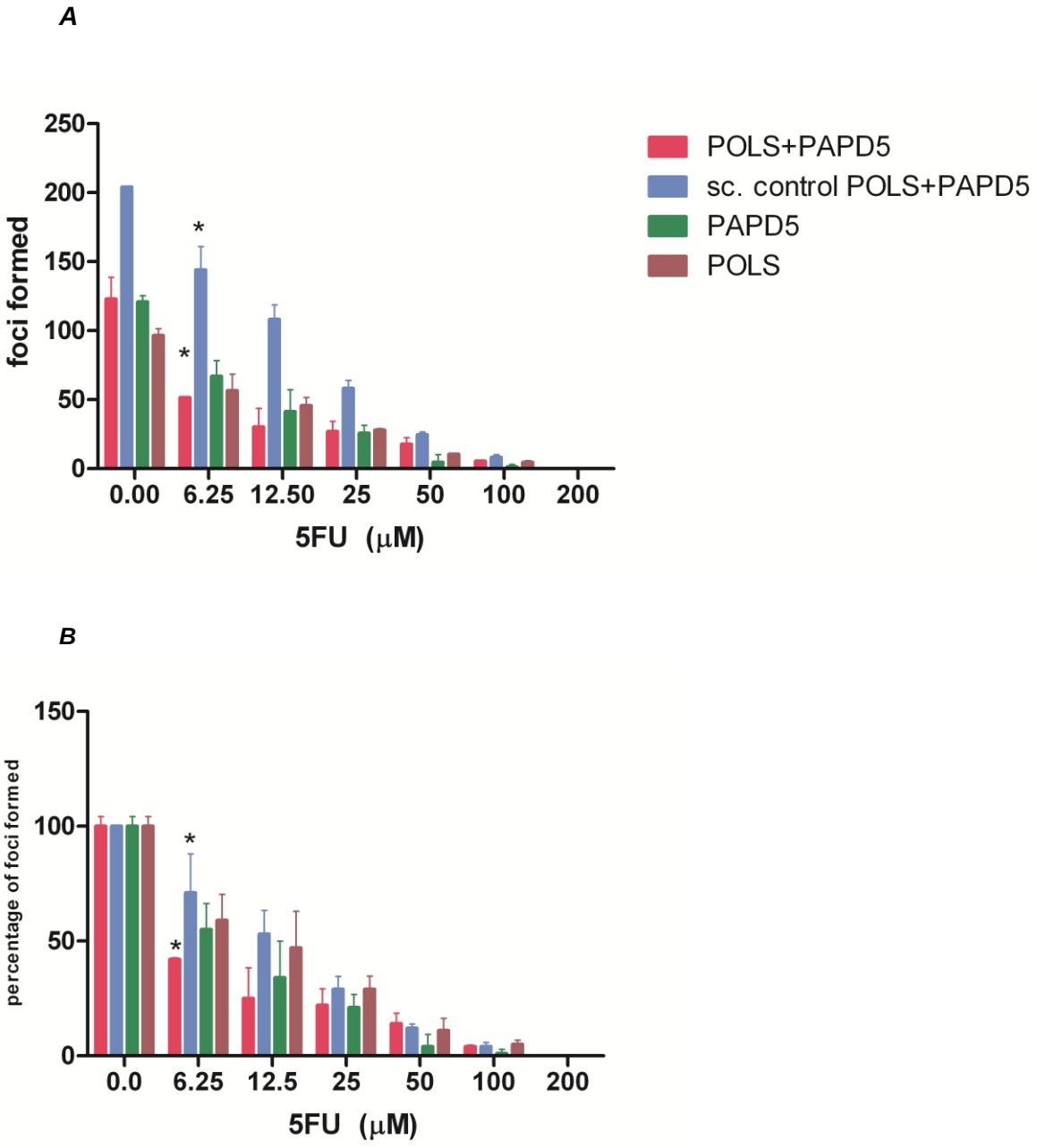


Figure 4.14 Is the sensitivity of cancer cells to 5-FU altered by the down-regulation of POLS and/or PAPD5?

A,B Colony formation assay. 24 hr after siRNA knockdown U2-OS cells were treated with 5-FU. The number of colonies were counted 10 days after the initial siRNA transfection. Each treatment was carried out in triplicate. * T-test p-value of 0.008 was obtained for control *vs* (POLS+PAPD5) DKD cells at 6.25 μM 5-FU. ***B*** Percentage foci formed with zero dosage normalised to 100%.

4.5 Discussion

Many polyadenylation events were detected using the 3'RACE technique in the human rRNA 3'ETS. To my knowledge this is the first record of detection of polyadenylation in metazoan rRNA 3'ETS.

By screening both human β-actin and rRNA 3'ETS tail compositions, it became apparent that many, and perhaps all, of the non-A residues incorporated in the A-rich tails are PCR-induced artefacts. The 3'ETS products are of very low abundance, necessitating a high number of PCR cycles (55) to be performed to obtain a product. With a low amount of actin template, 45 PCR cycles were required to obtain a product. Thus, non-A residues are incorporated when the abundance of template is low.

G and U nucleotide incorporation was greater than C nucleotide in actin or 3'ETS screened with Taq polymerase when 0.5 µg total RNA was used as the starting material for cDNA synthesis. Actin screened with Expand High fidelity polymerase resulted in a similar pattern of error. The preferences for G>U>C residue incorporation in oligoadenylated tails reported in other studies (LaCava *et al.* 2005,; Wlotzka *et al*., 2011) closely resemble the erroneous nucleotide incorporation for the relatively higher concentration of cDNA using Taq polymerase observed in this study (Table 4.1). However, this pattern of erroneous nucleotide incorporation was altered when 0.0025 µg of total RNA was used as the starting material for cDNA synthesis. Here, more erroneous C nucleotide incorporation was observed than G and U with Taq polymerase. Based on the results obtained in this study for both β-actin mRNA and 3'ETS rRNA, very stringent control measures need to be taken to ensure non-A residues incorporated within A-rich tails are not PCR-induced artefacts.

Results obtained in this study suggest that the cytoplasmic mRNA degradation intermediates containing heteropolymeric tails reported by the Schuster group could be poly(A) tails added

in the nucleus by the canonical poly(A) polymerase. The transcripts detected in the cytoplasm are likely to have been exported from the nucleus which may represent aberrant cleavage and polyadenylation rather than degradation intermediates.

No statistically significant difference in 3'ETS poly(A) tail length was observed between the DKD or PAPD5 knock-down compared to the control. Thus, there is no evidence that POLS or PAPD5 is involved in polyadenylating 3'ETS rRNA. The oligo-(dT)-adaptor used for cDNA synthesis contained 16 thymidine nucleotides, eliminating from the analysis any transcripts with fewer than 16 adenosines. As oligo-(dT) is capable of annealing to any position along the stretch of adenylated RNA, it is possible that the true tail length is under-represented. However, this would be true for both populations screened. In PCR and cloning reactions, major species would out-complete the minor species again in this instance this would apply to both populations. The total abundance of adenylated transcripts could not be quantified compared to the non-adenylated transcripts by 3' RACE.

Northern blotting would enable quantification of the polyadenylated pool of transcripts in the gene down-regulated compared to the control. However, as with all other oligo-dT selection methods non-adenylated transcripts would be eliminated. The use of high throughput sequencing might resolve this issue.

Other studies that implicate PAPD5 in polyadenylation slowed transcription by the addition of ActD to generate incomplete transcripts. 3'ETS rRNA turnover is likely to be very rapid under normal physiological conditions. To detect intermediate 3'ETS transcripts either the exonuclease and/or endonucleases responsible for degradation of these transcripts may need to be depleted.

The polyadenylated human 3'ETS transcripts identified appear to be representative of pre-cleavage transcripts that are still attached the 28S rRNA. Northern analysis indicates that

these transcripts are shorter than the 47S in the poly(A) selected fraction and have therefore presumably suffered 5'-3' degradation.

The colony formation assay indicated that U2-OS cells depleted of POLS and PAPD5 were hypersensitive to 5-FU at a concentration of 6.25 µM. In the nucleoli, the ribosomal RNA processing centres of the cell, an accumulation of POLS was noted at 5-FU concentrations of 2.5-25 µM. Human and murine PAPD5 have roles in adenylating incomplete rRNA transcripts. PAPD5 and POLS may have a redundant function which under normal circumstances is mainly performed by PAPD5. Trf5 is known to have a more significant role in rRNA processing than Trf4. If 5-FU targets PAPD5, POLS may be part of a mechanism that responds to PAPD5 inhibition and moves POLS into the nucleus. The addition of 5-FU altered the localization of POLS. At present the significance of this variation in localization POLS is unknown.

The vast amount of research performed on 5-FU links this drug to many target substrates such as rRNA, mRNA, tRNA, snRNA and snoRNA. Even splicing machinery is thought to be one of the drug targets. Yeast *rrp6* mutants display an accumulation of polyadenylated transcripts in the presence of 5-FU (Fang *et al.*, 2004). Targeting a polymerase such as Trf4 or Trf5 the drug could be acting on many substrates. According to the vast amount of research performed, 5-FU targets all the Trf4 and Trf5 substrates known to date. In human cells 5-FU may be targeting POLS or/and PAPD5, affecting many substrates.

5 Interacting protein partners of POLS

5.1 Introduction

Trf4 and Trf5 are known to target aberrant RNAs produced by different biogenesis pathways for exosome-mediated degradation *via* the TRAMP complexes. TRAMP4 (Trf4, Air2 and Mtr4) targets include non-nucleolar RNAs while TRAMP5 (Trf5, Air1/Air2 and Mtr4) is associated with nucleolar RNAs. *S.pombe* Cid14 was found to co-purify with homologues of Mtr4 and Air1 forming a *sp*TRAMP complex, which is involved in ribosomal RNA processing and silencing of heterochromatic genes *via* the exosome and/or the RNAi machinery.

Lubas *et al.* (2011) identify human TRAMP complex components PAPD5, ZCCHC7 and SKIV2L2/hMTR4 by SILAC/MS. PAPD5 and ZCCHC7 were reported to form a complex which is active in the nucleolus, targeting incomplete pre-rRNAs that accumulate following ActD treatment. A second Nuclear Exosome Targeting (NEXT) complex was also identified in the same study associated with the degradation of promoter-upstream transcripts (PROMPTs). Components of the NEXT complex include SKIV2L2, zinc finger protein ZCCHC8 and the RNA binding protein RBM7. This complex is active in the non-nucelolar region of the nucleus.

POLS, like most other non-canonical poly(A) polymerases, contains a catalytic domain and the PAP-associated domain but lacks an RNA-binding domain. Thus, it is very likely that POLS exerts its functions in association with other proteins.

5.2 Aim

I set out to identify interacting protein partners of TAP-tagged POLS and to validate the apparent interactions of POLS with BRCA1, RAD50 and Ku80. As a first step towards the characterization of interactions BRCA1, RAD50 and Ku80 were used as they were freely available antibodies in the department with no new purchase necessary. Based on U6 TUTase data, a potential interaction between POLS and CPSF-73 was also investigated.

5.3 Identification of interacting protein partners of POLS by tandem affinity purification.

To identify interacting protein partners of POLS, HEK-293T cells were transfected with pcDNA-*POLS*-TAP or pcDNA-TAP. HEK-293T cells were used because of the higher efficiency of transfection compared to U2-OS cells. Single-step affinity purification was used to generate protein complexes of sufficient purity to be analyzed by mass spectrometry (Figure 5.1). The TAP-tag method is based on the principle that two high affinity tags, Protein A of *Staphylococcus aureus* and the calmodulin binding peptide (CBP), would enable efficient recovery of fusion proteins. In this study the second affinity purification step was not performed, as material of sufficient purity for mass spectrometric analysis was recovered by IgG bead purification. The TEV protease cleavage site enables proteolytic release of the bound material under non-denaturing conditions ((Rigaut *et al*., 1999).

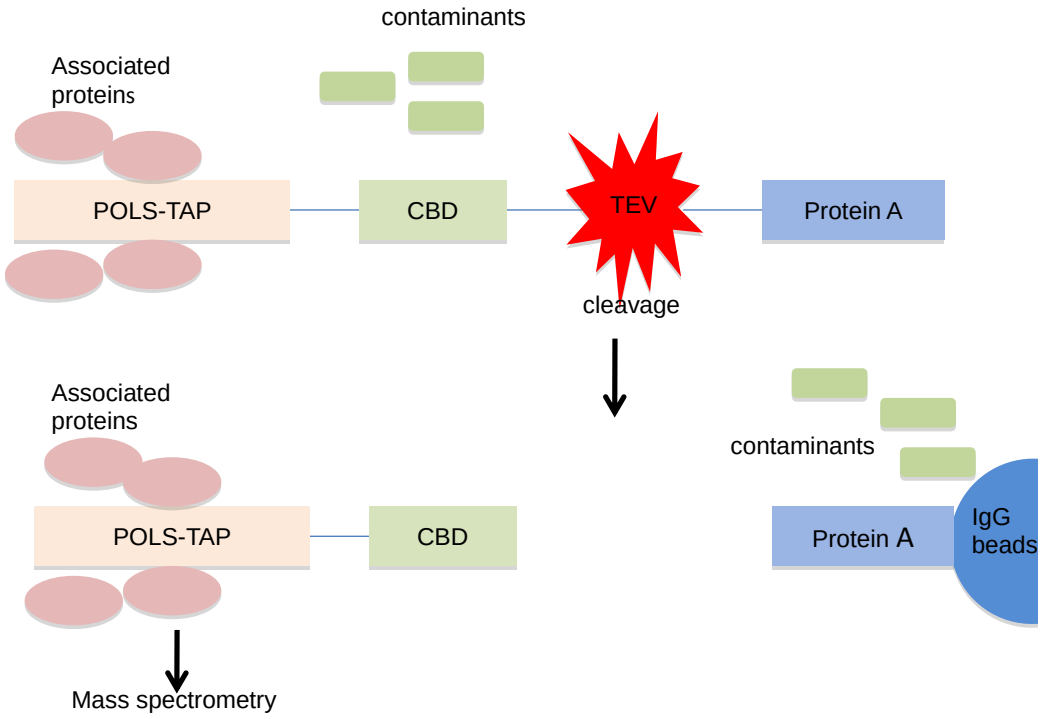


Figure 5.1 An overview of the TAP strategy.

The TAP-tag was fused to POLS and expressed in HEK-293T cells. The POLS-TAP and associated proteins were recovered by affinity selection on an IgG matrix. After washing, TEV protease was added to release the bound material, which was then subjected to mass spectrometric analysis.

