Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Lic. Marcelo Viegas
born in Buenos Aires, Argentina

Oral-examination:: .....

Physiological role and quantitative aspects of human Nonsense-mediated mRNA decay (NMD)

Referees: Prof. Dr. Iain Mattaj
Prof. Dr. Claus Bartram

"Reality is the only truth"

Aristoteles.

To Kumiko who taught me the deep meaning of the word “love”.
To my parents for whom fulfilling my wishes is the source of their happiness.
And for Uma, my first niece who is about to come.

Abstract

Nonsense-mediated mRNA decay (NMD) is a molecular pathway of mRNA surveillance that ensures the rapid degradation of mRNAs containing premature translation termination codons in all studied eukaryotes. Originally, NMD was thought of as a quality control pathway that targets non-functional mRNAs arising from mutations and splicing errors. More recently, NMD has been shown to also regulate normal gene expression and NMD thus emerged as one of the key post-transcriptional mechanisms of gene regulation. Despite the progress in the understanding of the role and mechanism of this pathway, the physiological impact of NMD on humans is not yet fully uncovered. To explore the functions of NMD in humans, I combined RNAi against the essential NMD factor UPF1 with genome-wide microarray analysis. My research indicate that NMD affects the expression of a large number of genes implicated in a wide diversity of functions although a majority seems to be affected indirectly and – consequently – do not represent legitimate NMD targets.

The validation of five *bona fide* NMD transcripts allowed me to develop an assay to quantitate differences in NMD efficiency. Using three different strains of HeLa cells as a simple model, I have systematically analysed the molecular mechanism underlying quantitative differences in NMD efficiency. The results of this analysis show that the quantitative differences in NMD efficiency represent a stable characteristic of the investigated strains. Low NMD efficiency is shown to be functionally related to the reduced abundance of the exon junction component RNPS1 in one of the analysed HeLa strains. Furthermore, restoration of functional RNPS1 expression, but not of NMD-inactive mutant proteins, also restores efficient NMD in the RNPS1 deficient cell line. I conclude that cellular concentrations of RNPS1 can modify NMD efficiency and propose that the cell type specific co-factor availability represents a novel principle that quantitatively controls NMD.

I also tested the hypothesis of NMD as a genetic modifier in the phenotypic expression of disease. A specific β-thalassemia – common in Mediterranean Asia – was assayed as a model. My results do not support a role of NMD for the variable severity of this specific mutation leading to anemia.

Zusammenfassung

Der Nonsens-vermittelte mRNA Abbau (Nonsense-mediated mRNA decay; NMD) ist ein Bestandteil der zellulären Qualitätskontrolle aller Eukaryonten durch den mRNAs mit einem vorzeitigen Translationsterminationskodon beschleunigt abgebaut werden. Ursprünglich dachte man, dass NMD nur nicht-funktionelle mRNAs zerstört, die z.B. durch Mutationen oder Spleißfehler entstehen. Kürzlich konnte gezeigt werden, dass NMD auch die normale Genexpression reguliert. Daher wird NMD heutzutage als einer der zentralen Mechanismen der post-transkriptionellen Regulation der Genexpression angesehen. Obwohl der molekulare Mechanismus des NMD zunehmend besser verstanden wird, ist noch wenig über die physiologische Bedeutung dieses Abbauwegs für den Menschen bekannt.

In der vorliegenden Arbeit habe ich die transkriptionellen Auswirkungen von RNAi gegen den essentiellen NMD-Faktor UPF1 mittels einer genomweiten Microarray-Analyse untersucht, um die Funktion von NMD im Menschen besser zu verstehen. Meine Untersuchungen zeigen, dass NMD die Expression einer großen Zahl von Genen beeinflusst, welche an einer Vielzahl unterschiedlicher zellulärer Prozesse beteiligt sind. Höchstwahrscheinlich wird die Mehrzahl der identifizierten mRNAs jedoch durch indirekte Effekte in ihrer Expression verändert und stellt somit keine legitimen (d.h. direkten) NMD-regulierten Transkripte dar. Die Identifizierung von fünf validierten direkten NMD-regulierten mRNAs ermöglichte mir, ein experimentelles System zur Quantifizierung von Unterschieden der NMD-Effizienz zu etablieren. Auf der Basis von drei verschiedenen HeLa-Zelllinien als Modellsystem konnte ich systematisch den molekularen Mechanismus unterschiedlicher NMD-Effizienzen untersuchen. Ich konnte zeigen, dass die unterschiedliche NMD-Effizienz eine unveränderliche Eigenschaft jeder dieser Zelllinien darstellt. Die relativ schlechte NMD-Effizienz einer der Zelllinien wird durch eine verringerte Expression des Proteins RNPS1 verursacht. Eine Erhöhung der RNPS1 Expression bewirkt eine deutliche Steigerung der NMD-Effizienz bis auf ein normales Niveau. NMD-inaktive Mutanten von RNPS1 zeigen jedoch nicht denselben Effekt. Die Konzentration von RNPS1 in der Zelle stellt also eine der Determinanten der NMD-Effizienz dar. Die Zelltyp-spezifische Verfügbarkeit von NMD-Faktoren ist somit ein neuartiger molekularer Mechanismus, der quantitativ die Wirklung von NMD durch seine Effizienz kontrolliert. Weiterhin habe ich untersucht, ob NMD generell die phenotypische Ausprägung von Erkrankungen als ein modifizierender genetischer Faktor verändern kann. Eine bestimmte Form der b-Thalassämie, die insbesondere im Nahen Osten verbreitet ist, wurde als experimentelles Modellsystem untersucht. Meine Ergebnisse zeigen jedoch, dass die unterschiedliche Schwere der

Erkrankung in den untersuchten Fällen wahrscheinlich nicht durch unterschiedliche NMD-Effizenz verursacht wird.

Table of contents

Abbreviations **1**

1. Introduction

1.1 Messenger RNA turnover in eukaryotes **3**

1.1.1 General pathways of mRNA turnover: the 5´and 3´decay pathways 3

1.1.2 mRNA decay via endonucleolytic cleavage 5

1.1.3 Regulated mRNA decay: AU-rich elements 8

1.1.4 mRNA surveillance pathways 10

1.1.5 Localisation of decay factors 11

1.2 Nonsense-mediated mRNA decay **13**

1.2.1 NMD in mammals 13

1.2.2 NMD in yeast, fly and worm 17

1.2.3 Physiological role of NMD 20

1.2.5 Role of NMD in disease 23

1.2.6 Variable NMD as a disease modifier –
The need of an assay system to quantify NMD efficiency 25

1.2.6.1 A case study for NMD variability: β-thalassemia (IVS1+6 T→C) 26

1.3 Aim of the project **28**

2. Materials and Methods

2.1 Materials **29**

2.1.1 Chemicals 29

2.1.2 Standard used buffers and media 29

2.1.3 Enzymes 31

2.1.4 Plasmids 32

2.1.5 Antibodies 33

2.1.6 Bacterial strains 33

2.1.7 Eukaryotic cell lines 33

2.1.8 Sequences of oligonucleotides used in PCR reactions 34

2.1.9 Sequences of siRNA used in knock-down experiments 36

2.1.10 Kits 36

2.1.11 Instrumental material 36

2.2 Methods 37

2.2.1 Standard methods 37

2.2.2 Bacterial techniques 37

2.2.3 DNA techniques 38

2.2.4 RNA techniques 41

2.2.5 Protein techniques 45

2.2.6 Cell culture techniques 47

2.2.7 Blood samples 49

2.2.8 Microarray techniques 50

3. Results

3.1 Physiological role of NMD in human cells 52

3.2 Characterisation of *bona fide* endogenous NMD targets – A key step towards a quantitation human NMD assay 58

3.3 Quantitative differences in human NMD: a HeLa cell model 64

3.3.1 Different HeLa strains display variations in NMD efficiency 64

3.3.2 RNPS1 is a potential modulator of NMD efficiency 66

3.4 Quantitative differences in NMD in human blood: A step towards the clinic 69

3.4.1 Estimation of NMD efficiency in blood samples 69

3.4.2 A case study: β-thalassemia caused by mutation in IVS1+6 of the β-globin gene 73

4. Discussion

4.1 Identification of *bona fide* NMD targets 78

4.2 Quantitative differences in cellular NMD efficiency 81

4.3 Variations in NMD efficiency – A potential new genetic modifier of disease 85

4.3.1 NMD in β-thalassemia 86

4.4 Concluding remarks 88

References 90

Acknowledgements 103

Appendix 104

List of Figures

Fig. 1	General mRNA decay pathways in eukaryotes	6
Fig. 2	Endonucleolytic mRNA decay pathways	8
Fig. 3	P-bodies	12
Fig. 4	PTC definition in mammals	14
Fig. 5	The Exon Junction Complex (EJC)	15
Fig. 6	A current model for NMD in mammalian cells	18
Fig. 7	Splicing variants of β-globin (IVS1+6)	27
Fig. 8	Transfer assembly for Northern-blots	43
Fig. 9	Transfer assembly for Western-blots	46
Fig. 10	Schematic representation of a biocoll gradient after centrifugation	50
Fig. 11	UPF1 depletion up-modulates the abundance of transfected and Endogenous NMD reporters	53
Fig. 12	List of significantly over-represented gene ontology categories for the up-modulated transcripts	57
Fig. 13	Differential expression analysis of 16 selected transcripts that were up-regulated by UPF1	59
Fig. 14	Analysis of potential off-target effects in 16 selected transcripts from the microarray analysis	60
Fig. 15	Pre-mRNA and mRNA analysis distinguishes potentially direct from indirect NMD targets in UPF1-depleted cells	61
Fig. 16	UPF1 depletion prolongs the half-lives of endogenous NMD targets	62
Fig. 17	UPF1 and UPF2 depletion cause similar degrees of up-modulation of NMD substrates	63
Fig. 18	Structure of the NMD sensitive isoforms for TBL2, GADD45B and NAT9	64
Fig. 19	Up-regulation of direct NMD substrates in UPF1-depleted cells measured according to distinct non-NMD transcripts	65
Fig. 20	The abundance of cellular NMD targets reflects the variability of NMD efficiency in HeLa cell strains	66
Fig. 21	RNPS1 is less abundant in HeLa strain B	67
Fig. 22	Over-expression of RNPS1 increases the degradation of a transfected NMD reporter in strain B cells	68

Fig. 23 The increased NMD efficiency of strain B cells is RNPS1-specific 69
Fig. 24 NMD transcripts abundance in blood of healthy donors 71
Fig. 25 NMD transcripts abundance in lymphoblastoid cell lines 72
Fig. 26 Detection of β-globin transcripts in patients´samples by RPA 75
Fig. 27 Quantification of the β-globin transcripts in patients´samples in the RPAs 76

List of Tables

Table 1 List of up-modulated transcripts in UPF1-depleted cells according to microarray data 54
Table 2 Distribution of targeted genes in the chromosomes 56

Abbreviations

%	per cent
approx.	approximately
APS	ammonium persulfate
ARE	Au-rich element
ATP	adenosine triphosphate
bp	base pairs
BrEt	ethidium bromide
cDNA	complementary DNA
C. elegans	*Caenorhabditis elegans*
CTP	cytosine triphosphate
D. melanogaster	*Drosophila melanogaster*
DNA	deoxyribonucleic acid
DSE	downstream sequence element
dsRNA	double-stranded RNA
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EJC	exon junction complex
Fig	figure
fL	femtoLiter
GTP	guanosine triphosphate
hs	hours
Ig	immunoglobulin
IRE	iron regulatory element
LB	Luria-Bertani
m^7GDP	7-methyl guanosine diphosphate
m^7GMP	7-methyl guanosine monophosphate
mA	milliamper
MCV	mean corpuscular volumen
mg	milligram
µg	microgram
min	minutes
mJ	milli-Joule

µl	microliter
mRNA	messenger RNA
nt(s)	nucleotide(s)
O.N.	overnight
ORF	open reading frame
P-body	processing body
PMSF	phenylmethylsulphonyl fluoride
Poly(A)	poly-adenylate
pre-mRNA	precursor mRNA
RNA	ribonucleic acid
RNase	ribonuclease
S. cerevisiae	*Saccharomyces cerevisiae*
RPA	RNase protection assay
SDS	sodium dodecylsulfate
siRNA	small interfering RNA
Taq	*Thermophilus aquaticus*
TEMED	N,N,N´,N´-tetramethylethylene diamine
tRNA	transfer RNA
uORF	upstream open reading frame
UTP	uracyl triphosphate
UTR	untranslated region
UV	ultra-violet
v/v	volume/volume
wt	wild type
w/v	weight/volume

1. Introduction

1.1 Messenger RNA turnover in eukaryotes

The control of messenger RNA (mRNA) degradation is an important component of the regulation of gene expression since the steady-state concentration of mRNA is determined both by the rates of synthesis and of decay. A high rate of turnover is actually one of the distinguishing features of mRNA. Equally rapid synthesis and degradation were proposed early on as key characteristic of a "messenger" carrying information from genes to proteins, based on the observation that both gene induction and repression ocurred within minutes (Jacob and Monod, 1961). The fast turnover of mRNA underlines the biological significance of the molecule's instability which allows a cell to adapt to changing physiological conditions (Ross, 1995).

In eukaryotes, there are several ways in which an mRNA can be degraded. Yet, the bulk of the transcripts are degraded by two general pathways that are constitutively active (see 1.1.1). Furthermore, pathways limited to subsets of mRNAs exist in which specific sequences trigger decay (see 1.1.2 and subsequent sections). The multiplicity of decay pathways is not redundant as it makes possible a differential regulation of individual mRNAs (reviewed by Beelman and Parker, 1995).

1.1.1 General pathways of mRNA turnover: the 5´and 3´decay pathways

The bulk of mRNAs decay appears to be catalyzed almost exclusively by exonucleases in eukaryotes (reviewed by Meyer et al., 2004). Probably for this reason mechanisms have evolved to chemically modify the ends of eukaryotic mRNAs and thus, to protect the transcripts. A m^7GpppN cap at the 5´end and a poly(A) tail at the 3´end assure mRNA stability.

Based on work primarily done in yeast and mammals, two general pathways of mRNA decay have been identified in eukaryotic cells (Tucker and Parker, 2000; Mitchell and Tollervey, 2001) (see Fig. 1).

Both pathways begin with the processive shortening of the poly(A) tail of the mRNA by a variety of deadenylases. So far, three different complexes have been identified as mRNA deadenylases and many of the proteins involved in these complexes have conserved

homologues among eukaryotes. In yeast, the so-called PAN complex consists of the proteins Pan2p and Pan3p. This complex is involved in an early step in poly(A) metabolism in which an initially long poly(A) tail is shortened to a length of 55-75 nucleotides (Brown and Sachs, 1998). Human homologues of Pan2p and Pan3p have been identified (Uchida et al., 2004), but their function has yet to be elucidated. The second and predominant deadenylase complex in yeast is constituted by two nucleases – Ccr4p and Pop2p – and several accessory proteins (Tucker et al., 2001; Denis and Chen, 2003). Several Ccr4p homologues have been identified in mammals. However, in vertebrates, the enzyme called PARN (poly(A) ribonuclease) seems to have the major deadenylase activity in the general pathways (Gao et al., 2000). In addition, PARN seems to be required for the rapid deadenylation induced by AU-rich elements (see 1.1.3) and can also affect the process of nonsense-mediated decay (see 1.1.4) (Lai et al., 2003; Lejeune et al., 2003). The various deadenylases described here differ in their substrate preference, biochemical properties and the way they are recruited. Even though the precise regulation of these enzymes is still poorly understood, the diversity of deadenylases would allow a tight control of the initial rate-limiting step of the mRNA turnover (reviewed by Parker and Song, 2004).

After deadenylation, the transcript is further degraded following one of two possible pathways. It is either degraded in the 3'-5' direction or from the 5' end by removal of the m^7GpppN cap and subsequent 5'-3' degradation (Anderson and Parker, 1998; Decker and Parker, 1993).

The 3´-5´ decay pathway might be the principal one in mammalian cells as suggested by *in vitro* experiments (Brewer, 1998; Mukherjee et al., 2002). A complex of conserved 3´exonucleases and associated polypeptides, together known as the exosome, is responsible for this pathway. Nine subunits are integrated into the exosome core forming a ring-like structure: six subunits share sequence similarity with bacterial 3´ to 5´exoribonucleases (Rrp41, Rrp42, Rrp46, PM/Scl-75, Mtr3, Oip2/Rrp43) and the other three (Rrp4, Rrp40 and Csl4) contain RNA-binding domains (Aloy et al., 2002; Symmons et al., 2002; Lehner and Sanderson, 2004; Liu et al., 2006). The exosome exists and has a role both in the nucleus and the cytoplasm; the complexes located in the two different compartments are distinguished by specific subunits and associated proteins (Butler, 2002) which regulate exosome activity. In the cytoplasm, this core structure interacts with Ski7 which, in turn, recruits the Ski2-3-8 complex (Araki et al., 2001). The whole extended complex is required to degrade cytoplasmic mRNA.

The 5´-3´ decay pathway (both in yeast and humans) is initiated after deadenylation with decapping by the heterodimer Dcp1/Dcp2. Dcp2 is the catalytic subunit while Dcp1 seems to be a co-activator (van Dijk et al., 2002; She et al., 2004). In yeast there is only one version of Dcp1 and Dcp2; in humans, there is only one known version of Dcp2 but two of Dcp1 (DCP1a and DCP1b) which might have different functional properties, thereby increasing the diversity of decapping activities (Lykke-Andersen, 2002). Moreover, it is increasingly evident that – apart from Dcp1 – a conserved group of proteins enhances decapping. The proteins in this group, which includes EDC3, Ge-1, the heptameric LSm1-7 complex, Dhh1 and Pat1, are collectively called decapping co-activators although the precise molecular function of most of them is unclear (reviewed by Eulalio et al., 2007).

Hydrolysis of the cap structure releases m^7GDP and leaves the transcript with a 5´monophosphate. This is the preferred substrate for the exonuclease Xrnp1 which is involved in the rapid 5´-3´ degradation of mRNAs in the cytoplasm, and this function is conserved between yeast and mammals (Bashkirov et al., 1997). Originally, the 5'-3' decay pathway was considered the major route of mRNA degradation in yeast (Muhlrad et al., 1995); however, more recently, a genome-wide survey of mRNA abundance showed that mutations affecting the 5´pathway changed the abundance of less than 20% of all yeast mRNAs, indicating that the 3´pathway might actually be more important (He et al., 2003).

In both pathways, the remaining cap structure is further broken down by the scavenger decapping enzyme DcpS (Wang and Kiledjian, 2001; van Dijk et al., 2003). DcpS decaps short cap oligonucleotides to release m^7GMP. It has also a second function in hydrolyzing the m^7GDP produced by Dcp1/Dcp2 to m^7GMP and phosphate (Liu et al., 2002).

1.1.2 mRNA decay via endonucleolytic cleavage

Certain eukaryotic mRNAs can be degraded via endonucleolytic cleavage prior to deadenylation (Fig. 2). The two fragments produced can then be degraded by the classical 5´ and 3´pathways. Evidence for the endonucleolytic mechanism comes from the analysis of transcripts such as mammalian IGF2, 9E3, the transferrin receptor (TfR) and *Xenopus* Xlhbox2B (Stoeckle and Hanafusa, 1989; Nielsen and Christiansen, 1992; Brown et al., 1993; Binder et al., 1994) where the selective degradation of the specific transcript is triggered in response to extracellular stimuli.

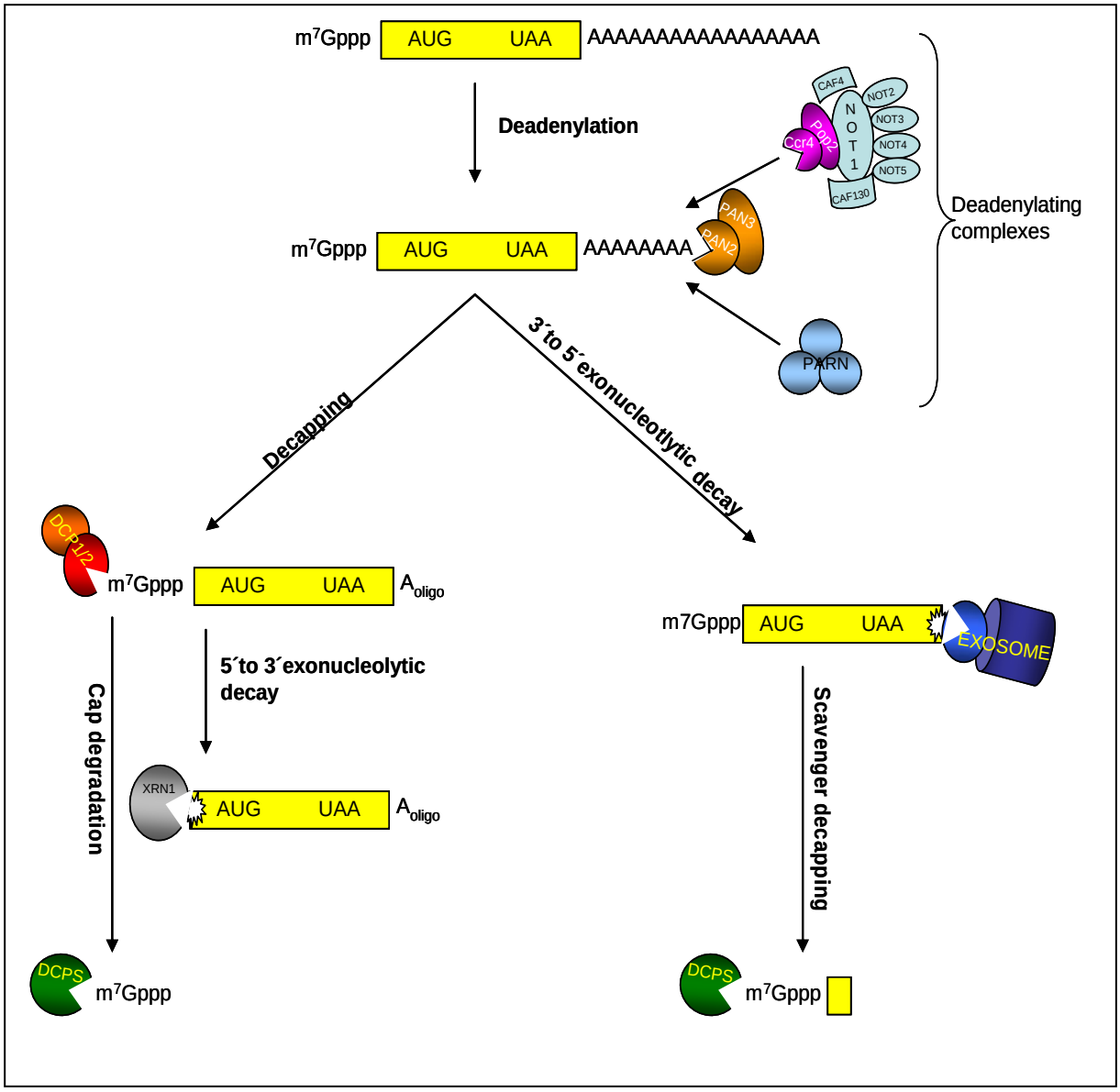

Figure 1. General mRNA decay pathways in eukaryotes. The constitutive turnover of most mRNAs is initiated by deadenylation, followed by either 5´ to 3´ or 3´ to 5´decay. Three different complexes are known for deadenylation. PAN2/3 may start the process followed by PARN (in higher eukaryotes) or CCR4/NOT (in yeast). The decapping complex involves DCP1 and DCP2 (which contains the enzymatic activity) plus other proteins which stimulate DCP2 (not shown). The exosome is responsible for the 3´ to 5´decay. The core exosome subunits form a barrel-like structure that interacts with the Ski complex (not shown). XRN1 degrades the uncapped transcript in the 5´ to 3 direction. DCPS has two related activities. In the 5´ decay pathway, it hydrolyzes the m^7GDP produced by DCP2 to m^7GMP. In the 3´decay pathway, it decaps short cap oligonucleotides (the final product of the exosome) to release m^7GMP. Adapted from (Parker and Song, 2004).

Some common mechanistic aspects have emerged from the examples of regulated endonucleolytic cleavage that have been characterised to some extent. Vertebrate endonucleases appear to act specifically at certain sites that are defined by their sequence or secondary-structure. However, there does not appear to be any similarity between the cleavage sites in these mRNAs. Consequently, it is possible that a wide variety of endonucleases with different cleavage specificities allows the control of the decay of limited

mRNAs or classes of mRNA (reviewed by Tourriere et al., 2002). Additionally, the presence of protective factors that bind at or near the cleavage site (thus, competing with the enzyme) further modulate the endonucleolytic activity. For instance, the binding of the iron response element-binding protein in the TfR 3´UTR in response to low intracellular iron concentrations inhibits the endonucleolytic degradation of this mRNA (Binder et al., 1994).

Decay via endonucleolytic cleavage has a more general role in *D. melanogaster* in a surveillance process known as nonsense-mediated decay, which is explained in the next chapter (see 1.2).

In the last ten years, a complete new chapter in endonucleolytic mRNA degradation has arisen with a specialised decay pathway that involves RNA interference (RNAi) (Meister and Tuschl, 2004; Mello and Conte, 2004). Two types of small RNAs can trigger RNAi: short interfering RNAs (siRNA) and fully-complementary micro-RNAs (miRNAs). Once processed, both types of RNAs are biochemically indistinguishable but they differ in their biogenesis (Carthew, 2006). In the case of siRNA, an initially longer dsRNA is cleaved into 21-23 nts double-stranded small interfering RNAs (siRNAs) by the RNase III-related enzyme Dicer (Bernstein et al., 2001; Elbashir et al., 2001). One strand of the siRNA is then incorporated into the RNA-induced silencing complex (RISC) where it serves as a guide for the selection of target mRNAs. Transcripts with complementary sequence to the siRNA are targeted and cleaved by an endonuclease, most likely the RISC-component Argonaute 2 (Ago2), in the middle of the recognised sequence (Pham et al., 2004; Liu et al., 2004; Song et al., 2004). This endonucleolytic cleavage is followed by the degradation of the mRNA fragments by the general exonucleolytic pathways (Montgomery, 2004).

MicroRNAs are generated by RNA polymerase II transcription which gives rise to a precursor that is subsequently processed by Drosha RNase III. The digested RNA has complementary sequences and, consequently, the molecule is folded in a hairpin-loop secondary structure. Although the exact silencing mechanism is not known, it is clear that this precursor undergoes a mode of silencing that is related to that employed by siRNAs (reviewed by Carthew, 2006).

RNA silencing has received an exceptional amount of attention due to its value for loss-of function studies. Importantly, siRNA transfection is used extensively in this study. However, the role of RNA silencing in the regulation of gene expression is only starting to emerge.

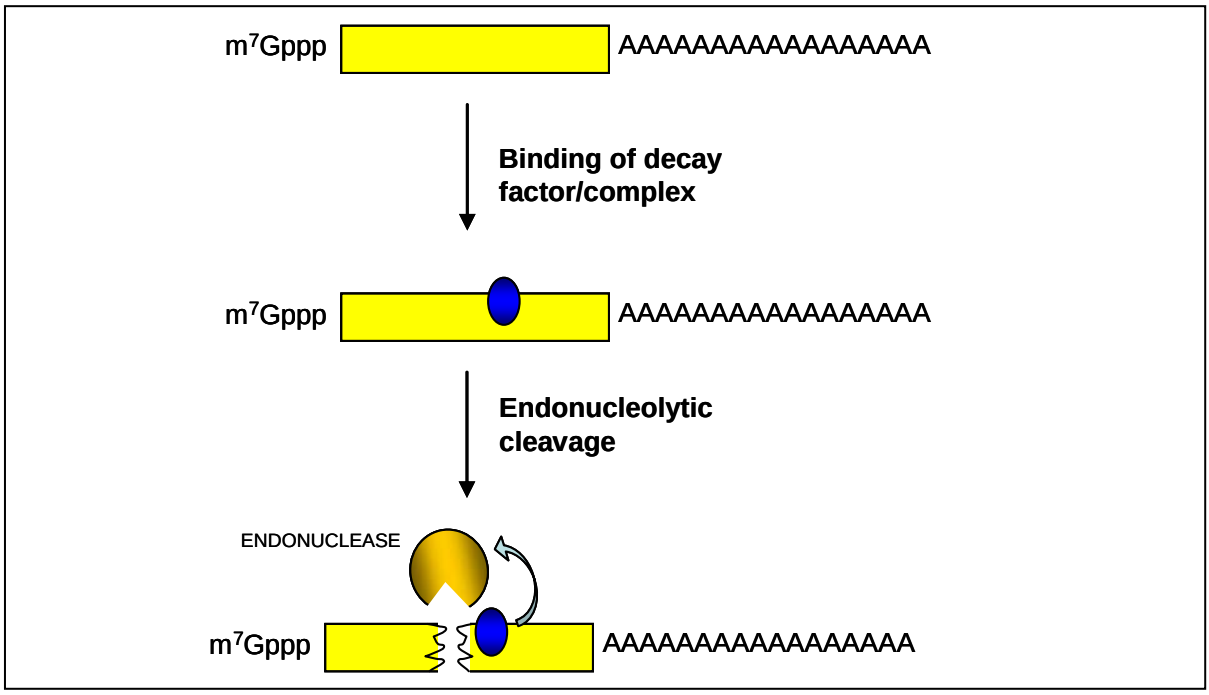

Figure 2. Endonucleolytic mRNA decay pathways. Some mRNAs are degraded by endonucleolytic cleavage as a first step. A generalisation of this pathway is illustrated here. A decay factor or complex (depicted in blue) recognises and binds to a sequence in the mRNA (generally but not exclusively in the 3´UTR). This factor/complex recruits – in turn – the endonuclease (interaction shown with an arrow). Examples of such a decay factor/complex are some ARE-binding proteins, IGFII-binding protein or, in fly, the NMD surveillance complex. It is also possible that a sequence in the transcript is recognised directly by the endonuclease (for instance, in RNAi where the RISC complex interacts with the mRNA through the siRNA) (not shown). Subsequently to this cleavage, the transcript can be further degraded by the 5´or 3´decay pathways showed in Fig. 1.

1.1.3 Regulated mRNA decay: AU-rich elements

In 1.1.2 some examples of transcripts bearing certain cis-acting sequences that serve to regulate their decay (through endonucleolytic attack) were shown. The cis-acting elements in those cases are specific for each mRNA. However, there also exits another type of elements that appears to be more general and that also modulates mRNA decay.

AU-rich elements (AREs) are sequence elements of 50-150 nts that are rich in adenosine and uridine bases (reviewed in Barreau et al., 2005). AREs are among the predominant cis-acting sequences that exist mainly in the 3'-UTR of mRNAs, regulating their stability. These elements usually contain AUUUA pentamers as the sequence motif. Based on the number and the distribution of these pentamers, AREs have been grouped into three classes (Bakheet et al., 2003). Class I AREs contain several dispersed copies of the AUUUA motif within U-rich regions. Class II AREs possess at least two overlapping UUAUUUA(U/A)(U/A) nonamers. Class III AREs are more loosely defined; they are U-rich regions but contain no AUUUA motif.

Originally, AREs were believed to be restricted to relatively few mRNAs, including those of interferons and cytokines, growth factors, and proto-oncogenes. More recent analysis, however, showed that ARE mRNAs represent as much as 8% of mRNAs transcribed from human genes that encode functionally diverse proteins important in many transient biologic processes (reviewed in Khabar, 2005). Among those processes are cell growth and differentiation, immune responses, transcriptional and translational control, hematopoiesis, apoptosis, signal transduction and nutrient transport.

Most of the ARE-containing transcripts are unstable (Frevel et al., 2003) and are degraded via the 3'-5' decay pathway, after deadenylation (Chen et al., 2001; Mukherjee et al., 2002). Nevertheless, degradation via endonucleolytic cleavage or the 5'-3' mRNA decay pathway also seem to contribute to the ARE-containing mRNA decay in some particular cases (Zhao et al., 2000; Gao et al., 2001). Many factors have been characterised to bind and regulate the stability of the ARE-containing mRNAs (reviewed by Barreau et al., 2005). Examples of these proteins are AUF1, TTP and KSRP which accelerate the turnover by recruiting deadenylases and the exosome to the transcript (Chen et al., 2001; Mukherjee et al., 2002; Tran et al., 2004). However, not all the ARE-binding factors have a destabilising effect. For instance, HuR and YB-1 are stabilising ARE-binding proteins that either compete with the destabilising factors for access to binding sites or protect the mRNA from decay by masking cleavage sites (Fan and Steitz, 1998; Capowski et al., 2001). Several of these trans-acting factors have been found to be targeted by signalling cascades. Growth factors and cytokines can affect the expression, post-translational modification and localisation of ARE-binding factors (reviewed by Shim and Karin, 2002). Thus, the abundance of ARE-containing transcripts can be rapidly adjusted in response to the extracellular stimuli.

Interestingly, recent studies also implicate a role of miRNAs in the regulation of mRNA stability by AREs. The human miRNA miR16 has been shown to contain an AU-rich sequence that is complementary to the ARE in the TNF-α transcript (Jing et al., 2005). This miRNA co-operates with the ARE-binding factor TTP in the destabilisation of the transcript. In contrast to this example, it has been shown that miR-122 induced translation repression of CAT-1 transcript is reversible and that HuR binding is required to overcome this inhibition (Bhattacharyya et al., 2006). These novel functions of miRNA and ARE elements (or ARE binding proteins) indicate that we are only starting to understand the complex molecular mechanisms involved in the regulation of mRNA turnover.

1.1.4 mRNA surveillance pathways

The flow of information from a gene to the protein it encodes is full of intermediate steps that allow a tight control of the gene expression at multiple levels. Splicing of pre-mRNA and other maturation processes, mRNA export and translation operate cooperatively and concurrently to maintain the fidelity of the genetic information. However, each of these levels is prone to errors and hence, quality control mechanisms are needed.

For instance, the processes of transcription, capping, splicing and nuclear export of mRNA are all closely connected by the use of common factors (reviewed by Reed, 2003; Kornblihtt et al., 2004; Proudfoot, 2004). In addition, the link between consecutive steps in the mRNA maturation limits the production of partially processed transcripts. Still, faulty processed transcripts are retained in the nucleus and degraded by the nuclear exosome (Bousquet-Antonelli et al., 2000; Burkard and Butler, 2000; Olesen et al., 2005).

In the cytoplasm, at least three surveillance mechanisms verify the translatability of the mRNAs. Transcripts that have no translation termination codons in their ORFs are the targets of the nonstop decay (NSD) pathway (Frischmeyer et al., 2002). During nonstop decay, ribosomes that stall at the 3' end of the mRNA without terminating properly cause the recruitment of factors involved in 3'-5' decay pathway, including the Ski complex and the exosome which ultimately degrades the transcript (van Hoof et al., 2002).

A different surveillance mechanism – termed no-go decay (NGD) – has been defined in yeast. In this case, messengers carrying a stable stem-loop structure which stall translation are degraded by an endonucleolytic cleavage (Doma and Parker, 2006). This pathway also depends on translation as a stop codon located upstream of the stem-loop inhibits degradation. Interesting, Hbs1 and Dom34 (interacting proteins with similarity to translation termination factors) are indispensable for NGD (Doma and Parker, 2006).

A third surveillance mechanism – nonsense-mediated mRNA decay (NMD) – targets transcripts that possess a premature translation termination codon (PTC) (for a review, see for instance Hentze and Kulozik, 1999; Conti and Izaurralde, 2005; Weischenfeldt et al., 2005). In mammals, the definition of stop codons as either 'normal' or 'premature' depends on their position relative to the last intron. Once a PTC has been recognised, the mRNA is rapidly degraded from either end. As this surveillance pathway is the subject of the present thesis, NMD is described in detail in the following chapter (see 1.2).

1.1.5 Localisation of decay factors

It has long been recognised that the different mRNA decay pathways are functionally interlinked. However, the notion that these processes are also physically connected has also recently emerged. Proteins that are involved in mRNA degradation, mRNA surveillance, translational repression and RNA-mediated gene silencing, along with their mRNA targets, co-localise within discrete cytoplasmic granular structures known as processing-bodies or P-bodies (see Fig. 3) (reviewed by Bruno and Wilkinson, 2006).

The first proteins shown to be present in P-bodies were those that function in the general 5´decay pathway such as DCP1/2, XRN1 and the Lsm proteins. (Cougot et al., 2004). Later on, the UPF proteins, SMG5 and SMG7 – all factors implied in mRNA surveillance (see 1.2) – were also found in these structures (Sheth and Parker, 2006). Moreover, Argonaute proteins and GW182, which are essential for miRNA-mediated silencing, also co-localise in P-bodies (Eystathioy et al., 2002; Liu et al., 2005; Behm-Ansmant et al., 2006). Interestingly, factors that integrate the exosome and Ski-complex are not detected in these granules (Sheth and Parker, 2003; Brengues et al., 2005).

In principle, P-bodies could simply represent storage sites for decay enzymes. However, there is increasing evidence that supports the hypothesis that P-bodies are actually sites of mRNA degradation. For instance, when cells are treated with ribonuclease A or with actinomycin D, P-bodies are lost, indicating that P-body assembly is dependent on RNA. If mRNA degradation is blocked at an early stage – for example, by preventing deadenylation in cells that lack the deadenylase CCR4 – P-bodies disappear. However, if the degradation is blocked at a later stage (for instance by depletion of XRN1 or DCP1) the P-bodies increase in size and number which suggests that transcripts in the P-bodies are undergoing decapping. At the same time, mRNA-decay intermediates are detected in P-bodies when XRN1 is depleted. Similarly, transcript-decay intermediates are detected when the progression of XRN1 along an mRNA is blocked by the insertion of specific RNA sequences (Sheth and Parker, 2003; Cougot et al., 2004; Andrei et al., 2005; Teixeira et al., 2005). Taken together, these observations definitively establish that P-bodies are storage-sites for mRNAs that are committed to degradation. Yet, it remains unclear whether the mRNA is responsible for recruiting the decay factors or there exist unknown interactions between the enzymes in the different pathways that bring these factors together in the P-bodies. In any case, the simultaneous presence of decay factors would make the mRNA degradation process more efficient and controlled.

