Post-transcriptional Regulation of Gene Expression by Rod1

Tiago França Brazão

The work presented in this thesis was performed at the department of Cell Biology at the Erasmus MC in Rotterdam.

The cover represents the alternative splicing of a transcript, mediated by a protein complex assembled at the polypyrimidine tract. The representation is fictional, and was inpired by the Erasmusbrug in Rotterdam.

Printed by: Off Page

ISBN: 9789490371371

Post-transcriptional Regulation of Gene Expression by Rod1

Post-transcriptionale Regulatie van Genexpressie door Rod1

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the

rector magnificus

Prof.dr. H.G. Schmidt

and in accordance with the decision of the Doctorate Board

The public defense shall be held on

Wednesday 6 October 2010 at 11.30 o'clock

by

Tiago França Brazão
born in Lisbon, Portugal

Doctoral Committee

Promotor:	Prof.dr. F.G. Grosveld
Other Members:	Prof.dr. J.N.J. Philipsen Dr. J. H. Gribnau Dr. M. W. J. Fornerod

Aos meus pais e aos meus irmãos

Contents

List of abbreviations

AS	alternative splicing
AS-NMD	alternative splicing-coupled nonsense-mediated mRNA decay
CLIP-seq	cross-linked immunoprecipitation-sequencing
CMV	cytomegalovirus
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EJC	exon junction complex
ESE	exonic splice enhancer
ESS	exonic splice silencer
HEK	human embryonic kidney
hnRNP	heterogeneous ribonucleoprotein
Ig	immunoglobulin
IP	immunoprecipitation
IRES	internal ribosome entry site
ISE	intronic splice enhancer
ISS	intronic splice silencer
mRNA	messenger ribonucleic acid
NMD	nonsense-mediated mRNA decay
NS	nonsense
ORF	open reading frame
pol	polymerase
polyA	poly-adenosine
PPT	polypyrimidine tract
PTC	premature termination codon
RNA	ribonucleic acid
RNAi	RNA interference
RRM	RNA recognition motif
RT-PCR	real time polymerase chain reaction
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
shRNA	short hairpin RNA
snRNP	small nuclear ribonucleoprotein
SS	splice site
UTR	untranslated region
wt	wild-type

Scope of this thesis

The process of gene expression consists of several interconnected steps which ensure proper protein production from a specific gene. However, an extensive amount of transcripts containing premature termination codons (PTCs) are generated in cells, mainly through alternative splicing events. If translated, these transcripts could generate truncated proteins which are potentially deleterious to the cell. Nonsense-Mediated mRNA Decay (NMD) is a mechanism of mRNA degradation which eliminates these faulty transcripts, thus protecting the cell against abnormal protein functions. Chapter 2 of this thesis describes experiments aimed at finding new protein factors involved in NMD. Rod1 is found to be required for NMD in HEK293 cells and to regulate several potential NMD targets.

In chapter 3, the role of Rod1 in erythroid cells, where it is highly expressed, is addressed. Rod1 is found to regulate Alternative Splicing (AS) together with Ptbp1 and Raver1. AS is a mechanism that regulates inclusion or skipping of thousands of exons and is therefore an important generator of protein diversity in metazoans. AS is regulated by a large number of factors, including core spliceosomal proteins and proteins that activate or repress specific exons. AS is regulated in a tissue-specific manner, and some AS factors, like Rod1, are expressed in restricted tissues as well. Chapter 3 also contains experiments aimed at finding RNAs associated with Rod1 and Ptbp1 in differentiated MEL (murine erythroleukemia) cells.

In chapter 4, results from chapters 2 and 3 are discussed, as well as its implications. Future directions are also presented.

Chapter 1

Introduction

Introduction

The biological characteristics of cells, tissues, organs and ultimately organisms, denominated as "phenotypes", result mostly from gene expression profiles. These are different across different tissues and thus make an erythrocyte, for example, different from a neuron. The qualitative and quantitative differences of protein expression among all cells in a given organism allow each tissue to have its own function and identity, and are ultimately responsible for the orchestrated functioning and behavior of each individual. The varied phenotypes seen among individuals of the same species are also explained by (subtle) qualitative and quantitative differences in gene expression. Finally, gene expression can also change within a specific tissue across time (during development or through ageing), or in response to environmental signals, making each organism a highly complex, dynamic and unique entity.

The central dogma of molecular biology, first proposed by Francis Crick in 1958 and re-stated in 1970 [1] had as a main concept the idea that genetic information flows from DNA into RNA into protein. This idea is in its essence still valid, but this flow of information is much more complex than thought at that time: we now know that a given gene can give rise to multiple proteins, as opposed to a single protein, as it was then believed. This is mostly the case with more complex organisms and particularly humans, for which there is evidence that 92-94% of genes undergo alternative splicing (AS)[2]. Furthermore, RNA molecules are not necessarily translated into protein; many non-coding RNA genes have been identified which play important cellular roles in translation [3-5], transcription [6-8], splicing [9-12] and gene silencing [13-15]. The RNA molecule is thus not only a messenger, i.e. a carrier of information, but also operates in other functions which were until recently assigned to proteins alone.

The process of gene expression in eukaryotes comprises a sum of several spatially and temporally interconnected steps which (in the vast majority of cases) ensure proper protein (or RNA) production from a specific gene. The main steps are DNA transcription, pre-mRNA splicing, mRNA export and translation. However, depending on the particular gene and tissue, we may add DNA recombination, RNA editing, alternative splicing, mRNA transport, mRNA degradation, translation arrest, post-translation modifications and protein degradation. These steps act in different ways and combinations upon different genes in a given cell or tissue, but they are also different for a specific gene in different tissues, thus permitting phenotypic diversity across the many tissues and organs that make up the

organism. Each of these steps, rather than being autonomous, is influenced by a variety of cellular mechanisms and even extra-cellular factors. For example, transcription is influenced by chromatin organization and condensation, which in turn is influenced by epigenetic marks, which in turn can be influenced by environmental factors, etc. Similarly, splicing is intimately associated with transcription: these two processes not only happen simultaneously but also share common regulatory factors.

This thesis will focus on a mechanism of mRNA degradation termed Nonsense-Mediated mRNA Decay (NMD) and Alternative Splicing (AS).

Nonsense-Mediated mRNA Decay

Nonsense-Mediated mRNA Decay (NMD) is a cellular mechanism responsible for the selected degradation of transcripts containing premature termination codons (PTCs). PTCs can arise not only through nonsense or frameshift mutations but also through transcription or splicing errors or, more often, through alternative splicing (AS) events. Since AS is regulated in a tissue-specific manner, NMD targets different subsets of transcripts in different tissues and thus contributes to proteome diversity across different cell types. As a general rule, NMD degrades transcripts containing PTCs located more than 50-55 nt upstream of the last exon-exon junction. An exception to this rule is found in transcripts containing PTCs located near the 5' end which escape NMD, an effect caused by the proximity of polyA-binding protein (PABP) to the PTC [16, 17]. NMD also degrades mRNAs containing upstream open reading frames (uORF's), transcripts with introns in the 3'UTR and transcripts with long 3'UTR's [18-21]. In the last two cases, although there is no PTC *per se*, the *wild type* termination codon is interpreted by the cell as being premature, thus rendering the transcript subject to degradation. NMD has an important physiological role since it prevents the production of truncated proteins, which might have deleterious effects on the cell, particularly the ones with dominant negative effects. Its importance *in vivo* is exemplified by the Upf1 knock-out mice, which are NMD-impaired and not viable [22]. Many genetic diseases have been described which are caused by mutations that target the corresponding transcripts to NMD; in some cases therapeutic approaches to impair NMD and restore protein production can be beneficial, so NMD has also been a target for pharmaceutical drugs and clinical trials.

Historical overview

The discovery of the Nonsense-Mediated mRNA Decay (NMD) mechanism had its origins in 1979, when yeast ura3 transcripts containing nonsense mutations were found to be destabilized [23]. In humans, the first observations related to NMD took place in the beginning of 1980's; back then, a β-thalassemia patient was observed to express lower levels of β-globin mRNA due to mRNA degradation [24]. This observation was soon correlated to the presence of a nonsense codon in the corresponding transcript [25]. Nonsense mutations or other mutations that give rise to premature termination codons (PTCs) have been since then identified in many genes, causing several genetic diseases [26]. The first elucidation of the NMD mechanism was in 1980, in a phenotypic screen for frameshift mutant supressors in yeast [27]; in this study, the authors identified mutations that increased frameshift suppression, which mapped to the Upf1 and Upf2 loci. Later, Upf1 was found to be necessary for the rapid turnover of mRNAs containing PTCs [28]. The roles of Upf2, Upf3 and several proteins that comprise the Exon-Junction Complex (EJC) were later characterized as essential players in NMD [29-33]. More recently, the group of Andreas Kulozik described two different subcomplexes triggering NMD, one being dependent on Upf2 and the other independent of Upf2 [34]. In the last 5 years, more details concerning the NMD mechanism have been elucidated [35, 36].

Several studies analyzing gene expression after RNA interference against components of the NMD machinery prompted a major change in the way scientists in the field viewed NMD [20, 37-40]. A considerable amount of naturally occurring transcripts in yeast, fruitfly and human cells were found to be targeted for degradation by NMD. Therefore, NMD could no longer be described only as a surveillance mechanism defending organisms against nonsense or frameshift mutations, which could give rise to potentially harmful truncated proteins, but had to be regarded as an important mechanism of gene regulation, affecting thousands of natural physiological transcripts. Indeed, it was estimated that around one third of the transcriptome contains PTCs. Furthermore, since NMD was found to be coupled to alternative splicing events, it acted in a tissue-specific manner [41, 42].

The mechanism of Nonsense-Mediated mRNA Decay

During transcription by RNA pol II the nascent transcripts are prevented from being degraded by several modifications. Upon transcription initiation, transcripts are capped at the 5' end with a methylated guanosine connected by a 5'-5' triphosphate linkage

(m^7GpppN) by 7-guanyl methyl-transferase, protecting them from 5' to 3' exonucleases. As the RNA polymerase progresses, a plethora of splicing factors, many of them bound to the C-terminal tail of RNA pol II, gradually assemble in and around introns to remove them in a multi-step process termed splicing. During transcription termination, the RNA molecule is cleaved and subsequentially polyadenylated by polyA adenylase. The polyA tail is recognized and bound by polyA-binding protein (PABP), which, by interacting with the nuclear cap-binding complex (composed of CBP20 and CBP80 in the nucleus and eIF4G in the cytoplasm), circularizes the mature transcript [43-45]. All these events not only contribute to the maturation of the transcript but also protect it from degradation.

Upon splicing, a high molecular weight (around 700 kDa) protein complex is assembled around 24 nucleotides upstream of each exon-exon boundary [46-48]. This complex, denominated Exon-junction complex (EJC), is composed of several protein factors involved in different processes such as mRNA export, polysome association and mRNA degradation [30, 31, 33, 49]. During the pioneer round of translation all the EJC's bound to the mRNA (not subjected to NMD) are displaced by the ribosomes [50]. When the ribosome reaches the termination codon, there are no further EJCs bound to the transcript and therefore NMD does not occur. After the pioneer round of translation the cap-binding complex is substituted by eIF4E, resulting in more efficient subsequent rounds of translation (fig.1). This ensures that only transcripts that have been subjected to quality control (i.e., the pioneer round of translation) are efficiently translated. In transcripts containing PTCs positioned more than ~50-55 nt upstream of the last exon-exon boundary, the ribosome will encounter a termination codon before reaching the last EJC, which remains bound to the mRNA. The eukaryotic release factors eRF3 and eRF1 are then recruited to the paused ribosome together with Upf1 and SMG1 (SURF complex) [51]; Upf1 binds the downstream EJC by interacting with Upf2, thus bridging the EJC and the ribosome complexes. The interaction of Upf1 with Upf2 is promoted by CBP80 [52]. Upf1 association with the EJC triggers phosphorylation of Upf1 by SMG1, a step promoted by both Upf2 and Upf3b and regulated by both SMG8 and SMG9 [40, 51, 53, 54]. Phosphorylated Upf1 both inhibits translation (by inhibiting eRF1 and eRF3) and recruits SMG7, which in turn recruits proteins necessary for the mRNA degradation steps [55-58]. SMG7, SMG6 and SMG5 are necessary for dephosphorylation of Upf1 through recruitment of PP2A (protein phosphatase 2A), thus allowing recycling of Upf1 for more rounds of NMD [59]. Transcript degradation involves de-adenylation and de-capping activities; both destabilize the transcript, which can then be degraded from the 3'end by the exosome or from the 5'end by Xrn1 [60]. De-capping is mediated by Dcp1 and Dcp2 which, similarly to Xrn1, are enriched in cytoplasmic bodies

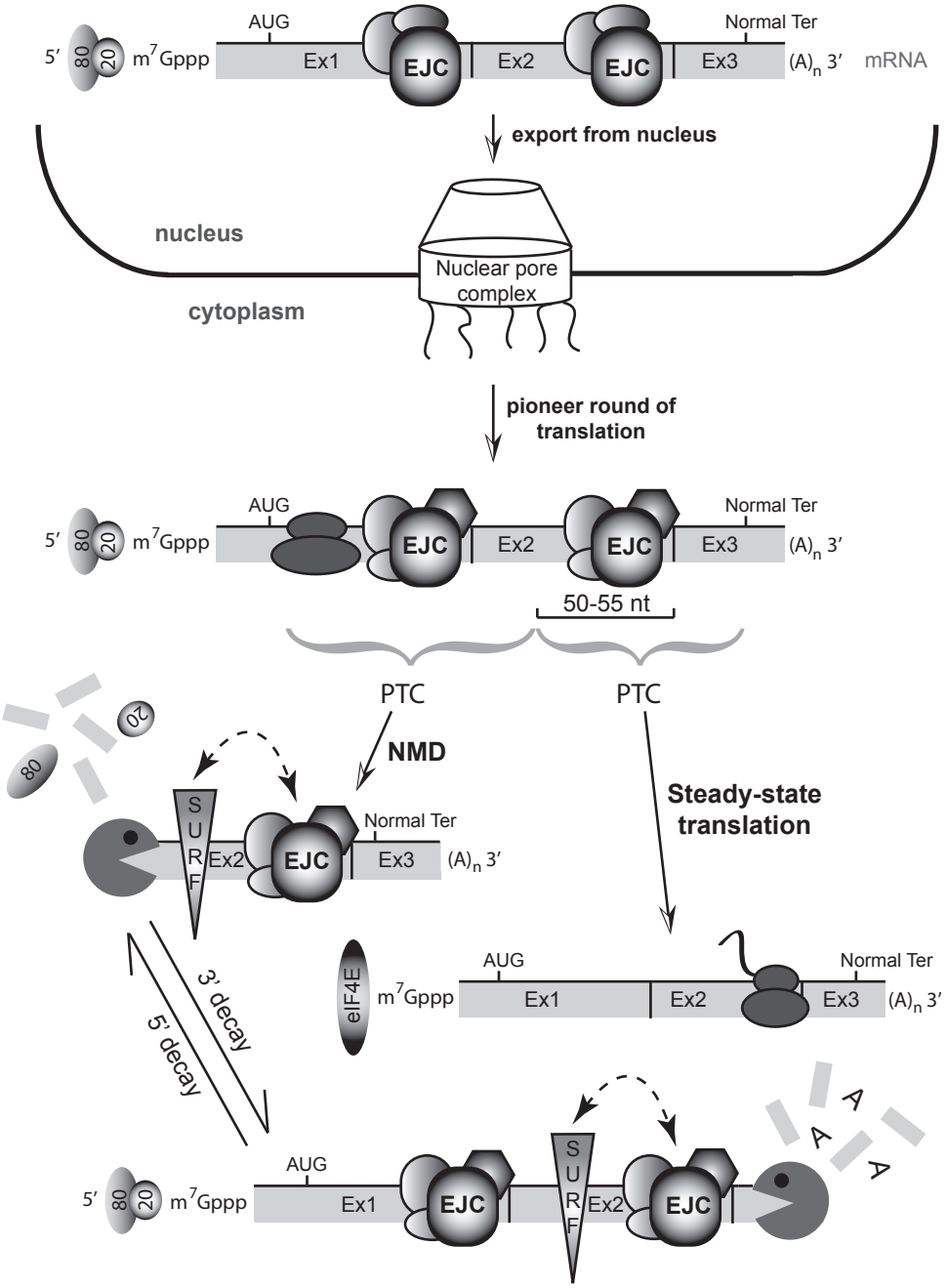


Figure 1 – Schematic representation of the Nonsense-Mediated mRNA Decay (NMD) mechanism; after mRNA export, transcripts are stripped off their EJC's in the pioneer round of translation and engage into steady-state

translation. PTC-containing transcripts can be discriminated as NMD targets or NMD-immune depending on the localization of the PTC. Figure adapted from [53]

termed P-bodies. This fact raises the question of whether these are sites of active transcript degradation by NMD. While this is still a possibility, P-bodies do not appear to be essential for NMD, as disruption of (visible) P-bodies has been shown not to affect NMD [61]. Recently, endonucleolytic cleavage of mRNA near the PTC has also been reported as a possible route for degradation; this pathway requires SMG6 and has been described in humans and Drosophila [62, 63].

The EJC complex

The Exon-Junction Complex, or EJC, associates with transcripts upon splicing around 24 nucleotides (nt) upstream of each exon-exon junction [64]. The assembly of this macromolecular complex occurs co-transcriptionally and requires proper capping and splicing reactions [65]. It functions as a molecular mark locating the positions of the spliced introns; this information is carried to the cytoplasm and is necessary for discrimination between non-PTC and PTC-containing transcripts. Apart from this, the EJC also enhances polysome association and facilitates mRNA export [30, 49]. For this reason, intronless genes are generally expressed in lower amounts than intron-containing ones, and are immune to (EJC-dependent) NMD. The EJC is composed of several proteins which together add up to around 700 kDa (Table 1). Y14, magoh, eIF4AIII, MLN51/Btz, RNPS1, SRm160, Aly/REF, UAP56, Pinin, TAP, PYM, Upf3a, Upf3b, Upf2 and Upf1 have all been reported to associate with the EJC [65, 67, 68]. However, it seems plausible that the composition of the EJC may vary depending on the cell type, transcript or even exon-exon junction; indeed, tethering experiments suggest that different EJC subcomplexes may exist, one including Upf2 (and RNPS1) and another excluding Upf2 [34]. Interestingly, the same group found a correlation between abundance of RNPS1 and efficiency of NMD [69]. This suggests that the two described NMD-triggering pathways (Upf2-dependent and Upf2-independent) could have different efficiencies. One could envisage that some cell types, with higher abundance of RNPS1, would preferentially use the more efficient Upf2-dependent pathway, while others, with lower expression of RNPS1, would preferentially use the less efficient Upf2-independent pathway. This could help to explain the observed varied NMD efficiencies in different cell lines [70]. Additional support for a Upf2-independent NMD pathway came from a recent finding in which a Upf1 mutant fails to interact with Upf2 and yet is still phosphorylated (presumably by SMG1) and functional in NMD [58].

EJC factor	Molecular Funtion	Ref.
Y14	together with magoh, locks the EJC to RNA by inhibiting ATPase activity of eIF4AIII; Y14 and magoh bind PYM in the cytoplasm, which triggers EJC disassembly	[71, 72]
Magoh	in Drosophila, it is required together with Btz for proper embryonic development by localizing oskar mRNA to the posterior pole of the oocyte; together with Y14, magoh binds PYM in the cytoplasm, which triggers EJC disassembly; magoh and Y14 "lock" the EJC to RNA by inhibiting the ATPase activity of eIF4AIII	[71-73]
Aly/REF	mRNA export factor through interaction with TAP; it is a member of the TREX complex which mediates mRNA export	[74-76]
eIF4AIII	necessary for the recruitment of Y14 and magoh to the RNA transcript; "locks" the EJC to the RNA	[71, 77]
Upf3a	competes with Upf3b for binding Upf2; weaker NMD activator than Upf3b	[78]
Upf3b	with Upf2, stimulates the RNA helicase activity of Upf1; competes with Upf3a for binding Upf2 and it's a stronger NMD activator than Upf3a	[78, 79]
RNPS1	promotes splicing by enhancing formation of the spliceosome A-complex; enhances polysome association	[49, 80]
SRm160	pre-mRNA splicing co-activator and regulator of alternative splicing; promotes transcript 3' end cleavage	[81-83]
UAP56	recruits Aly/REF to the EJC; UAP56, Aly/REF and the THO complex form the TREX complex which mediates mRNA export	[74, 75, 84]
MLN51/Btz	stimulates translation; in Drosophila, it is required together with magoh for proper embryonic development by localizing *oskar* mRNA to the posterior pole of the oocyte	[67, 68, 73]
Pinin	regulates alternative splicing; this activity is associated with a complex that includes RNPS1	[85, 86]
NFX1/TAP	interacts directly with Aly/REF and with FG-containing nucleoporins in the NPC (Nuclear Pore Complex); participates in mRNA export	[30]
Upf2	together with Upf3b, stimulates the RNA helicase activity of Upf1; in tethering experiments, Upf2 is necessary for RNPS1-mediated RNA degradation, but not for Y14 or eIF4AIII-mediated RNA degradation	[34, 79]
PYM	binds the Y14-magoh heterodimer in the cytoplasm, bridging the EJC to the 40S subunit of the ribosome; promotes translation of the mRNA and EJC disassembly	[72, 87]
Upf1	inhibits translation and recruits SMG7 during NMD, triggering mRNA degradation events; also participates in histone mRNA degradation, telomere maintenance, Staufen-mediated decay (SMD) and miRNA-mediated gene silencing	[88-92]

Table 1 - Protein factors described to associate with the Exon-Junction Complex (EJC) and their molecular function

The Upf factors (Upf3a, Upf3b, Upf2 and Upf1) play a crucial role in NMD by bridging the EJC to RNA degradation machineries. Not surprisingly, Upf1 is the most conserved factor [93, 94]; it is a predominantly cytoplasmic RNA helicase and has a nucleic-acid-dependent ATPase activity, essential to modulate structural changes in the RNA or RNA-protein complexes (fig. 2) [95]. It is a phosphoprotein, and cycles of phosphorylation/dephosphorylation are required for multiple rounds of NMD; its serine-glutamine (SQ) motifs near the C-terminal region are phosphorylated by SMG1 [53]. Like Upf1, Upf2 is a phosphoprotein, although the role of its phosphorylation is still unclear in mammals [96]. Despite being mostly a perinuclear protein, Upf2 contains nuclear localization signals (NLSs) in its N-terminus [97, 98]. Interestingly, although *in vivo* Upf2 is solely regarded as a molecular bridge between Upf3 and Upf1, *in vitro* it has RNA-binding activity [99]. Upf3 is the least conserved of the Upf factors [94]. In humans, Upf3 is encoded by two genes, Upf3a and Upf3b (also called Upf3X because it is localized in the X chromosome) [32, 98]. Both proteins are predominantly nuclear but shuttle to the cytoplasm, since both contain a conserved export signal located in the N-terminus. Despite being quite similar proteins, Upf3b is more efficient at promoting translation and triggering NMD than Upf3a [100]. Upf3a and Upf3b compete for binding Upf2; in the case of Upf3a, this interaction also stabilizes the protein, as "free" Upf3a is inherently unstable [78]. Binding to a composite site of the EJC core proteins Y14, magoh and eIF4AIII is mediated by the C-terminal region of both Upf3a and Upf3b [101, 102].

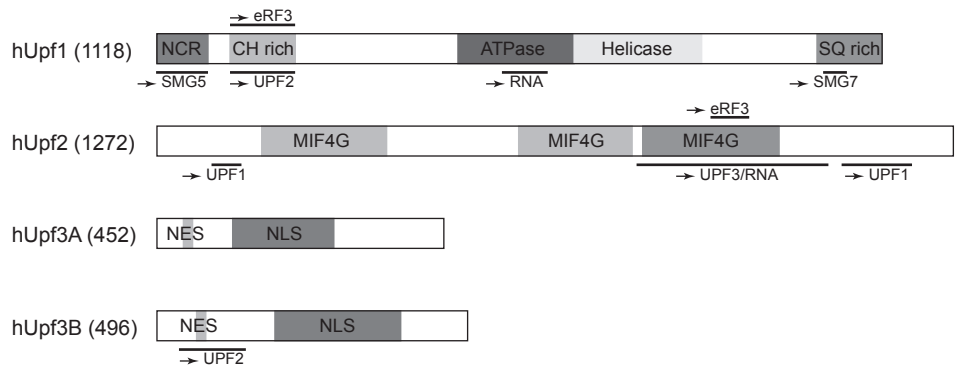


Figure 2 – Schematic view of protein domains from human Upf factors. Proteins are drawn to scale in respect to their aminoacid numbers, indicated between brackets; direct interactions with other proteins or RNA are indicated by black bars; reported interactions for which domains have not yet been mapped are not indicated (eg. Upf1-SMG1). hUpf1 contains an N-terminal conserved region (NCR), a CH-rich domain, a core domain comprising the ATPase and helicase, and an SQ-rich C-terminal domain. hUpf2 contains three mammalian eIF4G-like (MIF4G) domains. Both hUpf3a and hUpf3b contain nuclear export signals (NESs) and nuclear localization signals (NLSs). Figure adapted from [103].

NMD influences the clinical outcome of disease

Around one third of inherited or acquired genetic diseases are attributed to the presence of PTCs in the affected genes [104]. The selective degradation of transcripts carrying PTCs by NMD avoids synthesis of truncated polypeptides. These are generally unfavourable to the organism, but in some cases producing a truncated protein might be more beneficial than completely lacking the protein. Occasionaly a truncated polypeptide is still completely or partially functional, and therefore inhibiting NMD could alleviate the clinical outcome of disease. Patients with homozygous NMD-sensitive mutations could benefit from these treatments. Inhibition of NMD may be achieved by RNA interference (RNAi) against NMD effectors such as Upf2; RNAi against Upf1 could have unwanted negative consequences because Upf1 participates in mechanisms other than NMD, such as SMD or replication-dependent histone mRNA degradation. In a recent study, inhibition of NMD has also been used in mice tumors in order to stimulate immune responses to protein products which are normally under NMD control [105]. Inhibition of NMD is not an adequate therapy when PTCs give rise to truncated polypeptides that acquire dominant negative functions; for these cases, there have been attempts to suppress the nonsense codon, in order to restore synthesis of the full-length protein. NMD-sensitive PTCs cause degradation of the corresponding transcripts to about 5-25% of the levels of their *wild-type* counterparts [64], so combining inhibition of NMD with suppressing the PTC could be doubly beneficial. However, both these approaches would present several problems: global inhibition of NMD could have adverse effects, as many truncated proteins would be produced, since NMD targets many natural physiological transcripts (mostly derived from alternative splicing events or, in the case of cells from the immune system, from DNA rearrangements or hypermutations). Suppression of the nonsense codon, on the other hand, could suppress normal termination codons as well, giving rise to C-terminal extended proteins, which are potentially toxic [106]. Another problem from using nonsense-suppression therapies is the unknown consequence of failing to degrade naturally occurring NMD targets [64]. Nevertheless, these therapies have already been used in patients with some success. Particularly, the use of aminoglycoside antibiotics such as gentamicin has proved efficient in treating diseases like Cystic Fibrosis (CF) and has also been used to suppress PTCs within transcripts associated with muscular dystrophy, Hurler syndrome, cystinosis, polycystic kidney disease, X-linked nephrogenic diabetes insipidus and recessive spinal muscular atrophy [64].

NMD-insensitive PTCs, generally located in the last exon, can also give rise to dominant negative protein functions. Such is the case with some heterozygous β-thalassemia patients, who produce a non-functional β-globin polypeptide which causes toxic precipitation

of insoluble β-chains [107]. Heterozygous β-thalassemia patients with NMD-sensitive mutations, in contrast, are usually asymptomatic: the mutant allele produces no toxic protein and the remaining allele produces enough protein for the cell's needs. The same pattern of modulation of disease severity by NMD is seen in a variety of conditions including retinal degeneration [108, 109], factor X deficiency [110], susceptibility to mycobacterial infections [111, 112] and von Willebrand disease [113].

NMD, evolution and the emergence of highly complex proteomes

NMD has been positively selected during eukaryotic evolution to avoid the production of truncated proteins; these can be non-functional or even originate negative dominant phenotypes by acquiring new harmful functions. Disruption of NMD has been shown to affect normal development of several organisms. In *C. elegans*, the male bursa and hermaphrodite vulva develop abnormally resulting in decreased fecundity [114]. In zebrafish, embryonic development and survival is precluded upon depletion of NMD factors [115]. In mice, a Upf1 knock-out confers embryonic post-implantation lethality [22]. Similarly, Upf2 knock-out embryos are not viable at day 9,5 post coitus and conditional Upf2 knock-out cells from haematopoietic precursors fail to generate haematopoietic stem and progenitor cells [116]. Not surprisingly, the main NMD effectors are conserved across eukaryotes. In addition, NMD probably allowed other co-evolving cellular mechanisms such as alternative splicing (AS) to evolve into more complex and more flexible ways: AS events generating transcripts containing PTCs become "safe" with the emergence of NMD, since these mRNAs are targeted for degradation. Indeed, there is evidence in *Drosophila* that NMD has allowed intron gain across time [117]. Thus not only was NMD selected during evolution, but it also contributed to evolution by allowing AS to "experiment" more freely. This created the possibility to generate a wider variety of proteins from the same number of genes, a feature present in more complex organisms. To date, analysis of the homology of NMD factors such as the Upf proteins has shown that the organisms lacking NMD also lack AS [94].

Alternative Splicing

The coding information of metazoan genes is interrupted by introns, which are removed from the transcripts through the process of splicing [118, 119]. Alternative Splicing (AS) is the process by which multiple transcripts are generated from one single gene. AS was first described in 1980, when membrane-bound and secreted antibodies were found to be coded

by the same gene [120, 121]. In addition to the spliceosome, a macromolecular complex responsible for the removal of introns from the pre-mRNA, AS requires additional factors which can activate or inhibit splicing at certain locations, thus modifying the sequence of the mature transcript, giving rise to different protein isoforms. Around 75% of AS events occur in translated regions of mRNAs, modifying the primary structure of proteins [122, 123]. This can change properties such as binding affinity, intracellular localization, enzymatic activity or protein stability [124].

AS is widespread in metazoans and is differentially regulated across different cell types, contributing to the diversity of proteomes and phenotypes found in different tissues. Regulation of AS is also influenced by developmental or differentiation-specific cues [125-127]. Once regarded as a rare event, AS is now known to affect expression of most genes: particularly in humans, 92-94% of genes are subject to AS [2], some of them generating thousands of different transcripts [128, 129]. It seems likely that the known number of AS-derived transcripts is still underestimated, as more and more transcripts are identified with next generation sequencing (NGS) studies.

Rather than being an independent event, AS is intimately related to other nuclear processes, particularly transcription. Indeed, there is extensive evidence that the C-terminus of RNA polymerase II (RNAP II) provides a platform for numerous splicing factors [130-133]. Furthermore, the kinetics of RNAP II has been shown to influence exon inclusion or skipping: specific splicing enhancers or trans-acting factors which enhance the rate of RNAP II elongation favor exon skipping, while elongation inhibitor dichlororibofuranosylbenzimidazole (DRB) favors exon inclusion [134, 135]; it has also been observed that slowing down of transcription following UV-induced DNA damage changes patterns of AS [136]. Finally, nucleosome positioning correlates with intron-exon architecture, with exons displaying a 1.5 fold higher nucleosome occupancy than introns, presumably because of a higher GC content [137-139]. While the functional relevance of this observation needs further clarification, it is thought that nucleosome accumulation in exonic locations function as “speed bumps”, slowing down RNAP II and thus favoring exon inclusion [140].

AS generates transcript variants in several ways (fig. 3). In an estimated one third of AS events, PTCs are generated which render, in most cases, the corresponding transcripts subject to degradation by NMD [42]. This so called RUST (or AS-NMD) mechanism (Regulated Unproductive Splicing and Translation) is particularly common in splicing factors: an excess of the splicing factor shifts AS of its own transcript to generate the NMD-sensitive variant, thus decreasing the expression of the protein. This creates a feed-back

mechanism that keeps expression of the splicing factor within a narrow range [141-144]. In addition to these auto-regulatory feed-back loops, cross-repression of splicing factors through AS-NMD has also been reported for TIA-1, TIAR, PTBP1, nPTB and ROD1 [125, 126, 145]. This tight regulation of splicing and AS factors is physiologically important, as disruption of splicing or AS is associated with several disorders such as myotonic dystrophy, spinal muscular atrophy, retinitis pigmentosa, Frasier syndrome, atypical cystic fibrosis and several types of cancer [146-148].

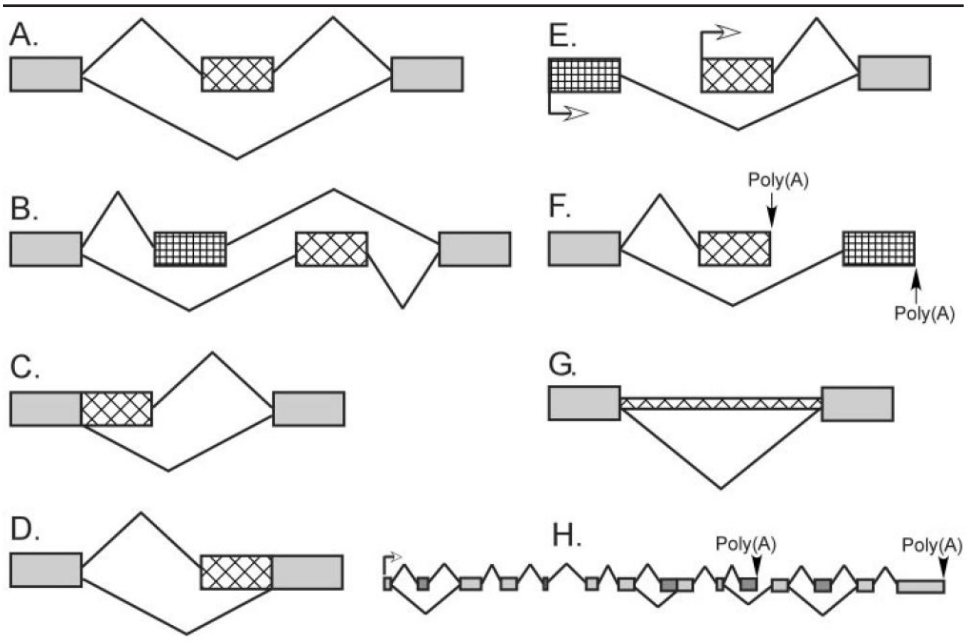


Figure 3 – Patterns of Alternative Splicing. Constitutive sequences present in all final mRNAs are shown as gray boxes. Alternative RNA segments that may or may not be included in the mRNA are shown as hatched boxes. (A) A cassette exon can be either included or excluded in the mRNA. (B) Mutually exclusive exons – only one exon in the group is included in the mRNA at a time. (C, D) Alternative 5' or 3' splice sites allow the lengthening or shortening of a particular exon. (E,F) Alternative promoters and alternative poly-A sites switch the 5' or 3' ends of transcripts. (G) Introns can be retained in the mRNA. (H) A single pre-mRNA can exhibit multiple sites of alternative splicing using different patterns of inclusion. These are often used in a combinatorial manner to produce many different final mRNAs. Figure taken from [149].