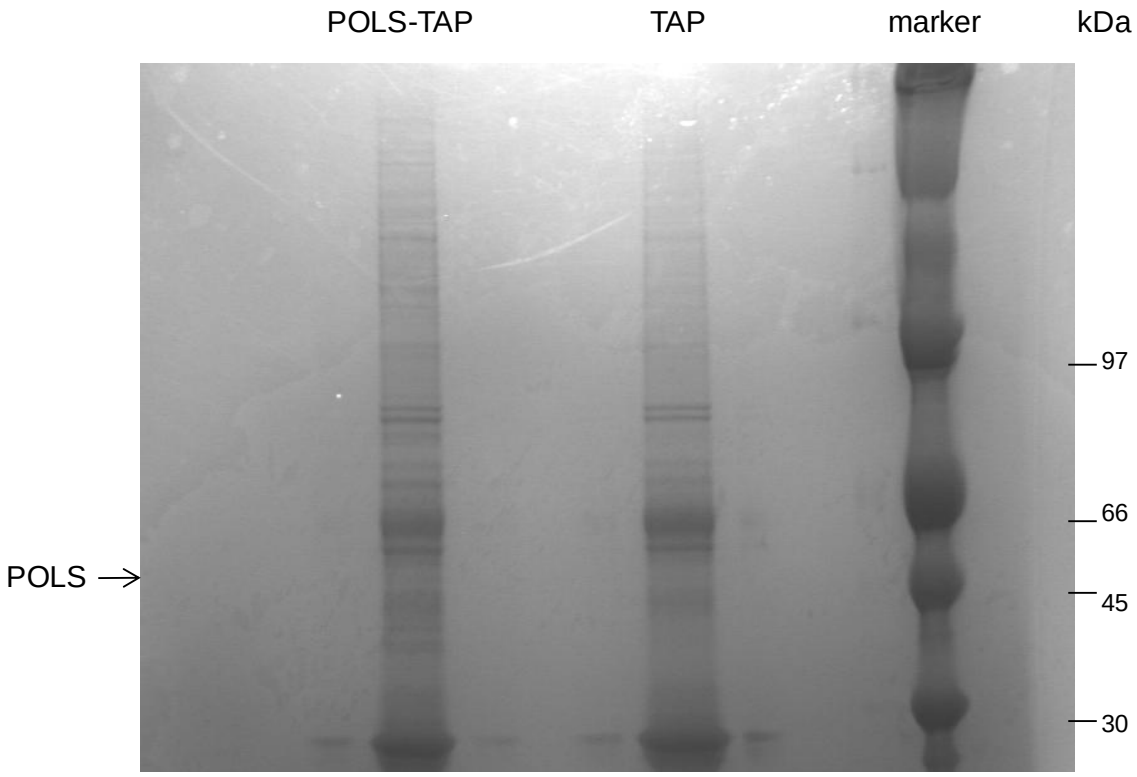


Figure 5.2 SDS gel analysis of proteins co-purified with POLS-TAP and TAP.
POLS-TAP and TAP, precipitated proteins are shown. Whole-gel analysis was performed for each. Marker; molecular weight markers.

SDS gel analysis of proteins co-purified with POLS-TAP and TAP are shown in Figure 5.2. The large number of interacting protein partners identified for POLS by MS analysis were placed in five main categories as follows:

1. RNA helicases and zinc finger proteins,
2. Other RNA metabolism-related proteins,
3. DNA repair-related proteins,
4. Cohesion, nuclear structure and nuclear pore-related proteins.
5. All other proteins, including uncharacterised proteins.

Proteins identified are listed in Appendix **D**.

DNA repair	Nuclear structure	RNA metabolism	RNA helicases and zinc finger proteins	Other proteins
DNA-PK RPA1 KU80 CHD4 MTA2 MRE11 RAD50 MSH6 KU70 PARP1	Importin5 Importin7 SMC1 NUP85 NUP214 NUP35 NUP205 SMC3 SMC2 KIF4A	BAT1 SWI/SNF complex Histone H1.4 Histone H1.2 HistoneH1.3 SF3B3 SF1 THO complex (SMARCC1) SUPT6H SFRS2 CF I_m 68 U2AF2 YWHAQ14-3-3 U4/U6 snRNP HDAC2 RRP42 XRN1 CCR4-NOT	ZNF143 ZNF828 ZC3H15 DDX17 DDX42 DHX9 SKIV2L2(hMtr4)	85

Table 5.1 Proteins identified as interacting partners of POLS by mass spectrometry. These included proteins functioning in DNA repair, cohesion, nuclear structure, nuclear pore complex and RNA metabolism. RNA helicases and zinc finger proteins are also listed. Eighty-five other proteins were identified as interactors including uncharacterized proteins. Proteins listed contained one or more unique peptides and an ion score of ≥45. Proteins are listed in descending order of ion score.

Proteins identified containing one or more unique peptides and an ion score of greater than 45 are listed in Table 5.1. Ion score is 10log(P), where P is the probability that the observed match is a random event. Non-specific hits present in the TAP control were eliminated. The TAP control contained proteins that were absent from the POLS-TAP. For instance when filtered with an ion score of ≥45, minimum peptide length 8 and a significance threshold $p< 0.05$, for POLS-TAP 130 proteins, for TAP-control 24 proteins and 47 proteins common to both were identified. Ion score was calculated using the Mascot software algorithm with strongest certainty of identification assigned the highest scores. The number of unique peptides corresponds to the peptides recognised with eight or more amino acids specific for the protein. It was striking that the interacting proteins identified exert their functions in the nucleus. Just three exceptions were noted; CCR4-NOT the major deadenylase, XRN1 and YWHAQ14-3-3 all reported to have solely cytoplasmic functions. It is possible that YWHAQ14-3-3, CCR4-NOT and XRN1 may have unidentified functions in the nucleus.

5.4 Is there a POLS-associated human TRAMP-like complex?

SKIV2L2, the human orthologue of Mtr4, the nuclear RNA helicase associated with the TRAMP complex, with 51% identity to the yeast protein (Houseley and Tollervey, 2008) was detected in this TAP-tag experiment. SKIV2L2 is also 33% identical with the yeast cytoplasmic RNA helicase Ski2. POLS may form a TRAMP-like complex in the nucleus and the cytoplasm.

ZCCHC7, a human protein which shares 34% identity to Air2 (Houseley and Tollervey, 2008) across the region containing five putative RNA binding motifs was not detected in this TAP-tag experiment. If the POLS-ZCCHC7 interaction is weak it may have been destroyed by the purification process. It is possible that evolutionary divergence has somewhat altered this pathway and there are many RNA binding proteins associated with the human equivalent of the TRAMP complex. Each substrate of the TRAMP-like complex could be recruited by its own specific zinc finger protein, with several TRAMP-like complexes forming in humans. Each substrate-specific TRAMP-like complex may assemble before degradation or processing occurs. It is possible that a specific zinc finger protein is associated with more than one function and more than one zinc finger protein may be involved in targeting a specific substrate increasing the complexity of the human TRAMP complex/es.

5.5 Protein interactions with POLS

A brief summary of the interacting proteins identified (in bold type) follows. This suggests an intricate network of protein interactions with POLS.

5.5.1 RNA helicases and zinc finger proteins identified in this study

In addition to **SKIV2L2/hMTR4**, described in section 5.4, RNA helicases belonging to the DEAD box family of proteins were identified which included **DDX17**, **DHX9** and **DDX42**

(Table 5.3). Pre-mRNA processing, ribosome biogenesis, RNA turnover, RNA export and translation are roles assigned to these proteins which involve the formation of large RNP complexes as well as the modification of large RNA structures. An emerging idea is that the DEAD box family of proteins are involved in coupling transcription with RNA processing especially with mature mRNA production. DEAD box proteins are feeble helicases (Cordin *et al., 2006)*. Apart from acting as *bona fide* unwindases these proteins are thought to have multiple functions within the cell which may be biochemically independent of their roles of unwindases. They are thought to act in key steps in RNA metabolism and their substrates are predicted to be short duplexes.

Three RNA binding zinc finger proteins **ZNF143**, **ZNF828** and **ZC3H15** of the CCHC-type, were also identified as interacting protein partners of POLS (Table 5.3). **ZNF143 (hStaf)** is capable of stimulating transcription from an mRNA-type pol II promoter. It is known to enhance the transcription of the snRNA, snoRNA-type genes (Schuster *et al.*, 1995). Expression of BubR1, a spindle checkpoint protein, is also known to be heavily dependent on ZNF143 (Myslinski *et al.*, 2007). No experimental data is available regarding **ZC3H15. ZNF828** is required for the alignment for chromosomes at metaphase and their accurate segregation. ZNF828 maintains the spindle microtubule attachment to the kinetochore during sister chromatid biorientation (Itoh *et al.*, 2011). The zinc finger proteins ZNF828, ZC3H15 and ZNF143 identified in this study as interacting proteins with POLS, show poor alignment with *S.cerevisiae* Air1 or Air2 (Table 5.2). Alignment scores tabulated in Table 5.2 were calculated by multiple alignments using the ClustalW programme. ClustalW multiple protein alignment scores are calculated by the following criteria;

1. All sequences are compared to each other (pairwise alignment).
2. A dendrogram is constructed describing the approximate grouping of the sequence by similarity.
3. The final multiple alignment is carried out, using the dendrogram as a guide.

The poor alignment of zinc finger proteins identified in this study is a possible indication that POLS may have different functions to TRAMP4 and TRAMP5. BLAST searches have not identified orthologues of Air1 and Air2 apart from ZCCHC7.

	Air 1	Air2	ZNF143	ZNF828	ZC3H15
Air1	-	44	3	9	2
Air2	44	-	8	9	10
ZNF143	3	8	-	3	3
ZNF828	9	9	3	-	11
ZC3H15	2	10	3	11	-

Table 5.2 Relatedness of POLS-interacting zinc finger proteins with yeast Air1 and Air2.

Alignment scores were calculated by multiple alignments using the ClustalW programme for *H.sapiens* ZNF828, ZC3H15 and ZNF143 identified in this study with *S.cerevisiae* Air1 and Air2.

RNA helicases and zinc finger proteins	Function
ZNF143	enhances transcription of sn/sno RNA genes
ZNF828	required for alignment of chromosomes in metaphase
ZC3H15	unknown to date
DDX17	splicing
DDX42	binds dsRNA
DHX9	transcriptional regulator
SKIV2L2	recruits nuclear exosome to its substrates

Table 5.3 RNA helicases and zinc finger proteins identified as interacting proteins with POLS by MS analysis
Proteins listed had one or more unique peptides and an ion score of ≥45. Proteins are listed in descending order of ion score.

The DEAD box RNA helicases DDX17, DHX9 and ZC3H15 identified as interactors may unwind DNA:RNA hybrids which would result in a substrate for POLS. Identification of the hMTR4/SKIV2L2 RNA helicase suggests the existence of a possible human TRAMP-like complex involving POLS.

5.5.2 RNA metabolism-related proteins

Among the proteins identified as interacting protein partners of POLS were factors associated with splicing, transcription elongation, pre-mRNA cleavage and polyadenylation and histone H1; these are tabulated in Table 5.4. The functions they are associated with are indicated.

Gene	Function
BAT1	spliceosome RNA helicase
SWI/SNF complex	splicing
Histone H1.4	condensation of nucleosome chains
Histone H1.2	condensation of nucleosome chains
Histone H1.3	condensation of nucleosome chains
SF3B3	splicing factor
SF1	splicing factor 1
THO complex (SMARCC1)	regulates transcription of certain genes
SUPT6H	elongation
SFRS2	splicing factor
CF I_m 68	cleavage and polyadenylation factor
U2AF2	splicing factor
YWHAQ14-3-3	mRNA turnover
U4/U6 snRNP	snRNP splicing
HDAC2	transcriptional repressor, deacetylates core histones
RRP42	core exosome component
XRN1	mRNA turnover-5' to 3'cytoplasmic exonuclease
CCR4-NOT	mRNA deadenylation
PRPF19	pre-mRNA processing and splicing factor
CF I_m 59	cleavage and polyadenylation factor
TAF4	transcription initiation factor
	.

Table 5.4 Transcription-related proteins identified as interacting proteins with POLS by MS analysis.
Proteins listed in black type had an ion score of ≥ 45 and more than one unique peptide. Proteins listed in green type had an ion score of ≥ 35 and one unique peptide. Proteins are listed in descending order of ion score.

U6 TUTase, also known as star-PAP, was shown previously to co-immunoprecipitate with CPSF-73 (Mellman *et al*., 2008). To investigate the possibility that POLS might be associated with cleavage and polyadenylation, immunoprecipitation of CPSF-73 from HEK-293T cells and immunoblotting with POLS was performed. A positive interaction was detected (Figure 5.3). As the efficiency of the target protein immunoprecipitation was not analyzed no conclusions could be drawn with regards to the strength of the interaction. The interaction between CPSF-73 and POLS may be so weak that it was not identified in the TAP-tag/MS data. CPSF-73 was used as it was an antibody available for performing a preliminary experiment.

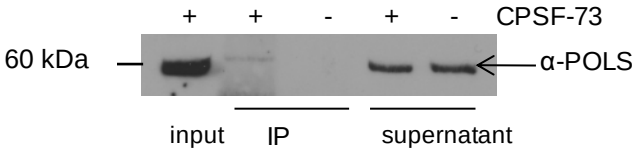


Figure 5.3 Analysis of POLS and CPSF-73 interaction.

Western blot analysis of anti-CPSF-73 immunoprecipitates with POLS antibody. Input corresponds to 5% of cell extract, IP; bound material and supernatant; unbound material.

My TAP tag data identified the proteins **cleavage factor [CF] I_m 68, (CPSF6)** and **CF I_m 59 (CPSF7).** CFI_m made up of subunits CF I_m 25**,** CF I_m 59 and CF I_m 68 which interacts with the splicing factor **U2AF2,** also identified here as an interacting protein of POLS. **DDX42** helicase is known to interact with splicing factor **SF3B3**; both were identified as interactors of POLS in this study. DDX42 is known to control two opposing activities; ATP triggers RNA strand separation while ADP triggers annealing of the complementary RNA strands.

5.5.3 DNA repair-related proteins

Most DNA damage responses are centred on a pair of related protein kinases called ataxia telangiectasia mutated (ATM) and ATR (ATM and Rad3 related). Association of ATR or ATM with damaged DNA leads to their phosphorylation of regulatory proteins and ultimately triggers the activation of two other protein kinases called Chk1 and Chk2. These initiate signalling pathways that inhibit cell cycle progression and stimulate the expression of large numbers of genes associated with DNA repair (Bartek and Lukas, 2007)

An unexpected finding was the association of POLS with proteins involved in DNA repair pathways. Proteins identified are listed in Table 5.5, with associated functions of these proteins.

DNA repair-related	Function
DNA-PK	NHEJ and V(D)J
RPA1	binds single stranded DNA.
Ku80	NHEJ and V(D)J
CHD4	associated with NuRD complex (transcriptional regulator)
MTA2	associated with NuRD complex (transcriptional regulator)
MRE11	component of the MRN complex
RAD50	component of the MRN complex
MSH6	involved in MMR
Ku70	NHEJ and V(D)J
PARP1	repairs single stranded breaks, interacts with BRCA1 and BRCA2
BRCA1	HR and transcriptional regulator
BRCA2	HR

Table 5.5 DNA repair-related proteins identified in this study as interacting proteins of POLS.
Proteins listed had an ion score of ≥45 and one or more unique peptides**.** Proteins are listed in descending order of ion score. BRCA1 and BRCA2 had 22 and 16 respectively, related peptides identified.