More recently, additional experimental evidence suggests that P-bodies could have a broader role in the regulation of mRNA turnover. The translation of CAT-1 mRNA is regulated by the microRNA miR-122 in hepatic cells. A study demonstrated that this transcript is present in the P-bodies while is repressed by miR-122. However, under stress conditions the repression is relieved (in a process that also involves the ARE-binding protein HuR), releasing CAT-1 mRNA from the P-bodies and allowing active translation (Bhattacharyya et al., 2006). Thus, P-bodies do not only represent sites of mRNA decay but also reservoirs of reversible pools of translationally inactive mRNAs.

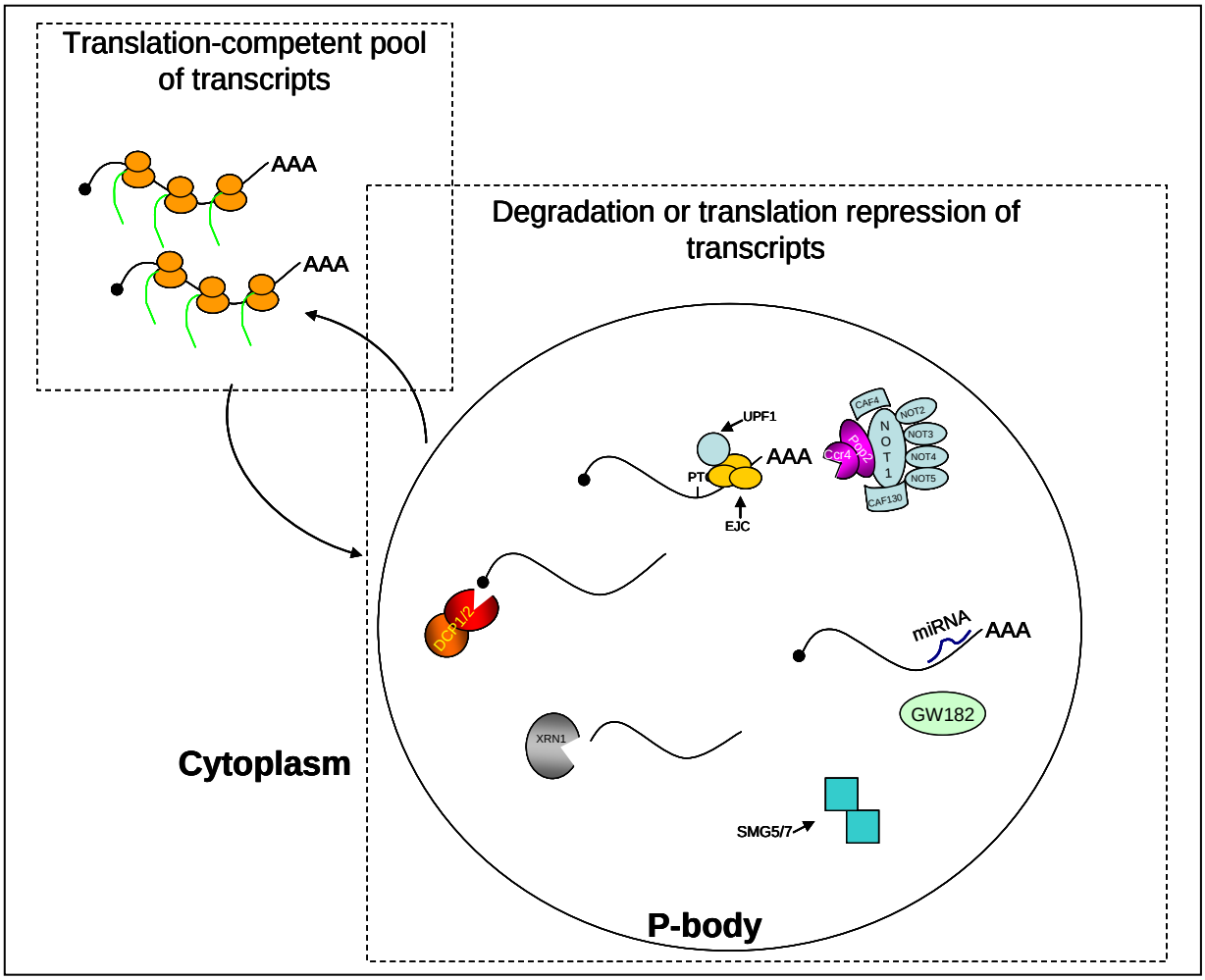

Figure 3. P-bodies. The P-bodies are granular structures that contain transcripts undergoing decay or translation-repression and factors involved in several decay pathways. Specifically, factors involved in NMD, deadenylation and the 5´-3´general decay pathways have been identified. In addition, proteins essential for miRNA silencing have also been found. Besides mRNA decay, it has been postulated that also reversible translation-repression occurs in P-bodies translation (Bhattacharyya et al., 2006; Bruno and Wilkinson, 2006). Not all the known components of P-bodies are represented in the scheme.

1.2 Nonsense-mediated mRNA decay

Nonsense-mediated mRNA decay (NMD) is an mRNA surveillance pathway that detects and degrades transcripts containing premature termination codons (PTCs) thus, preventing the synthesis of truncated proteins. Despite differences at the mechanistic level, NMD is functionally conserved in all studied eukaryotes including yeast (Leeds et al., 1991; Leeds et al., 1992), vertebrates (Maquat, 1995; Perlick et al., 1996), *C. elegans* (Pulak and Anderson, 1993; Cali and Anderson, 1998), *D. melanogaster* (Gatfield et al., 2003) and plants (Abler and Green, 1996).

NMD is always characterised by rapid mRNA decay, and three key trans-acting NMD factors – UPF1, UPF2 and UPF3 – are conserved from yeast to mammals (see for instance Culbertson and Leeds, 2003); however, the cis-acting elements and details in the pathway vary among the different species. For instance, mammals have a particular mechanism of recognition of premature nonsense codons that diverges from other organisms. For this reason, I will discuss NMD in mammals separately.

1.2.1 NMD in mammals

The history of NMD in mammals was opened with a pioneering study (Chang and Kan, 1979) where it was observed that the transcript steady-state levels of a mutated PTC-containing β-globin gene were lower than the wild type counterpart. From that moment intensive research has unravelled many details in the mechanism of NMD, although there are still several open questions.

In mammals, the definition of a PTC depends on the location of the termination codon in relation to the introns of the pre-mRNA. A translation termination codon is generally interpreted as “normal” if no intron follows more than 50-55 nucleotides downstream in the gene sequence (Nagy and Maquat, 1998). Thus, most normal stop codons are located in the last exon, while NMD-activating PTCs are usually situated further upstream (see Fig. 4). For this discrimination, both splicing and translation are critical. The requirement of splicing is demonstrated by the NMD insensitivity of intronless cDNA versions of PTC-containing genes normally subjected to NMD (Maquat and Li, 2001; Neu-Yilik et al., 2001). Additionally, the importance of translation has been demonstrated unequivocally since the pleiotropic (Carter et al., 1995) or specific (Thermann et al., 1998) inhibition of translation abrogates NMD of a PTC-mutated transcript.

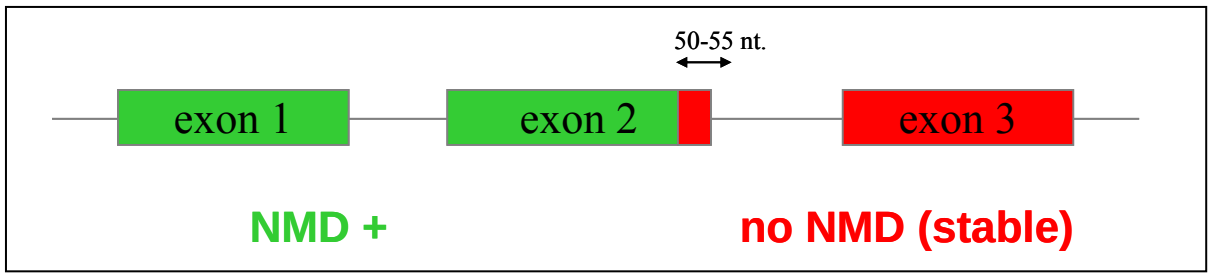

Figure 4. PTC definition in mammals. A nonsense codon is regarded as premature if it is in frame with the initial ATG and it is located at least 50-55 nt upstream of the 3´most intron. Depicted is an example of a gene with three exons. According to the "50 nt rule" a stop codon located in the green region would trigger NMD while a stop codon situated in the red region would be considered legitimate and would not trigger NMD.

The combined requirement of splicing (which takes place in the nucleus) and of translation (which occurs in the cytoplasm) represented a puzzle for the understanding of the NMD mechanism. The explanation of how the introns in a pre-mRNA function in NMD came from the finding of a protein complex that is deposited ~20-24 nucleotides upstream of each exon-exon junction (in a sequence-independent manner) after the intron is spliced out (Le Hir et al., 2000). This complex – called the "exon junction complex" (EJC) – provides a ´mark´ that communicates in the cytoplasm the position where an intron was previously located.

The EJC is a stable structure that assembles during splicing, and is remodelled during the subsequent steps of mRNA processing, with some components dissociating before or while the mRNP is exported to the cytoplasm and others joining the complex at later stages (reviewed by Tange et al., 2004). It is believed that the EJCs are removed from the mRNA by the translating ribosome during a first round of translation (Dostie and Dreyfuss 2002; Lejeune et al. 2002). While the EJC is still bound to the mRNA, it affects a number of post-transcriptional processes such as mRNA decay, localisation and translation (reviewed by Tange et al., 2004).

The stable core of the EJC contains at least four proteins: MAGOH and Y14 (which form a heterodimer), eIF4A3, and Barentsz (also known as MLN51) (reviewed by Tange et al., 2004). Recently, the crystal structure of the core complex has been revealed. The protein eIF4A3 has two helicase domains joined by a short linker. The arrangement of these domains creates a cleft that is where the RNA binds. Barentsz wraps around eIF4A3 and along with the heterodimer Y14/MAGOH stabilises the binding of the RNA (Bono et al., 2006). To this core complex, other components associate more peripherally or transiently. Such factors include the splicing co-activators SRm160 and RNPS1, the mRNA export factors UAP56, REF/Aly and TAP-p15 and – more recently uncovered – SAP18 and Acinus (reviewed by Tange et al., 2004 and Tange et al., 2005) (see Fig. 5).

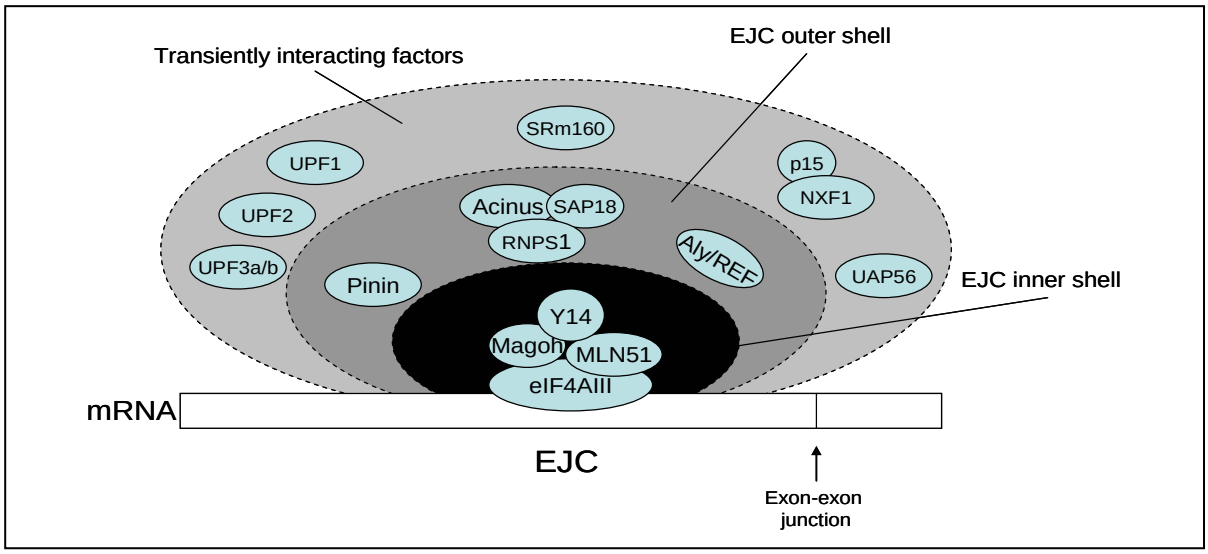

Figure 5. The Exon Junction Complex (EJC). The minimal EJC core (in black background) likely consists of a tetrameric complex containing eIF4AIII, MLN51 (also called Barentz), MAGOH, and Y14. All factors in this core are shuttling proteins and most likely follow the mRNA to the cytoplasm. In dark grey, the factors that constitute the outer shell are depicted. RNPS1, Acinus, and SAP18 can stably associate and may bind the EJC core as a trimeric complex. However, RNPS1 may also bind alone, e.g., via interactions with Pinin. In the light grey sphere are depicted those factors that bind transiently or in latter steps (as the Upf proteins). Modified from Tange et al., 2005.

The role of the EJC in identifying a stop codon as a PTC has been demonstrated by tethering EJC factors at least 50 nucleotides downstream of a stop codon to an otherwise normal mRNA. In this situation, a normal termination codon is redefined as premature and the transcript becomes unstable. RNPS1, Y14, MAGOH, Barentsz and eIF4AIII were shown to trigger NMD of a reporter construct in this system (Lykke-Andersen et al., 2001; Fribourg and Conti, 2003; Gehring et al., 2003; Palacios et al., 2004; Gehring et al., 2005). Additionally, depletion of Y14, Barentz and eIF4A3 by means of RNAi limits NMD in mammalian cells (Gehring et al., 2003; Ferraiuolo et al., 2004; Palacios et al., 2004; Shibuya et al., 2004).

Central to the NMD pathway are the UPF proteins: UPF1, UPF2 and UPF3. These proteins were originally identified as suppressors of the *his4-38* frameshift mutation (Culbertson et al. 1980) in yeast (UPF stands for "Up-Frameshift Suppressor"). Soon, homologues of these factors were identified in mammals (Sun et al., 1998; Serin et al., 2001). Unlike the EJC proteins which are essential for NMD only in mammals, the UPF family is crucial in all species where NMD is present (reviewed by Hentze and Kulozik, 1999; Maquat and Serin. 2001).

For UPF3, two closely related genes have been identified in mammals: UPF3a and UPF3b (Lykke-Andersen et al., 2000). These are shuttling proteins that are integrated into the EJC in the nucleus. Originally, it was thought that both isoforms played a role in NMD. More

recently, it was shown that only UPF3b participates in the surveillance pathway (Kunz et al., 2006). The NMD-function of UPF3b was shown by its ability to phenocopy NMD in tethering assays (Lykke-Andersen et al., 2000; Gehring et al., 2003).

UPF2 is preferentially located in the perinuclear region and it is also recruited by the EJC by interaction with UPF3 (Lykke-Andersen et al., 2000). The role of mammalian UPF2 in NMD was demonstrated by RNAi (Mendell et al., 2000), antisense approaches (Wang et al., 2002) and also by tethering (Lykke-Andersen et al., 2000) .

UPF1 is a predominantly cytoplasmic phospho-protein with ATPase and RNA helicase activity (Leeds et al., 1991; Czaplinski et al., 1995). It can interact directly with UPF2 (Lykke-Andersen et al., 2000) and has to be phosphorylated by the phosphatidyl-inositol 3-kinase-related kinase SMG1 to be NMD-competent (Yamashita et al., 2001). When UPF1 is down-modulated by RNAi or when a dominant negative mutant of UPF1 is over-expressed, NMD efficiency is markedly diminished (Sun et al., 1998; Mendell et al., 2002; Gehring et al., 2003). In addition, inhibition of UPF1 phosphorylation by wortmannin-treatment also impairs NMD (Pal et al., 2001; Yamashita et al., 2001).

Current NMD models postulate that during the first round of translation, the ribosome removes all the EJCs as it traverses the mRNA till it reaches the stop codon. If no EJC remains downstream of the ribosome position, a normal termination occurs (reviewed by Maquat 2004). Otherwise, some investigators have hypothesised that a linear pathway takes place in which the EJC sequentially recruits UPF3b, UPF2, and UPF1 to trigger NMD (Kim et al., 2001; Lykke-Andersen et al., 2001) (see Fig. 6). Recently, it has been shown that UPF1 forms a complex, named SURF, with SMG1 and the release factors eRF1 and eRF3. The interaction of this complex with the EJC is thought to be required for NMD (Kashima et al., 2006). In this model, SURF interacts with UPF2, UPF3b and additional EJC proteins bound to a downstream exon-exon boundary. This interaction would result in the formation of a decay-inducing complex (DECID) that – in turn – triggers UPF1 phosphorylation and the dissociation of eRF1 and eRF3 (Kashima et al., 2006). Finally, phosphorylated UPF1 is directly recognized by SMG5, SMG6 and SMG7 via a conserved 14-3-3 like domain present in all three proteins (Fukuhara et al., 2005). These proteins are thought to serve as adaptors between phosphorylated UPF1 and protein phosphatase 2A (PP2A), thereby triggering UPF1 dephosphorylation (Chiu et al., 2003; Ohnishi et al., 2003). However, the molecular mechanism of PP2A recruitment is still unknown.

New findings have lately challenged the linear model discussed here. In our group, it was found that two distinct subcomplexes of the EJC serve as entry points for the formation

of distinguishable NMD-activating mRNPs. One of these subcomplexes includes Y14, MAGOH, and eIF4A3 and can activate NMD in an UPF2-independent manner. The second subcomplex that can trigger NMD comprises RNPS1 and UPF2. The two branches of these nonlinear NMD model converge at a common requirement of UPF1 (Gehring et al., 2005).

How is the decay process triggered? Apparently, NMD in mammalian cells degrades mRNAs from both the 5' and 3' ends (Lejeune et al., 2003). Originally it was thought that this process was independent of prior deadenylation as it is the case in yeast (Hilleren and Parker, 1999) . However, additional evidence supports a rapid deadenylation step as the first decay event in humans (Chen and Shyu, 2003). There is some evidence that the UPF factors interact and recruit components of the general decay pathways. In particular, UPF proteins co-immunopurify with Dcp2, Xrn1, the exosomal components PM/Scl100, Rrp4, and Rrp41, and PARN (Lejeune et al., 2003).

It is worth noticing that UPF1, SMG5 and SMG7 localise in the P-bodies; the proximal location of the decay enzymes could accelerate the degradation of the PTC-containing transcript (reviewed by Eulalio et al. 2007). Moreover, it has recently been demonstrated that SMG6 has intrinsic nuclease activity *in vivo* (Glavan et al., 2006). This activity could constitute an alternative pathway that contributes to the rapid decay of transcripts that terminates translation prematurely.

1.2.2 NMD in yeast, fly and worm

Three genes (Upf1, Upf2 and Upf3) are essential for NMD in yeast. These genes are the orthologs of the respective mammalian factors (Gonzalez et al., 2001). In nematodes, seven genes named smg 1-7 (for suppressor with morphological effect on genitalia) are involved in this pathway (Pulak and Anderson, 1993). While smg 1,5, 6 and 7 are respectively homologous to SGM1, 5, 6 and 7 genes in mammals; Smg2, 3 and 4 are homologous to UPF1, 2 and 3. With the exception of smg7, *D. melanogaster* also has equivalent genes for all the sgm factors and, like in nematodes and mammals, sgm 1-6 are indispensable for NMD in fly (Gatfield et al., 2003).

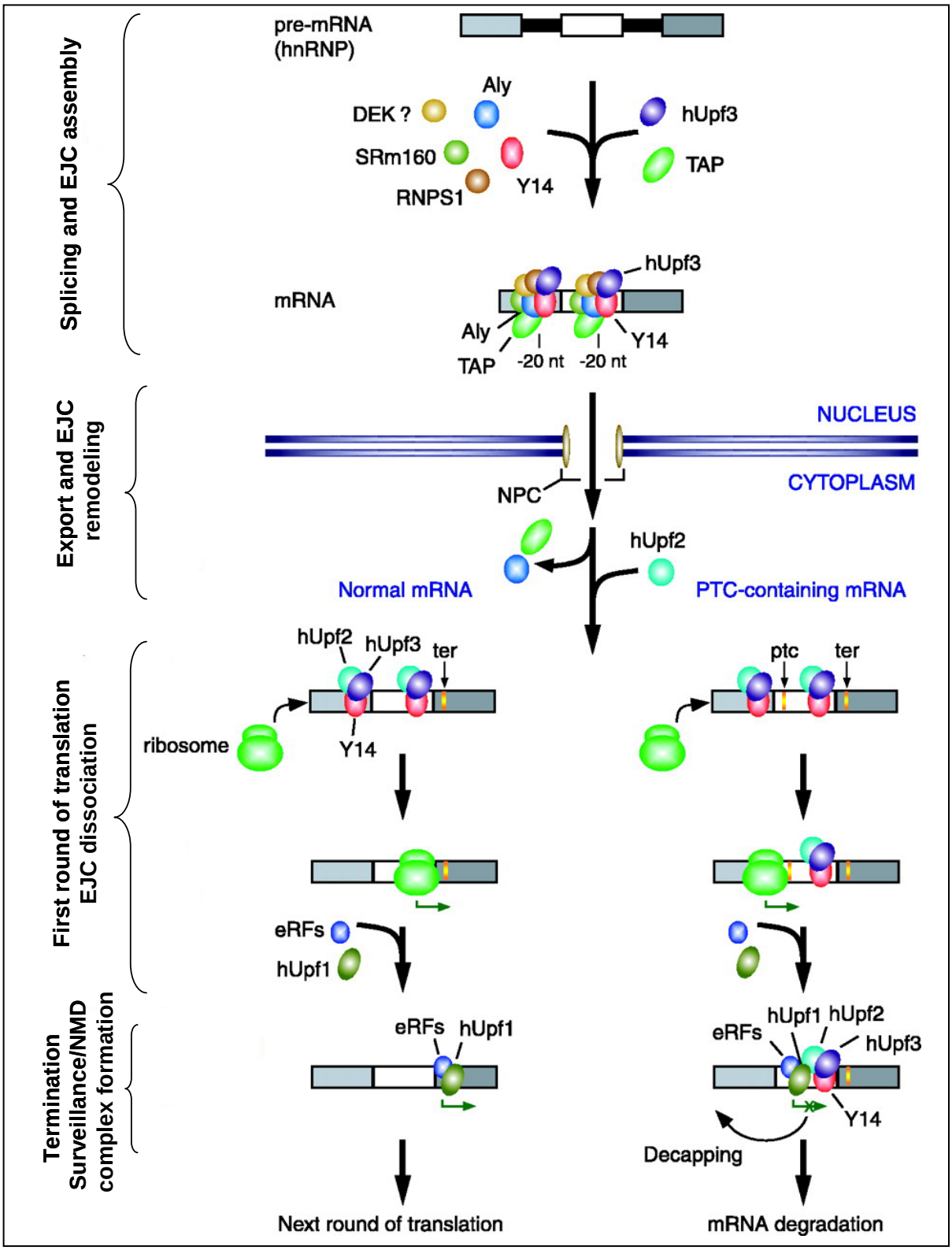

Figure 6. A current model for NMD in mammalian cells. The EJC assembles after splicing in the nucleus. The NMD factor UPF3 associates to the EJC at this stage. During the transcript export, the EJC is remodelled. In the perinucleus, UPF2 joins the complex. The EJCs are removed from the transcript during a pioneer round of translation by the ribosome. If translation terminates before the last EJC is removed, then the EJC is able to interact with UPF1 and the termination factors. An active surveillance complex is formed with UPF1, UPF2 and UPF3 which targets the transcript for decay in a yet unknown way. (Adapted from Kim et al, 2001).

One common feature of NMD in these lower eukaryotes is that – unlike the mammalian case – it is splicing-independent; consequently, there should be an alternative mechanism to define a PTC, distinct from the EJC position.

In yeast, this is not surprising since only 5% of yeast genes contain introns (Spingola et al., 1999) and none of the proteins found in the core of the mammalian EJC has an homologues in yeast (reviewed by Conti and Izaurralde, 2005). In *S. cerevisiae*, some mRNAs have been shown to harbour loosely defined downstream sequence elements (DSEs) with a function analogous to that of mammalian exon-exon junctions (Ruiz-Echevarria et al., 1998). The protein Hrp1p has been shown to bind to this element and to interact with the UPFs factors, thus associating the NMD complex and the DSE (Gonzalez et al., 2000). According to this model, once a translating ribosome encounters a PTC, the translation release factors eRF1 and eRF3 recruit Upf1 to the ribosome to assemble a surveillance complex. Subsequently, this complex searches 3' of a nonsense-codon for a DSE associated with RNA-binding proteins (Gonzalez et al., 2000). A direct interaction between the release factors and Upf1 has been proven (Czaplinski et al., 1998). Also, Upf1 ATPase and helicase activities haven been verified biochemically (Weng et al., 1996). However, the experimental evidence indicates there is a lack of a strong consensus among yeast DSEs which would suggest that multiple *cis*-acting sequence elements and *trans*-acting binding factors may exist.

Alternatively, it is possible that a generic feature of the mRNA, such as the poly(A)-tail or a mark deposited during the cleavage and polyadenylation reaction, provides the positional information needed to discriminate premature from natural stop codons in yeast. One model that considers the stop codon in the context of the 3´UTR and has recently received strong support is the so called “faux 3´UTR” (Amrani et al., 2004). It was shown that translation termination is aberrant at premature stop codons and that prematurely terminating ribosomes fail to release efficiently. This effect is abolished in strains lacking Upf1 or if the nonsense codon is flanked with a normal 3′ UTR. Moreover, tethering the poly(A)-binding protein (Pab1p) downstream of the PTC, which mimics a normal 3′UTR, leads to efficient translation termination and abolishes NMD (Amrani et al., 2004).

Although not experimentally proven, the faux 3´UTR model has also been proposed for nematodes and flies since NMD-sensitive transcripts do not possess any identifiable downstream sequence elements (Conti and Izaurralde, 2005).

As in mammals, studies in yeast indicate that the degradation of the targeted transcripts involves both the 5´-3´ and 3´-5´general decay pathways (although previous deadenylation is not required) (Hagan et al., 1995; Cao and Parker, 2003; Mitchell and

Tollervey, 2003). Which – if any – pathway is more relevant is still matter of debate (He et al., 2003).

In contrast to these two organisms, in *D. melanogaster* the degradation is initiated endonucleolytically, with a cleavage in the vicinity of the PTC (Gatfield and Izaurralde, 2004). The two generated fragments are further degraded by the general turnover mechanism. It will be interesting to see whether this pathway is an exception in fly or if it exemplifies a more general NMD decay pathway in other organisms.

1.2.3 Physiological role of NMD

NMD detects and degrades mRNAs containing PTCs; thus, avoiding the synthesis of C-terminally truncated proteins that might have a dominant-negative effect on the cell (reviewed by Hentze and Kulozik, 1999; Maquat, 2004). Originally, NMD was considered as a pathway that evolved to rid the cell of non-functional mRNAs arising from nonsense or frameshift mutations and splicing errors. As a matter of fact, NMD was first studied in cases of genes with nonsense mutations (Chang et al., 1979; Losson and Lacroute, 1979; Maquat et al., 1981). (In human, the role of NMD in controlling the expression of nonsense mutated genes is of particular interest because it might modulate the phenotypic expression of diseases caused by the genetic defect. For this reason, this aspect of NMD is treated separately in the next section (see 1.2.5)).

However, it was soon recognised that it is unlikely that NMD has evolved and maintained only to control the expression of mutated genes since the mutation frequency is actually very low in eukaryotes (Sueoka, 1993). Also, monitoring the expression of faulty spliced transcripts would not be expected to be a crucial function of this pathway as mRNAs containing introns are usually retained and degraded in the nucleus (reviewed by Vasudevan and Peltz, 2003).

A first role of NMD in the normal cellular physiology was uncovered in immunoglobin genes´ rearrangements, such as those of immunoglobulin (Ig) genes in B lymphocytes or the T cell receptor (TCR). In these cases, an exon encoding the antigen-binding domain is assembled by the joining of V (depending on the case also D) and J gene segments randomly selected from a large number of alternative sequences (Jung and Alt, 2004). In addition, the recombination site is not precise and several nucleotides can be added or lost at this point. The advantage of this mechanism to generate diversity in the immune

repertoire is balanced by the risk of changing the reading frame. As a consequence, many of the recombination events generate TCR or Ig genes that possess PTCs. In order to prevent the accumulation of truncated proteins, PTC-containing Ig and TCR transcripts are efficiently degraded by NMD (Carter et al., 1995; Li et al., 1997; Buhler et al., 2004; Delpy et al., 2004).

NMD of Ig or TCR transcripts is particularly efficient. The expression levels of the Ig or TCR mRNAs that possess a PTC are generally 30- to 100-fold lower than those of their wild-type counterparts (Buhler et al., 2004). In comparison, other PTC-containing transcripts are usually downregulated by a factor of 2 to 10 (Gudikote and Wilkinson, 2002). At least in the cases of the TCR and the Ig-μ gene, the high efficiency of NMD has been attributed to the presence of cis-acting sequence elements in the variable regions that function as enhancers of NMD upstream of a PTC (Gudikote and Wilkinson, 2002; Buhler et al., 2004).

Further insight into the physiological role of NMD was gained with the study of the regulation of the expression of splicing factors such as SC35 and the polypyrimidine tract binding protein (PTB). Transcription of both SC35 and PTB genes give rise to multiple isoforms by alternative splicing; some of these mRNAs have a PTC and are subjected to NMD. Notably, when these splicing factors are highly abundant, they can bind to their own pre-mRNAs and direct the splicing event toward the generation of NMD-sensitive isoforms. In this way, these factors can auto-regulate their own expression by means of the NMD machinery (Sureau et al., 2001; Wollerton et al., 2004).

Recently, a computer analysis of EST sequences resulted in the estimate that ~30% of the mRNAs that derive from alternative splicing contain a nonsense codon that can elicit NMD in humans (Lewis et al., 2003). This would indicate that NMD exerts significant regulation of gene expression through degradation of splice forms containing PTCs.

Several transcripts that encode selenoproteins constitute another group regulated by NMD. In this case, the mRNA contains a premature termination codon (UGA) that is not interpreted as such by the ribosome but as a codon encoding the rare amino acid selenocysteine. The incorporation of this amino acid depends on the secondary structure of the 3'UTR (reviewed by Namy et al., 2004) and on adequate levels of selenium. In a situation of low selenium concentration, a deficiency of selenocysteine-tRNA will make the translating ribosome stall at the premature UGA codon. In that case, NMD will eliminate a transcript that cannot be properly translated (Sun et al., 2000; Sun and Maquat, 2002).

An additional group of transcripts that are potential NMD substrates is made up of mRNAs that contain upstream open reading frames (uORFs). uORFs are short open reading

frames in the 5' leader sequence of a transcript, which often regulate the translation of the downstream major ORF (reviewed by Meijer and Thomas, 2002). uORFs can potentially affect the mRNA stability because the stop codon of an uORF can be regarded as a PTC (Ruiz-Echevarria et al., 1996; Linz et al., 1997). Experimentally, the NMD-sensitivity of uORF-containing transcripts has been verified in yeast (Ruiz-Echevarria and Peltz, 2000; He et al., 2003), worm (Lee and Schedl, 2004), fly (Rehwinkel et al., 2005) and mammals (Mendell et al., 2004). On the other hand, it has also been documented that the seventh uORF of the thrombopoietin (TPO) – a theoretical NMD target – is insensitive to NMD (Stockklausner et al., 2006). Moreover, reporter mRNAs bearing this TPO uORF escape NMD as well. This implies that mRNAs bearing uORFs cannot always be considered to represent NMD targets.

The increasing number of discovered physiological NMD targets stimulated a genome-wide analysis of the set of transcripts regulated by NMD. This has been successfully achieved in yeast (Lelivelt and Culbertson, 1999; He et al., 2003) and fly (Rehwinkel et al., 2005). To this end, a microarray-based expression profile analysis was done in both cases. In yeast, mutated strains lacking Upf1, Upf2 or Upf3 compared to a wild type strain were assessed; while for fly, NMD factors (Upf1, Upf2 and Upf3, Smg1, Smg5 and Smg6) were depleted by RNAi.

Several conclusions were drawn from these experiments. Firstly, about 5-10% of the transcriptome of these species were affected by NMD abrogation, confirming the significant role of this pathway. Secondly, the set of up-regulated transcripts was approximately the same regardless of which factor was defective. This result supports the idea that the NMD factors act principally in the surveillance complex and that their role in mRNA decay has not diverged substantially. Approximately 30% of the affected transcripts contain structural features that explain their NMD-sensitivity. These include mRNAs encoded by genes harbouring nonsense mutations; pre-mRNAs that retain their introns in the cytoplasm, mRNAs with uORFs, mRNAs subject to leaky scanning, mRNAs using frameshift in their translation, bicistronic mRNAs and mRNAs encoded by pseudogenes or transposable elements. Both in yeast and fly, the affected genes were associated with diverse functions including metabolism, cell cycle, DNA processing, protein synthesis and cellular transport (He et al., 2003; Rehwinkel et al., 2005). However, the specific genes involved in each functional group are not homologous in these species. This would indicate that there was not evolutionary pressure to conserve the identity of the set of regulated genes.

1.2.5 Role of NMD in disease

In humans, the study of NMD is of particular interest in relation to the clinic. An estimated 30% of inherited disorders and many forms of cancer are caused by frameshift or nonsense mutations which result in the generation of PTCs and hence, are potentially targeted by NMD (Frischmeyer and Dietz, 1999).

The role of NMD in disease was first appreciated in certain types of β-thalassemia where NMD has a protective role. Adult haemoglobin is composed of two α- and two β-globin subunits that form a quaternary complex necessary for oxygen transport. When the β-globin gene contains NMD-sensitive PTC mutations, aberrant β-globin mRNA is degraded by NMD. Thus the synthesis of truncated β-globin is restricted and these mutations cause a recessive form of β-thalassemia (Hall and Thein, 1994). Contrary to homozygotes, heterozygotes carrying these mutations generally synthesise enough β-globin from the remaining normal allele to support near-normal haemoglobin levels and consequently do not suffer from thalassemia. In contrast, rare nonsense mutations in the last exon of β-globin (which are NMD-insensitive) give rise to large amounts of truncated, non-functional β-globin that cannot be eliminated by the cell's proteolytic system and causes toxic precipitation of insoluble globin chains (Thein et al., 1990). The remarkably contrast between asymptomatic heterozygotes with NMD-competent mutations and affected heterozygotes with NMD-insensitive mutations indicates the protective effect of NMD in this disorder (Kugler et al., 1995).

NMD also seems to play an important role in certain types of cancer. Mutations in tumor-suppressor genes are common steps in the development and progression of this disease. NMD has been reported to reduce the abundance of PTC-containing transcripts of the *BRCA1* (Perrin-Vidoz et al., 2002), *p53* (Williams et al., 1998) and Wilms tumor (WT) genes (King-Underwood and Pritchard-Jones, 1998). In support of a protective effect of NMD in cancer, intronless cDNA versions of these mutated genes, that are NMD-insensitive, were expressed in cell lines or animal models. The uncontrolled synthesised of C-terminally truncated proteins exerted dominant negative effects, such as increased chemoresistance, increased tumorigenicity (Fan et al., 2001; Sylvain et al., 2002) or interference with transcription-activating ability and subcellular localization (Englert et al., 1995; Reddy et al., 1995). These studies underline the importance of NMD in the control of recessive oncoproteins. NMD may thus have a protective effect in heterozygous carriers of tumor-suppressor genes containing NMD-sensitive nonsense mutations from developing cancer.

Similar modulating effects of NMD can explain genotype-phenotype correlations in an increasing number of human disorders such as Robinow syndrome (Patton and Afzal, 2002), Willebrand disease (Schneppenheim et al., 2001), factor X deficiency (Millar et al., 2000) and retinal degeneration (Sung et al., 1991; Rosenfeld et al., 1992). In all of these examples, nonsense mutations located in the last exon of the gene of interest give rise to a dominant disorder while mutations at least 50 bp upstream of the last intron result in a recessive phenotype (reviewed by Holbrook et al., 2004).

However, it has to be noticed that NMD does not always exert a beneficial effect. In some instances, PTCs triggering NMD may result in more severe disease by abrogating the partial function of a mutant protein. A classical example of this type of diseases is represented by Duchenne muscular dystrophy (DMD). Dystrophin is a high-molecular-weight cytoskeletal protein that links the muscular fibres with the connective tissue (Worton, 1992). It is encoded by a gene containing 79 exons and spanning 2.4 megabases in the genome. As a consequence, more than 98% of DMD nonsense mutations result in a PTC capable of triggering NMD (Kerr et al., 2001). These mutations are associated with the most severe phenotype of DMD due to the absence of dystrophin. However, many of the truncated proteins that could be synthesised from the degraded mRNA retain partial or almost complete function as verified by frameshift mutations in similar positions that produce a change in the ORF but not a nonsense codon. In the latter case, patients carrying these mutations show a less severe phenotype called Becker muscular dystrophy (Kerr et al., 2001).

DMD is not the only case where NMD has a negative impact. Further examples are given by nonsense mutations in the CFTR and HEXA genes causing, respectively, Cystic Fibrosis and Tay-Sachs disease with a more severe phenotype when the nonsense mutations can trigger NMD (reviewed by Khajavi et al., 2006; Kuzmiak and Maquat 2006). The increasing understanding of the role of NMD in the clinic has allowed the development of therapies for nonsense-associated diseases. General antibiotics like aminoglycosides or newly and more specific ones like PTC124 promote the incorporation of an amino acid at a nonsense codon by a near-cognate aminoacylated tRNA, thus suppressing NMD. The effectiveness of these drugs is currently investigated although preliminary studies have shown variable results in patients´response (reviewed by Kuzmiak and Maquat, 2006).