Regulation of Alternative Splicing

The main steps of splicing are directed by conserved short sequence motifs called splice sites (fig. 4). These are the intronic 5' splice site (5' SS), which includes a GU dinucleotide, the branch point, followed by a polypyrimidine tract and the intronic 3' splice site (3' SS), with an AG dinucleotide. The U1 snRNP binds the 5' SS by basepairing between conserved

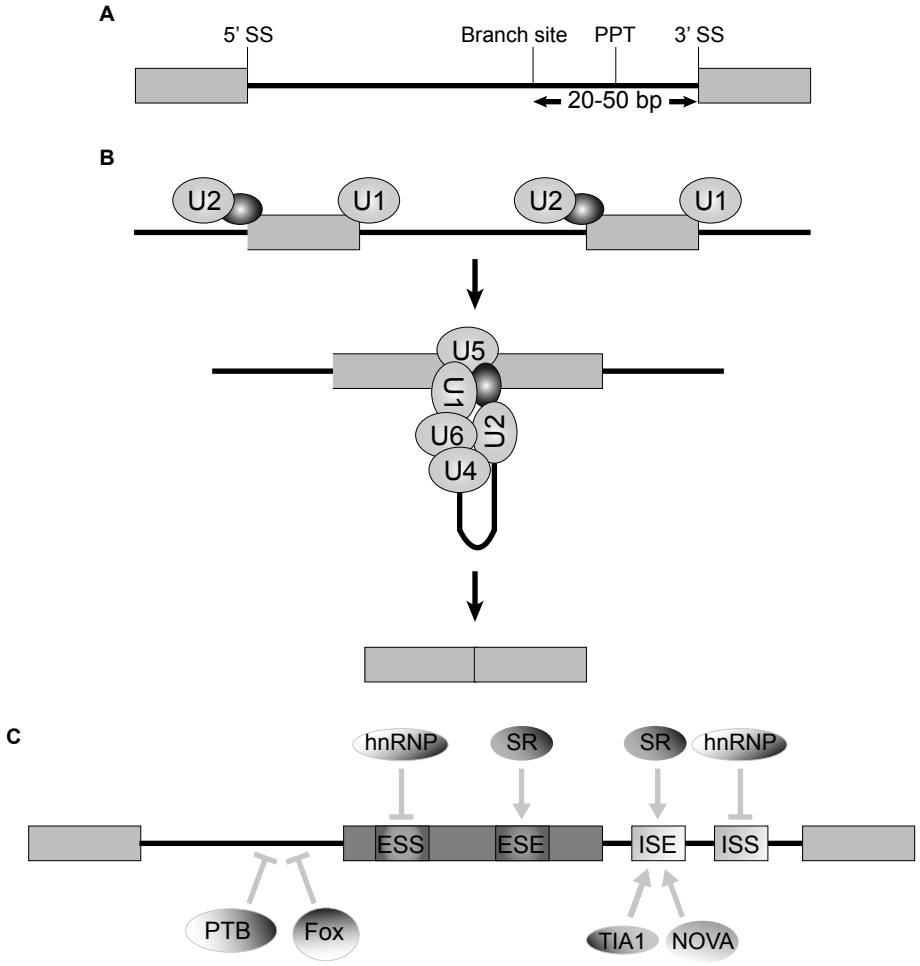


Figure 4 – The main splicing components and trans-acting factors. **a –** the intronic 5' splice site (5' SS), the branch point, polypyrimidine tract (PPT) and the 3' splice site (3' SS) direct binding of splicing factors and facilitate the splicing reaction steps. **b –** the U1 and U2 snRNPs bind to the 5' SS and branch site respectively; binding to the splice sites can be subject to regulation by other unspecified splicing components (unlabelled orange ovals). **c –** exons and introns contain exonic splicing enhancers (ESE), exonic splicing silencers (ESS), intronic splicing enhancers (ISE) and intronic splicing silencers (ISS); these regulatory sequences are bound by the serine/arginine (SR) proteins (regulating enhancers) and heterogenous ribonucleoproteins (hnRNPs – regulating silencers); TIA1 and the Nova proteins are additional regulators of ISEs, while the polypyrimidine tract-binding protein (PTB) and Fox proteins regulate binding of spliceosome components to the polypyrimidine tract and other locations of the intron. Figure adapted from [140].

bases in the 5' SS and the U1 snRNA. Splicing factor 1 (SF1) binds the branch point while the U2 auxiliary factor heterodimer (U2AF35 and U2AF65) binds the polypyrimidine tract (PPT) and the 3' SS [150-152]. This stage constitutes the E (early) complex, which assembles independently of ATP. SR proteins are known to facilitate assembly of the E

complex by binding intronic or exonic splicing enhancers (ISEs and ESEs) and mediating protein-protein interactions, which require phosphorylation and dephosphorylation cycles of their arginine/serine (RS) domains [153-155]. The spliceosomal A complex is formed upon substitution of SF1 at the branch point for U2 snRNP. Further recruitment of the U4/U6-U5 snRNP complex leads to the formation of the B complex. Through a number of rearrangements, including replacement of U1 snRNP for U6 snRNP at the 5' SS and loss of U1 and U4 snRNPs, the B complex is converted to the C (catalytic) complex, which will mediate the two trans-esterification reactions of splicing [149, 156]. In the first transesterification step, a nucleophylic attack involving the branch point and the 5' SS causes the cleavage of the 5' exon, with the subsequent ligation of the 5' SS to the branch point (lariat structure – fig. 5). The second transesterification step involves a nucleophylic attack of the 3' end (hydroxyl group) of the 5' exon to the 3' phosphate group of the intron (at the 3' SS); this leaves the two exons joined together, while the intron is released still in the lariat formation [149].

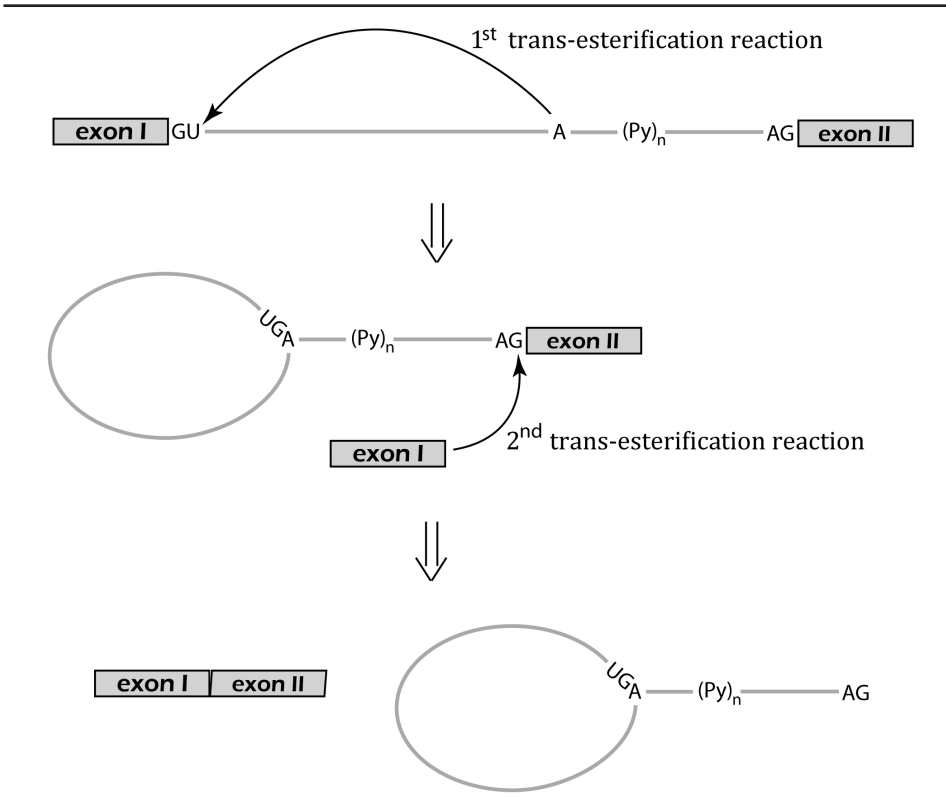


Figure 5 – The two transesterification reactions of splicing. The 1st reaction consists of a nucleophylic attack of the branch point to the 5' SS; in the 2nd reaction there is a nucleophylic attack of the 3' end of the 5' exon to the 3' SS, ligating the two exons and leaving the intron in a lariat conformation. Splicing factors are omitted for simplicity.

The assembly of the spliceosomal complex is regulated by several factors (table 2), some of them tissue-specific, which by promoting or inhibiting recruitment of spliceosomal components lead to skipping or inclusion of (alternative) exons. 158]. In contrast, polypyrimidine tract-binding protein (PTB or hnRNP I) inhibits splicing (promotes skipping) of many exons by outcompeting U2AF binding to the PPT in the intron [159], thus abrogating formation of the E complex. Combinations of activating and inhibitory effects will dictate whether a certain exon is skipped or included: for example, in the exon 5 of TNTT2 (troponin T type 2), PTB binds the PPT (inhibiting splicing), but ETR3 (which is tissue-specific) and CUGBP1 have been shown to displace PTB, thereby activating splicing [160]; in exon 2 of α-tropomyosin, 9G8 (SFRS7), hnRNP F and hnRNP H compete for binding the same element [161]. The concentration ratios of the regulating factors, which vary from tissue to tissue, are decisive as to whether a certain exon is skipped or included. Interestingly, some factors have been described which have both activating and inhibiting effects on splicing. Such is the case for NOVA1, NOVA2 (both are neuron-specific), FOX1, FOX2 (both are expressed specifically in muscle, heart and neurons), hnRNP L, hnRNP F and hnRNP H [162-169]. A closer inspection on individual AS events has shown that often the effect depends on the location of the factor's binding site relative to the regulated splice site. An illustrative example is the binding of hnRNP H to G-rich sequences (G runs) downstream of the 5' SS, which promotes formation of the A, B and C spliceosomal complexes, while binding to G runs in exons inhibits splicing [170]. Another example is the binding of NOVA1 and NOVA2 to both ISEs and ESSs, activating or inhibiting splicing, respectively [162].

Alternative Splicing and Disease

A survey from 1992 found that around 15% of point mutations which result in human genetic diseases affect classical splicing sites (i.e. 5' SS, branch site, PPT and 3' SS) [171]. Since this study did not include ESSs, ESEs, ISSs and ISEs, the proportion of point mutations interfering with splicing or AS is likely to be significantly higher. In addition, mutations involving splicing factors can also disrupt splicing or AS and cause or be associated with disease phenotypes. Examples of trans-acting splicing factors associated with disease are shown in table 3.

Mutations in *cis* can disrupt, enhance or create a (new) splice site, enhancer or silencer.

Serine/Arginine proteins (SR proteins)					
Name	**Domains**	**Binding motif**	**Name**	**Domains**	**Binding motif**
SRp20 (SFRS3)	RRM, RS	GCUCCUCUUC	**SRp54**	RRM, RS	Not defined
SC35 (SFRS2)	RRM, RS	UGCUGUU	**SRp46 (SFRS2B)**	RRM, RS	Not defined
ASF/SF2 (SFRS1)	RRM, RRMH, RS	RGAAGAAC	**RNPS1**	RRM, Ser-rich	Not defined
SRp40 (SFRS5)	RRM, RRMH, RS	AGGAGAAGGGA	**SRrp35**	RRM, RS	Not defined
SRp55 (SFRS6)	RRM, RRMH, RS	GGCAGCACCUG	**SRrp86 (SFRS12)**	RRM, RS	Not defined
SRp75 (SFRS4)	RRM, RRMH, RS	GAAGGA	**TRA2α**	RRM and two Arg-rich	GAAARGARR
9G8 (SFRS7)	RRM, zinc finger, RS	$(GAC)_n$	**TRA2β**	RRM and two RS	$(GAA)_n$
SRp30c (SFRS9)	RRM, RRMH, RS	CUGGAUU	**RBM5**	RRM, RS	Not defined
SRp38 (FUSIP1)	RRM, RS	AAAGACAAA	**CAPER (RBM39)**	RRM, RS	Not defined
Heterogenous ribonucleoproteins (hnRNPs)					
Name	**Domains**	**Binding motif**	**Name**	**Domains**	**Binding motif**
hnRNP A1	RRM, RGG, G	UAGGGA/U	**hnRNP H**	RRM, RGG, GYR, GY	GGGA and G-rich
hnRNP A2	RRM, RGG, G	$(UUAGGG)_n$	**hnRNP I (PTB)**	RRM	UCUU and CUCUCU
hnRNP C1	RRM	U-rich	**hnRNP L**	RRM	C and A-rich
hnRNP F	RRM, RGG, GY	GGGA and G-rich	**hnRNP M**	RRM, GY	Not defined
hnRNP G	RRM, SRGY	CC(A/C) and AAGU	**hnRNP Q**	RMM, RGG	Not defined

Table 2 – Factors involved in regulation of (Alternative) Splicing. In general, SR proteins activate splicing (i.e. promote exon inclusion) while hnRNPs inhibit splicing (i.e. promote skipping).

Gain-of-splicing-function mutations, which result in the creation or enhancement of a splicing element, have been described, for example, in β-thalassemia and Spinal muscular atrophy (SMA) [172-175]. Loss-of-splicing-function mutations disrupt an existent splice

site, enhancer or silencer and have been described in disorders such as β-thalassemia, hemophilia B, familial isolated growth hormone deficiency (IGHD II) or dementia [146, 176-179]. Both these two types of mutations often result in a loss-of-function phenotype, since the disruption of (alternative) splicing frequently generates PTCs which render the transcripts targets of NMD [147].

Alternative Splicing defects have also been regarded as a cause or modulator of cancer phenotypes. Trancripts from the nuclear hormone coactivator AIB1 (Amplified in Breast Cancer) have been shown to lack exon 3 in 7 out of 8 breast cancer samples, when compared to normal breast tissue; this alteration in AS of AIB1 generates a shorter protein which is more active at promoting estrogen receptor-mediated transcription than the full-length protein [180]. Hypomethylation of chromatin observed in liver tumors has been associated with changes in AS of Dnmt3b methyltransferase, in which skipping of exon 21 generates a frame-shift and a truncated protein [181]. A recent study reported that c-myc up-regulates PTB, hnRNP A1 and hnRNP A2, which cause alterations in AS leading to a high ratio of PKM2/PKM1 isoforms; high levels of PKM2 favor aerobic glycolysis, a hallmark of tumor cells, rather than oxidative phosphorylation. The same study demonstrates that human gliomas over-express c-myc, PTB, hnRNP A1 and hnRNP A2 in a manner that correlates with PKM2 expression [182]. PTB is also over-expressed in malignant astrocytomas and causes alterations is AS of FGFR1. Binding of PTB in the polypyrimidine tract upstream of the α exon of FGFR1 outcompetes U2AF, resulting in the exclusion of the exon from the mature transcript [183]. Similarly, FGFR2 is alternatively spliced, with at least 2 isoforms having different affinities to fibroblast growth factors. Changes in the FGFR2 isoform ratio have been correlated with prostate cancer progression and are associated with lung, skin and bone defects [184-186]. Many other examples of tumors where there is a disruption of AS have been described (reviewed in [148]); this fact has prompted the development of drugs or antisense oligonucleotides which, by correcting AS defects, delay tumor progression or alleviate some of its symptoms (reviewed in [147, 187]).

Another clinical field described to be modulated by AS is pharmacogenomics (reviewed in [188]). Several drug metabolizing genes have been described to be alternatively spliced; different patterns in AS result in different individual responsiveness or toxicity tolerance to certain prescribed drugs. Interestingly, many alterations in AS are caused by SNPs (single nucleotide polymorphisms); SNPs are geographically distributed in irregular patterns, which could contribute to the fact that certain populations metabolize certain drugs (or nutrients) more efficiently than others.

Splicing factor	Disease association	Comments	Ref.
CUGBP1	Myotonic Dystrophy	CUG-binding proteins may be affected in myotonic dystrophy.	[189]
CUGBP2 (ETR3)	Myotonic Dystrophy	CUG-binding proteins may be affected in myotonic dystrophy.	[190]
FUSIP1 (NSSR)	Leukemias and Sarcomas	FUSIP1 interacts with the C-terminal region of TLS and this interaction and perhaps the function of FUSIP1 is disrupted by the TLS-ERG fusion protein found in some leukemias and sarcomas.	[191]
FUS (TLS)	Liposarcomas, Acute Myeloid Leukemia (AML)	The FUS gene is translocated	[192, 193]
HMG-I (HMGA1a)	Alzheimer Disease (AD)	Overexpression of HMG-I was found to cause aberrant splicing of presenilin-2 transcripts, which is a feature of sporadic AD. Brains from AD patients show significant increases in HMG-I levels	[194]
MBNL1	Myotonic Dystrophy	Disease is likely caused by sequestration of MBNL1 to expanded repeat mutations in the transcripts	[195, 196]
NOVA1	Paraneoplastic syndrome	Autoantibodies to NOVA1 seen in patients with paraneoplastic syndrome	[197, 198]
PRPF3, PRPF8, PRPF31	Retinitis Pigmentosa (RP)	Mutations in PRPF3, PRPF8 or PRPF31 cause RP in some families	[199-201]
RBMY	Azospermia	Deletions of the RBMY coding gene(s) have been associated with azospermia	[202, 203]
HCC1	Hepatocellular carcinoma	Autoantibodies to HCC1 seen in patients, but pathophysiological consequences are not evident	[204]
SFPQ (PSF)	Papillary renal cell carcinoma	SFPQ was fused to the TFE3 gene product as a result of a translocation in papillary renal cell carcinomas	[205]
SMN1, SMN2	Spinal muscular atrophy (SMA)	Mutations in both SMN1 or SMN2 can cause SMA	[206, 207]
TP73L	Hay-Wells syndrome	Mutations in the alpha isoform of TP73L protein are associated with altered FGFR2 splicing and developmental abnormalities in Hay-Wells syndrome	[208]

Table 3 – Splicing trans-acting factors associated with human disease. Adapted from [147].

References

1. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.
2. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.
3. Steitz, T.A. and P.B. Moore, *RNA, the first macromolecular catalyst: the ribosome is a ribozyme.* Trends Biochem Sci, 2003. **28**(8): p. 411-8.
4. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function.* Cell, 2004. **116**(2): p. 281-97.
5. Lagos-Quintana, M., et al., *Identification of novel genes coding for small expressed RNAs.* Science, 2001. **294**(5543): p. 853-8.
6. Nguyen, V.T., et al., *7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes.* Nature, 2001. **414**(6861): p. 322-5.
7. Yang, Z., et al., *The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription.* Nature, 2001. **414**(6861): p. 317-22.
8. Guang, S., et al., *Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription.* Nature.
9. Tarn, W.Y. and J.A. Steitz, *Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns.* Science, 1996. **273**(5283): p. 1824-32.
10. Maniatis T, R.R., *The role of small nuclear ribonucleoproteins particles in pre-mRNA splicing.* Nature (Lond.), 1982. **325**: p. 673-678.
11. Vankan, P., C. McGuigan, and I.W. Mattaj, *Roles of U4 and U6 snRNAs in the assembly of splicing complexes.* EMBO J, 1992. **11**(1): p. 335-43.
12. Barabino, S.M., et al., *Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly.* Cell, 1990. **63**(2): p. 293-302.
13. Brown, C.J., et al., *The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus.* Cell, 1992. **71**(3): p. 527-42.
14. Brockdorff, N., et al., *The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus.* Cell, 1992. **71**(3): p. 515-26.
15. Sleutels, F., R. Zwart, and D.P. Barlow, *The non-coding Air RNA is required for silencing autosomal imprinted genes.* Nature, 2002. **415**(6873): p. 810-3.
16. Inacio, A., et al., *Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay.* J Biol Chem, 2004. **279**(31): p. 32170-80.
17. Silva, A.L., et al., *Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay.* RNA, 2008. **14**(3): p. 563-76.
18. Buhler, M., et al., *EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length.* Nat Struct Mol Biol, 2006. **13**(5): p. 462-4.
19. Eberle, A.B., et al., *Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region.* PLoS Biol, 2008. **6**(4): p. e92.

20. Mendell, J.T., et al., *Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise.* Nat Genet, 2004. **36**(10): p. 1073-8.
21. Singh, G., I. Rebbapragada, and J. Lykke-Andersen, *A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay.* PLoS Biol, 2008. **6**(4): p. e111.
22. Medghalchi, S.M., et al., *Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability.* Hum Mol Genet, 2001. **10**(2): p. 99-105.
23. Losson, R. and F. Lacroute, *Interference of nonsense mutations with eukaryotic messenger RNA stability.* Proc Natl Acad Sci U S A, 1979. **76**(10): p. 5134-7.
24. Maquat, L.E., et al., *Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia.* Cell, 1981. **27**(3 Pt 2): p. 543-53.
25. Kinniburgh, A.J., et al., *mRNA-deficient beta o-thalassemia results from a single nucleotide deletion.* Nucleic Acids Res, 1982. **10**(18): p. 5421-7.
26. Holbrook, J.A., et al., *Nonsense-mediated decay approaches the clinic.* Nat Genet, 2004. **36**(8): p. 801-8.
27. Culbertson, M.R., K.M. Underbrink, and G.R. Fink, *Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors.* Genetics, 1980. **95**(4): p. 833-53.
28. Leeds, P., et al., *The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon.* Genes Dev, 1991. **5**(12A): p. 2303-14.
29. Leeds, P., et al., *Gene products that promote mRNA turnover in Saccharomyces cerevisiae.* Mol Cell Biol, 1992. **12**(5): p. 2165-77.
30. Le Hir, H., et al., *The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay.* EMBO J, 2001. **20**(17): p. 4987-97.
31. Shibuya, T., et al., *eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay.* Nat Struct Mol Biol, 2004. **11**(4): p. 346-51.
32. Lykke-Andersen, J., M.D. Shu, and J.A. Steitz, *Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon.* Cell, 2000. **103**(7): p. 1121-31.
33. Gehring, N.H., et al., *Y14 and hUpf3b form an NMD-activating complex.* Mol Cell, 2003. **11**(4): p. 939-49.
34. Gehring, N.H., et al., *Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements.* Mol Cell, 2005. **20**(1): p. 65-75.
35. Behm-Ansmant, I., et al., *mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons.* FEBS Lett, 2007. **581**(15): p. 2845-53.
36. Isken, O. and L.E. Maquat, *The multiple lives of NMD factors: balancing roles in*

gene and genome regulation. Nat Rev Genet, 2008.
37. He, F., et al., *Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast.* Mol Cell, 2003. **12**(6): p. 1439-52.
38. Rehwinkel, J., et al., *Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets.* RNA, 2005. **11**(10): p. 1530-44.
39. Rehwinkel, J., J. Raes, and E. Izaurralde, *Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors.* Trends Biochem Sci, 2006. **31**(11): p. 639-46.
40. Wittmann, J., E.M. Hol, and H.M. Jack, *hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay.* Mol Cell Biol, 2006. **26**(4): p. 1272-87.
41. Green, R.E., et al., *Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes.* Bioinformatics, 2003. **19 Suppl 1**: p. i118-21.
42. Lewis, B.P., R.E. Green, and S.E. Brenner, *Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans.* Proc Natl Acad Sci U S A, 2003. **100**(1): p. 189-92.
43. Chiu, S.Y., et al., *The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex.* Genes Dev, 2004. **18**(7): p. 745-54.
44. Imataka, H., A. Gradi, and N. Sonenberg, *A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation.* EMBO J, 1998. **17**(24): p. 7480-9.
45. Kahvejian, A., et al., *Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms.* Genes Dev, 2005. **19**(1): p. 104-13.
46. Lykke-Andersen, J., M.D. Shu, and J.A. Steitz, *Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1.* Science, 2001. **293**(5536): p. 1836-9.
47. Kataoka, N., et al., *Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm.* Mol Cell, 2000. **6**(3): p. 673-82.
48. Kim, V.N., et al., *The Y14 protein communicates to the cytoplasm the position of exon-exon junctions.* EMBO J, 2001. **20**(8): p. 2062-8.
49. Nott, A., H. Le Hir, and M.J. Moore, *Splicing enhances translation in mammalian cells: an additional function of the exon junction complex.* Genes Dev, 2004. **18**(2): p. 210-22.
50. Ishigaki, Y., et al., *Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20.* Cell, 2001. **106**(5): p. 607-17.
51. Kashima, I., et al., *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, 2006. **20**(3): p. 355-67.

52. Hosoda, N., et al., *CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells.* Nat Struct Mol Biol, 2005. **12**(10): p. 893-901.
53. Kuzmiak, H.A. and L.E. Maquat, *Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges.* Trends Mol Med, 2006. **12**(7): p. 306-16.
54. Yamashita, A., et al., *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, 2001. **15**(17): p. 2215-28.
55. Yamashita, A., et al., *SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay.* Genes Dev, 2009. **23**(9): p. 1091-105.
56. Unterholzner, L. and E. Izaurralde, *SMG7 acts as a molecular link between mRNA surveillance and mRNA decay.* Mol Cell, 2004. **16**(4): p. 587-96.
57. Chen, C.Y. and A.B. Shyu, *Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway.* Mol Cell Biol, 2003. **23**(14): p. 4805-13.
58. Isken, O., et al., *Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay.* Cell, 2008. **133**(2): p. 314-27.
59. Ivanov, P.V., et al., *Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways.* EMBO J, 2008. **27**(5): p. 736-47.
60. Ohnishi, T., et al., *Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7.* Mol Cell, 2003. **12**(5): p. 1187-200.
61. Lejeune, F., X. Li, and L.E. Maquat, *Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities.* Mol Cell, 2003. **12**(3): p. 675-87.
62. Stalder, L. and O. Muhlemann, *Processing bodies are not required for mammalian nonsense-mediated mRNA decay.* RNA, 2009. **15**(7): p. 1265-73.
63. Huntzinger, E., et al., *SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan.* RNA, 2008. **14**(12): p. 2609-17.
64. Eberle, A.B., et al., *SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells.* Nat Struct Mol Biol, 2009. **16**(1): p. 49-55.
65. Le Hir, H., et al., *The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions.* EMBO J, 2000. **19**(24): p. 6860-9.
66. Custodio, N., et al., *In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei.* RNA, 2004. **10**(4): p. 622-33.
67. Ballut, L., et al., *The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity.* Nat Struct Mol Biol, 2005. **12**(10): p. 861-9.
68. Gehring, N.H., et al., *Disassembly of exon junction complexes by PYM.* Cell, 2009. **137**(3): p. 536-48.
69. Boswell, R.E., M.E. Prout, and J.C. Steichen, *Mutations in a newly identified*

Drosophila melanogaster gene, mago nashi, disrupt germ cell formation and result in the formation of mirror-image symmetrical double abdomen embryos. Development, 1991. **113**(1): p. 373-84.
70. Masuda, S., et al., *Recruitment of the human TREX complex to mRNA during splicing.* Genes Dev, 2005. **19**(13): p. 1512-7.
71. Strasser, K., et al., *TREX is a conserved complex coupling transcription with messenger RNA export.* Nature, 2002. **417**(6886): p. 304-8.
72. Stutz, F., et al., *REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export.* RNA, 2000. **6**(4): p. 638-50.
73. Zhang, Z. and A.R. Krainer, *Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11574-9.
74. Chan, W.K., et al., *A UPF3-mediated regulatory switch that maintains RNA surveillance.* Nat Struct Mol Biol, 2009. **16**(7): p. 747-53.
75. Chamieh, H., et al., *NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity.* Nat Struct Mol Biol, 2008. **15**(1): p. 85-93.
76. Trembley, J.H., et al., *Activation of pre-mRNA splicing by human RNPS1 is regulated by CK2 phosphorylation.* Mol Cell Biol, 2005. **25**(4): p. 1446-57.
77. Blencowe, B.J., et al., *A coactivator of pre-mRNA splicing.* Genes Dev, 1998. **12**(7): p. 996-1009.
78. Cheng, C. and P.A. Sharp, *Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion.* Mol Cell Biol, 2006. **26**(1): p. 362-70.
79. McCracken, S., M. Lambermon, and B.J. Blencowe, *SRm160 splicing coactivator promotes transcript 3'-end cleavage.* Mol Cell Biol, 2002. **22**(1): p. 148-60.
80. Luo, M.L., et al., *Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly.* Nature, 2001. **413**(6856): p. 644-7.
81. Degot, S., et al., *Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module.* J Biol Chem, 2004. **279**(32): p. 33702-15.
82. Tange, T.O., et al., *Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core.* RNA, 2005. **11**(12): p. 1869-83.
83. Sakashita, E., et al., *Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo.* Mol Cell Biol, 2004. **24**(3): p. 1174-87.
84. Wang, P., et al., *Modulation of alternative pre-mRNA splicing in vivo by pinin.* Biochem Biophys Res Commun, 2002. **294**(2): p. 448-55.
85. Diem, M.D., et al., *PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs.* Nat Struct Mol Biol, 2007. **14**(12): p. 1173-9.
86. Fukuhara, N., et al., *SMG7 is a 14-3-3-like adaptor in the nonsense-mediated*

mRNA decay pathway. Mol Cell, 2005. **17**(4): p. 537-47.
87. Jin, H., et al., *Human UPF1 participates in small RNA-induced mRNA downregulation.* Mol Cell Biol, 2009. **29**(21): p. 5789-99.
88. Kaygun, H. and W.F. Marzluff, *Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1.* Nat Struct Mol Biol, 2005. **12**(9): p. 794-800.
89. Kim, Y.K., et al., *Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay.* Cell, 2005. **120**(2): p. 195-208.
90. Azzalin, C.M., et al., *Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends.* Science, 2007. **318**(5851): p. 798-801.
91. Viegas, M.H., et al., *The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway.* Nucleic Acids Res, 2007. **35**(13): p. 4542-51.
92. Linde, L., et al., *The efficiency of nonsense-mediated mRNA decay is an inherent character and varies among different cells.* Eur J Hum Genet, 2007. **15**(11): p. 1156-62.
93. Page, M.F., et al., *SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast.* Mol Cell Biol, 1999. **19**(9): p. 5943-51.
94. Culbertson, M.R. and P.F. Leeds, *Looking at mRNA decay pathways through the window of molecular evolution.* Curr Opin Genet Dev, 2003. **13**(2): p. 207-14.
95. Cheng, Z., et al., *Structural and functional insights into the human Upf1 helicase core.* EMBO J, 2007. **26**(1): p. 253-64.
96. Chiu, S.Y., et al., *Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1.* RNA, 2003. **9**(1): p. 77-87.
97. Mendell, J.T., et al., *Novel Upf2p orthologues suggest a functional link between translation initiation and nonsense surveillance complexes.* Mol Cell Biol, 2000. **20**(23): p. 8944-57.
98. Serin, G., et al., *Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4).* Mol Cell Biol, 2001. **21**(1): p. 209-23.
99. Kadlec, J., E. Izaurralde, and S. Cusack, *The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3.* Nat Struct Mol Biol, 2004. **11**(4): p. 330-7.
100. Kunz, J.B., et al., *Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation.* RNA, 2006. **12**(6): p. 1015-22.
101. Bono, F., et al., *The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA.* Cell, 2006. **126**(4): p. 713-25.
102. Andersen, C.B., et al., *Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA.* Science, 2006. **313**(5795): p. 1968-72.
103. Stalder, L. and O. Muhlemann, *The meaning of nonsense.* Trends Cell Biol, 2008. **18**(7): p. 315-21.
104. Frischmeyer, P.A. and H.C. Dietz, *Nonsense-mediated mRNA decay in health and*

disease. Hum Mol Genet, 1999. **8**(10): p. 1893-900.
105. Pastor, F., et al., *Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay.* Nature. **465**(7295): p. 227-30.
106. Waggoner, S.A. and S.A. Liebhaber, *Regulation of alpha-globin mRNA stability.* Exp Biol Med (Maywood), 2003. **228**(4): p. 387-95.
107. Thein, S.L., et al., *Molecular basis for dominantly inherited inclusion body beta-thalassemia.* Proc Natl Acad Sci U S A, 1990. **87**(10): p. 3924-8.
108. Rosenfeld, P.J., et al., *A null mutation in the rhodopsin gene causes rod photoreceptor dysfunction and autosomal recessive retinitis pigmentosa.* Nat Genet, 1992. **1**(3): p. 209-13.
109. Sung, C.H., et al., *Rhodopsin mutations in autosomal dominant retinitis pigmentosa.* Proc Natl Acad Sci U S A, 1991. **88**(15): p. 6481-5.
110. Millar, D.S., et al., *Molecular analysis of the genotype-phenotype relationship in factor X deficiency.* Hum Genet, 2000. **106**(2): p. 249-57.
111. Jouanguy, E., et al., *Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guerin infection.* N Engl J Med, 1996. **335**(26): p. 1956-61.
112. Jouanguy, E., et al., *A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection.* Nat Genet, 1999. **21**(4): p. 370-8.
113. Schneppenheim, R., et al., *Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease.* Blood, 2001. **97**(7): p. 2059-66.
114. Hodgkin, J., et al., *A new kind of informational suppression in the nematode Caenorhabditis elegans.* Genetics, 1989. **123**(2): p. 301-13.
115. Wittkopp, N., et al., *Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival.* Mol Cell Biol, 2009. **29**(13): p. 3517-28.
116. Weischenfeldt, J., et al., *NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements.* Genes Dev, 2008. **22**(10): p. 1381-96.
117. Farlow, A., et al., *Nonsense-mediated decay enables intron gain in Drosophila.* PLoS Genet. **6**(1): p. e1000819.
118. Berget, S.M., C. Moore, and P.A. Sharp, *Spliced segments at the 5' terminus of adenovirus 2 late mRNA.* Proc Natl Acad Sci U S A, 1977. **74**(8): p. 3171-5.
119. Chow, L.T., et al., *An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA.* Cell, 1977. **12**(1): p. 1-8.
120. Alt, F.W., et al., *Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends.* Cell, 1980. **20**(2): p. 293-301.
121. Early, P., et al., *Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways.* Cell, 1980. **20**(2): p. 313-9.
122. Okazaki, Y., et al., *Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.* Nature, 2002. **420**(6915): p. 563-73.
123. Zavolan, M., et al., *Impact of alternative initiation, splicing, and termination*

on the diversity of the mRNA transcripts encoded by the mouse transcriptome. Genome Res, 2003. **13**(6B): p. 1290-300.
124. Stamm, S., et al., *Function of alternative splicing.* Gene, 2005. **344**: p. 1-20.
125. Boutz, P.L., et al., *A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons.* Genes Dev, 2007. **21**(13): p. 1636-52.
126. Makeyev, E.V., et al., *The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing.* Mol Cell, 2007. **27**(3): p. 435-48.
127. Sanchez, L., *Sex-determining mechanisms in insects.* Int J Dev Biol, 2008. **52**(7): p. 837-56.
128. Black, D.L., *Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology.* Cell, 2000. **103**(3): p. 367-70.
129. Graveley, B.R., *Alternative splicing: increasing diversity in the proteomic world.* Trends Genet, 2001. **17**(2): p. 100-7.
130. Fong, N., et al., *A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing.* EMBO J, 2003. **22**(16): p. 4274-82.
131. McCracken, S., et al., *The C-terminal domain of RNA polymerase II couples mRNA processing to transcription.* Nature, 1997. **385**(6614): p. 357-61.
132. Goldstrohm, A.C., A.L. Greenleaf, and M.A. Garcia-Blanco, *Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing.* Gene, 2001. **277**(1-2): p. 31-47.
133. de Almeida, S.F. and M. Carmo-Fonseca, *The CTD role in cotranscriptional RNA processing and surveillance.* FEBS Lett, 2008. **582**(14): p. 1971-6.
134. Kornblihtt, A.R., et al., *Multiple links between transcription and splicing.* RNA, 2004. **10**(10): p. 1489-98.
135. Nogues, G., et al., *Transcriptional activators differ in their abilities to control alternative splicing.* J Biol Chem, 2002. **277**(45): p. 43110-4.
136. Munoz, M.J., et al., *DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation.* Cell, 2009. **137**(4): p. 708-20.
137. Tilgner, H., et al., *Nucleosome positioning as a determinant of exon recognition.* Nat Struct Mol Biol, 2009. **16**(9): p. 996-1001.
138. Tillo, D. and T.R. Hughes, *G+C content dominates intrinsic nucleosome occupancy.* BMC Bioinformatics, 2009. **10**: p. 442.
139. Schwartz, S., E. Meshorer, and G. Ast, *Chromatin organization marks exon-intron structure.* Nat Struct Mol Biol, 2009. **16**(9): p. 990-5.
140. Keren, H., G. Lev-Maor, and G. Ast, *Alternative splicing and evolution: diversification, exon definition and function.* Nat Rev Genet. **11**(5): p. 345-55.
141. Wollerton, M.C., et al., *Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay.* Mol Cell, 2004. **13**(1): p. 91-100.
142. Jumaa, H. and P.J. Nielsen, *The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation.* EMBO J, 1997. **16**(16): p. 5077-

