

SUZ domain-containing proteins have multiple effects on nonsense-mediated decay target transcripts

Received for publication, January 31, 2023, and in revised form, July 13, 2023 Published, Papers in Press, July 26, 2023,
<https://doi.org/10.1016/j.jbc.2023.105095>

Mathias Halbout^{1,2}, Marina Bury^{1,2}, Aoife Hanet¹, Isabelle Gerin^{1,2}, Julie Graff^{1,2}, Theodore Killian³ ,
Laurent Gatto³ , Didier Vertommen⁴ , and Guido T. Bommer^{1,2,*}

From the ¹Department of Physiological Chemistry, ²WELBIO, ³Computational Biology Laboratory, and ⁴Protein Phosphorylation Unit, de Duve Institute, UCLouvain, Brussels, Belgium

Reviewed by members of the JBC Editorial Board. Edited by Karin Musier-Forsyth

Many transcripts are targeted by nonsense-mediated decay (NMD), leading to their degradation and the inhibition of their translation. We found that the protein SUZ domain-containing protein 1 (SZRD1) interacts with the key NMD factor up-frameshift 1. When recruited to NMD-sensitive reporter gene transcripts, SZRD1 increased protein production, at least in part, by relieving translational inhibition. The conserved SUZ domain in SZRD1 was required for this effect. The SUZ domain is present in only three other human proteins besides SZRD1: R3H domain-containing protein 1 and 2 (R3HDM1, R3HDM2) and cAMP-regulated phosphoprotein 21 (ARPP21). We found that ARPP21, similarly to SZRD1, can increase protein production from NMD-sensitive reporter transcripts in an SUZ domain-dependent manner. This indicated that the SUZ domain-containing proteins could prevent translational inhibition of transcripts targeted by NMD. Consistent with the idea that SZRD1 mainly prevents translational inhibition, we did not observe a systematic decrease in the abundance of NMD targets when we knocked down SZRD1. Surprisingly, knockdown of SZRD1 in two different cell lines led to reduced levels of the NMD component UPF3B, which was accompanied by increased levels in a subset of NMD targets. This suggests that SZRD1 is required to maintain normal UPF3B levels and indicates that the effect of SZRD1 on NMD targets is not limited to a relief from translational inhibition. Overall, our study reveals that human SUZ domain-containing proteins play a complex role in regulating protein output from transcripts targeted by NMD.

Nonsense-mediated decay (NMD) was initially discovered in yeast as a pathway leading to the degradation of transcripts with premature stop codons (1, 2). Later, similar processes were described in mammalian systems (3, 4), where it became apparent that this pathway degrades not only defective transcripts with premature stop codons but also a large number of physiological transcripts with particular architectural characteristics such as a long 3'UTR or upstream ORFs (5–8). This means that NMD not only serves as a quality control function but also is a way to regulate the expression of a subset of genes (9, 10).

Initiation and execution of NMD relies on a complex interplay between structural characteristics of the transcripts and multiple proteins. Recognition of premature stop codons relies in part on exon junction complexes (EJCs) containing the core components eukaryotic translation initiation factor 4A3 (eIF4A3), RBM8A, MLN51, and MAGOH, as well as several peripheral proteins such as up-frameshift 3A (UPF3A), UPF3B, and RNPS1 (11). These protein complexes are deposited concomitantly with splicing and are normally removed during the first round of translation (12). Yet, when premature stop codons are present more than 50 nucleotides upstream of an exon–exon junction, the EJC is not removed during translation (13). In this pathway, the eukaryotic release factor 3 (eRF3) within the terminated ribosome recruits the helicase UPF1 leading to the assembly of the so-called SURF complex, consisting of the kinase SMG1 (suppressor with morphogenetic effect on genitalia protein 1), UPF1, and the eukaryotic ribosome release factors eRF3 and eRF1 (14). Phosphorylation of UPF1 represents a central event in NMD and converts the SURF complex into the decay-inducing complex (DECID) (14). This transition can be facilitated by the helicase DHX34, which binds to both SMG1 and UPF1 (15, 16). This binding leads to conformational changes that allow the recruitment of UPF2 and additional phosphorylation of UPF1 by SMG1 (16). Close vicinity between a terminated ribosome and an EJC facilitates the UPF1–UPF2 interaction, since UPF2 binds to the EJC component UPF3B (17). However, it has been shown that UPF2 and UPF3B can also interact directly with eRF3 independently of the presence of an EJC (18). This allows initiation of NMD in the absence of an EJC, which can be observed on transcripts with very long 3'UTR (5, 6), albeit with a lower efficiency of target mRNA degradation (19, 20). In this constellation, the terminated ribosome is localized at a distance from the polyA tail where PABPC1 (polyA-binding protein cytoplasmic 1) binds. The resulting delayed removal of ribosomes seems to contribute to the initiation of NMD. Yet, reality might be more complicated than this, since recent work revealed that ribosome stalling at termination codons is not a prerequisite for NMD (21). Furthermore, NMD does not only occur during the first round of translation (22, 23). In fact, recent data indicate that the presence of architectural characteristics (e.g., long 3'UTRs or

* For correspondence: Guido T. Bommer, guido.bommer@uclouvain.be.

premature termination codons) increases the probability of degradation by NMD but that usually several rounds of translation are required (23).

Degradation of transcripts is the main consequence of NMD and depends on the recruitment of the endonuclease SMG6 and a complex of SMG5 and SMG7, which recruits the CCR4–NOT complex (10). However, several lines of evidence show that NMD can also lead to an inhibition of mRNA translation. Phosphorylated UPF1 has been found to interact with components of the eukaryotic translation initiation factor eIF3 (likely *via* eIF3a), thereby preventing the assembly of the 80S ribosome (24). Furthermore, the CCR4–NOT complex is not only a well-known deadenylase complex but has also been described as a translational repressor (25). Deadenylation of transcript by CCR4–NOT exonuclease leads to the 3′-5′ degradation of mRNA; thus, leading to a decrease in protein production. In addition, it has been demonstrated that the recruitment of the CCR4–NOT complex to transcripts can lead to targeted translational repression independently of mRNA degradation. In fact, several laboratories have demonstrated that miRNA-induced translation repression of mRNAs occurs *via* the recruitment of the CCR4–NOT complex (26–28). Of note, there are likely several additional mechanisms that might contribute to translational repression of NMD targets, for example, through the cap-binding protein 4E-T (29, 30). Moreover, NMD targets show a lower translational efficiency in comparison to non-NMD targets likely because of lower translational initiation and elongation efficiency (31). Yet, it is unclear whether this is causing NMD or whether this is a consequence of the occurrence of NMD.