1.2.6 Variable NMD as a disease modifier – The need of an assay system to quantify NMD efficiency

It is clear from the previous examples that NMD adds a layer of complexity to the genotype-phenotype relationship; the occurrence of both dominant and recessive mutant alleles in a single gene is made possible by this surveillance system. Yet, there is some evidence that points towards an even more complex role of NMD in disease. Theoretically, the efficiency of the NMD process could be variable. This variability could be thought of at the whole organism, tissue, cellular or even transcript level (Frischmeyer and Dietz, 1999). Experimentally, quantitative differences in NMD efficiency have been verified only in yeast (Kebaara et al., 2003). In that study, it was demonstrated that different strains of *S. cereviseae* degrades the pre-mRNA of CYH2 (an endogenous NMD target) to different extents attributing this phenomenon to differences in NMD efficiency in the tested strains. Furthermore, phenotype analysis of the progeny derived from crossing those strains suggested that the NMD efficiency is controlled pleiotropically in this species.

In humans, NMD variability has not been proved although it has been suggested by some clinical reports. Patients with diverse disease severity carrying identical nonsense mutations but expressing different levels of the resulting truncated protein have been reported for dystrophin and Jarid1c genes (Kerr et al., 2001; Jensen et al., 2005). In another study, it was verified that the nonsense mutated collagen X mRNA is not detectable in cartilage cells from two patients suffering from Schmid metaphyseal chondrodysplasia due to NMD degradation. However, the transcript was detected in lymphoblast and bone cells from the same patients suggesting that NMD efficiency could be tissue-specific (Bateman et al., 2003). In the same direction, when cells from foetuses diagnosed with Roberts syndrome were treated with an inhibitor of translation, the variable levels of nonsense mutant ESCO2 mRNA (responsible for the disorder) suggested the presence of inter-tissue and inter-individual differences in NMD efficiency (Resta et al., 2006).

However, the studies mentioned in the previous paragraph only describe casual correlations between the levels of a theoretical NMD-targeted transcript and the severity of a disease. As NMD is not the only mRNA decay pathway in the cells, pleiotropic effects could account for the reported results. The common limitation of these studies is the lack of a reliable assay to compare quantitative aspects of NMD. Such an assay would help to advance in the study of NMD efficiency variability.

From the clinical perspective, the study of NMD efficiency is important because variable NMD capacity would likely modulate the phenotypic manifestation of disease. In

concrete, it might explain why patients carrying identical nonsense mutations suffer from different degrees of disease severity. Additionaly, differences in NMD efficiency could provide a possible explanation for the variable response observed in human trials of aminoglycoside therapy.

1.2.6.1 A case study for NMD variability: β-thalassemia (IVS1+6 T→C)

In the previous paragraphs, several reasons were given that explain the need of an assay to quantify differences in NMD efficiency. Such an assay system would allow a better understanding of the relation of NMD and NMD-affected diseases. The study of β-thalassemia (IVS1+6 T→C) arises as a further and particular example of the importance of estimating variable NMD efficiency.

β-thalassemia is caused by mutations of the β-globin gene and represents the most common inherited disorder worldwide. In most families, is inherited as a simple monogenic autosomal recessive trait. In these cases, heterozygotes are phenotypically healthy and display minor abnormalities of red cell morphology (thalassemia minor). In contrast, homozygotes usually become symptomatic during the second half of their first year of life and are affected by a most severe transfusion dependent anemia (thalassemia major). However, in about 10-20% of homozygous patients the clinical phenotype is less severe and the anemia does not require (regular) transfusion treatment (thalassemia intermedia). In many of these cases, the clinical variability can be explained either by a high residual activity of the β-globin gene itself (Rosatelli et al., 1989), by the influence of identified genetic modifiers such as co-inherited α-thalassemia (Kulozik et al., 1993), or by co-existing point mutations or deletions that increase the activity of the γ-globin genes (Gilman et al.,1985). However, in many patients, the clinical variability cannot be explained by any of the known genetic modifiers.

A well defined group of such patients is affected by a mutation in position 6 of the splice consensus sequence of intron 1 of the β-globin gene (IVS1+6 T→C). This mutation (particularly common in the Eastern Mediterranean) impairs the efficiency of correct splicing and, as a result, four alternative variants can be produced (Treisman et al., 1983) (Fig. 7).

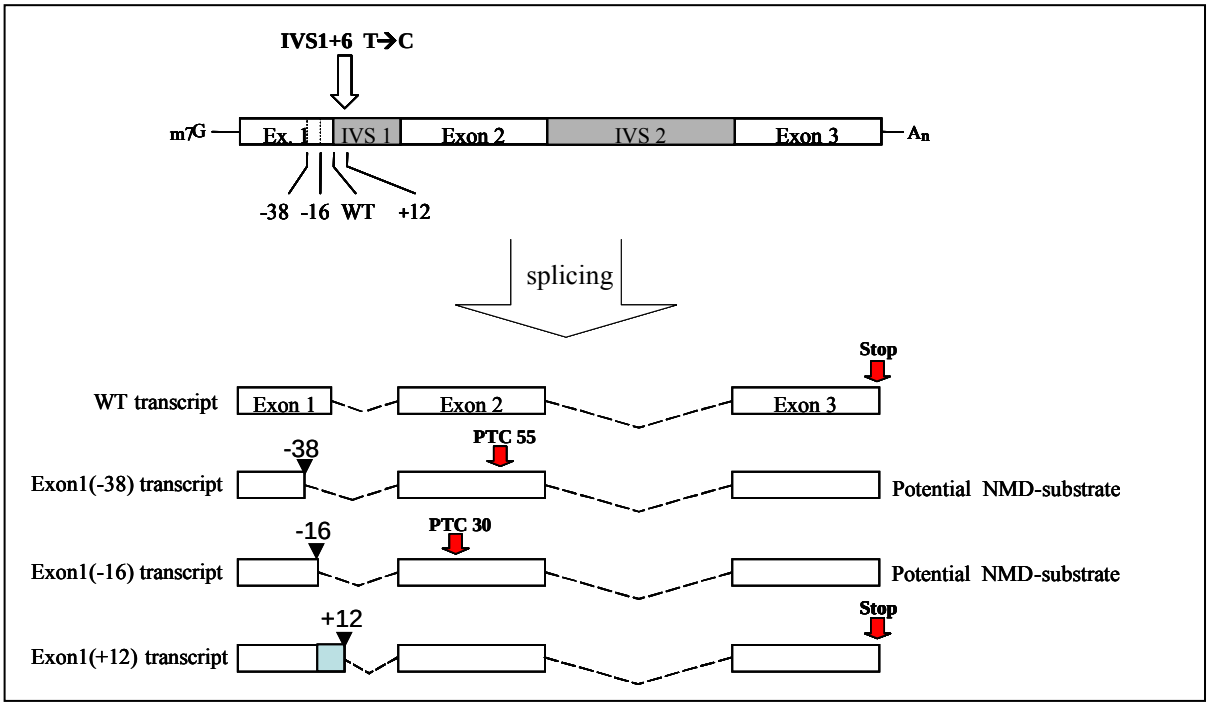

Figure 7. Splicing variants of β-globin (IVS1+6). Four alternative splicing variants have been described as a consequence of the mutation in position 6 of the splice consensus sequence of intron 1 and the use of alternative donor splicing sites (in position -38 and -16 of the end of the first exon, and in position +12 of the first intron). Two of these isoforms (exon1(-38) and exon1(-16)) are theoretical NMD targets due to the presence of a PTC as indicated. The exon1(+12) variant should not be affected by NMD.

The splicing variants arise from the use of 3 alternative splicing donor sites at positions -38 and -16 from the exon/intron boundary and at position +12 of the intron. The exon1(-38) and exon1(-16) variants contain a PTC due to a frameshift. Exon1(+12) maintains the normal ORF and hence, possesses the stop codon in the last exon (Fig. 7). NMD would be expected to down-regulate the exon1(-38) and exon1(-16) isoforms. A previous study from our laboratory (Danckwardt et al., 2002) demonstrated the NMD-sensitivity of the exon1(-16) transcript in HeLa cells; however, in the same study, the exon1(-38) was reported to be resistant to NMD, possibly due to the presence of a putative cis-acting element.

Due to the presence of PTC-containing transcripts, it is reasonable to investigate whether NMD may play a role in the phenotypic expression of this thalassemia. The hypothesis is that NMD efficiency displays differences within individuals and that these differences could account for the diverse severity of the disease. It would then be expected that patients carrying the IVS1+6 mutation and suffering from thalassemia major would have poor NMD activity. As a consequence, the β-globin NMD-sensitive isoform/s would not be down-modulated and truncated β-globin may exert a detrimental effect in the red cells. Conversely, those patients who suffer from thalassemia intermedia would exhibit more robust NMD efficiency. In that situation, truncated polypeptides will not be synthesized due to

degradation of the aberrant NMD-sensitive transcripts, and patients will have a milder phenotype.

It is likely that the characterization of the responsible mechanism for this anemia will not only be relevant for the understanding of the pathophysiology of β-thalassemia but could also have implications for the role of varying NMD capacity in other disorders caused by abnormalities of RNA processing in general or splicing in particular.

1.3 Aim of the project

The main purpose of this study was to test the hypothesis that there exist inter-cell and inter-individual stable differences in human NMD efficiency. To this end, three consecutive goals were persued:

Firstly, I aimed at identifying physiological bona fide NMD transcripts in human cells. This was achieved by abrogating the cellular NMD function by means of RNAi against the essential NMD factor UPF1 and analysing the impact of this depletion on gene expression by genome-wide microarray analysis. Complementary validations were made in order to confirm potential targets as genuine NMD-sensitive transcripts.

Secondly, I aimed at developing an assay system to quantify differences in NMD efficiency among cells and tissues. This assay was based on the relative quantification of the steady-state levels of a group of validated NMD mRNAs, characterised in the previous step. Moreover, I also intended to gain insight into the mechanisms that potentially control variable NMD capacity in the cell.

Finally, I intended to verify the hypothesis that variations in NMD efficiency may account for the phenotypic expression of disease. I investigated a particular β-thalassemia as a case study to test this hypothesis.

2. Materials and Methods

2.1 Materials

2.1.1 Chemicals

All the salts and standard chemicals were purchased from Roth (Karlsruhe, Germany) or Sigma (Steinheim, Germany). Cell culture media were purchased from Gibco (Invitrogen). Exceptions to this statement or particular chemicals are the following:

Agarose	Biozym Scientific (Oldendorf, Germany)
Bacto-agar	Becton, Dickinson and Co. (Sparks, USA)
Bromophenolblue	Serva (Heidelberg. Germany)
$CaCl_2$ (calcium chloride)	ICN Biomedicals (Ohio, USA))
Complete EDTA-free protease inhibitor	Roche (Mannheim, Germany)
DTT (dithiothreitol)	Promega (Madison, USA)
Formaldehyde	Merck (Darmstadt, Germany)
Immersion oil	Leica (Bensheim, Germany)
NaAc (sodium acetate)	Merck (Darmstadt, Germany)
Nucleotides (NTPs) and deoxy-nucleotides (dNTPs)	MBI Fermentas (Burlington, Canada) or Promega (Madison , USA)
Trizol (TriReagent)	MRC (Cincinnati, USA)
Tris-base	Merck (Darmstadt, Germany)

2.1.2 Standard used buffers and media

All solutions were prepared with milli-Q water.

Anode buffer I (Western Blot)	25 mM Tris pH 10.4
Anode buffer II (Western Blot)	25 mM Tris 20% methanol pH 10.4
BBS 2x	50mM N,N-bis(Hydroxyethyl)-2-Aminoethansulfonat 280mM NaCl

	1.5mM Na_2HPO_4 pH 6.96
Blotting buffer 1x	20mM Tris-Cl 150 mM glycin 0.02% (w/v) SDS
Cathode buffer I (Western Blot)	25 mM Tris 40 mM 6-amino-n-hexanoic acid 0.01% SDS
Cathode buffer II (Western Blot)	25 mM Tris 40 mM 6-amino-n-hexanoic acid 0.01% SDS 20% Methanol
Church Buffer	0.5M Na_2HPO_4 1mM EDTA 7% (w/v) SDS pH 7.2
LB agar (autoclaved)	1.8% (w/v) Bacto agar in LB medium
LB medium (autoclaved)	1% (w/v) Bacto tryptone 0.5% (w/v) Bacto yeast extract 170 mM NaCl pH 7.6
Lysis buffer for cytoplasmic protein and RNA isolation	50 mM Tris-HCl, ph 7.2 150 mM NaCl 0.5 % (v/v) NP-40 0.1 % Deoxycholate 5 mM Vanadyl-Ribosyl-complex 1 mM Dithiothreitol 0.5 mM PMSF 1x Complete protease inhibitor (Roche)
MOPS 1x	20 mM 3-(Nmopholino) propanesulfonic acid (MOPS) 8 mM NaAc 1mM EDTA pH 7.0
Total cell lysis buffer (for protein isolation)	50 mM Tris-HCl, ph 7.5 150 mM NaCl 1 mM EDTA 1 % Triton X-100 0.5 % Deoxycholate 0.1% SDS 1x Complete protease inhibitor (Roche)
Resolving gel (8% to 12%) solution	375 mM Tris-Cl, pH 8.8 8 to 12 % (w/v) acrylamide/bisacrylamide 0.1% (w/v) SDS
RNA denaturating buffer	50 % formamide 17.5 % formaldehyde (ph>4.0) 1 x MOPS
RNA loading buffer	50% (v/v) glycerol 1 mM EDTA 0.25% (w/v) BromoPhenol Blue
SDS Sample Buffer 6x	350 mM Tris-Cl, pH 6.8 600 mM DTT

	30 % (v/v) glycerol 10 % (w/v) SDS 0.012 % (w/v) Bromophenolblue
SDS Running Buffer (Laemmli buffer) 5x	1.5 % (w/v) Tris 7.2% (w/v) Glycine 0.5% (w/v) SDS
SSC buffer 20x	3 M NaCl 0.3 M Tri-sodium Citrate pH 7.0
Stacking gel (5%) solution	125 mM Tris-Cl, pH 6.8 5 % (w/v) acrylamide/bisacrylamide 0.1 % (w/v) SDS
TBE	1 M Tris-base 1 M Boric acid 20 mM EDTA
TBS	10 mM Tris-Cl pH 8.0 150 mM NaCl
TBS-T	10 mM Tris-Cl pH 7.6 150 mM NaCl 0.05% (v/v) Tween 20
Total cell lysis buffer (for protein isolation)	50 mM Tris-HCl, ph 7.5 150 mM NaCl 1 mM EDTA 1 % Triton X-100 0.5 % Deoxycholate 0.1% SDS 1x Complete protease inhibitor (Roche)
TE buffer	10 mM Tris-Cl 1 mM EDTA pH 7.5
TSS	85% LB medium 10% PEG (w/v, mw 8000) 5% DMSO 50 mM $MgCl_2$ pH 6.5

2.1.3 Enzymes

Here are only described the enzymes that do not belong to any commercial kit. TAQ polymerase ThermoPol and restriction enzymes were purchased from New England Biolabs (Ipswich, USA) along with their respective buffers. SP6 RNA polymerase and DNase I (RNase free) were purchased from Roche (Mannheim, Germany). RNasin® was purchased from Promega (Madison, USA). The reverse transcriptase (RevertAid H minus M-MuLV) was purchased from MBI Fermentas.

2.1.4 Plasmids

pCI-neo, pGEM (Promega) and pBlueskriptSK II+ (Stratagene) are the basic plasmids used to clone all the constructs used in the present study. Each construct was originated according to the table below. All PCR amplified inserts were validated by sequencing (ABI prism 377 DNA sequencer, Applied Biosystems). The following plasmids were used in the present thesis:

Name	Features
p(β-globin WT)	A 4.4 kb human β-globin gene with a linked SV40 enhancer inserted as a *Not*I fragment into pBlueskriptSK II+. (Thermann et al., 1998).
p(β-globin NS39)	Constructed in the same way as the WT isoforms but containing a nonsense mutation at position 39 of the second exon. This gene was derived from a patient with homozygous β-thalassemia. (Thermann et al., 1998).
p(β-globin exonIII)	Template for generation of radiolabeled probes for detection of β-globin mRNA in Northern Blots (Thermann et al., 1998).
pCIneo-βglobin WT	A 1423 bp β-globin gene fragment extending from the physiological translation initiation codon to the translation termination codon, was subcloned from p(β-globin WT) into the pCIneo vector at the *Xho*I–*Xba*I sites of the polylinker.
pCIneo-βglobin NS39	Idem pCIneo-βglobin WT but derived from the p(β-globin NS39).
pCIneo-βglobin(wt+300+e3)	An elongated human β-globin gene (wt+300+e3) that served as a control for transfection efficiency in all experiments. This was created by the insertion of β-globin exon 3 sequences into the SalI restriction site of β-globin wt+300 (Neu-Yilik et al., 2001).
pGEM(miniβ-globin)	Template for generation of radiolabeled probes for detection of β-globin mRNA in RNase protection assays. Created by subcloning of a ApaI-BamHI fragment from pCIneo-βglobin WT
pCIneo(βglobin-IVS1+6)	Created by site mutagenesis from a pCIneo(βglobin-IVS1+5) construct (Danckwardt et al. 2002)
pCIneo(−38 ΔIVS1)	Used for expression of exon1(-38) β-globin variant (Danckwardt et al. 2002)
pCIneo(−16 ΔIVS1)	Used for expression of exon1(-16) β-globin variant (Danckwardt et al. 2002)
pCIneo-Flag	Created by inserting the Flag-sequence into

	the NheI/XhoI sites of pCI-neo vector (Gehring et al., 2003).
pCIneo-FlagY14	The sequence of Y14 was subcloned from a pCI-λN-Y14 construct into the XhoI site of pCIneo-Flag (Gehring et al., 2003).
pCIneo-FlagRNPS1	RNPS1 sequence was amplified from HeLa cDNA and cloned like Y14 (Gehring et al., 2003).
pCIneo-FlagRNPS1Δ69–121	The truncated version of RNPS1 was generated by mutagenesis using overlapping PCR products. (Gehring et al., 2005).

2.1.5 Antibodies

The antibodies and corresponding dilutions that were used in this study are the following:

Antibody	Origin	Provided by	Working dilution
αBARENTZ	rabbit (polyclonal)	Tomasetto lab.	1:2000
αFlag	mouse (monoclonal)	Sigma-Aldrich	1:10000
αMAGOH	rabbit (polyclonal)	Izaurralde lab.	1:1000
α mouse IgG peroxidase coupled	goat (polyclonal)	Sigma-Aldrich	1:10000
α rabbit IgG peroxidase coupled	goat (polyclonal)	Sigma-Aldrich	1:10000
αRNPS1	rabbit (polyclonal)	Krainer lab.	1:4000
αTUBULIN	mouse (monoclonal)	Sigma-Aldrich	1:10000
αUPF1	rabbit (polyclonal)	EMBL Core facilities	1:5000
αUPF2	rabbit (polyclonal)	Lykke-Andersen lab.	1:5000
αUPF3b	rabbit (polyclonal)	Lykke-Andersen lab.	1:5000
αY14	rabbit (polyclonal)	Izaurralde lab.	1:1000

2.1.6 Bacterial strains

For the propagation of plasmid DNA, the *E. coli* strain XL1-Blue was used. The genotype of this strain is the following: hsdR17, supE44, recA1, endA1, gyrA96, thi, relA1, lac/F´[proAB^{+} lacI^{a}, lacZΔM15::Tn10(Tet^{R})].

2.1.7 Eukaryotic cell lines

In the present study, three strains of HeLa cells (which originate from a human cervical carcinoma) were used. I called them strain A, B and C. Strain A has been used by our laboratory for many years (Enssle et al., 1993; Thermann et al., 1998). Strain B (ACC 57) was

purchased at the German Repository of Cell lines (DSMZ). Strain C was kindly provided by Dr. Elisa Izaurralde (EMBL, Heidelberg).

The EBV-transformed B-cell lines used in the present study were purchased from the Coriell Institute for Medical Research (Camden,USA). The repository numbers corresponding to the tested lines are: GM07734, GM07852, GM07902, GM07904, GM14381, GM14382, GM14405, GM14406, GM14408 and GM14409.

2.1.8 Sequences of oligonucleotides used in PCR reactions

DNA oligonucleotides were ordered from Qiagen (Hilden, Germany) or Biomers (Ulm, Germany). PCR primers were obtained in desalted, lyophilised form and were dissolved in water. In addition, those primers used to amplify pre-mRNA were HPLC-purified. All primers were designed using OLIGO software with a Tm= 60ºC and a product size ranging from 180 to 300 bp, for optimal use in real-time PCR.

Code	Name	Sequence
1	gapdh-up	AAA CCT GCC AAA TAT GAT G
2	gapdh-lo	TAC CAG GAA ATG AGC TTG AC
3	sc35upper	GTG TCC AAG AGG GAA TCC AA
4	sc35lower	CTA CAC AAC TGC GCC TTT TC
5	Sepwupper	CCG AGT CGT TTA TTG TGG
6	Sepwlower	CCA CCA ACT TCA GAA ACT TG
7	phlda1upper	TTC TCC AAC ATG AAG ACC GTG
8	phlda1lower	TGC CGC GTG GAT TTGA C
9	itga6upper	GAT GGG TGG CAA GAT ATA GTT
10	itga6lower	TGC AAT GCC AAA CAT AGA A
11	Glrxupper	ATC AAG CCC ACC TGC CCG T
12	Glrxlower	CGC CTA TAC AAT CTT TAC CAA
13	gfpt2upper	CTC TGT GTC CAG GCT TGC T
14	gfpt2lower	TGG ACT TAT AGG AGG CAT AGG
15	il6upper	CCT CCA GAA CAG ATT TGA GA
16	il6lower	CGC AGA ATG AGA TGA GTT G
17	kcnj12upper	ATG TGG CGT GTG GGT AAC C
18	kcnj12lower	GCT GGC CTC GTC AAT CTC A
19	tnfrsf12aupper	ACA GAA AGG GAG CCT CAC G
20	tnfrsf12alower	CCC GTG GTG GAA TTT CA
21	rai3upper	CCT TCT TGC CTG GGT AGG AG
22	rai3lower	AAC CGG GCT TGT GCT AGT G
23	akr1c1upper	GTC CTG GCC AAG AGC TAC AA
24	akr1c1lower	GCA ATG CCC TCC ATG TTA AT
25	tbl2upper	GCA GTC ATT TAC CAC ATG C
26	tbl2lower	TAT TGT TTC TGC TTC TTG GAT
27	gadd45bupper	GAG TGA GAC TGA CTG CAA GC

28	gadd45blower	TCT TAT TAA TTC GCA AAC TGG
29	Tgfaupper	TTA ATG ACT GCC CAG ATT CCC
30	Tgfalower	CCT GGC AGC AGT GTA TCA GC
31	cdh19upper	GAA GAG GAT GAT TCG CAA ACA
32	cdh19lower	CAG ACA CCA CGC CTA CAA ATG
33	nat9upper	ATT GTG CTG GAT GCC GAG A
34	nat9lower	ACC TAG CGT GGT CAC TCC GTA
35	bcar3upper	ACC CAA ACT GCC ATT CTC TAT G
36	bcar3lower	CTC TCA GCA TTC ATC CGG TA
37	slc16a6upper	CAT ATG TAC GTC GCC ATC G
38	slc16a6lower	CGC TGG TCC TCT GAT AAT GAT
39	tbl2nmd-upper	GCC CGC CTG CCA AAA AG
40	tbl2nmd-lower	GGA GGC GGT GGG TGA AGT TG
41	sc35nmd-upper	GCC TGA AAC TGA AAC CAT
42	sc35nmd-lower	GGC AAA GCT TAA ACA AGT A
43	pre-sepwupper	CCA GAA GGT TCG AGA ATG T
44	pre-sepwlower	ATA CCC TTC CTG AGA CTT GC
45	pre-phlda1upper	CAT CAC GAC CGT CCT TGT
46	pre-phlda1lower	TGG AAA GAA ATG ACA CCG A
47	pre-itga6upper	GCT GAG GTG GTT GGT TGT
48	pre-itga6lower	ACC ATC CTA CCG AAA GAG TT
49	pre-glrxupper	GTT ATC TGC CCC TTG TTC TA
50	pre-glrxlower	CTA TTC GTA GCA AAT GGG AC
51	pre-gfpt2upper	TGA GCC TCA GTT GTC TTA CC
52	pre-gfpt2lower	GAA ACC GCA TCA CTT AGC
53	pre-il6upper	CAT CCT GGG AAA GGT ACT CTC
54	pre-il6lower	GAT GTT CTT CCT GCA CTC TTG
55	pre-kcnj12upper	GAC CAC CCA CCT GTT GAT
56	pre-kcnj12lower	CCA CCT TGT GTG AGA GTT GA
57	pre-tnfrsf12aupper	GGG CAG ACT TGA CAC TAG G
58	pre-tnfrsf12alower	CCT CCC CTC CAA ACT CTC
59	pre-rai3upper	TCT GGG AAG GAC TGC GTA
60	pre-rai3lower	GAC ATG GCC TTG ACT GAC A
61	pre-akr1c1upper	CAC CTT TCC CAG TAA CTT ACA
62	pre-akr1c1lower	GGT AAC ATG GGT CTC CAG TA
63	pre-glrxupper	GTT ATC TGC CCC TTG TTC TA
64	pre-glrxlower	CTA TTC GTA GCA AAT GGG AC
65	pre-tbl2upper	CAT GTC CTT AGT CGC TAT TTC
66	pre-tbl2lower	ATT AGC TGA GTA CAG TGG CA
67	pre-gadd45bupper	GAT GAA TGT GTG AGT CAG ACC
68	pre-gadd45blower	GCA GAC GAT ACA TCA GGA TAC
69	pre-tgfaupper	GAG ACC CGG ACT AGG TAG AA
70	pre-tgfalower	AAG ACA GAG GAG TGA ACG CT
71	pre-cdh19upper	TGT TCC CCT GAT ATG ATA GC
72	pre-cdh19lower	TAG TCT CCT CCA CGA TTC C
73	pre-nat9upper	GAC CCG AGA CAG GAT AGG A
74	pre-nat9lower	AGG CAT ACG TCT GAC ACC TT
75	pre-bcar3upper	ATG TGA CGA AGG CAG TAG TG
76	pre-bcar3lower	AGG GCT GGC TCG AAT AC

77	pre-slc16a6upper	CAG GAG TCT GAC AAA TCG TG
78	pre-slc16a6lower	CAG AGA TGG GAG GAT TGT T
79	pre-sc35upper	GAG CCG CAG GTA AAC G
80	pre-sc35lower	GGT CGC AGA CGG CGG AA
81	fos-upper	GCT GAC TGA TAC ACT CCA AGC G
82	fos-lower	TGA CAG GTT CCA CTG AGG GC
83	rpl32upper	TTG ACA ACA GGG TTC GTA G
84	rpl32lower	TTC TTG GAG GAA ACA TTG TG
85	cbfbupper	GCC CAT CTT TAC ATA CAC A
86	cbfblower	ACT TCA AAT TAT TAC TGG CTA C
87	hprt1upper	GAC CAG TCA ACA GGG GAC AT
88	hprt1lower	AAC ACT TCG TGG GGT CCT TTT C

2.1.9 Sequences of siRNA used in knock-down experiments

The siRNAs were purchased from Qiagen and obtained as dry annealed dsRNA duplexes which were redissolved in water.

Code	Name	Sequence
A	upf1-1	AAGAUGCAGUUCCGCUCCA-UU
B	upf1-2	CAGUCCUGGAGUGCUACAA-UU
C	upf2-1	CAACAGCCCUUCCAGAAUC-UU
D	upf2-2	UUACGUCUUUGACCAAACA-UU
E	Luciferase	CGUACGCGGAAUACUUCGA-UU

2.1.10 Kits

The following kits were purchased from Qiagen:

DNA preparation from *E.coli*	QIAprep (mini, midi and maxi) Kit
DNA extraction from agarose	QIAquick Gel Extraction Kit
RNA purification	RNeasy® Midi Kit
PCR purification	QIAquick® PCR Purification Kit

2.1.11 Instrumental material

Bioanalyzer (2100 Agilent)	Agilent (Waldbronn, Germany)
Agarose gel electrophoresis equipment	Peqlab Biotechnologie (Erlangen, Germany)
Acrylamide gel (SDS-PAGE) electrophoresis	Whatman-Biometra (Goettingen, Germany)

equipment	
Acrylamide gel (with urea) electrophoresis equipment	Owl Separation Systems (Portsmouth,uSA)
Blotting (Trans-Blot SD semi-dry transfer cell)	Bio-Rad (Hercules, USA)
Photometer (Biophotometer)	Eppendorf (Hamburg, Germany)
Real-time thermocycler (LightCycler)	Roche (Freiburg, Germany)
Thermocycler (T3000)	Whatman-Biometra (Goettingen, Germany)
PhospoImager FLA-3000	FujiFilm-LifeScience (Düsseldorf, Germany)

2.2 Methods

2.2.1 Standard methods

All standard methods of molecular biology that are not described in detail in this chapter such as e.g. gel electrophoresis of DNA were performed according to Sambrook et al., 1989.

2.2.2 Bacterial techniques

2.2.2.1 Preparation of competent E.coli for transformation

For the preparation of competent *E.coli* XL1-Blue, bacteria from a frozen stock were streaked onto LB agar and incubated at 37°C over night. A single colony was inoculated into 5 ml LB medium and allowed to proliferate over night at 37°C. One milliliter of the bacterial suspension was transferred to 100 ml LB medium and incubated at 37°C until the suspension reached an OD600 of 0.6. The culture was cooled on ice, and the bacterial cells were pelleted by centrifugation at 1200 g for 5 min in a pre-cooled rotor. The bacterial pellet was resuspended in 10 ml of ice-cold TSS, aliquoted in pre-cooled reaction tubes and stored at -80°C. Aliquots were tested by transformation with serial dilutions of plasmid DNA for transformation efficiency of at least 10^6 colonies per µg of plasmid DNA.

2.2.2.2 Transformation of competent bacteria

Fifty nanograms of plasmid DNA were mixed with 100 µl of competent *E.coli* XL1-blue, incubated on ice for 30 min and heat-shocked for 90 s at 42°C. The suspension was

chilled on ice, 500 µl of LB medium was added after 3 minutes and incubated for 30 min at 37°C with rotation. The transformed bacteria were then plated on LB agar plates containing 100 µg/ml ampicillin and incubated at 37°C over night.

2.2.2.3 Isolation of plasmid DNA from bacteria

Plasmids were propagated in *E.coli* XL1-blue. Single colonies were picked from an Agar plate and incubated in 3 ml LB medium containing 100 µg/ml ampicillin for mini-preparation or 100 ml LB medium for maxi-preparation of plasmids. Plasmid DNA was isolated using Qiagen plasmid purification kits according to the manufacturer's instructions. The DNA concentration was spectrophotometrically measured at 260 nm (A_{260}) in TE buffer. An A_{260} value of 1 corresponds to a double-stranded DNA concentration of 50 µg/ml. The plasmids were also visualised by a BrEt-stained agarose gel.

2.2.3 DNA techniques

2.2.3.1 Restriction digests

Plasmids and DNA fragments generated by PCR were digested with appropriate enzymes (supplied by NEB Biolabs). Optimal reaction buffers, as recommended in the NEB catalogue, were used. Reactions were typically carried out in a total volume of 50 µl, using 1 U of each enzyme per µg DNA. BSA was also added where required. The reaction was allowed to proceed for 3 to 5 hours in an incubator at 37ºC. Restriction fragments were visualised by agarose gel electrophoresis and purified using the QIAGEN Gel Extraction kit (see section 2.2.3.3 below).

2.2.3.2 Ligations

A linearised vector and a gene fragment with compatible ends were ligated using T4 DNA ligase (NEB Biolabs). Generally, 50 ng of linearised plasmid were ligated with different molar rations of insert. Ligations were performed in a 20 µl reaction containing 1 µl T4 DNA ligase (400 U/ml), and 2 µl 10x T4 DNA ligase buffer (100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP, 300 mM Tris-Cl, pH 7.8). Ligation reactions were performed for 2 hours at room temperature.

2.2.3.3 DNA Extraction from Agarose Gels

DNA fragments generated by PCR or by restriction digestion were purified depending on their sizes on 0.8% to 1.5% agarose gels. Subsequently, excised DNA fragments were extracted using the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. DNA was eluted from the silica-based purification column with 30 µl water and stored at -20°C.

2.2.3.4 Standard Polymerase Chain Reaction (PCR)

PCRs were carried out to verify gene expression using cDNA (see 2.2.3.6) as template. A typical reaction contained 2 µl of cDNA, 0.5 µM of the forward and reverse primers, dNTPs (Roth) at a final concentration of 200 µM each, as well as 1.5 units of ThermoPol TAQ Polymerase in a final volume of 50 µl. A usual PCR thermocycle protocol was as follows:

5 min	-	95º C - denaturation	
30 sec	-	95º C - denaturation	
45 sec	-	57-60º C - annealing (depending on primers´ Tm)	30 cycles
x sec	-	72º C - extension (depending on template's length, 1 kb/min)	
5 min	-	72º C - extension	

2.2.3.5 Real-time PCR

Real-time PCR was used to quantify gene expression levels. Purified cDNA (see 2.2.3.6) was used as template. A typical reaction contained 2 µl of cDNA, 4 mM $MgCl_2$, 1µM of each primer and 1X LightCycler® Fast-Start DNA Master SYBR green I (Roche) in a final volume of 20 µl. The real-time PCRs were performed in a LightCycler® machine using capillars provided by the same manufacturer. All the primers listed in 2.1.8 were used for quantification. A typical LightCycler protocol was as follows:

10 min	-	95º C - denaturation	
1 sec	-	95º C - denaturation	
7 sec	-	60º C - annealing	40 cycles for mRNA and 50 cycles for pre-mRNA analysis
14 sec	-	72º C - extension	
10 sec	-	40º C - cooling	

After each run, crossing points (CPs) and melting curves were analysed as indicated by the manufacturer to assess both, sample abundance and quality of the PCR. Amplified fragments were visualised in a BrEt-stained agarose gel to further control the PCR products**.**

2.2.3.6 Reverse Transcription

Total RNA (see 2.2.4.1 and 2.2.4.2 for extraction protocol) was used as a substrate for first strand cDNA synthesis. For pre-mRNA analysis, the RNA was further purified from DNA contamination (as explained in 2.2.4.3). Reverse transcription reactions were performed with 1-5 µg of total RNA and oligo(dT) primer or random hexamers (NEB Biolabs). The RNA was preincubated at 75°C for 10 min with 0.5 µg oligo(dT) primer or random hexamers, then chilled on ice. Reverse transcription was carried out with 200 U Revert Aid^{TM} H Minus M-MuLV Reverse Transcriptase and 1 µM dNTPs in 1x reaction buffer (Fermentas). The reaction mixture was incubated at 42°C for 1 hour. The reaction was stopped by heating at 70°C for 10 min. Following reverse transcription, the cDNA was purified with the Qiagen PCR purification kit according to the manufacturer's instructions.

2.2.3.7 DNA Sequencing

DNA sequencing was performed at the University of Heidelberg with the ABI Prism 377 DNA sequencer (Applied Biosystems). The raw data was analysed with Chromas (Griffith University, Australia). Sequencing reactions were performed with the BigDye Terminator v3.1 Sequencing Kit (Applied Biosystems). The protocol provided by the manufacturer was proportionally scaled-down to a final reaction volume of 10 µl.
The following primers were used for sequencing:

Name	Sequence
SP6	CATACATTTAGGTGACACTATAG
T7	TAATACGACATCACTATAGGGC
T3	CTTATCATGTCTGCTCGAAGC

2.2.3.7 DNA mutagenesis

A modified version of the Invitrogen Gene Tailor Kit was used to mutagenesise a pCIneo(βglobin-IVS1+5) into the pCIneo(βglobin-IVS1+6). 100 ng of the original plasmid was methylated with 4 U DNA methylase (Invitrogen) and 1X SAM (NEB) for 1 hour at 37°C in a final 16 µl reaction volume. 2µl of this reaction were subsequently used for a PCR set as follows:

10X High Fidelity PCR buffer	5 µl
10 mM dNTP	1.5 µl
10 mM $MgSO_4$	1 mM
Primers (10 µM each)	1.5 µl
Methylated DNA	2 µl
Platinum® Taq High Fidelity	0.5 µl (2.5 units)
Water	up to 50 µl

The cycling parameters were specified as follows:

2 min	94 °C	
30 sec	94 °C	┐
30 sec	55°C	├ 20 cycles
8 min	68°C	┘
10 min	68°C	

The primers were designed following the instructions of the Invitrogen Gene Tailor Kit.
Upper_primer: GTG AGG CCC TGG GCA GGT TGG CAT CAA GGT TAC
Lower_primer: CAA CCT GCC CAG GGC CTC ACC ACC AAC TTC

Finally, 2 µl of the PCR reaction were used to transform 100 µl of competent XL1-blue *E.coli* bacteria. Mini-preparations were done from single colonies to obtain plasmid DNA that was sequenced to confirm the mutagenesis.

2.2.4 RNA techniques

2.2.4.1 RNA extraction of cultured cells

Cells growing in six-well plates were washed twice with PBS. For whole cell RNA isolation, 1ml TriReagent (MRC, Cincinnati, USA) was applied per well. For cytoplasmic

RNA isolation, cells were scraped in 300 µl lysis buffer for cytoplasmic protein and RNA isolation (see 2.1.2 for composition). Nuclei and membranes were pelleted at 10000 g for 10 min. Total cytoplasmic RNA was isolated from the supernatant of homogenized cells with 750 µl TriReagent LS (MRC, Cincinnati, OH, USA). For both, whole cell and cytoplasmic RNA isolation, 200 µl of chloroform were mixed with the TriReagent by vigorous shaking. The aqueous and organic phases were separated by centrifugation for 15 min at 4°C, 12000 g. The upper aqueous phase was transferred to a new tube containing 40 µg of glycogen and the RNA was precipitated with 500 µl of isopropanol. After incubation at room temperature for 10 min, the RNA was pelleted by centrifugation at 12000 g for 12 min at 4°C. Subsequently, the pellet was washed with 70% ethanol and the pelleted RNA was resuspended in 11 µl RNase-free water. The concentration of the RNA was measured by determining the optical density of a 1:100 dilution at a wavelength of 260 nm (A_{260}). An A_{260} value of 1 corresponds to an RNA concentration of 40 µg/ml. The RNA was stored at -80°C until further analysis.

