Transcript-specific characteristics determine the contribution of endo- and exonucleolytic decay pathways during the degradation of nonsense-mediated decay substrates

FRANZISKA OTTENS,[1,4] VOLKER BOEHM,[1,4] CHRISTOPHER R. SIBLEY,[2,4] JERNEJ ULE,[3] and NIELS H. GEHRING[1]

[1]Institute for Genetics, University of Cologne, 50674 Cologne, Germany
[2]Division of Brain Sciences, Imperial College London, London W12 0NN, United Kingdom
[3]Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, United Kingdom

ABSTRACT

Nonsense-mediated mRNA decay (NMD) controls gene expression by eliminating mRNAs with premature or aberrant translation termination. Degradation of NMD substrates is initiated by the central NMD factor UPF1, which recruits the endonuclease SMG6 and the deadenylation-promoting SMG5/7 complex. The extent to which SMG5/7 and SMG6 contribute to the degradation of individual substrates and their regulation by UPF1 remains elusive. Here we map transcriptome-wide sites of SMG6-mediated endocleavage via 3′ fragment capture and degradome sequencing. This reveals that endogenous transcripts can have NMD-eliciting features at various positions, including upstream open reading frames (uORFs), premature termination codons (PTCs), and long 3′ UTRs. We find that NMD substrates with PTCs undergo constitutive SMG6-dependent endocleavage, rather than SMG7-dependent exonucleolytic decay. In contrast, the turnover of NMD substrates containing uORFs and long 3′ UTRs involves both SMG6- and SMG7-dependent endo- and exonucleolytic decay, respectively. This suggests that the extent to which SMG6 and SMG7 degrade NMD substrates is determined by the mRNA architecture.

Keywords: NMD; endonucleolytic cleavage; decay intermediates; degradome sequencing; SMG6; SMG7

INTRODUCTION

Nonsense-mediated mRNA decay (NMD) is a eukaryotic mRNA surveillance mechanism that maintains the fidelity of gene expression by targeting aberrant transcripts. NMD degrades transcripts containing premature termination codons (PTCs) and thereby prevents the synthesis of C-terminally truncated proteins with potentially unphysiological or deleterious functions (Holbrook et al. 2004; Chang et al. 2007; Nicholson et al. 2010). NMD also regulates many cellular mRNAs, which contain one or more NMD-activating features, for example, long 3′ UTRs or upstream open reading frames (uORFs). Thereby, NMD alters the expression levels of ∼5% of the transcriptome in eukaryotes (Kervestin and Jacobson 2012).

In mammalian cells, NMD is activated at termination codons that are located >50 nucleotides (nt) upstream of at least one downstream exon–exon junction that is marked by an exon-junction complex (EJC) (Nagy and Maquat 1998; Thermann et al. 1998; Neu-Yilik et al. 2001). In addition, NMD targets mRNAs with an aberrant architecture downstream from the termination codon, such as unusually long 3′ UTRs (Muhlrad and Parker 1999; Amrani et al. 2004; Bühler et al. 2006; Eberle et al. 2008; Singh et al. 2008). This EJC-independent NMD is activated in response to improper termination by the ribosome, for example, when the interaction of the cytoplasmic poly(A)-binding protein (PABPC1) with the eukaryotic release factors (eRFs) is impaired (Amrani et al. 2004; Eberle et al. 2008; Singh et al. 2008; Fatscher et al. 2014).

Activation of NMD by aberrant translation termination triggers the assembly of a surveillance complex onto the mRNA, consisting of the proteins UPF1, UPF2, and UPF3 (UPF3a and UPF3b in humans). During the initial phases of NMD,

[4]**These authors contributed equally to this work.**
Corresponding author: ngehring@uni-koeln.de
Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.059659.116.

the phosphoinositide 3-kinase-related kinase SMG1 phosphorylates the helicase UPF1 within its extended N- and C-terminal regions (Kashima et al. 2006), which function as binding sites for SMG6 and the SMG5/7 heterodimer (Okada-Katsuhata et al. 2012). It has been reported that SMG5, SMG6, and SMG7 interact preferentially with phosphorylated UPF1 via their 14-3-3-like domains (Fukuhara et al. 2005; Okada-Katsuhata et al. 2012; Kurosaki et al. 2014). However, a phosphorylation-independent interaction between an unstructured region of SMG6 and UPF1 has been recently identified (Chakrabarti et al. 2014; Nicholson et al. 2014).

The degradation of NMD targets occurs via independent pathways that either involve endonucleolytic cleavage or deadenylation and decapping (Mühlemann and Lykke-Andersen 2010). Deadenylation is executed by the SMG5/7 heterodimer, which directly interacts with POP2, a catalytic subunit of the CCR4–NOT deadenylase complex (Loh et al. 2013). Furthermore, deadenylation-independent decapping of NMD substrates is stimulated by interactions of NMD factors with decapping enzymes (Cho et al. 2009; Loh et al. 2013).

The endonucleolytic cleavage of NMD targets is mediated by the C-terminal PIN (PilT N terminus) domain of SMG6 (Glavan et al. 2006; Huntzinger et al. 2008; Eberle et al. 2009). Endocleavage occurs in the vicinity of the termination codon, and the 5′ and 3′ mRNA fragments are exonucleolytically degraded by the exosome and XRN1, respectively (Gatfield and Izaurralde 2004; Huntzinger et al. 2008; Eberle et al. 2009). Hence, 3′ decay intermediates are enriched in cells depleted of XRN1, which enables their amplification and mapping by high-throughput sequencing approaches (Boehm et al. 2014; Lykke-Andersen et al. 2014; Schmidt et al. 2015).

The existence of several degradation pathways used by the NMD machinery raises the question how individual NMD targets are selected for a particular degradation pathway. It has been suggested that NMD substrates are preferentially degraded via SMG6-mediated endonucleolytic cleavage (Gatfield and Izaurralde 2004). SMG5/7-dependent deadenylation and decapping activity was found to be elevated when SMG6 is inactive. However, for complete inhibition of the NMD machinery, both degradation routes have to be inactivated (Jonas et al. 2013; Lykke-Andersen et al. 2014; Schmidt et al. 2015). Hence, alternative degradation pathways are currently considered to ensure efficient degradation of NMD target mRNAs.

In this work, we map the sites of endocleavage and the NMD-eliciting features of many endogenous NMD substrates via degradome sequencing. We find that the endocleavage efficiency of NMD substrates with uORFs and long 3′ UTRs is inversely regulated by the activity of the SMG7-dependent exonucleolytic decay pathway. Interestingly, our results suggest a differential contribution of SMG6- and SMG7-dependent decay routes toward the degradation of different NMD target classes.

RESULTS

Global identification of endogenous 3′ fragments

To identify mRNAs cleaved by SMG6 in a transcriptome-wide manner, we adapted a method to capture and amplify polyadenylated RNAs with a 5′ monophosphate (Fig. 1A; Addo-Quaye et al. 2008; German et al. 2008; Gregory et al. 2008). We generated 3′ fragment libraries with RNA isolated from HeLa cells that were transfected with siRNA against XRN1 or XRN1 and SMG6, respectively (Supplemental Fig. S1A). Comparison of these two conditions allows the identification of XRN1-sensitive, SMG6-dependent cleavage products (Supplemental Fig. S1B). In addition, to determine the background of 5′ monophosphate RNAs, which are not degraded by XRN1, the same protocol was carried out on cells transfected with a negative control siRNA directed against Luciferase (Luc).

High-throughput sequencing and bioinformatics analysis was carried out for two replicate experiments, providing a total of 14,458,242 and 11,573,210 mapped reads for the siXRN1 and siXRN1/siSMG6 conditions, respectively (Supplemental Fig. S1C). The lowest percentage of mapped reads was detected in the control condition (Luc knockdown), as in this case only a small number of 3′ fragments is stable and can be captured by the transcriptome-wide mapping approach. The overall number of mapped reads corresponded to 149,346 low FDR degradation sites (see Materials and Methods), with replicate data sets revealing the most consistent degradation locations and normalized frequencies for each site within conditions (Supplemental Fig. S1D). This demonstrates the high reproducibility of the quantitative information contained within the degradome sequencing data. An average read length of 27 nt (Supplemental Fig. S1E) confirms that we specifically amplified EcoP15I-digested fragments. Annotation of genomic features revealed a strong enrichment of degradome sequencing fragments within processed mRNAs (Supplemental Fig. S1F), as SMG6-dependent reads preferentially map within the coding sequence (CDS) and 5′ and 3′ UTR, but not within introns or intergenic regions. Interestingly, in the library from XRN1-depleted cells alone, we observed a strong peak at the ORF—3′ UTR, but not 3′ UTR—intergenic border, demonstrating a preference for SMG6 endocleavage at the termination codon (Fig. 1B). In contrast, cleavage is enriched across the whole 5′ UTR, but without any positional preference. This could reflect cleavage in uORFs that can occur at various positions within 5′ UTRs of different genes, or generally indicate high cleavage across 5′ UTRs from other unknown phenomena.