BRCA1 and **BRCA2,** both associate with the HR machinery, although their roles are currently incompletely understood. In this study BRCA1 and BRCA2 (number of peptides 22 and 16 respectively) were identified as possible interacting partners of POLS. In order to test if an association between POLS and BRCA1 could be detected, co-immunoprecipitation was performed with BRCA1 from HEK-293T and immunoblotting with POLS antibody. An interaction was identified with POLS (Figure 5.4). As the efficiency of the target protein immunoprecipitation was not analyzed no conclusions could be drawn with regards to the strength of the interaction. This experiment was only performed once with an antibody recommended for western blotting. It is difficult estimate the level of interaction as a distinct band is not visible. It is possible that an antibody recommended to perform co-immunoprecipitations would yield a better interaction.

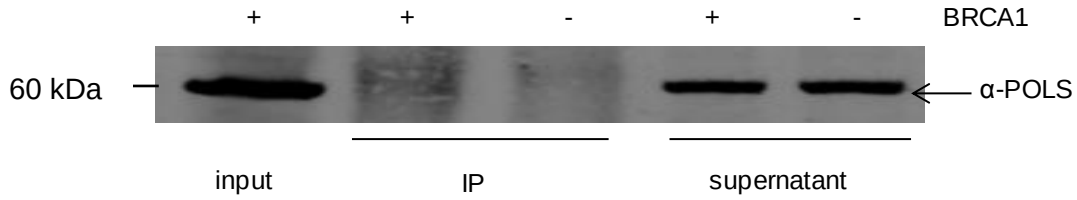


Figure 5.4 Analysis of POLS and BRCA1 interaction.
Western blot analysis of anti-BRCA1 immunoprecipitates with POLS antibody. Input corresponds to 5% of cell extract, IP; bound material and supernatant; unbound material.

Pol II has been shown to interact with the transcriptionally active C-terminal domain of BRCA1. BRCA1 is not itself a DNA-binding sequence-specific transcription factor, although it has been documented as regulating DNA repair, cell cycle and apoptosis. Thus, these functions could be attributed to BRCA1 acting as a coregulator probably *via* factors involved in transcriptional machinery.

The **Ku70/80** heterodimer binds to double stranded ends of the DSB creating a landing pad for **DNA-dependent protein kinase (DNA-PK)**. The interaction with POLS proved unspecific (Figure 5.5). In this experiment a transient transfection of POLS-TAP was performed to look at the interaction.

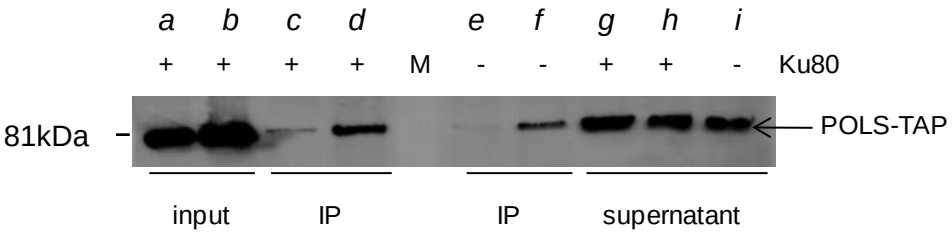


Figure 5.5 Analysis of POLS-TAP and Ku80 interaction.
Western blot analysis of anti-Ku80 immunoprecipitates with TAP antibody. Input corresponds to 5% of cell extract, IP; bound material and supernatant; unbound material. The individual lanes correspond to the following; a,c,e,g and i to pcDNA-*POLS*-TAP, b, d, f and h; pcDNA-TAP.

To analyse the POLS-RAD50 interaction a stable transfectant expressing HA-epitope tagged POLS was used. The HA antibody detected a band in the stable transfectant but not in U2-OS cell lysate and immunoprecipitate (Figure 5.6) confirming the specificity of the antibody.

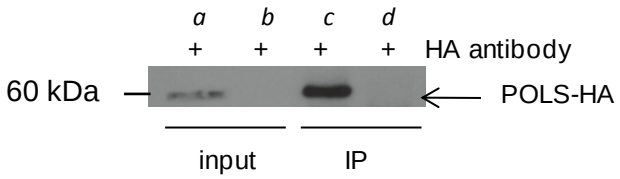


Figure 5.6 Detection of immunoprecipitated POLS-HA.
Western blot analysis of anti-HA immunoprecipitates with HA antibody. Input corresponds to 5% of cell extract and IP; bound material *a, c* ; stable clone K10 expressing pcDNA-POLS-HA, *b, d*; U2-OS cell lysates.

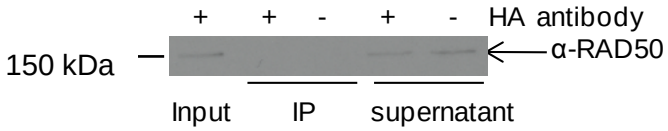


Figure 5.7 Analysis of POLS-HA and RAD50 interaction.
Western blot analysis of anti-HA immunoprecipitate with RAD50 antibody. Input corresponds to 5% of cell extract, IP; bound material and supernatant; unbound material.

No interaction of RAD50 with POLS was observed (Figure 5.7). This could be due to the epitope tag interfering with the interaction, differences in the sensitivity of the methods or differences in stringency.

5.5.4 Nuclear structure-related proteins.

In addition to transcription and DNA repair-related proteins, nuclear pore complex, cohesin-related and nuclear structural maintenance proteins were detected (Table 5.6).

Gene	Function
Importin 5	nuclear transport receptor
Importin 7	nuclear transport receptor
SMC1	structural protein forms part of cohesin
NUP85	forms part of the nuclear pore
NUP214	forms part of the nuclear pore
NUP35	forms part of the nuclear pore
NUP205	forms part of the nuclear pore
SMC3	structural protein forms part of cohesin
SMC2	structural protein forms part of the condensin
KIF4A	generate directional movement along microtubules
NUP50	forms part of the nuclear pore
NUP133	forms part of the nuclear pore
SMC4	structural protein forms part of condensin
TPR	nuclear pore complex related

Table 5.6 Cohesion, nuclear structure and nuclear pore related proteins identified as interacting proteins with POLS by MS analysis.
The proteins listed in black type had an ion score of ≥45 and more than one unique peptide. The proteins listed in green type had an ion score of ≥35 and one unique peptide. Proteins are listed in descending order of ion score.

One function of Cid14 is to maintain nucleolar structure (Win *et al.*, 2006) and Trf4 is associated with sister chromatid cohesion (Wang *et al.*, 2000). Therefore it is possible that POLS has roles in maintaining nuclear structure and sister chromatid cohesion.

Sister chromatid cohesion is required for chromosomes to bi-orient and segregate at random. **Smc1**, **Smc2** ,**Smc3** and **Smc4** identified in this study as interactors of POLS are members of a family of coiled-coil ATPases required for recombination, cohesion and condensation. Smc1 and 3 are components of cohesin; Smc2 and Smc4 are subunits of condensin.

5.6 Discussion

Many candidate protein interacting partners of POLS were identified in this study, which included proteins with roles in transcription, DNA repair, nuclear structure, nuclear pores and chromosome structure. It was striking that almost all of these proteins had a nuclear function, consistent with nuclear functions assigned to other known orthologues of POLS. The recent Lubas *et al.* (2011) study reported nuclear/nucleolar roles for PAPD5.

The mass spectrometry data obtained in this thesis indicate that at least three complexes are formed in association with POLS in the human nucleus. The first involved in RNA splicing which may include the BAT1 and/or DDX42 RNA helicase, zinc finger protein ZN143 and POLS. The second could be associated with DNA repair-related and possibly include DHX9, zinc finger protein ZC3H15 and POLS. A third complex involving POLS could have a function in chromosome structure including SKIV2L2 and zinc finger protein ZNF828. It is possible that the interactions identified perform individual functions in association with POLS, form sub complexes which interact with each other or function on their own. It is also possible that all the proteins are involved in a single complex.

The mass spectrometry data of this thesis indicate POLS involvement in pre-mRNA biology based on the numerous transcription-related proteins identified as interacting partners. It was surprising to find many splicing-related proteins in this study. These included splicing factors, splicing related RNA helicases and pre-mRNA processing factors. Trf4 is known to be involved in the degradation of very short-lived cryptic pol II transcripts (Wyers *et al.*, 2005) and in the removal of spliced out introns (San Paolo *et al.*, 2009). In the recent Lubas *et al.* (2011) study PAPD5 has been implicated in pre-mRNA biology, as splicing-associated proteins were identified in their SILAC/MS data.

Cleavage and polyadenylation factors, transcription initiation factors, elongation and mRNA maturation-associated proteins were also identified as interacting partners of POLS. The

interaction detected between POLS and CPSF-73 links POLS to cleavage and polyadenylation, further strengthening the case that POLS is involved in pre-mRNA biology.

CF I_m68 and CF I_m59, identified here as interactors of POLS, may have roles in alternative polyadenylation. The CF I_m complex acts with U2AF2 splicing factor. Splicing factors SF1, SFRS2, pre-mRNA spilicing and processing factors PRPF 19 were also identified as interacting proteins with POLS. Thus, POLS may have a role in alternative splicing and polyadenylation.

Identification of the major cytoplasmic deadenylase CCR4-NOT, the exonuclease XRN1 and YWHAQ14-3-3 proteins suggests a cytoplasmic role for POLS. POLS may be associated with mRNA deadenylation and decay in the cytoplasm. Identification of RRP42 core exosome component and exonuclease XRN1 indicates that the degradation may be 3'-5' and 5'-3'.

Another unexpected finding was the association of POLS with DNA repair proteins. The DNA repair proteins identified are involved in the base excision repair (BER), nucleotide excision repair (NER), homologous recombinational repair (HR), non-homologous end joining (NHEJ) and mismatch repair (MMS) pathways. A possible interaction of POLS with BRCA1 may have been identified by co-immunoprecipitation in this study. BRCA1 has been shown to have a transcriptional regulatory role. CHD4 and MTA2 also identified as interactors of POLS are associated with NuRD complex which regulates transcription. The DEAD box RNA helicases are capable of unwinding DNA:RNA hybrids, potentially creating an RNA substrate for POLS.

The identification of SMC1, SMC2, SMC3 and SMC4 cluster indicates strong links with sister chromatid cohesion and condensation, and POLS. Both Trf4 and Cid14 mutants show defects in chromosome segregation. It could be envisaged that with transcription read

through in convergent genes, the formation of dsRNA creates a substrate for POLS. Cohesion ensures faithful chromosome segregation to maintain genome integrity.

NUP proteins and POLS may have roles in maintaining nucleolar structure, as Cid14 is required to maintain nucleolar structure. POLS looks to be involved in nuclear transport by the importin family of proteins identified from MS data obtained in this study.

A possible model is proposed from the data obtained from this study Figure 5.8.

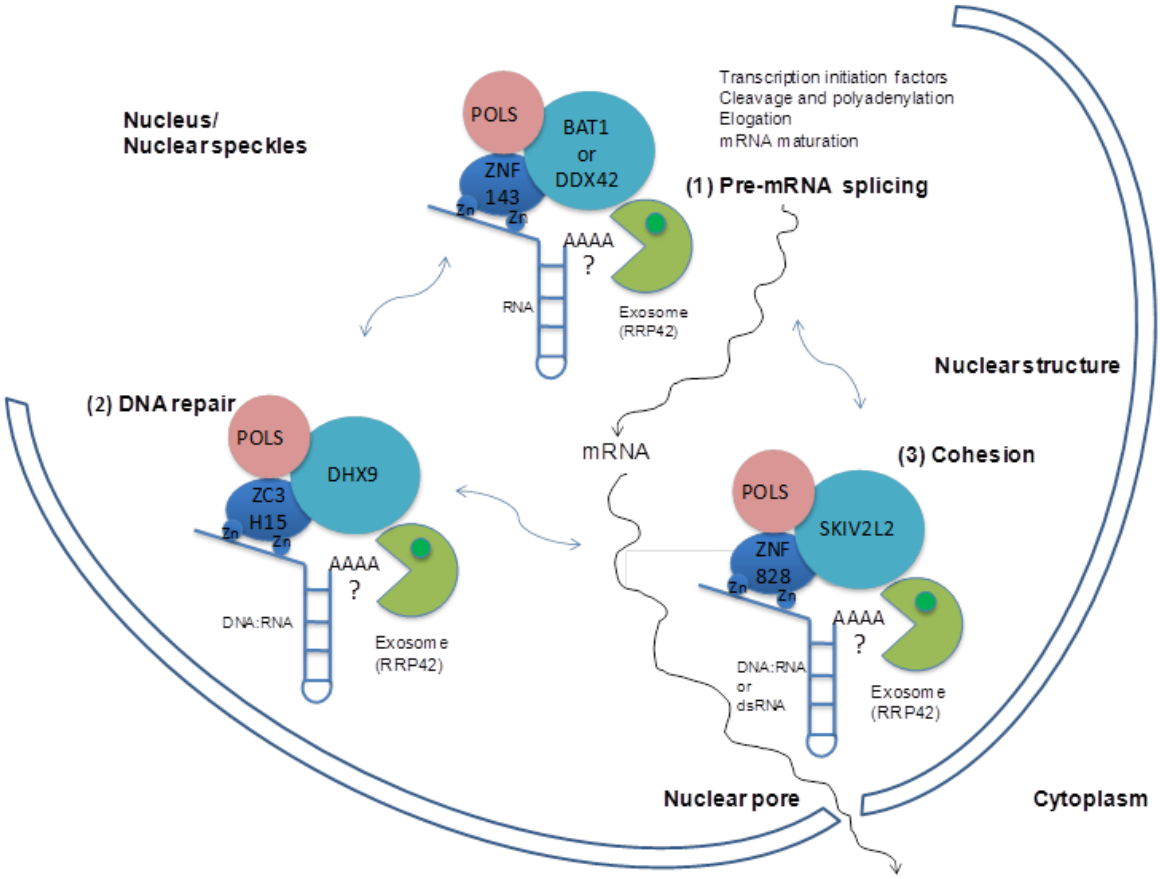


Figure 5.8 Speculative model of the roles of POLS.
This model proposes the involvement of human POLS in a minimum of three complexes in the nucleus. The first involved in pre-mRNA splicing which would include the RNA helicases BAT1 or DDX42, zinc finger ZN143 and POLS. The second POLS associated TRAMP-like complex would include the RNA helicase DHX9 and zinc finger ZC3H15 involved in DNA repair. A third TRAMP like complex associated with the RNA helicase SKIV2L2, zinc finger ZNF828 and POLS is proposed. It is unknown if the speculated polyadenylation activity of POLS is needed to perform the associated functions with these complexes. It is possible that the proteins identified act individually, in sub complexes or all interactions are brought together to perform a single function.

To date the only predicted RNA helicase associated with POLS is SKIV2L2 by analogy to the yeast Mtr4. Trf4 has been reported to be involved in sister chromatid cohesion. Thus it would be consistent to place SKIV2L2 in a TRAMP-like complex associated with cohesion. DNA repair and pre-mRNA splicing are two unidentified functions for Trf4/5 orthologues, although PAPD5 has been speculatively linked to pre-mRNA biology. Apart from these functions POLS seems to be associated with transcription initiation, cleavage and polyadenylation, elongation, mRNA export, nuclear structure and nuclear pores. At present it is unknown if POLS has a role in adenylation of these substrates.