Beyond UPF1, the requirement of individual NMD factors is much less clear. For example, transcript destabilization triggered by some EJC factors is dependent on UPF2, whereas destabilization by other EJC factors requires UPF3B (32, 33). At present, the mechanisms underlying these preferences is still unclear. Furthermore, we are only beginning to understand how NMD is regulated. Most of what is known about the regulation of NMD concerns the regulation of transcript degradation. Recruitment of the RNA-binding proteins PTBP1 or hnRNPL to the 3′UTR can protect transcripts from degradation *via* NMD (34–36). Curiously, there is little overlap between transcripts protected from NMD by hnRNPL and PTBP1, suggesting that these proteins protect a distinct subset of transcripts from degradation by NMD. Beyond this, additional *cis*-elements in the first 100 base pairs of long 3′UTRs have been shown to confer resistance to degradation by NMD (37). NMD also seems to be autoregulated since transcripts coding for UPF1, UPF3B, and SMG1 have all been described as NMD targets (38–40). Last but not least, while the initiation of NMD requires translation of transcripts, some factors that facilitate translation can counteract NMD. For example, the polyA-binding protein PABPC1 has been reported to block NMD activity *via* its interaction with eIF4G and eRF3 (6, 24, 41–46).

In the present study, we are investigating the functional and physical interaction between the protein SZRD1 and the NMD

machinery. Very little is known about the function of SZRD1. When overexpressed, SZRD1 can strongly activate mitogen-activated protein kinase–dependent reporter gene constructs (47), modulate a wide range of signaling pathways (48), and might have antineoplastic effects (49). However, how SZRD1 might exert these functions is unclear.

SZRD1 contains two conserved domains named SUZ and SUZ-C. These domains were identified based on their similarity to domains in SZY-20 (suppressor of *zyg-1* protein 20), a 530 amino acid *Caenorhabditis elegans* protein that is considerably larger than SZRD1 with its 153 amino acids (50). Based on the interaction with polyuridine-coated agarose beads (50), it was postulated that SZY-20 may bind to RNA, but this interaction was not further characterized. Functionally, SZY-20 limits centrosome size in Plk4-deficient worms (50).

The C-terminal SUZ-C domain consists of 30 amino acids. Despite the similarity in name to the SUZ domain, the two domains are completely unrelated in sequence. In humans, SUZ-C domains are present in the proteins LARP6, GEMIN7, and CSDE1 (51–53) that all interact with STRAP (serine/threonine kinase receptor–associated protein).

In humans, the SUZ domain contains approximately 70 amino acids and is only present in ARPP21, R3HDM1 and R3HDM2 (R3H domain–containing proteins 1 and 2), and SZRD1. ARPP21, R3HDM1, and R3HDM2 also contain an R3H domain, which is found in many proteins and has been described to mediate the interaction with RNA molecules in a nonsequence-specific manner (54, 55). Interestingly, ARPP21 has been shown to control dendritic branching, at least in part, *via* a post-transcriptional regulation of miR-128 target transcripts (56). Furthermore, while this article was in preparation, two studies indicated that orthologs of ARPP21, R3HDM1, and R3HDM2 might influence NMD. First, Kelliher *et al.* (57) reported that PRD2, the single *Neurospora crassa* ortholog, inhibits NMD of the transcript coding for casein kinase I. Second, the *Saccharomyces cerevisiae* ortholog Rbs1 was shown to interact with Upf1 and antagonize the degradation of specific mRNAs to which it is recruited (58).

To understand SZRD1 function, we first identified interaction partners of SZRD1 and determined the relative contribution of its domains to protein–protein interactions. We found that SZRD1 interacts with UPF1 in cells and that this interaction requires part of the SUZ domain. We then used reporter assay experiments to test the effect of SZRD1 on NMD activity. This revealed that the recruitment of SZRD1 to transcripts increases protein output from transcripts recognized by the NMD machinery. This effect was SUZ domain dependent and also observed with the protein ARPP21. Interestingly, the SUZ domain–dependent increase of protein production was accompanied by a much weaker (SZRD1) or no (ARPP21) increase in mRNA levels. This suggests that protein production was increased predominantly because of a relief from translational inhibition that occurs during NMD. To assess the effects of SZRD1 on endogenous NMD targets, we modulated SZRD1 levels and analyzed transcriptional changes. Consistent with a predominant effect of SZRD1 on protein translation, we did not observe decreases in the

abundance of NMD target transcripts, indicating that SZRD1 is not a general inhibitor of NMD function. In fact, we actually observed that knockdown of SZRD1 leads to a decrease in the abundance of some NMD components (*i.e.*, UPF3B) and a subtle increase in the abundance of a subset of NMD targets. This indicates that SZRD1 is required to maintain normal UPF3B levels. While the molecular mechanism for these observations remains to be elucidated, our work provides first insights into a network of SUZ domain-containing proteins that can modulate NMD function.

Results

SZRD1 protein isoforms interact with UPF1 and STRAP

The *SZRD1* gene codes for transcripts with up to four exons. Alternative splicing events can lead to four different transcripts that are predicted to produce four different protein isoforms (Fig. 1, A and B). Inclusion or exclusion of exon 2 results into long and short isoforms of SZRD1. The long and short isoforms have start codons in different ORFs converging in the same ORF in exons 3 and 4. Thus, these isoforms differ in their N-terminal part, encoded by exons 1 and 2, and are identical in the C-terminal part, encoded by exons 3 and 4. The long isoform is conserved in a wide range of organisms down to *C. elegans*, whereas the short form is restricted to a limited number of vertebrates (Fig. S1, A and B). Additional complexity of SZRD1 transcripts is caused by the inclusion of three additional nucleotides (CAG) on the 5' end of exon 3 during splicing in about half of the transcripts (59). This leads to the insertion or exclusion of one amino acid: an arginine in the short form of SZRD1 and a serine in the long form of SZRD1 (Fig. 1B).

All four transcripts are present at comparable levels in human tissues (59). Yet, it was unclear whether these isoforms were translated into proteins. This particularly concerns the short isoforms since productive translation of these isoforms is predicted to start with a downstream start codon, whereas usage of the first start codon of these transcripts would be expected to lead to a degradation by NMD. We therefore assessed endogenous SZRD1 protein levels with a polyclonal antibody expected to recognize both long and short isoforms. We observed two distinct bands that migrated at an apparent molecular weight of 20 and 23 kDa (Fig. 1C), close to 15 and 17 kDa predicted for the short and long isoforms, respectively. Next, we wanted to test whether these two bands were the consequence of the inclusion or not of exon 2. To this end, we knocked down either all SZRD1 isoforms using two different shRNAs targeting the common exon 4 (Fig. 1C, #1 and #2) or only the long isoform using four different shRNAs targeting exon 2 of SZRD1 (Fig. 1C, #3 to #6) in human embryonic kidney 293 (HEK293) cells. We observed that shRNAs #1 and #2 strongly decreased both bands, confirming that both bands are SZRD1 proteins. In contrast, shRNAs targeting exon 2 only decreased the band of higher predicted molecular weight with varying efficiency, whereas the lower bands remained unaffected. Of note, none of the shRNAs affected levels of the loading control β -actin. We conclude that the upper band

corresponds to long SZRD1 protein isoforms produced from transcripts containing exon 2, whereas the lower band corresponds to short SZRD1 protein isoforms that are produced from transcripts that do not contain exon 2.