To separate genes that had XRN1-sensitive, SMG6-dependent cleavage products associated with different features, we used a sliding window approach to identify 100-nt segments across the transcriptome that are most enriched in "XRN1 siRNA" reads relative to "XRN1/SMG6 siRNA" reads

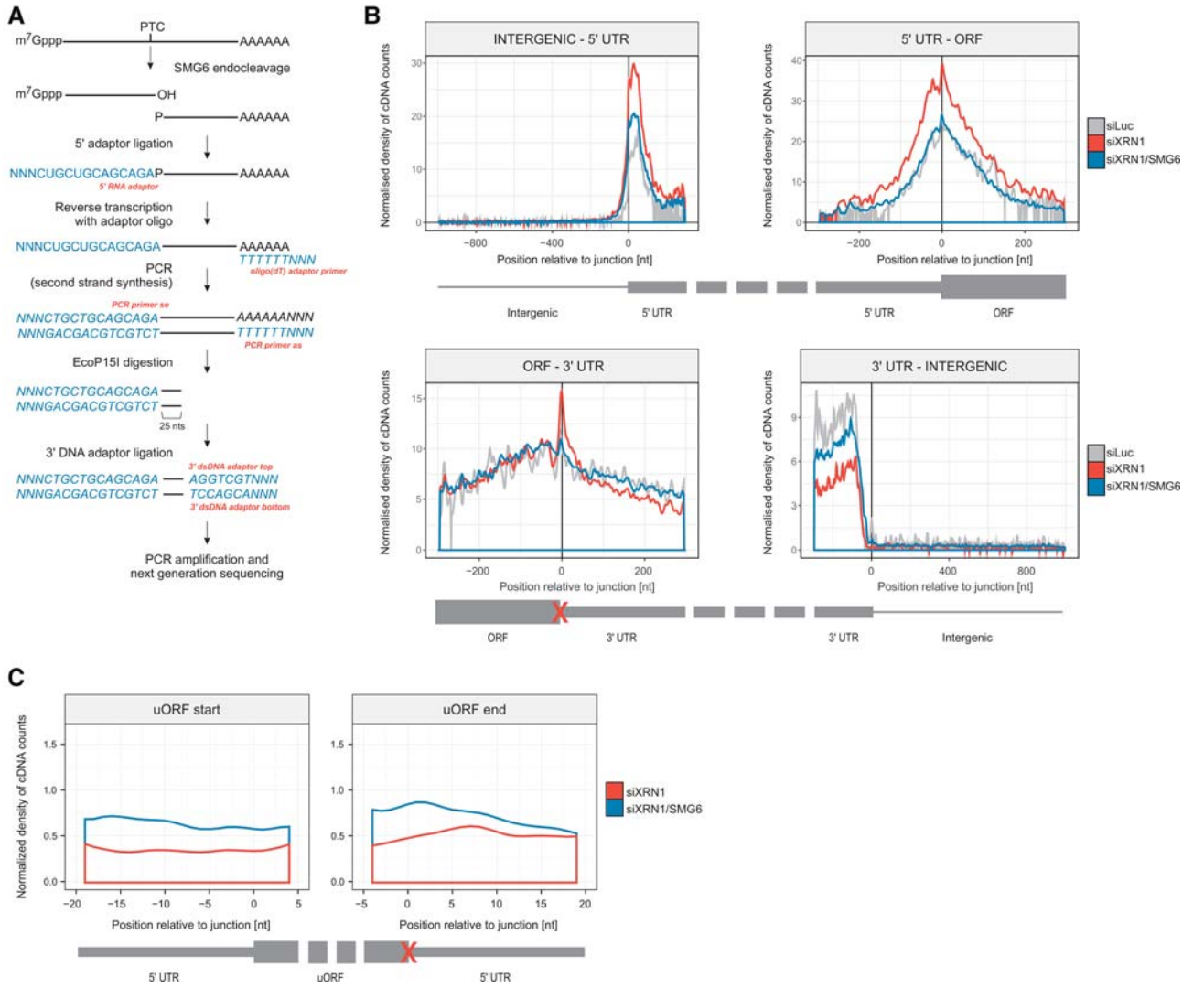

FIGURE 1. Computational analysis of 3′ fragment libraries. (*A*) Workflow of 3′ fragment library preparation for Illumina sequencing. Poly(A)$^+$ RNA was used for library construction. The free 5′ phosphate of 3′ fragments allows the ligation of a 5′ RNA linker. Subsequently, RT-PCR was performed with a linker oligo(dT) followed by second strand synthesis and EcoP15I digest to equally shorten 3′ fragment length down to 25–27 nt. Next, a 3′ dsDNA adaptor was ligated, harboring a fixed barcode and an index sequence to distinguish between knockdown conditions, followed by construct amplification for Illumina sequencing. (*B*) Normalized densities of cDNA reads plotted around indicated features for each degradome-sequencing library. The termination codon is indicated by a red cross. (*C*) Normalized density of cDNA reads plotted around uORF starts and uORF ends. All uORFs with at least one read in the siXRN1 condition were evaluated. A red cross indicates the position of the uORF termination codon. See Materials and Methods for full details.

(Supplemental Table S1). To assign genes to a feature (5′ UTR, ORF, 3′ UTR), we used a strict threshold of >50 reads in the window for the XRN1 siRNA condition and a fold-difference of >10 between the "XRN1 siRNA" and "XRN1/SMG6 siRNA" conditions. The overlapping feature (5′ UTR, ORF, 3′ UTR) of the window with the highest *z*-score in each mRNA was then assigned to the corresponding gene. In all identified SMG6-sensitive genes, the highest-scoring window was found within the following features: 1079 in ORFs, 271 in 5′ UTRs, 118 in 3′ UTRs. We additionally assessed how many of these identified target windows were cleaved in proximity to an annotated termination codon via the positional evaluation of each window's highest peak. This analysis determined that 19% of 3′ UTR sites were within 50 nt of an annotated termination codon, in contrast to 11% for ORF windows and 0% for 5′ UTR windows (Supplemental Fig. S1G). In agreement with genome-wide maps at the ORF-3′ UTR boundary (Fig. 1B), there was preference for max peaks to be located at or after the termination codons (Supplemental Fig. S1H).

To further investigate the previously observed overall enrichment of SMG6-dependent cleavage sites in the 5′ UTR (Fig. 1B), we probed uORF-containing mRNAs for SMG6-specific cleavage events. Analysis of all uORF-containing

mRNAs showed no enrichment of reads (Fig. 1C; Fritsch et al. 2012). This could be due to the described NMD-escaping behavior of uORFs (Stockklausner et al. 2006) or alternatively due to technical limitations associated with picking up 5′ UTR degradation products with a poly(A) enrichment method.

Implication of SMG6-specific cleavage events for different classes of endogenous NMD targets

Endogenous NMD targets can be classified by one or more NMD activating features, such as long 3′ UTRs, upstream open reading frames (uORFs), PTCs resulting from alternative or aberrant splicing events, or the presence of a selenocysteine codon (UGA). For each of these classes, we identified representative targets during our transcriptome-wide analysis of decay intermediates. The results from high-throughput target mapping also allowed us to obtain endocleavage patterns at single-nucleotide resolution. Previous studies have defined preferential sites of endocleavage at and downstream from termination codons (Eberle et al. 2009; Boehm et al. 2014). For selected transcripts, we find SMG6-dependent 3′ fragments mapping to positions at and downstream from termination codons located upstream of long 3′ UTRs (Fig. 2A), selenocysteine codons (Fig. 2C), or out-of-frame stop codons (Fig. 2D). Although we were not able to detect endocleavage at uORFs on a transcriptome-wide scale (Fig. 1C), individual examples of SMG6 activity around uORF stop codons could readily be detected (Fig. 2B). Hence, we reproduced the previously described pattern of SMG6-mediated cleavage of mRNAs in the vicinity of the NMD-activating termination codon for endogenous NMD substrates (Lykke-Andersen et al. 2014; Schmidt et al. 2015). Interestingly, some targets (C11orf31 and TMEM222) showed an additional enrichment of mapped reads upstream of the stop codon, but the relevance of this site remains to be determined (Fig. 2C,D). In order to validate that these mRNAs are bona fide NMD targets, we used a semiquantitative PCR-based approach to visualize 3′ decay intermediates. This PCR approach confirmed SMG6-dependent endocleavage of PDRG1, MED10, and SURF6 (long 3′ UTR, Fig. 2A), BAG1 and IFRD1 (uORF, Fig. 2B), and C11orf31 (selenocysteine codon, Fig. 2C) in the close proximity of the termination codon. The analysis of RPL10A and TMEM222 revealed SMG6-dependent 3′ fragments in an annotated, intronic region (RPL10A) or in an alternative exon with an in-frame stop codon (TMEM222), respectively (Fig. 2D). Endonucleolytic decay of all analyzed mRNAs was strictly UPF1-dependent, confirming that endocleavage occurred during NMD.

Endocleavage efficiency of several NMD targets is regulated by SMG7 levels

To study the regulation of NMD substrates, we constructed triosephosphate isomerase (TPI) reporters with 3′ UTRs of two endogenous targets, PDRG1 and SURF6, downstream from the TPI stop codon (Supplemental Fig. S2A). Other reporter constructs containing the shortened SMG5 3′ UTR, the UPF3b 3′ UTR, and the GFP coding sequence were characterized before and shown to elicit SMG6-dependent endocleavage (Boehm et al. 2014). We confirmed, in addition for the PDRG1 and SURF6 3′ UTR-containing reporters, that they undergo SMG6- and UPF1-mediated degradation and endocleavage (Supplemental Fig. S2B).