It is difficult to envisage a process that requires the proposed nucleotidyl transferase activity in the nucleus for the proposed model. A recent report indicated that TRAMP4 and TRAMP5 stimulate the degradation of spliced-out introns *via* a mechanism that does not require the polyadenylation activity of Trf4 (San Paolo *et al.*, 2009). Thus, for some POLS associated functions the catalytic activity may not be required.

The weak interaction with CPSF-73 detected by immunoprecipitation studies is encouraging while no firm conclusion can be reached regarding BRCA1 interaction.

6 Novel splice variants identified in an expression screen of human non-canonical poly(A) polymerases

6.1 Introduction

The human genome encodes seven non-canonical poly(A) polymerases belonging to the DNA polymerase β superfamily. The domain organization of the proteins encoded by these genes is shown in Figure 6.1.

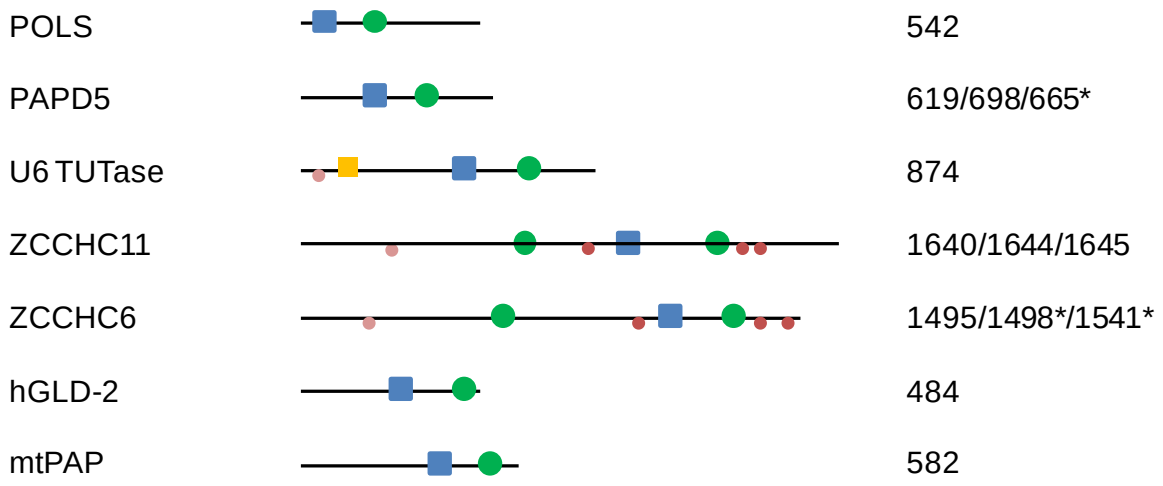


Figure 6.1 The domain organization of the human non-canonical poly(A) polymerases. Indicated are; the catalytic domain (blue), the PAP-associated domain (green), the RRM domain (yellow), C2H2 zinc fingers (pink) and the CCHC zinc fingers (deep red). The size of each protein (in aa) is indicated. The asterisks indicate novel splice variants identified in this study.

6.2 Aim

To determine whether the seven human non-canonical poly(A) polymerases are expressed at a detectable level in human cells.

6.3 Expression screen of the human non-canonical poly(A) polymerases

To investigate the expression of the human non-canonical poly(A) polymerases a semi-quantitative RT-PCR screen of the 3' regions of these genes were carried out in four human cell lines. Cervical carcinoma (Hela), osteosarcoma (U2-OS), non-small cell lung carcinoma (H1299) and large cell lung carcinoma (NCI-H460) cell lines were used in this study. A range of cycle numbers was used to quantify the product. For a detailed description of the assay performed see methods section 2.2.5. All primers used are tabulated in Table 2.2. All genes were found to be expressed at easily detectable levels.

For ZCCHC6 and PAPD5 more than one product was visible. By TA cloning and sequencing these products were found to be splice variants of these genes coding for different proteins. The splice variant of PAPD5 encoded a 665 aa protein (Figure 6.2) while two ZCCHC6 splice variants encoding for proteins of 1498/1541 aa were identified (Figure 6.3 and Figure 6.4 respectively).

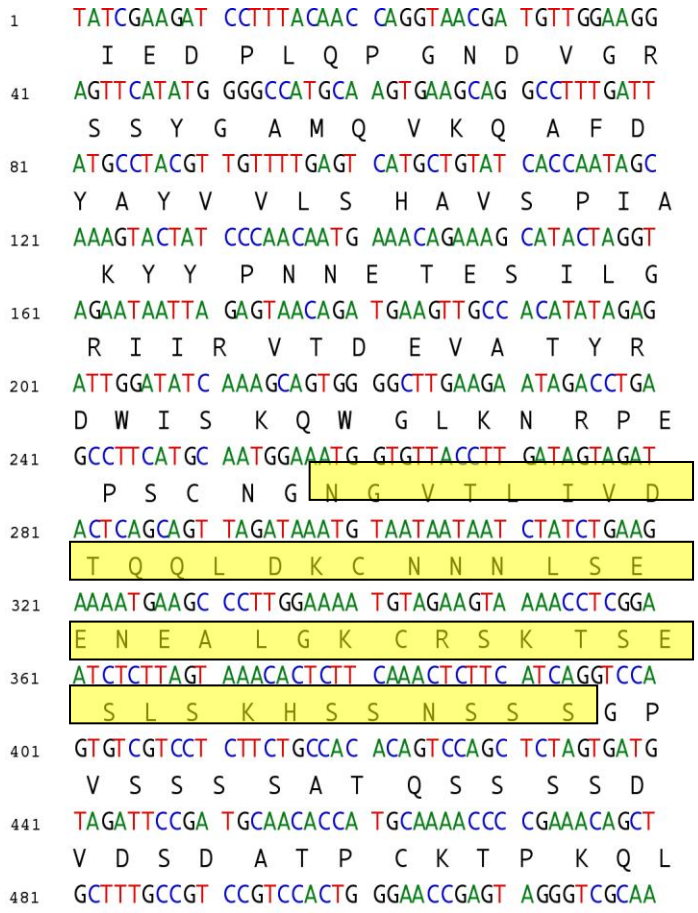


Figure 6.2 Splice variant of PAPD5 identified in this study.

ENSEMBL lists two sequences of 619 and 698 aa for PAPD5. In this study an additional splice variant coding for a protein of 665 aa was identified. Highlighted in yellow are the extra aa encoded by the novel splice variant.

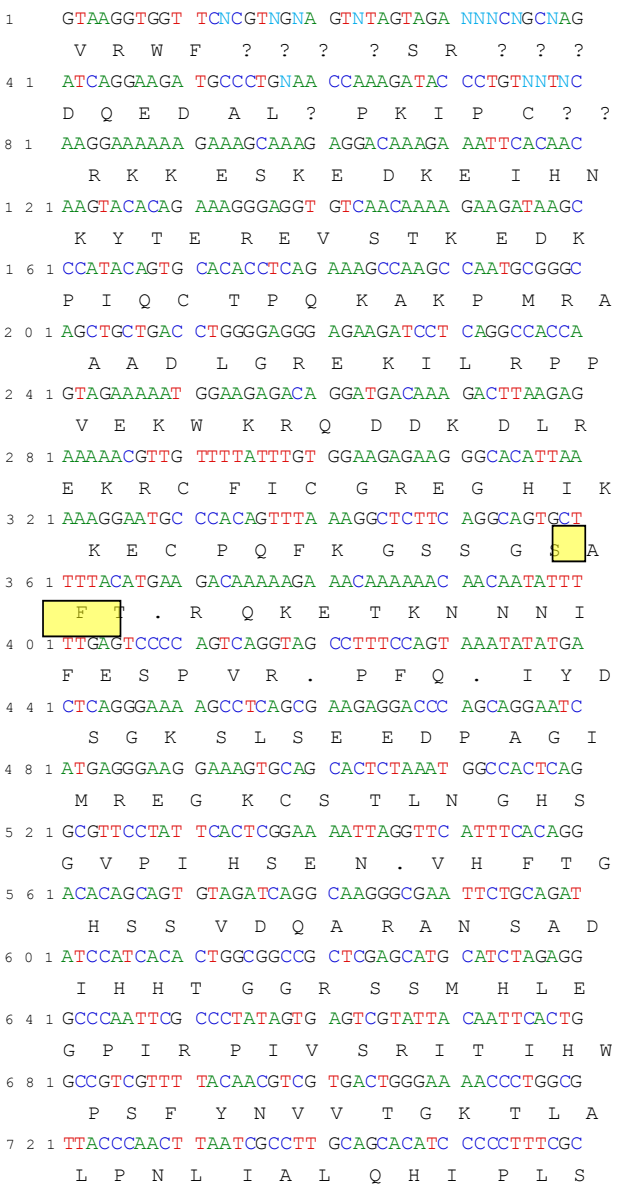


Figure 6.3 The first ZCCHC6 splice variant identified in this study.

ENSEMBL assigns a sequence of 1495 aa for ZCCHC6. In this study a splice variant coding for a protein of 1498 was identified. Although 63 additional bases were noted, due to the presence of an in frame stop codon only 3 aa are incorporated in the novel protein. Highlighted in yellow are the extra aa incorporated in the variant protein.



Figure 6.4 The second splice variant of ZCCHC6 identified in this study.

To the ENSEMBL assigned sequence of 1495 aa for ZCCHC6, this splice variant adds 46 additional aa, encoding a protein of 1541 aa. Highlighted in yellow are the extra aa encoded by the novel splice variant.

6.4 Discussion

The splice variants identified in this study are not reported to date in public databases. The splice variants identified for PAPD5 and ZCCHC6 contain no known protein domains. At present the function of ZCCHC6 is unknown. PAPD5 has been linked to rRNA processing and PROMTs degradation. The regions responsible these activities are unknown. Thus, as the novel splice variant contains no distinct proteins domains it is difficult to speculate on its function. Narry Kim's laboratory has reported three splice variants of ZCCHC11 giving rise to cDNAs of 1640, 1644 and 1645 bp. The uridylation activity of ZCCHC11 was reported to be the same for all three isoforms (Heo *et al.*, 2009). It is possible that there are more to date unidentified splice variants of these human non-canonical poly(A) polymerases. These splice variants could have distinct functions and activities. However, this possibility was not addressed in this present study.

7 Conclusions

This thesis has concentrated on POLS and PAPD5, assigned on the basis of multiple sequence alignments as the human orthologues of *S.cerevisiae* Trf4 and Trf5. Their cellular and biochemical functions were explored in some detail. Protein-protein interactions with POLS were identified by MS analysis. Co-IPs identified possible interactions of POLS with CPSF-73 and BRCA1.

At the outset of this study the localization of POLS was unknown. Thus as a first step towards the characterization of the cellular functions of POLS its localization was sought. In stable transfectant K4, expressing POLS-HA, POLS was found to localize to nuclear speckles, with an accumulation seen in the nucleolus in some cells. This is consistent with most of the interacting protein partners identified in the MS data, which have nuclear functions. Recently Lubas *et al*. (2011), claimed PAPD5 to be a nuclear/nucleolar protein based on nuclear factors found to be immunoprecipitating with PAPD5, along with the immunofluorescence localization detected with the protein encoded by their cDNA clone. Mullen and Marzluff (2007) report PAPD5 to be in the cytoplasm, with some staining visible in the nucleus. In this study for stable transfectant K10 expressing POLS-HA, POLS was found to be mainly cytoplasmic with significant but less intense staining in the nucleus. TAP- or HA-tagged transiently expressed POLS showed a similar localization. It is possible that when POLS is over-expressed, much of it is sequestered in the cytoplasm in an inactive state, while the minor nuclear fraction is biologically active. If stable transfectant K10 is representative of physiological expression of POLS, it may have functions in both the nucleus and cytoplasm. Three proteins, CCR4-NOT, XRN1 and YWHAQ14-3-3, were identified as POLS interactors in this study and to date have been described as having solely cytoplasmic functions. However, it is conceivable that the variation in localization is due to clonal differences between the stable transfectants K4 and K10 that are unrelated to the level of expression of POLS. The minor (<0.2%) population of cells displaying a nuclear speckled

pattern of POLS-HA localization in the transient transfection may represent a short-lived phase of the cell cycle, though the nuclear morphology gives no clues as to what this might be. Alternatively, these cells might have experienced some stochastic event such as spontaneous DNA damage. The confirmed specificity of the POLS antibody generated in this study opens the way to resolving the ambiguity with regard to the localization of POLS, as well as opening up other avenues of investigation.

The MS data obtained in this thesis indicate a potential involvement of POLS in pre-mRNA splicing. This is a surprising finding, but the nuclear speckle localization would be consistent with POLS involvement in splicing, as pre-mRNA splicing factors are stored and assembled in these structures. Lubas *et al*. (2011) tentatively implicated PAPD5 in pre-mRNA splicing based on their MS/SILAC data. The other orthologues of these proteins have not been associated with pre-mRNA splicing. Trf4 has been linked with spliced out intron removal (San Paolo *et al.*, 2009) and degradation of other Pol II transcripts (Wyers *et al.*, 2005). CF I_m 59 and CF I_m 68, identified here as interactors of POLS, have roles in canonical polyadenylation, and alternative polyadenylation. Thus, POLS may be associated with alternative polyadenylation and splicing. Alternative splicing and polyadenylation can vary the proteome output, hence understanding the molecular basis of this process is of prime medical importance. Many human diseases are associated with aberrant splicing (Cooper *et al.*, 2009; Dhir and Buratti, 2010). Alternative polyadenylation has been associated with tumorigenesis (Mayr and Bartel, 2009). The human THO complex, also identified as an interactor of POLS, co-localizes with splicing factors in nuclear speckles (Masuda *et al*., 2005). The identification of BAT1 spliceosome RNA helicase, SF3B3, SF1, SFRS2, U2AF2, PRPF19 splicing factors and the SWI/SNF complex all strengthen a role for POLS in pre-mRNA splicing.

Preliminary evidence of an interaction with cleavage and polyadenylation factor (CPSF)-73 was obtained in this study, along with the identification of the TAF4 transcription initiation

factor and SUPT6H, which is involved in transcription elongation, all suggestive of further roles for POLS in mRNA biology.

Another surprise finding was the association of POLS with DNA repair-related proteins. It is difficult to envisage a direct role for an RNA polymerase in DNA repair. A possible POLS interaction with BRCA1 was detected in this study. BRCA1 has been shown to be involved in regulation of transcription. CHD4 and MTA2 identified as interactors of POLS are associated with the NuRD complex, which regulates transcription. It is possible that the other DNA repair-related proteins identified here are also involved in as yet unidentified transcription regulation-related processes. It may be that POLS has an indirect role in DNA repair and it would be fascinating to identify the processes involved.

The identification of SMC1 and SMC3 indicate that POLS is involved in cohesion, while SMC2 and SMC4 link POLS to condensation of the chromosomes. A role in nuclear transport is suggested by the importin family members identified as interactors of POLS. Interactions with NUP proteins may reflect transport of POLS through the nuclear pore, of which the identified proteins are structural components, or a potential role of POLS in maintaining nuclear pore structure.