Very little is known about the function of SZRD1 and its isoforms. To glean insights into the function of SZRD1 proteins, we first decided to identify interacting protein partners. We overexpressed SZRD1 proteins with a C-terminal SFB tag consisting of the S-peptide, a FLAG tag, and the streptavidin-binding peptide (60, 61). SZRD1 and interacting proteins were purified in a tandem affinity purification consisting of binding to streptavidin-coated Sepharose beads, eluting with biotin and binding of eluted proteins to S-protein-coated agarose beads. Bands were excised, and putative interacting proteins were identified by mass spectrometry. Subsequently, we validated the interaction with the proteins with the highest spectral count by performing streptavidin-bead pulldowns followed by Western blot analysis (Fig. 1D). We confirmed that UPF1, STRAP, PABPC1 (polyadenylate-binding protein), and C1QBP (complement component 1Q subcomponent-binding protein) copurified with SFB-tagged SZRD1. These proteins were retained on streptavidin beads in cell lines expressing SFB-tagged SZRD1 isoforms but were absent in pulldowns performed in cell lines expressing an empty vector control (Fig. 1D). Expression levels of different SFB-tagged SZRD1 proteins were similar between cell lines facilitating the comparison of interaction efficiency between SZRD1 isoforms. UPF1 and PABPC1 interacted more readily with the short form of SZRD1, and C1QBP interacted mainly with the short form of SZRD1 containing the arginine insertion. Of note, we consistently observed that overexpression of SZRD1 led to an increase in UPF1 protein levels (Fig. 1, D and E), which was further investigated later in this work.

Taken together, at this stage, we concluded that long and short SZRD1 proteins can interact with UPF1, STRAP, and PABPC1, and that C1QBP interacts only with the short forms of SZRD1. In addition, the incorporation or not of the extra CAG sequence at the transcript level can affect the strength of some interactions. Of note, these interaction partners have known functions in RNA biology: UPF1 is best known for its central role in NMD among other functions (2), STRAP is implicated in different aspects of RNA metabolism (51–53), PABPC1 interacts with the polyA tail of mRNAs (62), and C1QBP is a multifunctional protein with a role in RNA splicing and homologous recombination (63–66).

To investigate whether the interaction of SZRD1 with these proteins is dependent on the presence of intact RNA, we purified SFB-tagged SZRD1 proteins from cells and treated (or not) the cell lysates with a mixture of RNase A and RNase T1. To simplify the experimental setup, we focused on the long and short isoforms of SZRD1 carrying the serine and arginine insertion, respectively (Fig. 1B, “+CAG”). In the presence of RNases, PABPC1 and C1QBP were copurified with SZRD1 to a weaker extent (Fig. 1E + *versus* – RNase A/T1). In contrast, the amounts of STRAP remained the same, and the amount of interacting UPF1 protein was increased upon RNase A/T1 treatment (Fig. 1E). UPF1 is an RNA helicase and binds to the