85.
143. Lejeune, F., Y. Cavaloc, and J. Stevenin, *Alternative splicing of intron 3 of the serine/arginine-rich protein 9G8 gene. Identification of flanking exonic splicing enhancers and involvement of 9G8 as a trans-acting factor.* J Biol Chem, 2001. **276**(11): p. 7850-8.
144. Sureau, A., et al., *SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs.* EMBO J, 2001. **20**(7): p. 1785-96.
145. Spellman, R., M. Llorian, and C.W. Smith, *Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1.* Mol Cell, 2007. **27**(3): p. 420-34.
146. Faustino, N.A. and T.A. Cooper, *Pre-mRNA splicing and human disease.* Genes Dev, 2003. **17**(4): p. 419-37.
147. Garcia-Blanco, M.A., A.P. Baraniak, and E.L. Lasda, *Alternative splicing in disease and therapy.* Nat Biotechnol, 2004. **22**(5): p. 535-46.
148. Venables, J.P., *Aberrant and alternative splicing in cancer.* Cancer Res, 2004. **64**(21): p. 7647-54.
149. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing.* Annu Rev Biochem, 2003. **72**: p. 291-336.
150. Berglund, J.A., et al., *The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC.* Cell, 1997. **89**(5): p. 781-7.
151. Zamore, P.D. and M.R. Green, *Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor.* Proc Natl Acad Sci U S A, 1989. **86**(23): p. 9243-7.
152. Nelson, K.K. and M.R. Green, *Mammalian U2 snRNP has a sequence-specific RNA-binding activity.* Genes Dev, 1989. **3**(10): p. 1562-71.
153. Kohtz, J.D., et al., *Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors.* Nature, 1994. **368**(6467): p. 119-24.
154. Wu, J.Y. and T. Maniatis, *Specific interactions between proteins implicated in splice site selection and regulated alternative splicing.* Cell, 1993. **75**(6): p. 1061-70.
155. Xiao, S.H. and J.L. Manley, *Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing.* Genes Dev, 1997. **11**(3): p. 334-44.
156. Chen, M. and J.L. Manley, *Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches.* Nat Rev Mol Cell Biol, 2009. **10**(11): p. 741-54.
157. Forch, P., et al., *The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites.* EMBO J, 2002. **21**(24): p. 6882-92.
158. Izquierdo, J.M., et al., *Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition.* Mol Cell, 2005. **19**(4): p. 475-84.
159. Valcarcel, J. and F. Gebauer, *Post-transcriptional regulation: the dawn of PTB.* Curr Biol, 1997. **7**(11): p. R705-8.
160. Charlet, B.N., et al., *Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing.* Mol Cell, 2002. **9**(3): p. 649-58.

161. Crawford, J.B. and J.G. Patton, *Activation of alpha-tropomyosin exon 2 is regulated by the SR protein 9G8 and heterogeneous nuclear ribonucleoproteins H and F.* Mol Cell Biol, 2006. **26**(23): p. 8791-802.
162. Ule, J., et al., *An RNA map predicting Nova-dependent splicing regulation.* Nature, 2006. **444**(7119): p. 580-6.
163. Yeo, G.W., et al., *An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells.* Nat Struct Mol Biol, 2009. **16**(2): p. 130-7.
164. Mauger, D.M., C. Lin, and M.A. Garcia-Blanco, *hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc.* Mol Cell Biol, 2008. **28**(17): p. 5403-19.
165. Hung, L.H., et al., *Diverse roles of hnRNP L in mammalian mRNA processing: a combined microarray and RNAi analysis.* RNA, 2008. **14**(2): p. 284-96.
166. Licatalosi, D.D., et al., *HITS-CLIP yields genome-wide insights into brain alternative RNA processing.* Nature, 2008. **456**(7221): p. 464-9.
167. Dredge, B.K., et al., *Nova autoregulation reveals dual functions in neuronal splicing.* EMBO J, 2005. **24**(8): p. 1608-20.
168. Wang, Z. and C.B. Burge, *Splicing regulation: from a parts list of regulatory elements to an integrated splicing code.* RNA, 2008. **14**(5): p. 802-13.
169. Martinez-Contreras, R., et al., *Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing.* PLoS Biol, 2006. **4**(2): p. e21.
170. Caputi, M. and A.M. Zahler, *Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family.* J Biol Chem, 2001. **276**(47): p. 43850-9.
171. Krawczak, M., J. Reiss, and D.N. Cooper, *The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences.* Hum Genet, 1992. **90**(1-2): p. 41-54.
172. Busslinger, M., N. Moschonas, and R.A. Flavell, *Beta + thalassemia: aberrant splicing results from a single point mutation in an intron.* Cell, 1981. **27**(2 Pt 1): p. 289-98.
173. Spritz, R.A., et al., *Base substitution in an intervening sequence of a beta+-thalassemic human globin gene.* Proc Natl Acad Sci U S A, 1981. **78**(4): p. 2455-9.
174. Kashima, T. and J.L. Manley, *A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy.* Nat Genet, 2003. **34**(4): p. 460-3.
175. Singh, N.N., E.J. Androphy, and R.N. Singh, *An extended inhibitory context causes skipping of exon 7 of SMN2 in spinal muscular atrophy.* Biochem Biophys Res Commun, 2004. **315**(2): p. 381-8.
176. Maquat, L.E., et al., *Processing of human beta-globin mRNA precursor to mRNA is defective in three patients with beta+-thalassemia.* Proc Natl Acad Sci U S A, 1980. **77**(7): p. 4287-91.
177. Rees, D.J., C.R. Rizza, and G.G. Brownlee, *Haemophilia B caused by a point mutation in a donor splice junction of the human factor IX gene.* Nature, 1985. **316**(6029): p. 643-5.

178. Ryther, R.C., et al., *Disruption of exon definition produces a dominant-negative growth hormone isoform that causes somatotroph death and IGHD II.* Hum Genet, 2003. **113**(2): p. 140-8.
179. Millar, D.S., et al., *Novel mutations of the growth hormone 1 (GH1) gene disclosed by modulation of the clinical selection criteria for individuals with short stature.* Hum Mutat, 2003. **21**(4): p. 424-40.
180. Reiter, R., A. Wellstein, and A.T. Riegel, *An isoform of the coactivator AIB1 that increases hormone and growth factor sensitivity is overexpressed in breast cancer.* J Biol Chem, 2001. **276**(43): p. 39736-41.
181. Saito, Y., et al., *Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis.* Proc Natl Acad Sci U S A, 2002. **99**(15): p. 10060-5.
182. David, C.J., et al., *HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer.* Nature. **463**(7279): p. 364-8.
183. Jin, W., et al., *Polypyrimidine tract-binding protein down-regulates fibroblast growth factor receptor 1 alpha-exon inclusion.* Cancer Res, 2003. **63**(19): p. 6154-7.
184. Naimi, B., et al., *Down-regulation of (IIIb) and (IIIc) isoforms of fibroblast growth factor receptor 2 (FGFR2) is associated with malignant progression in human prostate.* Prostate, 2002. **52**(3): p. 245-52.
185. De Moerlooze, L., et al., *An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis.* Development, 2000. **127**(3): p. 483-92.
186. Eswarakumar, V.P., et al., *The IIIc alternative of Fgfr2 is a positive regulator of bone formation.* Development, 2002. **129**(16): p. 3783-93.
187. van Alphen, R.J., et al., *The spliceosome as target for anticancer treatment.* Br J Cancer, 2009. **100**(2): p. 228-32.
188. Bracco, L. and J. Kearsey, *The relevance of alternative RNA splicing to pharmacogenomics.* Trends Biotechnol, 2003. **21**(8): p. 346-53.
189. Roberts, R., et al., *Altered phosphorylation and intracellular distribution of a (CUG)n triplet repeat RNA-binding protein in patients with myotonic dystrophy and in myotonin protein kinase knockout mice.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13221-6.
190. Lu, X., N.A. Timchenko, and L.T. Timchenko, *Cardiac elav-type RNA-binding protein (ETR-3) binds to RNA CUG repeats expanded in myotonic dystrophy.* Hum Mol Genet, 1999. **8**(1): p. 53-60.
191. Yang, L., L.J. Embree, and D.D. Hickstein, *TLS-ERG leukemia fusion protein inhibits RNA splicing mediated by serine-arginine proteins.* Mol Cell Biol, 2000. **20**(10): p. 3345-54.
192. Crozat, A., et al., *Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.* Nature, 1993. **363**(6430): p. 640-4.
193. Ichikawa, H., et al., *An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation.* Cancer Res,

1994. **54**(11): p. 2865-8.
194. Manabe, T., et al., *Induced HMGA1a expression causes aberrant splicing of Presenilin-2 pre-mRNA in sporadic Alzheimer's disease.* Cell Death Differ, 2003. **10**(6): p. 698-708.
195. Kanadia, R.N., et al., *A muscleblind knockout model for myotonic dystrophy.* Science, 2003. **302**(5652): p. 1978-80.
196. Miller, J.W., et al., *Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy.* EMBO J, 2000. **19**(17): p. 4439-48.
197. Buckanovich, R.J., J.B. Posner, and R.B. Darnell, *Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system.* Neuron, 1993. **11**(4): p. 657-72.
198. Jensen, K.B., et al., *Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability.* Neuron, 2000. **25**(2): p. 359-71.
199. Chakarova, C.F., et al., *Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa.* Hum Mol Genet, 2002. **11**(1): p. 87-92.
200. Vithana, E.N., et al., *A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11).* Mol Cell, 2001. **8**(2): p. 375-81.
201. McKie, A.B., et al., *Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13).* Hum Mol Genet, 2001. **10**(15): p. 1555-62.
202. Ma, K., et al., *A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis.* Cell, 1993. **75**(7): p. 1287-95.
203. Venables, J.P., et al., *RBMY, a probable human spermatogenesis factor, and other hnRNP G proteins interact with Tra2beta and affect splicing.* Hum Mol Genet, 2000. **9**(5): p. 685-94.
204. Imai, H., et al., *Novel nuclear autoantigen with splicing factor motifs identified with antibody from hepatocellular carcinoma.* J Clin Invest, 1993. **92**(5): p. 2419-26.
205. Clark, J., et al., *Fusion of splicing factor genes PSF and NonO (p54nrb) to the TFE3 gene in papillary renal cell carcinoma.* Oncogene, 1997. **15**(18): p. 2233-9.
206. Lefebvre, S., et al., *Identification and characterization of a spinal muscular atrophy-determining gene.* Cell, 1995. **80**(1): p. 155-65.
207. Srivastava, S., et al., *SMN2-deletion in childhood-onset spinal muscular atrophy.* Am J Med Genet, 2001. **101**(3): p. 198-202.
208. Fomenkov, A., et al., *P63 alpha mutations lead to aberrant splicing of keratinocyte growth factor receptor in the Hay-Wells syndrome.* J Biol Chem, 2003. **278**(26): p. 23906-14.

Chapter 2

Rod1, Upf2 and Upf1 interact and regulate common Nonsense-Mediated mRNA Decay target genes

work in progress

Summary

Nonsense-Mediated mRNA Decay (NMD) is a mechanism of mRNA degradation in eukaryotes which eliminates transcripts containing premature termination codons (PTCs). NMD targets include not only mutated transcripts containing PTCs, but also many natural physiologic transcripts containing long 3'UTRs, upstream ORFs, introns in the 3'UTR or transcripts which acquire PTC's through nuclear events such as alternative spicing. NMD relies on a large protein complex which assembles during splicing approximately 24 nucleotides upstream of exon-exon borders, the exon-junction complex (EJC). The EJC functions as a molecular mark locating intron positions and plays an important role in discriminating normal termination codons from premature ones. Tethering experiments suggest that at least two different EJC subcomplexes might exist, one containing Upf2 and another lacking Upf2. Here, we report a new interacting partner of the EJC-binding factors Upf2 and Upf1 in HEK293 cells called Rod1. Rod1 also interacted with Aly/REF, an EJC component involved in mRNA export. We show that Rod1, like Upf2 and Upf1, is required to down-regulate a β-globin reporter transcript containing a NMD-sensitive nonsense mutation. Furthermore, we show through exon array data that Rod1, Upf2 and Upf1 regulate hundreds of common genes, several of them upregulated upon knock-down of any of those factors, and thus potential NMD targets. Intriguingly, the pattern of gene regulation by Rod1 was strikingly similar to that of Upf2, suggesting that these two factors participate in the same mechanistic pathway(s). Overall our results indicate a new role for Rod1 in NMD.

Introduction

The generation of truncated polypeptides potentially changes the functional properties of proteins, with possible deleterious consequences for the normal functioning of the cell. The mechanism of Nonsense-Mediated mRNA Decay (NMD) prevents this by eliminating transcripts containing premature termination codons (PTCs) (reviewed in [1-3]). In addition to PTC-containing mRNAs, NMD targets natural physiological transcripts with long 3'UTRs, transcripts with introns in the 3'UTRs, transcripts with upstream ORFs and selenoprotein mRNAs [4-10]. PTC-containing transcripts, arising from nonsense or frameshift mutations, errors in transcription or splicing, DNA recombination or alternative splicing events, comprise a significant fraction of the transcriptome. Accordingly, inhibition of NMD has been reported to change expression of many hundreds of genes [9, 11, 12]. Mechanistically, NMD relies largely on a large molecular weight complex associated

with mature transcripts, the Exon-Junction Complex (EJC). The EJC is assembled during splicing around 24 nt upstream of exon-exon junctions and is composed of several factors which mediate RNA splicing (eg. Pinin, SRm160, RNPS1), export (eg. UAP56, Aly/REF, TAP), translation (eg. PYM, RNPS1, MLN51) and surveillance activities [13-20]. Upf3 (Upf3a or Upf3b), Upf2 and Upf1 constitute a core protein complex which associates with the EJC and is necessary for later steps in NMD. These factors are conserved from yeast to man [21]. EJCs are disassembled from the mRNA during the first round of translation, but if the ribosome encounters a PTC before displacing the last RNA-bound EJC, the transcript is degraded by NMD (there are some exceptions to this rule, namely when PTCs are located near the 5' end of the mRNA) [22, 23]. Pausing of the ribosome at a PTC triggers the formation of the SURF complex (SMG1, Upf1 and the eukaryotic release factors eRF1 and eRF3), which binds both the translation machinery and the downstream EJC [24]; Upf1 binds EJC-bound Upf2 which, together with Upf3b, promote phosphorylation and the RNA helicase activity of Upf1 [11, 25, 26]. Phosphorylation of Upf1 by SMG1 is thought to mediate recruitment of SMG7, which in turn brings in factors directly involved in degradation of the mRNA [27, 28]. mRNA decay is thought to involve three possible pathways: endonucleolytic RNA cleavage near the PTC triggered by SMG6, deadenylation by deadenylating enzymes (PARN, CCR4-NOT or PAN2-PAN3) and/or decapping by Dcp2 and Dcp1 [29-31]; these routes to mRNA degradation are not mutually exclusive. Any of these pathways exposes the mRNA to the cytoplasmic exosome (which degrades RNA in the 3' to 5' direction) or Xrn1 (a 5' to 3' exonuclease).

Tethering experiments have demonstrated that activation of NMD can be achieved by at least two different routes, one Upf2-dependent and one Upf2-independent. RNPS1 was shown to trigger NMD in an Upf2-dependent manner, while eIF4AIII, magoh, Upf3b or Y14 triggered NMD in an Upf2-independent manner [32]. This suggests there could be different EJC complexes functioning *in vivo*, the composition of which would depend on the tissue, transcript or even exon-exon junction.

In the present study, we use a proteomic approach to find Upf2-interacting factors in human embryonic kidney (HEK293) cells. We show that Rod1 interacts with Upf2, Upf1 and Aly/REF, suggesting that it might associate with the EJC. Using a PTC-containing β-globin reporter transcript, we demonstrate that Rod1, like Upf2 and Upf1, is required for its degradation. In addition, we performed an analysis of target genes regulated by Rod1, Upf2 and Upf1 by exon arrays. The results show that these three factors regulate hundreds of common genes, several of them potential NMD targets, since they are upregulated upon RNAi. Interestingly, genes downregulated/upregulated upon knock-down of Rod1 follow the same pattern upon knock-down of Upf2, suggesting there is a close functional relation

between these two factors. Overall, our results give evidence for a new Upf1 and Upf2-interacting partner with a role in NMD in HEK293 cells.

Results

Upf2 and Upf1 interact with Rod1: In order to find new Upf2-interacting factors in HEK293 cells, we immunoprecipitated Upf2-containing protein complexes from nuclear extracts with an anti-Upf2 antibody crosslinked to protein-G magnetic beads. Eluted complexes were separated by SDS-PAGE and analysed by mass spectrometry (table 1). As expected, Upf2 itself, Upf1 and Upf3b were present in purified complexes with high mascot scores (1214, 1079 and 680, respectively). Despite the low mascot score (67), we confirmed an interaction of Upf2 with Rod1 (fig. 1A). Rod1 is a Ptbp1 (hnRNP I) paralog expressed in several embryonic tissues and in adult haematopoietic cells [33]. Mass spectrometry data from Ptbp1-interacting proteins has indicated that Ptbp1 (PTB in humans) also interacts with Upf2 in murine erythroleukemia (MEL) cells (chapter 3). We further confirmed the Upf2-Rod1 interaction by analyzing Rod1-containing protein complexes in HEK293 nuclear extracts by mass spectrometry (table 2). Strikingly, both Upf2 and Upf1 were present with very high mascot scores (1203 and 654, respectively), in addition to Rod1 itself (663). Another EJC-associated factor, Aly/REF, was present with a mascot score of 110. These interactions were verified by immunoprecipitations (IPs) followed by western blotting (fig. 1A). To better characterize the interactions between Rod1 and both Upf1 and Upf2, we performed immunocytochemistry experiments in HEK293 cells (fig. 1B). Both Upf1 and Upf2 co-localized with Rod1 in nuclear speckles, where a significant fraction of Rod1 was localized. In contrast, both Upf2 and Upf1 show a wider cellular distribution; as opposed to what has been previously described for HeLa cells [34-36], both Upf2 and Upf1 are present in the nuclei of HEK293 cells in significant amounts. Rod1 is also present in the cytoplasm, although in lower amounts than in the nucleus, and in a fraction of cells it co-localizes with Upf1 in discrete cytoplasmic speckles (fig. 1B). This cytoplasmic co-localization is further supported by cytoplasmic mass spectrometry data showing that Upf1, but not Upf2, interacts with Rod1 (supplementary table 1). The cellular localization patterns of Rod1, Upf2 and Upf1, together with mass spectrometry data and IPs, indicate that a significant fraction of Rod1 is associated with Upf2, Upf1 or both, while the reverse is not valid: only a relatively small fraction of Upf2 or Upf1 appears to associate with Rod1. Altogether, our data provides clear evidence for as of yet unidentified nuclear interactions between Rod1, Upf2 and Upf1 in HEK293 cells. In addition, the cytoplasmic interaction of Rod1 with Upf1 is also demonstrated.

Protein	score Control	score Upf2	Protein	score Control	score Upf2
RRBP1	-	3510	**MAGED2**	-	123
CPSF1	457	1882	**TADA1**	-	121
RFX5	-	1679	**SMARCAD1**	29	121
SYMPK	111	1333	**LCLAT1**	-	118
Upf2 *	-	1214	**SACM1L**	-	118
TUBA1C	-	1191	**CALM1**	-	118
TUBB2A	-	1180	**ISY1**	-	116
SRP68	212	1098	**TADA3**	-	115
UPF1 *	-	1079	**CCAR1**	-	115
WDR33	-	972	**MDH2**	-	115
RUVBL2	-	968	**UCHL5**	-	113
FNBP1L	-	940	**BET1 ***	-	112
DHX16	-	933	**SEC11A**	-	111
DDX3Y	-	858	**CCDC75**	-	110
MYH11	-	852	**RBBP6**	-	110
SRP72	62	809	**TMEM205**	-	109
GPKOW	-	789	**HSDL2**	-	108
eEF1D	110	779	**snRPb2**	-	106
RUVBL1	-	731	**PRR5**	-	106
FIP1L1	63	726	**UBL4A**	-	105
CPSF4	148	720	**ZC3HAV1L**	-	100
UPF3B	36	680	**HLA B**	-	100
VDAC1	-	658	**FLOT1**	-	99
SRP54	-	654	**HEATR1**	-	97
hnRNP H1	128	653	**PEX16**	-	97
eEF1G	-	648	**GRB10**	-	95
DDX39	-	643	**GNPAT**	-	94
RFXAP	-	635	**MED6**	-	94
POLDIP3	-	618	**CLASP2**	-	91
CPSF3	-	613	**SMC1A**	-	90
CSTF2	108	584	**KPNA1**	-	90
CSTF1	-	561	**TOR1A**	-	90
SR140	105	530	**SPEN**	-	89
CTTN	-	496	**SYNJ2BP**	-	89
LUC7L	53	480	**RAB3L**	-	88
eEF1B2	99	464	**DNAH17**	-	86
CAT	-	456	**LDLRAP1**	-	85
PDIA6	73	408	**DHRS7B**	-	85
SEPT11	-	406	**DYNLT1**	-	84
MTA2	101	405	**eIF3M**	-	84
AIP	50	401	**PECR**	-	84
eIF3K	86	392	**ILK**	-	84
hnRNP K	-	390	**FAM98A**	-	83
SEPT7	-	388	**SMC4**	-	82
SERBP1	-	386	**PPIL3**	-	82
YKT6 *	-	369	**DTWD2**	-	81
FAR1	-	364	**CEP110**	-	81
CALR	65	356	**MOSC1**	-	81
RPL4	-	353	**ORC2L**	-	81
ACAA1	-	343	**CRKRS**	-	80

RFXANK	-	338	**UFD1L**	-	79
CIT	42	328	**CYB5R1**	-	79
SEPT9 *	-	324	**VAMP4**	-	79
HMOX2	51	323	**ACTR10**	-	79
ERLIN2	-	314	**C19ORF10**	-	78
AP2M1	70	311	**ACAD11**	-	77
SFRS15	-	308	**GGA2**	-	77
MLX	58	304	**MED16**	-	76
CRKL	-	301	**HCK**	-	76
NAPG	68	300	**POLR2I**	-	76
CLASP2	-	287	**NOSIP**	-	76
RAB5B	-	283	**TEX264**	-	75
MEF2A	-	274	**TLN1**	-	75
INTS1	-	270	**TERF2**	-	75
RAB8A	-	269	**UHRF1BP1L**	-	75
hnRP LL	54	261	**MXRA7**	-	75
RBM16	-	260	**ATP5C1**	-	75
ATP6V1D	-	253	**SCRIB**	-	75
PHB2	-	253	**PTPN1**	-	75
RPL14	61	248	**NME1**	-	74
STX5	-	240	**DDX23**	-	74
COPB2	42	238	**MGST3**	-	73
NR2F2	47	235	**BAZ1A**	-	73
MED20	-	234	**TMEM93**	-	73
DENR	-	234	**MED23**	-	72
TMEM43	-	231	**mTOR**	-	72
SMARCE1	-	228	**HAUS2**	-	71
WTAP	54	226	**SH3GLB1**	-	70
MORF4L1	-	224	**ZMPSTE24**	-	70
LOC402643	-	222	**MYCBP**	-	70
FAF2	-	221	**NKRF**	-	70
SLC25A6	-	218	**CXORF56**	-	70
PAF1	44	215	**ENO1**	-	70
ACOT8	48	214	**CWC15**	-	69
PEX14	-	213	**RAB9A**	-	69
DIDO1	-	212	**RAB31**	-	69
TTF2	-	210	**PCNP**	-	69
TMED4	-	210	**ATP6VOC**	-	68
DDX19B	36	209	**RBM4**	-	68
TRIP13	-	209	**ROD1 ***	-	67
DDX46	-	207	**POLR2A**	-	67
PSMB2	-	207	**YRDC**	-	67
MOBKL1B	49	204	**GIGYF2**	-	66
SSSCA1	-	196	**DGKD**	-	66
CDK9	-	195	**SFRS2**	-	66
RBMXL2	-	195	**CDS2**	-	65
LRBA	-	194	**INTS9**	-	65
ZC3H18	-	191	**PRKAG1**	-	65
EPHX1	-	191	**RAB24**	-	65
RALA	-	190	**FILIP1**	-	64
RBM17	-	186	**NUP160**	-	64
CBX3	46	185	**TP53**	-	63

RBM7	-	182	**PRDM2**	-	63
MAD2L1	-	181	**C10ORF58**	-	63
BCAS2 *	36	177	**CPNE8**	-	62
PRPF38A	-	175	**CLPTM1**	-	62
MED31	-	175	**SEP15**	-	62
EXOSC2	32	175	**NEK7**	-	62
PECI	-	172	**POLR3F**	-	61
PARVA	-	172	**KDM1A**	-	61
RAB18	-	167	**PABPC1**	-	60
PPIE	-	166	**DDX6**	-	60
ATXN10	-	165	**PEX3**	-	59
RXRA	-	165	**ORC5L**	-	59
ATP50	-	163	**TSEN15**	-	59
RBM22	-	161	**POP1**	-	58
CCNB1	-	160	**TOMM22**	-	58
FANCI	-	157	**PCYOX1**	-	58
SMC2	-	155	**SPC25**	-	57
SLC25A17	-	154	**ALDOA**	-	57
MCTS1	-	151	**LOC100128355**	-	55
SART1	-	151	**LIN37**	-	57
RPL21	30	150	**SHOC2**	-	57
TAF10	-	148	**IQSEC1**	-	57
TMEM85	-	146	**ENTPD6**	-	57
ORC4L	-	145	**THOC5**	-	56
SAFB	-	145	**APOBEC3D**	-	56
GSTK1	-	140	**SERPINH1**	-	56
PDIA4	-	140	**CCDC101**	-	56
OXSR1	-	138	**POP7**	-	56
DPM3	-	138	**PSMG1**	-	55
PHB	-	134	**ZWINT**	-	55
SLDH3A2	-	132	**DEGS1**	-	55
MTCH2	-	131	**PRPF3**	-	54
hnRNP F	-	130	**MEN1**	-	53
PRKAR2A	-	130	**HAUS5**	-	53
SEPT8	-	128	**ZBTB43**	-	53
LIMS1	-	128	**ERCC1**	-	53
SMARCB1	-	126	**CCNK**	-	53
ZBTB10	-	124	**TTN**	-	52
ATP5B	-	124	**MLF2**	-	52
DNAJC17	-	124	**WDR6**	-	52
SPTLC1	-	123	**NCSTN**	-	52

Table 1 – Mass Spectrometry data of nuclear Upf2-interacting factors in HEK293 cells. Numbers indicated represent mascot scores; only proteins with mascot scores higher than 50 were considered. The control immunoprecipitation was performed with normal IgG. * indicates proteins present in Rod1 nuclear mass spectrometry data (table 2).

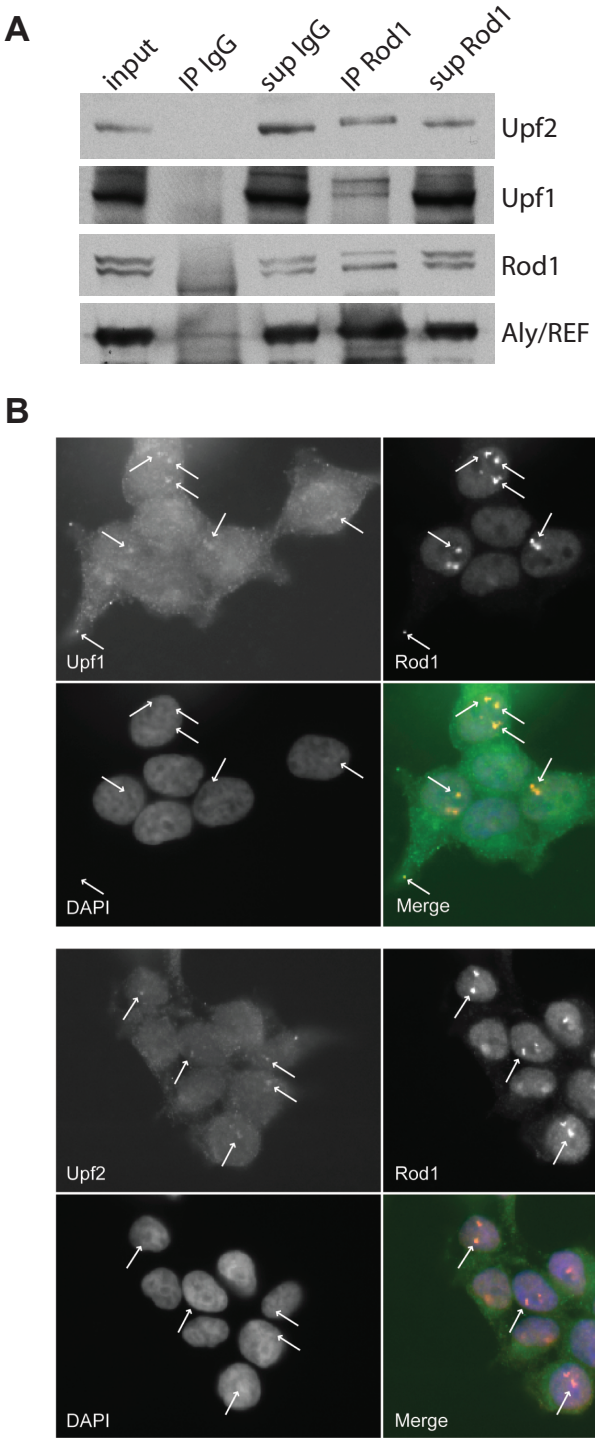


Figure 1 – Rod1 interacts with Upf1, Upf2 and Aly. **A –** Immunoprecipitations from nuclear HEK293 extracts showing interactions of Rod1 with Upf1, Upf2 and Aly. **B –** Rod1 co-localizes with Upf1 (upper set of panels) and Upf2 (lower set of panels) in nuclear speckles of HEK293 cells; in a fraction of cells, Rod1 co-localizes with Upf1 in cytoplasmic speckles. Speckles are indicated with arrowheads.