2.2.4.2 RNA extraction of blood samples

Pelleted total white cells or buffy coat from 10 ml. blood (see 2.2.5.1) were washed with 2ml pre-chilled PBS and centrifuged at 2500 g for 1 min, at 4 °C. The pellets were homogenised in 500-600 µl TriReagent, incubated at room temperature for 5 min and stored at –80°C O.N. Afterwards, 100 µl chloroform were mixed with the samples. From that point, the same steps used for RNA extraction of cultured cells (2.2.4.1) were followed.

2.2.4.3 DNase treatment of purified RNA

To remove remnants of genomic DNA from the isolated RNA (specially, for quantification of pre-mRNA), 4-10 µg of RNA were treated with 10 U DNaseI RNase-free (Roche) in 1x transcription buffer (40 mM Tris-Cl (pH 7.9), 6 mM $MgCl_2$, 10 mM NaCl, 2mM spermidine) in 40 µl total volume for 30 min, at 37 ºC. Subsequently, the samples were purified using the RNeasy® Midi Kit (Qiagen) following the manufacturer´s instructions. The RNA was recovered in 30 µl RNase-free water and stored at -80 ºC until use.

2.2.4.4 Northern Blotting

2.2.4.4.1 RNA Gel Electrophoresis

2-4 µg total cytoplasmic RNA were dissolved in 16 µl RNA denaturating buffer (see 2.1.2). Samples were denatured for 15 min at 65˚C and immediately chilled on ice. After adding RNA loading buffer, samples were loaded on an agarose/formaldehyde gel (1.3% (w/v) agarose, 10% (v/v) formaldehyde, 1x MOPS). The gels were run O.N at 52 V using a Biorad power supplier.

2.2.4.4.2 Transfer of RNA to nylon membranes

After electrophoresis, the agarose gel was rinsed twice with distilled water for 7 min. The RNA was blotted onto a positively charged nylon membrane (Nytran N) (Whatman Schleicher & Schuell) by upward capillary transfer in 20x SSC buffer (see 2.1.2). The transfer was assembled as shown below (Fig.8).

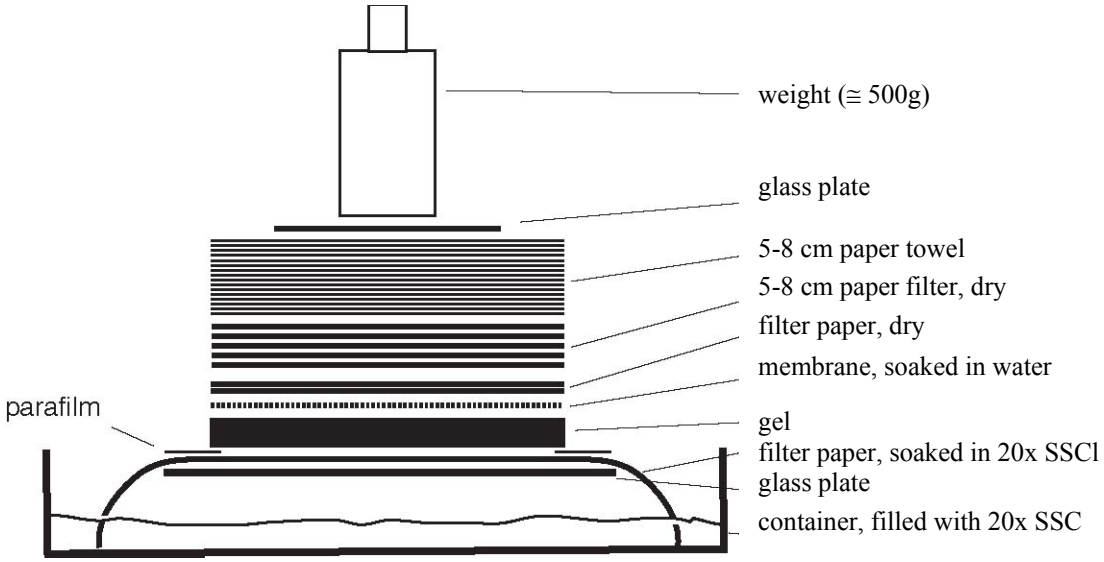

Figure 8. Transfer assembly for Northern-blots. Schematic representation of the assembly for the transfer of RNA to nylon membranes.

After at least 4 hours the RNA was cross-linked to the membrane in a UV Stratalinker apparatus (Stratagene) applying 120 mJ. Afterwards, the membranes were immediately used in Northern hybridisation.

2.2.4.4.3 Riboprobe synthesis

Riboprobes were obtained by in vitro transcription from 1µg of linearized p(β-globin exonIII) plasmid in a 22µl reaction containing 24 U of SP6 RNA-polymerase, 30 U of RNase

inhibitor RNasin, 50 µCi [α^{32}-P]GTP (800 mCi/mmol; NEN), 600 µM of ATP, CTP and UTP and 1x transcription buffer (Biolabs). The transcription was carried out at 40°C for 45 min. Subsequently, the DNA was digested with 10 U DNase I RNase-free (Roche) for 20 min at 37°C. The riboprobe was purified using G-25 Sephadex Quick Spin Columns for radiolabeled RNA purification (Roche) according to the manufacturer's recommendations. The ratio of radiolabelled RNA (flow-though) and non-incorporated nucleotides (column matrix) was estimated with a Geiger-Müller counter.

2.2.4.4.3 Northern Blot hybridisation

Membranes were pre-hybridised in Church buffer at 65°C for 2 hours. The ^{32}P-labelled β-globin riboprobe was added to 15 ml fresh Church buffer and the hybridisation took place O.N. at 65°C in a Biometra OV5 oven, with rotation. Unspecific hybridisation was removed by washing twice with 2x SSC, 0.1% SDS and twice with 0.2x SSC, 0.1% SDS at 65°C (15 min/wash). Radioactive signals were quantified by phosphoimaging in a FLA-3000 fluorescent image analyzer (Raytest, Fujifilm). In this study, the indicated expression levels were calculated after correction for transfection efficiency. For registration purposes, the membranes were then exposed to a Hyperfilm MP (Amersham) for 4 to 16 hours in a cassette with amplifying screen (Amersham) at -80°C.

2.2.4.5 RNase protection assay

The complementary probe for the β-globin mRNA was generated by *in vitro* SP6-driven transcription of pGEM(miniβ-globin) construct previously linearised with BamHI. The ribonuclease protection assay (RPA) was carried out using the RNase Protection Kit (Roche) according to manufacturer´s protocol. In brief, for the RPA, 4µg of total RNA was analysed by hibridisation with 500000 cpm radioactively labelled β-globin riboprobe. Hybridisation was carried out at 60°C over night. Next morning, samples were cooled to room temperature and mixed with ribonuclease T1 (2.5 U) and ribonuclease A (3.5 µg) in RNase digestion buffer. Ribonuclease treatment was carried out for 30 minutes at 30°C. Subsequently, RNases were digested by the addition of proteinase K (50 µg) and 10 µl of 20% SDS (15 min at 37°C). The RNA was isolated by phenol-chloroform extraction. The supernatant was precipitated with ethanol in the presence of 40 µg glycogen for 30 min at -20°C and pelleted by centrifugation at 13,200 rpm for 15 min at 4°C. The pellet was resuspended in 7 µl RNA loading buffer and denatured for 5 min at 95°C. The protected fragments were analysed on a

8% denaturing polyacrylamide urea sequencing gel. The gel was dried on Whatman 2MM paper (Whatman Biometra) and the radioactive bands were quantified by phosphoimaging in a FLA-3000 fluorescent image analyzer (Raytest, Fujifilm) and exposed to an autoradiograph film.

2.2.5 Protein techniques

2.2.5.1 Protein extraction

Cells growing in six-well plates were washed twice with PBS. For whole cell protein isolation, 250 µl of total cell lysis buffer were applied per well; for cytoplasmic protein isolation, 300 µl lysis buffer for cytoplasmic protein and RNA isolation were used (see 2.1.2). Cells were scraped on ice. Subsequently, the samples were centrifuged at 10000 g for 10 min (4 ºC) to pellet the undissolved material. Samples were stored at -80 ºC until use. Protein concentration was determined spectrophotometrically using the Bio-Rad protein assay (Biorad, Germany) and a BSA standard curve as indicated by the manufacturer.

2.2.5.2 Polyacrylamide gel electrophoresis

15-20 µg of proteins were separated according to size by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using a minigel system (BioRad). The samples were mixed with SDS sample buffer (1x final concentration) and heated for 3 min. at 95 ºC before loading. Polyacrylamide gels contained an upper stacking gel of fixed concentration (see 2.1.2) to ensure that all proteins in the sample enter the lower separating gel simultaneously. In the lower separating gel, proteins were resolved according to size. The concentration of acrylamide used was 8-12% resolving depending on the size of the proteins. The composition of the separating gel contains 0.375 M Tris, 0.1% SDS, 0.1% ammonium peroxodisulfate (APS), 0.04% TEMED and the desired concentration of acrylamide using as a stock an acrylamide:bisacrylamide (37.5:1) solution. Electrophoresis was carried out in 1x Laemmli buffer at constant current of 25 mA until the desired resolution was reached as judged by the separation of molecular weight markers (NEB and Fermentas).

2.2.5.3 Western blotting

To transfer proteins from a polyacrylamide gel onto a polyvinylidene fluoride (PVDF) membrane (Westran S, Whatman, Schleicher-Schuell), a minigel wet transfer system (BioRad) was used. The gel was placed on top of the membrane and 3 pieces of Whatman filter paper were soaked in cathode transfer buffer and other 3 pieces were soaked in anode transfer buffer and these pieces were cast with the gel and membrane as show in the following scheme:

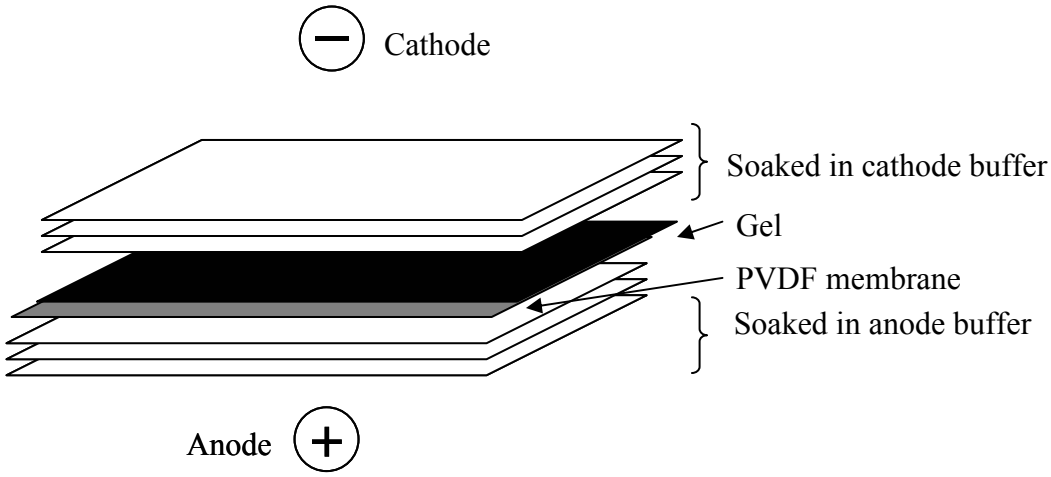

Figure 9. Transfer assembly for Western-blots. Schematic representation of the assembly for the transfer of proteins to a PVDF membrane. See text for details

For Y14 and MAGOH proteins, cathode and anode transfer buffers II (containing methanol; see 2.1.2) were used. For the other proteins, cathode and anode transfer buffers I were used. The blotting was carried out at constant current running at 1 mA per cm^2 for 1 hour.

After the transfer, the membrane was washed in TBS-T and blocked 1 hour in TBS containing 0.1 % Tween-20 and 5 % powder milk. Immunoblotting was performed according to instructions in the Western-Star chemiluminescent immunoblot detection system (Tropix, PE Biosystems). The system makes use of secondary antibodies coupled to alkaline phosphatase (AP). The immobilised AP enzyme dephosphorylates a substrate, which then decomposes, producing a prolonged emission of light that is imaged using photographic film.

For the immunoblot, the membrane was incubated with the primary antibody diluted in TBS containing 0.1 % Tween-20 and 5 % powder milk, on a shaking platform at room temperature for 1 hour or at 4 ºC, ON. The antibodies and dilutions used are listed in 2.1.5.

The membrane was then washed 3 times 10 min in TBS-T and subsequently incubated for 30 to 60 min with a secondary antibody coupled to alkaline phosphatase (Tropix) diluted in the same buffer as the first antibody. After 3 more washing steps with PBS-T, the membrane was developed with ECL Western Blotting Detection Reagent (Amersham Pharmacia Biotech) according to manufacturer´s instructions. Signals were visualized on autoradiograph ECL film (Amersham Pharmacia Biotech).

2.2.6 Cell culture techniques

2.2.6.1 Propagation of human cell lines

HeLa cells (strains A, B and C) were maintained in monolayer in 175 cm² tissue culture flasks (Sarsted, Germany) at 37°C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) (Gibco BRL) supplemented with 10% fetal bovine serum (Biochrom, Berlin, Germany), 100 U/ml penicillin and 100 µg/ml streptomycin (PAA laboratories, Austria). For subculturing, the cells were washed once with 20 ml of PBS and 5 ml Trypsin were added and the cells were allowed to detach at 37°C. 10 ml of cell culture medium was added and the cells were diluted 1:5 – 1:10 with fresh culture medium every 2-3 days.

All the EBV-transformed B-cell lines were maintained in suspension in 75 cm² tissue culture flasks (Sarsted, Germany) in upright position, at 37°C and 5% CO_2 in RPMI 1640 medium supplemented with 2mM L-glutamine and 10% fetal bovine serum. Cells were subcultured diluting the cells 1:3 – 1:4 with fresh medium, keeping the cell concentration at approx. $2*10^5$ cells/ml. The number of viable cells was counted before each passage dying the cells with Trypan Blue (Sigma) in a 1:1 dilution

2.2.6.2 Plasmid Transfection of Eukaryotic Cells

All the transfections carried out in this study were transient. Two different methods have been applied to transfect eukaryotic cell lines with plasmid DNA: the BBS/calciumphosphate method and the lipofectamine method. These methods are described separately.

2.2.6.2.1 Transfection using BBS/calciumphosphate

The 3 strains of HeLa cells used in the present thesis were transfected with this

method, except when they were transfected along with the other cell lines. Cells were subcultured 16 hours before transfection to a density of $2.4*10^5$ cells/ well in a 6 well plate. For each transfection, the DNA mix was supplemented with water to 90 µl and mixed with 10 µl of 2.5 M $CaCl_2$. 100 µl of 2x BBS were added and mixed, and the samples were incubated at room temperature for 15 min to allow formation of the complexes. The mixtures were applied dropwise to the cells. The cells were incubated with the DNA precipitate for 16hs under at 37 ºC and 3% CO_2. After removal of the precipitate, fresh medium was added to the cells and the cells were incubated for another 24h. Subsequently, the cells were harvested as described in 2.2.5.1.

For Northern Blot analysis, a typical transfection would contain 2 µg of β-globin reporter, 0.6 µg control plasmid (pCIneo-βglobin(wt+300+e3)), 0.2 µg GFP-expression vector for visual estimation of transfection efficiency and any expression vector required for the experiments.

2.2.6.2.2 Transfection using Lipofectamine

To compare NMD efficiency in the different cell lines used in this study, the cells were transfected with plasmid DNA using Lipofectamine™ (Invitrogen), a cationic lipid-based transfection reagent. The transfections were performed in 6-well plates (Nunc). 2.5-$3*10^5$ cells per well were seeded one day before transfection, and were transfected when they reached a confluence of 80-90%. For transfection in 6-well plates, a total amount of 5 µg plasmid DNA diluted in 250 µl of serum-free OPTIMEM (Invitrogen). In a separate tube, 10 µl Lipofectamine was diluted in 250 µl serum-free OPTIMEM. The two mixtures were incubated for 5 min at room temperature. Then they were combined, and incubated for another 20 min. During the incubation of the transfection complexes, the cells were washed twice with serum-free DMEM, and 2 ml DMEM (with serum) were added to the cell monolayer. Finally, the transfection mixture was added drop-wise to the cells, resulting in a final transfection volume of 2.5 ml. The cells were incubated for 16-20 hs at 37ºC, 5% CO_2.

2.2.6.3 RNAi Transfection of Eukaryotic Cells

Depletion of endogenous proteins by RNAi was carried out using siRNAs supplied by Dharmacon or Operon as listed in 2.1.9. Transient transfection of siRNAs was carried out using Oligofectamine reagent (Invitrogen) in Opti-MEM I reduced medium (Invitrogen)

without serum and antibiotics. Luciferase siRNAs was used as a negative control. HeLa cells were seeded in six-well plates at a density of $1.4*10^5$ cells per well. After O.N. incubation, the cell culture medium was replaced with 800 µl Opti-MEM I, and siRNA transfection was performed according to the manufacturer's recommendations using 10 µl of siRNAs (20 µM stock) and 3 µl Oligofectamine reagent. After 4 h, the medium was supplemented with 1 ml of DMEM containing 20% FCS. If subsequent plasmid transfection was required, the cells were transfected the following day with plasmid DNA using calciumphosphate precipitation as described in 2.2.6.2. Otherwise, the cells were harvested 72hs after the RNAi transfection.

2.2.6.4 Actinomycin D treatment and mRNA half-life estimation

The growth medium of HeLa cells (strain A) treated with Luciferase or UPF1 siRNA as described in the previous section was replaced 48 hs after the transfection for fresh medium containing 5 µg/ml actinomycin D (Sigma) to stop translation. Cells were harvested and the RNA was collected every hour. Transcript abundance was quantified by quantitative RT-PCR as described in 2.2.3. The log values of RNA concentration were plotted against time (in hours) and interpolation curves were calculated using Microsoft Excel software. The half-life of the *FOS* transcript was used to monitor efficient inhibition of transcription.

2.2.7 Blood samples

2.2.7.1 Whole blood

10 ml of blood were collected in a collection cartridge (S-Monovette, Sarstedt) containing 14 mg EDTA as anticoagulant and processed within 20 minutes. For total blood analysis, the blood was transferred to a Falcon tube containing 40 ml of pre-chilled lysis buffer (155 mM NH_4Cl, 10 mM NH_4HCO_3, 0.1 mM EDTA pH 8). The samples were mixed and incubated on ice for 10 min until the erythrocytes were lysated. The white cells were pelleted at 400 g, 10 min at 4ºC. Afterwards, the pellet was resuspended in 2 ml of lysis buffer, pelleted again at 2300g (1 min, at 4ºC) and washed with 2 ml pre-chilled PBS (PAA). After a new centrifugation step (1min, 2300g, at 4ºC), RNA was purified from the pellet as described in 2.2.4.2.

2.2.7.2 Blood mononuclear cell purification ("buffy coat")

10 ml of blood were collected as in 2.2.5.1 and diluted with 10 ml of PBS (PAA) in a Falcon tube. The sample was overlaid with an equal volume (20 ml) of Biocoll separating solution (Biochrom AG) and centrifuged at 800 x g for 20 minutes at room temperature in a swinging-bucket rotor (without brake). After the centrifugation, the different cells are separated according to their density. An interface enriched in lymphocytes and monocytes (the "buffy coat") appears as depicted in the following scheme.

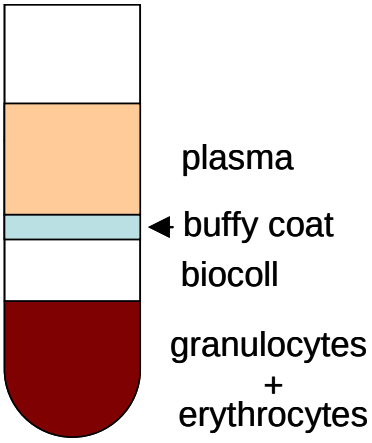

Figure 10. Schematic representation of a biocoll gradient after centrifugation. The layer called buffy coat contains the mononuclear blood cells (lymphocytes and monocytes).

The buffy coat was carefully separated and washed with 30 ml of PBS. The cells were pelleted by centrifugation at 500 x g for 10 minutes at room temperature (with brake). The pelleted cells were washed a second time with 2 ml PBS and centrifuged at 2300 g, 1min, at 4ºC. RNA was purified from the pellet as described in 2.2.4.2.

2.2.8 Microarray techniques

2.2.8.1 Complementary RNA preparation and microarray hybridization

The preparation and processing of labelled and fragmented cRNA targets for microarray hybridization has been performed according to the manufacturer's protocols (Affymetrix Inc.). Briefly, total RNAs were prepared from UPF1 or Luciferase-siRNA transfected HeLa cells (strain A) using the protocol explained in 2.2.4.1 for cytoplasmic RNA. Quality of the RNA was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies

GmbH,Waldbronn,Germany). First and second strand cDNAs were synthesized from 20μg of total RNA using the SuperScriptII double-stranded cDNA synthesis kit (Invitrogen) and oligo-dT24-T7primer. cRNA was synthesized and labelled with biotinylated UTP and CTP by *in vitro* transcription, starting with 0.5 μg of the T7 promoter coupled double stranded cDNA as template and the T7 RNA Transcript Labeling Kit (ENZO Diagnostics Inc.,Farmingdale,NY).

The labelled cRNA was separated from unincorporated ribonucleotides by passing through CHROMASPIN-100 columns (Clontech,PaloAlto,CA),precipitated, and fragmented by a brief alkaline treatment. The fragmented cRNA was hybridized for 16 hs at 45°C to HG_U133A GeneChips™ (Affymetrix, Inc.). Arrays were washed at 25°C and stained with phycoerythrin-conjugated streptavidin (Molecular Probes, Eugene, OR). Arrays were scanned using a laser confocal scanner (Gene Array™; Hewlett Packard, Palo Alto, CA).

2.2.8.2 Statistical microarray data analysis

Expression values for each gene were calculated using the Affymetrix microarray suite (MAS4.0) software. The signal intensity for each gene was calculated as the average intensity difference (AID), represented by Σ(PM-MM) / (number of probe pairs). PM and MM denote perfect match and mismatch probes, respectively. All probe sets that continuously resulted in absent calls were excluded from analyses. The GeneSpring4.2.1 software package (SiliconGenetics, RedwoodCity, CA) was used for normalization, background subtraction and scaling of all genes and arrays. Using the 50% percentile of each chip's intensity range, expression values were normalized across the sample set by scaling the average of the intensities of all genes to a constant target intensity. Relative expression data for each probe set was generated by normalization to itself over the median of the whole experiment set (per gene normalization). Using an initial simple statistical approach without multiple testing conditions (standard Student's *t*-test, $p<0.05$), I identified a group of genes that are differentially expressed by a factor of >2-fold between Luciferase and UPF1-siRNA transfected cells. The normalized data set was also analysed using the GeneCluster2 software package (White-head Institute, MIT, Boston, MA) for graphical purposes.

A comparative functional analysis of altered gene expression based on gene ontology was done using EASE 2.1 software (NIH, USA) described in (Hosack et al. 2003).

3. Results

3.1 Physiological role of NMD in human cells

In recent years, an emerging view of NMD as a key regulator of physiological gene expression has evolved. In yeast, NMD affects directly or indirectly ~10% of the transcriptome (He et al., 2003). The finding that transcripts encoding the splicing factors SC35 and PTB or the ribosomal proteins RPL3 and RPL12 are also regulated by NMD provided evidence for a role of this pathway in normal cellular function in humans, too (Sureau et al., 2001; Wollerton et al., 2004; Cuccurese et al., 2005).

At the starting point of this project, a transcriptome-wide analysis of NMD in humans was still lacking. Such a study would uncover new mRNAs controled by NMD and would also advance our understanding of the role of NMD at a global scale.

The microarray technology provides a way to approach such a comprehensive analysis. In fact, microarray analysis has been successfully applied to study NMD functions in yeast (He et al., 2003) and fly (Rehwinkel et al., 2005).

HeLa cells were treated in three independent experiments with siRNAs against UPF1 (a central factor in the NMD pathway) or Luciferase as a negative control (siRNA B and E, respectively; see Materials and Methods). To positively control for the effect of RNAi the efficient depletion of UPF1 was ascertained by immunoblotting (Fig. 11a). Furthermore, a classical NMD substrate (Gehring et al., 2003), the nonsense mutated β-globin mRNA (NS39) transfected into the siRNA-treated cells is stabilised more than 4 fold (Fig. 11b). Analogously, two NMD-sensitive splice variants of SC35 (referred to as SC35A and B) that represent two endogenous NMD targets (Sureau et al., 2001) were increased after UPF1 depletion (Fig. 11c and d).

RNA isolated from these cells was analysed on Affymetrix HG_U133A GeneChips™. Of 22,283 probe sets, representing approximately 14,500 human genes, 9,336 transcripts were expressed at a level of more than two standard deviations above background (as defined by the GeneSpring™ 4.2.1 software, Silicon Genetics) and were thus included in the analysis. 265 probe sets (2.8%) representing 227 genes were up-modulated more than 2-fold (the first 50 up-modulated transcripts are shown in Table 1), while 248 probe sets (2.6%) representing 202 genes were down-modulated more than 2-fold. These data indicate that a substantial number of genes are affected directly or indirectly by UPF1 activity. While transcripts that increase in abundance following UPF1 depletion are candidate NMD

substrates, endogenous transcripts with decreased expression possibly represent secondary effects.

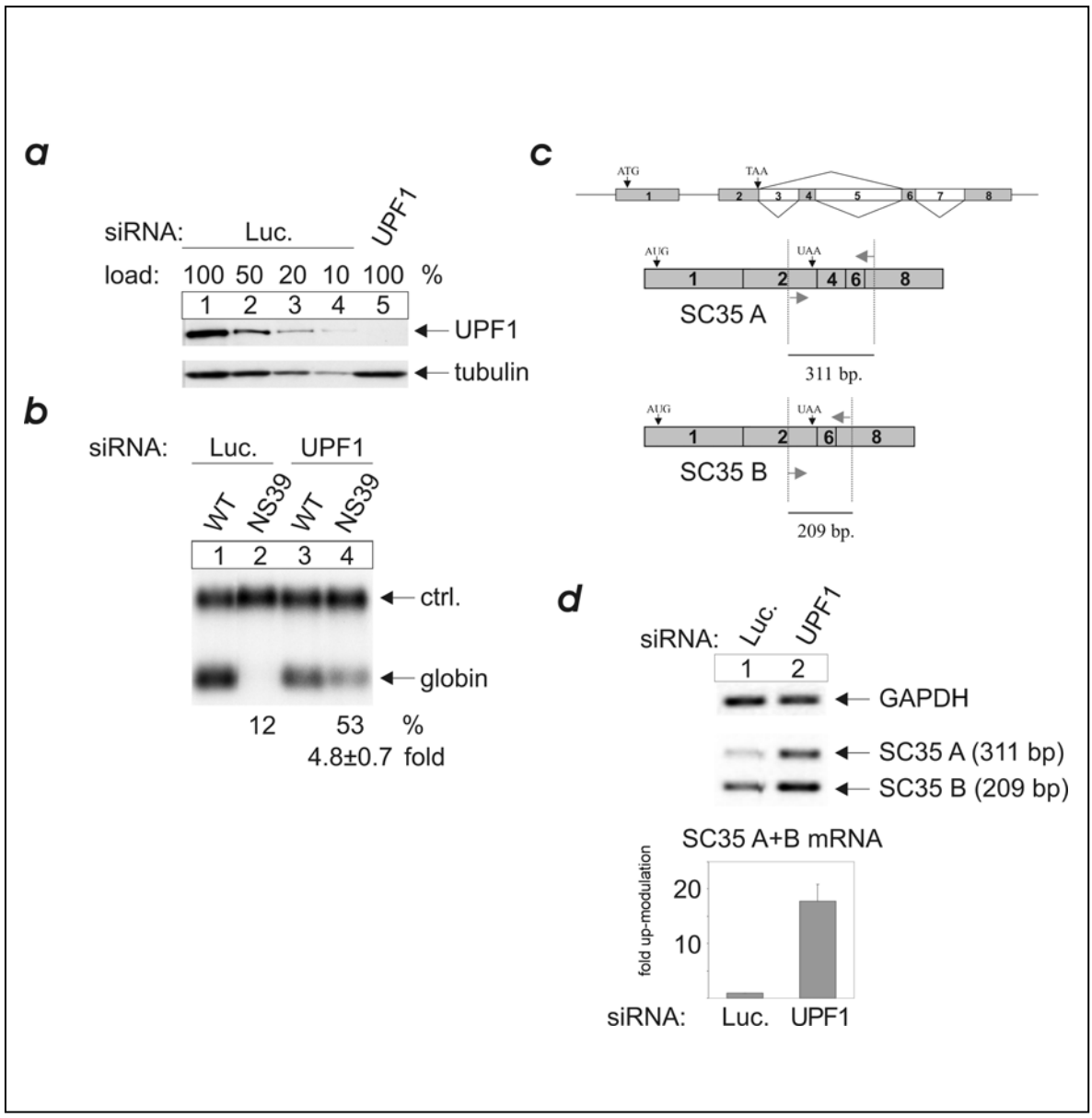

Figure 11. UPF1 depletion up-modulates the abundance of transfected and endogenous NMD reporters. ***a***, Immunoblot analysis of protein lysates from HeLa cells transfected with siRNAs against luciferase as a negative control or UPF1 using a UPF1 specific antibody. Serial dilutions corresponding to 100%, 50%, 20% or 10% (lanes 1-4) of the initial protein amount from luciferase-siRNA transfected cells were loaded to assess the efficiency of the UPF1 siRNA knock-down (lane 5). Reprobing with a tubulin specific antibody was performed as a loading control. ***b***, HeLa cells were transfected with siRNAs against luciferase or UPF1. 30 hours later the cells were co-transfected with the β-globin WT or NS39 reporter constructs and the control plasmid. The indicated percentages correspond to the expression levels of NS39 mRNAs compared to WT mRNAs after normalisation for transfection efficiency (ctrl.). The fold up-regulation represents the ratio of NS39 expression levels between UPF1 and Luc siRNA transfected cells. Values and standard errors were calculated from three independent experiments. ***c***, Gene structure of SC35 (SFRS2). Intron 1 is constitutively spliced, whereas exons 3 to 8 are subject to extensive alternative splicing. The position of the ATG and the stop codon are indicated. The exon-composition of the two SC35 splice-variants, SC35 A and SC35 B that are subjected to NMD are shown below. Positions of RT-PCR primer pairs for both species of SC35 are indicated. ***d***, Quantitative RT-PCR analysis of the NMD-sensitive SC35 (SFRS2) variants in cells transfected with either UPF1 or luciferase siRNAs and normalised against GAPDH expression. The bars indicate the integrated fold up-modulation of the isoforms A and B of SC35. The mean and the standard errors were calculated from 5 independent experiments. The amplified products are shown in a non-quantitative agarose gel.

Gene Symbol	Folds up-modulation	Description	Probe Set ID	Genbank	Putative NMD-inducing feature
SAT	14.14	spermidine/spermine N1-acetyltransferase	210592_s_at	M55580	unknown
PTGS2	12.08	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	204748_at	NM_000963	unknown
AMIGO2	10.78	amphoterin induced gene 2	222108_at	AC004010	uORF
PHLDA1	9.571	pleckstrin homology-like domain, family A, member 1	217996_at	AA576961	stop codon < 50 bases
TNFRSF12A	7.093	tumor necrosis factor receptor superfamily, member 12A	218368_s_at	NM_016639	unknown
DKK1	6.502	dickkopf homolog 1 (Xenopus laevis)	204602_at	NM_012242	unknown
CDKN1A	6.317	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	202284_s_at	NM_000389	uORF
PEA15	6.258	phosphoprotein enriched in astrocytes 15	200788_s_at	NM_003768	uORF
SLC16A6	5.77	solute carrier family 16 (monocarboxylic acid transporters), member 6	207038_at	NM_004694	uORF
GLRX	5.494	glutaredoxin (thioltransferase)	206662_at	NM_002064	stop codon < 50 bases
FN1	5.428	fibronectin 1	212464_s_at	X02761	unknown
ABCA1	5.183	ATP-binding cassette, sub-family A (ABC1), member 1	203505_at	AF285167	unknown
GFPT2	5.152	glutamine-fructose-6-phosphate transaminase 2 (GFPT2)	205100_at	NM_005110	unknown
GADD45A	5.105	growth arrest and DNA-damage-inducible, alpha	203725_at	NM_001924	unknown
ARHE	5.084	ras homolog gene family, member E	212724_at	BG054844	unknown
SERPINE2	4.927	serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	212190_at	AL541302	unknown
GABARAPL1	4.887	GABA(A) receptor-associated protein like 1	211458_s_at	AF180519	unknown
AKR1C1	4.877	aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase)	211653_x_at	M33376	uORF
BCAR3	4.797	breast cancer anti-estrogen resistance 3	204032_at	NM_003567	uORF
RAI3	4.794	retinoic acid induced 3	212444_at	AA156240	uORF
KCNJ12	4.782	potassium inwardly-rectifying channel, subfamily J, member 12	207110_at	NM_021012	unknown
TFPI2	4.599	tissue factor pathway inhibitor 2	209277_at	AL574096	unknown
ITGA6	4.485	integrin, alpha 6	201656_at	NM_000210	stop codon < 50 bases
RIOK3	4.144	RIO kinase 3 (yeast)	202131_s_at	NM_003831	alternative splicing creates PTC
PLAU	4.135	plasminogen activator, urokinase	211668_s_at	K03226	unknown

Table 1. List of up-modulated transcripts in UPF1-depleted cells according to microarray data. The 50 transcripts showing the highest up-modulation levels are shown in this table. The Gene Symbol corresponds to the Hugo name when it has been assigned. The probe set ID corresponds to the identification numbers for the Affymetrix HG_U133A. Putative features that may explain the NMD-sensitivity of the transcripts are described. Because some *bona fide* NMD transcripts do no obey the "50bp rule", transcripts with stop codons at a shorter distance than 50 bases are also indicated.

Gene Symbol	Folds up-modulation	Description	Probe Set ID	Genbank	Putative NMD-inducing feature
DNAJB9	4.109	DnaJ (Hsp40) homolog, subfamily B, member 9	202842_s_at	AL080081	unknown
ATF3	4.098	activating transcription factor 3	202672_s_at	NM_001674	alternative splicing creates PTC
SLC20A1	3.998	solute carrier family 20 (phosphate transporter), member 1	201920_at	NM_005415	uORF
OSMR	3.954	oncostatin M receptor	205729_at	NM_003999	unknown
C1QDC1	3.884	C1q domain containing 1	218456_at	NM_023925	alternative splicing creates PTC
IL6	3.85	interleukin 6 (interferon, beta 2)	205207_at	NM_000600	unknown
NAT9	3.764	N-acetyltransferase 9	204382_at	NM_015654	uORF
JAG1	3.639	jagged 1 (Alagille syndrome)	216268_s_at	U77914	unknown
EPAS1	3.577	endothelial PAS domain protein 1	200878_at	AF052094	uORF
TNFSF9	3.472	tumor necrosis factor (ligand) superfamily, member 9	206907_at	NM_003811	unknown
STC2	3.414	stanniocalcin 2	203439_s_at	BC000658	uORF
LRRC17	3.392	leucine rich repeat containing 17	205381_at	NM_005824	stop codon < 50 bases
ARHGEF2	3.359	rho/rac guanine nucleotide exchange factor (GEF) 2	207629_s_at	NM_004723	unknown
KIAA1245	3.323	chomosome one amplified sequence 1 cyclophilin	214693_x_at	BE732345	unknown
ATF4	3.266	activating transcription factor 4 (tax-responsive enhancer element B67)	200779_at	NM_001675	uORF
C1orf27	3.203	chromosome 1 open reading frame 27	218721_s_at	NM_017847	uORF
AIM1	3.186	absent in melanoma 1	212543_at	U83115	unknown
CARS	3.156	cysteinyl-tRNA synthetase	212971_at	AI769685	unknown
TGFA	3.143	transforming growth factor, alpha	205016_at	NM_003236	unknown
TIMP1	3.119	tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	201666_at	NM_003254	uORF

Table 1 (cont.). List of up-modulated transcripts in UPF1-depleted cells according to microarray data.

Structural analysis of the up-modulated targets revealed that approximately 40% of these transcripts possesses a stop codon regarded as a PTC due to the presence of an intron in the 3´UTR, alternative splicing or an uORF (Table 1). One additional transcript encoding a selenoprotein was also up-regulated, consistent with the alternative recognition of the UGA selenocysteine codon as a signal for translation termination. The remaining ~60% of the transcripts have no classical NMD feature according to database information.

The chromosomal position of genes encoding the putative NMD transcripts was also examined. The results show that the number of targeted genes per chromosome is approximately proportional to the total number of genes per chromosome as would be expected by a random distribution (Table 2). A χ^2-test comparing the expected versus the actual distribution of these genes confirms that there is not significant bias toward the regulation of genes in any specific chromosome (P-value>0.05).

Chromosome	Number of hits	Total number of genes
1	25	2782
2	18	1888
3	9	1469
4	6	1154
5	12	1268
6	12	1505
7	21	1452
8	1	984
9	7	1148
10	11	1106
11	11	1848
12	16	1370
13	5	551
14	6	1275
15	6	945
16	10	1109
17	18	1469
18	4	432
19	10	1695
20	8	737
21	1	352
22	4	742
X	6	1336

Table 2. Distribution of targeted genes in the chromosomes. Per each chromosome the number of hits and the total number of genes (according to NCBI database) are indicated.