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

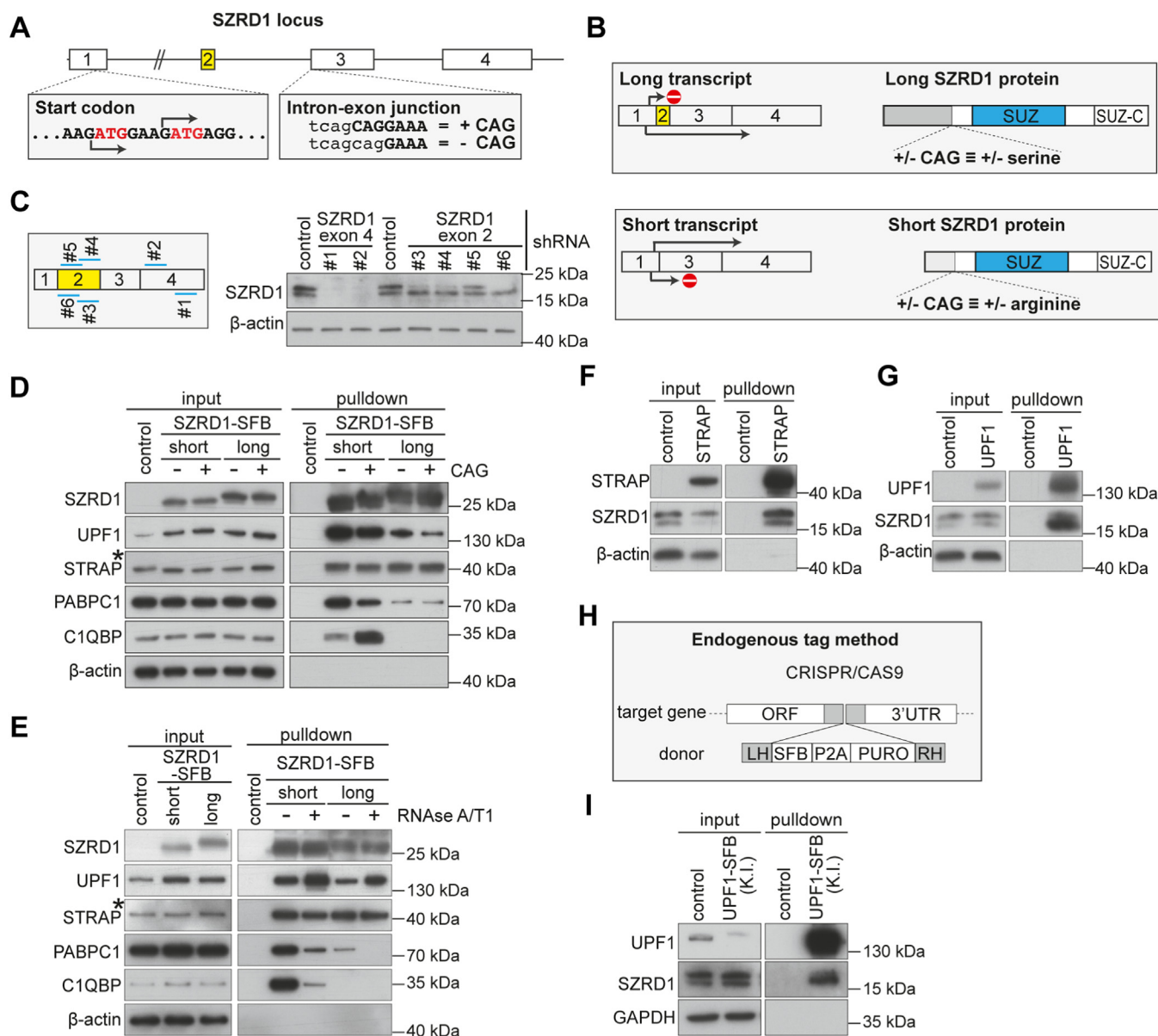


Figure 1. SZRD1 isoforms interact with UPF1 and STRAP. A, schematic representation of the *SZRD1* gene locus and the resulting transcripts. Inclusion or exclusion of exon 2 (in yellow) lead to long and short isoforms, respectively. B, schematic representation of *SZRD1* transcripts and the resulting proteins. C, Western blot analysis with antibodies predicted to recognize short and long *SZRD1* isoforms as well as β -actin was performed in HEK293 cells expressing shRNAs specifically targeting the long (exon 2; “#3,” “#4,” “#5,” and “#6”) or all *SZRD1* isoforms (exon 4, “#1” and “#2”). Samples labeled with “control” express a nonsilencing control shRNA. D, lysates from HEK293 cells overexpressing the indicated *SZRD1* proteins (“SZRD1”) with a C-terminal SFB tag or an empty vector (“control”) were subjected to affinity purification and analyzed by Western blot using the indicated antibodies. Input samples correspond to 10% of the amount used in the pull-down. Exposure time was identical for input and pull-down samples except for Western blots analysing STRAP, where pull-down exposure was five times shorter (*). E, lysates from HEK293 cells overexpressing *SZRD1* protein isoforms (containing the additional arginine or lysine residue) with a C-terminal SFB tag identical to D were subjected to affinity purification in the presence (“+”) or the absence (“-”) of RNase A/T1. Western blot analysis and presentation are identical to D. F and G, lysates from HEK293 cells overexpressing UPF1 with an N-terminal SFB tag (“UPF1”) (F), STRAP with an N-terminal SFB tag (“STRAP”) (G) or an empty vector (“control”) were subjected to affinity purification followed by Western blotting as described in D. H, schematic representation of the selection cassette used to insert an SFB tag before the stop codon of the endogenous coding sequence of UPF1. I, protein lysates from HEK293 cells carrying a C-terminal SFB tag knocked into the endogenous loci of UPF1 or an unmodified control cell line (“ctrl”) were subjected to affinity purification followed by Western blot analysis with the indicated antibodies. Input corresponds to 10% of the amount used in the pull-down. Exposure time was identical for input and pull-down samples except for Western blots analysing UPF1, where pull-down exposure was five times shorter (*). HEK293, human embryonic kidney 293 cell line; STRAP, serine/threonine kinase receptor-associated protein; SZRD1, SUZ domain-containing protein 1; UPF1, up-frameshift 1.

3'UTR of many mRNAs (67, 68). Hence, the increased interaction of SZRD1 with UPF1 upon RNase treatment might be caused by a liberation of mRNA-bound UPF1, which can then bind to SZRD1. Overall, these experiments indicated that the interactions of SZRD1 with PABPC1 and C1QBP might be

indirect *via* an RNA scaffold, whereas the interactions with UPF1 and STRAP are maintained even when cellular RNA is degraded.

Our initial experiments demonstrated that epitope-tagged SZRD1 interacts with endogenous UPF1 and STRAP. To

complement these data, we also performed the reverse experiment where we overexpressed UPF1 and STRAP with an N-terminal SFB tag (Fig. 1, F and G). In these experiments, we recovered endogenous SZRD1 only in cells that express epitope-tagged STRAP (Fig. 1F) or UPF1 (Fig. 1G) but not in cell lines transduced with an empty vector ("control"). In all these experiments, UPF1 was preferentially recovered with the short isoform, whereas STRAP was pulled down with both isoforms (Fig. 1, D–F). We next investigated whether both proteins interact at physiological levels. Unfortunately, we encountered two problems. First, we did not succeed to immunoprecipitate endogenous SZRD1. Second, when we immunoprecipitated UPF1, the immunoglobulin bands overlapped with the SZRD1 band. To avoid this problem, we inserted a sequence coding for C-terminal SFB tag into the endogenous locus of UPF1 (Fig. 1H). Affinity purification of cellular lysates from these cells using streptavidin-coated beads clearly revealed copurifying SZRD1, whereas no signal was detectable in wildtype cells (Fig. 1I). This indicated that SZRD1 interacts with UPF1 when expressed at physiologic levels.

The SUZ domain is required for the interaction with UPF1, and the SUZ-C domain is required for the interaction with STRAP

Next, we investigated to what extent the conserved SUZ and SUZ-C domains in SZRD1 are required for its protein–protein interactions. The SUZ-C domain is approximately 30 amino acids long, and, besides SZRD1, the human genome codes for three other proteins with this domain (GEMIN7, CSDE1, and LARP6) (Fig. 2A). Interestingly, all three proteins have been described to interact with STRAP (51–53). Hence, we hypothesized that the SUZ-C domain would be required for the interaction between SZRD1 and STRAP. To test this, we overexpressed SFB-tagged mutant SZRD1 proteins in HEK cells. We decided to either delete the entire SUZ-C domain ("ΔC") or to replace three conserved amino acids ("GPD") by three alanine residues ("AAA") (Fig. 2A). Affinity purifications with streptavidin beads were performed and analyzed by Western blot. This revealed that only wildtype SZRD1 copurified with STRAP, whereas the deletion of the SUZ-C domain and the indicated three amino acid substitution abolished the interaction with STRAP (Figs. 2C and S2A). Of note, mutations in the SUZ-C domain abolished the interaction with STRAP but did not perturb the interaction with UPF1 (Figs. 2C and S2A). These experiments demonstrate that the SUZ-C domain of SZRD1 is required for its interaction with STRAP in cells but is dispensable for the interaction with UPF1.

The SUZ domain comprises approximately 70 amino acids and has been suggested to be involved in RNA binding (50). Besides SZRD1, there are three other human proteins with an SUZ domain (ARPP21, R3HDM1, and R3HDM2; Fig. 2B). We decided to test whether the SUZ domain of SZRD1 was required for the interactions with STRAP or UPF1. To do this, we overexpressed SFB-tagged SZRD1 proteins carrying mutations of conserved amino acids in the SUZ domain. We generated four different mutations (Fig. 2B): two deletions of

four amino acids ("Δ1" and "Δ2") and two single amino acid substitutions ("Y→F" and "R→Q"). The mutations were introduced based on the sequence conservation among the SUZ domains in the human ARPP21, R3HDM2/2, and SZRD1 (Fig. 2B). We then analyzed whether these proteins interacted with UPF1 and STRAP using affinity purification with streptavidin beads followed by Western blot analysis. In this case, mutations in the SUZ domain did not affect the interaction with STRAP indicating that the SUZ domain of SZRD1 is dispensable for its interaction with STRAP (Figs. 2D and S2B). In contrast, deletion of the conserved sequence ILKR ("Δ1") abolished the interaction of the long and short SZRD1 isoforms with UPF1 (Figs. 2D and S2B). Mutations in the C-terminal part of the SUZ domain reduced the interaction of the long SZRD1 isoform, whereas the interaction with the short SZRD1 isoform was less affected (Fig. S2B). As already noted in Figure 1, D and E, we observed an increase in UPF1 protein levels (Fig. 2, C and D) when SZRD1 was overexpressed. In reverse, knockdown of SZRD1 (shown in Fig. 2F) led to a reduction in UPF1 protein levels, which was further investigated later. Interestingly, the introduction of mutations in the SUZ domain of SZRD1 abolished this increase in UPF1 protein levels (Figs. 2D and S2B), although only deletion of the ILKR motif (*i.e.*, mutation Δ1) abolished interaction with UPF1. Overall, our data indicated that the four residues ILKR in the SUZ domain are required for the interaction between SZRD1 and UPF1, and that SZRD1 *via* its SUZ domain might somehow affect UPF1 protein levels.