Protein	score Control	score Rod1	Protein	score Control	score Rod1
MYH10	465	2019	CPSF6	-	129
UPF2 *	-	1203	MARK2	-	126
TUBA	-	1147	TIM44	-	125
TUBA6	-	1055	FLYWCH	-	121
SEC16A	-	953	CTP synthase	-	121
TUBB	-	911	LOC493753	-	120
hnRNP A1	-	841	TRF2	-	120
TIA1	-	809	RER1	-	120
HMMR	-	800	AP-3	-	119
ROD1 *	-	663	RAC3	-	118
UPF1 *	-	654	ZMAT5	-	117
DGCR14	42	603	hCG1782167	-	116
HSP90AA1	-	467	RGSip1	-	115
KIAA1741	102	461	PRPF19	-	115
ARL10C	84	450	snRP E	-	113
RNP1	-	427	Ki-67	-	112
KIAA1826	-	419	DYNLL2	-	112
CSE1	-	414	mtSSB	-	112
RAB1A	-	397	IER3ip1	-	111
PRC1	90	381	ALY	-	110
RAB8B	-	344	APC10	-	108
TDP43	-	328	SEPT9 *	-	108
TAF15	-	327	PPIL1	-	107
Transportin	70	313	TRIP230	-	106
U2AF	-	297	DES	-	104
CLIP-ap1	-	282	RPS11	-	103
Exportin 1	-	278	MGC2803	-	103
DAZap1	63	277	RCBTB1	-	102
ARL1	-	272	RPL11	-	101
MAP7D2	50	260	GFAP	-	100
RPS10	-	257	HOXC8	-	100
RPS14	40	252	LSF	-	99
CDC42	-	244	MARCKS	-	99
RAB10	-	242	LDH-B	-	98
TMED10	-	237	SRI	-	98
SNAP29	45	232	CAD	-	96
PCD7	-	231	SMARCA5	-	93
MAZ	52	230	SSRP1	-	93
NACA	-	226	Cofilin 1	-	92
TMPO	54	226	ASF1	-	91
HOXB9	49	218	HMG-1	-	91
MYL6	47	214	XPO5	-	90
RPS4X	-	212	NUP155	-	87
CUGbp2	-	200	FAM83D	-	87
Annexin A2	-	200	PRA1	-	86
SRP19	-	191	RNA pol II	-	85
RBM9	47	189	RPS27	-	84
BTEbp4	-	186	TRX-1	-	84
Transportin 2	-	183	p53	-	83
VAT1	-	182	GCN1	-	83

U11/U12 snRNP	-	180	YIP1	-	82
GST-PI	-	178	DNAJ	-	81
FIP1I1	-	175	RAMA1	-	79
PCD6	-	170	SEC24B	-	78
CDK2	-	169	COX5B	-	77
RBM39	-	168	HOXA9B	-	76
BCR/ABL	-	168	RPS28	-	76
COL1A1	-	168	HOX7	-	74
CEP55	-	167	TNRC6C	-	74
MYO12A	-	163	U4/U6 snRNP	-	73
C1ORF35	-	161	RPL35A	-	73
RPS19BP1	-	161	HSPC137	-	72
T-PLASTIN	-	157	ARCN1	-	72
TOPBP1	-	155	DNAJC9	-	71
ARP1	-	154	BCL7C	-	71
LSM2	-	154	B99	-	69
HSPC128	-	153	THOC2	-	68
GNL3L	-	150	RAE1	-	67
BACH1	-	147	PRMT1	-	66
NEZHA	-	147	RPL38	-	66
ZFP768	-	145	C11ORF73	-	66
Transferrin	-	145	ASH2L2	-	65
BCAS2 *	-	144	WDR41	-	65
CUL-2	-	144	CAF1B	-	65
LENG1	-	144	SNF5/INI1	-	61
TRIM33B	-	143	BET1 *	-	61
CGI-135	-	142	CDC23	-	60
RAN	-	141	TRAP25	-	59
GATAd2A	-	140	NUP50	-	59
H1d	-	139	MED8	-	58
YKT6 *	-	137	SERPINE1	-	58
RPR1A	-	136	FAM128B	-	55
WDR48	-	134	NCOR1	-	54
Importin 5	-	133	SDP3	-	53
HEY1	-	133	RPS7	-	51
HOXC9	-	132	RPL30	-	51
snRP A	-	131	SF3B10	-	50
LSM14	-	130	NEFM	-	50

Table 2 – Mass Spectrometry data of nuclear Rod1-interacting factors in HEK293 cells. Numbers indicated represent mascot scores; only proteins with mascot scores higher than 50 were considered. The control immunoprecipitation was performed with normal IgG. * indicates proteins present in Upf2 nuclear mass spectrometry data (table 1).

Rod1 is necessary for destabilizing a reporter β-globin PTC-containing transcript: Given the interactions with the NMD factors Upf1 and Upf2, as well as with the EJC factor Aly/REF, we next addressed the possibility of Rod1 involvement in NMD. We transiently transfected HEK293 cells with a reporter β-globin construct containing a NMD-sensitive nonsense mutation in codon 39 (NS39 - [37]), or the *wild-type* construct as control (wt), and performed RNAi against Rod1, Upf2, Upf1 or control. Levels of the NS39 transcript were quantified by northern blotting (fig. 2). The NS39 transcript levels were 19% of the wt in the control RNAi experiment, confirming the degradation of the PTC-carrying transcript (fig. 2C, lanes 1-2). NS39 levels increased to 37%, 35% and 26% upon knock-down of Rod1, Upf2 and Upf1, respectively (fig. 2C, lanes 3-8). We performed two other similar independent experiments with quantification of β-globin transcript levels by RT-PCR and confirmed the stabilization of the NS39 transcript upon knock-down of Rod1 (supplementary fig. 1). These results indicate that Rod1, similarly to Upf2 and Upf1, is required for the NMD-dependent degradation of the NS39 transcript in HEK293 cells.

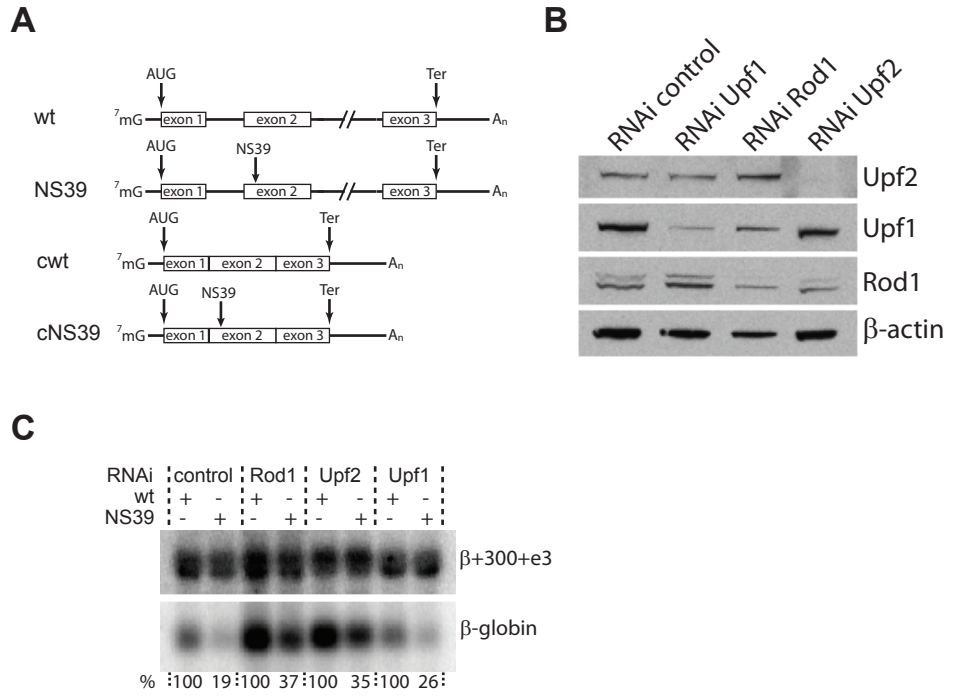


Figure 2 – Rod1 is required for the NMD-dependent destabilization of a reporter β-globin PTC-containing transcript in HEK293 cells. **A –** Reporter β-globin constructs transfected into HEK293 cells consist of the human β-globin coding region containing introns 2 and 3, driven by the CMV promoter; a construct containing a nonsense mutation in codon 39 was used as the NMD-sensitive reporter, while the *wild-type* was used as control. **B –** RNA interference against Rod1, Upf2 and Upf1 in HEK293 cells, confirmed by western blotting. **C –** Northern blot

showing stabilization of the NS39 transcript upon knock-down of Rod1, Upf2 or Upf1; β-globin transcript levels were normalized to the transfection control β+300+e3. Bands were quantified with ImageQuant 5.2.

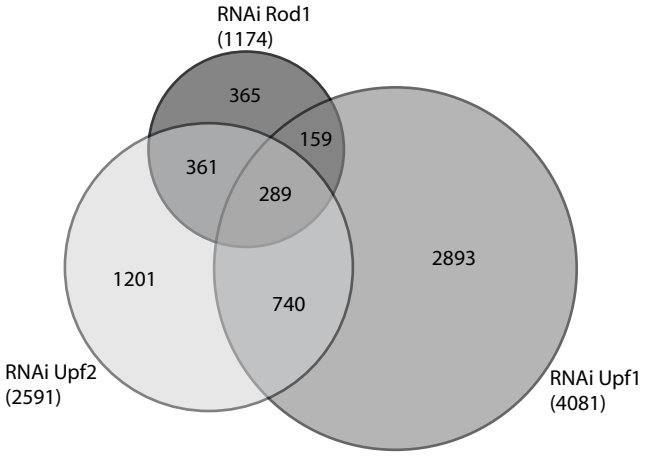
Rod1, Upf2 and Upf1 regulate common potential Nonsense-Mediated mRNA Decay targets: The interactions of Rod1 with Upf2, Upf1 and Aly and the stabilization of a β-globin transcript containing a nonsense mutation upon Rod1 knock-down suggest that Rod1 may regulate physiological NMD targets in HEK293 cells. To verify this, we performed RNAi experiments targeting Rod1, Upf2, Upf1 and a control knock-down, and isolated total RNA which was hybridized to a human exon array platform interrogating over one million exon clusters within the human genome. We performed a Principal Component Analysis (PCA) with the four sample triplicates, which clustered in separate regions in a PCA map; one sample from RNAi against Upf1 was considered an outlier and was excluded from further analysis (supplementary fig. 2A). Several hundreds of transcripts were up or down-regulated more than 1.5 fold upon RNAi against Rod1, Upf2 or Upf1, compared to control (figs. 3A and 3B). Among these, 289 (25% of Rod1-regulated transcripts) were commonly mis-regulated after knock-down of Rod1, Upf2 or Upf1; Rod1 and Upf2 shared 650 (55% of Rod1-regulated transcripts), while Rod1 and Upf1 shared 448 (38% of Rod1-regulated transcripts) target transcripts. Not surprisingly, Upf2 and Upf1 shared an extensive number of target transcripts (1029). Among common target transcripts, there were up-regulated, down-regulated transcripts as well as transcripts regulated in an antagonistic manner, i.e. transcripts down-regulated by one or two factors and up-regulated by the other(s) (fig. 3A). Strikingly, only 8 transcripts were antagonistically regulated among common Rod1 and Upf2 targets, strongly suggesting that both operate in the same mechanistic pathway(s). Potential NMD targets are expected to be up-regulated upon RNAi against Upf1 (not necessarily up-regulated upon RNAi against Upf2). Therefore, 121 transcripts are potential Rod1-mediated NMD targets (10% of Rod1-regulated transcripts), 241 transcripts are potential Upf2-mediated NMD targets (9% of Upf2-regulated transcripts) and 1321 transcripts are potential NMD targets not regulated by Rod1 nor Upf2 (figs. 3A and 3C). We confirmed the up-regulation of several Rod1 and Upf2-regulated potential NMD targets by RT-PCR (supplementary fig. 2B), some of them being previously described NMD targets (GADD45B and ATF3 - [9]). A gene ontology (GO) analysis was made of Rod1, Upf2, Upf1, as well as common target genes to find out whether any particular cellular function was over-represented (supplementary fig. 3C). In line with previous studies which report that NMD targets are involved in diverse cellular functions [12], no specific cellular or developmental function was over-represented in the GO analysis. Altogether, these results show that Rod1, Upf2 and Upf1 regulate hundreds of common transcripts in HEK293 cells. Furthermore, around 10% of Rod1-regulated transcripts are potential NMD targets, since

they are up-regulated upon knock-down of Rod1 or Upf1.

A

Group	Mis-regulated transcripts	Up-regulated	Down-regulated	Antagonistic effect
Rod1	1174	598	576	-
Upf2	2591	1132	1459	-
Upf1	4081	1683	2398	-
Common Rod1-Upf2	650	283	359	8
Common Rod1-Upf1	448	121	179	148
Common Upf2-Upf1	1029	241	507	281
Common Rod1-Upf2-Upf1	289	76	131	82*

B



C

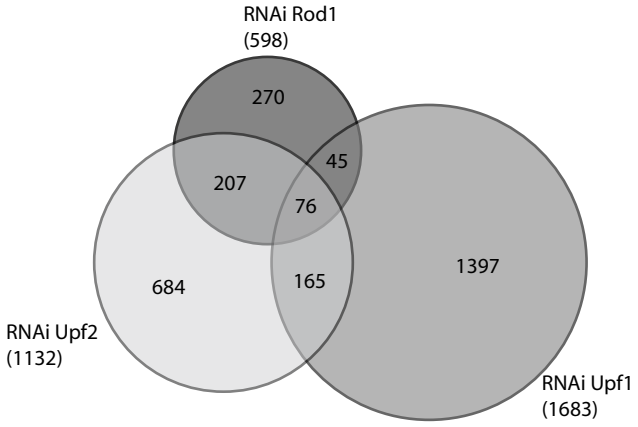


Figure 3 – Rod1, Upf1 and Upf2 regulate common genes. **A –** Numbers of mis-regulated genes (>1,5X) upon RNAi against Rod1, Upf1 or Upf2; numbers of common mis-regulated genes are also indicated; ***** - from these, 76 were antagonistically regulated by Upf1 (relative to both Rod1 and Upf2), 5 by Rod1 (relative to both Upf1 and Upf2) and 1 by Upf2 (relative to both Upf1 and Rod1). **B –** Venn diagram showing numbers of mis-regulated genes upon knock-down of Rod1, Upf1 and Upf2 (areas are proportional to numbers of genes). **C -** Venn diagram showing numbers of up-regulated genes upon knock-down of Rod1, Upf1 and Upf2 (areas are proportional to numbers of genes).

CLIP-seq analysis of Rod1-RNA complexes in HEK293 cells: The interactions of Rod1 with Aly/REF, Upf2 and Upf1 suggest that Rod1 could be part of the EJC. If so, Rod1 is expected to associate with RNA at EJC sites, i.e. approximately 24 nt upstream of each exon-exon boundary. We analysed Rod1-RNA complexes by CLIP-seq (Cross-Linked ImmunoPrecipitation-sequencing – see fig. 3A, chapter 3) to map Rod1 binding sites. Rod1-RNA complexes are visible as a smear above the molecular weight of Rod1 (59 kDa) after SDS-PAGE (fig. 4).

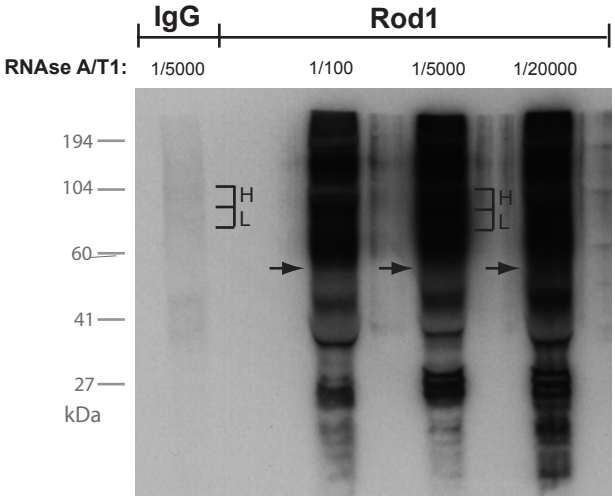


Figure 4 - CLIP-seq (Cross-Linked ImmunoPrecipitation-sequencing) of Rod1-RNA complexes in HEK293 cells. Different RNAse A/T1 mix concentrations were used to confirm that complexes contain RNA (signal is slightly decreased with higher concentration of RNAse A/T1); no significant amount of protein-RNA complexes was isolated with IgG control (left lane); arrowheads indicate the size of free Rod1 bands; two segments of nitrocellulose membrane from the 1/5000 RNAse A/T1 lanes were cut for sequencing, high (H) and low (L).

Discussion

The results presented in this chapter demonstrate a role for Rod1 in Nonsense-Mediated mRNA Decay (NMD) in HEK293 (Human Embryonic Kidney) cells. Although Rod1 is expressed in several embryonic tissues, in adults its expression is restricted to haematopoietic lineages [33]. This fact suggests that NMD functions through different mechanisms, or

requires a different set of factors, across development and tissues. Different branches of the NMD mechanism have been previously described [32, 38]. In addition, the efficiency of NMD is known to vary across different tissues, as does the cellular abundance of diverse NMD effectors [39-41]. Our results therefore support the concept of heterogeneity of the NMD mechanism.

How does Rod1 contribute to NMD? Rod1-interacting factors derived from mass spectrometry data (table 2 and supplementary table 1) include several translation and ribosomal proteins, suggesting that requirement of Rod1 for NMD could mirror a translational effect. Another possibility emerges from the cellular distribution of Rod1: the cytoplasmic speckles observed in some cells co-localize with Upf1 and could constitute sites of mRNA degradation. Rod1 could thus operate as a molecular linker, connecting NMD target transcripts to sites of mRNA decay. Further experiments are needed to identify the nature of these speckles, namely testing whether they coincide with P-body components, such as GW182, Dcp1 or Dcp2 [42-45]. In support of this hypothesis, the cytoplasmic mass spectrometry data (supplementary table 1) suggests that Rod1 is intimately associated with the cytoplasmic cytoskeleton, as are stress granules and P-bodies [46, 47]. Finally, given the presence of nuclear speckles containing Rod1 and Upf1 and/or Upf2 in HEK293 cells (fig. 1B), a nuclear NMD pathway must be considered. Nuclear NMD has been described in some studies [48-55], but its mechanism remains elusive and difficult to reconcile with the fact that NMD requires a stop codon "reader", associated with translation. Nevertheless, this possibility deserves further investigation. Inhibiting mRNA export and verifying whether NMD still occurs or not could provide an answer.

Several lines of evidence, namely mass spectrometry data, immunoprecipitations and immunocytochemistry (tables 1 and 2, fig. 1), indicate that a relatively small fraction of cellular Upf2 and Upf1 interact with Rod1, while a significant fraction of Rod1 interacts with Upf2 and Upf1. This may be related to the fact that at least Upf1 participates in several other mechanisms, such as histone degradation, telomere maintenance or Staufen-mediated decay (SMD) [56-59]. The relatively high number of transcripts regulated in an antagonistic manner among common Upf1-Upf2 and Upf1-Rod1-regulated transcripts also underscores the Upf1 participation in other cellular functions, besides NMD. Our Upf2 mass spectrometry results suggest that Upf2 may also participate in other cellular events. The exon array data (fig. 3B) further highlights the notion that Rod1 is intimately associated with Upf1 and Upf2: ~69% of Rod1-regulated genes are regulated by Upf1, Upf2, or both. Conversely, ~11% of Upf1-regulated genes and ~25% of Upf2-regulated genes are also regulated by Rod1. The profile of exon array data therefore mirrors the interaction patterns between Rod1, Upf2 and Upf1. An intriguing result from the exon

array data is that only ~23% of Upf1 and Upf2 co-regulated genes were up-regulated upon RNAi, despite both proteins being best characterized as NMD factors; a striking 49% of co-regulated genes were down-regulated. While these could reflect secondary events, it is possible that Upf1 and Upf2 function together in a distinct mechanism which stabilizes mRNAs. Finally, our exon array data suggests that Rod1 may not regulate all cellular NMD target genes: 241 genes were commonly up-regulated upon knock-down of Upf1 or Upf2, but not Rod1 (fig. 3A).

Materials and Methods

Plasmid Constructs and antibodies: pTre-pur-β*wt*, pTre-pur-β39 and pΤ-TA (required for β-globin expression from the pTre-pur constructs – tet-off system) were previously described in [23]; pCI-neo-β+300+e3 was kindly provided by Niels Gehring [60]. The following antibodies were used in this study: mouse monoclonal anti-Rod1 (Santa Cruz cat. no. sc-100845), goat polyclonal anti-Upf2 (Santa Cruz cat. no. sc-20227), goat polyclonal anti-Upf1 (Bethyl cat. no. A300-036A) and mouse monoclonal anti-Aly/REF (Millipore cat. no. 05-1510).

Cell culture, plasmid transfections and RNAi: HEK293 cells were grown in DMEM supplemented with 10% fetal calf serum; transfections were done with Lipofectamine 2000 according to manufacturer's instructions; for the NMD assay, transfections were performed in 6-well plates with 0,8 μg of pTre-pur-β (*wt* or NS39), 0,8 μg of peYFP-C1 (Clontech) and 2,4 μg of pΤ-TA in combination with RNAi. RNAi was performed using lentiviral particles containing shRNA driven by the U6 promotor (pLK0.1-shRNA); lentivirus were produced by transfections in 293T cells according to standard protocols [61]; for transfection assays, transduction of cells with lentivirus was done 72h before transfection; cells were harvested 48h after transfection for RNA or protein extractions. shRNA target sequences are as follows: ROD1: 5'-GCCGTTACTATGGTGAATTAT-3'; UPF1: 5'-GCTGAGTTGAACTTCGAGGAA-3'; UPF2: 5'-GCGAGATACGTCACAATGGTA-3'; control: 5'-ATTCTCCGAACGTGTCACG-3'.

RNA extractions and northern blotting: Total RNA extracts were done with Trizol (Invitrogen) according to the manufacturer's instructions; 15 μg RNA per sample were mixed with equal volume of loading buffer (1X MOPS, 18.5% formaldehyde, 50% formamide, 4% ficoll400, bromophenolblue) and incubated at 90ºC for 5 min.; a 1.3% agarose gel (in 1X MOPS buffer – 40mM MOPS, 10mM Na-Acetate, 1mM EDTA, pH7.2) with 0.6M formaldehyde was run in 1X MOPS buffer until loading dye reached the bottom

of the gel; the gel was rinsed with 50mM NaOH before being soaked in 100mM NaCl (20 min.), 100mM Tris-HCl (20 min.) and 2X SSC (0.3M NaCl, 0.03M Na-citrate) (20 min.); capillary transfer to a Hybond-XL nylon membrane (GE Healthcare) was done overnight in 10X SSC; UV crosslinking was performed using the Stratalinker 1800 UV crosslinker (Stratagene) followed by incubation in hybridization buffer (50mM Tris-HCl pH 7.5, 1M NaCl, 50% formamide, 5X Denhardt's, 0.1% SDS, 5% dextransulfate, 100 µg/ml salmon sperm DNA) for 30 min. at 65ºC; radioactive probes were added followed by incubation overnight at 65ºC; the membrane was washed in 3X SSC + 0.5% SDS followed by 0.3X SSC + 0.5% SDS until background radioactive signal was low; membrane exposure was left for 72h and analysed in a Typhoon 2000 phosphoimager (GE Healthcare). Bands were quantified with Image Quant, version 5.2. Radioactive probes were prepared by PCR using α-^{32}P-dATP.

Protein extractions, immunoprecipitations and mass spectrometry: HEK293 cells were washed in cold PBS and protein extracts were performed as described in [62], followed by dialysis in buffer C-100 (20 mM HEPES pH 7.6, 0.2 mM EDTA, 1,5 mM $MgCl_2$, 100 mM KCl, 20% glycerol). Extracts were centrifuged at 16000g prior to immunoprecipitations to remove insoluble precipitates. For immunoprecipitations, Dynabeads coupled with protein G (Invitrogen) were washed several times with PBS and blocked with 200 µg/ml chicken egg albumin for 1 hr at room temperature; 10 µg control IgG, mouse anti-Rod1 (F-30 clone from Santa Cruz) or goat anti-Upf2 (clone C18 from Santa Cruz) were then incubated with blocked beads for 30 min. at R.T.; nuclear or cytoplasmic extracts were incubated at 4°C with benzonase (150 u/ml) and protease inhibitors (Complete, Roche). Extracts were incubated with beads and antibody for 2 hr at 4°C, beads were washed 5X with buffer C-100 + Complete + 0.02% NP-40 and immunoprecipitated protein complexes eluted in Laemli buffer at 99°C for 5 min. For Mass Spectrometry, samples were loaded into a 10% SDS-PA gel followed by Colloidal Blue Staining (Invitrogen); lanes were cut and submitted to in-gel digestion with trypsin (Promega) as previously described [63]. Nanoflow liquid chromatography-tandem mass spectrometry was performed on a 1100 series capillary liquid chromatography system (Agilent Technologies) coupled to an LTQ-orbitrap mass spectrometer (Thermo) as previously described [64]. Matching peptide fragmentation spectra to databases was performed with Mascot as previously described [64].

Exon arrays: Total RNA extracted with Trizol was purified using the Qiagen RNeasy kit; RNA quality was assessed using the Agilent Bioanalyser. Further processing of samples was performed according to the Affymetrix GeneChip WT Sense Target Labeling Assay. Affymetrix GeneChip Human Exon 1.0 ST arrays were used to determine the expression

level of virtually all exons in the human genome (1.4 million probe sets covering >1 million exon clusters). Data was analysed with Partek Genomics Suite 6.4; background correction and normalization of probe set intensities were done using the Robust Multi-array Analysis and GC content was taken into account (GC-RMA) [65]; probe set summarization was performed with median polish settings. Exon-level data was filtered to include only those probe sets that are in the "core" meta-probe list. Differentially expressed genes were detected using ANOVA (ANalysis Of VAriance); a cutoff value of 1.5X was used to filter for up or downregulated genes. Up and downregulated genes were classified according to molecular function using the Ingenuity Pathway Analysis software.

CLIP-seq: CLIP-seq experiments were performed essentially as described in [66] with the following modifications: RNAse A/T1 mix (Fermentas) was used at a final concentration of 0.5 µg/ml RNAse A and 0.25 U/ml RNAse T1 instead of RNAse A; 5' and 3' linkers were DNA/RNA oligonucleotide hybrids described in [67]; we used an anti-Rod1 antibody (Santa Cruz) for immunoprecipitation of protein-RNA complexes. We used an Illumina GAIIx sequencer.

References

1. Nicholson, P., et al., *Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors.* Cell Mol Life Sci, 2009.
2. Isken, O. and L.E. Maquat, *The multiple lives of NMD factors: balancing roles in gene and genome regulation.* Nat Rev Genet, 2008.
3. Rehwinkel, J., J. Raes, and E. Izaurralde, *Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors.* Trends Biochem Sci, 2006. **31**(11): p. 639-46.
4. Moriarty, P.M., C.C. Reddy, and L.E. Maquat, *Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA.* Mol Cell Biol, 1998. **18**(5): p. 2932-9.
5. Buhler, M., et al., *EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length.* Nat Struct Mol Biol, 2006. **13**(5): p. 462-4.
6. Amrani, N., et al., *A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay.* Nature, 2004. **432**(7013): p. 112-8.
7. Eberle, A.B., et al., *Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region.* PLoS Biol, 2008. **6**(4): p. e92.
8. Muhlrad, D. and R. Parker, *Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance.* RNA, 1999. **5**(10): p. 1299-307.
9. Mendell, J.T., et al., *Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise.* Nat Genet, 2004. **36**(10): p. 1073-8.
10. Singh, G., I. Rebbapragada, and J. Lykke-Andersen, *A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay.* PLoS Biol, 2008. **6**(4): p. e111.
11. Wittmann, J., E.M. Hol, and H.M. Jack, *hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay.* Mol Cell Biol, 2006. **26**(4): p. 1272-87.
12. Rehwinkel, J., et al., *Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets.* RNA, 2005. **11**(10): p. 1530-44.
13. Diem, M.D., et al., *PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs.* Nat Struct Mol Biol, 2007. **14**(12): p. 1173-9.
14. Kataoka, N., et al., *Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm.* Mol Cell, 2000. **6**(3): p. 673-82.
15. Le Hir, H., et al., *The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay.* EMBO J, 2001. **20**(17): p. 4987-97.
16. Le Hir, H., et al., *The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions.* EMBO J, 2000. **19**(24): p. 6860-9.

17. Luo, M.L., et al., *Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly.* Nature, 2001. **413**(6856): p. 644-7.
18. Nott, A., H. Le Hir, and M.J. Moore, *Splicing enhances translation in mammalian cells: an additional function of the exon junction complex.* Genes Dev, 2004. **18**(2): p. 210-22.
19. Shibuya, T., et al., *eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay.* Nat Struct Mol Biol, 2004. **11**(4): p. 346-51.
20. Zhou, Z., et al., *The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans.* Nature, 2000. **407**(6802): p. 401-5.
21. Isken, O. and L.E. Maquat, *Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function.* Genes Dev, 2007. **21**(15): p. 1833-56.
22. Ishigaki, Y., et al., *Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20.* Cell, 2001. **106**(5): p. 607-17.
23. Inacio, A., et al., *Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay.* J Biol Chem, 2004. **279**(31): p. 32170-80.
24. Kashima, I., et al., *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, 2006. **20**(3): p. 355-67.
25. Ivanov, P.V., et al., *Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways.* EMBO J, 2008. **27**(5): p. 736-47.
26. Yamashita, A., et al., *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, 2001. **15**(17): p. 2215-28.
27. Fukuhara, N., et al., *SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway.* Mol Cell, 2005. **17**(4): p. 537-47.
28. Unterholzner, L. and E. Izaurralde, *SMG7 acts as a molecular link between mRNA surveillance and mRNA decay.* Mol Cell, 2004. **16**(4): p. 587-96.
29. Lejeune, F., X. Li, and L.E. Maquat, *Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities.* Mol Cell, 2003. **12**(3): p. 675-87.
30. Eberle, A.B., et al., *SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells.* Nat Struct Mol Biol, 2009. **16**(1): p. 49-55.
31. Huntzinger, E., et al., *SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan.* RNA, 2008. **14**(12): p. 2609-17.
32. Gehring, N.H., et al., *Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements.* Mol Cell, 2005. **20**(1): p. 65-75.
33. Yamamoto, H., et al., *Isolation of a mammalian homologue of a fission yeast differentiation regulator.* Mol Cell Biol, 1999. **19**(5): p. 3829-41.
34. Mendell, J.T., et al., *Novel Upf2p orthologues suggest a functional link between*

translation initiation and nonsense surveillance complexes. Mol Cell Biol, 2000. **20**(23): p. 8944-57.

35. Serin, G., et al., *Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4).* Mol Cell Biol, 2001. **21**(1): p. 209-23.
36. Mendell, J.T., C.M. ap Rhys, and H.C. Dietz, *Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts.* Science, 2002. **298**(5592): p. 419-22.
37. Maquat, L.E., et al., *Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia.* Cell, 1981. **27**(3 Pt 2): p. 543-53.
38. Chan, W.K., et al., *An alternative branch of the nonsense-mediated decay pathway.* EMBO J, 2007. **26**(7): p. 1820-30.
39. Linde, L., et al., *The efficiency of nonsense-mediated mRNA decay is an inherent character and varies among different cells.* Eur J Hum Genet, 2007. **15**(11): p. 1156-62.
40. Viegas, M.H., et al., *The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway.* Nucleic Acids Res, 2007. **35**(13): p. 4542-51.
41. Maquat, L.E. and G. Serin, *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, 2001. **66**: p. 313-20.
42. Sheth, U. and R. Parker, *Decapping and decay of messenger RNA occur in cytoplasmic processing bodies.* Science, 2003. **300**(5620): p. 805-8.
43. Eystathioy, T., et al., *A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles.* Mol Biol Cell, 2002. **13**(4): p. 1338-51.
44. van Dijk, E., et al., *Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures.* EMBO J, 2002. **21**(24): p. 6915-24.
45. Eystathioy, T., et al., *The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies.* RNA, 2003. **9**(10): p. 1171-3.
46. Aizer, A., et al., *The dynamics of mammalian P body transport, assembly, and disassembly in vivo.* Mol Biol Cell, 2008. **19**(10): p. 4154-66.
47. Loschi, M., et al., *Dynein and kinesin regulate stress-granule and P-body dynamics.* J Cell Sci, 2009. **122**(Pt 21): p. 3973-82.
48. Buhler, M., M.F. Wilkinson, and O. Muhlemann, *Intranuclear degradation of nonsense codon-containing mRNA.* EMBO Rep, 2002. **3**(7): p. 646-51.
49. Aoufouchi, S., J. Yelamos, and C. Milstein, *Nonsense mutations inhibit RNA splicing in a cell-free system: recognition of mutant codon is independent of protein synthesis.* Cell, 1996. **85**(3): p. 415-22.
50. Baserga, S.J. and E.J. Benz, Jr., *Beta-globin nonsense mutation: deficient accumulation of mRNA occurs despite normal cytoplasmic stability.* Proc Natl Acad Sci U S A, 1992. **89**(7): p. 2935-9.
51. Belgrader, P. and L.E. Maquat, *Nonsense but not missense mutations can decrease the abundance of nuclear mRNA for the mouse major urinary protein, while both*

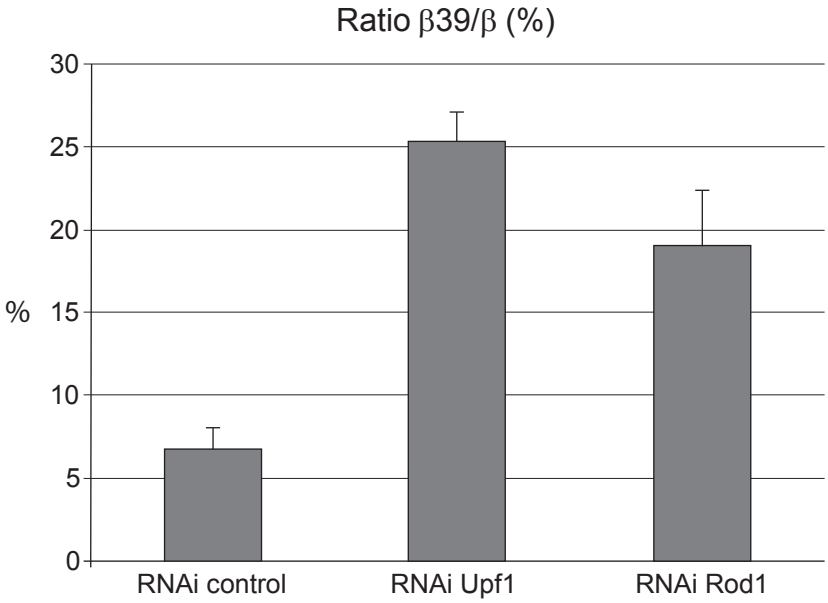
types of mutations can facilitate exon skipping. Mol Cell Biol, 1994. **14**(9): p. 6326-36.
52. Carter, M.S., et al., *A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro.* J Biol Chem, 1995. **270**(48): p. 28995-9003.
53. Kessler, O. and L.A. Chasin, *Effects of nonsense mutations on nuclear and cytoplasmic adenine phosphoribosyltransferase RNA.* Mol Cell Biol, 1996. **16**(8): p. 4426-35.
54. Li, S., D. Leonard, and M.F. Wilkinson, *T cell receptor (TCR) mini-gene mRNA expression regulated by nonsense codons: a nuclear-associated translation-like mechanism.* J Exp Med, 1997. **185**(6): p. 985-92.
55. Urlaub, G., et al., *Nonsense mutations in the dihydrofolate reductase gene affect RNA processing.* Mol Cell Biol, 1989. **9**(7): p. 2868-80.
56. Azzalin, C.M. and J. Lingner, *The human RNA surveillance factor UPF1 is required for S phase progression and genome stability.* Curr Biol, 2006. **16**(4): p. 433-9.
57. Azzalin, C.M., et al., *Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends.* Science, 2007. **318**(5851): p. 798-801.
58. Kaygun, H. and W.F. Marzluff, *Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1.* Nat Struct Mol Biol, 2005. **12**(9): p. 794-800.
59. Kim, Y.K., et al., *Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay.* Cell, 2005. **120**(2): p. 195-208.
60. Gehring, N.H., et al., *Y14 and hUpf3b form an NMD-activating complex.* Mol Cell, 2003. **11**(4): p. 939-49.
61. Zufferey, R., et al., *Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo.* Nat Biotechnol, 1997. **15**(9): p. 871-5.
62. Dignam, J.D., R.M. Lebovitz, and R.G. Roeder, *Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.* Nucleic Acids Res, 1983. **11**(5): p. 1475-89.
63. Wilm, M., et al., *Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry.* Nature, 1996. **379**(6564): p. 466-9.
64. Sanchez, C., et al., *Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor.* Mol Cell Proteomics, 2007. **6**(5): p. 820-34.
65. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data.* Biostatistics, 2003. **4**(2): p. 249-64.
66. Ule, J., et al., *CLIP: a method for identifying protein-RNA interaction sites in living cells.* Methods, 2005. **37**(4): p. 376-86.
67. Mamanova, L., et al., *FRT-seq: amplification-free, strand-specific transcriptome sequencing.* Nat Methods. **7**(2): p. 130-2.