Two pathways are involved in the degradation of NMD targets, SMG5/7-dependent deadenylation and decapping (Loh et al. 2013), as well as endonucleolytic degradation via SMG6 (Glavan et al. 2006; Huntzinger et al. 2008; Eberle et al. 2009). Recent high-throughput analyses showed an increase in decapped mRNAs upon XRN1 and SMG6 co-depletion, which was attributed to enhanced deadenylation-mediated by SMG7 (Lykke-Andersen et al. 2014; Schmidt et al. 2015). Therefore, we aimed to investigate if SMG7 expression levels influence the efficiency of SMG6-dependent endocleavage. In XRN1 and SMG7 co-depleted cells expressing a TPI-SURF6 reporter construct we readily detected 3′ fragments, which were reduced when siRNA-resistant SMG7 was expressed (Fig. 3A). It is conceivable that 3′ decay intermediates are stabilized due to decreased deadenylation activity upon SMG7 knockdown. To this end, we co-depleted XRN1 and CNOT1, a crucial scaffolding component of the CCR4–NOT deadenylase complex. We detected comparable 3′ fragment levels in XRN1 and XRN1/CNOT1 depletion conditions, indicating that the 3′ fragments are not directly degraded via deadenylation. Furthermore, monitoring 3′ fragment stability after transcriptional shut-off of stable cell lines expressing the TPI-SURF6 reporter showed that the block of deadenylase activity (XRN1/CNOT1 knockdown) did not alter the degradation rate of 3′ decay intermediates (Supplemental Fig. S2C,D). The influence of SMG7 on 3′ fragment levels was also observed using TPI constructs with a SMG5, UPF3b, or GFP 3′ UTR, which behaved similarly to the TPI-SURF6 construct (cf. Fig. 3A and Supplemental Fig. S2E). The inverse correlation of 3′ fragments with SMG7 expression levels was also detected by semiquantitative PCR for many of the endogenous NMD substrates However, the 3′ fragments of both targets with PTCs, incorporated due to alternative splicing (TMEM222 and RPL10A), were not affected by SMG7 levels (Fig. 3B). Additionally, we performed the PCR assay with cDNA that was generated by a poly(A)-independent method (Supplemental Fig. S2F). This showed that the observed effect on 3′ fragment abundance indeed results from differential SMG7 expression levels and not from enhanced or impaired deadenylation of the 3′ endocleavage products. Taken together, these observations indicate a differential mode of degradation of target mRNAs, depending on their NMD-inducing features. While PTC-containing RNAs mainly undergo efficient endocleavage, the decay of transcripts with uORFs, selenocystein codons, and long 3′ UTRs is regulated by both SMG6-mediated endocleavage and SMG7-induced deadenylation/decapping.

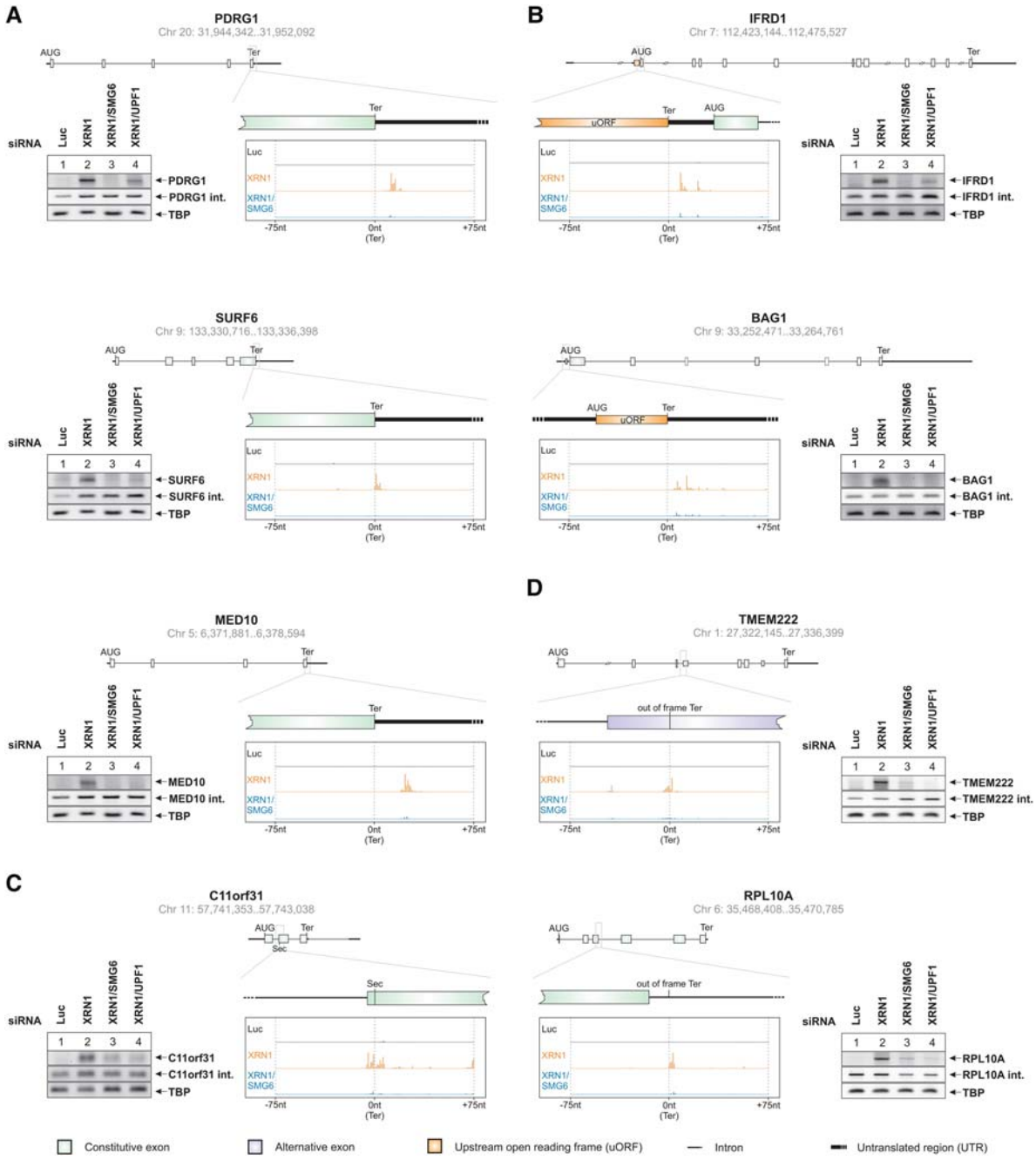

FIGURE 2. Functional analysis of potential endogenous NMD targets identified during high-throughput sequencing. (*A–D*) A PCR-based approach was used to quantify 3′ fragment levels upon knockdown of NMD effectors. HeLa cells were transiently transfected with the indicated siRNAs and poly(A)$^+$ RNA was extracted. 3′ Decay intermediates were ligated to an RNA linker, followed by reverse transcription with an oligo(dT) primer and gene-specific PCR. Overall transcript levels were determined with primer pairs located downstream from the estimated endocleavage site (second panel each, indicated as "gene int."). PCR with TATA-Box binding protein (TBP) primers was used for cDNA level determination. For each class of NMD targets (*A*, *B*, *C*, *D*, respectively), the same set of cDNA was used and therefore the same TBP profiles are shown for each target type. Degradome-sequencing reads were plotted against their position on the indicated mRNAs. Endocleavage events within the selected targets were visualized in an enlarged view spanning 150 nt (±75 nt). The position of the second nucleotide of the respective stop codon is set as zero. Mapped reads per nucleotide were plotted against the mRNA length for each knockdown condition (Luciferase [Luc, black], XRN1 [orange], XRN1/SMG6 [blue]). (*A*) PDRG1, SURF6, and MED10 transcripts with a long 3′ UTR. (*B*) IFRD1 and BAG1 transcripts with an uORF. (*C*) C11orf31, encoding for Selenoprotein H, containing a selenocysteine (Sec) codon. (*D*) TMEM222, incorporation of an alternative exon (indicated in purple), harboring a PTC as well as RPL10A, a transcript gaining a PTC probably due to alternative splice site usage within the third intron.