The DEAD box RNA helicases DDX17, DHX9 and ZC3H15, identified as POLS interactors, may unwind DNA:RNA hybrids which would result in a substrate for POLS. Identification of the hMTR4/SKIV2L2 RNA helicase suggests the existence of a possible human TRAMP-like complex. However the zinc finger protein ZCCHC7, which shares 34% identity with Air2, was not detected in this experiment. BLAST searches have not identified any other human orthologues of Air1 or Air2. Thus, the zinc finger proteins identified in this study, ZNF828, ZNF143 and ZC3H15, may perform functions different from those of Air1 and Air2.

The identified interactors could interact individually with POLS, form sub-complexes which interact with each other (Figure 5.8) or could be involved in a single complex. Pre-rRNA transcripts polyadenylated within the human rRNA 3'ETS were detected for the first time in this study. In the course of investigating POLS and PAPD5 involvement in the oligoadenylation of transcripts it became apparent that the other nucleotides incorporated within the A-rich tails were PCR-induced artefacts. Thus, POLS and PAPD5 do not appear to add non-A residues to the tail sequences found within the human 3'ETS. Based on the data obtained from this thesis, human β-actin tails contain pure A tails. Other nucleotides incorporated within the A-rich tails were found to be PCR-induced artefacts for β-actin too.

Data from 3' RACE experiments in this study provided no evidence that POLS and/or PAPD5 are responsible for adding poly(A) tails to the 3'ETS transcripts, though the possibility that this is due to POLS- and/or PAPD5-related transcript turnover being very rapid cannot be ruled out. It might be necessary to inhibit exosome-mediated degradation and processing of transcripts in order to detect intermediate transcripts in greater abundance. The exosome components EXOC 10 (hRrp6), EXOC 7(hRrp42) and EXOC 8 (hRrp43) were found to interact with POLS, indicating the possible involvement of POLS in the 3'-5' exosome-mediated degradation pathway.

Down-regulation of POLS and PAPD5 did not appear to alter the positioning of the poly(A) tails found within the 3'ETS transcript, as no difference in mean residual ETS length was detected in the gene down-regulated *vs* control populations.

Northern analysis indicated that the polyadenylated 3'ETS transcripts detected had not been cleaved at the 3' end of the 28S rRNA. They were smaller than the 47S transcript, presumably due to 5'-3' degradation. XRN1, which was identified as an interactor of POLS, may be associated with 5'-3' degradation of these transcripts. It is possible that POLS may have a role in regulating 5'-3' degradation.

Individual down-regulation of POLS or PAPD5 and the DKD (POLS+PAPD5) resulted in a slow growth phenotype. In the individual knock-down no increase in dead cells was recorded, although a difference in biomass was recorded by the MTT-type cell growth assay. No difference in cell cycle profile was noted for POLS down-regulated *vs* control cells. In the DKD a lower biomass was recorded, along with an increase in dead cells as quantified by the trypan blue and flow cytometric analyses. An accumulation of cells in G2/M was visible in the DKD. The accumulation of dead cells in the DKD, as with the budding yeast *trf4 trf5* double mutant, indicates that *POLS* and *PAPD5* share an essential redundant function.

5-FU treated cells showed an accumulation of POLS in the nucleolus at concentration of 2.5-25 µM. If 5-FU is targeting a redundant function of POLS and/or PAPD5, POLS may be recruited to the nucleolus *via* a mechanism that could restore this function. Both murine and human PAPD5 are associated with the degradation of incomplete Pol I transcripts (Shcherbik *et al.*, 2010; Lubas *et al.,* 2011). 5-FU has been shown to affect the processing of the 47S transcript. Thus, it is possible that a role played by PAPD5 in pre-rRNA processing is affected by 5-FU. The DKD cells were hypersensitive to 5-FU at a concentration of 6.25 µM, which may be the crucial concentration where POLS is unable to support the function of PAPD5 if the above hypothesis is correct.

During the expression screen of the human non-canonical poly(A) polymerases a novel splice variant of PAPD5 coding for a 665 aa protein was identified. For ZCCHC6 two novel splice variants coding for proteins of 1498 and 1541 aa were detected. These splice variants may have different functions but this possibility was not further investigated in this study.

7.1 Future directions

This thesis has suggested many exciting and novel roles for POLS. Many possible interactions have been identified which could be of medical importance.

The confirmed specificity of the polyclonal antibody should enable clarification of the POLS subcellular localization. This would enable biological functions of POLS to be assigned with greater confidence. Exploring the following topics would lead into many exciting areas of research and solve many unanswered questions. Confirmation of the interactions and elucidating the possible mechanisms involved will be of prime interest from a basic science perspective, as well as identifying potential drug targets relevant to a range of diseases. Suggested areas of research based on the data obtained from this thesis would include:

1. POLS interactions with pre-mRNA splicing,
2. POLS interactions with DNA repair-related proteins,
3. POLS interactions with cohesin and condensin subunits.
4. POLS interactions with histone H1 and HDACs.
5. Further investigating the relationship between POLS, PAPD5 and 5-FU.

From the results obtained in this thesis it is clear that while some functions are conserved from yeast to human cells, POLS has diverged considerably from its yeast counterparts. This opens up a range of very exciting areas of research which would be of imperative medical importance.

References

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999a).Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J *18*, 5399-5410.

Allmang, C., Mitchell, P., Petfalski, E., and Tollervey, D. (2000). Degradation of ribosomal RNA precursors by the exosome. Nucleic Acids Res *28*, 1684-1691.

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999b). The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. Genes Dev *13*, 2148-2158.

Amarasinghe, G.K., De Guzman, R.N., Turner, R.B., Chancellor, K.J., Wu, Z.R., and Summers, M.F. (2000). NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. J Mol Biol *301*, 491-511.

Anderson, J.T., and Wang, X. (2009). Nuclear RNA surveillance: no sign of substrates tailing off. Crit Rev Biochem Mol Biol *44*, 16-24.

Aphasizhev, R. (2005). RNA uridylyltransferases. Cell Mol Life Sci *62*, 2194-2203.

Armstrong, R.D., Lewis, M., Stern, S.G., and Cadman, E.C. (1986). Acute effect of 5-fluorouracil on cytoplasmic and nuclear dihydrofolate reductase messenger RNA metabolism. J Biol Chem *261*, 7366-7371.

Bai, Y., Srivastava, S.K., Chang, J.H., Manley, J.L., and Tong, L. (2011). Structural basis for dimerization and activity of human PAPD1, a noncanonical poly(A) polymerase. Mol Cell *41*, 311-320.

Barnard, D.C., Ryan, K., Manley, J.L., and Richter, J.D. (2004). Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. Cell *119*, 641-651.

Benoit, P., Papin, C., Kwak, J.E., Wickens, M., and Simonelig, M. (2008). PAP- and GLD-2-type poly(A) polymerases are required sequentially in cytoplasmic polyadenylation and oogenesis in Drosophila. Development *135*, 1969-1979.

Bilger, A., Fox, C.A., Wahle, E., and Wickens, M. (1994). Nuclear polyadenylation factors recognize cytoplasmic polyadenylation elements. Genes Dev *8*, 1106-1116.

Briggs, M.W., Burkard, K.T., and Butler, J.S. (1998). Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J Biol Chem *273*, 13255-13263.

Buhler, M., Haas, W., Gygi, S.P., and Moazed, D. (2007). RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. Cell *129*, 707-721.

Castano, I.B., Heath-Pagliuso, S., Sadoff, B.U., Fitzhugh, D.J., and Christman, M.F. (1996). A novel family of TRF (DNA topoisomerase I-related function) genes required for proper nuclear segregation. Nucleic Acids Res *24*, 2404-2410.

Chen, C.C., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr Biol *15*, 378-383.

Chicoine, J., Benoit, P., Gamberi, C., Paliouras, M., Simonelig, M., and Lasko, P. (2007). Bicaudal-C recruits CCR4-NOT deadenylase to target mRNAs and regulates oogenesis, cytoskeletal organization, and its own expression. Dev Cell *13*, 691-704.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA *13*, 998-1016.

Cristodero, M., and Clayton, C.E. (2007). Trypanosome MTR4 is involved in rRNA processing. Nucleic Acids Res *35*, 7023-7030.

Dez, C., Houseley, J., and Tollervey, D. (2006). Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of Saccharomyces cerevisiae. EMBO J *25*, 1534-1546.

Dolnick, B.J., and Pink, J.J. (1985). Effects of 5-fluorouracil on dihydrofolate reductase and dihydrofolate reductase mRNA from methotrexate-resistant KB cells. J Biol Chem *260*, 3006-3014.

Doong, S.L., and Dolnick, B.J. (1988). 5-Fluorouracil substitution alters pre-mRNA splicing in vitro. J Biol Chem *263*, 4467-4473.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science *326*, 544-550.

Dziembowski, A., Lorentzen, E., Conti, E., and Seraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat Struct Mol Biol *14*, 15-22.

Edmonds, M. (2002). A history of poly A sequences: from formation to factors to function. Prog Nucleic Acid Res Mol Biol *71*, 285-389

Egecioglu, D.E., Henras, A.K., and Chanfreau, G.F. (2006). Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA *12*, 26-32.

Etheridge, R.D., Clemens, D.M., Gershon, P.D., and Aphasizhev, R. (2009). Identification and characterization of nuclear non-canonical poly(A) polymerases from Trypanosoma brucei. Mol Biochem Parasitol *164*, 66-73.

Fang, F., Hoskins, J., and Butler, J.S. (2004). 5-fluorouracil enhances exosome-dependent accumulation of polyadenylated rRNAs. Mol Cell Biol *24*, 10766-10776.

Fodor, E., Mikulasova, A., Mingay, L.J., Poon, L.L., and Brownlee, G.G. (2000). Messenger RNAs that are not synthesized by RNA polymerase II can be 3' end cleaved and polyadenylated. EMBO Rep *1*, 513-518.

Fox, C.A., Sheets, M.D., and Wickens, M.P. (1989). Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. Genes Dev *3*, 2151-2162.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol Cell *14*, 571-583.

Ghoshal, K., and Jacob, S.T. (1994). Specific inhibition of pre-ribosomal RNA processing in extracts from the lymphosarcoma cells treated with 5-fluorouracil. Cancer Res *54*, 632-636.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nat Rev Mol Cell Biol *9*, 337-344.

Grzechnik, P., and Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Mol Cell *32*, 247-258.

Hadjiolova, K.V., Nicoloso, M., Mazan, S., Hadjiolov, A.A., and Bachellerie, J.P. (1993). Alternative pre-rRNA processing pathways in human cells and their alteration by cycloheximide inhibition of protein synthesis. Eur J Biochem *212*, 211-215.

Hamill, S., Wolin, S.L., and Reinisch, K.M. (2010). Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. Proc Natl Acad Sci U S A *107*, 15045-15050.

Haracska, L., Johnson, R.E., Prakash, L., and Prakash, S. (2005). Trf4 and Trf5 proteins of Saccharomyces cerevisiae exhibit poly(A) RNA polymerase activity but no DNA polymerase activity. Mol Cell Biol *25*, 10183-10189.

Hartung, S., and Hopfner, K.P. (2007). The exosome, plugged. EMBO Rep *8*, 456-457.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696-708.

Hoskins, J., and Butler, J.S. (2008). RNA-based 5-fluorouracil toxicity requires the pseudouridylation activity of Cbf5p. Genetics *179*, 323-330.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nat Rev Mol Cell Biol *7*, 529-539.

Houseley, J., and Tollervey, D. (2006). Yeast Trf5p is a nuclear poly(A) polymerase. EMBO Rep *7*, 205-211.

Houseley, J., and Tollervey, D. (2009). The many pathways of RNA degradation. Cell *136*, 763-776.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225-229.

Jia, H., Wang, X., Liu, F., Guenther, U.P., Srinivasan, S., Anderson, J.T., and Jankowsky, E. (2011). The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. Cell *145*, 890-901.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol *11*, 1157-1163.

Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. Genes Dev *18*, 1227-1240.

Kadaba, S., Wang, X., and Anderson, J.T. (2006). Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. RNA *12*, 508-521.

Kammler, S., Lykke-Andersen, S., and Jensen, T.H. (2008). The RNA exosome component hRrp6 is a target for 5-fluorouracil in human cells. Mol Cancer Res *6*, 990-995.

Kashiwabara, S., Nakanishi, T., Kimura, M., and Baba, T. (2008). Non-canonical poly(A) polymerase in mammalian gametogenesis. Biochim Biophys Acta *1779*, 230-238.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., and Baba, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev *23*, 433-438.

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

Keller, C., Woolcock, K., Hess, D., and Buhler, M. (2010). Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. RNA *16*, 1124-1129.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517-522.

Kwak, J.E., Wang, L., Ballantyne, S., Kimble, J., and Wickens, M. (2004). Mammalian GLD-2 homologs are poly(A) polymerases. Proc Natl Acad Sci U S A *101*, 4407-4412.

Kwak, J.E., and Wickens, M. (2007). A family of poly(U) polymerases. RNA *13*, 860-867.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713-724.

Lafontaine, D.L., and Tollervey, D. (2001). The function and synthesis of ribosomes. Nat Rev Mol Cell Biol *2*, 514-520.

LaRiviere, F.J., Cole, S.E., Ferullo, D.J., and Moore, M.J. (2006). A late-acting quality control process for mature eukaryotic rRNAs. Mol Cell *24*, 619-626.

Lebreton, A., Tomecki, R., Dziembowski, A., and Seraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993-996.

Lenz, H.J., Manno, D.J., Danenberg, K.D., and Danenberg, P.V. (1994). Incorporation of 5-fluorouracil into U2 and U6 snRNA inhibits mRNA precursor splicing. J Biol Chem *269*, 31962-31968.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223-1237.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2010). Interaction profiling identifies the human nuclear exosome targeting complex. Mol Cell *43*, 624-637.

Malet, H., Topf, M., Clare, D.K., Ebert, J., Bonneau, F., Basquin, J., Drazkowska, K., Tomecki, R., Dziembowski, A., Conti, E.*, et al.* (2010). RNA channelling by the eukaryotic exosome. EMBO Rep *11*, 936-942.

Martin, G., and Keller, W. (1996). Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. EMBO J *15*, 2593-2603.

Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138*, 673-684.

Mellman, D.L., Gonzales, M.L., Song, C., Barlow, C.A., Wang, P., Kendziorski, C., and Anderson, R.A. (2008). A PtdIns4,5P2-regulated nuclear poly(A) polymerase controls expression of select mRNAs. Nature *451*, 1013-1017.

Minoda, Y., Saeki, K., Aki, D., Takaki, H., Sanada, T., Koga, K., Kobayashi, T., Takaesu, G., and Yoshimura, A. (2006). A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. Biochem Biophys Res Commun *344*, 1023-1030.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M., and Tollervey, D. (2003). Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. Mol Cell Biol *23*, 6982-6992.

Mitchell, P., and Tollervey, D. (2003). An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation. Mol Cell *11*, 1405-1413.

Morozov, I.Y., Jones, M.G., Razak, A.A., Rigden, D.J., and Caddick, M.X. (2010). CUCU modification of mRNA promotes decapping and transcript degradation in Aspergillus nidulans. Mol Cell Biol *30*, 460-469.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes Dev *22*, 50-65.

Murthy, K.G., and Manley, J.L. (1992). Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. J Biol Chem *267*, 14804-14811.