Supplementary data

Protein	score Control	score Rod1	Protein	score Control	score Rod1
FAM83H	-	2312	CTNNA1	-	93
HMMR	-	1762	FRG1	-	92
HLTF	-	1484	GPR98	-	92
MATR3	-	1184	TARDBP	-	92
CAMSAP1L1	-	990	PPIH	-	90
HSPA1A	-	934	TCF20	-	85
CCDC77	-	741	OGT	-	82
UNC45A	-	695	XRCC6	-	80
FAM83D	-	689	hnRNP UI1	-	80
BACH1	-	664	SUPT16H	-	80
FLII	-	659	VIPAR	-	79
HSP90AA1	-	568	snRNP48	-	74
TRIM33	-	513	FAM110B	-	72
CDC14A	-	508	AXIN1	-	72
ROD1	-	500	FAM64A	-	70
SF3b1	92	467	Plectin	-	69
TRIM21	47	426	ERI1	-	69
PTBP1	-	424	Cyclin T1	-	68
CELF1	-	423	RALY	-	68
hnRNP H1	62	392	CDC42	-	67
RNP1	-	381	SCN3A	-	67
CSNK1A1	43	365	IQCB1	-	67
CLASP2	-	354	GIPC1	-	66
APC	71	350	HIRA	-	66
ARHGEF2	-	345	MSH6	-	66
RPL9	-	344	hnRNP F	-	65
KIAA1543	-	323	AIMP2	-	64
SACS	-	321	PCMT1	-	63
CDK17	-	305	PCM1	-	63
MAP2	-	304	RB1CC1	-	62
LARS	47	275	MYBbp1A	-	61
SF3b2	-	253	YWHAH	-	60
UPF1	-	249	TMEM33	-	60
RRP36	-	228	SF3a1	-	59
VPS33B	-	211	BANF1	-	59
XRCC5	-	209	IQGAP3	-	59
DYNLL2	-	198	KBTBD5	-	58
Hist1H1e	-	185	PRPF3	-	57
TAF15	-	180	GRIA4	-	57
HSD17B4	-	179	ARHGEF17	-	56
CSNK1E	-	176	GSK3b	-	56
CTNNB1	-	175	NOL8	-	55
RBM4	-	168	TEX15	-	55
RBM14	-	168	DIAPH3	-	55
PUF60	-	163	FYTTD1	-	54
PUM1	-	159	SRRM1	-	54
STIP1	-	152	RBMS1	-	54

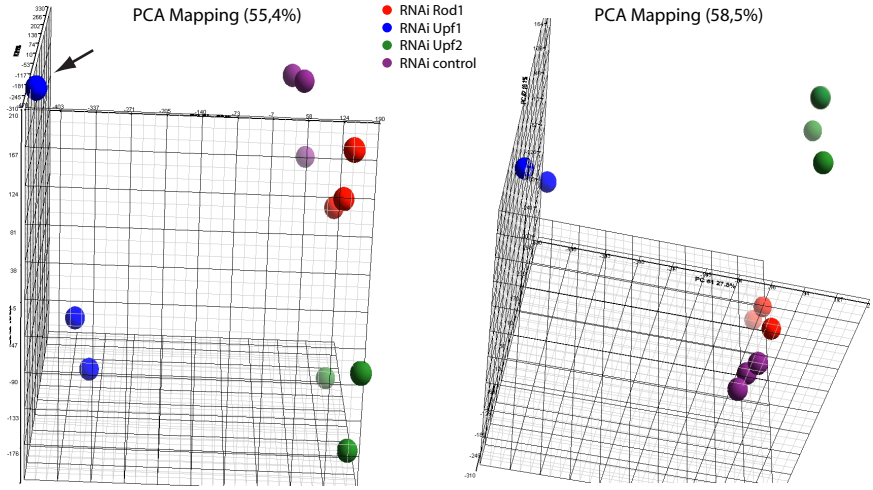
PDCD7	-	148	**LOC100291593**	-	53
LRRFIP2	-	148	**HERC1**	-	53
MYH2	-	146	**MAGOH**	-	53
PTBP2 (nPTB)	-	126	**LOC100132738**	-	53
CSNK1D	-	125	**LOC100288473**	-	52
NARF	-	124	**LAGE3**	-	52
C16ORF48	-	117	**OTUD4**	-	52
SFRS14	-	116	**BBS9**	-	52
SETX	-	116	**CROCC**	-	52
ZMAT5	-	115	**XIRP1**	-	51
HuR	-	115	**XPC**	-	51
CELF2	-	113	**eIF4G2**	-	51
HSPB1	-	105	**PAPD7**	-	51
PHF5A	-	100	**TMTC2**	-	51

Supplementary table 1 – Mass Spectrometry data of cytoplasmic Rod1-interacting factors in HEK293 cells. Numbers indicated represent mascot scores; only proteins with mascot scores higher than 50 were considered. The control immunoprecipitation was performed with normal IgG.

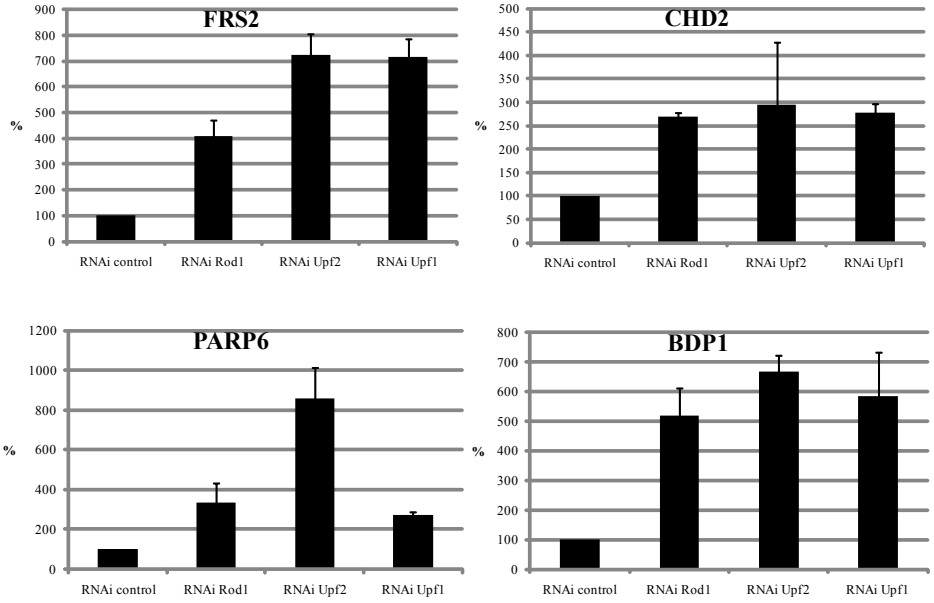


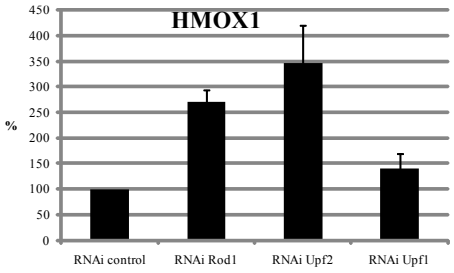
Supplementary figure 1 – Rod1 is necessary for destabilization of a reporter β-globin transcript carrying a NMD-sensitive nonsense mutation in codon 39 (NS39). While NS39 levels are ~7% of the *wild-type* in the control RNAi experiment, they increase to ~25% and ~19% upon knock-down of Upf1 and Rod1, respectively. Transcript levels were quantified by RT-PCR.

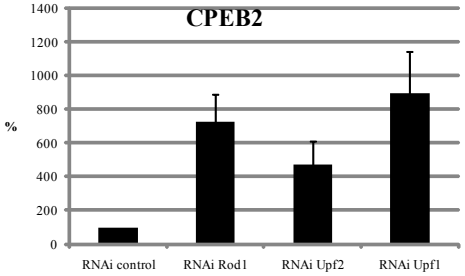
A

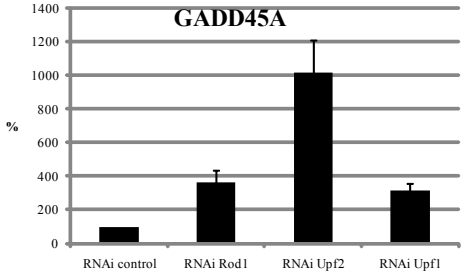

PCA Mapping (55,4%)
RNAi Rod1
RNAi Upf1
RNAi Upf2
RNAi control
PCA Mapping (58,5%)

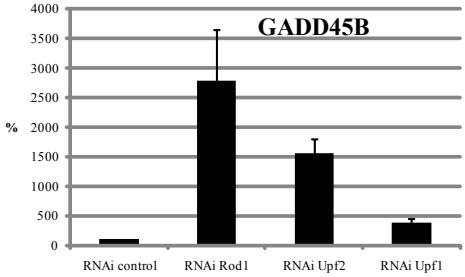
B

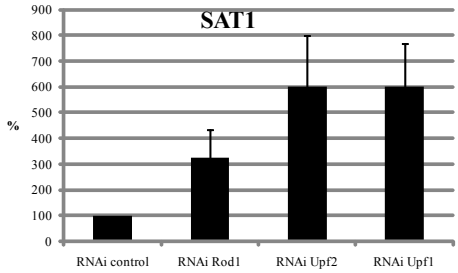

FRS2
CHD2
PARP6
BDP1
%
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1

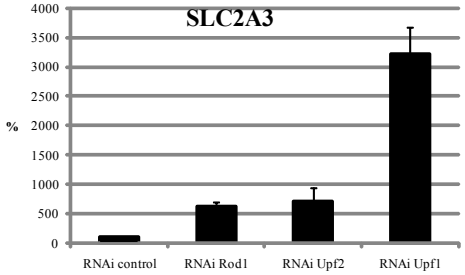

HMOX1
%
450
400
350
300
250
200
150
100
50
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1

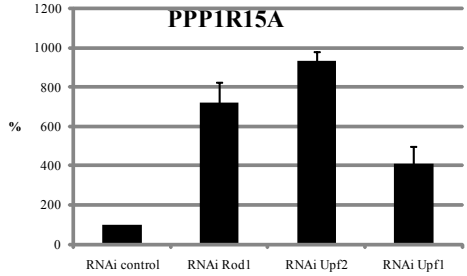

CPEB2
%
1400
1200
1000
800
600
400
200
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1

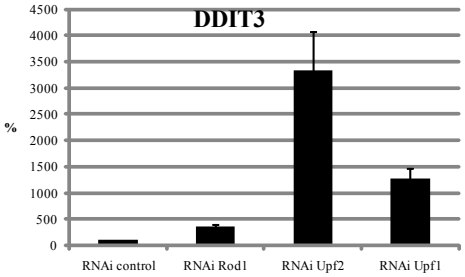

GADD45A
%
1400
1200
1000
800
600
400
200
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1

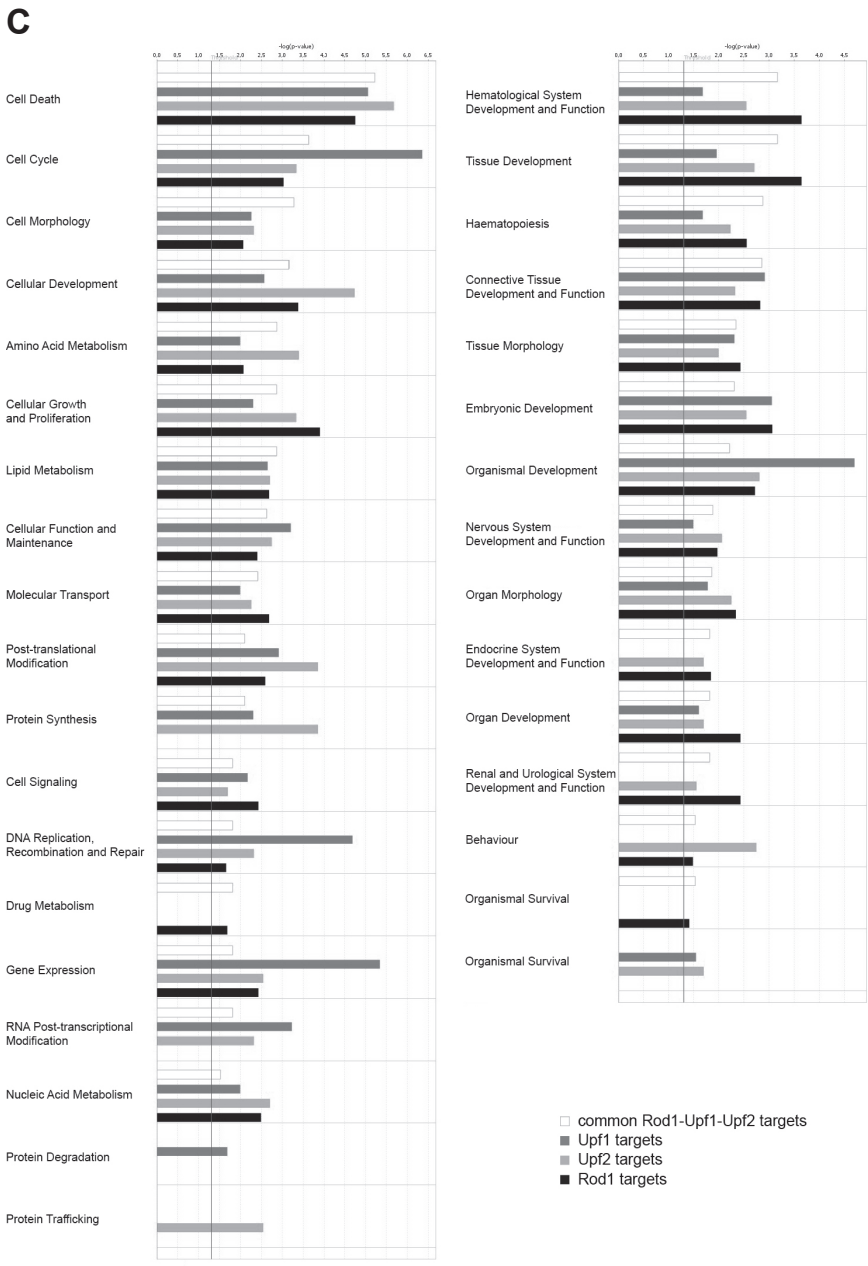

GADD45B
%
4000
3500
3000
2500
2000
1500
1000
500
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1


SAT1
%
900
800
700
600
500
400
300
200
100
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1


SLC2A3
%
4000
3500
3000
2500
2000
1500
1000
500
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1


PPP1R15A
%
1200
1000
800
600
400
200
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1


DDIT3
%
4500
4000
3500
3000
2500
2000
1500
1000
500
0
RNAi control
RNAi Rod1
RNAi Upf2
RNAi Upf1



Supplementary figure 2 – Rod1, Upf2 and Upf1 regulate common (potential) Nonsense-Mediated mRNA Decay target genes. **A –** Principal Component Analysis (PCA) map of the exon array sample triplicates; one sample of the Upf1 RNAi (arrowhead) was interpreted as an outlier and excluded from further analysis. The resultant PCA map is shown on the right. **B –** Verification of (potential) NMD targets, up-regulated upon RNAi against Rod1, Upf2 or Upf1, by RT-PCR. **C –** Gene Ontology (GO) analysis of Rod1, Upf1 and Upf2-regulated genes, as well as common regulated genes; both cellular (left) and developmental (right) functions were considered.

Chapter 3

A global analysis of Alternative Splicing regulated by Rod1, Ptbp1 and Raver1 in erythroid cells

work in progress

Summary

Alternative Splicing (AS) is responsible for the generation of tissue-specific protein isoforms which contribute to cellular identity. Numerous ubiquitous and tissue-specific AS factors have been described which either activate or inhibit splicing of specific exons. Rod1 is a paralog of Ptbp1 predominantly expressed in haematopoietic cells in adult organisms. Rod1 and Ptbp1 share similar protein architectures, with four RNA recognition motifs (RRMs), and a significant homology at the amino acid level. This suggests that, like Ptbp1, Rod1 may regulate AS in tissues where it is expressed. Here, we present mass spectrometry data showing that Rod1 interacts with several splicing and AS factors, including Ptbp1 and Raver1, in mouse erythroleukemia (MEL) cells. Although these interactions exist in both undifferentiated and differentiated MEL cells, we observe through exon array analysis that Rod1 AS targets overlap with those of Ptbp1 and Raver1 only in differentiated MEL cells. In addition to Rod1, Ptbp1 and Raver1 regulating the same exons, they also have the same effect on AS events, i.e. either promoting or inhibiting exon splicing. Altogether, these results demonstrate that many AS events in differentiated MEL cells require Rod1, Ptbp1 and Raver1.

Introduction

Alternative Splicing (AS) is a major generator of protein diversity in metazoans. A recent study estimates that 92-94% of human genes show alternative splicing [1]. A large variety of AS factors act in concert with the basal splicing machinery to regulate the splicing of a vast number of exons, promoting exon inclusion or skipping. Removal or inclusion of exonic sequences potentially changes structural or functional properties of proteins and in some occasions precludes protein production altogether. Indeed, AS events often generate transcripts containing premature termination codons (PTCs) which are targeted for degradation by the mechanism of Nonsense-Mediated mRNA Decay (NMD) [2, 3]. While some AS factors are ubiquitous, others are expressed only in specific tissues and therefore contribute to tissue identity. Examples of tissue-specific AS factors include nPTB, NOVA1, NOVA2 and Hu/ELAV proteins (expressed in neurons), RBM35a and RBM35b (expressed in epithelial cells), FOX1 and FOX2 (expressed in muscle, heart and neurons) or MBNL (expressed in muscle, uterus and ovaries) [4-10].

Rod1 is a Ptbp1 (polypyrimidine tract binding protein 1) paralog expressed in some embryonic tissues (mainly kidney, thymus, liver and lung) and adult haematopoietic cells [11, 12]. Rod1 and Ptbp1 are highly homologous at the amino acid level (>70%) and both

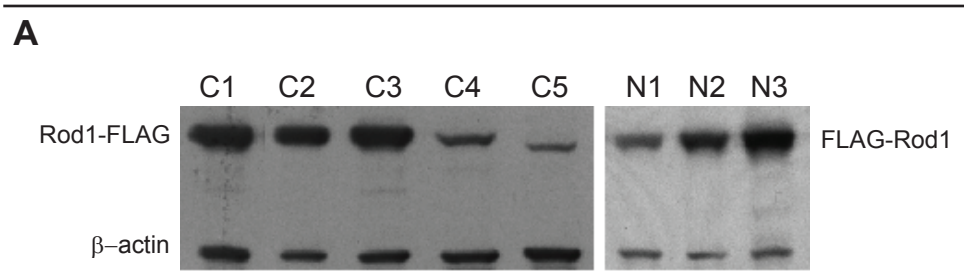
contain four RRM-type domains, suggesting a high functional overlap. Ptbp1 has been reported to regulate several AS events in different cell types, mainly by promoting exon skipping [13-15]; some of these Ptbp1-regulated AS events are likely to require previous recruitment and binding of Raver1 to the RNA, as has been shown for the Tpm1 gene [16, 17].

Here we evaluate the role of Rod1 in the regulation of AS in mouse erythroleukemia (MEL) cells. By mass spectrometry analysis of Rod1-containing complexes, we show that Rod1 interacts with a vast number of splicing and AS factors, including Ptbp1 and Raver1. Interactions with these factors were verified in both undifferentiated and differentiated MEL cells. We analysed AS events by knocking down Rod1, Ptbp1 and Raver1 and performing exon arrays. The results show that there is a large overlap between Rod1, Ptbp1 and Raver1-regulated AS events in differentiated, but not undifferentiated MEL cells. In addition, we observed that the vast majority of AS events are regulated agonistically by these three factors (i.e. all repress or all promote exon inclusion). We performed CLIP-seq experiments to map Rod1 and Ptbp1 RNA-binding sites in differentiated MEL cells. In combination with RNAi against Ptbp1 or Raver1, CLIP-seq experiments show that the amount of Rod1-RNA complexes is significantly reduced upon knock-down of Ptbp1, indicating that, at least for some sites, Rod1 binding to RNA is dependent on Ptbp1. We conclude that many AS events are co-regulated by Rod1, Ptbp1 and Raver1 in erythroid cells.

Results

Rod1 interacts with Ptbp1 and Raver1 in erythroid cells: Rod1 is expressed mainly in haematopoietic cells in adult organisms. To better understand its role in these cells we searched for Rod1-interacting partners in erythroleukemia (MEL) cells. We generated two cell lines stably expressing an N-terminal or C-terminal FLAG-tagged Rod1 (clones C3 and N2, fig. 1A). From these two cell lines, nuclear extracts were made from both undifferentiated and differentiated (by adding 2% DMSO) MEL cells, followed by FLAG pull-downs and analysis of the eluted protein complexes by mass spectrometry. Results are shown in table 1. Pull-downs of N-terminal and C-terminal FLAG-Rod1 fusion proteins generated very similar results, with an overlap of 70% in undifferentiated cells and 83% in differentiated cells. There was also a significant overlap between Rod1-interactors in undifferentiated and differentiated cells (63%). A high number of splicing factors were present in both, altogether representing a significant proportion of Rod1-interacting

partners (fig. 1B). Interestingly, Ptbp1, a Rod1 paralog, and Raver1, a Ptbp1-interacting protein, were present with significantly high mascot scores. These interactions were further confirmed by immunoprecipitations (fig. 1C). The pattern of the Rod1-Ptbp1 interaction in both undifferentiated and differentiated MEL cells is heterogeneous across a given cell population, as judged by immunocytochemistry, suggesting it may be subject to regulation throughout the cell cycle (supplementary fig. 1). To complement these results, we immunopurified Ptbp1 complexes and analysed these by mass spectrometry (table 2). Rod1 was present in both undifferentiated and differentiated cells, further confirming the Rod1-Ptbp1 interaction. Raver1 was present with a high mascot score, but equally so in the IgG control immunoprecipitation. Nevertheless, the Ptbp1-Raver1 interaction was verified by immunoprecipitation followed by SDS-PAGE (fig. 1C). Binding of Raver1 to Ptbp1 has previously been reported in HeLa and skeletal muscle cells [16-18]. Interestingly, many factors interacting with Rod1 also interact with Ptbp1: in undifferentiated MEL cells 38 out of 98 (39%) Rod1-interactors interact also with Ptbp1, while in differentiated cells 15 out of 35 (43%) do so. This fact can either reflect a close structural similarity between Rod1 and Ptbp1, with both proteins independently interacting with common factors, or it could mean that Rod1 and Ptbp1 exist together in complexes with these proteins. Altogether, these results indicate that Rod1 interacts with multiple splicing factors in both undifferentiated and differentiated MEL cells. In particular, interactions with Ptbp1 and Raver1 suggest a role for Rod1 in regulation of AS in erythroid cells.



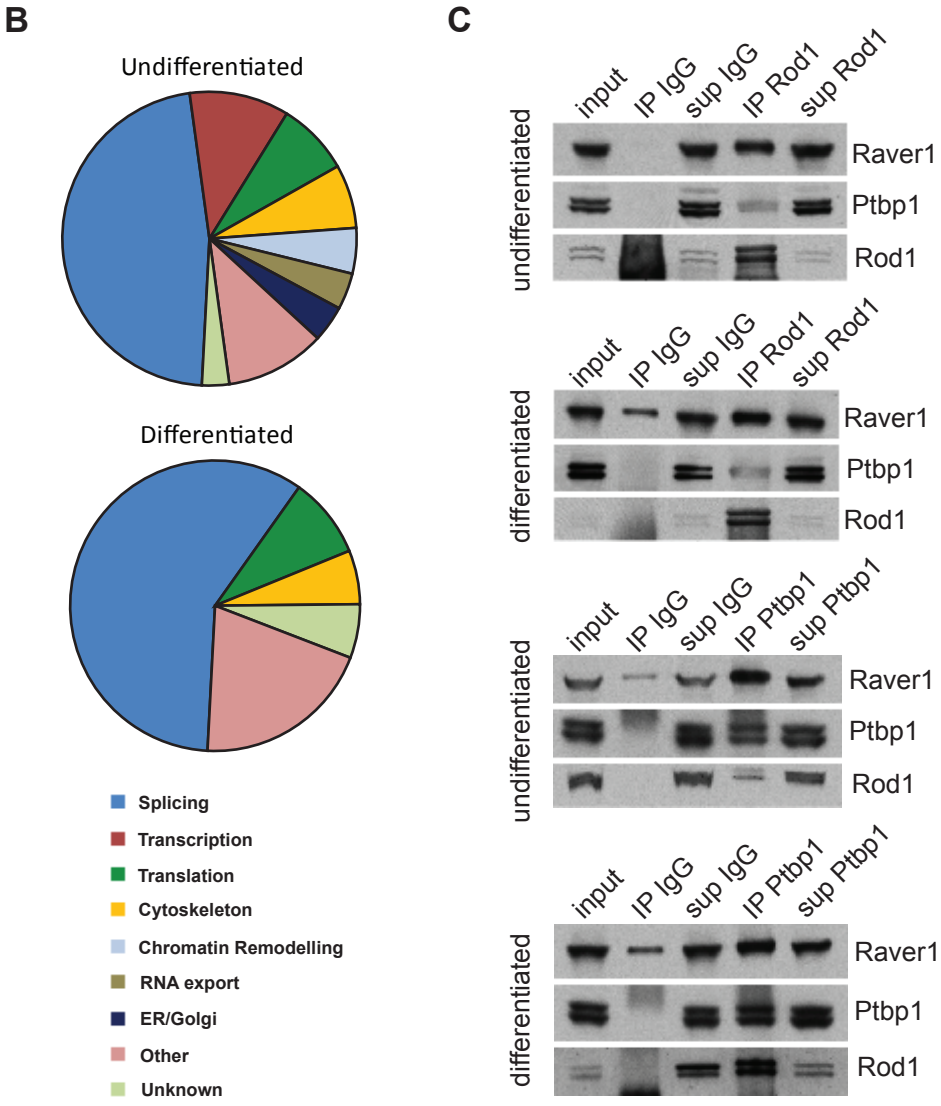


Figure 1 – FLAG-tagged Rod1 interacts with multiple splicing factors in both undifferentiated and differentiated murine erythroleukemia (MEL) cells, among which Ptbp1 and Raver1. **A –** MEL clones stably expressing either N or C-terminal FLAG-tagged Rod1 were created; clones C3 and N2 were used in this study. **B –** Functional analysis of Rod1-interacting factors in undifferentiated and differentiated MEL cells; splicing factors constitute a major fraction of Rod1 interactors. **C –** Western blots confirming mutual interactions among Rod1, Ptbp1 and Raver1 in both differentiated and undifferentiated MEL cells.

Undifferentiated					Differentiated				
Protein	score control	score FLAG-Rod1	score control	score Rod1 - FLAG	Protein	score control	score FLAG-Rod1	score control	score Rod1-FLAG
hnRNP M +	-	2408	-	1340	Raver1 +	-	1926	-	1487
Matr3 *+	-	2394	-	1995	hnRNP M +	-	1777	-	849
Raver1 +	-	2116	-	1655	Rod1 *+	-	1653	-	1577
Ilf3	-	1815	-	197	hnRNP L *+	-	1553	-	-
Rod1 *+	-	1785	-	1505	Matr3 *+	-	1406	-	1143
Ptbp1 *+	-	1766	-	952	hnRNP A1 +	-	1280	-	455
Dhx9 *+	-	1655	-	496	Ptbp1 *+	-	1268	-	656
hnRNP A2b1 *	-	1480	-	-	hnRNP A3 +	50	1231	-	749
hnRNP L +	-	1453	-	1227	Ddx17 *	-	1212	49	456
Tubb2c +	-	1330	-	575	Ddx5	-	1177	-	643
hnRNP UI2	-	1298	-	234	hnRNP R +	-	1167	-	485
hnRNP A1 *+	-	1273	-	639	Rbm14 *+	-	1000	-	574
Sf3b3	-	1259	-	117	hnRNP K	74	979	-	785
hnRNP R *+	-	1212	-	325	Hspa5 (BiP) *+	53	953	-	-
hnRNP A3 +	-	1205	124	830	PCbp2 +	-	810	-	889
Rbm14 *+	-	1181	-	365	Tubb2c *+	-	676	-	-
hnRNP C +	-	1155	137	707	hnRNP AB +	-	669	-	-
Flii	54	1148	-	98	Raly +	-	623	-	345
Myef2	-	1122	-	188	hnRNP H2	-	601	-	220
Elavl1 (HuR) *	-	1034	-	817	Zfr *+	-	547	-	-
Hspa5 (BiP) *+	-	1007	-	671	Dhx9 +	-	496	-	84
PCbp2 +	-	900	-	1027	hnRNP A0 *+	-	477	-	-
hnRNP U	-	874	-	413	hnRNP F *	-	449	-	98
Prpf31 *	46	807	-	-	PCbp1 +	-	434	-	645
Ilf2	-	795	-	232	Tardbp *	-	377	-	335
Sf3b2 *	-	791	-	222	Khdrbs1 *+	-	252	-	-
Zfr +	-	768	-	525	Lmnb1	-	191	-	-
hnRNP AB +	-	731	-	362	Hspd1	-	111	-	433
Raly *+	-	691	-	335	HMGb2 *	-	100	-	-
Golga3 *	-	662	-	-	Magoh *	-	83	-	-
PCbp1 +	70	642	-	431	snRNP200 +	-	59	-	-
Ybx1 *	-	623	-	-	hnRNP C +	-	-	-	441
CUGbp1	-	617	-	126	Ubc	-	-	-	99
hnRNP A0 +	-	558	-	660	Pf4	-	-	-	82
Ruvbl2	-	479	-	153	Mobkl1b +	-	-	-	57
Sf3a3	-	456	-	65					
SnRPa1 *	-	450	-	160					
Rps14	48	443	-	57					
Rbm10 *	-	406	-	-					
Srrt *	-	369	-	-					
Tyms	-	305	-	51					
Rbm7	-	301	-	189					
Sfrs9 (SRp30c)*	-	289	-	-					
Khdrbs1 +	-	272	-	60					
Rpl23 *	-	270	-	-					
Fusip1 (SRp38) *	-	270	-	-					
Sf3a2	-	234	-	-					
Ddx3x *	-	230	-	128					
SnRPa	-	230	-	54					

Y14 *	-	213	-	-
Sfrs7 *	-	207	-	-
Prpf6 *	-	199	-	-
Zcchc8	-	199	-	56
Mbnl3	-	186	-	99
RBbp4	-	182	-	78
Mobkl1b +	-	178	-	179
Thoc6 *	-	172	-	-
U2af65 *	-	159	-	-
Rpl22	-	158	-	-
SnRNP200 *+	-	157	-	-
Rbm17 *	-	152	-	-
Sfrs3 (SRp20)	-	145	-	-
SnRPc	-	135	-	-
SnRNP70 *	-	115	-	-
Rqcd1	-	108	-	-
Rbm39 *	-	106	-	58
Ncbp2 (CBP20)	-	102	-	-
Rps5	-	101	-	51
Prpf19 *	-	92	-	-
Gar1 (NOLA1)	-	88	-	-
Prpf8 *	-	87	-	-
Runx1	-	86	-	86
Eef1d *	-	76	-	-
Thoc1	-	73	-	-
Ddx39 (UAP56)*	-	72	-	-
Cbfb *	-	70	-	172
Cnot1 *	-	67	-	-
Ppih *	-	63	-	-
Myh1	-	-	-	923
Myh4	-	-	-	879
Myh8	-	-	-	634
RpsA	-	-	-	383
GATAd2b	-	-	-	266
hnRPdl	-	-	-	194
Smarcc1	-	-	-	193
Myh13	-	-	-	133
Mrlc2	-	-	-	86
Aldoa	-	-	-	83
GATAd2a	-	-	-	81
Dynll2	-	-	-	76
Ppia	-	-	-	71
Smarce1	-	-	-	60
4930485B16Rik	-	-	-	57
Csnk2a1	-	-	-	57
Smarca4	-	-	-	54
Cit	-	-	-	53
Akap9	-	-	-	52
Cc2d1a	-	-	-	50

Table 1 – Mass Spectrometry data showing Rod1-interacting factors in undifferentiated and differentiated MEL cells. Numbers indicated represent mascot scores; only proteins with mascot scores higher than 50 were considered. Results derive from two MEL clones (N2 and C3) stably expressing an N-terminal (FLAG-Rod1) and C-terminal (Rod1-FLAG) FLAG-fused proteins. Control FLAG pull-downs were performed from *wild-type* MEL cells lacking any FLAG-fused proteins. * indicates proteins present in the Ptbp1 mass spectrometry data (table 2); + indicates proteins present in both the differentiated and undifferentiated sets of data.