It had previously been reported that UPF1 and STRAP interact with each other (69). Given that both proteins also interacted with SZRD1 *via* separate distinct domains, we considered that SZRD1 may link UPF1 to STRAP (Fig. 2E, upper panel). To test this, we knocked down SZRD1 in cell lines expressing SFB-tagged UPF1 or STRAP. When we pulled down proteins with streptavidin beads, we copurified SFB-tagged STRAP with endogenous UPF1 and SFB-tagged UPF1 with endogenous STRAP (Fig. 2, F and G). However, this interaction was not affected by a reduction of SZRD1 levels in cell lines expressing SZRD1-specific shRNA.

We conclude that STRAP and UPF1 interact with SZRD1 *via* distinct domains in cells. The interaction of these proteins with SZRD1 is not required for the previously published interaction between STRAP and UPF1 (Fig. 2E, lower panel) (69).

SZRD1 increases protein levels from reporter gene constructs targeted by the NMD machinery

Next, we wondered whether SZRD1 could modulate NMD, a process where UPF1 plays a central role. To do this, we generated reporter constructs with six BoxB sites in the 3'UTR of transcripts similar to what has been described by Gehring *et al.* (70) (Fig. 3A). The presence of these sites in transcripts allows the recruitment of proteins carrying a λN-tag. As reporter gene, we used *Renilla* luciferase fused on its C terminus with the ORF as the human genomic β-globin gene sequence (Fig. 3A, "WT-BoxB-noARE") (71). To have a quantifiable readout of NMD in cells, we used a similar reporter gene

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

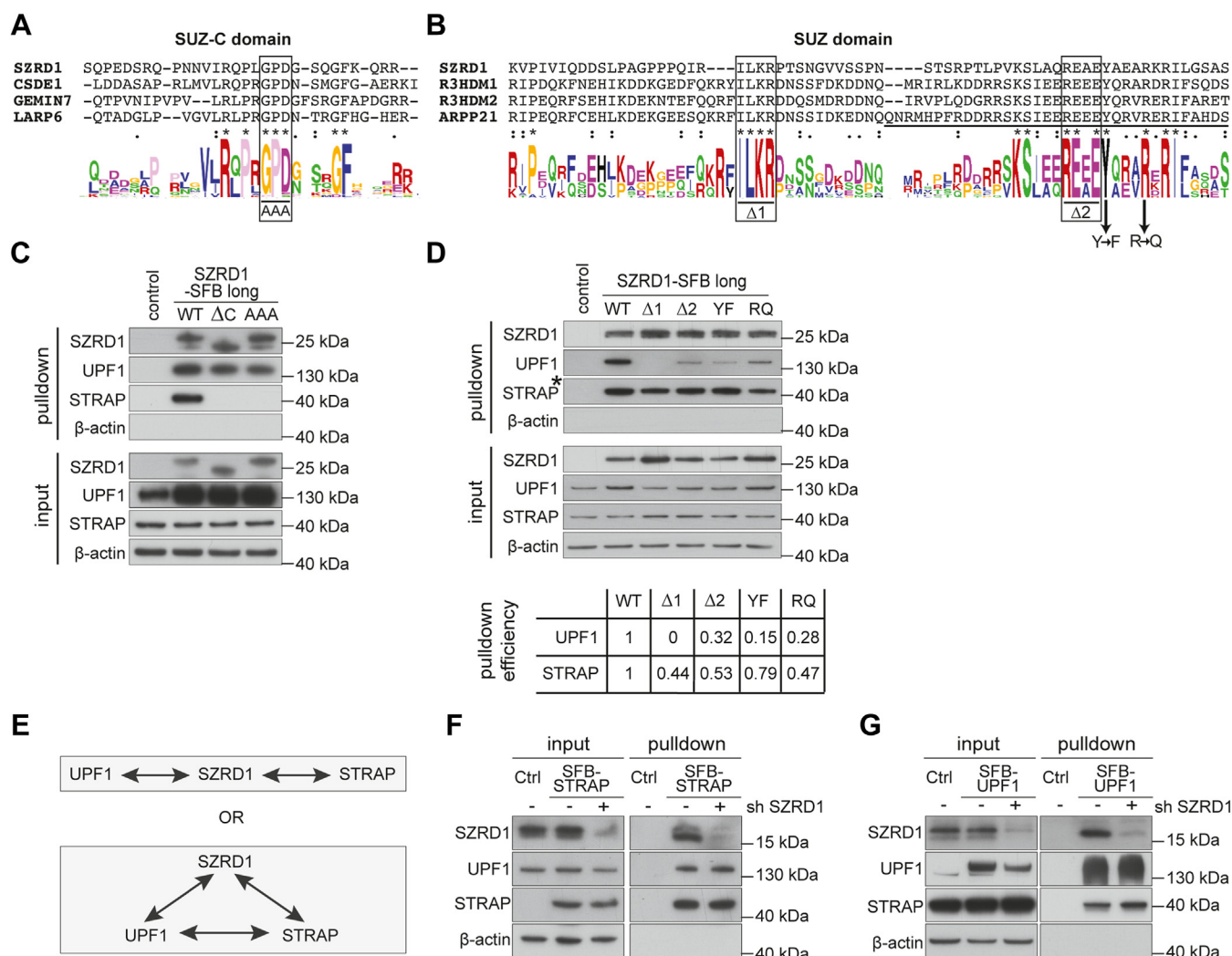


Figure 2. The SUZ domain is required for the interaction with UPF1, and the SUZ-C domain is required for the interaction with STRAP. A and B, alignment of human SUZ-C domain (A) and SUZ domain (B) sequences of the indicated proteins (top panel) and graphical representation of the conservation of these sequences across evolution (humans, mouse, chicken, and frog) (lower panel). Mutations and deletions (Δ) used in the other panels are indicated. C and D, lysates from HEK293 cells overexpressing the C-terminally SFB-tagged SZRD1 long isoform (carrying the mutations illustrated in A and B) were subjected to affinity purification with streptavidin beads followed by Western blot analysis. Input corresponds to 10% of the amount used in the pulldown. Cell lines labeled with "control" were transduced with an empty control vector. Exposure time was identical for input and pulldown samples except for Western blots analyzing STRAP, where pulldown exposure was five times shorter (*). Pulldown efficiency in D corresponds to the signal obtained in the pulldown relative to the signal obtained in the input sample. To facilitate comparisons, pulldown efficiency was normalized to the wildtype construct. E, schematic representation of SZRD1 interacting with UPF1 and STRAP. F and G, HEK293 cells expressing STRAP (F) or UPF1 (G) with an N-terminal SFB tag or empty vector control cells ("ctrl") were engineered to express an shRNA targeting SZRD1 ("+") or a nonsilencing shRNA ("–"). Protein lysates were subjected to affinity purification followed by Western blot analysis. Input corresponds to 10% of the amount used in the pulldown. HEK293, human embryonic kidney 293 cell line; STRAP, serine/threonine kinase receptor-associated protein; SZRD1, SUZ domain-containing protein 1; UPF1, up-frameshift 1.

construct that carries a premature stop codon at position 39 of the β -globin sequence (71), which triggers degradation by NMD (Fig. 3A, "NMD-BoxB-noARE") (72). As a control, we used a construct carrying an AU-rich element in the 3'UTR, which drives rapid degradation of the reporter mRNA (Fig. 3A, "WT-BoxB-ARE") (73). Furthermore, in all experiments, we cotransfected a firefly luciferase expression construct driven by the same promoter but lacking the indicated elements in the 3'UTR. Since the best-characterized outcome of NMD activity is mRNA degradation, we quantified mRNA levels of the *Renilla* and firefly luciferase transcripts by RT-quantitative PCR (qPCR) in parallel to the luciferase activity measurements at steady state (Fig. 3A, cyan arrows).