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

Nagaike, T., Suzuki, T., Katoh, T., and Ueda, T. (2005). Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase. J Biol Chem *280*, 19721-19727.

Nakamura, R., Takeuchi, R., Takata, K., Shimanouchi, K., Abe, Y., Kanai, Y., Ruike, T., Ihara, A., and Sakaguchi, K. (2008). TRF4 is involved in polyadenylation of snRNAs in Drosophila melanogaster. Mol Cell Biol *28*, 6620-6631.

Parker, W.B., and Cheng, Y.C. (1990). Metabolism and mechanism of action of 5-fluorouracil. Pharmacol Ther *48*, 381-395.

Petronczki, M., Chwalla, B., Siomos, M.F., Yokobayashi, S., Helmhart, W., Deutschbauer, A.M., Davis, R.W., Watanabe, Y., and Nasmyth, K. (2004). Sister-chromatid cohesion mediated by the alternative RF-CCtf18/Dcc1/Ctf8, the helicase Chl1 and the polymerase-alpha-associated protein Ctf4 is essential for chromatid disjunction during meiosis II. J Cell Sci *117*, 3547-3559.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851-1854.

Read, R.L., Martinho, R.G., Wang, S.W., Carr, A.M., and Norbury, C.J. (2002). Cytoplasmic poly(A) polymerases mediate cellular responses to S phase arrest. Proc Natl Acad Sci U S A *99*, 12079-12084.

Reis, C.C., and Campbell, J.L. (2007). Contribution of Trf4/5 and the nuclear exosome to genome stability through regulation of histone mRNA levels in Saccharomyces cerevisiae. Genetics *175*, 993-1010.

Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol Cell Biol *27*, 3612-3624.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol *16*, 616-623.

Rouhana, L., Wang, L., Buter, N., Kwak, J.E., Schiltz, C.A., Gonzalez, T., Kelley, A.E., Landry, C.F., and Wickens, M. (2005). Vertebrate GLD2 poly(A) polymerases in the germline and the brain. RNA *11*, 1117-1130.

Rouquette, J., Choesmel, V., and Gleizes, P.E. (2005). Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. EMBO J *24*, 2862-2872.

Sadoff, B.U., Heath-Pagliuso, S., Castano, I.B., Zhu, Y., Kieff, F.S., and Christman, M.F. (1995). Isolation of mutants of Saccharomyces cerevisiae requiring DNA topoisomerase I. Genetics *141*, 465-479.

Saitoh, S., Chabes, A., McDonald, W.H., Thelander, L., Yates, J.R., and Russell, P. (2002). Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. Cell *109*, 563-573.

San Paolo, S., Vanacova, S., Schenk, L., Scherrer, T., Blank, D., Keller, W., and Gerber, A.P. (2009). Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. PLoS Genet *5*, e1000555.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends Biochem Sci *33*, 501-510.

Schmid, M., Kuchler, B., and Eckmann, C.R. (2009). Two conserved regulatory cytoplasmic poly(A) polymerases, GLD-4 and GLD-2, regulate meiotic progression in C. elegans. Genes Dev *23*, 824-836.

Schmidt, M.J., West, S., and Norbury, C.J. (2010). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA *17*, 39-44.

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science *300*, 805-808.

Sheth, U., and Parker, R. (2006). Targeting of aberrant mRNAs to cytoplasmic processing bodies. Cell *125*, 1095-1109.

Silverstein, R.A., Gonzalez de Valdivia, E., and Visa, N. (2011). The incorporation of 5-fluorouracil into RNA affects the ribonucleolytic activity of the exosome subunit Rrp6. Mol Cancer Res *9*, 332-340.

Slomovic, S., Fremder, E., Staals, R.H., Pruijn, G.J., and Schuster, G. (2010). Addition of poly(A) and poly(A)-rich tails during RNA degradation in the cytoplasm of human cells. Proc Natl Acad Sci U S A *107*, 7407-7412.

Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2006). Polyadenylation of ribosomal RNA in human cells. Nucleic Acids Res *34*, 2966-2975.

Staals, R.H., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J., Raijmakers, R., and Pruijn, G.J. (2010) Dis3-like 1: a novel exoribonuclease associated with the human exosome. EMBO J *29*, 2358-2367.

Stead, J.A., Costello, J.L., Livingstone, M.J., and Mitchell, P. (2007). The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. Nucleic Acids Res *35*, 5556-5567.

Takimoto, C.H., Tan, Y.Y., Cadman, E.C., and Armstrong, R.D. (1987). Correlation between ribosomal RNA production and RNA-directed fluoropyrimidine cytotoxicity. Biochem Pharmacol *36*, 3243-3248.

Thiebaut, M., Kisseleva-Romanova, E., Rougemaille, M., Boulay, J., and Libri, D. (2006). Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Mol Cell *23*, 853-864.

Thomson, E., and Tollervey, D. (2010). The final step in 5.8S rRNA processing is cytoplasmic in Saccharomyces cerevisiae. Mol Cell Biol *30*, 976-984.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res *33*, 201-212.

Trippe, R., Guschina, E., Hossbach, M., Urlaub, H., Luhrmann, R., and Benecke, B.J. (2006). Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA *12*, 1494-1504.

van Dijk, E.L., Schilders, G., and Pruijn, G.J. (2007). Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways. RNA *13*, 1027-1035.

van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol Cell Biol *20*, 441-452.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol *3*, e189.

Vasiljeva, L., and Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Mol Cell *21*, 239-248.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008a). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol *15*, 795-804.

Vasiljeva, L., Kim, M., Terzi, N., Soares, L.M., and Buratowski, S. (2008b). Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. Mol Cell *29*, 313-323.

Walowsky, C., Fitzhugh, D.J., Castano, I.B., Ju, J.Y., Levin, N.A., and Christman, M.F. (1999). The topoisomerase-related function gene TRF4 affects cellular sensitivity to the antitumor agent camptothecin. J Biol Chem *274*, 7302-7308.

Wang, S.W., Norbury, C., Harris, A.L., and Toda, T. (1999). Caffeine can override the S-M checkpoint in fission yeast. J Cell Sci *112 (Pt 6)*, 927-937.

Wang, S.W., Stevenson, A.L., Kearsey, S.E., Watt, S., and Bahler, J. (2008). Global role for polyadenylation-assisted nuclear RNA degradation in posttranscriptional gene silencing. Mol Cell Biol *28*, 656-665.

Wang, S.W., Toda, T., MacCallum, R., Harris, A.L., and Norbury, C. (2000a). Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. Mol Cell Biol *20*, 3234-3244.

Wang, Z., Castano, I.B., De Las Penas, A., Adams, C., and Christman, M.F. (2000b). Pol kappa: A DNA polymerase required for sister chromatid cohesion. Science *289*, 774-779.

West, S., Gromak, N., Norbury, C.J., and Proudfoot, N.J. (2006). Adenylation and exosome-mediated degradation of cotranscriptionally cleaved pre-messenger RNA in human cells. Mol Cell *21*, 437-443.

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

Wilkinson, D.S., and Pitot, H.C. (1973). Inhibition of ribosomal ribonucleic acid maturation in Novikoff hepatoma cells by 5-fluorouracil and 5-fluorouridine. J Biol Chem *248*, 63-68.

Win, T.Z., Draper, S., Read, R.L., Pearce, J., Norbury, C.J., and Wang, S.W. (2006). Requirement of fission yeast Cid14 in polyadenylation of rRNAs. Mol Cell Biol *26*, 1710-1721.

Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S.*, et al.* (2002). The genome sequence of Schizosaccharomyces pombe. Nature *415*, 871-880.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B.*, et al.* (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725-737.

Appendix **A**

Oligo/poly-adenylated 3'ETS tails were amplified from $(dT)_{16}$-adaptor primed cDNA with primers 3'ETSF and $(dT)_{16}$-adaptor followed by a second nested PCR with 3'ETSFF and $(dT)_{16}$-adaptor (Table 2.4) and cloned into pCR2.1(Invitrogen). Randomly selected clones were sequenced. Tail lengths were calculated after deducting 16 adenosine residues corresponding to the $(dT)_{16}$ primer.

siRNA DKD (POLS-70% KD+PAPD5-55% KD)			
Clone	Tail	Poly(A) site	Tail-length
1_2	A22	101	22
4_1	A6TA13GA12GA2GATA12TA24	106	76
3_1	A6GA4TA34TA4GA3TA7GA45GA3GA22	108	136
7_5	A91	110	91
2_11	A39GA5GTGA11GA2GA2GA6GA3GA27	111	104
1_10	A43GA5GTGA2GA12GA2GA2GA6GA3GA7GA17	112	110
6_2	TA7GA2TA8TA9TA28	115	58
10_6	TA5T5A39GA34	119	85
1_3	A17GA58	121	76
SI_3	ATA24	123	26

Table (*a*) 3'ETS tails identified in the POLS and PAPD5 gene down-regulated cells.

Control for DKD (POLS+PAPD5)			
Clone	Tail	Poly(A) site	Tail-length
7_E	ATAGA9TA3TA41	102	59
1_C	A4TA17	103	22
5_A	A2GA23	104	26
8_C	A7T2A57	107	66
7_A	A6TA4TA58	109	70
3_C	A2GA68GA11G2A14GA119	110	219
C_K	A2GA5GA8CA7TA8TA29	111	64
10_F	A7TATA2T2A45	112	59
6_I	A29	115	29
9_A	A3GA5GA2TA59	118	72
4_C	TA7GA4TA5GA4CA72	119	96
5_D	A8TA7GA2TA61GA65	120	147
5_K	A2TA4TA2GAnTA24GA40nGA52	123	170
C_O	A43TA18GA23	127	86
C_T	A4TA4T3A55	150	67
C_G	A23	162	23

Table (*b*) 3' ETS tails identified in the control cells.

Appendix **B**

Oligo/poly-adenylated β-actin tails were amplified from $(dT)_{16}$-adaptor primed cDNA with primers actinFF and $(dT)_{16}$-adaptor followed by a second nested PCR with actinF and $(dT)_{16}$-adaptor (Table 2.4) and cloned into pCR2.1 (Invitrogen). Randomly selected clones were sequenced.

high template concentration				
clone	tail	tail description	length	tail length-$(dT)_{16}$
A41	AAAAAAAAAAAAAAAAAAAA	A21	21	5
A14	AAAAAAAAAAAAAAAAAAAAAA	A23	23	7
A38	AAAAAAAAAAAAAAAAAAAAAA	A23	23	7
A44	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
A47	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
A8	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
A12	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
A23	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
A42	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
A35	AAAAAAAAAAAAAAAAAAAAAAAAAA	A26	26	10
A49	AAAAAAAAAAAAAAAAAAAAAAAAAA	A26	26	10
A9	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	A28	28	12
A30	AAAAAAGAAAAAGAAAAAAAAAAAAAAAAA	A6GA5GA17	30	14
A34	AAAAAAAAAAAAAAAAAAAAAAAA	A34	34	18
A19	AAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A5TGA33	40	24
A16	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A42	42	26
A17	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A42	42	26
A24	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A42	42	26
A10	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A43	43	27
A40	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A43	43	27
A4	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A45	45	29
A13	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A47	47	31

Continue...

A37	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A47	47	31
A20	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A48	48	32
A22	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A52	52	36
A3	AAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAA	A23TA30	54	38
A36	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC	A54	54	38
A18	AAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAA	A25TA29	55	39
A32	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A62	62	46
A26	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A63	63	47
A48	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A64	64	48
A39	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAA	A37TA27	65	49
A46	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A65	65	49
A29	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A101	101	85
				889

Table (*a*) β-actin mRNA tails identified with a high concentration of cDNA as the starting material for the PCR with Taq polymerase.

low template concentration				
clone	tail	tail description	length	tail length-$(dT)_{16}$
C33	AAAAAAAAAAAAAAAAAAAAA	A22	22	6
C55	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
C48	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A57	57	41
C27	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A68	68	52
C23	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCAAAAAAAAAAAAAAAAAAAA	A46C2A21	69	53
C29	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAT CAAAAAAAAAAAAAAGTCAAGCCAAAAAAAAAAAAA	A67TCA14GTCA 2GCCA13	104	88
C40	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAANAAAAAAAAAACCAAGCTAAAAAAAAAAAAAAA	A75NA10C2A2G CTA16	109	93
				341

Table (*b*) β-actin mRNA tails identified with a low concentration of cDNA as the starting material for the PCR with Taq polymerase.

high template concentration				
clone	tail	tail description	length	tail length-$(dT)_{16}$
H1_25	AAAAAAAAAAAAAAAAAA	A19	19	3
H1_5	AAAAAAAAAAAAAAAAAAAAA	A22	22	6
H1_16	AAAAAAAAAAAAAAAAAAAAA	A22	22	6
H1_17	AAAAAAAAAAAAAAAAAAAAA	A22	22	6
H1_19	AAAAAAAAAAAAAAAAAAAAAAA	A23	23	7
H1_2	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
H1_8	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
H1_13	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
H1_3	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
H1_7	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	A28	28	12
H1_21	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	A28	28	12
H1_4	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A29	29	13
H1_9	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A30	30	14
H1_6	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A31	31	15
H1_11	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A31	31	15
H1_15	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A33	33	17
H1_18	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A34	34	18
H1_12	AAAAAAAAAAAAAAAAAAAA

low template concentration				
clone	tail	tail description	length	tail length-$(dT)_{16}$
H5_16	AAAAAAAAAAAAAAAA	A17	17	1
H5_8	AAAAAAAAAAAAAAAAAAAA	A21	21	5
H5_41	AAAAAAAAAAAAAAAAAAAAA	A22	22	6
H5_5	AAAAAAAAAAAAAAAAAAAAAAA	A23	23	7
H5_14	AAAAAAAAAAAAAAAAAAAAAAA	A23	23	7
H5_26	AAAAAAAAAAAAAAAAAAAAAAA	A23	23	7
H5_25	AAAAAAAAAAAAAAAAAAAAAAA	A23	23	7
H5_24	AANAAAAAAAAAAAAAAAAAAA	A23	23	7
H5_13	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
H5_17	AAAAAAAAAAAAAAAAAAAAAAAA	A24	24	8
H5_18	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
H5_21	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
H5_22	AAAAAAAAAAAAAAAAAAAAAAAAA	A25	25	9
H5_29	AAAAAAAAAAAAAAAAAAAAAAAAAA	A26	26	10
H5_36	AAAAAAAAAAAAAAAAAAAAAAAAAA	A26	26	10
H5_23	AAAAAAAAAAAAAAAAAAAAAAAAAA	A26	26	10
H5_6	AAAAAAAAAAAAAAAAAAAAAAAAAAA	A27	27	11
H5_2	ANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	ANA34	36	20
H5_1	AAAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A5TGA32	39	23
H5_30	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A41	41	25
H5_20	AAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAACAAGAAAAAAA	A28TA3CA2GA7	43	27
H5_40	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A48	48	32
H5_31	AAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A23TA29	53	37
				295

Table (*d*) β-actin mRNA tails identified with a low concentration of cDNA as the starting material for the PCR with Expand High Fidelity polymerase.