Undifferentiated			Differentiated		
Protein	**Score control**	**Score Ptbp1**	**Protein**	**Score control**	**Score Ptbp1**
Ptbp1 *+	179	2192	**Ptbp1 *+**	80	1188
SnRNP200 *	41	1697	**PCBP2**	199	832
Dhx9 *	130	1610	**Hspa5 (BiP) *+**	124	706
Kifc1	-	1375	**Matr3 *+**	-	615
Prpf19 *	219	1229	**Env +**	-	510
Matr3 *+	264	1152	**Myh9**	-	446
Prpf8 *	-	1084	**Rod1 *+**	-	422
Env +	-	1042	**Ddx17 ***	-	402
Hspa5 (BiP) *+	-	1001	**Atp5c1**	-	394
Elavl1 (HuR) *	-	999	**Gapdh**	-	344
Wtap	-	900	**Tardbp ***	-	336
Ptb-af1	-	831	**Tubb2c ***	-	334
eIF4A3	54	822	**Trim21 +**	-	304
Eftud2	-	814	**Tpx2**	-	282
hnRNP L +	-	787	**Nusap1**	-	260
Arid3a	-	754	**Numa1 +**	-	251
Ddx39 (UAP56) *	90	674	**Fkbp11**	-	248
hnRNP A2b1 *	80	535	**Khsrp**	-	240
Rod1 *+	-	525	**hnRNP L *+**	-	221
Thoc2	-	520	**Atp1a1**	-	215
Sfrs1 (ASF)	-	516	**Magmas**	-	208
Rbm15	-	491	**Mosc2**	-	208
Cnot1 *	-	460	**Ttn**	-	208
PPP1CA	-	400	**Snw1 +**	-	205
hnRNP R *	-	397	**Man1a2**	-	185
Thoc6 *	-	393	**Aurkb**	-	183
Ddx3x *	-	379	**Fam162a**	-	181
Rbm14 *+	-	379	**Rbm14 *+**	-	172
Rab11	-	355	**Entpd6**	-	159
Thoc1	-	345	**Sltm**	-	156
Sept7 (cdc10)	-	336	**Slc25a1**	-	155
Rab8b	-	305	**Cisd2 +**	-	155
SnRPa1 *	-	304	**Aifm1**	-	154
Nhp2 (NOLA2)	-	291	**Csnk1d**	-	149
Znf512	-	288	**Brp44**	-	143
Rab5c	-	288	**Mrpl32**	-	138
Sf3b2 *	-	282	**Ndufa4**	-	138
Snw1 +	-	274	**Aldh2**	-	138
Y14 *	-	264	**Slc35b2**	-	135
hnRNP A1 *	-	259	**hnRNP A0 ***	-	133
Trim21 +	-	249	**Zfr ***	-	131
Dido1	-	244	**Magt1**	-	128

Rqcd1	-	243	**hnRNP F ***	-	125
Rpl23 *	-	239	**Actb**	-	124
Thoc5	-	233	**Lman2 +**	-	122
Rab14	-	231	**Csnk1a1**	-	120
Zc3h11a	-	221	**Gosr1**	-	119
Sfrs7 *	-	218	**Atp5h**	-	116
Npm1 +	-	215	**Lmnb1**	-	115
SnRNP70 *	-	207	**Txndc14**	-	113
Sfrs3 (SRp20)	-	206	**Arid3b**	-	113
Prpf6 *	-	201	**Ccdc79**	-	112
Ppih *	-	192	**Preb**	-	110
Rpl22	-	191	**Tmem65**	-	109
Pre-mRNA cleavage factor I	-	190	**Sccpdh**	-	108
Pak6	-	182	**Slc3a2**	-	107
Prpf40a	-	178	**Tomm20**	-	104
SnRP C	-	177	**Cyc1**	-	101
Acin1	-	177	**Rbm33 (Prr8)**	-	98
Sipa1l3	-	171	**Cenpf**	-	97
CUGbp1	-	170	**Canx**	-	97
ERH	-	169	**Slc25a13**	-	96
TAF15	-	167	**Pxmp4**	-	94
Ncbp2 (CBP20)	-	167	**HMGb2 ***	-	94
SnRP G	-	166	**Ssr3**	-	93
Bcas2	-	166	**Glipr2**	-	92
Raly *	-	164	**Atp2a2**	-	91
Fip1l1	-	163	**ORMdl2**	-	89
Eef1d *	-	160	**5033414D02Rik**	-	87
Mpg	-	157	**Cyb5 +**	-	85
Numa1 +	-	155	**Syne2**	-	82
U2AF65 *	-	154	**Npm1 +**	-	82
U2AF35	-	146	**Slc45a4**	-	81
40-2-3	-	144	**Magoh ***	-	80
Prkdc +	-	141	**Tmem55a +**	-	79
Rab6b	-	140	**Clasp1**	-	78
Dnmt1	-	140	**Aspm**	-	78
Cpsf3	-	140	**Dnahc10**	-	78
Fusip1 (SRp38) *	-	137	**Dnajc1**	-	78
Rps3a	-	133	**Prkdc +**	-	78
Rps4x	-	132	**RAP1A**	-	78
Ybx1 *	-	132	**Ccdc101**	-	78
14-3-3	-	130	**Zfp462**	-	77
Sfrs15	-	129	**BC030307**	-	77
Cbfb *	-	129	**Bcl2**	-	76
Dazap1	-	125	**Mgat2**	-	76
Asf1b	-	124	**Fmo9**	-	76

Plrg1	-	122	**Rbm27**	-	76
Cyb5 +	-	121	**Atp5k**	-	75
Srp14	-	119	**Tap2**	-	75
Cisd2 +	-	115	**Bcap31**	-	75
DNMTap1	-	113	**Col5a3**	-	74
Khdrbs1 +	-	113	**Actn4**	-	73
Lman2 +	-	112	**Ckap5**	-	72
Bsn	-	111	**Brd2**	-	72
Sfrs9 (SRp30c) *	-	109	**2900010M23Rik**	-	72
Tmem33	-	108	**Racgap1**	-	72
Ogt	-	107	**Mki67**	-	71
Rbm17 *	-	106	**Nup107**	-	70
Rbm25	-	105	**Casc5**	-	70
Rer1	-	105	**Polr1e**	-	70
Tmem55a +	-	104	**Cpt1a**	-	69
Golga3 *	-	104	**Ckap4**	-	69
Ryr3 +	-	103	**Prkaca**	-	69
Nfkb1	-	102	**Khdrbs1 *+**	-	68
Sp2	-	101	**Rapgef5**	-	68
Rbm39 *	-	99	**Ppp4r4**	-	68
Bax	-	98	**Uaca**	-	67
SMN1	-	97	**Cttnbp2**	-	67
Smarca1 (ISWI)	-	97	**Atp5l**	-	67
Rbm10 *	-	96	**Rims2**	-	67
USP39	-	96	**Dst**	-	67
Pbrm1 (BAF180)	-	96	**Jtb**	-	67
Pex16	-	95	**Ptprd**	-	66
SnRP A	-	91	**Ccdc110**	-	66
Slc25a5	-	91	**Thap11**	-	66
Slc25a31	-	91	**Btaf1 +**	-	65
SnRP B2	-	90	**Tnni3**	-	65
Srrt *	-	89	**Atp8**	-	64
Zfp362	-	89	**Gcdh**	-	63
Yif1b	-	87	**Rufy2**	-	63
Acsl4	-	86	**Rbbp4**	-	63
Scand3	-	84	**Nxf1 (Tap)**	-	63
Cpsf1	-	83	**Ncor1**	-	63
SMNdc1	-	82	**hnRNP AB**	-	62
Tmed10	-	82	**Mtr**	-	62
Btaf1 +	-	82	**Oxr1**	-	62
Scn5a	-	81	**Mela**	-	62
Snx25	-	81	**Med8**	-	62
Sf3a2	-	80	**Eif2c4**	-	62
Zc3h13	-	78	**Wwp2**	-	62
Supt16h	-	78	**LepR**	-	61
Prpf31 *	-	77	**Ryr3 +**	-	61

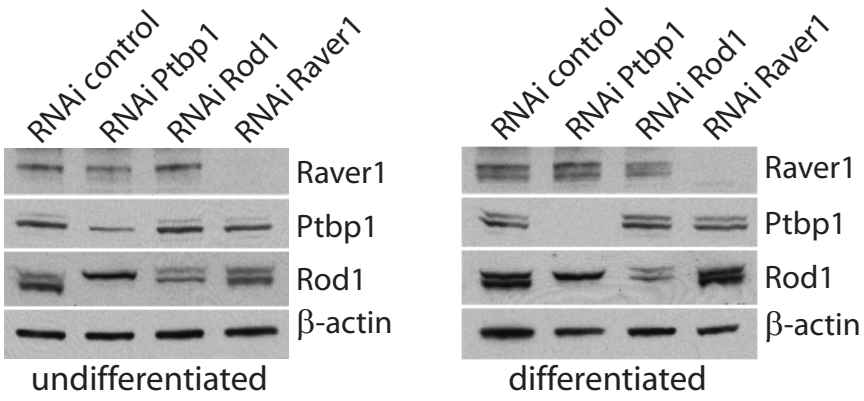
Blm	-	77	**OTTmusG00000015752**	-	60
Ints3	-	74	**Hs2st1**	-	59
Slc2a3	-	74	**H2-eb2**	-	59
Sbf1	-	74	**1190002N15Rik**	-	59
Ndufa8	-	74	**2500003M10Rik**	-	58
Upf2	-	68	**Meox1**	-	58
Erc2	-	68	**Caskin1**	-	58
Aifm2	-	68	**Pnma5**	-	58
Cdc2a	-	66	**Ccdc132**	-	58
Net1	-	65	**Atp7b**	-	57
Galnt10	-	65	**Tep1**	-	57
Cdc23	-	65	**Arap3**	-	56
Gar1 (NOLA1)	-	64	**Galnt11**	-	56
Hdac2	-	64	**Usp51**	-	56
Fh1	-	62	**Tbc1d17**	-	56
Thoc7	-	61	**Exoc2**	-	56
Slc1a1	-	60	**Pank1**	-	56
Cdkl5	-	57	**2310079N02Rik**	-	55
Cdc6	-	56	**Mtap7**	-	55
Rbm5	-	54	**Tmco4**	-	54
Eif2ak4	-	54	**Maea**	-	54
Dip2c	-	53	**Pik3c2g**	-	54
			Sytl2	-	54
			Bcap29	-	53
			Wiz	-	53
			Akap9	-	53
			Ankrd1	-	53
			Ssh1	-	53
			Atp5f1	-	52
			Tcp1	-	52
			Phlpp	-	52
			2010107E04Rik	-	52
			Iqca	-	52
			Reln	-	52
			Ano10	-	52
			Selp	-	51
			Egfl7	-	51
			Mov10l1	-	51
			Gm221	-	51

Table 2 – Mass Spectrometry results showing Ptbp1-interacting factors in undifferentiated and differentiated MEL cells. Numbers indicated represent mascot scores; only proteins with mascot scores higher than 50 were considered. Control experiments were performed with normal IgG. * indicates proteins present in the FLAG-tagged Rod1 Mass Spectrometry results (table 1); $^{+}$ indicates proteins present in both the undifferentiated and differentiated sets of data.

Rod1, Ptbp1 and Raver1 regulate common AS events: The interactions with Ptbp1, Raver1 and several other splicing factors suggest that Rod1 may regulate AS in MEL cells. To verify this, we analysed AS events in both undifferentiated and differentiated MEL cells by exon arrays. We performed RNAi against Rod1, Ptbp1 and Raver1 (fig. 2A) and extracted total RNA, which we hybridized to a mouse exon array covering over one million exons of the mouse genome. Each sample duplicate clustered in a distinct spatial region in a Principal Component Analysis (PCA) map; however, the pattern of clustering indicates that overall differences in exon expression between samples are more prominent in differentiated cells, compared to undifferentiated (supplementary fig. 2A). An FDR (False Discovery Rate) of 0.05 was set as a threshold for AS-regulated transcripts. Strikingly, while in undifferentiated MEL cells very few transcripts were subject to AS regulated by Rod1 (2), Ptbp1 (37) and Raver1 (6), in differentiated cells there were hundreds, many of which were common AS targets (fig. 2B and 2C). Of the 618 transcripts undergoing Rod1-mediated AS events, 528 (85%) were also regulated by Ptbp1, Raver1 or both (fig. 2C). We performed a manual review of "Gene View" profiles for these 528 transcripts and observed that, for the vast majority of them, AS events were agonistically co-regulated (examples in supplementary fig. 2B). In addition, these agonistic AS events show that Rod1, like Ptbp1, is predominantly a splicing repressor. AS events may result in up or down-regulation of the corresponding transcripts, for example through AS-NMD. Interestingly, while Ptbp1 and Raver1-mediated AS events are significantly coupled to mis-regulation of mRNAs, the same is not valid for Rod1: only 21 transcripts (21/618, i.e. 3.4%) were mis-regulated in the set of Rod1 AS-regulated mRNAs, while for Ptbp1 and Raver1 there were 687 (26.9%) and 625 (30.6%), respectively.

Ptbp1 requires its interacting partner Raver1 to repress exon 3 of the α-tropomyosin (Tpm1) gene [16, 17]. Our results demonstrate that a vast number of AS-regulated transcripts are indeed both Ptbp1 and Raver1 targets in differentiated MEL cells (figs. 2B and 2C). In addition, these AS events involve mainly Ptbp1 and Raver1-mediated exon co-repression, as seen in "Gene View" profiles (examples in supplementary fig. 2C), strongly suggesting that the mechanism described for Tpm1 exon 3 repression is a common one. In summary, these results indicate that there is a significant overlap between Rod1, Ptbp1 and Raver1-mediated AS events in differentiated MEL cells. In particular, most Rod1-mediated AS events in differentiated MEL cells are agonistically regulated by Ptbp1, Raver1 or both.

A



B

AS-regulated transcripts	Undifferentiated	Differentiated
Rod1	2	618
Ptbp1	37	2554
Raver1	6	2041
common Rod1-Ptbp1	1	477
common Rod1-Raver1	0	435
common Ptbp1-Raver1	0	1467
common Rod1-Ptbp1-Raver1	0	384

C

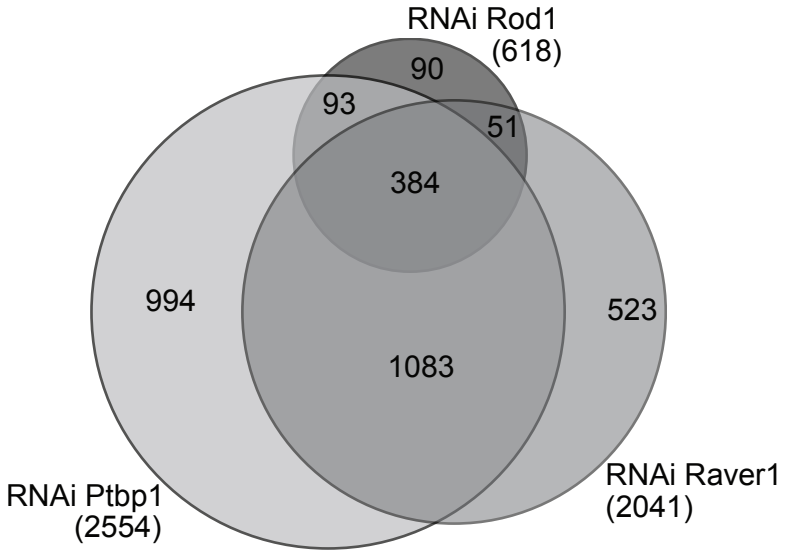
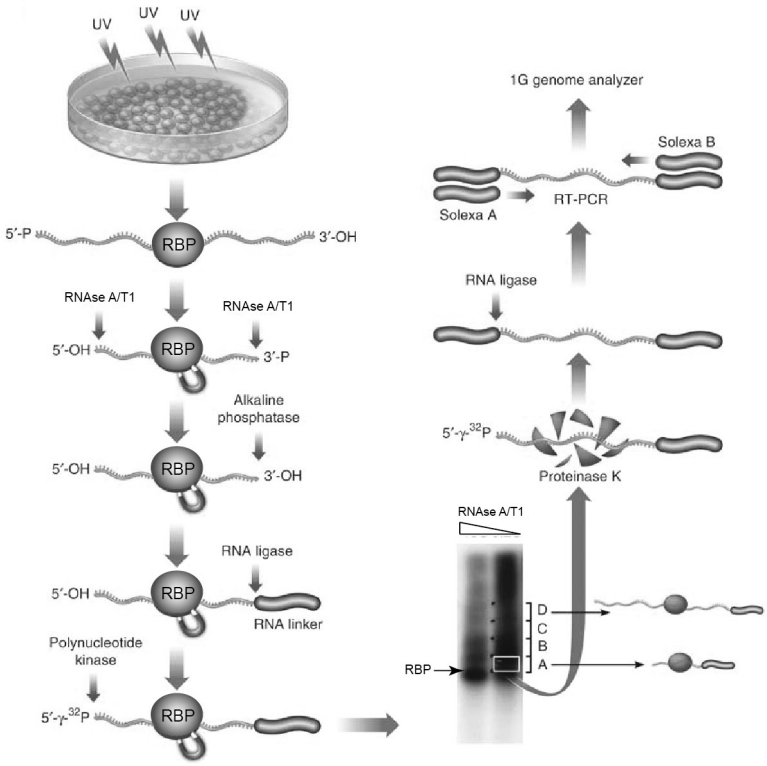


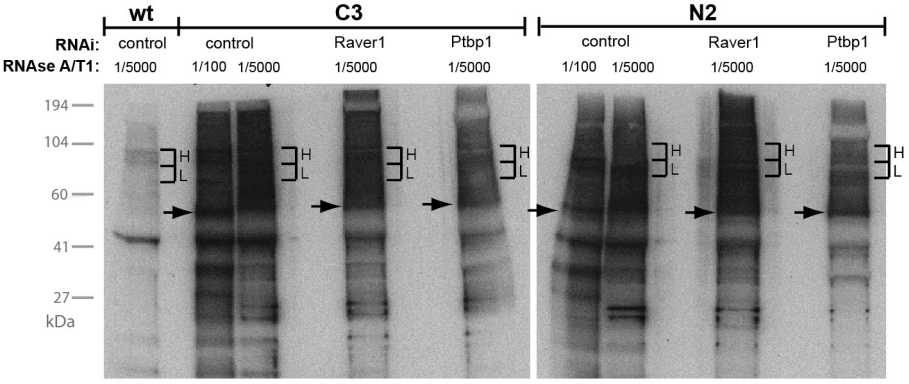
Figure 2 – Rod1, Ptbp1 and Raver1 co-regulate hundreds of transcripts through Alternative Splicing (AS). **A –** Western blots showing RNA interference against Rod1, Ptbp1 and Raver1 in undifferentiated and differentiated murine erythroleukemia (MEL) cells. **B –** Numbers of Rod1, Ptbp1 and Raver1 AS-regulated transcripts in undifferentiated and differentiated MEL cells. **C –** Venn diagram showing overlap among Rod1, Ptbp1 and Raver1 AS targets in differentiated MEL cells. Areas are proportional to numbers of transcripts.

CLIP-seq analysis of Ptbp1 and Rod1-bound transcripts: To gain insights into the mechanism of regulation of AS by Rod1 in differentiated MEL cells, Rod1-RNA interactions were mapped by CLIP-seq (fig. 3A). An anti-FLAG antibody was used to purify Rod1-RNA complexes from total cell extracts from differentiated MEL clones N2 and C3. As a control, *wild-type* induced MEL total extracts lacking any FLAG-fused proteins were used. Protein-RNA complexes in the control experiment were essentially absent, indicating that protein-RNA complexes purified from clones N2 and C3 are specifically due to FLAG-tagged Rod1 (fig. 3B). To determine the set of RNA Rod1-binding sites which is dependent on Ptbp1 or Raver1 a CLIP-seq was performed in parallel with RNA interference. Interestingly, a significant decrease in the amount of Rod1-RNA complexes was observed upon knock-down of Ptbp1, compared to both control and Raver1 RNAi lanes, suggesting that Rod1 depends on Ptbp1 for binding some RNA sites (fig. 3B). We took a similar approach to map Ptbp1-RNA interactions in control, Raver1 and Rod1 knock-down CLIP-seq experiments (fig. 3C). We used an anti-Ptbp1 antibody or control IgG to purify protein-RNA complexes; no visible differences in the amounts of Ptbp1-RNA complexes were detected upon knock-down of Rod1 or Raver1 compared to control (fig. 3C).

A


UV
UV
UV
5′-P
RBP
3′-OH
RNAse A/T1
RNAse A/T1
5′-OH
RBP
3′-P
Alkaline phosphatase
5′-OH
RBP
3′-OH
RNA ligase
5′-OH
RBP
RNA linker
Polynucleotide kinase
5′-γ-32P
RBP
1G genome analyzer
Solexa B
RT-PCR
Solexa A
RNA ligase
5′-γ-32P
Proteinase K
RNAse A/T1
D
C
B
A
RBP

B


wt
C3
N2
RNAi:
control
control
Raver1
Ptbp1
control
Raver1
Ptbp1
RNAse A/T1:
1/5000
1/100
1/5000
1/5000
1/5000
1/100
1/5000
1/5000
1/5000
194
104
60
41
27
kDa
H
L

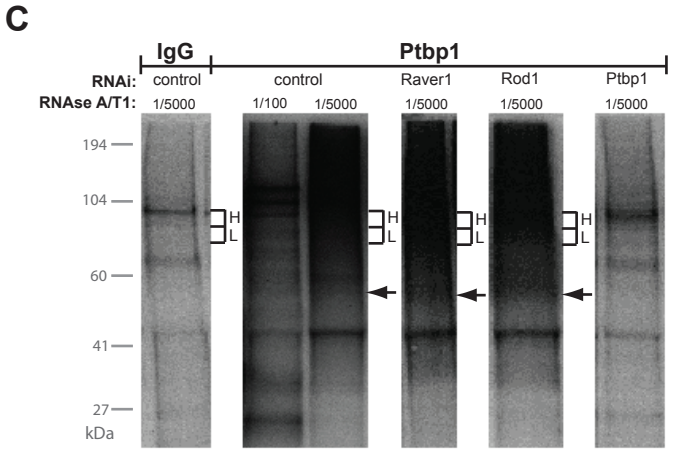


Figure 3 – CLIP-seq (Cross-Linked ImmunoPrecipitation-sequencing) of Rod1 (FLAG-tagged) and Ptbp1-RNA complexes. **A –** the CLIP-seq method: cells are UV-irradiated to cross-link protein-RNA complexes and subject to total lysis; extracts are digested with RNAse A/T1 mix to trim protein-bound RNA molecules and immunoprecipitated with a specific antibody; a DNA-RNA hybrid linker oligonucleotide is then ligated to the 3' end of RNA molecules and subsequently transcripts are phosphorylated with γ-^{32}P by polynucleotide kinase and run in a 10% polyacrylamide gel to separate protein-RNA complexes; these are then transferred to a nitrocellulose membrane and segments of membrane containing appropriate sizes of protein-RNA complexes are cut and digested with proteinase K to separate RNA from protein; RNA is then ligated to the 5' DNA-RNA hybrid oligonucleotide, purified and sent for sequencing; RBP – RNA Binding Protein. **B –** FLAG-tagged Rod1-RNA complexes isolated from differentiated MEL clones N2 and C3 subject to RNAi (control, Ptbp1 and Raver1); two different RNAse A/T1 mix concentrations were used in the control RNAi to confirm that complexes contain RNA (signal is decreased with higher concentration of RNAse A/T1); no significant amount of protein-RNA complexes were isolated with IgG control (left lane, left gel); arrowheads indicate the size of free Rod1 bands; for each lane, two segments of nitrocellulose membrane were cut for sequencing, high (H) and low (L). **C –** Ptbp1-RNA complexes isolated from differentiated MEL cells subject to RNAi (control, Rod1, Raver1); no protein-RNA complexes are visible upon RNAi against Ptbp1, confirming that complexes are specific to Ptbp1; no significant amount of protein-RNA complexes were isolated with IgG control (left lane).

Discussion

The results presented here demonstrate a novel role for Rod1 in the regulation of AS in erythroid cells. Our exon arrays results show that Rod1 itself is regulated by AS in both undifferentiated (Rod1 and Ptbp1-mediated AS) and differentiated (Raver1 and Ptbp1-mediated AS) MEL cells (supplementary fig. 2D). In both cases, the regulatory factors promote skipping of exon 2 of Rod1 (supplementary fig. 2D). Rod1 exon 2 has similar features to exon 11 of Ptbp1 and for both exons skipping leads to the introduction of a PTC and subsequent NMD [19]. Indeed, RNAi against Ptbp1 in both undifferentiated and differentiated MEL cells visibly shifted expression of Rod1 isoforms, favoring the longer isoform which presumably includes exon 2 (fig. 2A). Since the smaller isoform is down-regulated, it is possible that some AS events co-regulated by Rod1 and Ptbp1 are mediated by this Rod1 isoform. An analogous role has been described for Ptbp1 isoforms, which are

reportedly associated with different splicing activities [20]. Repression of exon 2 of Rod1 has been described before in HeLa cells: knock-down of either Ptbp1 or nPTB promoted partial exon inclusion, and inclusion was greatly enhanced with the double knock-down [11]. This auto and cross-regulatory feed-back mechanism by AS coupled to NMD has been reported for several splicing factors [21-29], and keeps protein levels within a narrow range. Such a tight regulation is likely to be physiologically important for Rod1; indeed, Rod1 over-expression increases cell proliferation in 293T cells [30] and inhibits differentiation of both K562 and MEL haematopoietic cell lines ([12] and our data – supplementary fig. 3). Whether this phenotype is mediated through AS or other possible functions of Rod1 is not known. In addition to regulation of Rod1 AS-NMD by Ptbp1, our data show that the reverse is likely valid: Ptbp1 mRNA levels were increased upon knock-down of Rod1 (and Raver1) in differentiated MEL cells (supplementary fig. 4). Similarly, nPTB (Ptbp2) levels were elevated upon knock-down of Ptbp1 in both undifferentiated and differentiated cells (supplementary fig. 4C). Altogether, our data fully support previous evidence for auto and cross-regulation of Ptbp1 and its paralogs Rod1 and nPTB [11].

Individual knock-down of Rod1, Ptbp1 or Raver1 failed to significantly impact AS in undifferentiated MEL cells, despite all proteins being expressed at high levels. In contrast, in differentiated MEL, hundreds of transcripts underwent changes in AS patterns upon RNAi against Rod1, Raver1 or Ptbp1. This discrepancy in AS activity could be related to the different set of interacting partners observed for both Rod1 and Ptbp1 upon erythroid differentiation (tables 1 and 2). It's possible that certain interactions, absent in undifferentiated cells, are required for AS activity. Another explanation, not mutually exclusive, is that the patterns of AS of the Rod1 transcript could change upon differentiation; alternative Rod1 isoforms could have different AS activities, possibly through changes in their interacting partners. While post-translational modifications of Rod1, Ptbp1 and/or Raver1 could provide yet another explanation, none are known so far for any of these factors in mammals.

The Rod1 paralog nPTB (PTBP2) contributes to neuronal and muscle development by repressing an extensive number of tissue-specific exons [6, 31, 32]. In neurons, the ratio of nPTB to PTB expression levels constitutes a molecular switch which regulates neuronal specificity. Repression of PTB leads to increased protein levels of nPTB with the resultant expression of neuronal-specific isoforms [6]. It is plausible that Rod1 plays a similar role in haematopoietic cells. The fact that over-expression of Rod1 blocks differentiation of haematopoietic cell lines supports this hypothesis. In addition, a gene ontology analysis shows that genes involved in haematopoietic development are better represented in the set

of Rod1-regulated AS targets than Ptbp1-regulated ones (supplementary fig. 5).

The regulation of AS by Rod1 in differentiated MEL cells and its homology to Ptbp1 suggest that Rod1 binds RNA, presumably on sites where AS events take place. Our CLIP-seq experiments show that Rod1 is present in protein-RNA complexes, which partially disappear with high concentrations of RNAse A/T1 mix (fig. 3B). Interestingly, RNAi against Ptbp1 caused a reduction in the amount of (visible) Rod1-RNA complexes, indicating that Ptbp1 is required for Rod1 to bind a fraction of these RNA sites. Since Rod1 and Ptbp1 interact, it is plausible that Rod1 is only able to bind some sites when heterodimerised with Ptbp1. In support of this, a *Saccharomyces cerevisiae* Rod1 functional homolog, Nrd1, binds RNA as a heterodimer with Nab3 [33, 34]. Furthermore, Nab3 binds UCUU motifs, while Ptbp1 binds optimally UCUU and, with less affinity, other U and C-rich motifs [35, 36]. Although the quantity of Ptbp1-RNA complexes is not visibly altered after knock-down of Rod1 (fig. 3C), this does not exclude the possibility that a Rod1-Ptbp1 heterodimer is necessary for binding some RNA sites, since only ~19% of alternatively spliced Ptbp1-regulated transcripts are also regulated by Rod1 (fig. 2C); in contrast, ~77% of alternatively spliced Rod1-regulated transcripts are regulated by Ptbp1 (fig. 2C).

Ptbp1, like Rod1, is predominantly a splicing repressor [16, 37-39]. How Ptbp1 mediates exon skipping is not fully understood. A model was proposed in which Ptbp1 competes with U2AF for binding RNA and consequently abrogate the formation of the spliceosomal E complex [37, 38, 40, 41]. However, a simple competition mechanism does not always explain the splicing repression activity of Ptbp1, since for several reported cases multiple Ptbp1-binding sites far from the polypyrimidine tract are required for exon exclusion [13, 35, 42-46]. An alternative model has been proposed in which Ptbp1 binding to these multiple sites loops out the RNA molecule and prevents the full assembly of a functional spliceosome [39]. However, this model implies Ptbp1 dimerization, which has been shown not to be strictly necessary for splicing repression [41]. Furthermore, Ptbp1 is present as a monomer in solution [37, 47]. In this context, the data presented here suggests that AS events co-regulated by Rod1 and Ptbp1 could involve looping of the RNA through RNA-bound Rod1 and Ptbp1 interactions (fig. 4). Rod1 binding to RNA would possibly require pre-binding of Ptbp1 to a nearby site; this model does not imply dimerization of Ptbp1 and explains the significant decrease in Rod1 binding to RNA after Ptbp1 knock-down (fig. 3B). The precise mapping of Rod1 and Ptbp1 RNA binding sites around co-regulated exons would provide important clues as to how exactly these AS events take place.

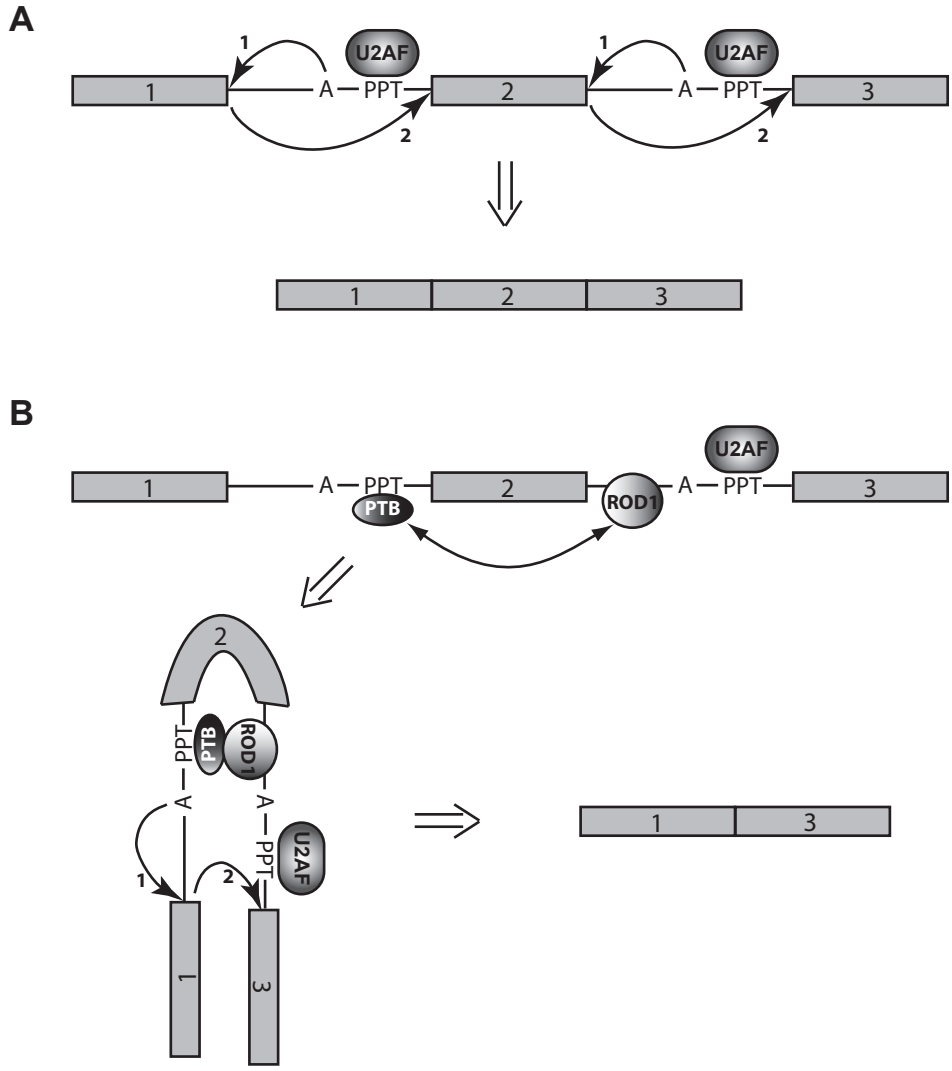


Figure 4 – Proposed mechanism of exon skipping mediated by Rod1 and Ptbp1. **A –** Binding of U2AF to the polypyrimidine tract (PPT) in both introns leads to exon 2 inclusion; the two trans-esterification splicing reactions are indicated by "1" and "2"; "A" indicates the branch point. **B –** Binding of Ptbp1 (PTB) to the PPT and Rod1 to a proximal site in the downstream intron leads to looping out of exon 2 and its subsequent exclusion from the mature RNA.

Materials and Methods

Plasmid constructs and antibodies: pcBA-FLAG-Rod1 was cloned by inserting a PCR-amplified mouse Rod1 cDNA into a plasmid containing two FLAG motifs (2XFLAG) driven by the chicken beta-actin promoter; the resulting construct has an ORF consisting of 2XFLAG followed by 3 glycine residues and the mouse Rod1 cDNA. A similar procedure resulted in the cloning of pcBA-Rod1-FLAG. The following antibodies were used in this study: mouse monoclonal anti-FLAG (Sigma cat. no. F3165), goat polyclonal anti-Rod1 (Santa Cruz cat. no. sc-15251), mouse monoclonal anti-Raver1 (described in [18]) and rabbit polyclonal anti-Ptbp1 (kindly provided by Douglas L. Black).

Cell culture, transfections and RNAi: Murine Erythroleukemia (MEL) cells were grown in DMEM supplemented with 10% fetal calf serum at 37ºC with 5% CO_2; for erythroid differentiation, 2% dimethylsulfoxide (DMSO) was added to cultures during 96-120 hours. Transfections of MEL cells were performed with ~$10X10^6$ cells and ~50 µg plasmid in 500 µl PBS using a Gene Pulser II electroporator (Bio-Rad); cells were left for 5 min. to rest and pipetted into 20 ml DMEM + 10% FCS; serial dilutions (up to 10X) were done and cells were plated into 96-well plates (100 µl each well) and incubated at 37ºC and 5% CO_2; 24 hours after transfection, selection antibiotic was added; clones were harvested and screened for expression by SDS-PAGE. RNAi was performed using lentiviral particles containing shRNA driven by the U6 promotor (pLK0.1-shRNA); lentivirus were produced by transfections in 293T cells according to standard protocols [48]. shRNA target sequences are as follows: Rod1: 5'-GCTGCTGTTACTATGATAA-3'; Ptbp1: 5'-CCAAAGCCTCTTTATTCTCTT-3'; Raver1: 5'-CACAACCCTTACTACCACCAT-3'; control: 5'-ATTCTCCGAACGTGTCACG-3'.