In these experiments, we focused on the long SZRD1 isoform since it is more conserved during evolution. When we compared the relative reporter gene mRNA levels (i.e., *Renilla* relative to firefly luciferase mRNA) of the NMD-sensitive ("NMD-BoxB-noARE") and the ARE ("WT-BoxB-ARE") with the control construct ("WT-BoxB-noARE"), we observed 75% and 60% lower relative reporter mRNA levels, respectively (Fig. 3B). Cotransfection with a plasmid driving expression of the λ N-tagged long SZRD1 isoform did not change reporter mRNA levels for the ARE construct, and we only observed a nonsignificant 1.5-fold increase in relative reporter mRNA levels produced from the NMD-sensitive reporter gene construct (Fig. 3B). At first sight, this

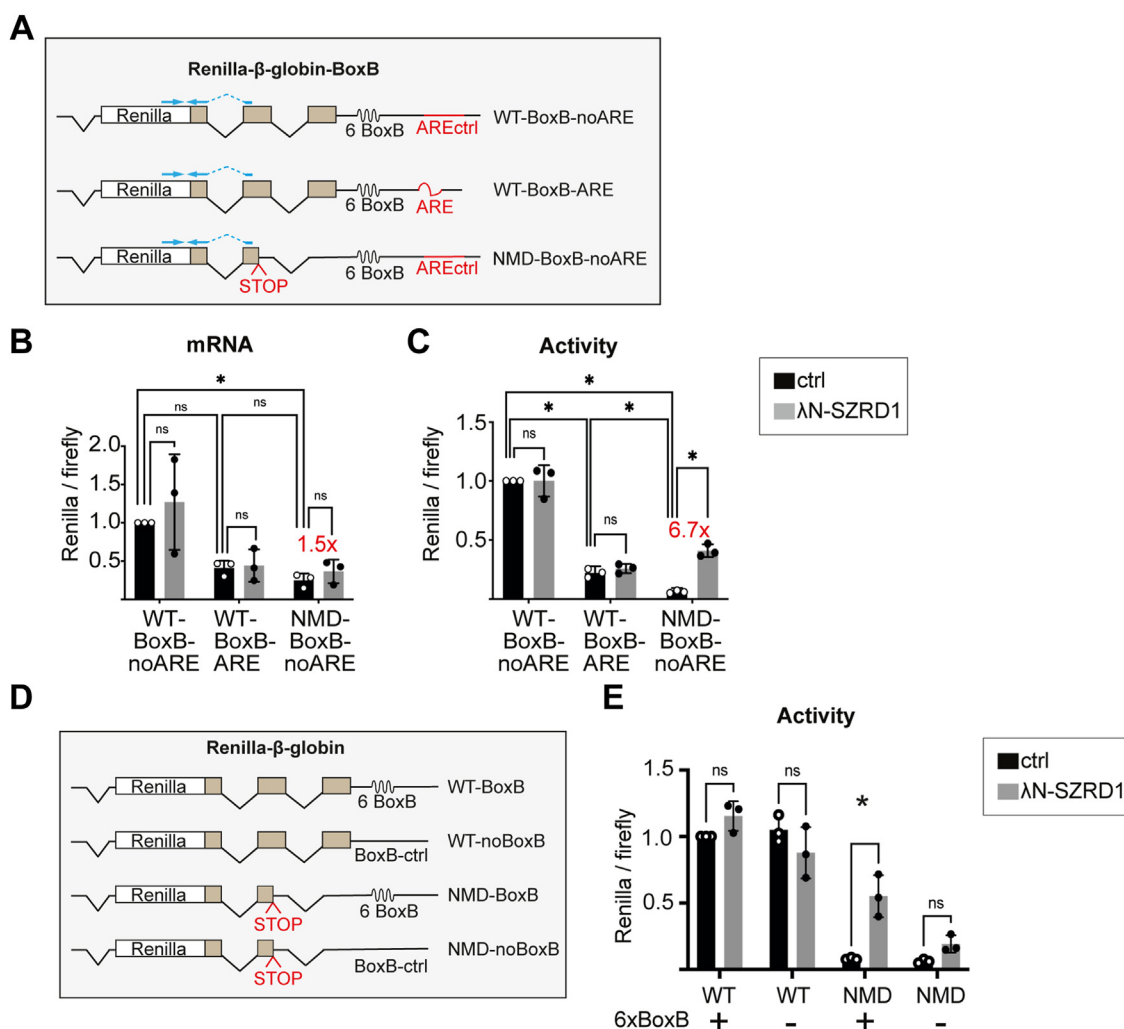


Figure 3. Recruitment of SZRD1 increases protein production from NMD-sensitive reporter constructs. A, schematic representation of *Renilla* reporter gene constructs that allow recruitment of λ N-tagged proteins. Boxes represent the ORFs, which contains the *Renilla* ORF (in white) followed by a genomic β -globin sequence (in brown). The "V" shape illustrates intron positions. In the 3'UTR, six BoxB sites ("BoxB") and an AU-rich element ("ARE") or a control sequence ("AREctrl") were inserted. Locations of the primers used in the RT-qPCR are indicated in cyan, and the premature stop codon is indicated with "STOP". A constitutively active firefly construct with the same promoter was used as transfection control. B and C, relative luciferase mRNAs (B, assessed by RT-qPCR) and activities (C, assessed by luminescent assay) in HEK293T cells after transfection of the indicated reporter gene constructs in the presence ("N-SZRD1") or the absence ("control" = empty vector) of plasmids driving expression of the long form of SZRD1 with a λ N tag in the N terminus. Values represent the ratio of *Renilla* to firefly luciferase activities or mRNA levels and are means $\pm$ SD of three independent experiments, where each condition was performed in triplicates. Asterisks indicate $p < 0.05$ in post hoc testing after two-way ANOVA; ns denotes not significant. D, schematic representation of *Renilla* luciferase reporter gene constructs used in (E) that allow ("WT-BoxB" and "NMD-BoxB") recruitment of λ N-tagged SZRD1 via six BoxB sites or not ("WT-noBoxB" and "NMD-noBoxB"). Each construct contains the *Renilla* luciferase ORF (in white) followed by a genomic β -globin sequence (in brown) containing two introns. In the 3'UTR, six BoxB sites ("BoxB") or not ("BoxB-ctrl") were inserted. A premature stop codon ("STOP") in the NMD constructs triggers NMD. E, relative luciferase activities observed for the reporter gene constructs ("WT" and "NMD") with 6x BoxB sites (+) or not (-), when cotransfected in the presence (gray bars) or the absence (black bars, "ctrl" = empty vector) of plasmid driving expression of the λ N-tagged long form of SZRD1 ("N-SZRD1"). Values represent the ratios of *Renilla* to firefly luciferase activities and are means $\pm$ SD of three independent experiments, where each condition was transfected in triplicates. Asterisks indicate $p < 0.05$ in post hoc testing after two-way ANOVA; ns denotes not significant. HEK293T, human embryonic kidney 293 cell line; NMD, nonsense-mediated decay; qPCR, quantitative PCR; SZRD1, SUZ domain-containing protein 1.