To understand the cellular functions affected by NMD, a comparative functional analysis of altered gene expression based on gene ontology (Ashburner et al., 2000) demonstrated that genes which are functionally involved in amino acid metabolism, amino acid transport, organic transport, apoptosis and cell adhesion are significantly overrepresented

amongst the up-modulated genes (Fig. 12). However, transcripts involved in a wide variety of functions were also affected including (but not limited to) enzymes with oxidoreductase activity and kinases, interleukin biosynthesis, DNA repair, Golgi to plasma membrane transport and carbohydrate metabolism. NMD thus emerges as a posttranscriptional mechanism that, in addition to its proofreading function, likely contributes to the regulation of physiological transcripts and important pathways in man.

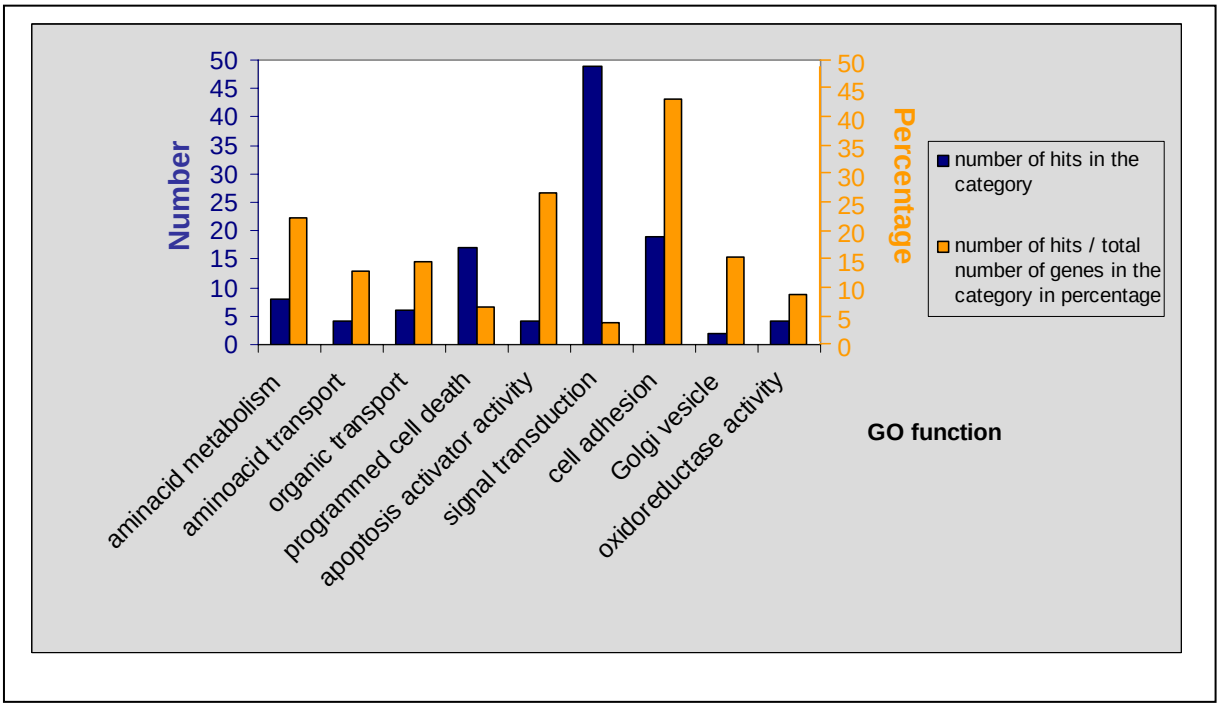

Figure 12. List of significantly over-represented gene ontology categories for the up-modulated transcripts ($p<0.05$). The number of transcripts found in each category (blue bars) and the percentage this number represents in the total number of genes in the category (white bars) are indicated.

In yeast, a similar analysis showed over-representation of only two functional categories: cell rescue and defense and transport facilitation by NMD-regulated mRNAs (He et al., 2003). In addition, genes involved in telomere maintenance, thiamine biosynthesis, pre-mRNA splicing, peroxisomal function, nitrogen metabolism, and in DNA repair were found in the same study. A comparison of the GO categories found in yeast and humans reveals that while there are some common regulated functions to both organisms (such as transport facilitation or DNA repair), most of the functions affected by NMD are different.

Taken these results together, NMD affects directly or indirectly the expression of genes involved in a wide spectrum of functions. This regulation is not tightened to chromosomal location and is not – in general – conserved between yeast and man.

3.2 Characterisation of *bona fide* endogenous NMD targets – A key step towards a quantitation human NMD assay

The main goal of the present study is to develop an assay to estimate differences in NMD efficiency among cells and organisms. Such assay would be based on the quantitation of the steady-state levels of endogenous *bona fide* NMD transcripts. Relatively high abundance of NMD target mRNAs likely correlates with comparatively poor NMD efficiency, provided that this correlation holds for several mRNAs involved in distinct metabolic pathways. The use of unrelated NMD target transcripts makes improbable a common expression control of these targets by an alternative mechanism to NMD. Consequently, the following step in the present study was to define and characterise a group of *bona fide* NMD transcripts that would integrate the before mentioned assay and woud be used to estimate NMD efficiency in HeLa cells and human blood (see 3.3 and 3.4). For this reason, I selected an initial set of 16 mRNAs that were up-modulated by UPF1 RNAi in the microarrays and that are expressed in both HeLa cells and blood. Many of the selected targets presented a strong differential expression in the microarray analysis; however, some other targets that showed lower levels of up-modulation were also randomly chosen. Interestingly, half of the targets do not contain any classical feature that explains their NMD-sensitivity (Fig. 13).

Two considerations have to be addressed when interpreting the microarray results. The first issue is the potential problem of off-target effects. This effect is due to unexpected complementarity of the UPF1-siRNA to other transcripts. Thus, some of the affected mRNAs in the microarray data may not be regulated by UPF1 but by off-targets. To validate my set of up-modulated transcripts, HeLa cells were treated with one of two sequence-unrelated siRNAs against UPF1 (siRNAs A or B, see Materials and Methods) or against Luciferase as a negative control. Again, the depletion of UPF1 was determined by immunoblotting (Fig. 14a). In two independent experiments, the siRNAs B showed a stronger effect than siRNAs A in the reduction of UPF1 protein (~10% vs. less than 5% remaining protein, respectively). This result has been consistently reproduced in our laboratory and is possibly attributed to a different intrinsic efficiency of these siRNAs. I determined the up-modulation levels for the 16 chosen transcripts by quantitative RT-PCR in RNA isolated from these cells (Fig. 14b).

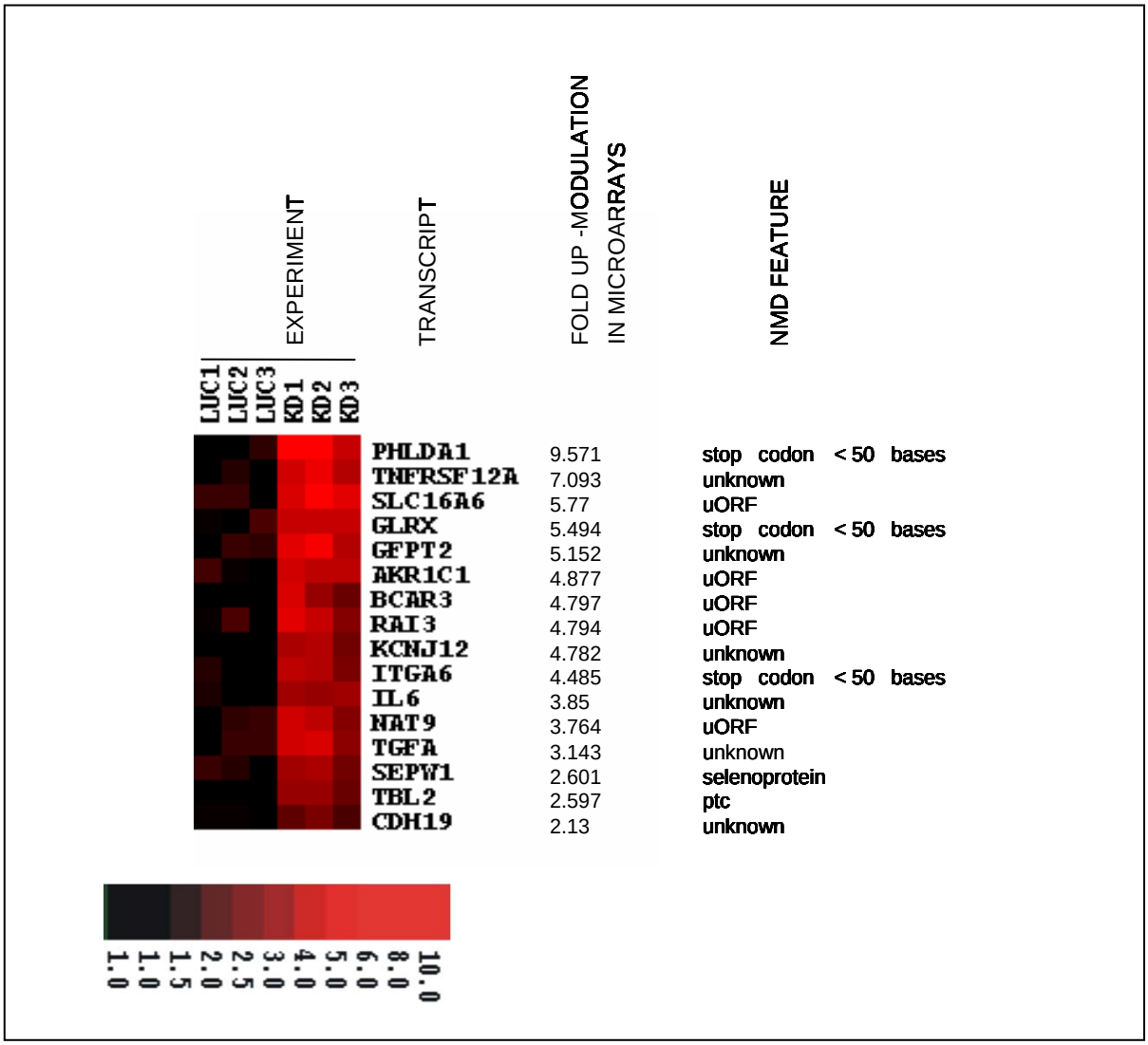

Figure 13. Differential expression analysis of 16 selected transcripts that were up-regulated by UPF1 RNAi. *TreeView* program was used to illustrate the normalized expression levels of 16 selected transcripts in 3 independent experiments where cells were treated with UPF1-siRNA (KD1,KD2 and KD3) or luciferase-siRNA (LUC1, LUC2 and LUC3). The color bar indicates relative values for normalized expression intensity in the samples (red= high expression, black= low expression). In horizontal labelling, the HUGO gene symbol, average up-modulation level and structural NMD features are given for each transcript.

It was possible to verify that for many transcripts the up-modulation levels tended to be higher in cells treated with siRNA B, consistent with the lower remaining UPF1 levels observed using these siRNAs. Nevertheless, taken a threshold level of 2 fold (see Fig. 14) as a limit of significant up-modulation, the results indicate that the 16 transcripts were up-modulated when cells were treated with siRNA B (the same oligo used for the microarray analysis) and 15 out of these 16 mRNAs were also affected using siRNA A. Because the sequences of the two siRNAs against UPF1 are completely unrelated, these results indicate that only TGFA is a possible off-target in the initial panel of 16 transcripts. As a fraction of the panel was randomly chosen (with the only restriction of confirmed expression in HeLa cells and blood cells) an extrapolation of these results would suggest that most of these targets in the microarray are not due to off-target effects.

The second consideration is the potential presence of indirect targets in the microarray analysis. Abrogation of NMD by UPF1-depletion likely affects legitimate NMD transcripts that -in turn- affect other non-NMD mRNAs. In fact, the microarray data showed that the numbers of down-regulated and up-modulated transcripts were comparable. Indirect effects are a plausible explanation for this observation.

NMD only affects spliced, translated mRNA. Thus, the pre-mRNA levels of direct targets are expected to remain constant when comparing Upf1- and Luciferase-siRNA transfected cells. I used this strategy in order to exclude transcripts that are indirectly affected by UPF1 depletion. In several independent experiments performed on UPF1-depleted HeLa cells that showed efficiently inhibited NMD function (see Fig. 11), pre-mRNA and mRNA levels for the selected 16 transcripts were quantified by RT-PCR (Fig. 15).

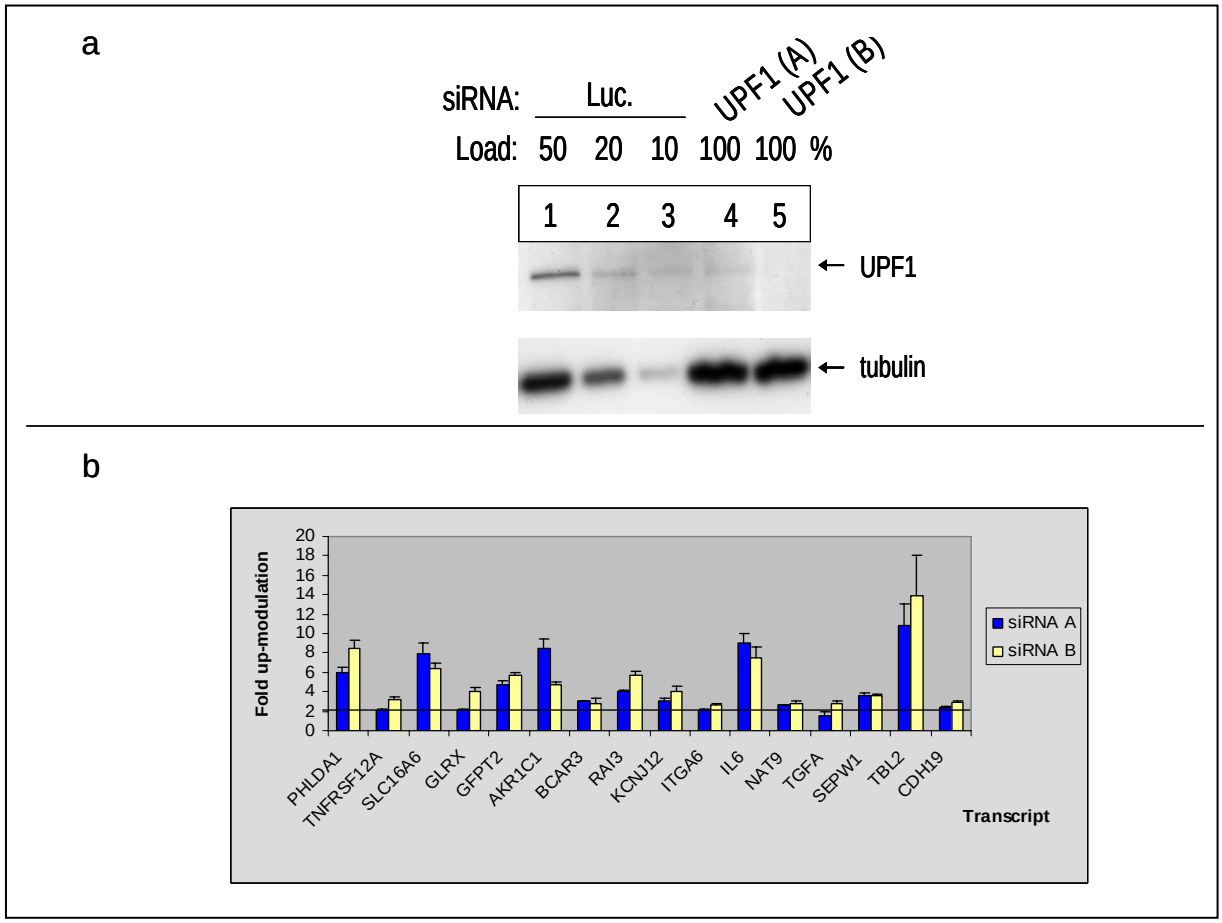

Figure 14. Analysis of potential off-target effects in 16 selected transcripts from the microarray analysis. ***a***, Representative immunoblot analysis of protein lysates from HeLa cells transfected with siRNAs against luciferase as a negative control or UPF1 using a UPF1 specific antibody. Serial dilutions corresponding to 50%, 20% or 10% (lanes 1-3) of the initial protein amount from luciferase-siRNA transfected cells were loaded to assess the efficiency of the UPF1 depletion in cells treated with either siRNA A (lane 4) or siRNA B (lane 5). Reprobing with a tubulin specific antibody was performed as a loading control. ***b***, Quantitative RT-PCR analysis of the 16 selected transcripts in cells transfected with either siRNAs A, B or luciferase and normalised against GAPDH expression. The bars indicate the mean fold up-modulation values for siRNA A- or siRNA B-treated cells as indicated in the figure. A 2-fold up-modulation value (indicate by a horizontal line) is considered the threshold of significance due to reproducibility of the PCR method (this value also coincides with the microarray threshold). The mean and the standard errors were calculated from 2 independent experiments.

While up-modulation of all 16 transcripts was confirmed, only in the case of TBL2 the abundance of the pre-mRNA remained unchanged while the abundance of the mRNA was up-modulated approximately 8-fold. In the case of NAT9 these differences were marginal. In all other 14 mRNAs, the abundance of the pre-mRNA and the mRNA did not differ significantly, although in two (KCNJ12, SEPW1) the pre-mRNA remained below the threshold of 2-fold up-regulation, whereas the mRNA was up-regulated to a level of > 2-fold. These data strongly suggest that most of these mRNAs are likely up-modulated transcriptionally and do not represent *bona fide* NMD targets. By implication, these data also suggest that a substantial fraction, likely most of the almost 230 transcripts that are up-modulated by UPF1 depletion in the microarray data represent indirect targets.

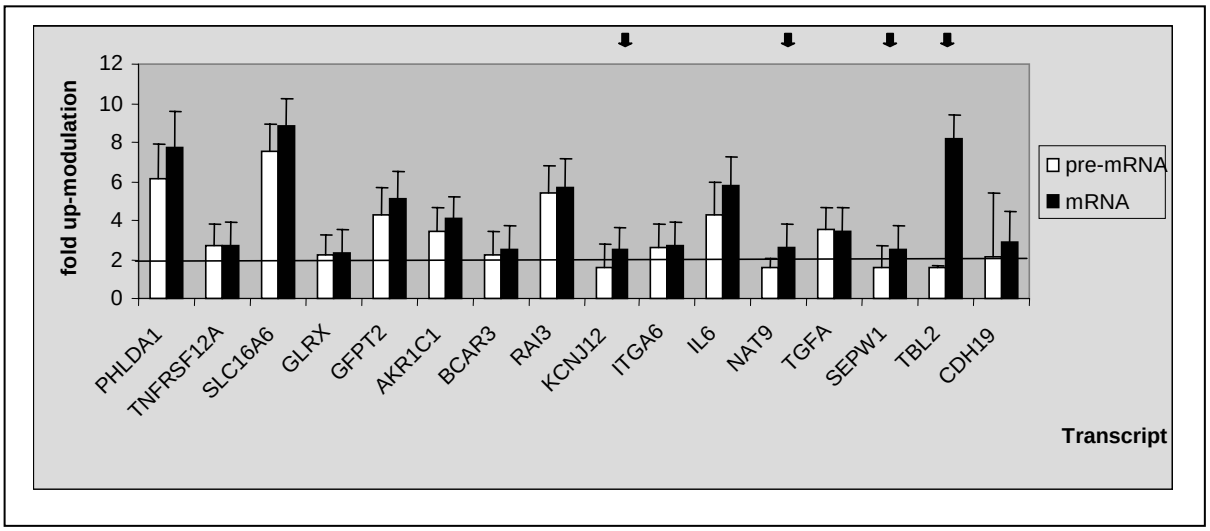

Figure 15. Pre-mRNA and mRNA analysis distinguishes potentially direct from indirect NMD targets in UPF1-depleted cells. Quantitative RT-PCR for 16 UPF1-sensitive transcripts from cells transfected with luciferase (negative control) or UPF1 siRNAs. The abundance of *GAPDH* was used for normalization. The fold up-modulation of pre-mRNAs and mRNAs by UPF1 depletion (mean +/- SE) were calculated from 5-7 independent GAPDH normalized and luciferase controlled experiments. Potential direct NMD targets (arrows) were defined as those mRNAs with a mean up-modulation > 2-fold and with a mean pre-mRNA up-modulation < 2-fold.

Transcripts that are targeted by NMD are expected to be stabilised by an inhibition of this pathway. I thus analysed the decay rates of the *KCNJ12*, *NAT9*, *SEPW1* and *TBL2* mRNAs. At this point, the *GADD45B* transcript was also included, which had previously been suggested to represent an endogenous NMD target by in-silico analysis (Hillman et al., 2004) and is experimentally shown to be upmodulated by UPF1-depletion here (see below). Actinomycin D was added to cells pre-treated with siRNA against UPF1 or Luciferase. The short-lived *FOS* transcript was used as a positive control to assess the block of transcription (Fig. 16a). Prolonged half-lives in UPF1-depleted cells were detected for GADD45B, TBL2 and NAT9 confirming that UPF1 depletion increases the abundance of these transcripts by

reducing degradation (Fig. 16b-d). The stability of SEPW1 and KCNJ12 did not show any effect on UPF1 depletion (Fig. 16e-f). These transcripts were thus excluded from further analysis.

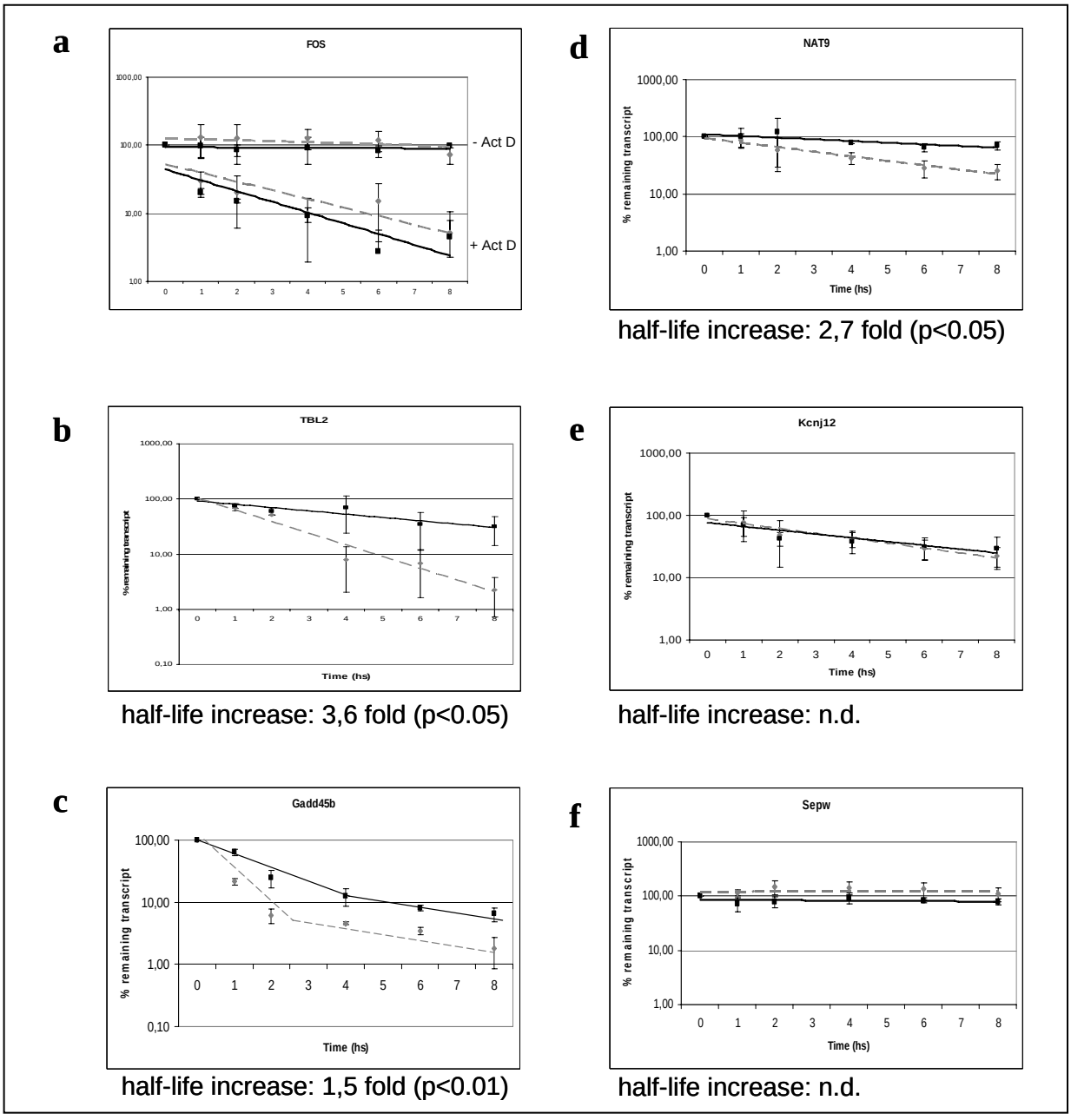

Figure 16. UPF1 depletion prolongs the half-lives of endogenous NMD targets. Decay rates of endogenous transcripts were measured in HeLa cells that were transfected with Upf1 siRNA (solid line) or Luciferase siRNA (dashed line) as control. 48 hours later, the cells were treated with actinomycin D (5ug/ml). Samples were taken every hour. mRNA levels were determined by RT-PCR quantification. The results represent the mean and standard deviation of three independent experiments. **a:** The positive control FOS mRNA is stable in the absence of actinomycin D (- act.D) but it decays rapidly following a block of transcription (+act.D). **b-f:** mRNA decay of the selected transcripts. The increase of the transcript´s half-life is indicated below the corresponding plot.

The role of NMD in directly modulating the abundance of the *TBL2*, *NAT9* and *GADD45B* transcripts was further analysed by depleting UPF2, which interacts with UPF1 in the NMD pathway (Singh and Lykke-Andersen, 2003; Maquat, 2004; Kashima et al., 2006). The efficient depletion of UPF2 to less than 10% was confirmed by immunoblotting (Fig. 17a) and, as a functional control, I assessed the abundance of SC35(A) and SC35(B) isoforms. The degree of up-modulation of the *SC35* and *TBL2* transcripts was stronger in UPF1-depleted than in UPF2-depleted cells while the effects were similar for NAT9 and GADD45B (Fig. 17b). Taken together, these results indicate that *SC35* (A + B isoforms) *TBL2*, *NAT9* and *GADD45B* are *bona fide* NMD targets that depend on both, UPF1 and UPF2.

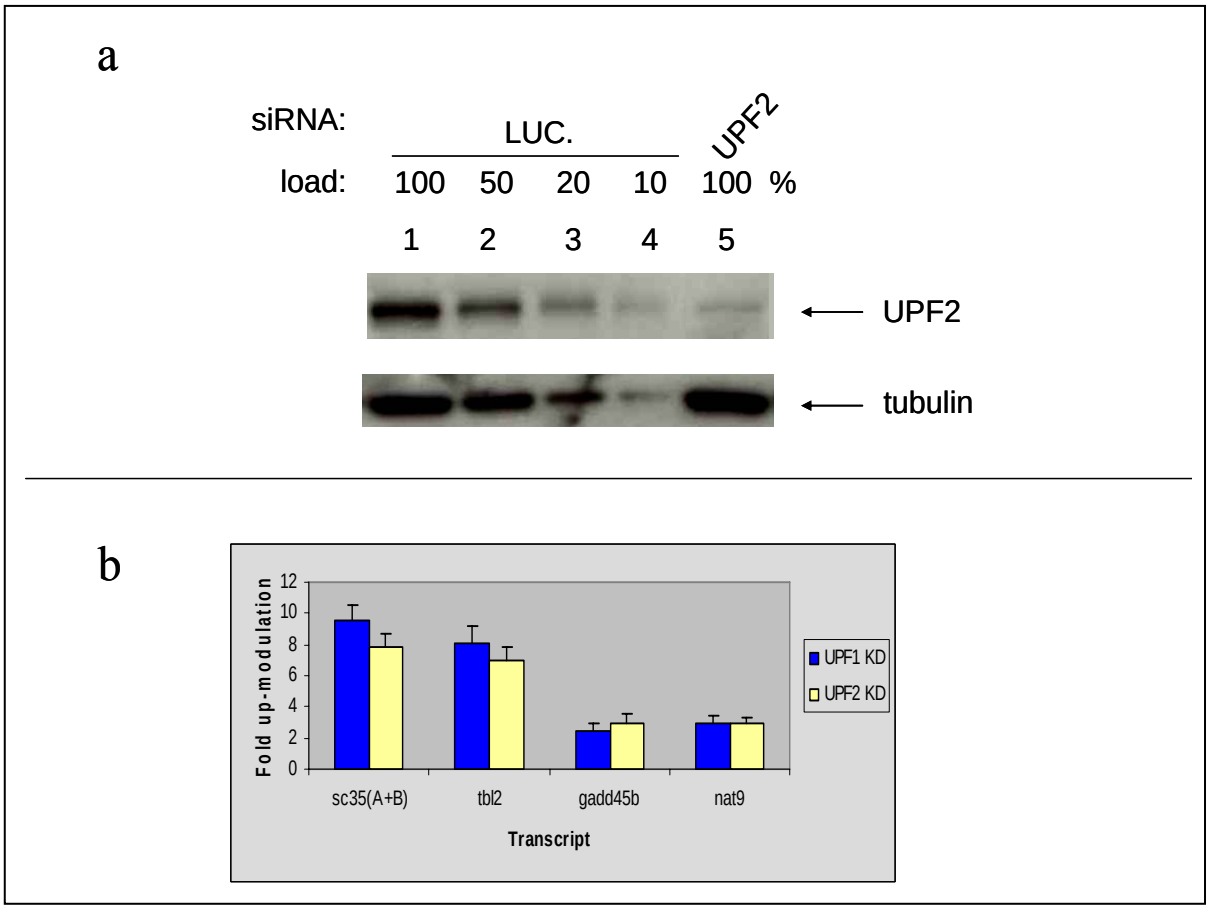

Figure 17. UPF1 and UPF2 depletion cause similar degrees of up-modulation of NMD substrates. **a:** Immunoblot analysis of protein lysates from HeLa cells transfected with siRNAs against luciferase as a negative control or UPF2 using a UPF2 specific antibody. Serial dilutions corresponding to 100%, 50%, 20% or 10% (lanes 1-4) of the initial protein amount from luciferase-siRNA transfected cells were loaded to assess the efficiency of the UPF2 siRNA depletion (lane 5). Reprobing with a tubulin specific antibody was performed as a loading control. **b:** Quantitative RT-PCR analysis of SC35 NMD-sensitive variants, TBL2, GADD45B and NAT9 in cells transfected with UPF1 or UPF2 siRNAs. The UPF1 and UPF2 siRNA treatments were controlled by luciferase siRNA and normalized against GAPDH. Mean and standard errors were calculated from 3 independent experiments.

Analysis of the structure of these validated transcripts using sequence databases show that TBL2 and GADD45B possess a premature termination codon located more than 55 bases from the last exon-exon junction, while NAT9 contains an upstream open reading frame (Fig. 18). These structural features are typical for cellular NMD targets (Nagy and Maquat, 1998; Mendell et al., 2004), which may explain the sensitivity of these endogenous mRNAs to cellular NMD activity.

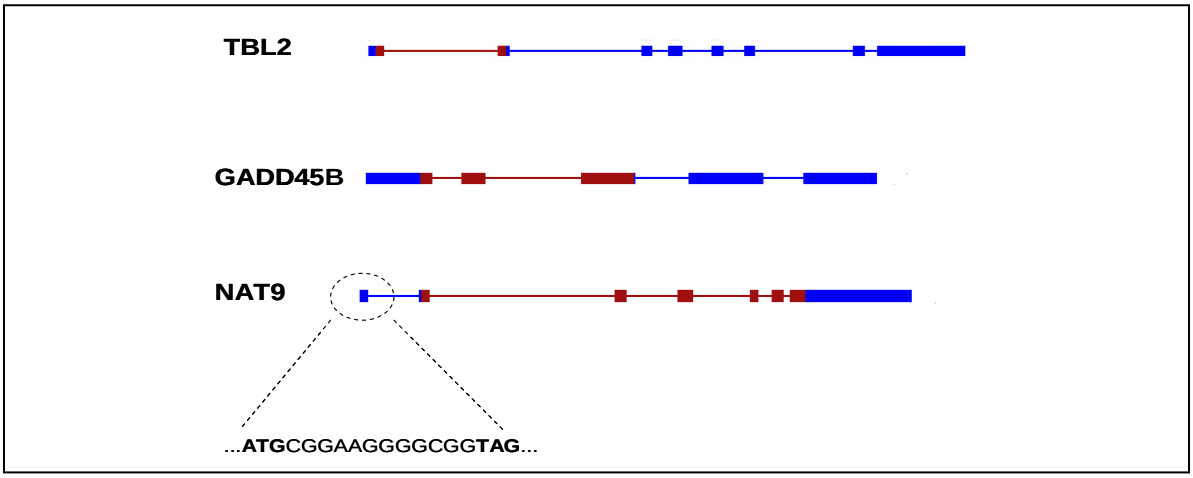

Figure 18. Structure of the NMD sensitive isoforms for TBL2, GADD45B and NAT9. The scheme shows the coding regions in red and the untranslated regions in blue according to the NCBI and ENSEMBL databases. NAT9 does not contain a PTC in the 3´UTR but has an uORF in the first exon which sequence is shown. The transcripts were not depicted in the same scale.

3.3 Quantitative differences in human NMD: a HeLa cell model

3.3.1 Different HeLa strains display variations in NMD efficiency

I next analysed if a panel of the five validated endogenous NMD transcripts consisting of SC35 (A+B isoforms), TBL2, NAT9 and GADD45B can be used to estimate differences of NMD efficiency between cells in a simple model. Unpublished observations in our laboratory have previously suggested that different human cell lines may differ in their NMD capacity.

To estimate NMD efficiency in different cell lines I developed an assay based on the quantification of the 5 transcripts that had been validated to be controlled by NMD against appropriate standard controls. To avoid the potential bias of quantification against one single housekeeping gene in different cell lines, 4 different transcripts (HPRT1, CBFB, GAPDH and RPL32) were selected for normalisation purposes (Jin et al., 2004; Zhang et al., 2005). This group of control transcripts was selected because (1) they showed less than 10% variability in all of my microarray experiments; (2) they were expressed at different steady-state levels, and

(3) they belong to different metabolic pathways and are thus unlikely to be co-regulated. The comparison of the degree of up-modulation following UPF1 depletion showed similar results for all transcripts that were used for normalization (Fig. 19), which indicated that all of these housekeeping genes can be used as standards.

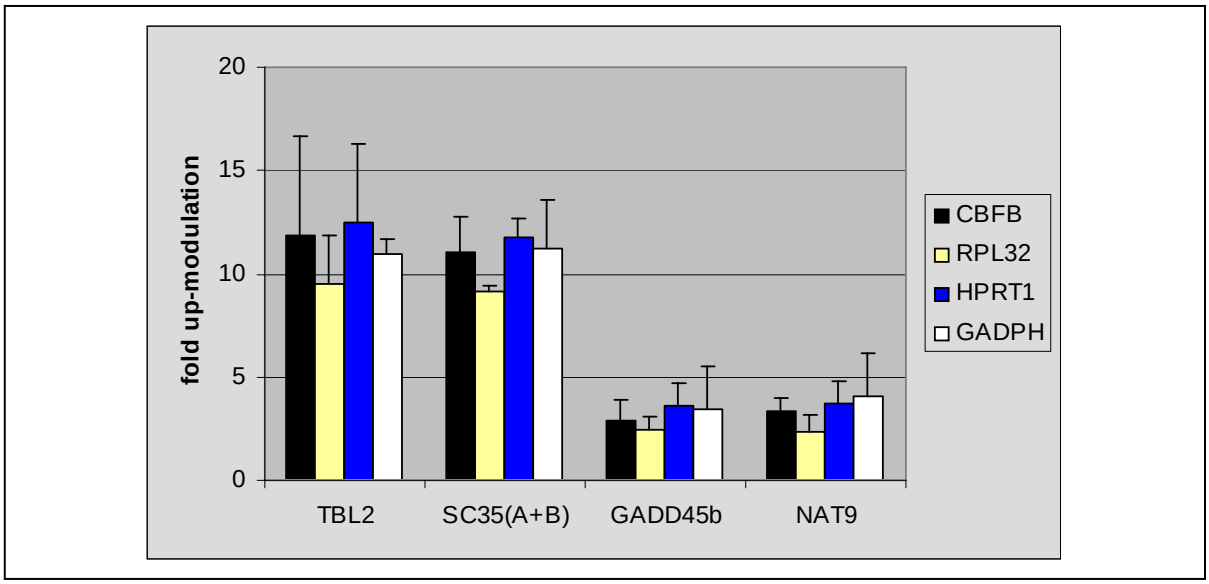

Figure19. Up-regulation of direct NMD substrates in UPF1-depleted cells measured according to distinct non-NMD transcritps. NMD endogenous mRNAs levels were quantified in normal and Upf1-depleted cells from HeLa cells (strain A) using GAPDH, CBFB, RPL32 and HPRT1 transcripts as normalisation controls. The values represent fold up-modulated levels. Mean and standard errors were calculated from 3 independent experiments.

Three different strains of HeLa cells (referred to as A, B and C) were chosen to systematically analyse the NMD efficiency. These cells were transfected with the β-globin (NS39) reporter in 3 independent experiments and the abundance of this NMD-sensitive transcript was compared in the 3 strains (Fig. 20a and b). The down-modulation of the NS39 reporter was approximately 3-fold stronger in the HeLa strain A than in the strains B and C, while C tended to be approximately 1.5 fold stronger than B (Fig. 20a). I next analysed whether these differences in NMD efficiency of transfected β-globin (NS39) between the HeLa strains were reflected by differences in abundance of the panel of cellular NMD targets.

The quantification of these cellular NMD reporters demonstrated that the differences of NMD efficiency between strains A, B and C, as estimated by the transfected NS39 reporter, were also reflected by the abundance of the endogenous reporter panel (Fig. 20c). Moreover, when the data of the five NMD transcripts were combined together, the estimation of NMD efficiency by the endogenous targets gave very similar results to that based on the tranfected β-globin (NS39) reporter (Fig. 20d).