Appendix **C**

Generation of POLS polyclonal antibody

In order to visualize the endogenous POLS protein and hence avoid potential artefacts introduced by the HA or TAP tagging, a polyclonal antibody was sought. To generate antigens for raising polyclonal antisera a recombinant fragment of POLS was expressed in *E.coli*. As *POLS* and *PAPD5* share a very similar central region (Figure 1.6), the GST-fusion proteins constructed avoided this region to minimize potential cross reactivity. A fragment from the amino-terminus were used to generate polyclonal antibody. A distinct band in the region of 55 kDa was identified with the crude serum for POLS in western blots of HEK-293T and U2-OS cell lysates. The specificity of the POLS antibody was confirmed by siRNA knock-down (Figure 3.7). A 73% or 94% down-regulation of POLS was seen 48 hr post-transfection in HEK-293T or U2-OS cells respectively as quantified by densitometric scanning after western blotting. The mRNA knockdown 48 hr post-transfection of the same cells was quantified to be 70% and 83% in HEK-293T and U2-OS cells respectively, as monitored by q-PCR. For the q-PCR triplicate samples were analyzed while the down-regulation of protein was quantified from one western blot for each cell type.

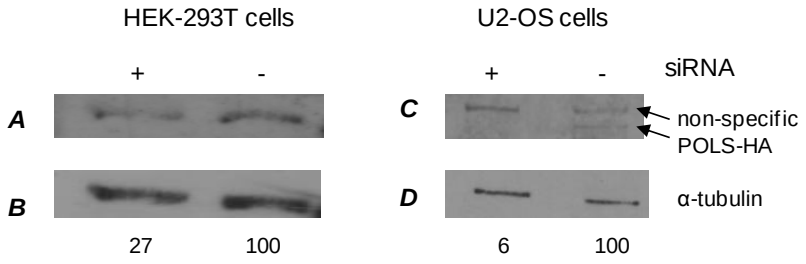


Figure *(a)* Demonstration of specificity of anti-POLS antibody by western blotting.

A,C Polyclonal antibody raised to detect POLS was used in western blot analysis of protein extracts from HEK-293T cells and U2-OS cells, respectively. ***B,D*** α-tubulin served as a loading control. Amount of POLS present is indicated below the blot. For the POLS knock-down in U2-OS cells, twice the quantity of protein lysate was analyzed compared to the control. In this instance a very efficient down-regulation of POLS was achieved with the non-specific band visible serving as a loading control.

Appendix **D**

Interacting protein partners of POLS

My TAP-tag/MS data identified the proteins tabulated in Table (*a*) as interacting partners of POLS.

Gene Names	Protein IDs	Unique Peptides control	Unique Peptides POLS	Mol. Weight [kDa]	Intensity control	Intensity POLS	Signal specificity	Protein abundance
PRKDC	IPI00296337	35	35	469.08	2.23E+07	2.29E+07	4.339189	4.688038
IPO5	IPI00939304	10	19	125.54	2452400	1.25E+08	6.609196	5.998787
HSC70	IPI00003865	12	13	70.897	4.32E+07	1.89E+08	5.790977	6.426913
HSP90AB1	IPI00414676	4	13	83.263	1.13E+07	3.53E+07	5.572193	5.627151
HSP60	IPI00784154	10	13	61.054	5.16E+07	1.33E+08	5.625046	6.338007
DKFZp686C11235	IPI00930124	0	13	52.12	9.94E+08	6.93E+08	5.125991	7.123535
HSPA9	IPI00007765	9	12	73.68	8566600	6.11E+07	5.985847	5.918656
TFRC	IPI00022462	12	12	84.87	2.86E+07	2.60E+07	5.029053	5.485935
HNRNPK6	IPI00216746	9	12	51.028	3.23E+07	1.43E+08	5.937869	6.44695
DHC1	IPI00456969	9	12	532.4	1985100	2071600	4.292285	3.590068
IGHG1	IPI00785084	0	12	51.082	2.09E+07	6316000	4.77261	5.092174
TUBA1A	IPI00180675	1	11	50.135	2.64E+07	3.35E+07	5.402865	5.824255
HSP90A	IPI00382470	6	11	98.16	5502300	2.00E+07	5.569072	5.309616
DKFZp686N02209	IPI00384938	1	11	52.851	258920	0	_	_
HSPA4	IPI00002966	8	10	94.33	1.06E+07	2.51E+07	5.399266	5.425594
PRO1400	IPI00022463	11	10	77.049	3.39E+07	4.63E+07	5.249075	5.778711
DKFZp686I15212	IPI00969547	3	10	56.813	3.79E+07	3.02E+07	5.14678	5.725041
ALB protein	IPI00216773	1	10	45.159	0	0	_	_
HSPA1	IPI00304925	9	10	70.051	6.34E+07	2.62E+08	5.770245	6.572074
DKFZp686I04196	IPI00399007	0	10	46.06	4.70E+08	2.82E+08	5.115151	6.78711

KPNB1	IPI00001639	9	9	97.169	1.26E+07	3.62E+07	5.470453	5.570893
CCT7	IPI00018465	6	9	59.366	3630300	1.65E+07	5.884898	5.44484
CCT1	IPI00290566	8	9	60.343	8902400	1.23E+07	5.360899	5.310407
CCT2	IPI00297779	6	9	57.488	8673800	1.93E+07	5.587996	5.526205
XRCC6	IPI00644712	5	9	69.842	2255300	2.12E+07	6.129384	5.482588
HSPA5	IPI00003362	7	8	72.421	2657500	1.72E+07	5.952199	5.376673
IQGAP1	IPI00009342	5	8	189.25	288800	1708800	5.495058	3.955655
CCT6	IPI00027626	6	8	58.024	3188300	1.48E+07	5.903212	5.406771
HSP105	IPI00939163	5	8	96.864	2778500	6185500	5.361402	4.805212
NONO	IPI00304596	4	8	54.231	3646800	2.36E+07	6.077322	5.639235
EEF1A	IPI00396485	8	8	50.14	3.95E+07	4.06E+07	5.311756	5.908716
TRIM28	IPI00438229	7	8	88.549	3963100	2.29E+07	5.814389	5.412424
MCCB	IPI00784044	5	8	61.332	1011200	1.24E+07	6.302052	5.306889
CHD4	IPI00455210	7	7	220.83	778010	2010700	5.068304	3.959289
TUBB2C	IPI00007752	1	7	49.83	1.87E+07	1.13E+07	5.082843	5.355126
PTBP1	IPI00183626	6	7	59.632	1.08E+07	4.07E+07	5.799139	5.834328
NCL	IPI00604620	7	7	76.613	5408700	1.08E+07	5.41627	5.149363
FUBP2	IPI00479786	4	7	73.114	823270	7535600	6.097575	5.013117
IGKC	IPI00550731	0	7	26.234	590570	1376900	5.948767	4.720038
TUBB5	IPI00909140	2	7	52.048	2.29E+07	1.99E+07	5.223168	5.582776
TPR	IPI00742682	4	7	267.29	939670	3326600	5.122043	4.095018
LRP130	IPI00783271	3	7	157.9	81126	2338800	6.261451	4.170611
KIAA0002	IPI00784090	5	7	59.62	3671900	2.30E+07	6.021181	5.586071
IGK@	IPI00784865	1	7	25.773	124470	294750	5.963224	4.058289
IGK@	IPI00889156	0	7	25.834	6.01E+08	2.62E+08	5.227589	7.006275
CCT5	IPI00010720	4	6	59.67	392270	1.88E+07	6.904424	5.498009
HCFC1	IPI00641743	5	6	213.47	2471400	8094900	5.185932	4.578875
BAP135	IPI00054042	6	6	112.42	1702000	3237500	5.228406	4.459366
IGLC217	IPI00154742	1	6	24.792	1.79E+08	1.19E+08	5.431005	6.682838
EEF2	IPI00186290	5	6	95.337	6055900	5768800	4.999645	4.781824

MTHFC	IPI00794900	7	6	110.5	4386300	3562100	4.866245	4.508344
CCT4	IPI00302927	6	6	57.924	4245500	8122600	5.518908	5.146837
DKFZp686K18196	IPI00386879	2	6	53.088	6048700	1.14E+07	5.550817	5.332479
NOTCH1	IPI00412982	6	6	272.44	3070600	3212900	4.584403	4.071626
PAPD7;POLS;TRF4	IPI00418630	1	6	59.873	23508	7121000	7.704094	5.07531
PPP2R1A	IPI00554737	4	6	65.308	864540	2859800	5.704584	4.641369
VL4	IPI00829640	1	6	24.854	1.71E+07	8849900	5.318013	5.551542
DDX9	IPI00844578	3	6	140.96	261930	3373500	5.9608	4.378985
IGHG4	IPI00930442	2	6	52.42	2.26E+07	3.93E+07	5.520556	5.874376
PGDH3	IPI00011200	3	5	56.65	1333900	5168700	5.835058	4.960181
ACAC	IPI00396015	3	5	270	785310	7881300	5.570193	4.465234
LMN1	IPI00021405	3	5	74.139	320270	3321000	6.145706	4.651222
ACTB	IPI00021440	7	5	41.792	3.42E+07	7.89E+07	5.741518	6.276165
HBP	IPI00022228	2	5	141.45	40711	774270	6.128578	3.73829
VCP	IPI00022774	2	5	89.321	519980	1976700	5.629001	4.344987
CKAP5	IPI00921422	4	5	226.25	254030	1168800	5.308267	3.713152
NICE4	IPI00514856	5	5	114.53	954010	4896500	5.651413	4.630967
HNRNPM	IPI00171903	4	5	77.515	557530	2.25E+07	6.71649	5.462758
LMNB1	IPI00217975	0	5	66.408	0	2754400	_	4.617806
GAPDH	IPI00788737	5	5	38.791	7124600	2.20E+07	5.900813	5.753573
ILF3	IPI00418313	4	5	95.807	629590	1774700	5.46867	4.267728
XRCC5	IPI00220834	3	5	82.704	5213800	7971300	5.266848	4.984003
SMC1	IPI00291939	2	5	143.23	144030	1613800	5.893363	4.051816
PC	IPI00299402	4	5	129.63	1.77E+07	2.34E+08	6.008291	6.255749
SF3B3	IPI00300371	5	5	135.58	1204200	5397300	5.519282	4.599981
NUP93	IPI00397904	3	5	93.487	137290	1434800	6.048401	4.18604
VIM	IPI00418471	3	5	53.651	3403000	7201600	5.595989	5.127851
TD60	IPI00465044	1	5	56.084	23434	935590	6.8524	4.222247
OIP3	IPI00479186	1	5	57.936	1762200	1.18E+07	6.061034	5.307089
RP11-11N7.3-002	IPI00883857	3	5	90.583	4338200	9550900	5.385688	5.022997

IGHM	IPI00896380	5	5	51.79	3.42E+07	4.71E+07	5.425414	5.958932
HSPC075	IPI00782992	4	5	299.61	201970	282280	4.668837	2.974124
SUPT6H	IPI00784161	4	5	199.07	190060	976570	5.411807	3.690698
Uncharacterized	IPI00807428	0	5	24.746	0	1129400	_	4.659343
SPTA2	IPI00879810	1	5	284.94	51957	1953400	6.120394	3.836038
SRP1	IPI00002214	1	4	57.861	61600	6428700	7.256156	5.045737
OGT	IPI00005780	4	4	116.92	1074800	1939500	5.188473	4.219801
RPLP0	IPI00008530	4	4	34.273	2.41E+07	2.77E+07	5.525557	5.908185
LMNB2	IPI00009771	1	4	69.948	262200	2435800	6.123234	4.541866
MCM3	IPI00013214	3	4	95.907	1389500	2090400	5.19552	4.338379
EPRS	IPI00013452	1	4	170.59	98804	2.19E+07	7.114469	5.109243
SF3A1	IPI00017451	4	4	88.885	2176200	1.70E+07	5.942899	5.280597
DLAT	IPI00021338	3	4	68.996	1738600	8370000	5.843702	5.083901
EEF1D	IPI00642971	4	4	71.407	1564400	5443500	5.687789	4.882137
VDAC2	IPI00902560	4	4	33.371	3198200	6855700	5.807777	5.312683
MCCA	IPI00024580	3	4	80.472	639650	1.25E+07	6.385497	5.191439
RPN1	IPI00025874	3	4	72.777	125200	1403800	6.187707	4.285311
FAS	IPI00026781	2	4	273.42	414590	247620	4.339336	2.956955
NUP50	IPI00026940	0	4	50.144	0	2446000	_	4.688238
TP53BP1	IPI00742743	2	4	214.11	169720	440540	5.083615	3.313348
DDX6	IPI00030320	2	4	54.416	268880	1.81E+07	7.092873	5.522431
U2AF2	IPI00031556	2	4	53.5	636450	7453900	6.340265	5.14403
CAPG	IPI00106495	4	4	114.33	964750	1553600	5.148764	4.133179
HSPC272	IPI00179298	1	4	482.14	63782	1002600	5.513257	3.317955
HIST1H1C	IPI00217465	1	4	21.364	160310	1310500	6.582794	4.787755
SMC3	IPI00219420	0	4	141.54	0	779230	_	3.740787
SMARCC1	IPI00234252	2	4	122.87	866820	6066900	5.755592	4.693521
NUP210	IPI00291755	2	4	205.11	103430	846980	5.60124	3.615886
RANGAP1	IPI00294879	1	4	63.541	35548	1674800	6.870094	4.420909
TPR2	IPI00329629	2	4	56.44	350330	3600600	6.260311	4.804788

FBP11	IPI00337386	4	4	112.4	2378400	3.86E+07	6.15912	5.535405
MSH6	IPI00384456	2	4	152.78	972190	1828300	5.09023	4.077981
PCMT1	IPI00411680	3	4	24.65	1438700	1588100	5.651091	4.809061
ASCC3L1	IPI00420014	5	4	244.5	286010	907820	5.11334	3.569721
ATP5A	IPI00440493	4	4	59.75	2291500	1516800	5.044471	4.404591
PARP1	IPI00449049	3	4	113.08	2745000	4531000	5.164266	4.602808
UBA1	IPI00645078	4	4	117.85	1467600	1124200	4.812906	3.979514
PERQ2	IPI00647635	3	4	152.53	641220	2242000	5.360273	4.16728
VH	IPI00782983	1	4	18.58	1.88E+07	6632300	5.27793	5.552619
Uncharacterized	IPI00940245	0	4	44.786	356700	365410	5.359335	3.911638
AIFM1	IPI00000690	3	3	66.9	139560	1040800	6.04718	4.191941
U5-116KD	IPI00003519	1	3	109.43	118680	622650	5.68073	3.755108
PRSS15	IPI00005158	3	3	106.49	300530	897870	5.448017	3.925905
SLC25A5	IPI00007188	2	3	32.895	4135000	1.21E+07	5.949897	5.566373
PRPF8	IPI00007928	3	3	273.6	101060	590090	5.329223	3.333802
DDX48	IPI00009328	3	3	46.871	611900	612540	5.32955	4.11623
TRIM33	IPI00010252	2	3	122.52	36981	521610	6.06116	3.629139
PDIA1	IPI00010796	1	3	57.116	103090	1599000	6.433874	4.447091
RPS2	IPI00013485	3	3	31.324	1.91E+07	8479600	5.150806	5.432498
CSTF3	IPI00015195	2	3	82.921	791380	3620700	5.741743	4.640128
GLUD	IPI00016801	0	3	61.397	0	444490	_	3.859715
PHB	IPI00017334	3	3	29.804	1899400	2902300	5.709851	4.988468
DDX5	IPI00017617	0	3	69.147	199370	2162100	6.195443	4.495102
RPA1	IPI00020127	3	3	68.137	743600	2084100	5.614196	4.485536
ACLY	IPI00394838	0	3	125.14	0	1352100	_	4.033613
PCNA	IPI00021700	3	3	28.768	1106600	3625300	6.056444	5.100434
GPATC5	IPI00024255	2	3	52.228	392700	1184300	5.761497	4.355558
ERP57	IPI00025252	1	3	56.782	53075	704130	6.368552	4.093442
NUP155	IPI00026625	1	3	155.2	43874	862800	6.102811	3.745018
HNRNPL	IPI00027834	1	3	64.132	156790	1687600	6.224876	4.420195