indicated that mRNA degradation triggered by NMD was weakly affected by SZRD1 recruitment. Yet, the situation was different, when we quantified luciferase activities in the same experiments to assess protein production from reporter transcripts. Similarly to the situation on mRNA level, recruitment of the long SZRD1 isoform did not lead to any effect on the control reporters (Fig. 3C). In contrast, we observed a 6.7-fold increase in relative luciferase activity from the NMD-sensitive reporter gene construct. Thus, recruitment of SZRD1 led to a larger increase in reporter protein

production in comparison to mRNA production (6.7-fold versus 1.5-fold). Almost identical observations were made in another cell line (U2OS, Fig. S3, A and B) and with the short isoform (data not shown). Of note, this discrepancy was only observed for the NMD-sensitive reporter construct. Given that NMD not only leads to transcript degradation but also to an inhibition of translation (24, 42, 44, 69, 74), this observation suggests that SZRD1 might predominantly relieve the translational repression and only slightly affect the transcript degradation occurring during NMD.

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

Next, we investigated whether recruitment of SZRD1 to the transcript was actually required for its effect on the reporter gene construct. To do this, we compared the effect of λ N-tagged SZRD1 on reporter constructs that either contain six BoxB sites or not (Fig. 3D). Cotransfection of λ N-tagged SZRD1 led to an increase of *Renilla* luciferase activity from both NMD-sensitive reporters (Fig. 3E). However, identical amounts of λ N-tagged SZRD1 led to a threefold larger increase in luciferase activity produced from the BoxB site containing reporter gene transcript in comparison to the transcript that lacks these sites. This difference was even stronger in U2OS cells where no increase was observed for the reporter constructs lacking BoxB sites (Fig. S3B). Overall, we conclude that SZRD1 increases protein production from reporter transcripts recognized by the NMD machinery, and that this effect is facilitated by the recruitment of SZRD1 to these transcripts.

Degradation of transcripts is the main consequence of NMD. However, NMD machinery can also lead to an inhibition of mRNA translation (24, 42, 44, 69, 74). A combination of transcript degradation and inhibition of translation has also been observed when miRNAs bind to mRNA transcripts (75). Thus, we wondered whether SZRD1 might affect protein production from transcripts targeted by miRNAs. To test this, we inserted four let7a-binding sites ("let7a_bs") or binding sites where key nucleotides had been mutated ("let7a_mu") in the 3'UTR of reporter constructs containing six BoxB sites (Fig. S3C). When we transfected these constructs into HEK293T cells, we observed approximately 50% less *Renilla* luciferase activity from the let7a reporter construct in comparison to the construct where the binding sites were mutated at steady state (Fig. S3D). Coexpression of long λ N-tagged SZRD1 isoform did not change the activity of the let-7a reporter construct (Fig. S3D). This indicates that recruitment of SZRD1 does not block post-transcriptional regulation by let-7 in this reporter gene construct.

Next, we were interested to find out whether SZRD1 also affects protein levels produced from endogenous NMD-sensitive transcripts. In many cases, NMD-sensitive transcripts represent one out of several splice variants, which all contribute to protein production, making it difficult to understand the contribution of NMD-sensitive transcript to overall protein levels. To select proteins that might change in abundance depending on NMD activity, we turned to a quantitative proteomics study that identified proteins that are translated from transcripts targeted by NMD and that were upregulated upon UPF1 knockdown (76). Among the top four candidates, antibodies recognizing GABARAPL1 and PEA15 were commercially available. We then analyzed protein levels by Western blot and mRNA levels by RT-qPCR in cell lines expressing different SZRD1 shRNAs or a nonsilencing control shRNA (Fig. 4A). PEA15 protein was undetectable in HEK293 cells. Yet, we observed a 50% reduction in GABARAPL1 protein levels (significant with one shRNA and not significant with the other shRNA despite comparable fold change). In these conditions, GABARAPL1 mRNA levels were not reduced (Fig. 4, B and C), consistent with a model where

SZRD1 helps maintain protein production from some NMD target transcripts. Unfortunately, we were unable to pulldown GABARAPL1 transcripts in an RNA immunoprecipitation approach, which leaves the question unanswered whether SZRD1 directly binds to these transcripts. Furthermore, these observations are clearly context dependent, since a comparable analysis in U2OS cells revealed no change in GABARAPL1 and a 50% reduction in PEA15 protein levels (Fig. 4, D–H).

Taken together, we conclude that the recruitment of SZRD1 to the NMD-sensitive reporter transcript increases protein production with only minor effects on mRNA levels. Consistent with a contribution of SZRD1 to the protein production from some endogenous NMD-sensitive transcripts, we observed a context-dependent reduction of GABARAPL1 and PEA15 protein levels when SZRD1 was knocked down. In contrast, no effect of SZRD1 was observed on reporter gene constructs with an ARE or let-7 binding sites in the 3'UTR, suggesting that the action of SZRD1 has some specificity for transcripts targeted by NMD. Further studies will be needed to determine which subset of NMD targets is affected by SZRD1.

SZRD1 does not interact with the EJC

EJCs remaining on transcripts can trigger NMD, and different EJC complex components have variable contributions to NMD (32). Hence, we wondered whether SZRD1 might interact with components of the EJC and thereby exert its effect on NMD target transcripts. To avoid artifacts because of the overexpression of proteins, we used cell lines, where an SFB tag had been knocked in after the ORF of the genetic loci coding for the EJC components eIF4A3 and RNPS1 as well as UPF1 (Fig. S4A). This also allowed us to ensure that all proteins were pulled down with comparable efficiency, which is particularly important in the case of RNPS1 where reliable antibodies are unavailable. When we pulled down SFB-tagged endogenous proteins with streptavidin beads, we saw that endogenous UPF1 copurified with endogenous SZRD1 (Fig. S4B). In contrast, we did not observe any copurification of SZRD1 with the EJC components RNPS1 or eIF4A3. At the same time, the EJC components MLN51 and Y14 (also named RBM8A) did interact with eIF4A3 and RNPS1 suggesting that EJC components could be identified using our approach (Fig. S4B). Likewise, UPF1 did copurify with MLN51, UPF2, and UPF3B suggesting that the interactions between NMD and EJC components were also maintained. Thus, we did not find evidence that SZRD1 would modulate NMD by an interaction with EJC components.