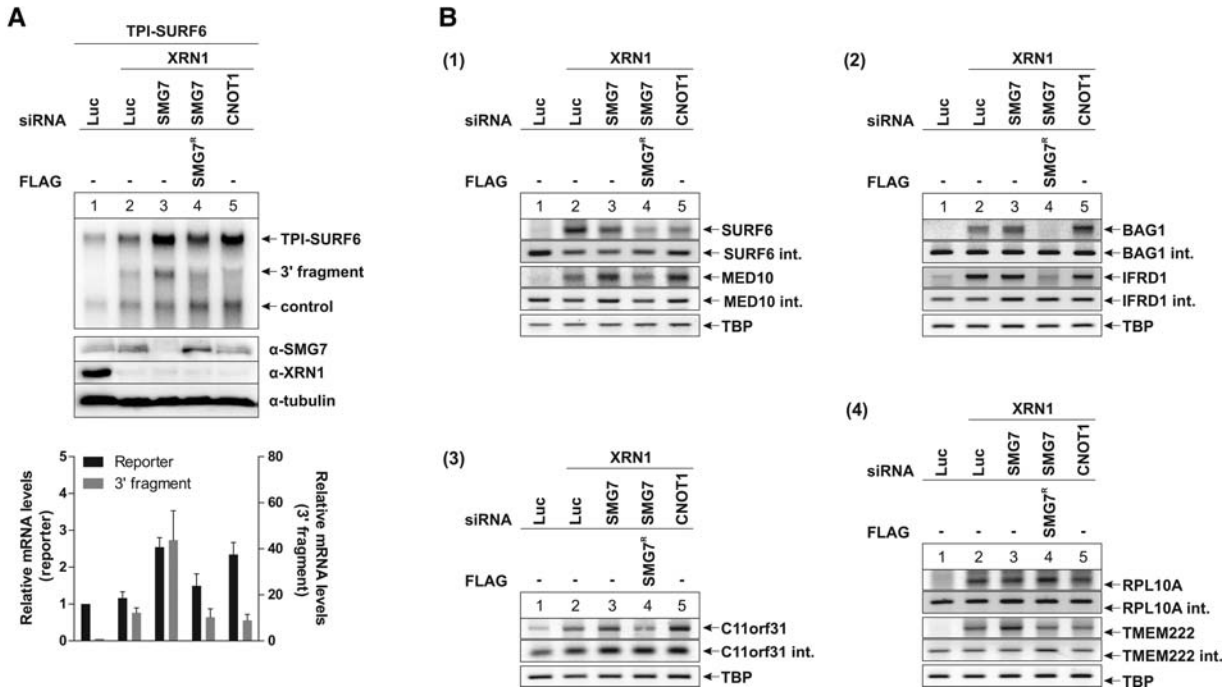

FIGURE 3. Endocleavage of endogenous targets is regulated SMG7-dependently. (*A*) Northern blots of RNA samples extracted from HeLa cells transfected with the indicated siRNAs, plasmids, and reporter constructs using β-globin as control mRNA. Endocleavage products (3′ fragments) are indicated. A representative Western blot is shown at the *bottom*, using tubulin as loading control. Mean values of reporter and 3′ fragment signal ± SD ($n = 3$) were quantified and normalized to the Luc control knockdown. (*B*) Detection of endogenous 3′ fragments for different classes of NMD targets ([1] long 3′ UTR, [2] uORF, [3] selenocysteine codon, [4] alternative splicing) dependent on SMG7 expression levels. HeLa cells were transiently transfected with the indicated siRNAs and plasmids. The same PCR-based approach as shown in Figure 2 was applied. The cDNA was transcribed from total RNA with an oligo(dT) primer, and TBP served to determine overall cDNA levels. For each class ([1]–[4]), the same set of cDNA was used.

Endonucleolytic cleavage is not saturated by increasing levels of NMD substrates

Our results suggest substantial differences in the requirement of NMD factors and the mechanism of degradation for different endogenous and transfected NMD substrates. Two redundant degradation pathways may act as a fail-safe mechanism to ensure efficient degradation of target mRNAs even if the integrity of one of the pathways is compromised. We speculate that such a fail-safe mechanism may be relevant for cells that produce large amounts of NMD substrates, and increasing levels of NMD substrates may activate the alternative SMG7-dependent NMD pathway. To this end, we transfected variable amounts of plasmids expressing long 3′ UTR- or PTC-containing mRNAs (Fig. 4A,B). We observed no change of endocleavage efficiency with increasing amounts of NMD substrates, suggesting that the NMD machinery and specifically SMG6 is able to cope with large quantities of target mRNA.

We also investigated whether increasing amounts of PTC-containing reporter mRNA, which is cleaved efficiently, can influence the degradation of constant amounts of long 3′ UTR mRNA, which is on average a less efficient SMG6 substrate (Boehm et al. 2014). However, we observed neither an increase in reporter levels nor a decrease in 3′ fragments generated from the TPI-SMG5 reporter when the levels of TPI-PTC160 reporter were increased (Fig. 4C). Interestingly, overexpression of SMG6 WT increased endocleavage efficiency when TPI-SMG5, but not TPI-PTC160, was used as reporter mRNA (Fig. 4). This observation further indicates a differential contribution of SMG6- and SMG7-dependent decay routes toward the degradation of different NMD target classes.

SMG7-mediated regulation of endocleavage depends on functional UPF1 binding

To better understand the interplay between SMG6 and SMG7, we performed overexpression studies using different SMG7 constructs (Fig. 5A). Specifically, we asked whether SMG7 regulates endocleavage via its interaction with UPF1 or by a UPF1-independent mechanism. To this end, we compared the effects of full-length wild-type SMG7, full-length SMG7 with a mutation in the 14-3-3-like domain (14-3-3^{Mut}; unable to interact with UPF1), and a C-terminally truncated SMG7 variant (1-497) (Fukuhara et al. 2005). SMG7 (1-497) consists only of the 14-3-3-like and helical domain and therefore lacks the region required for interaction with the deadenylase POP2 (Loh et al. 2013). The

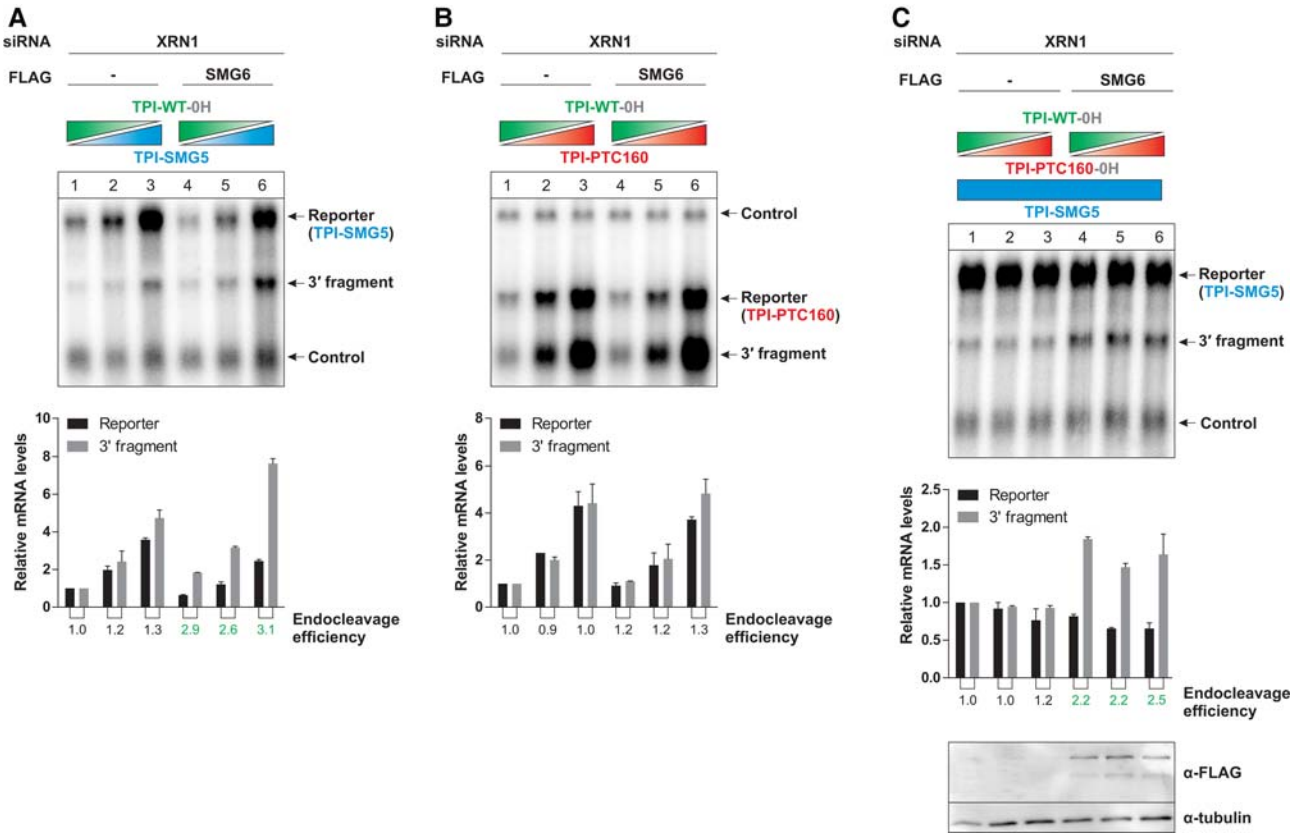

FIGURE 4. Robustness of the NMD system activated via different decay-inducing features. (*A–C*) Northern blots of RNA samples extracted from HeLa cells transfected with the indicated siRNAs, plasmids, and reporter constructs. Cotransfected β-globin (*A*,*C*) or LacZ-H4 (*B*) served as control mRNA. Endocleavage products (3′ fragments) are indicated. To eliminate effects caused by altered transfection efficiency or availability of the general gene expression machinery, decreasing amounts of TPI-WT-0H ([*A*,*B*] 3, 2, and 0 µg; [*C*] 4, 2, and 0 µg) were cotransfected with increasing amounts of reporter ([*A*,*B*] 1, 2, and 4 µg; [*C*] constant 3 µg) or competition vector ([*C*] 0, 2, and 4 µg). The 0H constructs lack the Northern blot probe binding sites and are not detected by the 3′ probe. Mean values of reporter and 3′ fragment signal ± SD ($n = 3$) were quantified and normalized to the control. The ratio of 3′ fragment to reporter mRNA levels is indicated *below* the bars. A representative Western blot is shown at the *bottom*, using tubulin as loading control (*C*).