Protein extractions, immunoprecipitations, FLAG pull-downs and mass spectrometry: MEL cells were washed in cold PBS and protein extracts were performed as described in [49], followed by dialysis in buffer C-100 (20 mM HEPES pH 7.6, 0.2 mM EDTA, 1,5 mM $MgCl_2$, 100 mM KCl, 20% glycerol). Extracts were centrifuged at 16000g prior to immunoprecipitations or FLAG pull-downs to remove insoluble precipitates. For immunoprecipitations, Dynabeads coupled with protein G (Invitrogen) were washed several times with PBS and blocked with 200 µg/ml chicken egg albumin for 1 hr at room temperature; 10 µg goat IgG or goat anti-Ptbp1 (N20 clone from Santa Cruz) were then incubated with blocked beads for 30 min. at R.T.; nuclear extracts were incubated at 4°C with benzonase (150 u/ml) and protease inhibitors (Complete, Roche). Extracts were incubated with beads and antibody for 2 hr at 4°C, beads were washed 5X with

buffer C-100 + Complete + 0.02% NP-40 and immunoprecipitated protein complexes eluted in Laemli buffer at 99°C for 5 min. For FLAG pull-downs, 60 µl of anti-FLAG M2 agarose beads (Sigma), equilibrated in buffer C-100, were added to 1 ml of nuclear extract pre-treated with benzonase (150 u/ml), incubated for 2 hr at 4°C, washed 5X with buffer C-100 + Complete + 0.02% NP-40 and eluted four times with C-100 containing 0.2 mg/ml FLAG tripeptide (Sigma) for 15 min at 4°C. Aliquots from eluted fractions were analysed by SDS-PAGE for amounts of FLAG-Rod1 or Rod1-FLAG fused proteins and fractions containing higher amounts were precipitated with TCA-DOC for 30 min. at 4ºC, washed with -20ºC acetone, ressuspended in Laemli buffer and incubated at 99ºC for 5 min. For Mass Spectrometry, samples were loaded into a 10% SDS-PA gel followed by Colloidal Blue Staining (Invitrogen); lanes were cut and submitted to in-gel digestion with trypsin (Promega) as previously described [50]. Nanoflow liquid chromatography-tandem mass spectrometry was performed on a 1100 series capillary liquid chromatography system (Agilent Technologies) coupled to an LTQ-orbitrap mass spectrometer (Thermo) as previously described [51]. Matching peptide fragmentation spectra to databases was performed with Mascot as previously described [51].

Exon arrays: Total RNA extracted with Trizol was purified using the Qiagen RNeasy kit; RNA quality was assessed using the Agilent Bioanalyser. Further processing of samples was performed according to the Affymetrix GeneChip WT Sense Target Labelling Assay. Affymetrix GeneChip Mouse Exon 1.0 ST arrays were used to determine the expression level of exons in the mouse genome. Data was analysed with Partek Genomics Suite 6.4; background correction and normalization of probe set intensities were done using the Robust Multi-array Analysis and GC content was taken into account (GC-RMA) [52]; probe set summarization was performed with median polish settings. Exon-level data was filtered to include only those probe sets that are in the "core" meta-probe list. Within this gene set, the Analysis of Variance (ANOVA) and multi-test correction for P-values were used to identify alternative splicing events. A list of genes with significant AS events was generated using a 0.05 FDR criterion as a significant cutoff. A manual review of gene view plots was performed to identify common AS events shared between Rod1, Ptbp1 and Raver1. Alternatively spliced genes were classified according to molecular function using the Ingenuity Pathway Analysis software.

CLIP-seq: CLIP-seq experiments were performed essentially as described in [53] with the following modifications: RNAse A/T1 mix (Fermentas) was used at a final concentration of 0.5 µg/ml RNAse A and 0.25 U/ml RNAse T1 instead of RNAse A; 5' and 3' linkers were DNA/RNA oligonucleotide hybrids described in [54]; we used anti-FLAG or anti-

Ptbp1 antibodies for immunoprecipitation of protein-RNA complexes and an Illumina GAIIx for sequence analysis.

References

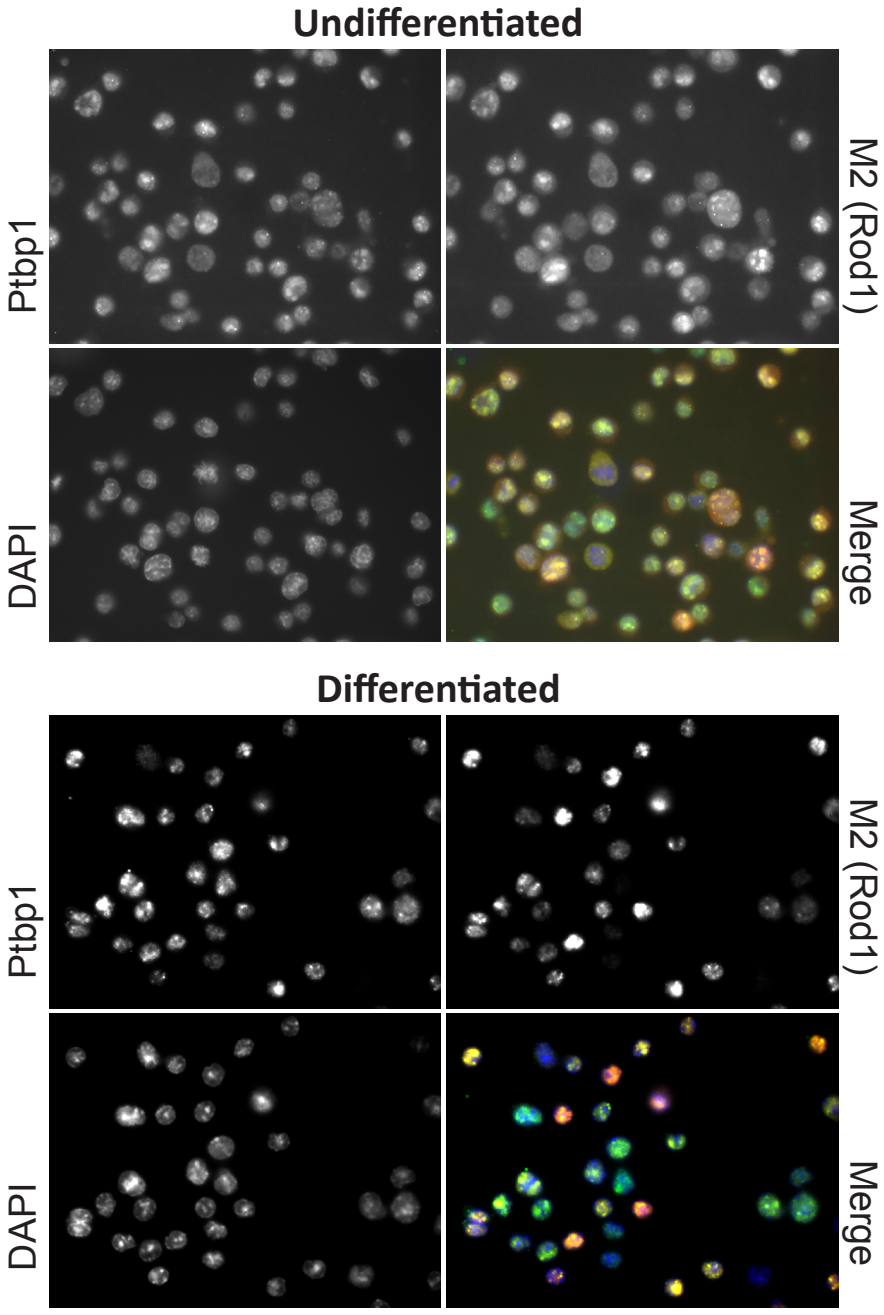
1. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.
2. Green, R.E., et al., *Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes.* Bioinformatics, 2003. **19 Suppl 1**: p. i118-21.
3. Lewis, B.P., R.E. Green, and S.E. Brenner, *Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans.* Proc Natl Acad Sci U S A, 2003. **100**(1): p. 189-92.
4. Jin, Y., et al., *A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG.* EMBO J, 2003. **22**(4): p. 905-12.
5. Warzecha, C.C., et al., *ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing.* Mol Cell, 2009. **33**(5): p. 591-601.
6. Boutz, P.L., et al., *A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons.* Genes Dev, 2007. **21**(13): p. 1636-52.
7. Ule, J., et al., *Nova regulates brain-specific splicing to shape the synapse.* Nat Genet, 2005. **37**(8): p. 844-52.
8. Zhu, H., et al., *A nuclear function of Hu proteins as neuron-specific alternative RNA processing regulators.* Mol Biol Cell, 2006. **17**(12): p. 5105-14.
9. Underwood, J.G., et al., *Homologues of the Caenorhabditis elegans Fox-1 protein are neuronal splicing regulators in mammals.* Mol Cell Biol, 2005. **25**(22): p. 10005-16.
10. Kalsotra, A., et al., *A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20333-8.
11. Spellman, R., M. Llorian, and C.W. Smith, *Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1.* Mol Cell, 2007. **27**(3): p. 420-34.
12. Yamamoto, H., et al., *Isolation of a mammalian homologue of a fission yeast differentiation regulator.* Mol Cell Biol, 1999. **19**(5): p. 3829-41.
13. Gooding, C., G.C. Roberts, and C.W. Smith, *Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon.* RNA, 1998. **4**(1): p. 85-100.
14. Raponi, M., et al., *Polypyrimidine tract binding protein regulates alternative splicing of an aberrant pseudoexon in NF1.* FEBS J, 2008. **275**(24): p. 6101-8.
15. Xue, Y., et al., *Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping.* Mol Cell, 2009. **36**(6): p. 996-1006.
16. Gromak, N., et al., *The PTB interacting protein raver1 regulates alpha-tropomyosin alternative splicing.* EMBO J, 2003. **22**(23): p. 6356-64.
17. Rideau, A.P., et al., *A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain.* Nat Struct Mol Biol, 2006. **13**(9): p. 839-

48.
18. Huttelmaier, S., et al., *Raver1, a dual compartment protein, is a ligand for PTB/hnRNPI and microfilament attachment proteins.* J Cell Biol, 2001. **155**(5): p. 775-86.
19. Wollerton, M.C., et al., *Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay.* Mol Cell, 2004. **13**(1): p. 91-100.
20. Wollerton, M.C., et al., *Differential alternative splicing activity of isoforms of polypyrimidine tract binding protein (PTB).* RNA, 2001. **7**(6): p. 819-32.
21. Dredge, B.K., et al., *Nova autoregulation reveals dual functions in neuronal splicing.* EMBO J, 2005. **24**(8): p. 1608-20.
22. Jumaa, H. and P.J. Nielsen, *The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation.* EMBO J, 1997. **16**(16): p. 5077-85.
23. Sureau, A., et al., *SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs.* EMBO J, 2001. **20**(7): p. 1785-96.
24. Chabot, B., et al., *An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1.* Mol Cell Biol, 1997. **17**(4): p. 1776-86.
25. Kumar, S. and A.J. Lopez, *Negative feedback regulation among SR splicing factors encoded by Rbp1 and Rbp1-like in Drosophila.* EMBO J, 2005. **24**(14): p. 2646-55.
26. Lareau, L.F., et al., *Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements.* Nature, 2007. **446**(7138): p. 926-9.
27. Ni, J.Z., et al., *Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay.* Genes Dev, 2007. **21**(6): p. 708-18.
28. Stoilov, P., et al., *Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA.* Hum Mol Genet, 2004. **13**(5): p. 509-24.
29. Saltzman, A.L., et al., *Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay.* Mol Cell Biol, 2008. **28**(13): p. 4320-30.
30. Sadvakassova, G., et al., *Regulator of differentiation 1 (ROD1) binds to the amphipathic C-terminal peptide of thrombospondin-4 and is involved in its mitogenic activity.* J Cell Physiol, 2009. **220**(3): p. 672-9.
31. Makeyev, E.V., et al., *The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing.* Mol Cell, 2007. **27**(3): p. 435-48.
32. Boutz, P.L., et al., *MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development.* Genes Dev, 2007. **21**(1): p. 71-84.
33. Arigo, J.T., et al., *Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3.* Mol Cell, 2006. **23**(6): p. 841-51.
34. Thiebaut, M., et al., *Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance.*

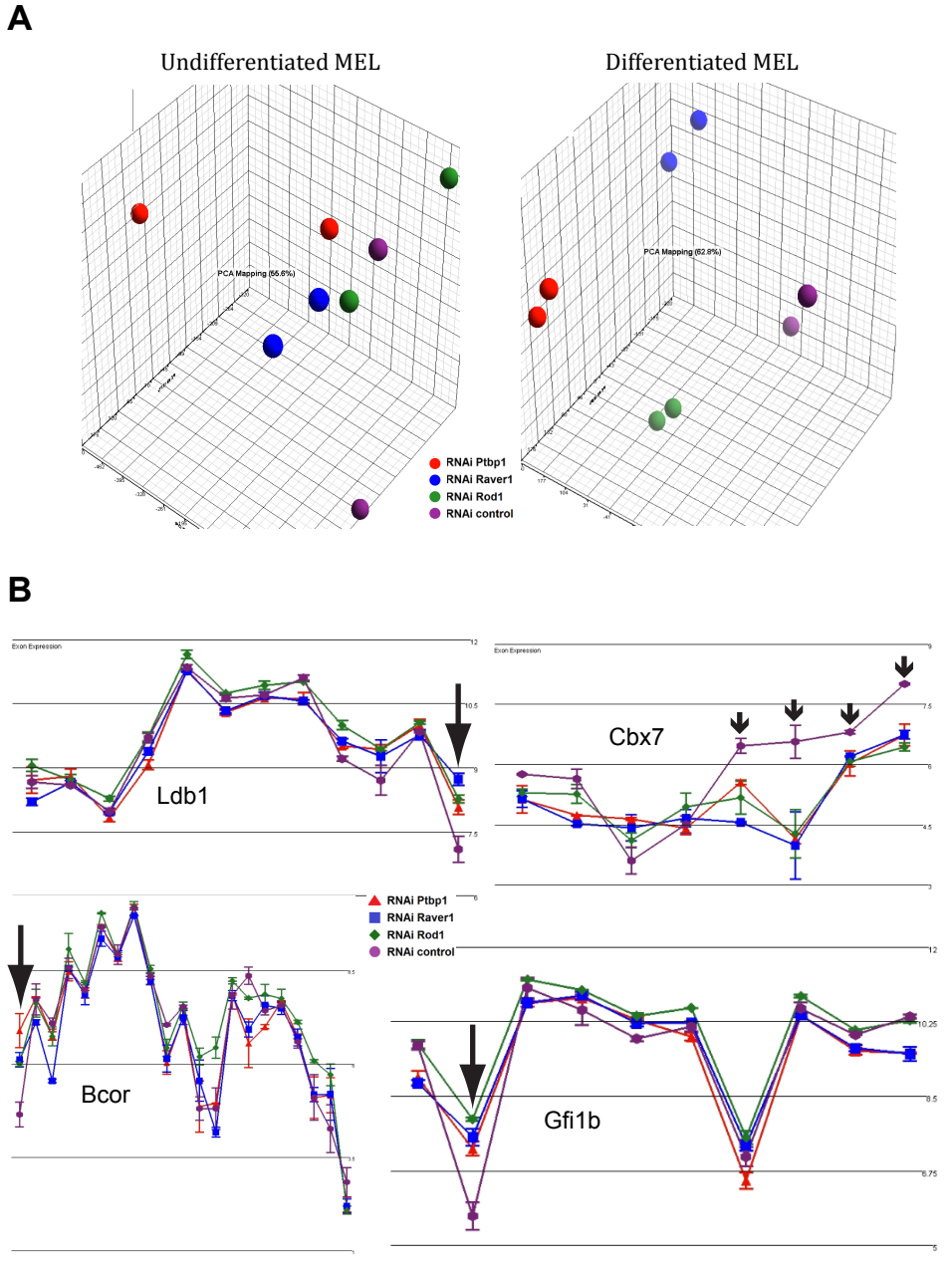
Mol Cell, 2006. **23**(6): p. 853-64.
35. Perez, I., et al., *Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo.* RNA, 1997. **3**(7): p. 764-78.
36. Singh, R., J. Valcarcel, and M.R. Green, *Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins.* Science, 1995. **268**(5214): p. 1173-6.
37. Amir-Ahmady, B., et al., *Exon repression by polypyrimidine tract binding protein.* RNA, 2005. **11**(5): p. 699-716.
38. Sharma, S., A.M. Falick, and D.L. Black, *Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex.* Mol Cell, 2005. **19**(4): p. 485-96.
39. Wagner, E.J. and M.A. Garcia-Blanco, *Polypyrimidine tract binding protein antagonizes exon definition.* Mol Cell Biol, 2001. **21**(10): p. 3281-8.
40. Liu, H., et al., *Mutations in RRM4 uncouple the splicing repression and RNA-binding activities of polypyrimidine tract binding protein.* RNA, 2002. **8**(2): p. 137-49.
41. Izquierdo, J.M., et al., *Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition.* Mol Cell, 2005. **19**(4): p. 475-84.
42. Chan, R.C. and D.L. Black, *Conserved intron elements repress splicing of a neuron-specific c-src exon in vitro.* Mol Cell Biol, 1995. **15**(11): p. 6377-85.
43. Gooding, C., et al., *Smooth muscle-specific switching of alpha-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon.* EMBO J, 1994. **13**(16): p. 3861-72.
44. Southby, J., C. Gooding, and C.W. Smith, *Polypyrimidine tract binding protein functions as a repressor to regulate alternative splicing of alpha-actinin mutally exclusive exons.* Mol Cell Biol, 1999. **19**(4): p. 2699-711.
45. Chan, R.C. and D.L. Black, *The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream.* Mol Cell Biol, 1997. **17**(8): p. 4667-76.
46. Chou, M.Y., et al., *Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing.* Mol Cell, 2000. **5**(6): p. 949-57.
47. Simpson, P.J., et al., *Structure and RNA interactions of the N-terminal RRM domains of PTB.* Structure, 2004. **12**(9): p. 1631-43.
48. Zufferey, R., et al., *Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo.* Nat Biotechnol, 1997. **15**(9): p. 871-5.
49. Dignam, J.D., R.M. Lebovitz, and R.G. Roeder, *Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei.* Nucleic Acids Res, 1983. **11**(5): p. 1475-89.
50. Wilm, M., et al., *Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry.* Nature, 1996. **379**(6564): p. 466-9.
51. Sanchez, C., et al., *Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor.* Mol Cell Proteomics, 2007. **6**(5): p. 820-34.

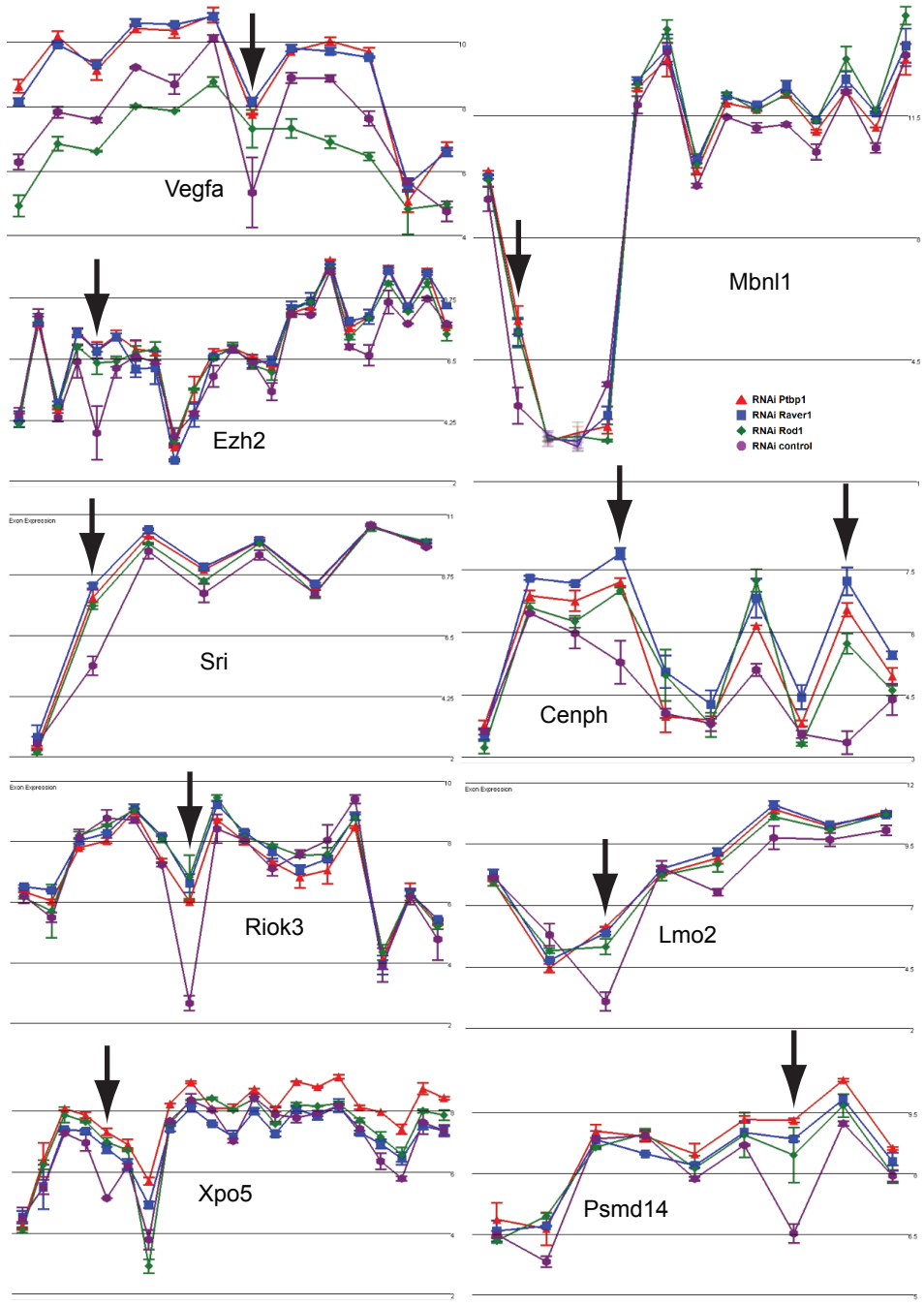
52. Irizarry, R.A., et al., *Exploration, normalization, and summaries of high density oligonucleotide array probe level data.* Biostatistics, 2003. **4**(2): p. 249-64.
53. Ule, J., et al., *CLIP: a method for identifying protein-RNA interaction sites in living cells.* Methods, 2005. **37**(4): p. 376-86.
54. Mamanova, L., et al., *FRT-seq: amplification-free, strand-specific transcriptome sequencing.* Nat Methods. **7**(2): p. 130-2.

Supplementary Data

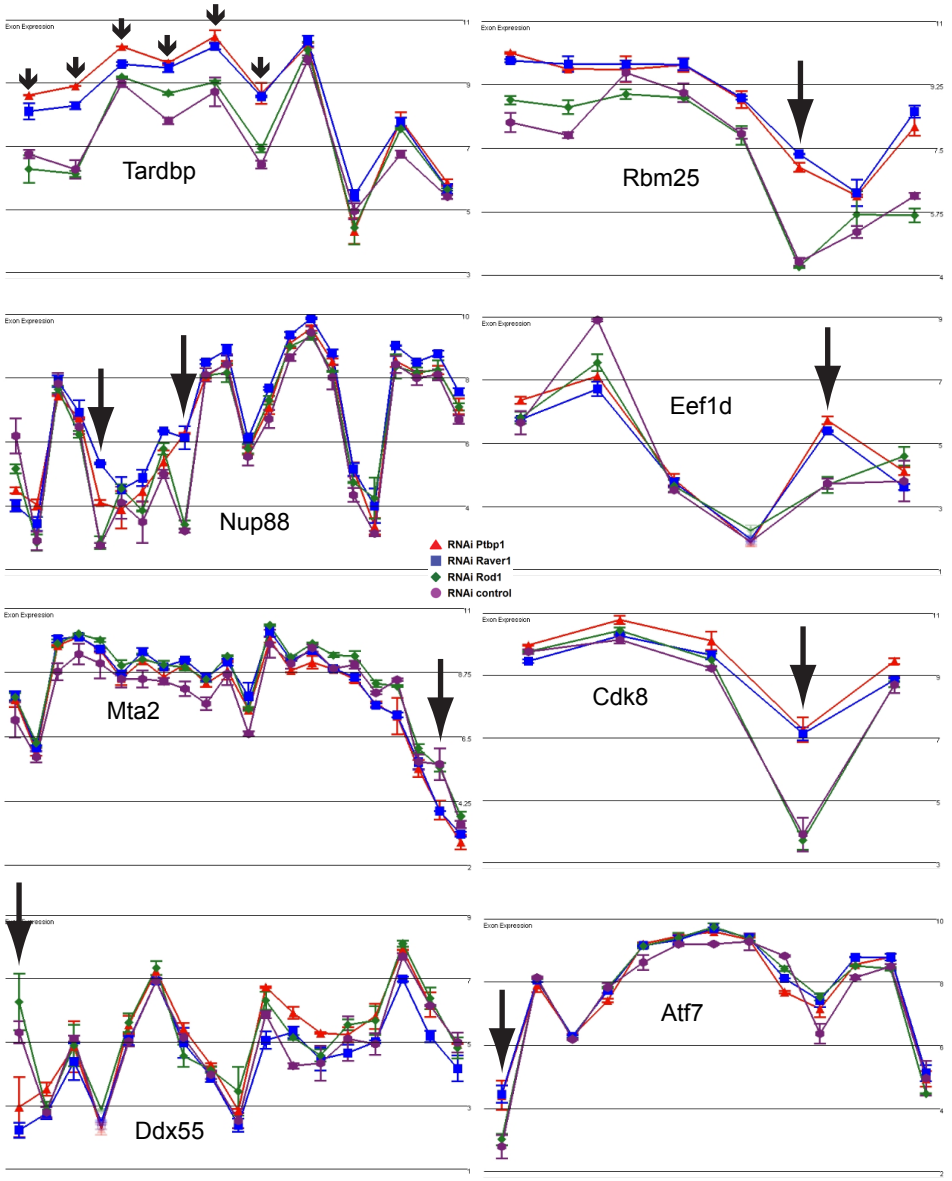


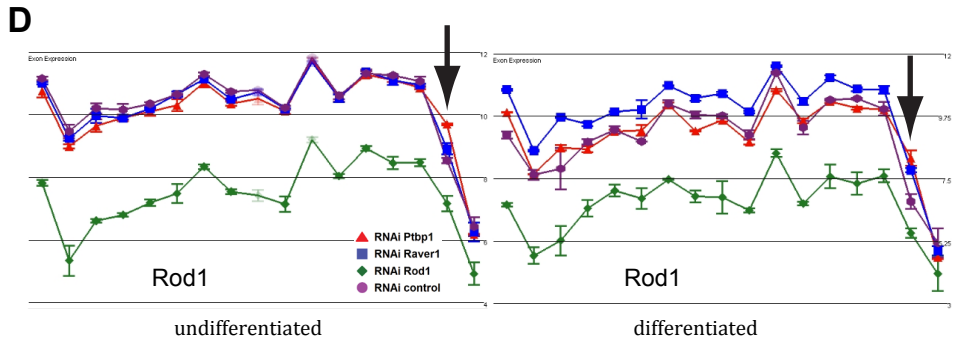
Supplementary figure 1 – Immunocytochemistry showing Rod1-Ptbp1 co-localization in undifferentiated (upper set of panels) and differentiated (lower set of panels) murine erythroleukemia (MEL) cells.



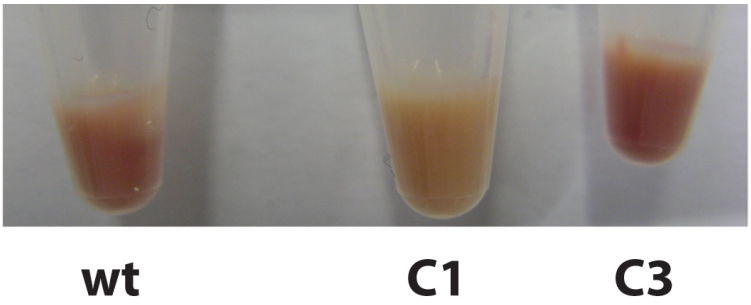

Vegfa
Mbnl1
RNAi Ptbp1
RNAi Raver1
RNAi Rod1
RNAi control
Ezh2
Sri
Cenph
Riok3
Lmo2
Xpo5
Psmd14

C

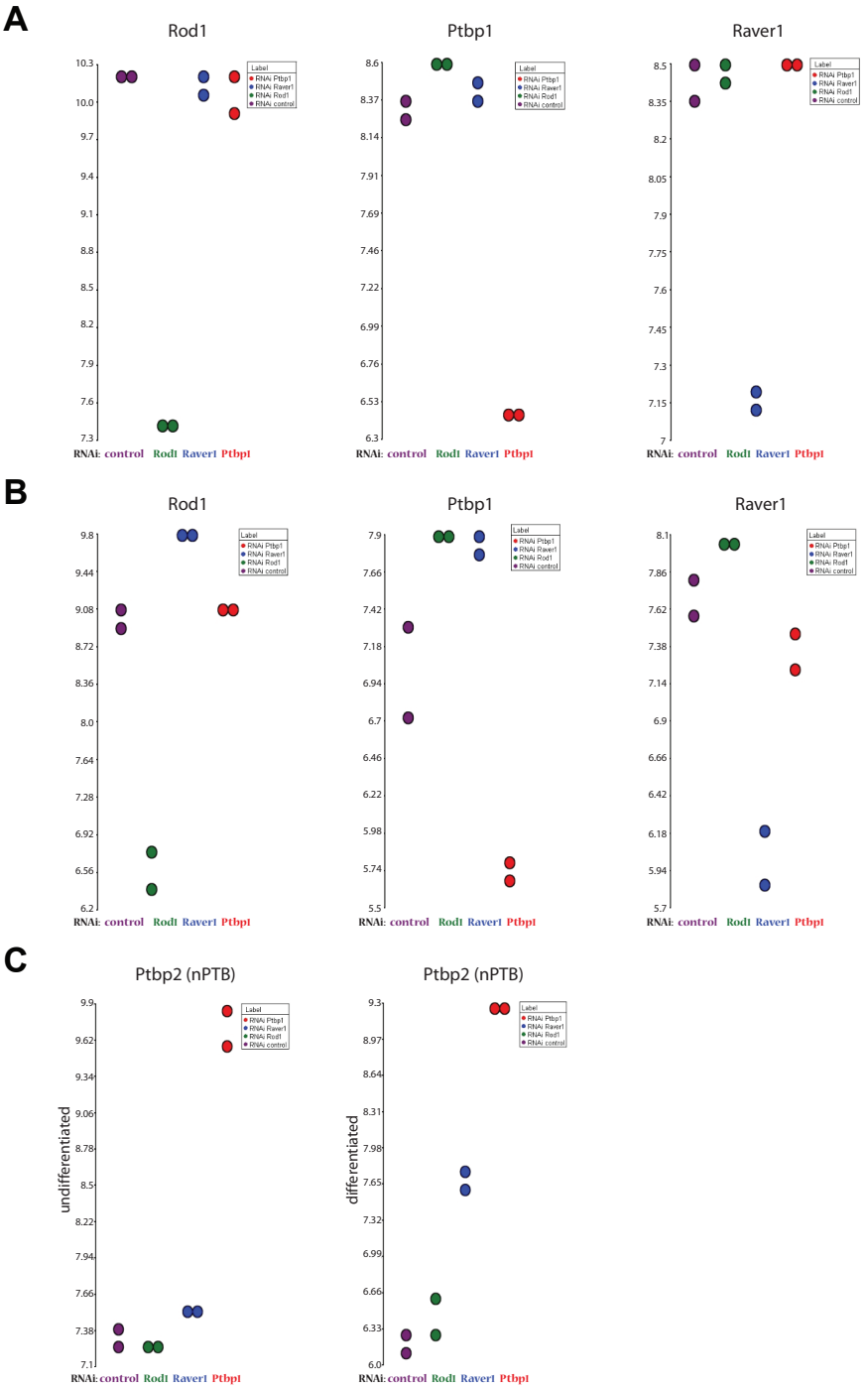

Exon Expression
Tardbp
Rbm25
Nup88
Eef1d
RNAi Ptbp1
RNAi Raver1
RNAi Rod1
RNAi control
Mta2
Cdk8
Ddx55
Atf7



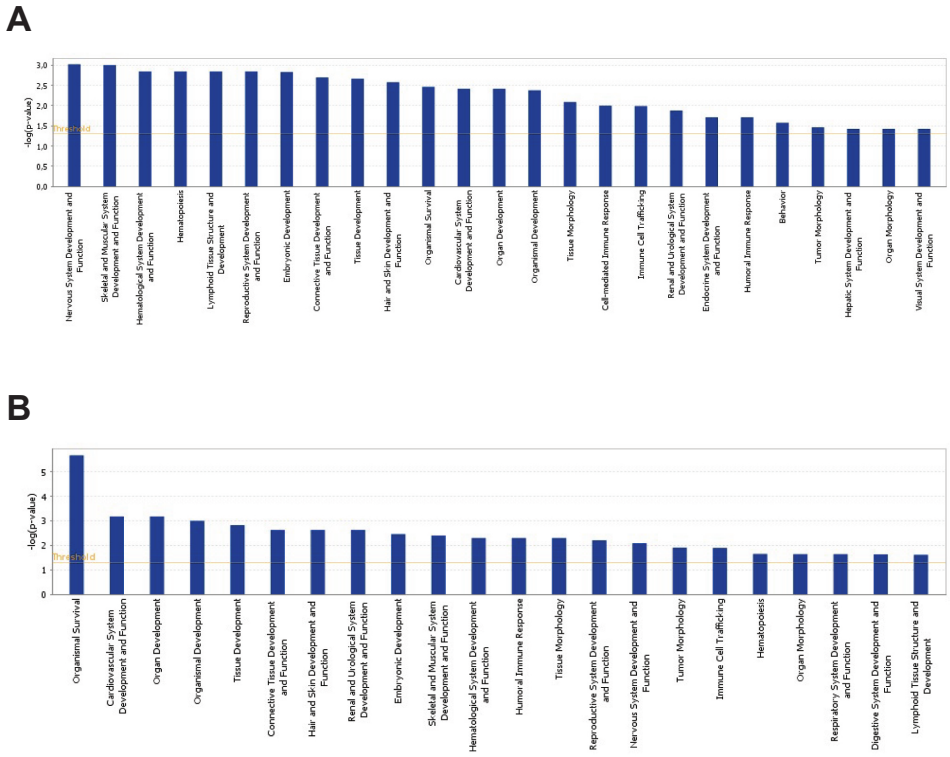
Supplementary Figure 2 – A – Principal Component Analysis (PCA) mapping of Exon Array sample duplicates in undifferentiated and differentiated MEL cells. **B –** Gene Views showing AS events co-regulated by Rod1, Ptbp1 and Raver1; AS events are indicated with arrowheads. **C –** Gene Views showing AS events co-regulated by Ptbp1 and Raver1. **D –** Gene Views of Rod1 showing AS of exon 2 (arrowheads); both Rod1 and Ptbp1 promote skipping of exon 2 in undifferentiated MEL cells, while Ptbp1 and Raver1 promote skipping of exon 2 in differentiated MEL cells. The horizontal axis represents the probe position along the gene and the vertical axis indicates the relative hybridization signal (log2 scale).