These data demonstrate that differences in NMD efficiency between human cell lines can be estimated by measuring the abundance of a carefully validated panel of cellular NMD targeted transcripts.

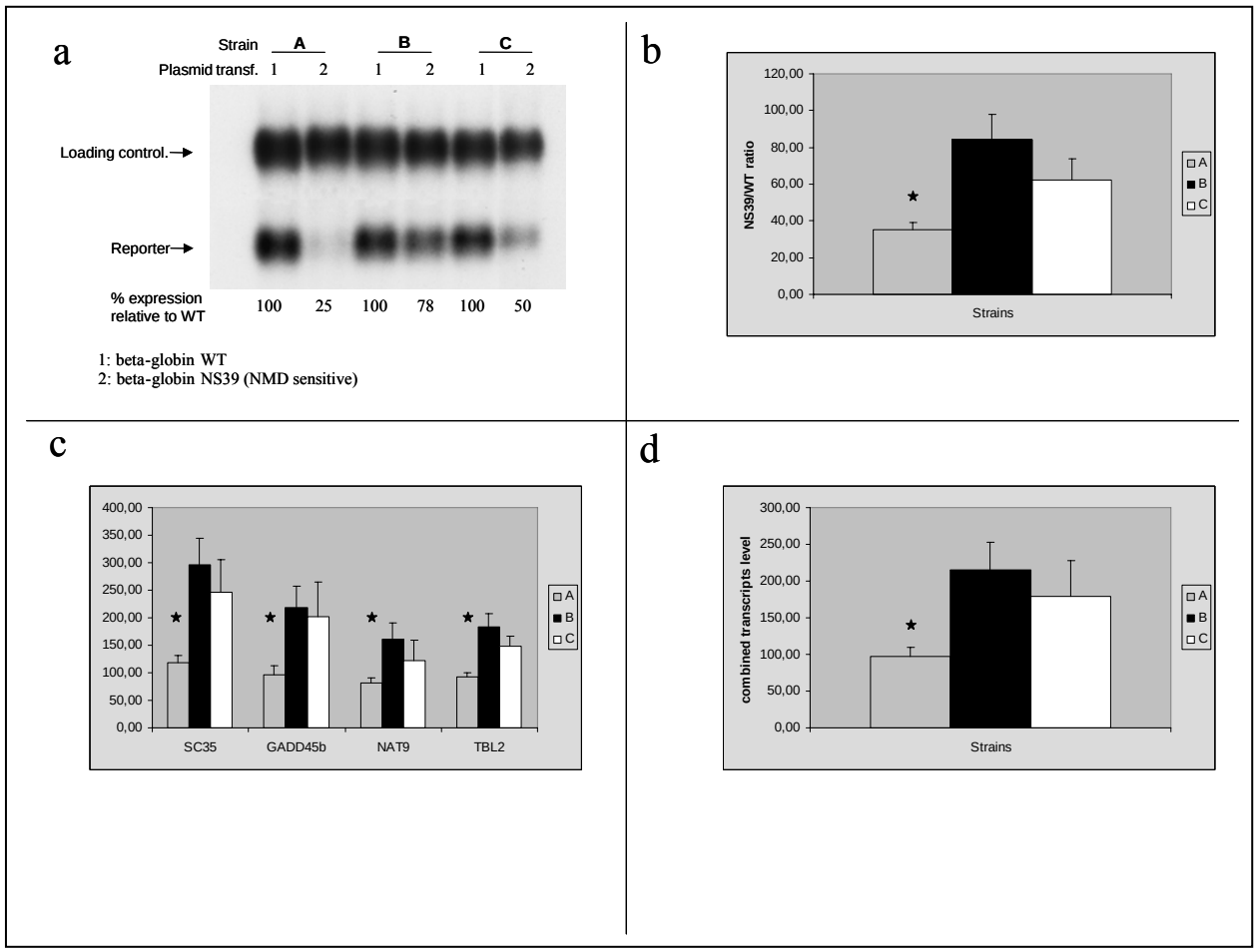

Figure 20. The abundance of cellular NMD targets reflects the variability of NMD efficiency in HeLa cell strains. a: Representative Northern-blot of RNA from 3 strains of HeLa cells that were transfected with β-globin WT or β-globin NS39. **b:** Quantification of the NS39/WT ratio for the Northern-blots analysis. The bar diagram expresses the mean and standard deviation for 3 independent experiments. **c:** Quantification of the 5 endogenous NMD transcritps (SC35 A and B are combined) by real-time PCR in untransfected cells. The values plotted represent the average of the quantification obtained using each of the non-NMD targets for normalization plus standard deviation. **d:** Average and standard deviation of the integrated data for the 4 NMD reporters (as shown in c) combined together. Significant differences ($p<0.05$) are indicated with a star.

3.3.2 RNPS1 is as a potential modulator of NMD efficiency

Subsequently, I aimed at understanding why these HeLa strains display differing NMD efficiencies. As a starting point, I analysed by immunoblotting the abundance of the key NMD proteins UPF1, UPF2 and UPF3b (Fig. 21, upper row) and of the functionally

critical exon junction complex components Y14, MAGOH, and RNPS1 (Fig. 21, lower row). The abundance of the UPF proteins, Y14 and MAGOH did not differ between the three strains. In contrast, RNPS1 is much less abundant in cells of the strain B (Fig. 21), which correlates with the poor NMD efficiency of this strain.

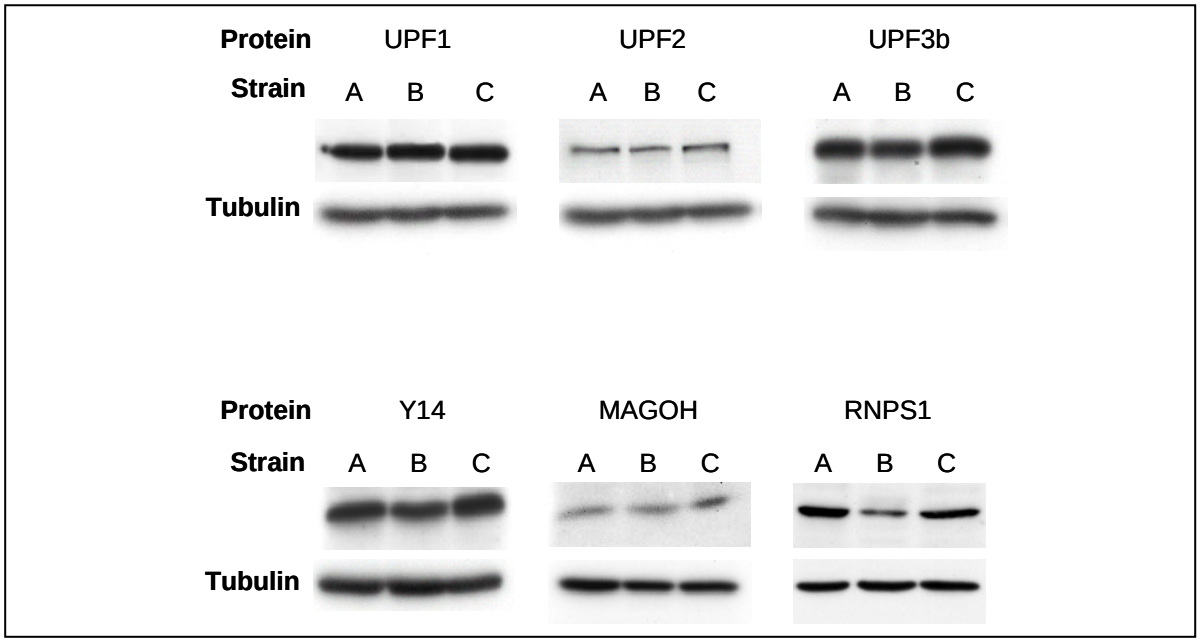

Figure 21. RNPS1 is less abundant in HeLa strain B. Representative Western-blots of cytoplasmic lysates following staining with specific antibodies against UPF proteins (upper row), Y14, MAGOH and RNPS1 (lower row) in HeLa strains A,B and C. Reprobing with a tubulin specific antibody was performed as a loading control.

I next functionally analysed if RNPS1 might be the limiting factor for NMD in these cells and over-expressed functional RNPS1 (Gehring et al., 2005) in cells that were transfected with β-globin reporter genes. Increasing amounts of RNPS1 increased the degradation of the β-globin NS39 reporter up to 4 times in cells of strain B but had no effect in cells of strains A or C (Fig. 22a).

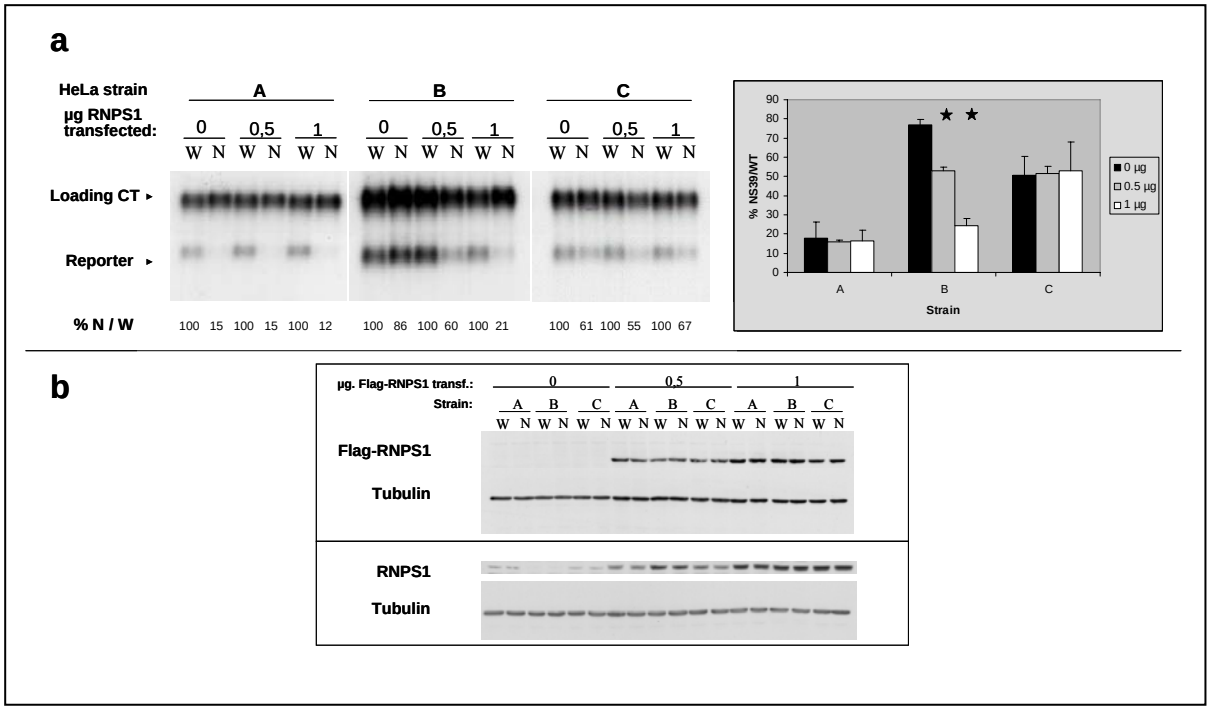

Figure 22. Over-expression of RNPS1 increases the degradation of a transfected NMD reporter in strain B cells. a: Left: Representative Northern-blot of RNA from the 3 strains of HeLa cells that were transfected with β-globin WT (W) or β-globin NS39 (N) and 0, 0.5 or 1 µg of pCI-NEO-FlagRNPS1. Right: The diagram shows the quantification of the Northern-blots. The bars indicate the mean and s.d. of 3 independent experiments. **b:** Western-blots for the transfected HeLa cells used in these experiments. The same extracts were run in two gels. After blotting, the upper membrane was probed using anti-flag antibody and the lower one was probed using anti-RNPS1 antibody. Both membranes were reprobed with anti-tubulin antibody to control the load.

This effect is specific for RNPS1, because over-expression of RNPS1Δ69-121 (a truncated version of RNPS1 known to be non-functional in NMD (Gehring et al., 2005)) does not affect the down-modulation of the NS39 reporter (Fig. 23a and c). Furthermore, the over-expression of Y14 – another protein of the EJC which is present in equal levels in the 3 strains – does not have any significant effect in the reporter's stability (Fig. 23b and c). I confirmed that the transfection of pCI-NEO-Flag has no effect on the abundance of endogenous RNPS1 in any strain (Fig. 22b, lower Western-blot) and that the plasmid is expressed at similar levels in the three cell lines (Fig. 22b, upper Western-blot).

These results demonstrate that the abundance of RNPS1 is limiting for NMD efficiency in strain B. The results also show that in strain C the low NMD efficiency is caused by an apparently different mechanism.

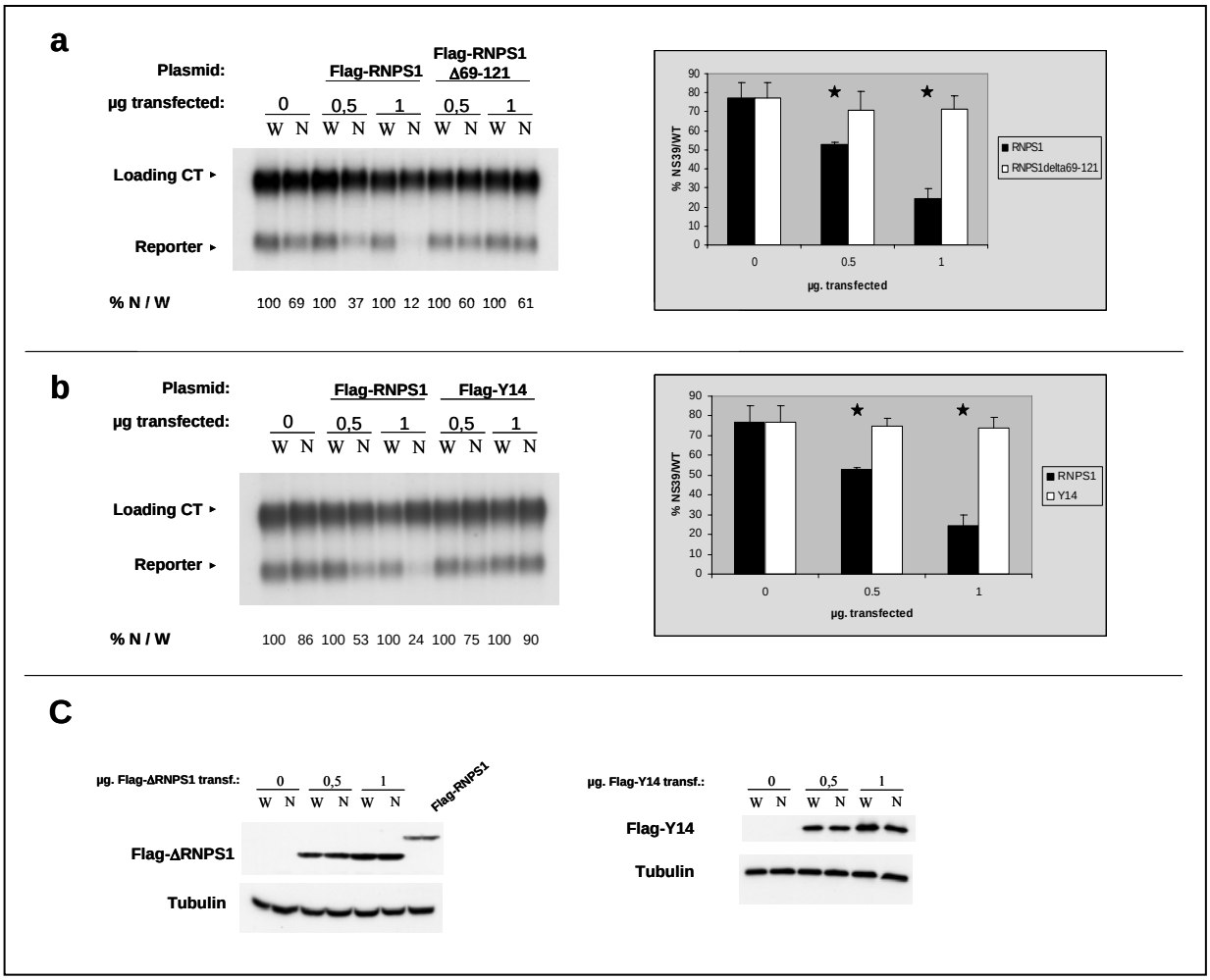

Figure 23. The increased NMD efficiency of strain B cells is RNPS1-specific. a: Left: Representative Northern-blot of RNA from strain B cells that were transfected with β-globin WT or β-globin NS39 and 0, 0.5 or 1 µg of pCI-NEO-FlagRNPS1 or pCI-NEO-FlagRNPS1Δ69-121 as a negative control. Right: The diagram shows the quantification of Northern-blots. The bars indicate the mean and s.d. of 3 independent experiments. **b:** idem b) but transfecting pCI-NEO-FlagY14 as a negative control. **c:** Western-blot for the transfected strain B HeLa cells. Left: An anti-flag antibody was used to verify the expression of flag-RNPS1Δ69-121. As a further control, an extract of cells transfected with the full length RNPS1 was also loaded. Right: An anti-flag antibody was used to verify the expression of flag-Y14. The blots were reprobed with anti-tubulin to control the load.

3.4 Quantitative differences in NMD in human blood: A step towards the clinic

3.4.1 Estimation of NMD efficiency in blood samples

The results presented in 3.3 support the hypothesis of distinct NMD efficiency at the cellular level. As a step forward, I then asked whether it is also possible to characterise NMD efficiency at a tissue or even whole organism level. To this end, NMD efficiency was estimated in human blood samples by applying the assay already established in the HeLa model.

The analysis in HeLa cells demonstrated that NMD efficiency is a relatively stable parameter in the different strains (Fig. 20). To address the question of whether NMD

efficiency also represents a constant parameter in blood, I analysed samples from five healthy volunteers who donated blood four times on a weekly basis. Total RNA was extracted either from whole blood or mononuclear cells (mainly lymphocytes and some monocytes) after a Ficoll gradient. The gene expression values of donor 1 on day 1 were arbitrarly set to 100% and the other values were normalised against this reference. For each NMD transcript, its abundance was calculated against each standard gene and the mean and standard deviation were plotted for comparison.

Several observations can be drawn from the outcome of this analysis (Fig. 24). In general, it has to be noted that: 1) for each donor, there are considerable variations in the abundance of each individual NMD transcript in different samples. The fluctuations seem to move in a wider range in the case of RNA extracted from whole blood than for the RNA extracted from the purified mononuclear cells. 2) There is no correlation between the abundance of different NMD targets in a single sample. For example, in the quantification of mononuclear cell fractions, donor 4 presents the highest levels of SC35 and GADD45B on the fourth day; however, this donor presents the lowest value for Nat9 in the same sample (Fig.24b). 3) Finally, although the standard deviation represents only 20% of the transcript mean abundance, the difference in abundance of each NMD transcript per donor is relatively low (30-40%) and consequently it is not possible to distinguish whether these donors have a characteristic baseline NMD efficiency.

These results constrast with those obtained in the HeLa model and they could be attributed to the higher complexity of blood, composed of many different cell types that might have a distinct NMD capacity. The observation that the range of fluctuation in the abundance of the NMD targets tends to be narrower after a purification step in Ficoll (where the granulocytes are removed from the samples) might support this idea.

To test whether the variability in the transcript abundance is due to working with crude samples, I applied the same assay to lymphoblastoid cell lines obtained from the Coriell Cell Repositories. These lines were established by Epstein-Barr virus transformation of peripheral B lymphocytes from donors. Eight different lines originating from four pairs of identical twin siblings were tested. The abundance of the set of NMD transcripts was assessed in three different passages of each line. The analysis indicates that the NMD transcripts´ levels are relatively stable in the lymphoblasts (Fig. 25).

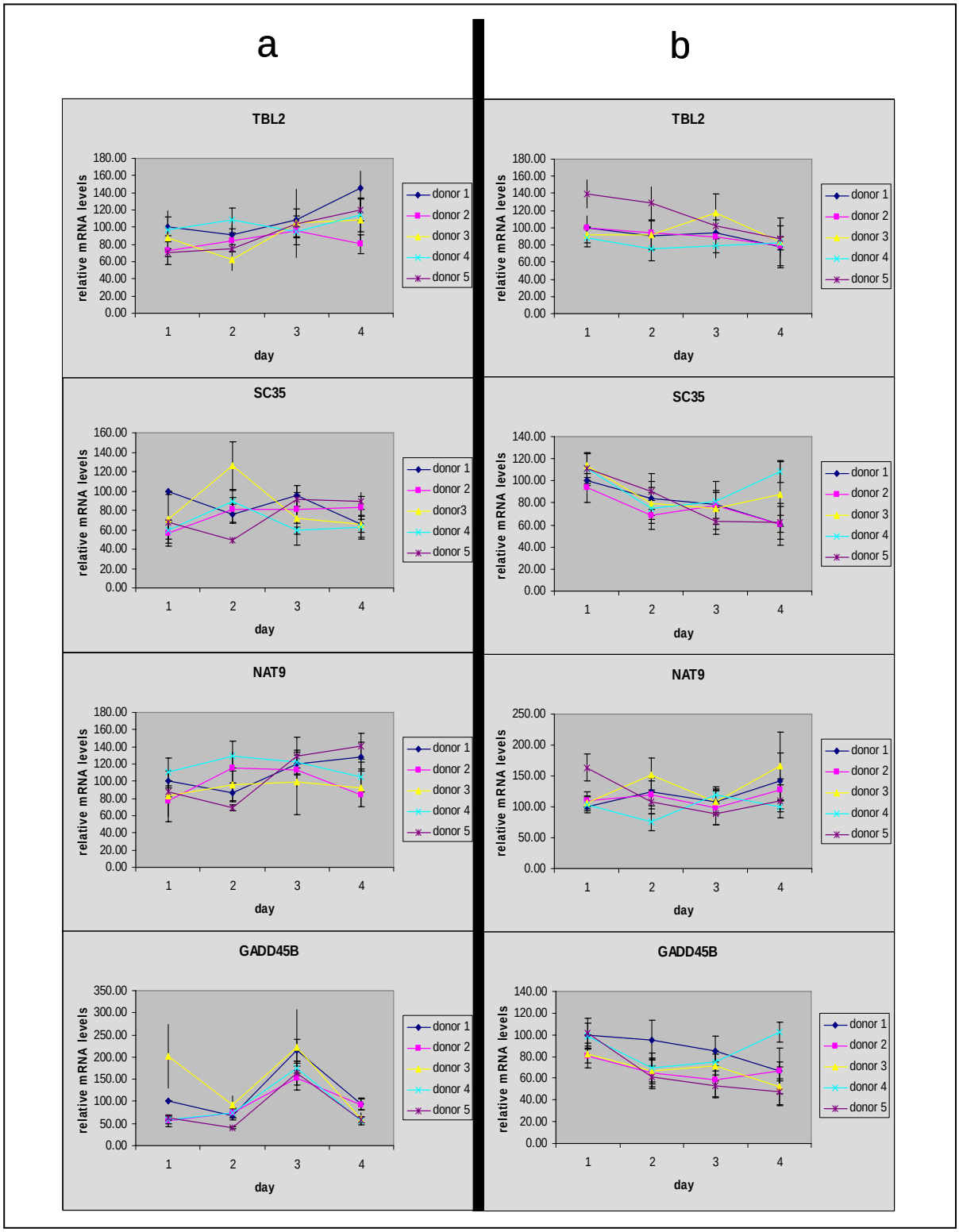

Figure 24. NMD transcripts abundance in blood of healthy donors. RNA extracted from whole blood (panel a) or from the mononuclear fraction after a Ficoll grandient (panel b) was used to quantifiy the abundance of a panel of NMD transcripts. The values plotted represent the average and standard deviation of the values obtained upon normalisation against each of the reference genes. Each donor donated blood on 4 different days as indicated.

For SC35, TBL2 and NAT9 the mRNA abundance varies in a ± 20% range. For GADD45B, the range of variability is higher but limited to a maximum of ± 50% (Fig. 25, lower left).

Nevertheless, the abundance of these transcripts is very similar in the eight lymphoblastoid cell lines. None of the mean values is significantly different fron the others, as assessed by the magnitude of their standard deviations (for example, GADD45B level of line L2-2 in the second passage; Fig. 24). This lack of diversity in the transcript levels prevents me from verifying any correlation in the NMD efficiency of genetically homogeneous cell lines. According to this assay, there is no distinction in NMD capacity between genetically related or unrelated cell lines.

Taken together, these analyses do not support the idea of quantitative differences of NMD efficiency in blood from healthy donors, as least as assessed by this assay system. Alternatively, it is possible that my set of NMD/standard transcripts is not suitable to estimate NMD efficiency in these systems. For instance, a potentially strong expression control at the transcriptional level could mask the NMD effect in these cells.

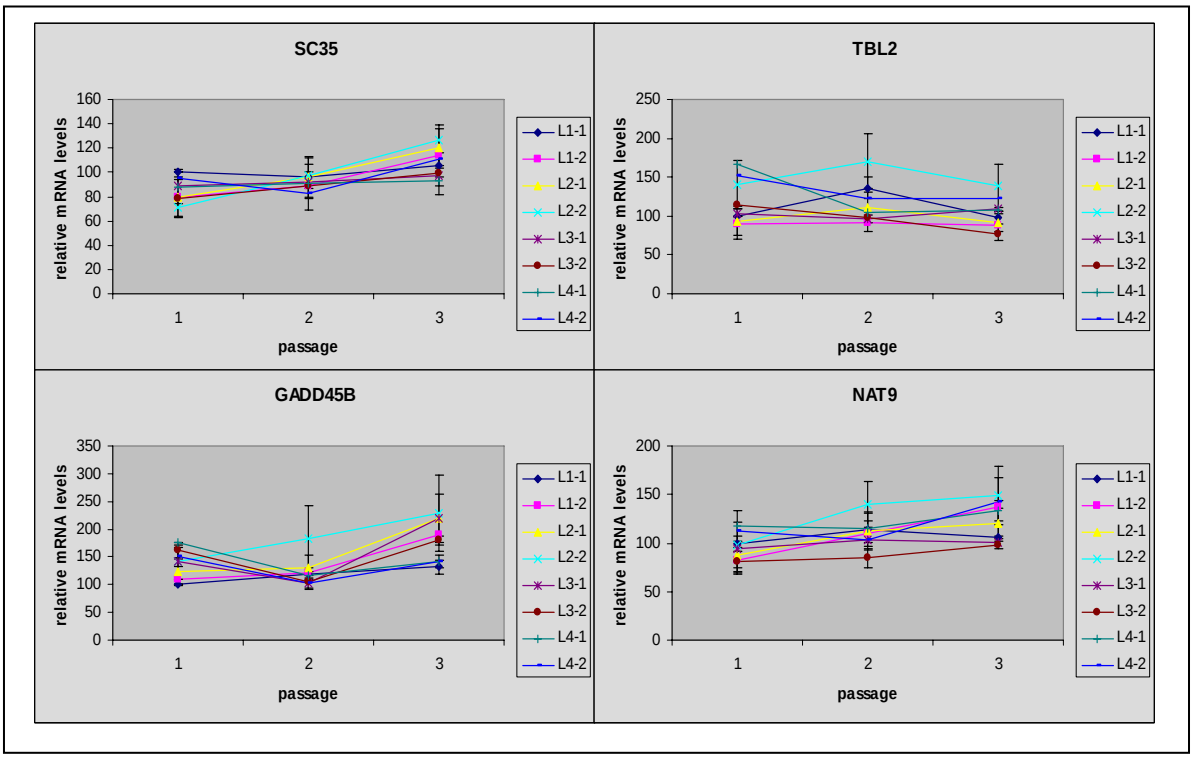

Figure 25. NMD transcripts abundance in lymphoblastoid cell lines. RNA extracted from EBV-transformed B lymphocytes was used to quantify the abundance of my panel of NMD transcripts. The values plotted represent the average and standard deviation of the values obtained upon normalisation against each of the 4 reference genes. 8 different lines were analyses coming from 4 pair of twin siblings: L1-1 and L1-2, L2-1 and L2-2, L3-1 and L3-2, L4-1 and L4-2.

3.4.2 A case study: β-thalassemia caused by mutation in IVS1+6 of the β-globin gene

As commented in the introduction, a mutation in the position 6 of the first intron of the β-globin gene impairs the efficiency of correct splicing of the β-globin transcript (Treisman et al., 1983). Consequently, three alternative isoforms are produced in addition to the wild type as a result of the recognition of cryptic 5´splicing donor sequences. I will refer to these alternative spliced variants as exon1(+12), exon1(-16) and exon1(-38). While exon1(+12) does not contain any PTC, exon1(-16) and exon1(-38) are theoretically NMD targets. Studies done in transfected HeLa cells have confirmed the NMD-sensitivity of the exon1(-16) variant; however, in the same studies, the exon1(-38) isoform proved to be resistant to NMD, presumably due to an unknown cis-acting determinant that influences the surveillance pathway.

As part of an on-going collaboration with one group in Palestine (Dr.Kanaan, Bethlehem University) and one group in Israel (Dr. Filon, Hadassah University Hospital), I examined the expression of β-globin transcript variants in a group of patients from Eastern Mediterranean homozygous for the IVS1+6 (T→C) mutation in the β-globin gene. While all the patients carry the same mutation and consequently suffer from β-thalassemia, the phenotypic expression of the disease is highly variable within the group. Our collaborators have classified the patients into two categories according to disease severity: those who suffer from thalassemia major and those who suffer from thalassemia intermedia. Several parameters were taken into account for the classification. In general, patients were considered to suffer from thalassemia major when were transfused for the first time at an early age (< 2 years old), required regular transfusions (every 1-3 months) and have significantly low MCV (< 60 fL).

All well-studied genetic modifiers that could potentially explain the clinical variability have been explored (concomitant α-thalassemia, α-globin gene triplication and $^{G}\gamma$-globin -158 (C→T) mutation). Despite these efforts, the reason of variable severity in these patients remains elusive. Therefore, the analysis of the β-globin splicing variants could shed some light on a potential role of NMD in the clinical manifestation of this thalassemia.

RNA extracted from mononuclear cells (previously purified by Ficoll gradient) from these patients´ blood was analysed by RNase protection assay (RPA). For this purpose, a riboprobe containing most of exon1 and intron1 (carrying the IVS1+6 (T→C) mutation) was used to detect the four expected splice variants (Fig. 26a).

In addition, four constructs: -38ΔIVS1, -16ΔIVS1, β-globin WT and β-globin IVS1+6 (T→C) were transfected separately into HeLa cells. The -38ΔIVS1 and -16ΔIVS1 constructs express the exon1(-38) and exon1(-16) isoforms, respectively. The β-globin WT construct encodes the normally spliced β-globin mRNA and the β-globin IVS1+6 (T→C) gives rise to the same 4 transcripts expected in the patients´ samples (see Materials and Methods). Total RNA of these transfected cells and RNA extracted from blood of healthy volunteers were used to control the position of the expected bands in the RPAs.

In these assays, the four β-globin transcripts could be detected although the normal isoform alone accounts for 95-98% of the total expression of the β-globin gene in this group of patients (see Fig. 25b for a representative RPA). On the other hand, the exon1(-16) variant represents the least abundant isoform in all the samples, accounting for ~0.5% of the total β-globin mRNA.

A total of 24 patients (14 suffering from thalassemia major and 10 from thalassemia intermedia) was analysed by this assay and the relative abundance of each abnormal splicing variant was quantified against the wild type isoform (Fig. 26).

This analysis indicates that the aberrant isoforms are expressed at very variable levels in the different individuals and that there is no significant difference in the expression levels between the two groups of patients (thalassemia major vs. thalassemia intermedia). For example, the exon1(-16) isoform represents 0.52 ± 0.5 % of the wild type β-globin in the "thalassemia major" group and 0.31 ± 0.14 % of the same isoform in the "thalassemia intermedia" group (Fig. 27).

The levels of the exon1(-16) variant are extremely low in comparison to those of the wild type. As this may produce unreliable quantifications, the abundance of the exon1(-16) isoform was also compared against the exon1(+12) – which is not an NMD target – and against exon1(-38) – that is NMD resistant in HeLa cells. All the aberrant isoforms are expressed at similar levels; however, these additional comparisons confirmed that both groups are indistinguishable (Fig. 27).

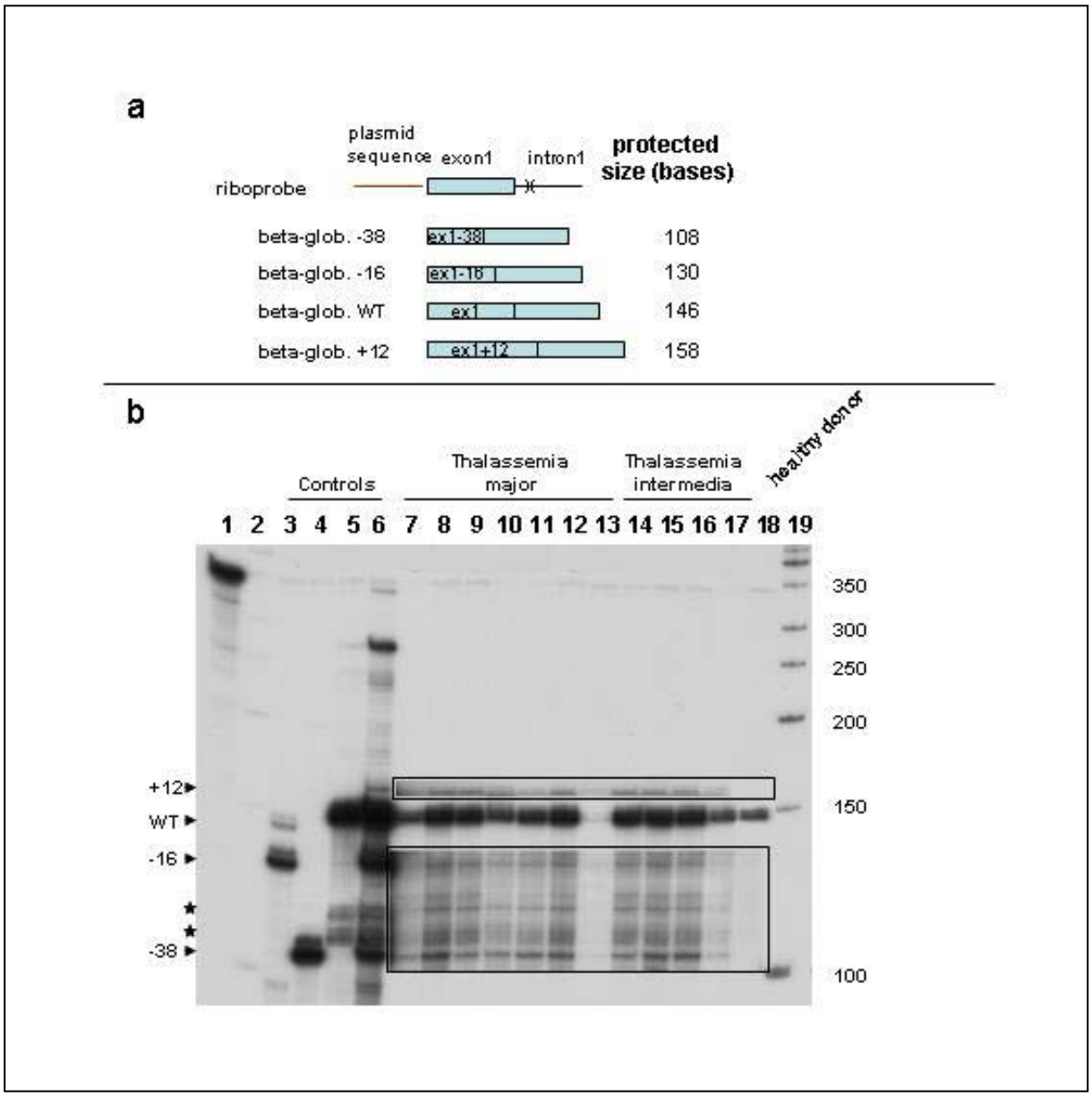

Figure 26. Detection of β-globin transcripts in patients´ samples by RPA. a: Scheme showing the sequence of the riboprobe used which consist of the last 146 bases of exon1 and intron 1; the position 6 of intron1(represented by a cross) has been mutated to a T to match the sequence in the patients. The different expected isoforms are aligned with the riboprobe and the predicted protected size is indicated. **b:** Representative RPA for the patients samples. Line 1: full riboprobe; lines 2 and 19: 50 bp marker. To control the position of the expected bands, RNA of HeLa cells transfected with the -16ΔIVS1 (line 3), -38ΔIVS1 (line 4), β-globin WT (line 5) and β-globin IVS1+6 (T→C) (line 6) was also loaded. Patients´ samples were grouped according to the severity of their anemia: lanes 7-13: thalassemia major; lanes 14-17: thalassemia intermedia. Lane 18: sample from a healthy donor. The position of the transcripts is indicated by arrows. Exon1(+12), β-globin WT, exon1(-16) and exon1(-38) are described as +12, WT, -16 and -38, respectively. Two unidentified bands –indicated by stars– appear in the samples but they also appear in the β-globin WT control. The exon1(+12) variant is also present in the normal blood sample although at lower levels. The parts enclosed in squares have a longer exposition due to the difference in intensity between the β-globin WT and the aberrant variants.

Although the variability in the quantifications prevents me from stating any statistically proven conclusions, it is interesting to observe that the average abundance of all the three aberrant isoforms is higher in the "thalassemia major" group. If this observation could be validated in a larger number of samples, then a difference in splicing could be responsible for the phenotypic variability of this thalassemia rather than variable NMD efficiency. A consequence of this hypothesis is that – under equal transcription rates – the β-globin wild type mRNA would be less abundant in the more severe group due to the synthesis of a higher proportion of the alternative isoforms. Unfortunately, the abundance of β-globin wild type mRNA has not been assessed yet and it remains an important experiment to define a modifier of the severity of this disease.