overexpression of wild-type SMG7 (SMG7 WT) or 1-497 resulted in reduced 3′ fragment generation from TPI-SMG5, whereas overexpression of SMG7 14-3-3^{Mut} did not alter endocleavage activity (Fig. 5B). Notably, elevated reporter mRNA levels were observed for SMG7 1-497, which is not able to activate substrate deadenylation (compare lanes 1 and 4). These results suggest that some NMD substrates may be degraded either by the SMG6- or SMG7-dependent pathways, and the cellular availability of SMG6 and SMG7 therefore determines which decay mechanism prevails. Interestingly, we did not detect reduced endocleavage of a PTC reporter mRNA (TPI-PTC160) (Fig. 5C) when SMG7 was overexpressed, indicating that SMG7 does not influence SMG6-mediated endocleavage of this substrate (Fig. 5D).

The UPF1 C-terminal extension is required for endocleavage of long 3′ UTR targets

We next focused on UPF1, which represents a central key molecule of the NMD machinery and serves as a binding platform for different decay factors. The kinase SMG1 phosphorylates UPF1 at SQ and TQ motifs in the unstructured N- and C-terminal regions in order to recruit SMG5, SMG6, and SMG7 (Yamashita et al. 2001; Chakrabarti et al. 2014; Kurosaki et al. 2014). Previously, T28 in the UPF1 N terminus was identified as the binding site for SMG6, whereas the phosphorylated UPF1 C terminus mediates the interaction with the SMG5/7 heterodimer (Okada-Katsuhata et al. 2012). As revealed by SMG7 overexpression, the interaction of overexpressed SMG7 with UPF1 reduced endocleavage of long 3′ UTR targets (Fig. 5B). We therefore studied the role of the SMG7 binding sites (C-terminal extension of UPF1) during endocleavage by analyzing truncated UPF1 proteins in complementation assays (Fig. 6A). Both, deletion of the complete C terminus (N-A2) as well as the very C-terminal cluster of SQ/TQ motifs (N-1030), which retains the helicase regulating SQ region of UPF1 (Fiorini et al. 2013), increased TPI-PTC160 reporter mRNA levels in complementation assays (Fig. 6B). Although the endocleavage efficiency was reduced due to

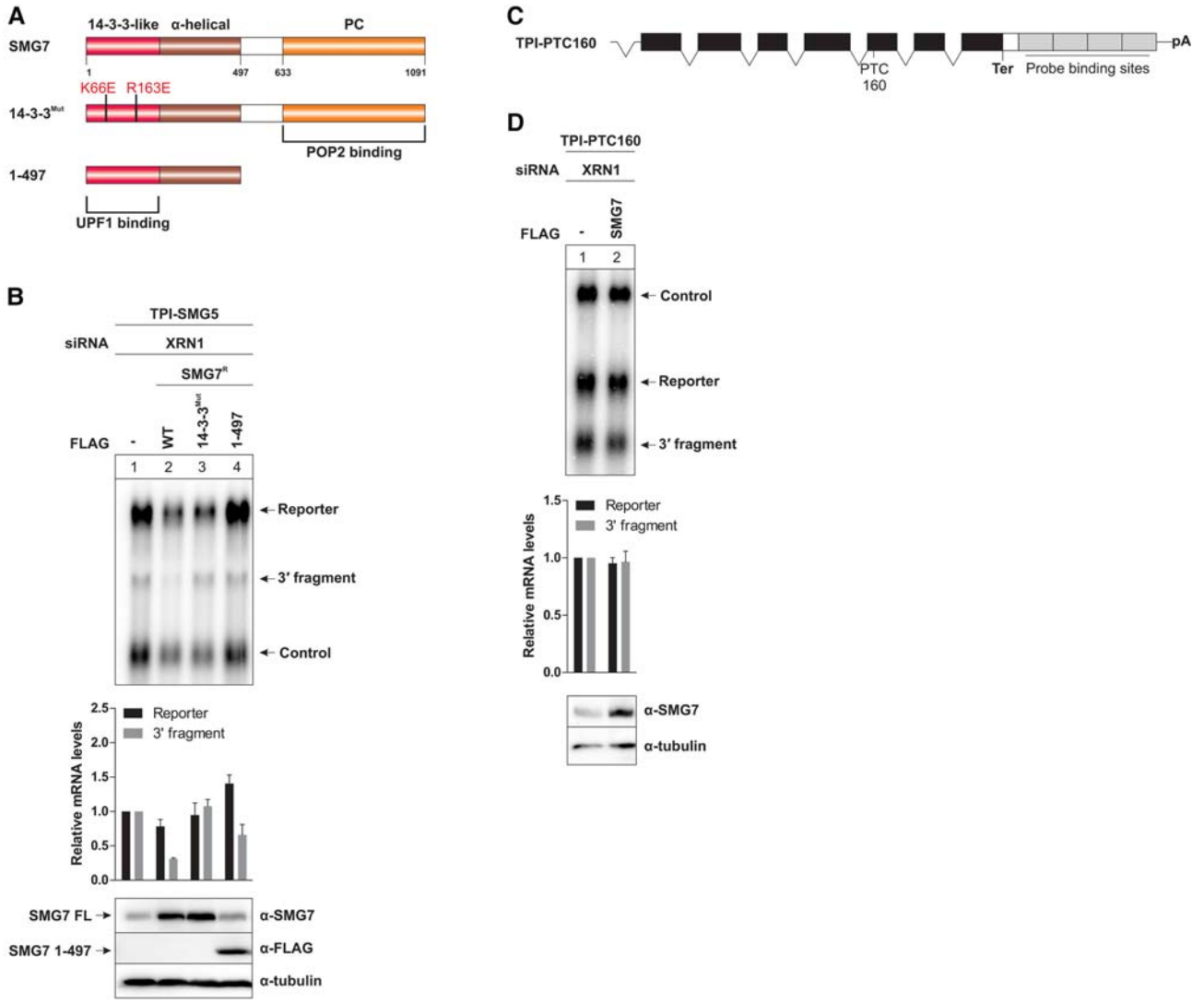

FIGURE 5. SMG7 abundance influences endocleavage efficiency for long 3′ UTR targets. (*A*) Domain structure of SMG7, showing the N-terminal 14-3-3-like domain (interacts with UPF1) and α-helical extensions as well as the C-terminal PC region (interacts with POP2). (*B*) Northern blot of RNA samples extracted from HeLa cells transfected with the indicated siRNAs and reporter constructs. Cotransfected β-globin served as control mRNA. Endocleavage products (3′ fragments) are indicated. Mean values of reporter and 3′ fragment signal ± SD ($n = 3$) were quantified and normalized to the XRN1 control knockdown. A representative Western blot is shown at the *bottom*, using tubulin as loading control. (*C*) Schematic representation of the transfected triosephosphate isomerase (TPI) reporter with a PTC at amino acid position 160. Exons are depicted as white (untranslated) or black (translated) boxes, introns as two connecting black lines, and the Northern probe binding sites as light gray boxes. Vector derived 5′ UTR intron and SV40 poly(A) signal (pA) are shown. (*D*) Northern blot of RNA samples extracted from HeLa cells transfected with the indicated siRNAs, plasmids, and reporter constructs using LacZ-4H as control mRNA. Endocleavage products (3′ fragments) are indicated. Mean values of reporter and 3′ fragment signal ± SD ($n = 3$) were quantified and normalized to the XRN1 control knockdown. A representative Western blot is shown at the *bottom*, using tubulin as loading control.

the accumulation of reporter mRNA, the total amount of 3′ fragments changed only moderately, indicating that endocleavage is maintained in spite of decreased NMD activity. In contrast, deletions of the C terminus strongly reduced endocleavage of TPI-SMG5 mRNA, suggesting that the truncated UPF1 proteins are inactive on NMD substrates with long 3′ UTRs (Fig. 6C). This finding is in accordance with the results of the SMG7 overexpression assay (Fig. 5) and thereby further strengthens the view that the combined activity of SMG6 and SMG7 is implicated in the degradation of non-PTC NMD targets and that the mode of degradation depends on the mRNA architecture and the availability of SMG6 and SMG7.