Supplementary Figure 3 – Over-expression of Rod1 in MEL clone C1 inhibits DMSO-induced erythroid differentiation. *Wild-type* MEL cells, as well as cells from MEL clones C1 and C3 over-expressing FLAG-tagged Rod1 were induced to erythroid differentiation with DMSO; cells from the MEL clone C1 fail to differentiate, as judged by haemoglobin (red colour) production.


A
Rod1
Ptbp1
Raver1
RNAi: control Rodl Raver1 Ptbp1
B
Rod1
Ptbp1
Raver1
RNAi: control Rodl Raver1 Ptbp1
C
Ptbp2 (nPTB)
Ptbp2 (nPTB)
undifferentiated
differentiated
RNAi: control Rodl Raver1 Ptbp1

Supplementary figure 4 – Dot plots of Rod1, Ptbp1, Raver1 and nPTB (Ptbp2). **A –** Transcript levels of Rod1, Ptbp1 and Raver1 in samples of undifferentiated MEL cells. **B –** Transcript levels of Rod1, Ptbp1 and Raver1 in samples of differentiated MEL cells. **C –** Transcript levels of nPTB (Ptbp2) in samples of undifferentiated and differentiated MEL cells.



Supplementary figure 5 – Gene Ontology Analysis of Rod1 (**A**) and Ptbp1 (**B**) AS-regulated target genes.

Chapter 4

Discussion

Discussion

The results presented in chapters 2 and 3 demonstrate the role of Rod1 in two distinct post-transcriptional regulation mechanisms: Nonsense-Mediated mRNA Decay (NMD) and Alternative Splicing (AS). This multi-functionality is common among splicing factors: for example, the Rod1 paralog Ptbp1 is involved not only in splicing [1], but also in poly-A adenylation [2, 3], RNA export [4] and IRES-mediated translation initiation [5-8]. In fact, many Ptbp1 interacting proteins (chapter 3, table 2) support the role of Ptbp1 in these processes: UAP56 and the THOC subunits Thoc1, Thoc2, Thoc5, Thoc6 and Thoc7 are part of the TREX complex involved in mRNA export [9-11], PCBP2, Ybx1 and HuR participate in IRES-mediated translation [12-15] and both Cpsf1 and Fip1l1 are involved in poly-A adenylation [16, 17]. Other multi-functional splicing factors include several Exon Junction Complex (EJC) proteins, which mediate mRNA export, polysome association and NMD [18-23]. From an evolutionary perspective, it is tempting to speculate that RNA-binding proteins which were initially regulating one mechanism, such as splicing, eventually evolved to regulate other aspects of RNA metabolism. The potential to evolve in such manner may be due to the ability of a protein to bind RNA and remain bound throughout part of the transcript's history. The two different mechanisms of gene expression (NMD and AS) regulated by Rod1 were described in two different cell types, which derive from different species and different developmental stages. Whether Rod1 regulates AS in HEK293 cells and/or NMD in MEL cells remains to be seen; several splicing factors interact with Rod1 in HEK293 cells (ex. hnRNP A1, U2AF, U11/U12 snRNP, U4/U6 snRNP), suggesting that Rod1 could indeed participate in splicing or AS regulation. In MEL cells it seems unlikely that Rod1 participates in NMD; both magoh and UAP56 were present in mass spectrometry data, but interactions with Upf factors or other EJC proteins were absent.

A common feature in chapters 2 and 3 is that Rod1 transcripts are regulated by NMD (AS-NMD in chapter 3, regulated by Rod1 itself and/or Ptbp1 and/or Raver1; in chapter 2, knock-down of Upf1 results in slightly increased Rod1 protein levels, indicating that NMD targets at least one Rod1 isoform – fig. 2B). This tight auto-regulation of Rod1 is evident in both over-expression and RNAi experiments, both of which being difficult to achieve. In fact, due to the relatively inefficient knock-down of Rod1, the number of Rod1 targets, through both NMD and AS, is likely to be higher than presented in chapters 2 and 3, respectively. On the other hand, the observation that a ~50% decrease in Rod1 protein levels is sufficient to significantly affect NMD (chapter 2) and AS (chapter 3) indicates that small differences in Rod1 expression lead to physiologically big effects. It is thus not

surprising that levels of Rod1 protein are kept in check *in vivo*. In addition to the phenotype resulting from over-expression of Rod1 in haematopoietic cells (chapter 3), which inhibits differentiation, over-expression of Rod1 in both skin fibroblasts and 293T cells results in increased basal cell proliferation [24]. Over-expression of Rod1 has also been associated with leukemia [25] and described as a poor prognosis in breast and prostate cancers [26, 27], while treatments which inhibit tumor growth are reportedly associated with down-regulation of Rod1 [28]. To complement these studies, it would be interesting to study the effect of absence of Rod1 in the context of the whole organism. In this regard, one could anticipate a role for Rod1 in embryonic development; Rod1 is expressed in several embryonic tissues, as opposed to adult organisms, where it is expressed in haematopoietic tissues only [29]. Furthermore, Rod1 regulates NMD in embryonic kidney (HEK293) cells and both the NMD factors Upf1 and Upf2 have important roles in development [30-32].
Several similarities can be found between Rod1 and its yeast homolog Nrd1: both have ~57 kDa and contain four RRM domains; in *Saccharomyces cerevisiae*, Nrd1 regulates mRNA stability [33], while in *Saccharomyces pombe* Nrd1 regulates differentiation, specifically in response to metabolic changes [34]. Interestingly, Rod1 is up-regulated upon inhibition of mTOR, a nutrient-responsive protein kinase, suggesting that this mechanism may be at least partially conserved [35, 36]. Finally, there is evidence that Rod1, like Nrd1, may bind the C-terminal tail of RNA polymerase II [24]. A variety of splicing factors are known to bind the C-terminus of RNA pol II in response to its phosphorylations/dephosphorylations during transcription [37, 38]. This ensures the timely processing of the nascent RNA, including capping, splicing, cleavage and adenylation.

Most of the Rod1 and Ptbp1-interacting splicing factors in differentiated MEL cells described in chapter 3 typically promote exon skipping. This indicates that both Rod1 and Ptbp1 mediate their predominantly repressive splicing activity by interacting with other repressors. There are however descriptions of Ptbp1 as a splicing activator [39-41], and our data show that also Rod1 promotes splicing of some exons. Several studies of AS regulation by hnRNPs underscore the role of the positioning of RNA binding sites relative to the regulated exon in its inclusion or exclusion [42-44]: in general, binding of hnRNPs within the exon or in the proximal flanking intronic regions leads to exon skipping, while binding in the intron preceding the regulated exon results in exon inclusion [45]. This position effect has also been observed for Ptbp1 [40] and Fox2 [46], and it probably also applies to Rod1. The CLIP-seq mapping of Rod1 and Ptbp1 RNA binding sites should help to clarify this effect. An interesting common feature between Rod1 and Ptbp1 is that both proteins interact with U2AF in undifferentiated MEL cells (Rod1 interacts with U2AF65,

while Ptbp1 interacts with both U2AF65 and U2AF35) and these interactions are lost upon differentiation. In contrast, some interactions are gained upon differentiation, namely with some hnRNPs. These alterations in Rod1 and Ptbp1-binding partners upon erythroid differentiation may contribute to binding at different RNA sites with the consequent repression of several exons.

In summary, the results presented in chapters 2 and 3 establish a role for Rod1 in two distinct but interconnected post-transcriptional mechanisms of gene regulation: NMD and AS. These are integrated in a more complex system of regulatory mechanisms such as transcription, mRNA localization, translation, post-translational modifications, etc (fig. 1). Some studies support the view that certain cellular events are more dependent on a particular step of gene expression than others. For example, translational regulation is particularly important under conditions which require rapid and precise changes in protein levels, such as stress, apoptosis, cell division or synaptic activity [47-49]. Similarly, NMD is particularly important in the process of immunoglobulin gene recombination in B cells, during which a vast amount of PTC-containing transcripts is generated [32, 50].

In the future, it would be interesting to investigate the role of Rod1 in development. A knock-out mouse could provide useful information into the role of Rod1 in the embryonic stages of development; it could also address the importance of Rod1 in adult tissues, namely haematopoietic cells (for this purpose, a conditional knock-out would potentially be required). As already mentioned, further experiments are necessary to understand the precise molecular function of Rod1 in NMD. Future experiments should also address the role of cytoskeletal Rod1-interacting proteins in Rod1 cellular localization and function. Regarding the participation of Rod1 in AS, one could study the role of its interactions with other splicing factors, particularly Ptbp1 and Raver1, in exon inclusion or exclusion. Some possible experiments include creating interaction-deficient mutants of Rod1 and verifying if these can rescue the AS alterations induced by Rod1 knock-down.

Finally, the knowledge of the molecular functions of Rod1 and the mechanisms in which it participates could in the future be used for designing drugs in order to modulate AS or NMD of disease-causing genes.

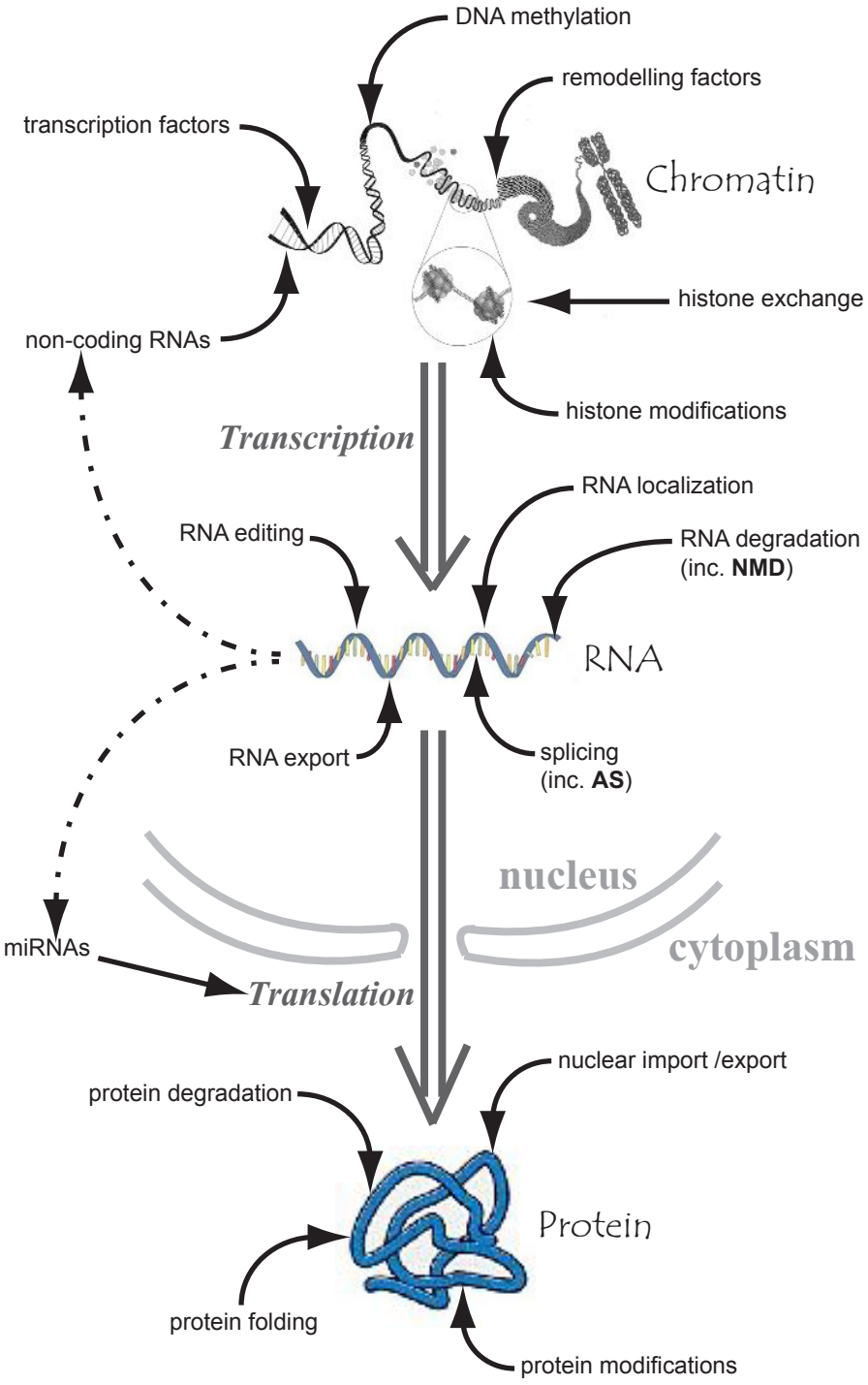

DNA methylation
remodelling factors
transcription factors
Chromatin
histone exchange
non-coding RNAs
histone modifications
Transcription
RNA localization
RNA editing
RNA degradation
(inc. NMD)
RNA
RNA export
splicing
(inc. AS)
nucleus
cytoplasm
miRNAs
Translation
nuclear import /export
protein degradation
Protein
protein folding
protein modifications

Figure 1 – Nonsense-Mediated mRNA Decay (NMD) and Alternative Splicing (AS) are integrated in a complex network of regulatory steps, which together mediate protein production from a gene.

References

1. Garcia-Blanco, M.A., S.F. Jamison, and P.A. Sharp, *Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns.* Genes Dev, 1989. **3**(12A): p. 1874-86.
2. Castelo-Branco, P., et al., *Polypyrimidine tract binding protein modulates efficiency of polyadenylation.* Mol Cell Biol, 2004. **24**(10): p. 4174-83.
3. Lou, H., et al., *Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3'-terminal exon.* Mol Cell Biol, 1999. **19**(1): p. 78-85.
4. Zang, W.Q., et al., *Role of polypyrimidine tract binding protein in the function of the hepatitis B virus posttranscriptional regulatory element.* J Virol, 2001. **75**(22): p. 10779-86.
5. Jang, S.K. and E. Wimmer, *Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein.* Genes Dev, 1990. **4**(9): p. 1560-72.
6. Hellen, C.U. and E. Wimmer, *Translation of encephalomyocarditis virus RNA by internal ribosomal entry.* Curr Top Microbiol Immunol, 1995. **203**: p. 31-63.
7. Vagner, S., et al., *Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor.* J Biol Chem, 1995. **270**(35): p. 20376-83.
8. Bushell, M., et al., *Polypyrimidine tract binding protein regulates IRES-mediated gene expression during apoptosis.* Mol Cell, 2006. **23**(3): p. 401-12.
9. Luo, M.L., et al., *Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly.* Nature, 2001. **413**(6856): p. 644-7.
10. Strasser, K., et al., *TREX is a conserved complex coupling transcription with messenger RNA export.* Nature, 2002. **417**(6886): p. 304-8.
11. Cheng, H., et al., *Human mRNA export machinery recruited to the 5' end of mRNA.* Cell, 2006. **127**(7): p. 1389-400.
12. Rivas-Aravena, A., et al., *The Elav-like protein HuR exerts translational control of viral internal ribosome entry sites.* Virology, 2009. **392**(2): p. 178-85.
13. Walter, B.L., et al., *Differential utilization of poly(rC) binding protein 2 in translation directed by picornavirus IRES elements.* RNA, 1999. **5**(12): p. 1570-85.
14. Cobbold, L.C., et al., *Upregulated c-myc expression in multiple myeloma by internal ribosome entry results from increased interactions with and expression of PTB-1 and YB-1.* Oncogene. **29**(19): p. 2884-91.
15. Evdokimova, V., et al., *Translational activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition.* Cancer Cell, 2009. **15**(5): p. 402-15.
16. Kaufmann, I., et al., *Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase.* EMBO J, 2004. **23**(3): p. 616-26.
17. Murthy, K.G. and J.L. Manley, *The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation.* Genes Dev, 1995. **9**(21): p. 2672-83.
18. Le Hir, H., et al., *The exon-exon junction complex provides a binding platform for*

factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J, 2001. **20**(17): p. 4987-97.
19. Mayeda, A., et al., *Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing.* EMBO J, 1999. **18**(16): p. 4560-70.
20. Nott, A., H. Le Hir, and M.J. Moore, *Splicing enhances translation in mammalian cells: an additional function of the exon junction complex.* Genes Dev, 2004. **18**(2): p. 210-22.
21. Tange, T.O., A. Nott, and M.J. Moore, *The ever-increasing complexities of the exon junction complex.* Curr Opin Cell Biol, 2004. **16**(3): p. 279-84.
22. Viegas, M.H., et al., *The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway.* Nucleic Acids Res, 2007. **35**(13): p. 4542-51.
23. Zhou, Z., et al., *The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans.* Nature, 2000. **407**(6802): p. 401-5.
24. Sadvakassova, G., et al., *Regulator of differentiation 1 (ROD1) binds to the amphipathic C-terminal peptide of thrombospondin-4 and is involved in its mitogenic activity.* J Cell Physiol, 2009. **220**(3): p. 672-9.
25. Covell, D.G., et al., *Molecular classification of cancer: unsupervised self-organizing map analysis of gene expression microarray data.* Mol Cancer Ther, 2003. **2**(3): p. 317-32.
26. Dressman, H.K., et al., *Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy.* Clin Cancer Res, 2006. **12**(3 Pt 1): p. 819-26.
27. Bacac, M., et al., *A mouse stromal response to tumor invasion predicts prostate and breast cancer patient survival.* PLoS One, 2006. **1**: p. e32.
28. Zhao, A.G., et al., *Effects of Wei Chang An on expression of multiple genes in human gastric cancer grafted onto nude mice.* World J Gastroenterol, 2008. **14**(5): p. 693-700.
29. Yamamoto, H., et al., *Isolation of a mammalian homologue of a fission yeast differentiation regulator.* Mol Cell Biol, 1999. **19**(5): p. 3829-41.
30. Thoren, L.A., et al., *UPF2 Is a Critical Regulator of Liver Development, Function and Regeneration.* PLoS One. **5**(7): p. e11650.
31. Medghalchi, S.M., et al., *Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability.* Hum Mol Genet, 2001. **10**(2): p. 99-105.
32. Weischenfeldt, J., et al., *NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements.* Genes Dev, 2008. **22**(10): p. 1381-96.
33. Thiebaut, M., et al., *Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance.* Mol Cell, 2006. **23**(6): p. 853-64.
34. Tsukahara, K., H. Yamamoto, and H. Okayama, *An RNA binding protein negatively controlling differentiation in fission yeast.* Mol Cell Biol, 1998. **18**(8): p. 4488-98.
35. Wullschleger, S., R. Loewith, and M.N. Hall, *TOR signaling in growth and*

metabolism. Cell, 2006. **124**(3): p. 471-84.
36. Bandhakavi, S., et al., *Quantitative nuclear proteomics identifies mTOR regulation of DNA damage response.* Mol Cell Proteomics. **9**(2): p. 403-14.
37. Das, R., et al., *SR proteins function in coupling RNAP II transcription to pre-mRNA splicing.* Mol Cell, 2007. **26**(6): p. 867-81.
38. Lin, S., et al., *The splicing factor SC35 has an active role in transcriptional elongation.* Nat Struct Mol Biol, 2008. **15**(8): p. 819-26.
39. Xing, Y., et al., *MADS: a new and improved method for analysis of differential alternative splicing by exon-tiling microarrays.* RNA, 2008. **14**(8): p. 1470-9.
40. Xue, Y., et al., *Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping.* Mol Cell, 2009. **36**(6): p. 996-1006.
41. Boutz, P.L., et al., *A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons.* Genes Dev, 2007. **21**(13): p. 1636-52.
42. Hung, L.H., et al., *Diverse roles of hnRNP L in mammalian mRNA processing: a combined microarray and RNAi analysis.* RNA, 2008. **14**(2): p. 284-96.
43. Martinez-Contreras, R., et al., *Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing.* PLoS Biol, 2006. **4**(2): p. e21.
44. Mauger, D.M., C. Lin, and M.A. Garcia-Blanco, *hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc.* Mol Cell Biol, 2008. **28**(17): p. 5403-19.
45. Konig, J., et al., *iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution.* Nat Struct Mol Biol. **17**(7): p. 909-15.
46. Yeo, G.W., et al., *An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells.* Nat Struct Mol Biol, 2009. **16**(2): p. 130-7.
47. Holcik, M. and N. Sonenberg, *Translational control in stress and apoptosis.* Nat Rev Mol Cell Biol, 2005. **6**(4): p. 318-27.
48. Jorgensen, P. and M. Tyers, *How cells coordinate growth and division.* Curr Biol, 2004. **14**(23): p. R1014-27.
49. Bramham, C.R. and D.G. Wells, *Dendritic mRNA: transport, translation and function.* Nat Rev Neurosci, 2007. **8**(10): p. 776-89.
50. Wang, J., et al., *A quality control pathway that down-regulates aberrant T-cell receptor (TCR) transcripts by a mechanism requiring UPF2 and translation.* J Biol Chem, 2002. **277**(21): p. 18489-93.

Summary

Gene expression encompasses several tightly regulated steps which ensure proper protein production from a gene. During transcription a plethora of protein factors assemble at the nascent transcript to initiate the process of intron removal, denominated splicing. Splicing is a flexible mechanism regulated by a high number of factors. For the vast majority of human genes (92-94%), splicing results in the production of more than one mature transcript, and so contributes to the generation of protein diversity within the cell. This process is denominated Alternative Splicing (AS). Many factors regulate AS by promoting inclusion or exclusion of specific exons. We demonstrate through exon array data that Rod1 is a splicing factor which participates in the repression of hundreds of exons in murine erythroleukemia (MEL) cells. It does so together with its paralog Ptbp1 and Raver1, which are also predominantly splicing repressors. CLIP-seq (Cross-linking Immunoprecipitation-sequencing) analysis of Rod1 and Ptbp1-RNA complexes indicates that both proteins associate with RNA. Furthermore, the CLIP-seq data suggest that Rod1 depends on Ptbp1 to associate with at least some RNA sites. Although the interactions between Rod1, Ptbp1 and Raver1 are present in both undifferentiated and differentiated MEL cells, no AS events were detected in undifferentiated cells, indicating that other factors play an important role in the Rod1-mediated AS. Overall these results show that Rod1, Ptbp1 and Raver1 are necessary for the repression of an extensive number of exons in differentiated MEL cells.

Alternative splicing events often generate transcripts containing premature termination codons. These may also originate from nonsense or frameshift mutations. These mRNAs, if translated, result in the production of truncated proteins which may acquire dominant negative functions. The mechanism of Nonsense-Mediated mRNA Decay (NMD) prevents this by eliminating transcripts containing premature termination codons. NMD also targets transcripts with introns in the 3'UTR, transcripts with upstream ORFs and transcripts with long 3'UTRs. NMD is based on a protein complex which assembles during splicing ~24 nucleotides of each exon-exon boundary, the exon junction complex (EJC). The EJC is composed of proteins which participate in splicing, mRNA export, polysome association and NMD. Upf3, Upf2 and Upf1 associate with the EJC and are required for the later stages of NMD. Through immunoprecipitations and mass spectrometry we identified a new Upf2-interacting factor in HEK293 cells denominated Rod1. Rod1 interacts also with Upf1 and with the EJC factor Aly, involved in mRNA export. We demonstrate that Rod1 is required for the NMD-dependent destabilization of a reporter β-globin transcript containing a nonsense mutation in codon 39 (NS39). In additional support of a Rod1 role in NMD, we

show that Rod1, Upf2 and Upf1 regulate hundreds of common genes, several of these potential NMD targets. Approximately 10% of Rod1-regulated genes were up-regulated upon Rod1 or Upf1 knock-down and thus are possibly regulated by NMD. In summary, these results indicate a new role for Rod1 in the mechanism of Nonsense-Mediated mRNA Decay.

Samenvatting

Genexpressie omvat een aantal strikt gereguleerde stappen, die zorgdragen voor juiste eiwitproductie vanaf een gen. Tijdens transcriptie verzamelt zich een overvloedigheid van eiwitten op het opkomend transcript om het proces van intron-verwijdering, dat bekend staat als splicing, te initiëren. Splicing is een flexibel mechanisme dat wordt gereguleerd door een groot aantal factoren. Voor de grote meerderheid van humane genen (92-94%) resulteert splicing in de productie van meer dan één transcript, waarmee het bijdraagt aan de eiwitdiversiteit in een cel. Dit proces wordt ook wel alternatieve splicing (AS) genoemd. Veel factoren reguleren AS middels het bevorderen van in- of uitsluiting van specifieke exonen. Wij laten met behulp van exon array data zien dat Rod1 een splicing factor is die een rol speelt bij de repressie van honderden exonen in muizen erythroleukemische (MEL) cellen. Rod1 werkt hierin samen met zijn paralogen, Ptbp1 en Raver1, welke ook voornamelijk functioneren als onderdrukkers van splicing. CLIP-seq (Cross-linking Immunoprecipitatie Sequencing) analyse van Rod1- en Ptbp1-RNA complexen wijst erop dat beide eiwitten associëren met RNA. De CLIP-seq data suggereert verder dat Rod1 afhankelijk is van Ptbp1 voor binding aan tenminste enkele RNA plaatsen. Hoewel de interacties tussen Rod1, Ptbp1 en Raver1 gedetecteerd worden in zowel ongedifferentieerde als gedifferentieerde MEL cellen, werden er geen voorvallen van AS gedetecteerd in ongedifferentieerde cellen, wat erop duidt dat andere factoren een belangrijke rol spelen in Rod1-gemedieerde AS. Samengenomen laten deze resultaten zien dat Rod1, Ptbp1 en Raver1 vereist zijn voor de repressie van een groot aantal exonen in gedifferentieerde MEL cellen.

Gevallen van alternatieve splicing brengen vaak transcripten voort die vroegtijdige terminatie codons bevatten. Deze kunnen ook ontstaan als gevolg van nonsense of frameshift mutaties. Als deze mRNAs getransleerd worden, resulteren ze in de productie van getrunceerde eiwitten, die mogelijk dominant negatieve eigenschappen vergaren. Dit laatste wordt voorkomen door het Nonsense-Mediated mRNA Decay (NMD) mechanisme, dat transcripten met vroegtijdige terminatie codons elimineert. NMD richt zich ook op transcripten met intronen in de 3'UTR, met upstream ORFs en op transcripten met een lange 3'UTR. NMD is gebaseerd op een eiwitcomplex, het exon junction complex (EJC), dat gedurende splicing wordt opgebouwd op ongeveer 24 nucleotiden afstand van iedere exon-exon overgang. Het EJC bestaat uit eiwitten die betrokken zijn bij splicing, mRNA export, polysoom associatie en NMD. Upf3, Upf2 en Upf1 vervoegen zich bij het EJC en zijn nodig tijdens de latere stadia van NMD. Door middel van immunoprecipitaties en massaspectrometrie hebben we een nieuwe Upf2-interacterende factor, genaamd Rod1,

geïdentificeerd in HEK293 cellen. Rod1 bindt ook aan Upf1 en aan de EJC factor Aly, welke betrokken is bij mRNA export. We laten zien dat Rod1 nodig is voor de NMD-afhankelijke destabilisatie van een β-globine transcript met een nonsense mutatie in codon 39 (NS39). Als verdere onderbouwing voor een rol van Rod1 in NMD laten we zien dat Rod1, Upf2 en Upf1 enkele honderden gemeenschappelijke genen reguleren, waarvan een aantal mogelijk ook gereguleerd worden door NMD. Het expressieniveau van ongeveer 10% van de Rod1-gereguleerde genen ging omhoog wanneer eiwitniveaus van Rod1 of Upf1 werden verlaagd; deze genen worden dus mogelijk gereguleerd door NMD. Samenvattend wijzen deze resultaten op een nieuwe rol voor Rod1 in NMD.

Curriculum Vitae

Name:	Tiago França Brazão
Date of birth:	18-02-1977, in Lisbon, Portugal

Education

1989-1995	Pre-University Education Escola Secundária D. Filipa de Lencastre, Lisbon
1995-2000	Master Biochemistry, Faculty of Sciences, University of Lisbon
2001-2002	Research project Instituto de Tecnologia Química e Biológica (ITQB), Oeiras Analysis of exchanges of genetic material between natural populations of *Streptococcus pneumoniae* (Dr. Mário Ramirez)
2002-2003	PhD research Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisbon Analysis of trans-acting protein factors binding to the NS15 and NS39 β-globin nonsense transcripts in erythroid cells using 2D Gel Electrophoresis (Dr. Luísa Romão)
2003-2010	PhD research Department of Cell Biology, Erasmus MC, Rotterdam Post-transcriptional regulation of gene expression by Rod1 Promotor: Prof.dr. Frank Grosveld

List of Publications

Maranga L, **Brazão TF**, Carrondo MJ. Virus-like particle production at low multiplicities of infection with the baculovirus insect cell system. Biotechnol Bioeng. 2003 Oct 20;84(2):245-53

Acknowledgements

During these years as a PhD student I was lucky to have the support of friends, colleagues and family which made my work possible and easier. The following lines express my gratitude to all of them.

In first place, I want to thank Prof. Frank Grosveld. Frank, you are an inspiring supervisor. I have learned a lot from you, and enjoyed your company inside and outside the lab. I really appreciate the variety of scientific activity which you bring into the department, through seminars, students and facilities. I think it's a good place for PhD students to acquire a solid scientific background. Thank you very much.

John, thank you for your support, patience and supervising, particularly in the first years. I was very inexperienced when I came to the department, but you were very patient and taught me a lot of things, like cloning and biotin pull-downs. I appreciate the effort you did, specially because I was working on quite a different subject than you were.

Luísa, thank you for your help and support. Several times you provided me very useful DNA constructs or antibodies. You were always clearly worried when things were not going well (which unfortunately was a great deal of time). Thank you for the useful discussions we've had when I visited you in Lisbon. Obrigado também à Rute, Ângela, Francisco, Ana, Ana Luísa, Isabel Picanço e à Paula pelo que me ensinaram, pela ajuda que me prestaram e pelo convívio dentro e fora do laboratório.

To Sjaak, Joost, Maarten, Dies and Carmo, thank you very much for being part of the committee, for the corrections, critics and comments which made my thesis better. Sjaak, thank you also for organizing the winter school meetings in Kleinwalsertal. I've learned a lot and had lots of fun there!

Thank you to Jeroen Demmers, Wilfred and Justine Peeters for contributing actively to my work.

A big, big thank you to the members of lab 706. Debbie, Alireza, Erik, Ernie, Maureen, Umut, Maaike and Lilaye, you are exceptional colleagues. I am very lucky to be in a lab which values social contact as much as you do. Our dinners, movie nights and game nights are guaranteed fun! Thank you all so much. Alireza, thank you also for being my paranimf. I always enjoy chatting with you, either about science or something else, and sharing our successes and frustrations in the lab. Raymond, thank you for helping me with some technical issues in my experiments (you are quite good with getting the details right), and for giving me useful critics or comments in many occasions.

Thank you as well to all the past members of 706, with whom I shared fantastic moments. Patrick, Katy, Vesna, Javier (I really miss our half-marathons), Marianne, Jennifer, Edgar,

Miyata, I will always remember you.

To everyone in 702 and 710, thank you for being such nice and helpful lab neighbors. Thank you as well for hanging out with 706 in numerous occasions. Thanks also to Noorie (pity you moved out of Rotterdam, I hope we can walk again in the future!), Frank Sleutels, Sanja, Jeffrey, Dorota, Susana, Chris, Andreia, Anna, Stavros, Akiko, Iris, Eveline, Filippo, Karl, Sander, Ekim, Prashant, Siska, for sharing so many entertaining social events.

Thank you Marike and Jasperina for your help and efficiency.

Thank you so much Eugin, Chris, Harald, Wendy and Audrey for the great moments we shared, movies, poker nights, etc.

Thank you Maaike, Erik and Geert for the physically tiring but mentally refreshing tennis evenings. I'm sure now I can beat Federer... :-)

Debbie, meeting you was the best thing that happened to me in Rotterdam. Thank you for your love and support, and for sharing a life together. Thank you Martin, Anita, Olaf, Erik and Geg for receiving me so well and for being so kind.

Um obrigado especial aos meus amigos Paulo Gonçalves, à Catarina e à Rita. Rita, és uma amiga única, obrigado por seres minha paranimf (como será que se diz isto em português?!...).

Um muito obrigado a toda a minha família. Tenho muita sorte por ter tantos tios, tias, primos e primas com quem gosto tanto de conviver e de quem tenho saudades por estar emigrado. Um agradecimento muito especial aos meus pais e aos meus irmãos (e também ao Tiago, João, Diogo e à Luz), a quem dedico esta tese.

Tiago

Erasmus MC
Universitair Medisch Centrum Rotterdam

PhD Portfolio

Department of Cell Biology	2003-2010
Promotor: Prof.dr. F.G. Grosveld	

1. PhD training

General courses	Year
- Experimental Approach to Molecular and Cell Biology	2004
- From Development to Disease	2005
- In vivo Imaging	2005
- Partek Training Course	2010
- NGS (Next Generation Sequencing) Data Analysis	2010
Seminars and Workshops	
- MGC PhD student workshop poster presentation oral presentation	2003-2008 2003-2006 2007-2008
- Winterschool Transcriptional Control of Developmental Processes oral presentation	2005-2010
- Spetses Summer School poster presentation	2008
International Conferences	
RNA 2003, Vienna, Austria poster presentation	2003
33rd FEBS Congress, Athens, Greece oral presentation	2008
EMBO Meeting, Amsterdam, Netherlands	2009