Total number of samples : 24

Number of samples from : 14
T.major patients

Number of samples from : 10
T. Intermedia patients

Isoforms proportion	Thalassemia major	Thalassemia intermedia
(-16 / wt) %:	0.52 ± 0.5	0.31 ± 0.14
(-38 / wt) %:	1.42 ± 0.4	1.13 ± 0.5
(+12 / + wt) %:	0.79 ± 0.4	0.63 ± 0.38
(-16 / -38) %:	31.3 ± 27.7	29.6 ± 11.7
(-16 / +12) %:	85 ± 12	62 ± 35

Figure 27. Quantification of the β-globin transcripts in patients´ samples in the RPAs. The values are given in percentage and indicate the mean and standard deviation of the abundance of one isoform compared to a second one. Exon1(-16), exon1(-38), exon1(+12) and β-globin WT are indicated by -16, -38, +12 and wt, respectively.

Taken together, my results so far do not support a general role of NMD in the phenotypic expression of β-thalassemia IVS1+6 (T→C). This conclusion is based on the observation that there is no significant difference in the abundance of an NMD-targeted β-

globin transcript in patients suffering from mild or severe anemia. A larger number of patients should be studied to obtain statistically significant conclusions. Additionally, quantification of wild type β-globin transcript should be carried out in order to test possible alternative explanations for the disease variability.

4. Discussion

4.1 Identification of *bona fide* NMD targets

Messenger RNA stability represents a parameter targeted by post-transcriptional mechanisms that ensure quality control over gene expression: mRNAs that contain PTCs and thus encode C-terminally truncated polypeptides are generally degraded by NMD, a system that is conserved from yeast to humans (Gonzalez et al., 2001; Mango, 2001; Gatfield et al., 2003; Maquat, 2004).

In humans, NMD involves splicing-dependent marking of the exon junctions by the EJC (Kataoka et al., 2000; Le Hir et al., 2000). According to the current mechanistic understanding of NMD, the position of the translation stop codon relative to the position of the EJCs is sensed by the ribosome and factors that associate with the ribosome during translation termination. Transcripts with an exon-exon junction following more than ~50 nucleotides downstream from the stop codon are generally interpreted as improper and subjected to degradation (Nagy and Maquat, 1998; Thermann et al., 1998). NMD also requires UPF1, UPF2 and UPF3b factors (Gonzalez et al., 2001). According to a current model, the latter binds UPF2, which in turn interacts with UPF1 in a linear pathway. Binding of the translation release factors eRF1 and eRF3 to Upf1 biochemically connects these NMD factors to translation termination (Kashima et al., 2006).

Originally, NMD was thought to be a surveillance mechanism that detects and degrades aberrant mRNA derived from the expression of nonsense or frameshift mutations in genes or from faulty splicing. The down-modulation of immunoglobulin PTC-containing transcripts that result from programmed V(D)J rearrangements foreshadowed a role for NMD also in normal cellular function (Baumann et al., 1985; Li and Wilkinson, 1998). Moreover, NMD is implicated in the auto-regulation of the expression of some genes subjected to alternative splicing (Sureau et al., 2001; Wollerton et al., 2004; Cuccurese et al., 2005). A wider physiological role of NMD was uncovered in yeast and fly where genome-wide analyses revealed that 5-10% of the transcriptomes of these organisms are affected by NMD (He et al., 2003; Rehwinkel et al., 2005).

In the present study, a first goal to achieve was the identification of physiological *bona fide* NMD transcripts. A genome-wide microarray analysis of UPF1-depleted HeLa cells indicated that this central NMD factor seems to be implicated in the control of the levels of a wide variety of transcripts in humans. Around 230 genes representing 2.8% of the total number of tested mRNAs were up-modulated according to this microarray data. UPF1

depletion impairs NMD activity, thus stabilising NMD targets (Lelivelt and Culbertson, 1999; Page et al., 1999; Mendell et al., 2002; Gehring et al., 2003). Thus, it is assumed that most of the up-modulated transcripts likely represent direct or indirect NMD targets. Under this assumption, the data suggest that NMD particularly affects certain functions including amino acid metabolism and transport, apoptotic activity, signal transduction and cell adhesion (Fig. 12). However, the diversity of functional classes rather than the number of over-represented gene ontology categories is the most striking. Transcripts encoding for enzymes with oxidoreductase activity and kinases, interleukin biosynthesis, DNA repair, Golgi to plasma membrane transport and carbohydrate metabolism are other examples of the wide spectrum of activities regulated by NMD.

Parallel to my analysis, a similar study was published by another group that also reported multiple effects of NMD in normal cellular function (Mendell et al., 2004). Approximately 40% of the up-regulated transcripts found in this second study were also up-regulated in my investigation. I attribute the differences in the remaining 60% to the use of different microarray platforms and strains of HeLa cells. In Mendell et al. (2004) only genes involved in amino acid transport were found to be significantly over-represented. This apparent discrepancy, however, is likely explained by the fact that a larger microarray containing more than 10000 extra probe-sets was used in my study. Consistently, a larger number of affected transcripts was found in my investigation.

Interestingly, there is no evident overlap in the functions affected by NMD between humans and yeast. Recently, this observation has also been reported and expanded to *D. melanogaster* (Rehwinkel et al., 2005). The conclusion is that although NMD regulates a large number of cellular activities, these functions are not conserved from yeast to human.

As it was already mentioned, these conclusions are based on the assumption of UPF1 having an (almost) exclusive role in NMD. Nevertheless, it is possible that a fraction of the transcripts described in this study is actually targeted by unknown NMD-independent functions of UPF1. Indeed, a role of UPF1 in non-NMD functions has been demonstrated in the last two years. UPF1 is involved in a novel mRNA decay mechanism called Staufen-mediated mRNA decay (SMD). For this pathway, translation and the interaction of UPF1 with the RNA binding protein Staufen (Stau1) are essential. In contrast, pre-mRNA splicing and the NMD factors UPF2 and UPF3 are dispensable, features that distinguish SMD from NMD (Kim et al., 2005).

In addition, UPF1 (but not UPF2) is required for DNA replication and S phase progression (Azzalin and Lingner, 2006a) and is implicated in the regulated degradation of

histone mRNAs at the end of the S phase (Kaygun and Marzluff, 2005). Consistent with these functions, UPF1 interacts with DNA polymerase delta (Azzalin and Lingner, 2006a) and with the histone-binding stem-loop binding protein (SLBP) (Kaygun and Marzluff, 2005). Hence, UPF1 is emerging as a key player in different cellular processes that exerts multiple functions depending on the complexes to which it associates (Azzalin and Lingner, 2006b). Consequently, caution is requiered when interpreting the physiological role of NMD on the basis of microarray analysis. Functions that are attributed to NMD might be related to UPF1 NMD-independent activities. For instance, although not experimentally proven, it is tempting to associate the up-regulation of transcripts involved in apoptotic activity in UPF1-depleted cells with the requirement of UPF1 for S phase progression.

More recently, a genome-wide microarray analysis in UPF2-depleted HeLa cells has also been accomplished (Wittmann et al., 2006). That study revealed that while there are a number of transcripts regulated by both UPF1 and UPF2, there are also mRNAs affected exclusively by either UPF1 or UPF2, suggesting the possibility of UPF1- and UPF2-specific NMD complexes (or non-NMD functions) with only a partial target overlap. Indeed, a study from our laboratory showed biochemical evidence of a UPF2-independent NMD pathway (see below).

Nevertheless, a common limitation of all the discussed studies is the lack of an estimation of the number of *bona fide* NMD targets in the microarray datasets. Non-discrimination between direct versus indirect effects is always a limitation in large-scale expression analysis. Consequently, validation of microarray data has to be done in order to advance in the understanding of both, the impact and the mechanism of NMD. The present study is the first one to attempt a discrimination of secondary effects in a group of putative NMD-targeted mRNAs. As NMD is a posttranslational mechanism, it only acts on spliced mRNAs. Thus, pre-mRNA abundance of a *bona fide* NMD transcript should not be affected by this pathway. I made use of this reasoning as a basis for distinguishing non-legitimate from genuine NMD substrates. My analysis revealed that 12 out of 16 selected transcripts from the microarray data were affected at the transcriptional level and consequently, likely represent indirect targets (Fig. 15). If we extrapolate these results to the entire dataset, this would imply that around 75% of the up-modulated mRNAs found in my analysis are not primary targets for NMD.

Indirect effects could be a reason for the large number of down-modulated transcripts found in my microarray data. The large proportion of indirect targets might also explain why approximately 60% of the transcripts I found to be up-modulated in UPF1-depleted cells have

no classical NMD feature that could account for their putative NMD sensitivity (Table 1). Interestingly, the published studies in humans (Mendell et al., 2004; Wittmann et al., 2006), yeast (He et al., 2003) and fly (Rehwinkel et al., 2005) all show a significant proportion of putative NMD-regulated transcripts that also lack such structural features. Hence, the significant contribution of secondary effects imposes a limitation in the interpretation and comparison of these microarray datasets.

Nevertheless, valuable information can be gained from the microarray analysis when it is complemented with further assays such as quantification of pre-mRNA/mRNA abundance and mRNA stability. Through this carefully validated process I characterised two new bona fide NMD targets, namely *TBL2* and *NAT9* (Fig.15 and 16). This procedure allowed me also to confirm another target (*GADD45B*) predicted by in-silico analysis (Hillman et al., 2004).

To conclude, the microarray technique is a powerful tool that permits the identification (on a global scale) of transcripts affected by depletion of an NMD factor. However, this tool has innate limitations represented by the indiscrimination of *bona fide* from indirect or off-target effects (Alonso, 2005). These limitations can be overcome by complementing the microarray analysis with suitable assays that allow validation of the results. I applied this procedure to a small subset of transcripts of my microarray dataset. A comprehensive validation of all the microarray dataset should be carried out to uncover the set of legitimate NMD transcripts. A similar task should be accomplished with the datasets found in the literature, including those that deal with NMD in other organisms. Comparison of such revised sets exclusively composed by genuine NMD transcripts will, in turn, shed light on the mechanism, the physiological role and the conservation of NMD in eukaryotes.

4.2 Quantitative differences in cellular NMD efficiency

It has long been observed that NMD does not reduce to zero levels of transcripts containing PTCs. Moreover, the efficiency of the transcript down-regulation seems to be particular for each mRNA. For instance, aberrant isoforms of TPI (triose phosphate isomerase) and β-globin are usually down-regulated 2 to 5-fold whereas DHFR (dihydrofolate reductase) and APRT (adenine phosphoribosyl transferase) NMD-sensitive variants are reduced 5-10 fold (Gudikote and Wilkinson, 2002). Likewise, different mutant mRNAs were found to be degraded due to NMD to variable levels in a colorectal cancer cell line (El-Bchiri et al., 2005). In agreement with these observations, I found disparate up-modulation levels in the endogenous NMD transcripts used in this study. While SC35 and TBL2 were up-

modulated 10-12 fold in cells lacking UPF1, GADD45B and NAT9 were up-modulated only three fold under the same conditions (Fig 19). This transcript-dependent NMD efficiency might be explained in terms of differences in the transcriptional or translational rates in some cases. Alternatively, the presence of cis-acting sequences may modulate the NMD activity on a particular mRNA. In support of the latter idea, it has been demonstrated that the Ig μ and TCR-β transcripts possess a sequence that is required for their highly efficient degradation. In both cases, the introduction of the corresponding sequence element enhances degradation of an otherwise poor NMD reporter (Gudikote and Wilkinson, 2002; Buhler et al., 2004).

It has also been suggested that innate NMD capacity may vary among different cells, tissues or even entire multicellular organisms (Frischmeyer and Dietz, 1999). NMD variability has been demonstrated in yeast, in which different strains of *S. cerevisiae* possess distinct NMD efficiencies (Kebaara et al., 2003). The mechanism responsible for the variable strain-dependent NMD efficiency has not been unravelled. However, phenotype analysis of spore clones from a cross of cells with differing NMD capacities suggested that at least two genes, and probably more, are implicated in the variability of the pathway.

To my knowledge, no systematic analysis on NMD efficiency has been done in humans or higher eukaryotes so far. This is surprising since NMD can regulate the phenotypic expression of disease (see introduction). However, some clinical reports suggest that there might exist differences in NMD efficiency. Patients with diverse disease severity carrying identical nonsense mutations but expressing different levels of the resulting truncated protein have been reported for dystrophin and Jarid1c genes (Kerr et al., 2001; Jensen et al., 2005). Furthermore, tissue-specific NMD has been suggested in a case of Schmid metaphyseal chondrodysplasia (Bateman et al., 2003). Yet, direct experimental proof of quantitative differences in NMD efficiency is absent.

The present study is the first one that provides evidence of stable differences in NMD efficiency in human cells. Five endogenous NMD-sensitive transcripts (SC35 A, SC35 B, TBL2, NAT9 and GADD45B) encoded by four different genes showed differences in their steady-state levels in three different strains of HeLa cells. In particular, the abundance of these NMD targets was reproducibly and significantly lower in cells of strain A than in cells of strain B or C (Fig. 20c and d). Consistent with these results, the steady-state levels of ectopically expressed β-globin NS39 were significantly lower in cells of strain A (Fig. 20a and b). The four endogenous gene reporters are expressed at different levels and apparently perform different roles in the cell: SC35 is a splicing factor that regulates alternative splicing (Fu, 1995); GADD45B is thought to function in co-operation with other genes in cell growth

inhibition (Takekawa and Saito, 1998). The functions of TBL2 and NAT9 are not known yet; however, sequence similarity analysis predicts that TBL2 could function as a small G protein (Perez Jurado et al., 1999) and NAT9 could be an N-acetyltransferase (Strausberg et al., 2002). It is therefore likely that these genes are involved in different metabolic pathways and are most probably subjected to different regulatory controls, being NMD the only known pathway that controls all of them. Still, I observed a strict correlation in the abundance of these transcripts in a strain-dependent manner which was also independently reflected by a transfected NMD reporter assay. Consequently, I interpret these findings as an indication of distinct NMD efficiency in the tested strains. Other potential explanations (such as variations in transcription, mRNA exportation, etc) seem unlikely in this context.

To gain insight into the mechanism of NMD variability, I analysed the abundance of key NMD factors in the HeLa strains. Surprisingly, I found that cells of strain B, which have the lowest NMD efficiency, display a lower abundance of the factor RNPS1 (Fig. 21). I confirmed that a deficiency of RNPS1 is at least partially responsible for the low NMD activity in these cells because over-expression of RNPS1 enhances the degradation of an NMD reporter in cells of strain B but it has no effect in cells of the other strains (Fig 22).

RNPS1 is an EJC component initially identified as a general activator of splicing (Mayeda et al., 1999). Y14 is another EJC component which is stably associated to the complex both in the nucleus and in the cytoplasm (Kim et al., 2001; Le Hir et al., 2001). Unlike RNPS1, Y14 was shown to be equally abundant in the three cell strains (Fig. 21). To verify that the effect of RNPS1 over-expression is specific to this protein and not related to overexpression of EJC proteins in general, Y14 was also over-expressed. Contrary to the results obtained with RNPS1, the over-expression of Y14 did not have any effect on the reporter levels in strain B cells (Fig. 23b). Moreover, the over-expression of a truncated NMD-defective RNPS1 protein (Gehring et al., 2005) could not reproduce the effects of the wild type protein (Fig. 23a).

RNPS1 contains an RNA-recognition motif (RRM) and an arginine/serine/proline rich domain (RS/P) that resembles the RS domain of the so called SR protein family of splicing factors (Mayeda et al., 1999; Sakashita et al., 2004). Interestingly, over-expression of one of these factors –SF2/ASF (SFRS1) – was also reported to enhance NMD activity in HeLa cells (Zhang and Krainer, 2004). However, while RNPS1 shuttles between the nucleus and the cytoplasm, SF2/ASF appears to be exclusively located in the nucleus, making it difficult to explain how this factor can affect NMD. Nevertheless, it will be interesting to address a general potential role of SR proteins in NMD. However, preliminary results of our laboratory

suggest that over-expression of SFRS4 and SFRS7 splicing factors do not have any effect in strain B cells which would favour a particular role of RNPS1 as an NMD-modulator.

Recently, the model that suggests the stepwise interaction of the UPF factors in a linear pathway necessary for NMD has been challenged (Gehring et al., 2005). It has been demonstrated that different EJC subcomplexes can trigger NMD. In a tethering assay, Y14, MAGOH and eIF4A3 were shown to activate NMD in an UPF2-independent manner, whereas UPF2-induced NMD required RNPS1. The two proposed pathways meet in a common requirement for UPF1.

The NMD-sensitive transcripts used in my assays are all UPF1- and UPF2-sensitive (Fig. 17). Consistent with the branched NMD model, this set of mRNAs was affected by a deficiency in RNPS1 in one of the strains of HeLa cells. In addition, these findings might help to test the new model further. If the deficiency of RNPS1 were the only significant factor responsible for the low efficiency of NMD in strain B compared to A, a prediction would be that the abundance of UPF2-independent NMD-sensitive mRNAs should be the same for both of these strains. In any case, the absolute requirement of RNPS1 for the UPF2-dependent NMD pathway and the role of RNPS1 uncovered in the present study suggest that relative changes in RNPS1 levels may act as a genetic modifier in NMD.

As stated above, in yeast the NMD phenotype is pleiotropic. Given the increased complexity of NMD in higher eukaryotes, it is surprising that the product of a single gene appears to account –at least to a significant extent– for the low NMD capacity of a human cell line. However, such a simple explanation of variable NMD efficiency is not general to all human cells because RNPS1 seems not to be involved in the low efficiency of strain C cells (Fig. 21). As a matter of fact, the reasons for the low NMD efficiency of this strain remain unknown.

These findings let us hypothesise that the availability of single or combined NMD factors may account for the efficiency of this pathway. The proteins involved might belong to the surveillance complex directly or could be part of the degradation machinery. For instance, it has been proposed that Vps16 – a protein that enhances decapping but that it is not essential – might be a candidate to regulate NMD in yeast (Zhang et al., 1999; Kebaara et al., 2003). Also in humans, dispensable factors that enhance mRNA turnover could play this role.

In conclusion, I have provided evidence that NMD varies in its efficiency in different human cell lines. In one particular case, a cause of a poor NMD phenotype was shown to be associated with the low abundance of an NMD factor (RNPS1) which is also involved in the UPF2-dependent NMD pathway. This finding provides a conceptual frame to understand

phenotypic variations in NMD-modulated diseases. It also suggests that NMD can be physiologically regulated by the accessibility of one or more NMD factors.

4.3 Variations in NMD efficiency – A potential new genetic modifier of disease

Different cell lines display differences in NMD efficiency. It is not surprising then to find important variations in the abundance of the five NMD reporter transcripts in blood samples (Fig. 24). Even for a single donor, the quantified variability may be attributed to the differences in cellular composition in each sample. In fact, it has long been known that many analytes such as uric acid, sugar and triglycerides vary greatly from one blood sample to another (Marnini and Zulian, 1976). For this reason, I decided to estimate NMD in lymphoblastoid cell lines, as this system allows a tighter control. Despite a lower variability and stronger correlation in the expression of the NMD targets, the assay did not demonstrate differences in NMD efficiency in these lymphocytic lines (Fig. 25).

One possible hypothesis to explain the variability of transcripts levels in the latter case would be that the chosen NMD-affected genes are not suitable for quantification in B-lymphocytes, possibly due to transcriptional regulation of the chosen genes. At present, no report confirms this hypothesis; however, GADD45B is known to be regulated in T-cells where it is critical for T-cell lineage development as well as for perpetuating cognate and inflammatory signals in these lymphocytes (Lu et al., 2004; Schwartz et al., 2006). Alternatively, the selected housekeeping genes could be suitable for HeLa but not for blood cells. Although some studies support my selection (Jin et al., 2004; Zhang et al., 2005), a more recent publication states that particularly GAPDH is not an appropriate housekeeping gene in EBV-transformed B-cells due to the variable quantification results obtained in microarray chips (de Brouwer et al., 2006). It remains for future work to assess the levels of the NMD transcripts using other genes for normalisation.

On the other hand, it is interesting to notice the similar mean quantified values for the different lymphocytic cell lines regardless of the aforementioned variability in the data (Fig. 24). From this point, an alternative potential reason for the failure to measure differences in NMD capacity would be that they simply do not exist in the tested lines. It has already been noticed that aberrant immunoglobulin transcripts are down-regulated very efficiently in lymphocytes. It has been estimated that two-thirds of immunoglobulin rearrangements are unproductive and generate PTCs (Li and Wilkinson, 1998). In this context, it is possible that lymphocytes have evolved under the selective pressure of conserving a strong NMD

efficiency. It has been also mentioned before that the effective down-modulation of immunoglobulin mRNAs is due to the presence of cis-acting elements. This observation is compatible with my hypothesis because an "NMD-enhancer" would be likely futile in cells with intrinsically low NMD efficiency. At the moment this hypothesis is mere speculation. However, it has been reported that while variable up-modulation levels of an NMD-sensitive transcript were quantified in different cells of two sibling foetus diagnosed with Roberts syndrome after treatment with puromycin, lymphocytic cell lines generated from their parents (who are genetically unrelated) showed similar levels in the same treatment (Resta et al., 2006). A comparison of NMD efficiency in different cell lines (probably by transfection of a common NMD reporter) would help to advance in the understanding of this issue.

4.3.1 NMD in β-thalassemia

β-thalassemia is caused by diverse mutations in the β-globin gene. One particular mutation, β-thalassemia (IVS1+6 T→C) was the focus of my study. A group of Palestinean and Israeli patients affected by β-thalassemia and who are all homozygous for this mutation was included in the present study. Interestingly, despite carrying the same mutation in the β-globin gene, these patients display a striking variability in anemia severity.

Two genetic modifiers of β-thalassemia, α-globin and γ-globin have been identified years ago. High fetal ($\alpha_2\gamma_2$) haemoglobin production can compensate for the lack of β-globin, thus ameliorating the disease (Rund et al., 1997). Aberrations in α-globin may modify the severity of β-thalassemia in both directions. Reduction in the number of functional α-globin genes usually causes a milder condition (Kulozik, 1992; Rund et al., 1997). On the other hand, additional functional α-globin genes in a β-thalassemia carrier may cause severe thalassemia intermedia (Kulozik et al., 1987; Oron et al., 1994). The presence of these modifiers was assayed in the selected group of patients by our collaborators. Their results indicated that the clinical variability of the anemia in these patients cannot be explained by these genetic modifiers and the causes of the phenotypic variability of the disease remain unknown.

The mutation that produces this anemia impairs the correct splicing of the first intron of the β-globin gene. As a consequence, 4 splice isoforms can be produced: β-globin wild type and 3 aberrant variants (Treisman et al., 1983). One such spliced variants was demonstrated to be NMD-sensitive (Danckwardt et al., 2002) and a second one is also a potential NMD target. As a consequence, it was envisioned that inter-interindividual

differences in NMD efficiency could account for the phenotypic variability of this particular thalassemia, with a more efficient NMD expected to lead to milder disease.

As the assay based in the quantification of NMD targets did not allowed estimation of varying NMD efficiency in normal blood samples, I had to rely exclusively in the correlation between the relative abundance of the β-globin variants and the severity of the anemia in each patient to determine whether NMD could be a potential modifier in this disease. The four expected isoforms were detected in patients´ blood samples by RPA analysis. The first observation was that the normal β-globin isoform accounted for 95-98% of the total β-globin mRNA while both exon1(-16) and exon1(-38) variants represented less than 1% and 1.5%, respectively. In principle, such minor proportions of aberrant isoforms could be considered irrelevant for the pathology of the disease. However, a recent study describes a PTC-containing transcript of the high-affinity immunoglobulin E (IgE) receptor (FcεRIβ) arising from alternative splicing which is expressed at very low levels compared to the full-length transcript, as would be expected for an NMD target. Nonetheless, the truncated protein is not only detectable but also it competes effectively with the full-length protein to control FcεRIβ expression on the cell surface (Donnadieu et al., 2003). Thus, even low endogenous expression levels of NMD targets can suffice to generate a product with a dominant negative function. In addition, the detected β-globin transcripts derived from anucleated red cells (mainly reticulocytes) present in blood. As there is no transcription in these cells and as NMD-sensitive β-globin isoforms decay faster than the normal transcript, the quantified proportions of β-globin variants may not correspond to those present in the nucleated red cell precursors at earlier stages.

This analysis also showed that the aberrant β-globin isoforms are expressed at highly variable levels in different patients regardless of the severity of their anemia (Fig. 26 and 27). Moreover, there was no evident correlation between the levels of these isoforms in individual patients either. Consequently, these results do not grant a general role of NMD in the phenotypic expression of β-thalassemia IVS1+6 (T→C). However, as my analysis is restricted to peripheral blood cells, it is not possible to exclude a potential role of NMD in normoblasts where transcription of the β-globin gene takes place. It is possible that the circulating reticulocytes represent a surviving fraction of the total original population due to their relatively low levels of aberrant transcripts. To test this hypothesis, bone marrow samples should be analysed; however, the invasiveness of the procedure of bone marrow sampling imposes a limit to the investigation.

As the causes of variable severity in the thalassemia remain elusive, alternative hypotheses should also be tested. For instance, a new possible genetic modifier has lately been described. The alpha-hemoglobin stabilising protein (AHSP) is a chaperone that binds and stabilises free α-globin, preventing its precipitation in the cell (Viprakasit et al., 2004). AHSP appears to modulate the clinical outcome of β-thalassemia in a murine model (Kong et al., 2004) and to be also a relevant contributory factor in some patients (Lai et al., 2006). However, AHSP expression in this set of patients has not been assayed.

My results also indicate that patients with severe thalassemia tend to express higher levels of all the aberrant isoforms (Fig. 26), albeit a larger number of cases should be studied to confirm this observation. In the context of this hypothesis, a change in the production of different β-globin splicing variants could account for differying abundances of normal and aberrant β-globin protein regardless of (or in addition to) differences in NMD activity. Unfortunately, abundance of mRNA (against a reference gene) and protein for each isoform has not been quantified yet.

In summary, postulating a broad role of NMD in the phenotypic outcome of β-thalassemia (IVS1+6) is unwarranted according to our studies. However, a larger number of cases should be analysed to obtain firm conclusions on the potential function of NMD in this disease. A more robust assay to estimate NMD activity, such as the one developed in this thesis, was not applied since it could not detect quantitative differences in NMD efficiency in normal blood samples. Nevertheless, since the application of this assay to immortalized B-lymphocytes gave reproducible results (although no difference in NMD efficiency was detected in the analysed lines), it could be used to assess NMD variability in EBV-transformed lymphocytes of these patients, as a complement to the RPAs.

Additionally, alternative factors that might act as novel genetic modifiers of β-thalassemia severity should be explored. Tentatively among these factors, AHSP and splicing factors could result promising candidates.

4.4 Concluding remarks

In the present study I aimed at developing an assay system to estimate quantitative differences in human NMD efficiency. Such a system was based in the quatification of the abundance of *bona fide* NMD transcripts. To this end, I undertook a combined RNAi/genome-wide microarray approach to identify human genes with increased expression

following depletion of the essential NMD factor UPF1. The outcome of this analysis suggests that NMD affects an ample number of activities in normal cellular function, although a large proportion of the genes identified is likely indirectly affected or is controlled by non-NMD UPF1-related novel functions

A careful validation of the microarray data allowed me to uncover and confirm novel *bona fide* NMD mRNAs. These transcripts were used to establish the before mentioned assay system. Using this assay, I demonstrate that diverged HeLa cell strains display large and stable differences in NMD efficiency. Moreover, one molecular mechanism that quantitatively affects NMD was uncovered, which represents a first example to explain variations in mammalian NMD efficiency.

A similar analysis in blood did not result in clearly interpretable data, possibly due to the higher complexity of this system. In addition, quantitative analysis of NMD-sensitive β-globin transcripts was performed in blood samples of patients suffering from β-thalassemia (IVS1+6). The results obtained so far do not support a role of NMD as a potential genetic modifier of this disease.

References

Abler, M. L. and P. J. Green (1996). "Control of mRNA stability in higher plants." Plant Mol Biol **32**(1-2): 63-78.

Alonso, C. R. (2005). "Nonsense-mediated RNA decay: a molecular system micromanaging individual gene activities and suppressing genomic noise." Bioessays **27**(5): 463-6.

Aloy, P., F. D. Ciccarelli, C. Leutwein, A. C. Gavin, G. Superti-Furga, P. Bork, B. Bottcher and R. B. Russell (2002). "A complex prediction: three-dimensional model of the yeast exosome." EMBO Rep **3**(7): 628-35.

Amrani, N., R. Ganesan, S. Kervestin, D. A. Mangus, S. Ghosh and A. Jacobson (2004). "A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay." Nature **432**(7013): 112-8.

Andrei, M. A., D. Ingelfinger, R. Heintzmann, T. Achsel, R. Rivera-Pomar and R. Luhrmann (2005). "A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies." Rna **11**(5): 717-27.

Araki, Y., S. Takahashi, T. Kobayashi, H. Kajiho, S. Hoshino and T. Katada (2001). "Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast." Embo J **20**(17): 4684-93.

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin and G. Sherlock (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet **25**(1): 25-9.

Azzalin, C. M. and J. Lingner (2006a). "The double life of UPF1 in RNA and DNA stability pathways." Cell Cycle. **5**(14): 1496-8. Epub 2006 Jul 17.

Azzalin, C. M. and J. Lingner (2006b). "The human RNA surveillance factor UPF1 is required for S phase progression and genome stability." Curr Biol **16**(4): 433-9.

Bakheet, T., B. R. Williams and K. S. Khabar (2003). "ARED 2.0: an update of AU-rich element mRNA database." Nucleic Acids Res **31**(1): 421-3.

Barreau, C., L. Paillard and H. B. Osborne (2005). "AU-rich elements and associated factors: are there unifying principles?" Nucleic Acids Res **33**(22): 7138-50.

Bashkirov, V. I., H. Scherthan, J. A. Solinger, J. M. Buerstedde and W. D. Heyer (1997). "A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates." J Cell Biol **136**(4): 761-73.

Bateman, J. F., S. Freddi, G. Nattrass and R. Savarirayan (2003). "Tissue-specific RNA surveillance? Nonsense-mediated mRNA decay causes collagen X haploinsufficiency in Schmid metaphyseal chondrodysplasia cartilage." Hum Mol Genet **12**(3): 217-25.

Baumann, B., M. Potash and G. Köhler (1985). "Consequences of frameshift mutations at the immunoglobulin heavy chain locus of the mouse." EMBO J. **4**: 351-359.

Beelman, C. A. and R. Parker (1995). "Degradation of mRNA in Eukaryotes." Cell **81**: 179-183.

Behm-Ansmant, I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork and E. Izaurralde (2006). "mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes." Genes Dev **20**(14): 1885-98.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon (2001). "Role for a bidentate ribonuclease in the initiation step of RNA interference." Nature **409**(6818): 363-6.

Bhattacharyya, S. N., R. Habermacher, U. Martine, E. I. Closs and W. Filipowicz (2006). "Relief of microRNA-mediated translational repression in human cells subjected to stress." Cell **125**(6): 1111-24.

Binder, R., J. A. Horowitz, J. P. Basilion, D. M. Koeller, R. D. Klausner and J. B. Harford (1994). "Evidence that the pathway of transferrin receptor mRNA degradation involves an endonucleolytic cleavage within the 3' UTR and does not involve poly(A) tail shortening." Embo J **13**(8): 1969-80.

Bono, F., J. Ebert, E. Lorentzen and E. Conti (2006). "The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA." Cell. **126**(4): 713-25.

Bousquet-Antonelli, C., C. Presutti and D. Tollervey (2000). "Identification of a regulated pathway for nuclear pre-mRNA turnover." Cell **102**(6): 765-75.

Brengues, M., D. Teixeira and R. Parker (2005). "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies." Science **310**(5747): 486-9.

Brewer, G. (1998). "Characterization of c-myc 3' to 5' mRNA decay activities in an in vitro system." J. Biol. Chem. **273**: 34770-34774.

Brown, B. D., I. D. Zipkin and R. M. Harland (1993). "Sequence-specific endonucleolytic cleavage and protection of mRNA in Xenopus and Drosophila." Genes Dev **7**(8): 1620-31.

Brown, C. E. and A. B. Sachs (1998). "Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation." Mol Cell Biol **18**(11): 6548-59.

Bruno, I. and M. F. Wilkinson (2006). "P-bodies react to stress and nonsense." Cell **125**(6): 1036-8.

Buhler, M., A. Paillusson and O. Muhlemann (2004). "Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5'-half of the VDJ exon." Nucleic Acids Res **32**(11): 3304-15.

Burkard, K. T. and J. S. Butler (2000). "A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p." Mol Cell Biol **20**(2): 604-16.

Butler, J. S. (2002). "The yin and yang of the exosome." Trends Cell Biol **12**(2): 90-6.

Cali, B. M. and P. Anderson (1998). "mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans." Mol Gen Genet **260**: 176-184.

Cao, D. and R. Parker (2003). "Computational modeling and experimental analysis of nonsense-mediated decay in yeast." Cell **113**(4): 533-45.

Capowski, E. E., S. Esnault, S. Bhattacharya and J. S. Malter (2001). "Y box-binding factor promotes eosinophil survival by stabilizing granulocyte-macrophage colony-stimulating factor mRNA." J Immunol **167**(10): 5970-6.

Carter, M. S., J. Doskow, P. Morris, S. Li, R. P. Nhim, S. Sandstedt and M. F. Wilkinson (1995). "A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts *in vivo* is reversed by protein synthesis inhibitors *in vitro*." J. Biol. Chem. **270**: 28995-29003.

Carthew, R. W. (2006). "Gene regulation by microRNAs." Curr Opin Genet Dev **16**(2): 203-8.

Chang, J. C. and Y. W. Kan (1979). "Beta 0 thalassemia, a nonsense mutation in man." Proc Natl Acad Sci USA **76**(6): 2886-2889.

Chang, J. C., G. F. Temple, R. F. Trecartin and Y. W. Kan (1979). "Suppression of the nonsense mutation in homozygous beta 0 thalassaemia." Nature **281**(5732): 602-3.

Chen, C. Y., R. Gherzi, S. E. Ong, E. L. Chan, R. Raijmakers, G. J. Pruijn, G. Stoecklin, C. Moroni, M. Mann and M. Karin (2001). "AU binding proteins recruit the exosome to degrade ARE-containing mRNAs." Cell **107**(4): 451-64.

Chen, C. Y. and A. B. Shyu (2003). "Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway." Mol Cell Biol **23**(14): 4805-13.

Chiu, S. Y., G. Serin, O. Ohara and L. E. Maquat (2003). "Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1." Rna **9**(1): 77-87.

Conti, E. and E. Izaurralde (2005). "Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species." Curr Opin Cell Biol **17**(3): 316-25.

Cougot, N., S. Babajko and B. Seraphin (2004). "Cytoplasmic foci are sites of mRNA decay in human cells." J Cell Biol **165**(1): 31-40.

Cuccurese, M., G. Russo, A. Russo and C. Pietropaolo (2005). "Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression." Nucleic Acids Res **33**(18): 5965-77.

Culbertson, M. R. and P. F. Leeds (2003). "Looking at mRNA decay pathways through the window of molecular evolution." Curr Opin Genet Dev **13**(2): 207-14.

Culbertson, M. R., K. M. Underbrink and G. R. Fink (1980). "Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors." Genetics **95**(4): 833-53.

Czaplinski, K., M. J. Ruiz.Echevarria, S. V. Paushkin, X. Han, Y. Weng, H. A. Perlick, D. H.C., M. D. Ter-Avanesyan and S. W. Peltz (1998). "The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs." Genes Dev. **12**: 1665-1677.

Czaplinski, K., Y. Weng, K. W. Hagan and S. W. Peltz (1995). "Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation." Rna **1**(6): 610-23.

Danckwardt, S., G. Neu-Yilik, R. Thermann, U. Frede, M. W. Hentze and A. E. Kulozik (2002). "Abnormally spliced beta-globin mRNAs: a single point mutation generates transcripts sensitive and insensitive to nonsense-mediated mRNA decay." Blood **99**(5): 1811-1816.

de Brouwer, A. P., H. van Bokhoven and H. Kremer (2006). "Comparison of 12 reference genes for normalization of gene expression levels in Epstein-Barr virus-transformed lymphoblastoid cell lines and fibroblasts." Mol Diagn Ther **10**(3): 197-204.

Delpy, L., C. Sirac, E. Magnoux, S. Duchez and M. Cogne (2004). "RNA surveillance down-regulates expression of nonfunctional kappa alleles and detects premature termination within the last kappa exon." Proc Natl Acad Sci U S A **101**(19): 7375-80.

Denis, C. L. and J. Chen (2003). "The CCR4-NOT complex plays diverse roles in mRNA metabolism." Prog Nucleic Acid Res Mol Biol **73**: 221-50.

Doma, M. K. and R. Parker (2006). "Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation." Nature **440**(7083): 561-4.

Donnadieu, E., M. H. Jouvin, S. Rana, M. F. Moffatt, E. H. Mockford, W. O. Cookson and J. P. Kinet (2003). "Competing functions encoded in the allergy-associated F(c)epsilonRIbeta gene." Immunity **18**(5): 665-74.

Dostie, J. and G. Dreyfuss (2002). "Translation is required to remove Y14 from mRNAs in the cytoplasm." Curr Biol **12**(13): 1060-7.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl (2001). "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells." Nature **411**(6836): 494-8.

El-Bchiri, J., O. Buhard, V. Penard-Lacronique, G. Thomas, R. Hamelin and A. Duval (2005). "Differential nonsense mediated decay of mutated mRNAs in mismatch repair deficient colorectal cancers." Hum Mol Genet **14**(16): 2435-42.

Englert, C., M. Vidal, S. Maheswaran, Y. Ge, R. M. Ezzell, K. J. Isselbacher and D. A. Haber (1995). "Truncated WT1 mutants alter the subnuclear localization of the wild-type protein." Proc Natl Acad Sci U S A **92**(26): 11960-4.

Enssle, J., W. Kugler, M. W. Hentze and A. E. Kulozik (1993). "Determination of mRNA fate by different RNA polymerase II promoters." Proc Natl Acad Sci U S A **90**(21): 10091-5.