DISCUSSION

NMD of many substrate mRNAs is initiated by endonucleolytic cleavage, and SMG6 represents an important NMD factor in human cells. In this work, we obtained insights into the

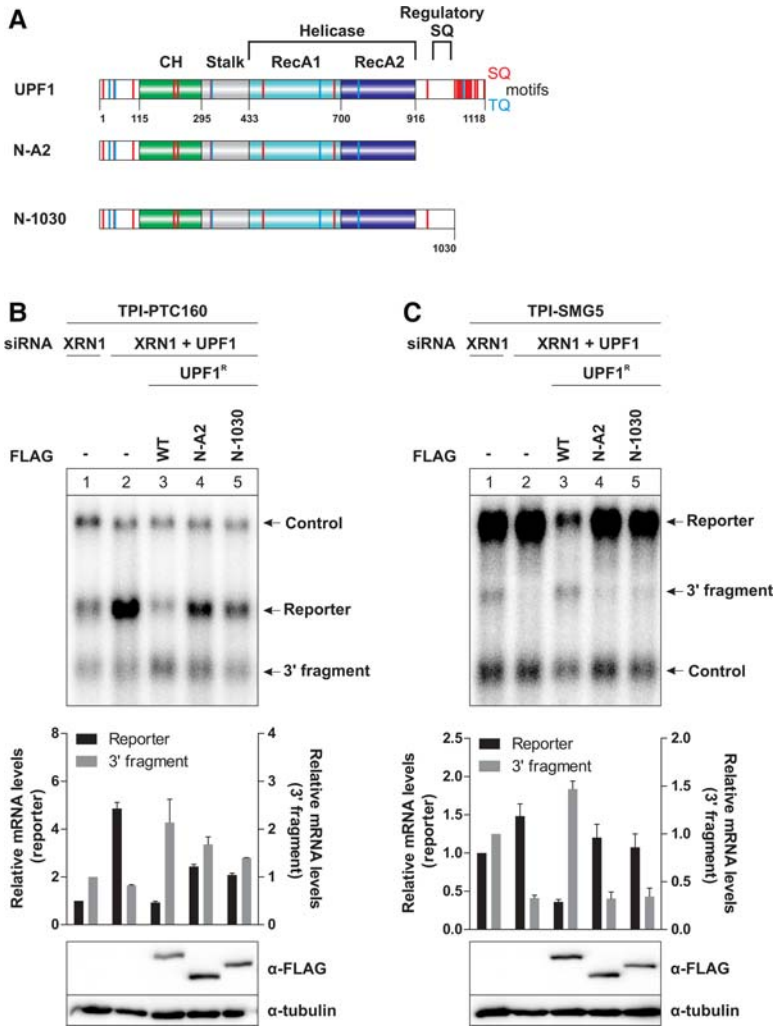

FIGURE 6. Deletions of C-terminal UPF1 phosphorylation sites impair NMD differently. (*A*) UPF1 protein architecture is depicted schematically. All structural and functional domains are indicated; the presence of potential phosphorylation sites (SQ/TQ) are shown in red and blue, respectively. (*B*,*C*) Northern blots of RNA samples extracted from HeLa cells transfected with the indicated siRNAs and reporter constructs. Cotransfected LacZ-H4 (*B*) or β-globin (*C*) served as control mRNA. Endocleavage products (3′ fragments) are indicated. Mean values of reporter and 3′ fragment signal ± SD ($n = 3$) were quantified and normalized to the XRN1 control knockdown. Representative Western blots are shown at the *bottom*, using tubulin as loading control.

molecular mechanism of endonucleolytic cleavage by SMG6 and its regulation in the context of NMD. Furthermore, we identified endogenous NMD targets on a transcriptome-wide scale and mapped sites of SMG6 endocleavage at nucleotide resolution. This analysis confirmed that endocleavage of NMD substrates occurs mainly at and downstream from termination codons. In fact, our unbiased approach enabled us to identify termination codons that are recognized as aberrant and to discriminate between NMD targets with uORFs, long 3′ UTRs, selenocysteine codons, and those produced by alternative splicing. Interestingly, the "degradative" NMD of some transcripts is dominated by rapid endocleavage, while "regulatory" NMD of other mRNAs uses endo- and exonucleolytic decay pathways. Hence, the identification of two partially redundant degradation pathways is an important aspect of this work. We note that our method would in principle allow the detection of transcripts that are targeted for degradation due to transcription or splicing errors. However, this biologically important class of NMD substrates, which escapes quality control during production or processing, is currently not recorded because of the low abundance of individual transcripts. Hence, it will be important to further improve the methodology to identify the whole spectrum of SMG6 substrates in human cells.

We observe that some of the endocleavage sites we identify here overlap with previously reported sites (Schmidt et al. 2015). In contrast, there seems to be little overlap with another study, in which HEK 293 cells were used (Lykke-Andersen et al. 2014). Although this variation could be due to a different sequencing technique that was used, it may also indicate that certain NMD substrates are cell-type specific and others ubiquitous. A comparison of NMD substrates of different cells and/or organisms will be required to identify a core set of NMD targets.

Recently, an increased decapping of NMD substrates was reported when SMG6, but not UPF1, was depleted (Lykke-Andersen et al. 2014; Schmidt et al. 2015). This suggested that mRNAs, which are normally degraded by SMG6, can be targeted by an alternative decay pathway involving SMG5/7-mediated deadenylation and decapping. This observation is consistent with reports that NMD is strongly inhibited by the combined depletion of SMG6 with either SMG5 or SMG7 (Jonas et al. 2013). This raises the question how individual NMD substrates are assigned to different degradation pathways, i.e., what determines whether an mRNA is degraded preferentially by endocleavage or deadenylation and decapping. To address this question, we altered the expression levels of SMG7 by overexpression or depletion in cultured cells. We chose SMG7, because it interacts directly with phosphorylated UPF1, and several SNPs have been found to be associated with differences in the expression of SMG7 (Nguyen et al. 2014). Hence, SMG7 may show highly diverse expression levels in different individuals. Interestingly, the efficiency of endocleavage of a substrate mRNA with a long 3′ UTR inversely correlated with the

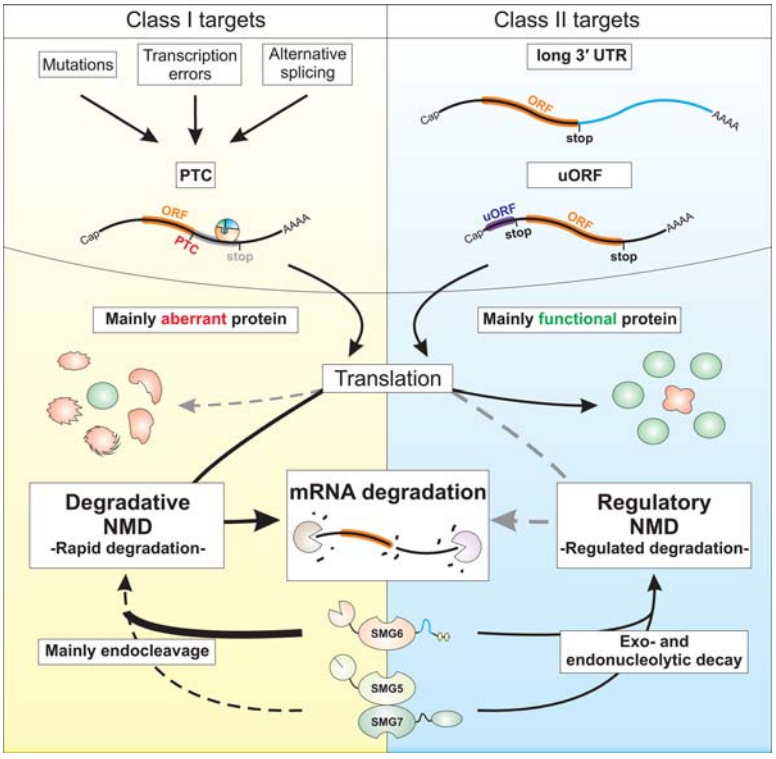

FIGURE 7. Model of degradative and regulatory branches of NMD. Depending on the mRNA architecture, NMD substrates are either regulated (long 3′ UTR and uORF targets; encoding for mostly functional protein) or degraded (PTC targets; encoding for aberrant proteins). During degradative NMD, the robust elimination of PTC-containing mRNAs is mainly achieved by endonucleolytic decay via SMG6, which is supported by exonucleolytic degradation induced by SMG5/7. Alternatively, the combined activity of SMG5/7 and SMG6 (exo- and endonucleolytic degradation) during the decay of long 3′ UTR- and uORF-containing targets allows for regulated degradation ("down-regulation"). For further details, see Discussion.

expression levels of SMG7. This suggests that the degradation of these NMD substrates uses both endo- and exonucleolytic decay and their contribution depends on the levels of SMG6 and SMG7, respectively. Therefore, the complete lack of a SMG7 gene in the *Drosophila melanogaster* genome may explain why endocleavage of NMD targets has been originally identified in cultured cells from *D. melanogaster* (Gatfield and Izaurralde 2004).