URB2	IPI00028980	1	3	170.54	25641	624190	6.154556	3.563491
EIF3A	IPI00029012	0	3	166.57	0	963720	_	3.762354
SF3A3	IPI00029764	2	3	58.848	895290	1.24E+07	6.370077	5.322041
LAP2	IPI00030131	1	3	50.67	759190	6848200	6.250475	5.130826
PSME3	IPI00219445	3	3	30.886	8955100	3371900	5.086042	5.038113
HADH	IPI00031522	3	3	82.999	366840	1055100	5.539744	4.104221
LSF	IPI00037599	3	3	57.255	638870	2282900	5.795261	4.600674
FUBP3	IPI00377261	1	3	61.64	189430	2496400	6.330003	4.607451
KIAA1352	IPI00103994	0	3	134.46	0	1944800	_	4.160282
LUC7L3	IPI00967258	2	3	58.22	1387000	8961600	6.045237	5.187314
HNRNPA1	IPI00215965	2	3	38.746	677230	2290000	5.940872	4.771609
DDX3	IPI00215637	0	3	73.243	0	264160	_	3.557101
LAP2	IPI00216230	0	3	75.491	0	139590	_	3.266959
HIST1H1E	IPI00217467	1	3	21.865	104430	2378400	7.01771	5.036535
RBM12B	IPI00871780	3	3	118.1	271190	745380	5.366854	3.800128
MRE11	IPI00940798	1	3	80.592	178250	1622100	6.052756	4.303786
U1AP1	IPI00290204	2	3	51.556	2430200	8002900	5.805326	5.190968
TAF2S	IPI00247871	1	3	123.9	36362	757890	6.225887	3.786535
KPNA3	IPI00299033	2	3	57.81	262280	2943200	6.288052	4.706817
CDC73	IPI00300659	1	3	60.576	136960	1781200	6.331819	4.468412
FLN	IPI00333541	2	3	280.74	214620	350840	4.765135	3.096805
HGRG8	IPI00306043	2	3	62.333	70093	398630	5.960177	3.805852
NUP35	IPI00329650	3	3	34.773	1786000	1437500	5.364484	4.616366
FUBP1	IPI00641948	1	3	69.993	103160	1962700	6.434288	4.447799
LSM12	IPI00867714	3	3	21.999	3.17E+07	2.35E+08	6.527176	7.028591
RPS3A	IPI00419880	2	3	29.945	1049400	1.03E+07	6.515361	5.536302
MPB1	IPI00465248	2	3	47.168	1093400	2.14E+07	6.617703	5.656482
API5	IPI00855856	1	3	59.004	92823	997960	6.260576	4.228232
POU2F1	IPI00640177	3	3	72.082	538890	583740	5.176892	3.908393
SKIV2L2	IPI00647217	3	3	117.8	647670	1138800	5.173948	3.985302

DDX17	IPI00940000	0	3	80.457	0	94929	_	3.071835
RP11-151A6.1-004	IPI00744115	0	3	80.058	0	4250800	_	4.725066
scFv	IPI00830088	0	3	12.839	0	88596	_	3.838883
IGHV	IPI00916434	2	3	25.145	2.85E+07	1.41E+07	5.293331	5.747688
QARS	IPI00026665	2	3	89.656	2351900	3839400	5.260265	4.631684
STAU1	IPI00000001	0	2	63.182	0	176500	_	3.446151
RPL5	IPI00000494	2	2	34.362	1145400	4374600	6.045903	5.10486
GCN1L1	IPI00001159	0	2	292.75	0	114710	_	2.593104
RBM25	IPI00004273	1	2	100.18	290060	1707200	5.769015	4.231503
C1orf107	IPI00004290	2	2	87.054	170720	440430	5.471804	3.704088
PYGB	IPI00783313	2	2	97.147	254790	250630	5.005421	3.411604
KDP	IPI00004472	2	2	250.79	113730	343340	5.080539	3.136414
PRPF19	IPI00004968	1	2	55.18	30585	770740	6.659618	4.145126
UBP1	IPI00005018	2	2	60.49	879710	2148100	5.606031	4.550371
FACT80	IPI00005154	1	2	81.074	130370	660660	5.795919	3.911096
SPTBN1	IPI00005614	0	2	274.61	0	215110	_	2.893944
RAB1A	IPI00005719	2	2	22.677	80740	3835400	7.321136	5.228225
NUP98	IPI00006038	2	2	197.58	309470	703180	5.060705	3.551323
EIF3D	IPI00791086	2	2	63.902	381280	820070	5.527092	4.108336
PCCB	IPI00946876	2	2	60.736	3545300	2.01E+07	5.969059	5.518712
SMC2	IPI00007927	2	2	135.65	42897	682450	6.069224	3.701651
MARS	IPI00008240	2	2	101.11	318610	337070	5.019667	3.522926
CRDBP	IPI00008557	2	2	63.48	442780	873710	5.492542	4.138731
MAP1B	IPI00008868	2	2	270.62	393340	165970	4.192902	2.78767
CFIM68;CPSF6	IPI00030654	1	2	63.47	22892	479700	6.518717	3.878401
ZNF207	IPI00792837	1	2	52.693	27199	961410	6.826603	4.261156
HDAC1	IPI00013774	2	2	55.102	153840	4256400	6.700806	4.887875
KARS	IPI00307092	1	2	71.496	17673	1515900	7.079079	4.326389
ZNF631	IPI00878557	0	2	62.157	0	78486	_	3.101302
DDX21	IPI00015953	2	2	87.343	213410	187800	5.003253	3.332468

SAR1	IPI00015954	1	2	22.367	4.71E+07	4.43E+08	6.624476	7.297266
PLC1	IPI00383849	2	2	148.66	970740	1.19E+07	5.915848	4.902951
MATR3	IPI00789551	1	2	99.966	331110	1191700	5.556342	4.076314
HNRPQ	IPI00018140	2	2	69.602	1400400	2325200	5.377586	4.523839
PSMC3;TBP1	IPI00018398	2	2	49.203	179940	1102700	6.095338	4.350466
TROVE2	IPI00019450	1	2	60.67	232540	292440	5.316565	3.683063
MAP3K7IP1	IPI00019459	0	2	54.643	0	304370	_	3.745867
DEK	IPI00020021	2	2	42.674	478660	1534500	5.875776	4.555804
RPS6	IPI00021840	2	2	28.68	1386100	559380	5.148333	4.290128
CLTCL2	IPI00024067	2	2	191.61	254380	978680	5.30274	3.708223
BTAF1	IPI00024802	1	2	206.89	23632	82098	5.225092	2.598593
STUB1	IPI00025156	2	2	34.856	5844700	2159100	5.025233	4.791995
SF3B1	IPI00026089	1	2	145.83	14091	214520	6.018679	3.167621
NPEPPS	IPI00026216	2	2	103.28	289210	366590	5.088951	3.550164
GNB1	IPI00026268	2	2	37.377	928790	1293900	5.571379	4.539296
GADS	IPI00026928	2	2	37.909	2.03E+08	3.97E+08	5.711634	7.019643
SUPT16H	IPI00026970	0	2	119.91	0	662830	_	3.742547
HSP90B1	IPI00027230	1	2	92.468	1264100	1607900	5.138486	4.240268
BAP	IPI00027252	2	2	33.296	3429300	2329600	5.309684	4.844889
PRIM2	IPI00027705	1	2	58.805	78211	611180	6.123487	4.016755
RPN2	IPI00028635	1	2	69.283	48788	329260	5.988599	3.676912
PSMA6	IPI00029623	2	2	27.399	1972300	2.03E+07	6.574339	5.869312
TRAP1	IPI00030275	1	2	80.109	84736	181440	5.426984	3.355052
ZNF828	IPI00064212	1	2	89.098	514450	1319800	5.459297	4.17064
TIP120	IPI00100160	0	2	136.37	0	392260	_	3.458855
PUF60	IPI00100716	1	2	58.171	1262300	1628800	5.345999	4.447161
SRRF	IPI00103247	1	2	60.082	73676	259110	5.767414	3.63474
AMOT	IPI00163085	1	2	118.08	23828	149350	5.724941	3.102029
DAZAP1	IPI00165230	1	2	43.383	191260	623560	5.875935	4.157558
NUP75;NUP85	IPI00790530	0	2	75.019	0	593850	_	3.898505

MTA2	IPI00171798	1	2	75.022	71477	1270200	6.374517	4.228683
MNAR;PELP1	IPI00787089	0	2	135.49	0	707310	_	3.717703
KIF4A	IPI00178150	1	2	139.88	12563	366070	6.318715	3.417809
DIS3	IPI00746351	1	2	109	92919	85806	4.927987	2.896091
PSMD12	IPI00185374	0	2	52.904	0	2422300	_	4.660739
RPL18	IPI00215719	2	2	21.634	1074000	2710300	6.066876	5.09788
PPP1CB	IPI00218236	2	2	37.186	174360	316540	5.688602	3.930049
EIF2A	IPI00219678	2	2	36.112	144490	279700	5.729203	3.889041
ARS2	IPI00375144	2	2	100.67	912660	814680	4.947778	3.908087
COPB2	IPI00220219	0	2	102.49	0	343300	_	3.524992
NPM1	IPI00549248	2	2	32.575	9764000	3889300	5.087359	5.076987
SF3B2	IPI00221106	1	2	100.23	39895	695520	6.240394	3.841312
AP2A1	IPI00304577	2	2	107.54	215150	351970	5.182194	3.514936
TAP	IPI00900293	0	2	281.63	0	181360	_	2.808863
HDAC2	IPI00289601	1	2	65.537	259450	1162000	5.834666	4.24872
CTPS	IPI00290142	1	2	66.69	862090	608920	5.024947	3.9605
CCT3	IPI00553185	2	2	60.533	741300	5286200	6.071157	4.941152
NUP133	IPI00291200	1	2	128.98	97877	405910	5.507227	3.497907
ANT3	IPI00291467	1	2	32.866	2844100	1167800	5.096677	4.550622
GEMIN5	IPI00291783	0	2	168.59	0	219480	_	3.114563
NUP62	IPI00293533	1	2	53.254	401180	454610	5.327948	3.931287
KRT6B	IPI00293665	1	2	60.066	465640	0	_	_
ZNF162	IPI00386114	1	2	68.329	45551	780430	6.399231	4.057729
LCP2	IPI00297169	1	2	60.187	1064400	2657500	5.617866	4.64497
NPA1	IPI00297241	1	2	254.39	21008	140610	5.420132	2.742516
EIF5B	IPI00299254	0	2	138.83	0	398220	_	3.45764
CAPD2	IPI00299524	0	2	157.18	0	389340	_	3.393932
VAV3	IPI00910879	0	2	100.81	0	5512000	_	4.737805
CAD	IPI00301263	2	2	242.98	221360	117180	4.338184	2.683283
MIOS	IPI00301987	2	2	98.583	130170	219350	5.232825	3.347336

ATP5B	IPI00303476	2	2	56.559	1339200	742560	4.991385	4.11823
ZNF143	IPI00305316	2	2	68.895	799350	684830	5.094658	3.997395
ADTB2	IPI00784366	1	2	105.69	68801	308690	5.627894	3.465489
SMC4	IPI00411559	2	2	147.18	78210	707740	5.788763	3.682025
IGM	IPI00383953	1	2	16.315	334690	31873	4.766193	3.290836
KTN1	IPI00328753	0	2	156.27	0	960520	_	3.788631
RPL6	IPI00790342	2	2	32.872	623940	2240800	6.038434	4.833577
PPP2R2A	IPI00798041	1	2	52.999	474840	4133600	6.215514	4.892061
;HNRNPAB	IPI00334587	2	2	35.967	4357300	2358300	5.177477	4.816695
NUP358	IPI00921207	1	2	198.39	40626	106630	5.121556	2.73036
Uncharacterized	IPI00382436	0	2	11.392	0	208470	_	4.262444
HSPC099	IPI00384541	1	2	156.02	34139	634230	6.075816	3.609066
Uncharacterized	IPI00385252	1	2	11.83	3031300	1238600	5.538317	5.019946
SRPR	IPI00385267	1	2	69.81	118480	677820	5.913552	3.987197
EIF4G1	IPI00956559	1	2	176.23	46285	176600	5.335471	3.000911
MAP4	IPI00880007	2	2	245.44	88651	446330	5.312027	3.259711
CHD8	IPI00398992	0	2	290.51	0	56312	_	2.28744
DDX42	IPI00409671	2	2	102.97	667990	958780	5.144238	3.969008
GTF3C5	IPI00879210	0	2	60.368	0	721640	_	4.077514
NOC2L	IPI00411886	1	2	84.904	12766	142970	6.120262	3.226317
IPOA7	IPI00747764	2	2	60.579	524260	2004900	5.800224	4.519771
CRSP3	IPI00479534	0	2	157.11	0	190890	_	3.084579
CDC5L	IPI00465294	2	2	92.25	195010	163190	4.95767	3.247727
RPL15	IPI00470528	1	2	24.146	108400	355310	6.132733	4.167762
STAG2	IPI00470883	0	2	145.75	0	248980	_	3.232556
TMEM214	IPI00477118	1	2	77.15	63136	720690	_	3.970412
CPSF7	IPI00915356	1	2	56.374	176480	999540	6.002026	4.248721
BAT1;UAP56;DDX39	IPI00644431	2	2	53.696	779570	1703300	5.609494	4.501349
IARS	IPI00644127	1	2	144.5	323350	7954100	6.23105	4.740723
IGVH	IPI00735451	2	2	12.674	1006600	172570	5.131195	4.134052

Uncharacterized	IPI00783024	1	2	14.142	1198400	1377200	5.909884	4.988486
NUP205	IPI00783781	2	2	227.92	130120	662840	5.349282	3.463626
V<kappa>3	IPI00784430	2	2	13.798	5223100	1996000	5.442416	5.160345
EIF2A	IPI00012462	1	2	67.851	117120	222360	5.446869	3.5155

Table (*a*). Interacting protein partners of POLS identified in this study.
Quantification was performed using MAxQuant software (version 3.8, 2013). Proteins listed contained two or more unique peptides in the POLS fraction. Relative peptide intensity was divided by the molecular weight (MW) to calculate the protein abundance. Signal specificity is the ratio between the intensity of peptides originating from POLS *vs* control.