Eulalio, A., I. Behm-Ansmant and E. Izaurralde (2007). "P bodies: at the crossroads of post-transcriptional pathways." Nat Rev Mol Cell Biol **8**(1): 9-22.

Eystathioy, T., E. K. Chan, S. A. Tenenbaum, J. D. Keene, K. Griffith and M. J. Fritzler (2002). "A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles." Mol Biol Cell **13**(4): 1338-51.

Fan, S., R. Yuan, Y. X. Ma, Q. Meng, I. D. Goldberg and E. M. Rosen (2001). "Mutant BRCA1 genes antagonize phenotype of wild-type BRCA1." Oncogene **20**(57): 8215-35.

Fan, X. C. and J. A. Steitz (1998). "Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs." Embo J **17**(12): 3448-60.

Ferraiuolo, M. A., C. S. Lee, L. W. Ler, J. L. Hsu, M. Costa-Mattioli, M. J. Luo, R. Reed and N. Sonenberg (2004). "A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay." Proc Natl Acad Sci U S A **101**(12): 4118-23.

Frevel, M. A., T. Bakheet, A. M. Silva, J. G. Hissong, K. S. Khabar and B. R. Williams (2003). "p38 Mitogen-activated protein kinase-dependent and -independent signaling of mRNA stability of AU-rich element-containing transcripts." Mol Cell Biol **23**(2): 425-36.

Fribourg, S. and E. Conti (2003). "Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15." EMBO Rep **4**(7): 699-703.

Frischmeyer, P. A. and H. C. Dietz (1999). "Nonsense-mediated mRNA decay in health and disease." Hum Mol Genet **8**: 1893-1900.

Frischmeyer, P. A., A. van Hoof, K. O'Donnell, A. L. Guerrerio, R. Parker and H. C. Dietz (2002). "An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination Codons." Science **295**(5563): 2258-2261.

Fu, X.-D. (1995). "The superfamily of arginine/serine-rich splicing factors." RNA **1**: 663-680.

Fukuhara, N., J. Ebert, L. Unterholzner, D. Lindner, E. Izaurralde and E. Conti (2005). "SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway." Mol Cell **17**(4): 537-47.

Gao, M., D. T. Fritz, L. P. Ford and J. Wilusz (2000). "Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro." Mol Cell **5**(3): 479-88.

Gao, M., C. J. Wilusz, S. W. Peltz and J. Wilusz (2001). "A novel mRNA-decapping activity in HeLa cytoplasmic extracts is regulated by AU-rich elements." Embo J **20**(5): 1134-43.

Gatfield, D. and E. Izaurralde (2004). "Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila." Nature **429**(6991): 575-8.

Gatfield, D., L. Unterholzner, F. D. Ciccarelli, P. Bork and E. Izaurralde (2003). "Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways." Embo J **22**(15): 3960-70.

Gehring, N. H., J. B. Kunz, G. Neu-Yilik, S. Breit, M. H. Viegas, M. W. Hentze and A. E. Kulozik (2005). "Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements." Mol Cell **20**(1): 65-75.

Gehring, N. H., G. Neu-Yilik, T. Schell, M. W. Hentze and A. E. Kulozik (2003). "Y14 and hUpf3b Form an NMD-Activating Complex." Mol Cell **11**(4): 939-49.

Glavan, F., I. Behm-Ansmant, E. Izaurralde and E. Conti (2006). "Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex." Embo J **25**(21): 5117-25.

Gonzalez, C. I., A. Bhattacharya, W. Wang and S. W. Peltz (2001). "Nonsense-mediated mRNA decay in Saccharomyces cerevisiae." Gene **274**(1-2): 15-25.

Gonzalez, C. I., M. J. Ruiz-Echevarria, S. Vasudevan, M. F. Henry and S. W. Peltz (2000). "The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay." Mol Cell **5**(3): 489-99.

Gudikote, J. P. and M. F. Wilkinson (2002). "T-cell receptor sequences that elicit strong down-regulation of premature termination codon-bearing transcripts." Embo J **21**(1-2): 125-34.

Hagan, K. W., M. J. Ruiz-Echevarria, Y. Quan and S. W. Peltz (1995). "Characterization of *cis*-acting sequences and decay intermediates involved in nonsense mediated mRNA turnover." Mol. Cell. Biol. **15**(2): 809-823.

Hall, G. W. and S. Thein (1994). "Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia." Blood **83**: 2031-2037.

He, F., X. Li, P. Spatrick, R. Casillo, S. Dong and A. Jacobson (2003). "Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast." Mol Cell **12**(6): 1439-52.

Hentze, M. W. and A. E. Kulozik (1999). "A perfect message: RNA surveillance and nonsense-mediated decay." Cell **96**(3): 307-10.

Hilleren, P. and R. Parker (1999). "mRNA surveillance in eukaryotes: Kinetic proofreading of proper translation termination as assessed by mRNP domain organization." RNA **5**: 711-719.

Hillman, R. T., R. E. Green and S. E. Brenner (2004). "An unappreciated role for RNA surveillance." Genome Biol **5**(2): R8.

Holbrook, J. A., G. Neu-Yilik, M. W. Hentze and A. E. Kulozik (2004). "Nonsense-mediated decay approaches the clinic." Nat Genet **36**(8): 801-8.

Hosack, D. A., G. Dennis, Jr., B. T. Sherman, H. C. Lane and R. A. Lempicki (2003). "Identifying biological themes within lists of genes with EASE." Genome Biol **4**(10): R70.

Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J Mol Biol **3**: 318-56.

Jensen, L. R., M. Amende, U. Gurok, B. Moser, V. Gimmel, A. Tzschach, A. R. Janecke, G. Tariverdian, J. Chelly, J. P. Fryns, H. Van Esch, T. Kleefstra, B. Hamel, C. Moraine, J. Gecz, G. Turner, R. Reinhardt, V. M. Kalscheuer, H. H. Ropers and S. Lenzner (2005). "Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation." Am J Hum Genet **76**(2): 227-36.

Jin, P., Y. Zhao, Y. Ngalame, M. C. Panelli, D. Nagorsen, V. Monsurro, K. Smith, N. Hu, H. Su, P. R. Taylor, F. M. Marincola and E. Wang (2004). "Selection and validation of endogenous reference genes using a high throughput approach." BMC Genomics **5**(1): 55.

Jing, Q., S. Huang, S. Guth, T. Zarubin, A. Motoyama, J. Chen, F. Di Padova, S. C. Lin, H. Gram and J. Han (2005). "Involvement of microRNA in AU-rich element-mediated mRNA instability." Cell **120**(5): 623-34.

Jung, D. and F. W. Alt (2004). "Unraveling V(D)J recombination; insights into gene regulation." Cell **116**(2): 299-311.

Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss and S. Ohno (2006). "Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex

(SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay." Genes Dev **20**(3): 355-67.

Kataoka, N., J. Yong, V. N. Kim, F. Velazquez, R. A. Perkinson, F. Wang and G. Dreyfuss (2000). "Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm." Mol Cell **6**(3): 673-82.

Kaygun, H. and W. F. Marzluff (2005). "Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1." Nat Struct Mol Biol **12**(9): 794-800.

Kebaara, B., T. Nazarenus, R. Taylor and A. L. Atkin (2003). "Genetic background affects relative nonsense mRNA accumulation in wild-type and upf mutant yeast strains." Curr Genet **43**(3): 171-7.

Kerr, T. P., C. A. Sewry, S. A. Robb and R. G. Roberts (2001). "Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay?" Hum Genet **109**(4): 402-7.

Khabar, K. S. (2005). "The AU-rich transcriptome: more than interferons and cytokines, and its role in disease." J Interferon Cytokine Res **25**(1): 1-10.

Khajavi, M., K. Inoue and J. R. Lupski (2006). "Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease." Eur J Hum Genet **14**(10): 1074-81.

Kim, V. N. (2001). "Role of the nonsense-mediated factor hUpf3 in the splicing-dependent exon-exon junction complex." Science **293**: 1833-1836.

Kim, V. N., J. Yong, N. Kataoka, L. Abel, M. D. Diem and G. Dreyfuss (2001). "The Y14 protein communicates to the cytoplasm the position of exon-exon junctions." Embo J **20**(8): 2062-8.

Kim, Y. K., L. Furic, L. Desgroseillers and L. E. Maquat (2005). "Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay." Cell. **120**(2): 195-208.

King-Underwood, L. and K. Pritchard-Jones (1998). "Wilms' tumor (WT1) gene mutations occur mainly in acute myeloid leukemia and may confer drug resistance." Blood **91**(8): 2961-8.

Kong, Y., S. Zhou, A. J. Kihm, A. M. Katein, X. Yu, D. A. Gell, J. P. Mackay, K. Adachi, L. Foster-Brown, C. S. Louden, A. J. Gow and M. J. Weiss (2004). "Loss of alpha-hemoglobin-stabilizing protein impairs erythropoiesis and exacerbates beta-thalassemia." J Clin Invest **114**(10): 1457-66.

Kornblihtt, A. R., M. de la Mata, J. P. Fededa, M. J. Munoz and G. Nogues (2004). "Multiple links between transcription and splicing." Rna **10**(10): 1489-98.

Kugler, W., J. Enssle, M. W. Hentze and A. E. Kulozik (1995). "Nuclear degradation of nonsense mutated beta-globin mRNA: a post-transcriptional mechanism to protect heterozygotes from severe clinical manifestations of beta-thalassemia?" Nucleic Acids Res **23**(3): 413-8.

Kulozik, A. E. (1992). "Beta-thalassaemia: molecular pathogenesis and clinical variability." Eur J Pediatr **151**(2): 78-84.

Kulozik, A. E., S. L. Thein, J. S. Wainscoat, R. Gale, L. A. Kay, J. K. Wood, D. J. Weatherall and E. R. Huehns (1987). "Thalassaemia intermedia: interaction of the triple alpha-globin gene arrangement and heterozygous beta-thalassaemia." Br J Haematol **66**(1): 109-12.

Kunz, J. B., G. Neu-Yilik, M. W. Hentze, A. E. Kulozik and N. H. Gehring (2006). "Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation." Rna **12**(6): 1015-22.

Kuzmiak, H. A. and L. E. Maquat (2006). "Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges." Trends Mol Med **12**(7): 306-16.

Lai, M. I., J. Jiang, N. Silver, S. Best, S. Menzel, A. Mijovic, S. Colella, J. Ragoussis, C. Garner, M. J. Weiss and S. L. Thein (2006). "Alpha-haemoglobin stabilising protein is

a quantitative trait gene that modifies the phenotype of beta-thalassaemia." Br J Haematol **133**(6): 675-82.

Lai, W. S., E. A. Kennington and P. J. Blackshear (2003). "Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease." Mol Cell Biol **23**(11): 3798-812.

Le Hir, H., D. Gatfield, E. Izaurralde and M. J. Moore (2001). "The exon-exon junction provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay." EMBO J **10**(17): 4987-4997.

Le Hir, H., E. Izaurralde, L. E. Maquat and M. J. Moore (2000). "The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions." EMBO J **19**(24): 6860-9.

Lee, M. H. and T. Schedl (2004). "Translation repression by GLD-1 protects its mRNA targets from nonsense-mediated mRNA decay in C. elegans." Genes Dev **18**(9): 1047-59.

Leeds, P., S. W. Peltz, A. Jacobson and M. R. Culbertson (1991). "The product of the yeast *UPF1* gene is required for rapid turnover of mRNAs containing a premature translational termination codon." Genes Dev. **5**: 2303-2314.

Leeds, P., J. M. Wood, B.-S. Lee and M. R. Culbertson (1992). "Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*." Mol. Cell. Biol. **12**(5): 2165-2177.

Lehner, B. and C. M. Sanderson (2004). "A protein interaction framework for human mRNA degradation." Genome Res **14**(7): 1315-23.

Lejeune, F., Y. Ishigaki, X. Li and L. E. Maquat (2002). "The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling." Embo J **21**(13): 3536-45.

Lejeune, F., X. Li and L. E. Maquat (2003). "Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities." Mol Cell **12**(3): 675-87.

Lelivelt, M. J. and M. R. Culbertson (1999). "Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome." Mol Cell Biol **19**(10): 6710-9.

Lewis, B. P., R. E. Green and S. E. Brenner (2003). "Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans." Proc Natl Acad Sci U S A **100**(1): 189-92.

Li, S., D. Leonard and M. F. Wilkinson (1997). "T cell receptor (TCR) mini-gene mRNA expression regulated by nonsense codons: a nuclear-associated translation-like mechanism." J.Exp.Med. **185**: 985-992.

Li, S. and M. F. Wilkinson (1998). "Nonsense surveillance in lymphocytes?" Immunity **8**: 135-141.

Linz, B., N. Koloteva, S. Vasilescu and J. E. McCarthy (1997). "Disruption of ribosomal scanning on the 5'-untranslated region, and not restriction of translational initiation per se, modulates the stability of nonaberrant mRNAs in the yeast Saccharomyces cerevisiae." J. Biol. Chem. **272**: 131-140.

Liu, H., N. D. Rodgers, X. Jiao and M. Kiledjian (2002). "The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases." Embo J **21**(17): 4699-708.

Liu, J., M. A. Valencia-Sanchez, G. J. Hannon and R. Parker (2005). "MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies." Nat Cell Biol **7**(7): 719-23.

Liu, Q., J. C. Greimann and C. D. Lima (2006). "Reconstitution, activities, and structure of the eukaryotic RNA exosome." Cell **127**(6): 1223-37.

Losson, R. and F. Lacroute (1979). "Interference of nonsense mutations with eukaryotic messenger RNA stability." Proc Natl Acad Sci U S A **76**(10): 5134-7.

Lu, B., A. F. Ferrandino and R. A. Flavell (2004). "Gadd45beta is important for perpetuating cognate and inflammatory signals in T cells." Nat Immunol **5**(1): 38-44.

Lykke-Andersen, J. (2002). "Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay." Mol Cell Biol **22**(23): 8114-21.

Lykke-Andersen, J., M. D. Shu and J. A. Steitz (2001). "Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1." Science **293**: 1836-1839.

Lykke-Andersen, J., M.-D. Shu and J. A. Steitz (2000). "Human Upf Proteins Target an mRNA for Nonsense-Mediated Decay When Bound Downstream of a Termination Codon." Cell **103**: 1121-1131.

Mango, S. E. (2001). "Stop making nonSense: the C. elegans smg genes." Trends Genet **17**(11): 646-53.

Maquat, L. E. (1995). "When cells stop making sense: effect of nonsense codons on RNA metabolism in vertebrate cells." RNA **1**: 453-465.

Maquat, L. E. (2004). "Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics." Nature Reviews Molecular Cell Biology **5**: 89-99.

Maquat, L. E., A. J. Kinniburgh, E. A. Rachmilewitz and J. Ross (1981). "Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia." Cell **27**(3 Pt 2): 543-53.

Maquat, L. E. and X. Li (2001). "Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay." RNA **7**(3): 445-456.

Maquat, L. E. and G. Serin (2001). "Nonsense-mediated mRNA decay: insights into mechanism from the cellular abundance of human Upf1, Upf2, Upf3, and Upf3X proteins." Cold Spring Harb Symp Quant Biol **66**: 313-20.

Marnini, P. and C. Zulian (1976). "[Relation between age, sex and various metabolic parameters: blood uric acid (4776 specimens), blood sugar (7211 specimens), blood cholesterol (5600 specimens) and blood triglycerides (4438 specimens)]." Quad Sclavo Diagn **12**(4): 416-30.

Mayeda, A., J. Badolato, R. Kobayashi, M. Q. Zhang, E. M. Gardiner and A. R. Krainer (1999). "Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing." Embo J **18**(16): 4560-70.

Meijer, H. A. and A. A. Thomas (2002). "Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA." Biochem J **367**(Pt 1): 1-11.

Meister, G. and T. Tuschl (2004). "Mechanisms of gene silencing by double-stranded RNA." Nature **431**(7006): 343-9.

Mello, C. C. and D. Conte, Jr. (2004). "Revealing the world of RNA interference." Nature **431**(7006): 338-42.

Mendell, J. T., C. M. ap Rhys and H. C. Dietz (2002). "Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts." Science **298**(5592): 419-22.

Mendell, J. T., S. M. Medghalchi, R. G. Lake, E. R. Noensie and H. C. Dietz (2000). "Novel Upf2p Orthologues Suggest a Funcional Link between Translation Initiation and Nonsense Surveillance Complexes." Molecular and Cellular Biology **20**(23): 8944-8957.

Mendell, J. T., N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo and H. C. Dietz (2004). "Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise." Nat Genet **36**(10): 1073-8.

Meyer, S., C. Temme and E. Wahle (2004). "Messenger RNA turnover in eukaryotes: pathways and enzymes." Crit Rev Biochem Mol Biol **39**(4): 197-216.

Millar, D. S., L. Elliston, P. Deex, M. Krawczak, A. I. Wacey, J. Reynaud, H. K. Nieuwenhuis, P. Bolton-Maggs, P. M. Mannucci, J. C. Reverter, P. Cachia, K. J. Pasi,

D. M. Layton and D. N. Cooper (2000). "Molecular analysis of the genotype-phenotype relationship in factor X deficiency." Hum Genet **106**(2): 249-57.

Mitchell, P. and D. Tollervey (2001). "mRNA turnover." Curr Opin Cell Biol **13**(3): 320-5.

Mitchell, P. and D. Tollervey (2003). "An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation." Mol Cell **11**(5): 1405-13.

Montgomery, M. K. (2004). "RNA interference: historical overview and significance." Methods Mol Biol **265**: 3-21.

Muhlrad, D., C. J. Decker and R. Parker (1995). "Turnover mechanisms of the stable yeast *PGK1* mRNA." Mol. Cell Biol. **15**: 2145-2156.

Mukherjee, D., M. Gao, J. P. O'Connor, R. Raijmakers, G. Pruijn, C. S. Lutz and J. Wilusz (2002). "The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements." Embo J **21**(1-2): 165-74.

Nagy, E. and L. Maquat (1998). "A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance." TIBS **23**: 198-199.

Namy, O., J. P. Rousset, S. Napthine and I. Brierley (2004). "Reprogrammed genetic decoding in cellular gene expression." Mol Cell **13**(2): 157-68.

Neu-Yilik, G., N. H. Gehring, R. Thermann, U. Frede, M. W. Hentze and A. E. Kulozik (2001). "Splicing and 3' end formation in the definition of nonsense-mediated decay-competent human beta-globin mRNPs." Embo J **20**(3): 532-40.

Nielsen, F. C. and J. Christiansen (1992). "Endonucleolysis in the turnover of insulin-like growth factor II mRNA." J Biol Chem **267**(27): 19404-11.

Ohnishi, T., A. Yamashita, I. Kashima, T. Schell, K. R. Anders, A. Grimson, T. Hachiya, M. W. Hentze, P. Anderson and S. Ohno (2003). "Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7." Mol Cell **12**(5): 1187-200.

Olesen, J. R., D. Libri and T. H. Jensen (2005). "A link between transcription and mRNP quality in Saccharomyces cerevisiae." RNA Biol. **2**(2): 45-8. Epub 2005 Apr 18.

Oron, V., D. Filon, A. Oppenheim and D. Rund (1994). "Severe thalassaemia intermedia caused by interaction of homozygosity for alpha-globin gene triplication with heterozygosity for beta zero-thalassaemia." Br J Haematol **86**(2): 377-9.

Page, M., B. Carr, K. A. Anders, A. Grimson and P. Anderson (1999). "SMG-2 is a phosphorylated protein required for mRNA surveillance in *Caenorhabditis elegans* and related to Upf1p of yeast." Mol. Cell. Biol. **19**: 5943-5951.

Pal, M., Y. Ishigaki, E. Nagy and L. E. Maquat (2001). "Evidence that phosphorylation of human Upfl protein varies with intracellular location and is mediated by a wortmannin-sensitive and rapamycin-sensitive PI 3-kinase-related kinase signaling pathway." Rna **7**(1): 5-15.

Palacios, I. M., D. Gatfield, D. St Johnston and E. Izaurralde (2004). "An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay." Nature **427**(6976): 753-7.

Parker, R. and H. Song (2004). "The enzymes and control of eukaryotic mRNA turnover." Nat Struct Mol Biol **11**(2): 121-7.

Patton, M. A. and A. R. Afzal (2002). "Robinow syndrome." J Med Genet **39**(5): 305-10.

Perez Jurado, L. A., Y. K. Wang, U. Francke and J. Cruces (1999). "TBL2, a novel transducin family member in the WBS deletion: characterization of the complete sequence, genomic structure, transcriptional variants and the mouse ortholog." Cytogenet Cell Genet **86**(3-4): 277-84.

Perlick, H. A., S. M. Medghalchi, F. A. Spencer, R. J. Kendzior and H. C. Dietz (1996). "Mammalian orthologues of a yeast regulator of nonsense transcript stability." Proc. Natl. Acad. Sci. USA **93**: 10928-10932.

Perrin-Vidoz, L., O. M. Sinilnikova, D. Stoppa-Lyonnet, G. M. Lenoir and S. Mazoyer (2002). "The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons." Hum Mol Genet **11**(23): 2805-14.

Pham, J. W., J. L. Pellino, Y. S. Lee, R. W. Carthew and E. J. Sontheimer (2004). "A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila." Cell **117**(1): 83-94.

Proudfoot, N. (2004). "New perspectives on connecting messenger RNA 3' end formation to transcription." Curr Opin Cell Biol **16**(3): 272-8.

Pulak, R. and P. Anderson (1993). "mRNA surveillance by the Caenorhabditis elegans smg genes." Genes Dev **7**(10): 1885-97.

Reddy, J. C., J. C. Morris, J. Wang, M. A. English, D. A. Haber, Y. Shi and J. D. Licht (1995). "WT1-mediated transcriptional activation is inhibited by dominant negative mutant proteins." J Biol Chem **270**(18): 10878-84.

Reed, R. (2003). "Coupling transcription, splicing and mRNA export." Curr Opin Cell Biol **15**(3): 326-31.

Rehwinkel, J., I. Letunic, J. Raes, P. Bork and E. Izaurralde (2005). "Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets." Rna **11**(10): 1530-44.

Resta, N., F. C. Susca, M. C. Di Giacomo, A. Stella, N. Bukvic, R. Bagnulo, C. Simone and G. Guanti (2006). "A homozygous frameshift mutation in the ESCO2 gene: evidence of intertissue and interindividual variation in Nmd efficiency." J Cell Physiol **209**(1): 67-73.

Rosenfeld, P. J., G. S. Cowley, T. L. McGee, M. A. Sandberg, E. L. Berson and T. P. Dryja (1992). "A null mutation in the rhodopsin gene causes rod photoreceptor dysfunction and autosomal recessive retinitis pigmentosa." Nat Genet **1**(3): 209-13.

Ross, J. (1995). "mRNA stability in mammalian cells." Microbiol. Rev. **59**: 423-450.

Ruiz-Echevarria, M. J., K. Czaplinski and S. W. Peltz (1996). "Making sense of nonsense in yeast." Trends Biochem Sci **21**: 433-438.

Ruiz-Echevarria, M. J., C. I. Gonzalez and S. W. Peltz (1998). "Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA." EMBO J **17**: 575-589.

Ruiz-Echevarria, M. J. and S. W. Peltz (2000). "The RNA binding protein Pub 1 modulates the stability of transcripts containing upstream open reading frames." Cell **101**: 741-751.

Rund, D., V. Oron-Karni, D. Filon, A. Goldfarb, E. Rachmilewitz and A. Oppenheim (1997). "Genetic analysis of beta-thalassemia intermedia in Israel: diversity of mechanisms and unpredictability of phenotype." Am J Hematol **54**(1): 16-22.

Sakashita, E., S. Tatsumi, D. Werner, H. Endo and A. Mayeda (2004). "Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo." Mol Cell Biol **24**(3): 1174-87.

Schneppenheim, R., U. Budde, T. Obser, J. Brassard, K. Mainusch, Z. M. Ruggeri, S. Schneppenheim, R. Schwaab and J. Oldenburg (2001). "Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease." Blood **97**(7): 2059-66.

Schwartz, R., I. Engel, M. Fallahi-Sichani, H. T. Petrie and C. Murre (2006). "Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development." Proc Natl Acad Sci U S A **103**(26): 9976-81.

Serin, G., A. Gersappe, J. D. Black, R. Aronoff and L. E. Maquat (2001). "Identification and characterisation of human orthologues to *Sacharomyces cerevisae* Upf2 proteins and

Upf3 protein (*Caenorhabditis elegans* SMG-4)." Molecular and Cellular Biology **21**(1): 209-223.

She, M., C. J. Decker, K. Sundramurthy, Y. Liu, N. Chen, R. Parker and H. Song (2004). "Crystal structure of Dcp1p and its functional implications in mRNA decapping." Nat Struct Mol Biol **11**(3): 249-56.

Sheth, U. and R. Parker (2003). "Decapping and decay of messenger RNA occur in cytoplasmic processing bodies." Science **300**(5620): 805-8.

Sheth, U. and R. Parker (2006). "Targeting of aberrant mRNAs to cytoplasmic processing bodies." Cell **125**(6): 1095-109.

Shibuya, T., T. O. Tange, N. Sonenberg and M. J. Moore (2004). "eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay." Nat Struct Mol Biol. **11**(4): 346-51. Epub 2004 Mar 21.

Shim, J. and M. Karin (2002). "The control of mRNA stability in response to extracellular stimuli." Mol Cells **14**(3): 323-31.

Singh, G. and J. Lykke-Andersen (2003). "New insights into the formation of active nonsense-mediated decay complexes." Trends Biochem Sci **28**(9): 464-6.

Spingola, M., L. Grate, D. Haussler and M. Ares, Jr. (1999). "Genome-wide bioinformatic and molecular analysis of introns in Saccharomyces cerevisiae." Rna **5**(2): 221-34.

Stockklausner, C., S. Breit, G. Neu-Yilik, N. Echner, M. W. Hentze, A. E. Kulozik and N. H. Gehring (2006). "The uORF-containing thrombopoietin mRNA escapes nonsense-mediated decay (NMD)." Nucleic Acids Res **34**(8): 2355-63.

Stoeckle, M. Y. and H. Hanafusa (1989). "Processing of 9E3 mRNA and regulation of its stability in normal and Rous sarcoma virus-transformed cells." Mol Cell Biol **9**(11): 4738-45.

Strausberg, R. L., E. A. Feingold, L. H. Grouse, J. G. Derge, R. D. Klausner, F. S. Collins, L. Wagner, C. M. Shenmen, G. D. Schuler, S. F. Altschul, B. Zeeberg, K. H. Buetow, C. F. Schaefer, N. K. Bhat, R. F. Hopkins, H. Jordan, T. Moore, S. I. Max, J. Wang, F. Hsieh, L. Diatchenko, K. Marusina, A. A. Farmer, G. M. Rubin, L. Hong, M. Stapleton, M. B. Soares, M. F. Bonaldo, T. L. Casavant, T. E. Scheetz, M. J. Brownstein, T. B. Usdin, S. Toshiyuki, P. Carninci, C. Prange, S. S. Raha, N. A. Loquellano, G. J. Peters, R. D. Abramson, S. J. Mullahy, S. A. Bosak, P. J. McEwan, K. J. McKernan, J. A. Malek, P. H. Gunaratne, S. Richards, K. C. Worley, S. Hale, A. M. Garcia, L. J. Gay, S. W. Hulyk, D. K. Villalon, D. M. Muzny, E. J. Sodergren, X. Lu, R. A. Gibbs, J. Fahey, E. Helton, M. Ketteman, A. Madan, S. Rodrigues, A. Sanchez, M. Whiting, A. C. Young, Y. Shevchenko, G. G. Bouffard, R. W. Blakesley, J. W. Touchman, E. D. Green, M. C. Dickson, A. C. Rodriguez, J. Grimwood, J. Schmutz, R. M. Myers, Y. S. Butterfield, M. I. Krzywinski, U. Skalska, D. E. Smailus, A. Schnerch, J. E. Schein, S. J. Jones and M. A. Marra (2002). "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc Natl Acad Sci U S A **99**(26): 16899-903.

Sueoka, N. (1993). "Directional mutation pressure, mutator mutations, and dynamics of molecular evolution." J Mol Evol **37**(2): 137-53.

Sun, X. and L. E. Maquat (2002). "Nonsense-mediated decay: assaying for effects on selenoprotein mRNAs." Methods Enzymol **347**: 49-57.

Sun, X., P. M. Moriatry and L. E. Maquat (2000). "Nonsense-mediated decay of gluthatione peroxidase 1 mRNA in the cytoplasm depends on intron position." EMBO **19**(17): 4734-4744.

Sun, X., H. A. Perlick, H. C. Dietz and L. E. Maquat (1998). "A mutated human homologue to yeast Upf1 protein has a dominant-negative effect on the decay of nonsense-containing mRNAs in mammalian cells." Proc Natl Acad Sci U S A **95**(17): 10009-14.

Sung, C. H., C. M. Davenport, J. C. Hennessey, I. H. Maumenee, S. G. Jacobson, J. R. Heckenlively, R. Nowakowski, G. Fishman, P. Gouras and J. Nathans (1991). "Rhodopsin mutations in autosomal dominant retinitis pigmentosa." Proc Natl Acad Sci U S A **88**(15): 6481-5.

Sureau, A., R. Gattoni, Y. Dooghe, J. Stevenin and J. Soret (2001). "SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs." Embo J **20**(7): 1785-96.

Sylvain, V., S. Lafarge and Y. J. Bignon (2002). "Dominant-negative activity of a Brca1 truncation mutant: effects on proliferation, tumorigenicity in vivo, and chemosensitivity in a mouse ovarian cancer cell line." Int J Oncol **20**(4): 845-53.

Symmons, M. F., M. G. Williams, B. F. Luisi, G. H. Jones and A. J. Carpousis (2002). "Running rings around RNA: a superfamily of phosphate-dependent RNases." Trends Biochem Sci **27**(1): 11-8.

Takekawa, M. and H. Saito (1998). "A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK." Cell **95**(4): 521-30.

Tange, T. O., A. Nott and M. J. Moore (2004). "The ever-increasing complexities of the exon junction complex." Curr Opin Cell Biol **16**(3): 279-84.

Tange, T. O., T. Shibuya, M. S. Jurica and M. J. Moore (2005). "Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core." Rna. **11**(12): 1869-83.

Teixeira, D., U. Sheth, M. A. Valencia-Sanchez, M. Brengues and R. Parker (2005). "Processing bodies require RNA for assembly and contain nontranslating mRNAs." Rna **11**(4): 371-82.

Thein, S. L., C. Hesketh, P. Taylor, I. J. Temperley, R. M. Hutchinson, J. M. Old, W. G. Wood, J. B. Clegg and D. J. Weatherall (1990). "Molecular basis for dominantly inherited inclusion body beta-thalassemia." Proc Natl Acad Sci USA **87**: 3924-3928.

Thermann, R., G. Neu-Yilik, A. Deters, U. Frede, K. Wehr, C. Hagemeier, M. W. Hentze and A. E. Kulozik (1998). "Binary specification of nonsense codons by splicing and cytoplasmic translation." EMBO J **17**(12): 3484-3494.

Tourriere, H., K. Chebli and J. Tazi (2002). "mRNA degradation machines in eukaryotic cells." Biochimie **84**(8): 821-37.

Tran, H., M. Schilling, C. Wirbelauer, D. Hess and Y. Nagamine (2004). "Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU." Mol Cell **13**(1): 101-11.

Treisman, R., S. H. Orkin and T. Maniatis (1983). "Specific transcription and RNA splicing defects in five cloned β-thalassaemia genes." Nature **302**: 591-596.

Tucker, M. and R. Parker (2000). "Mechanisms and control of mRNA decapping in Saccharomyces cerevisiae." Annu Rev Biochem **69**: 571-95.

Tucker, M., M. A. Valencia-Sanchez, R. R. Staples, J. Chen, C. L. Denis and R. Parker (2001). "The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae." Cell **104**(3): 377-86.

Uchida, N., S. Hoshino and T. Katada (2004). "Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein." J Biol Chem **279**(2): 1383-91.

van Dijk, E., N. Cougot, S. Meyer, S. Babajko, E. Wahle and B. Seraphin (2002). "Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures." Embo J **21**(24): 6915-24.

van Dijk, E., H. Le Hir and B. Seraphin (2003). "DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway." Proc Natl Acad Sci U S A **100**(21): 12081-6.

van Hoof, A., P. A. Frischmeyer, H. C. Dietz and R. Parker (2002). "Exosome-mediated recognition and degradation of mRNAs lacking a termination codon." Science **295**(5563): 2262-4.

Vasudevan, S. and S. W. Peltz (2003). "Nuclear mRNA surveillance." Curr Opin Cell Biol **15**(3): 332-7.

Viprakasit, V., V. S. Tanphaichitr, W. Chinchang, P. Sangkla, M. J. Weiss and D. R. Higgs (2004). "Evaluation of alpha hemoglobin stabilizing protein (AHSP) as a genetic modifier in patients with beta thalassemia." Blood **103**(9): 3296-9.

Wang, J., V. M. Vock, S. Li, O. R. Olivas and M. F. Wilkinson (2002). "A quality control pathway that down-regulates aberrant T-cell receptor (TCR) transcripts by a mechanism requiring UPF2 and translation." J Biol Chem **277**(21): 18489-93.

Wang, Z. and M. Kiledjian (2001). "Functional link between the mammalian exosome and mRNA decapping." Cell **107**(6): 751-62.

Weischenfeldt, J., J. Lykke-Andersen and B. Porse (2005). "Messenger RNA surveillance: neutralizing natural nonsense." Curr Biol **15**(14): R559-62.

Weng, Y., K. Czaplinski and S. Peltz (1996). "Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the upf1 protein." Mol. Cell. Biol. **16**: 5477-5490.

Wittmann, J., E. M. Hol and H. M. Jack (2006). "hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay." Mol Cell Biol **26**(4): 1272-87.

Wollerton, M. C., C. Gooding, E. J. Wagner, M. A. Garcia-Blanco and C. W. Smith (2004). "Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay." Mol Cell **13**(1): 91-100.

Worton, R. G. (1992). "Duchenne muscular dystrophy: gene and gene product; mechanism of mutation in the gene." J Inherit Metab Dis **15**(4): 539-50.

Yamashita, A., T. Ohnishi, I. Kashima, Y. Taya and S. Ohno (2001). "Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay." Genes Dev. **15**(17): 2215-2228.

Zhang, S., C. J. Williams, K. Hagan and S. W. Peltz (1999). "Mutations in VPS16 and MRT1 stabilize mRNAs by activating an inhibitor of the decapping enzyme." Mol Cell Biol **19**(11): 7568-76.

Zhang, X., L. Ding and A. J. Sandford (2005). "Selection of reference genes for gene expression studies in human neutrophils by real-time PCR." BMC Mol Biol **6**(1): 4.

Zhang, Z. and A. R. Krainer (2004). "Involvement of SR proteins in mRNA surveillance." Mol Cell **16**(4): 597-607.

Zhao, Z., F. C. Chang and H. M. Furneaux (2000). "The identification of an endonuclease that cleaves within an HuR binding site in mRNA." Nucleic Acids Res **28**(14): 2695-701.

Acknowledgements

First of all, I want to thank my supervisors Prof. Dr. Andreas Kulozik and Prof. Dr. Matthias Hentze (MMPU-EMBL) for giving me the opportunity of undertaking this challenging project and for their scientific advice.

I am also grateful to all the members of the laboratory. I am especially in debt to Dr. Niels Gehring who supervised my daily work, provided me with plasmid constructs and protocols and, most important, with advice in our frequent scientific discussions. He also supported me in non-scientific related matters such as translating documents and letters from German into English. I also want to acknowledge Dr. Sven Danckwardt and Dr. Pavel "Pasha" Ivanov with whom I had interesting discussions (both scientific and non-scientific ones) and a lot of support, particularly in the most difficult moments of this project. Special thanks go also to Dr. Jill Holbrook for providing me with a lot of literature at the beginning of my project and for critically reading the thesis and to Dr. Gabriele Neu-Yilik who provided me with "logistic" information.

I would also like to thank Prof. Dr. Iain Mattaj, Prof. Dr. Claus Bartram and Dr. Eileen Furlong who accomplished an excellent task as members of my thesis advisory committee.

During these almost four years I had the opportunity of making new friends in Heidelberg. Unfortunately, I cannot acknowledge all of them here. Nevertheless, I want to say "thank you" at least to Carmelo Lopez-Portilla, Luis Vacs, Angel Chong, Dr. Andres Gaytan de Ayala Alonso, Dr. Eduardo Garcia Urdiales and Dr**.** Maximiliano Gutierrez from EMBL, and Elsa Harraez-Hernandez and Ton Garreta from DKFZ. I also want to mention my Italian friends Federico Moretti, Valentina Giovacchini and Dr. Martino Donnini and my Argentinean "super-friends" Alejandro Goldemberg, Carolina Mansilla, Dr. Valeria Rivarola, Dr. Miguel Urga, Dr. Florencia Labombarda and Lionel Rossi. Thank you all for your friendship!

Last but not least, many thanks go to my whole beloved family without which this doctoral thesis would have been impossible. Thanks you all for helping me to fulfil a dream.

Appendix

Parts of the thesis were published in:

Viegas MH, Gehring NH, Breit S, Hentze MW, Kulozik AE. "RNPS1 abundance determines variability of NMD efficiency". Submitted.

Gehring NH, Kunz JB, Neu-Yilik G, Breit S, Viegas MH, Hentze MW, Kulozik AE. "Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements". Mol Cell. 2005 Oct 7;20(1):65-75.

Poster or oral presentations including parts of the thesis:

"Physiological human NMD targets display differential requirements for hUPF1, hUPF2 and hUPF3a/b". Viegas MH, Gehring NH, Neu-Yilik G, Breit S, Hentze MW, Kulozik AE. 9th Annual Meeting of the RNA Society, p.487, 2004.

"RNPS1 abundance determines variability of NMD efficiency". Viegas MH, Gehring NH, Breit S, Hentze MW, Kulozik AE. Submitted to the 11th Annual Meeting of the RNA Society (Madison, USA).