Moreover, we find that different classes of endogenous NMD substrates show different levels of partitioning between SMG6- and SMG7-dependent decay. Transcripts containing uORFs, long 3′ UTRs, or selenocysteine codons undergo more efficient endocleavage when SMG7 levels are low, while endocleavage efficiency of PTC-containing transcripts is not affected by SMG7 abundance. This suggests a previously unrecognized complexity of NMD and indicates that endogenous targets are either rapidly degraded by the SMG6-mediated endonucleolytic degradation, or slowly degraded ("down-regulated") by the weaker activity of SMG6, and the slower SMG7-dependent decay (Fig. 7). The difference between these degradative and regulatory effects of NMD is consistent with the different likelihood of NMD targets to encode functional proteins. PTCs are most often caused by mutations or aberrant splicing in the ORF that can produce aberrant proteins. SnoRNA host transcripts are also rapidly degraded by the SMG6-mediated endocleavage, and their mature mRNAs that result from splicing-coupled snoRNA production do not encode functional proteins (Lykke-Andersen et al. 2014). In contrast, mRNAs with uORFs or long 3′ UTRs are most often canonical transcript isoforms that encode functional proteins. It is thus appropriate that functional mRNAs are not efficiently degraded by the SMG6-mediated endocleavage, and are instead more amenable to the slower degradation by SMG7. Although we clearly observe differences between the mRNA targets of "degradative" and "regulatory" NMD pathways, the molecular characteristics of the two classes have not been explicitly determined and need to be investigated in the future. However, we speculate that NMD caused by defective ribosome recycling may be a common feature of regulated substrates (Fatscher et al. 2014; Joncourt et al. 2014).

The existence of two partially redundant degradation pathways might be required when cells need to cope with increased amounts of NMD targets. This could occur when other quality control systems are impaired or when cells are exposed to highly mutagenic conditions. However, our experimental data (Fig. 4) demonstrate that SMG6 catalyzes endocleavage of overexpressed NMD substrates as efficiently as small amounts thereof. Hence, we suggest that alternative degradation mechanisms have not been implemented to process an excess of NMD substrates, but to regulate NMD in a transcript-specific manner (Fig. 7). Inherent features of the transcript, such as mRNA architecture, determine whether the NMD machinery executes a degradative or down-regulating function. While NMD has been originally considered as a mechanism to degrade faulty mRNAs, our results underline its growing importance as a master regulator of gene expression (Nicholson et al. 2010; Karam et al. 2013; Lykke-Andersen and Jensen 2015).

MATERIALS AND METHODS

Plasmids and cell culture

Plasmid constructs β-globin, LacZ-4H, pCI-Flag, pCI-mVenus, WT, and PTC-containing TPI and expression vectors for UPF1

and SMG6 were described previously (Gehring et al. 2009; Boehm et al. 2014). SMG7 was cloned from HeLa cDNA and mutated to confer siRNA insensitivity (targeting sequence 5′-CGATTTGGAA TACGCTTTA-3′ replaced by 5′-TGACCTTGAGTATGCCCTG-3′). Point or deletion mutants of UPF1 and SMG7 were generated by site-directed mutagenesis, cloned in the designated vector, and verified by sequencing. Standard protocols were used to generate the stable HeLa Flp-In T-REx TPI-SURF6 cell line (HeLa Flp-In T-REx initially established by Elena Dobrikova and Matthias Gromeier, Duke University Medical Center). Expression of stable cell lines was induced with 1 µg/mL doxycycline for 24 h. All cell lines were cultured in DMEM (Gibco) supplemented with 9% fetal bovine serum (FBS; Gibco) and 1× Pen Strep (Gibco), and the cells were incubated at 37°C, 5% CO_2, and 90% humidity.

siRNA transfections

HeLa Tet-Off cells (5×10^5) (Clontech) were grown overnight in 6 cm plates and transiently transfected with 200 pmol siRNA for single or 400 pmol total siRNA for double knockdowns using Lipofectamine RNAiMAX (Life Technologies). Twenty-four hours post-transfection, the cells were split 1:2 in 10 cm plates, and the day after were transfected again with 600 pmol (single knockdown) or 1200 pmol (double knockdown) siRNA. For triple knockdown, 400 pmol of the single siRNAs were used (1200 pmol in total). For library preparation and verification, 1×10^6 HeLa cells were cultured overnight in 10 cm dishes and transiently transfected with 600 pmol total siRNA following the protocol for RNAiMAX (Life Technologies). Twenty-four hours after transfection, cells were split 1:8 in 10 cm dishes. Twenty-four hours later, cells again were transfected with 600 pmol (300 pmol each for double knockdowns) siRNA and incubated for 48 h. Stable cell lines were reverse transfected using 2.5 µL Lipofectamine RNAiMAX and 60 pmol siRNA per 2×10^5 HeLa cells. The following siRNA target sequences were used for luciferase, 5′-CGTACGCGGAATACTTCGA-3′; for XRN1, 5′-AGATGAACTTACCGTAGAA-3′; for SMG6, 5′-GGGTCACA GTGCTGAAGTA-3′; for UPF1, 5′-GATGCAGTTCCGCTCCA TT-3′; for CNOT1, 5′-GGAACUUGUUUGAAGAAUA-3′; and for SMG7, 5′-CGATTTGGAATACGCTTTA-3′.

Plasmid transfections

HeLa cells were split to six-well plates the day after siRNA transfection and transfected by calcium phosphate precipitation with 0.5 µg of a mVenus expression plasmid, 2 µg control plasmid (LacZ-4H), and 1.5 µg plasmid encoding for TPI-PTC160 reporter mRNA. For long 3′ UTR reporters, 3 µg reporter and 0.75 µg control plasmid (β-globin) were transfected. For rescue assays, 1 µg of Flag-tagged expression plasmid was included in the transfection mix.

RNA extraction and Northern blotting

Total RNA was extracted with peqGOLD TriFast (Peqlab), resolved on a 1% agarose and 0.4 M formaldehyde gel using a tricine–triethanolamine buffer system and analyzed by Northern blotting. pSP65-globin plasmid was linearized with BamHI and used as template for in vitro transcribed [α-^{32}P]-GTP body-labeled RNA probes, which were used for the detection of all reporter and control RNA. 7SL endogenous RNA was detected using a 5′-^{32}P-labeled oligonucleotide (5′-TGCTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3′). Signals were scanned using a Typhoon FLA 7000 (GE Healthcare). For time-course assays, the stable HeLa Flp-In T-REx cells were treated with 5 µg/mL actinomycin D for the indicated time prior to harvesting.

Immunoblot analysis and antibodies

SDS-polyacrylamide gel electrophoresis and immunoblot analysis was performed using protein samples derived from TriFast extractions. The antibodies against tubulin (T6074) and Flag (F7425) were from Sigma-Aldrich, the antibody against SMG6 (ab87539) was from Abcam, the antibodies against XRN1 (A300-443A) and SMG7 (A302-170A) were from Bethyl, and the antibody against UPF1 was kindly provided by Jens Lykke-Andersen.

3′ Fragment library preparation

siRNA-treated cells were used for poly(A)$^+$ RNA isolation, utilizing the magnetic mRNA Isolation Kit (New England Biolabs). Two hundred picomoles of 5′ RNA linker, containing a recognition site for EcoP15I (5′ linker_EcoP15I), was ligated to 3 µg poly(A)$^+$ RNA using T4 RNA Ligase I. After DNase I digest and RNA purification (RNA Clean & Concentrator Kit, Zymo Research), RT-PCR was performed using a linker oligo(dT)$_{18}$V (Harigaya and Parker 2012). The cDNA was PCR amplified for nine cycles, digested with EcoP15I overnight, which cleaves 25–27 nt downstream from the recognition site, then gel-purified and ligated to a 3′ dsDNA adaptor. After column purification of the ligation reaction, the ligated material was PCR-amplified for subsequent Illumina sequencing, using P3 and P5 primers and Accuprime Taq Polymerase (Life Technologies). For oligonucleotide sequences see Supplemental Table S2.

Library processing

Single-end sequencing reads had their barcodes and adapter trimmed with fastx_clipper from the fastx_toolkit (v. 0.0.13) before being mapped to the human genome (hg19) using TopHat (v. 2.0.11) and known splice junctions from ENSEMBL annotated transcripts using the following settings: -g 1 –p 8 –library-type fr-secondstrand. Sorted bam files were converted to bed files with the bedtools bamtobed command (v. 2.16.2), before single-nucleotide resolution cleavage sites were determined using awk (v. 3.1.5) and sed (v. 4.1.5). Specifically, based on the protocols design, the site of cleavage was considered the nucleotide that was antisense to the nucleotide immediately following the 3′ end of the read. For evaluation of replicate consistency, we initially applied a peak-finding algorithm that identified clusters of degradation sites with significant enrichment of degradation events relative to the local environment (König et al. 2010; Zarnack et al. 2013). We used all replicates from all conditions for this analysis, and a flank size of 15 nt on either side to define significant clusters with FDR <5%. This yielded a total of 299,442 degradation sites. We then evaluated the number of degradation events mapping to each degradation site for each replicate in order to compare how individual libraries contributed to each defined degradation cluster.

RNA maps

Maps were created through intersection of single-nucleotide resolution cleavage files with indicated features using the bedtools intersect command. Features were determined based on their positioning in annotation files used for library mapping with the exception of uORFs (see below). Counts at individual loci were normalized to the corresponding libraries total read count. Similarly, for each junction type, counts were divided by the maximum value across all libraries to allow for the comparison across different features. Maps were then created from normalized values around indicated features using R (v. 3.1.1). Enrichment analysis of the different features of the transcriptome was performed using the Homer annotatePeaks.pl command (Heinz et al. 2010), utilizing the same annotation file as that used for mapping.

Separation into 5′ UTR, 3′ UTR, and ORF categories was achieved by first applying a 100-nt sliding window across the transcriptome in order to identify windows most enriched between the siRNA XRN1 and siXRN1/SMG6 conditions. To then assign genes to a given feature, we applied a strict threshold of >50 reads in the window for the XRN1 siRNA condition, a fold-difference of >10 between the XRN1 siRNA and XRN1/SMG6 siRNA conditions, and, using a ranked list of *z*-scores, we took the first occurrence of each gene and assigned it to the feature in which that window overlapped. Thereby windows were assigned based on their dominant overlap with genomic features. In cases where windows overlapped features equally, the hierarchical order for assignment is "ncRNA; ORF; 3′ UTR; 5′ UTR; intron; telomere; intergenic." To then determine proximity of cleavage sites to annotated stop codons, we intersected the coordinates of all annotated stop codons with 50-nt windows surrounding the highest peak in each SMG6-sensitive window. The highest peak was used under the assumption that it represented the dominant cleavage site within the indicated window.

Coordinates of experimentally verified uORFs were taken from ribosome footprinting experiments (Fritsch et al. 2012). This included 2107 unique uORF start and end coordinate pairs, although starts and ends could be shared between different uORFs for the same gene. Using these sites, we determined 50-nt windows surrounding all unique uORF end coordinates and intersected these with single-nucleotide resolution cleavage files. We removed all windows that had zero counts in the XRN1-sensitive condition prior to drawing maps.

PCR-based analysis of potential NMD targets

Two hundred picomoles of a 5′ RNA linker (5′ RNA linker_VB2) were ligated to 3 µg of poly(A)$^+$ RNA or 10 µg of total RNA, respectively. Reverse transcription was performed using an oligo $(dT)_{20}$VNN primer or random hexamers. To functionally analyze NMD targets identified from the high-throughput sequencing data, PCR was carried out using a primer directed against the 5′ linker (VB2_linker_se) and a target-specific antisense primer. Total amounts of target RNA were determined using a gene-specific sense primer located downstream from the estimated endocleavage site. For all targets, PCR fragments were verified via Sanger sequencing. All used oligonucleotides are listed in Supplemental Table S2.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Heidi Thelen and Juliane Hancke for technical assistance, the Leptin and Schnetz laboratory for sharing equipment, and members of the Gehring laboratory for useful discussions. We are grateful to Jens Lykke-Andersen for antibodies against UPF1. We thank Tomaž Curk and Nejc Haberman for bioinformatics advice and assistance. This research was funded by a grant from the Deutsche Forschungsgemeinschaft (GE2014/4-1) to N.H.G.

Author contributions: N.H.G., F.O., and V.B. designed the study. V.B. and F.O. performed the experiments. C.S. and J.U. carried out high-throughput sequencing. C.S. and J.U. analyzed and interpreted the resulting data, V.B., F.O., and N.H.G. analyzed and interpreted all other data. All authors wrote the manuscript, discussed the results and implications, and commented on the manuscript at all stages.

Received October 19, 2016; accepted April 17, 2017.

REFERENCES

Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. 2008. Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Curr Biol* **18:** 758–762.

Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, Jacobson A. 2004. A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432:** 112–118.

Boehm V, Haberman N, Ottens F, Ule J, Gehring NH. 2014. 3′ UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. *Cell Rep* **9:** 555–568.

Bühler M, Steiner S, Mohn F, Paillusson A, Mühlemann O. 2006. EJC-independent degradation of nonsense immunoglobulin-µ mRNA depends on 3′ UTR length. *Nat Struct Mol Biol* **13:** 462–464.

Chakrabarti S, Bonneau F, Schüssler S, Eppinger E, Conti E. 2014. Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. *Nucleic Acids Res* **42:** 9447–9460.

Chang YF, Imam JS, Wilkinson MF. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76:** 51–74.

Cho H, Kim KM, Kim YK. 2009. Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. *Mol Cell* **33:** 75–86.

Eberle AB, Stalder L, Mathys H, Orozco RZ, Mühlemann O. 2008. Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region. *PLoS Biol* **6:** e92.

Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol* **16:** 49–55.

Fatscher T, Boehm V, Weiche B, Gehring NH. 2014. The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. *RNA* **20:** 1579–1592.

Fiorini F, Boudvillain M, Le Hir H. 2013. Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains. *Nucleic Acids Res* **41:** 2404–2415.

Fritsch C, Herrmann A, Nothnagel M, Szafranski K, Huse K, Schumann F, Schreiber S, Platzer M, Krawczak M, Hampe J, et al. 2012. Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Res* **22:** 2208–2218.

Fukuhara N, Ebert J, Unterholzner L, Lindner D, Izaurralde E, Conti E. 2005. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol Cell* **17:** 537–547.

Gatfield D, Izaurralde E. 2004. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* **429:** 575–578.

Gehring NH, Lamprinaki S, Hentze MW, Kulozik AE. 2009. The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS Biol* **7:** e1000120.

German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, et al. 2008. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol* **26:** 941–946.

Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J* **25:** 5117–5125.

Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR. 2008. A link between RNA metabolism and silencing affecting *Arabidopsis* development. *Dev Cell* **14:** 854–866.

Harigaya Y, Parker R. 2012. Global analysis of mRNA decay intermediates in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* **109:** 11764–11769.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38:** 576–589.

Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. 2004. Nonsense-mediated decay approaches the clinic. *Nat Genet* **36:** 801–808.

Huntzinger E, Kashima I, Fauser M, Saulière J, Izaurralde E. 2008. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14:** 2609–2617.

Jonas S, Weichenrieder O, Izaurralde E. 2013. An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes Dev* **27:** 211–225.

Joncourt R, Eberle AB, Rufener SC, Mühlemann O. 2014. Eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay by two genetically separable mechanisms. *PLoS One* **9:** e104391.

Karam R, Wengrod J, Gardner LB, Wilkinson MF. 2013. Regulation of nonsense-mediated mRNA decay: implications for physiology and disease. *Biochim Biophys Acta* **1829:** 624–633.

Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S. 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev* **20:** 355–367.

Kervestin S, Jacobson A. 2012. NMD: a multifaceted response to premature translational termination. *Nat Rev Mol Cell Biol* **13:** 700–712.

König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* **17:** 909–915.

Kurosaki T, Li W, Hoque M, Popp MW, Ermolenko DN, Tian B, Maquat LE. 2014. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev* **28:** 1900–1916.

Loh B, Jonas S, Izaurralde E. 2013. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev* **27:** 2125–2138.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* **16:** 665–677.

Lykke-Andersen S, Chen Y, Ardal BR, Lilje B, Waage J, Sandelin A, Jensen TH. 2014. Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes Dev* **28:** 2498–2517.

Mühlemann O, Lykke-Andersen J. 2010. How and where are nonsense mRNAs degraded in mammalian cells? *RNA Biol* **7:** 28–32.

Muhlrad D, Parker R. 1999. Aberrant mRNAs with extended 3′ UTRs are substrates for rapid degradation by mRNA surveillance. *RNA* **5:** 1299–1307.

Nagy E, Maquat LE. 1998. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* **23:** 198–199.

Neu-Yilik G, Gehring NH, Thermann R, Frede U, Hentze MW, Kulozik AE. 2001. Splicing and 3′ end formation in the definition of nonsense-mediated decay-competent human β-globin mRNPs. *EMBO J* **20:** 532–540.

Nguyen LS, Wilkinson MF, Gecz J. 2014. Nonsense-mediated mRNA decay: inter-individual variability and human disease. *Neurosci Biobehav Rev* **46**(Pt2): 175–186.

Nicholson P, Yepiskoposyan H, Metze S, Zamudio Orozco R, Kleinschmidt N, Mühlemann O. 2010. Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell Mol Life Sci* **67:** 677–700.

Nicholson P, Josi C, Kurosawa H, Yamashita A, Mühlemann O. 2014. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Res* **42:** 9217–9235.

Okada-Katsuhata Y, Yamashita A, Kutsuzawa K, Izumi N, Hirahara F, Ohno S. 2012. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res* **40:** 1251–1266.

Schmidt SA, Foley PL, Jeong DH, Rymarquis LA, Doyle F, Tenenbaum SA, Belasco JG, Green PJ. 2015. Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. *Nucleic Acids Res* **43:** 309–323.

Singh G, Rebbapragada I, Lykke-Andersen J. 2008. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS Biol* **6:** e111.

Stockklausner C, Breit S, Neu-Yilik G, Echner N, Hentze MW, Kulozik AE, Gehring NH. 2006. The uORF-containing thrombopoietin mRNA escapes nonsense-mediated decay (NMD). *Nucleic Acids Res* **34:** 2355–2363.

Thermann R, Neu-Yilik G, Deters A, Frede U, Wehr K, Hagemeier C, Hentze MW, Kulozik AE. 1998. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J* **17:** 3484–3494.

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

Zarnack K, König J, Tajnik M, Martincorena I, Eustermann S, Stévant I, Reyes A, Anders S, Luscombe NM, Ule J. 2013. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152:** 453–466.