The SUZ domain is required for the effect of SZRD1 and ARPP21 on an NMD-sensitive reporter transcript

To elucidate the mechanism of action of SZRD1, we decided to identify the domains of SZRD1 that are required for its effect on the NMD-sensitive reporter construct. To achieve this, we generated expression constructs for λ N-tagged versions of the SZRD1 mutants described in Figure 2, A and B. As observed before, the λ N-tagged long SZRD1 isoform increased *Renilla* activity from an NMD-sensitive reporter gene

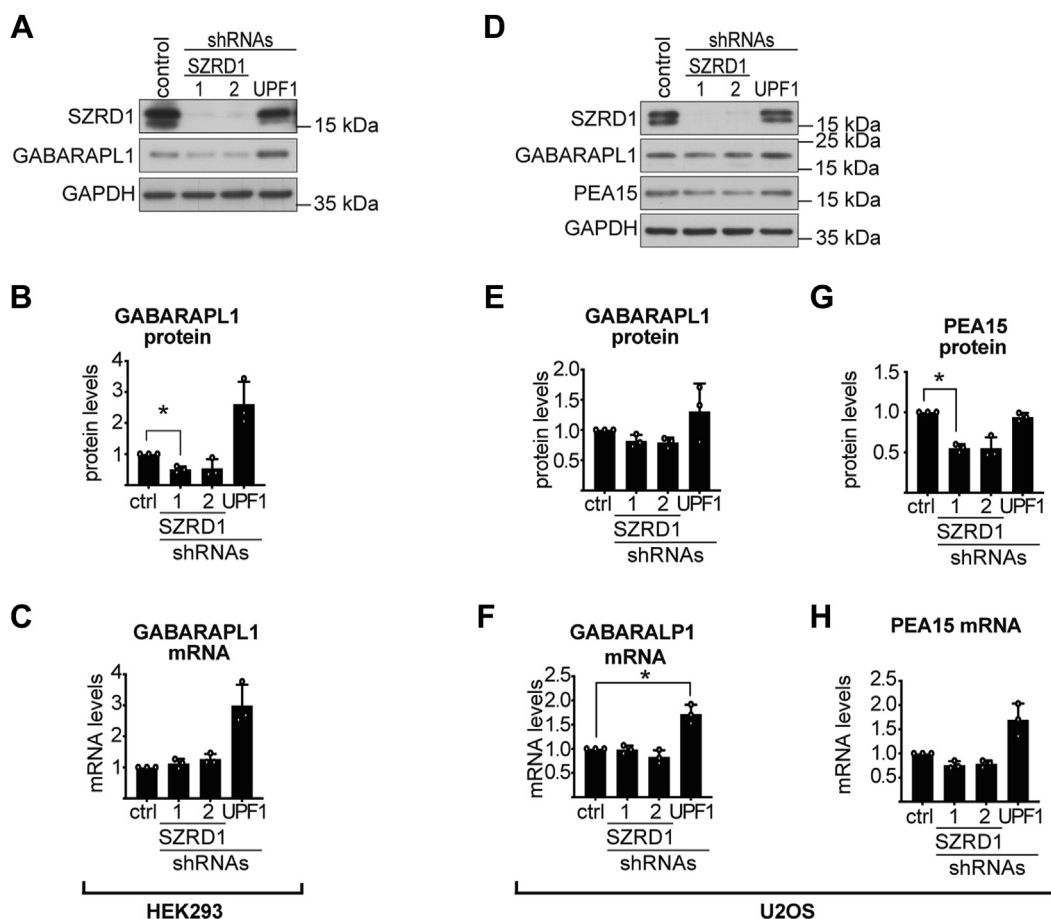


Figure 4. SZRD1 knockdown decreases protein production from the NMD-sensitive mRNAs GABARAPL1 and PEA15. A, Western blot analysis using the indicated antibodies in HEK293 cell lines expressing a nonsilencing shRNA ("control"), two different shRNAs targeting SZRD1 ("1" and "2"), or one shRNA targeting UPF1. B and C, quantitative analysis of Western blots for GABARAPL1 normalized to GAPDH (B) and qPCR analysis for GABARAPL1 mRNA levels normalized to TBP and β 2-microglobulin (C) in cell lines described in A. Data represent means $\pm$ SD from three independent experiments and were normalized to the control shRNA condition. D, Western blot analysis of U2OS cell lines expressing a nonsilencing shRNA ("control"), two different shRNAs targeting SZRD1 ("1" and "2"), or one shRNA targeting UPF1. E–H, quantification of GABARAPL1 (E and F), PEA15 (G and H) protein levels (E and G), and mRNA levels (F and H) in U2OS cell lines were determined and are presented as shown in B and C. Post hoc testing after one-way ANOVA was performed in comparison to the control condition. Asterisks indicate comparisons with $p < 0.05$; not significant comparisons are not indicated. Blots presented in A are one of three experiments that were also used to generate Figure 7, A and F. As such, panels for SZRD1 and GAPDH in A and D are identical to the ones in Figure 7, A and F. HEK293, human embryonic kidney 293 cell line; NMD, nonsense-mediated decay; qPCR, quantitative PCR; SZRD1, SUZ domain-containing protein 1; UPF, up-frameshift.

construct ("NMD-BoxB") (Fig. 5, A and B) but did not affect the control construct ("WT-BoxB", Fig. S5A). Mutations in the SUZ domain almost completely abolished the effect of SZRD1 on the NMD-sensitive reporter, with the exception of the mutant "R→Q," which still showed a residual effect (Fig. 5B, bar 6). Of note, expression levels of the mutant proteins were not lower than the levels of the wildtype protein (Figs. 5B and S5B), suggesting that difference in protein expression cannot explain the differential effects in reporter assays. The effect of SZRD1 was not abolished but even stronger when we introduced a mutation in the SUZ-C domain that abolishes the interaction with STRAP ("AAA") (Fig. 5B, bar 7). Thus, our data indicate that the SUZ domain is required for the increase of *Renilla* luciferase levels generated by NMD-sensitive reporter gene constructs, whereas the SUZ-C domain is not.

Interestingly, some of the SZRD1 mutants still showed some interaction with UPF1 (Fig. 2D) but were unable to increase *Renilla* luciferase levels produced from NMD-sensitive reporter gene constructs (Fig. 5B). This suggests that the

interaction with UPF1 might not suffice for the effect of SZRD1 but that an interaction with other players is required.

Besides in SZRD1, SUZ domains have only been found in a group of related proteins that comprises R3HDM1, R3HDM2, and ARPP21. Recently, Rehfeld *et al.* (56) demonstrated that ARPP21 can increase protein production from some transcripts when bound to uridine-rich sequences in the 3'UTR. Some data were presented that this occurs by an increase in protein translation in part *via* an interaction with eIF4A1 and eIF4G, but a combined effect on protein translation and transcript abundance could not be excluded.

In analogy to our observations concerning SZRD1, we hypothesized that recruitment of ARPP21 might lead to a stronger effect on an NMD-sensitive reporter than on control constructs, potentially by relieving translational repression (Fig. 5C). This hypothesis was also supported by previous publications that reported UPF1 as a potential interaction partner of R3HDM1 and ARPP21 ((69) and supporting information of Ref. (56)). First, we tested whether ARPP21

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

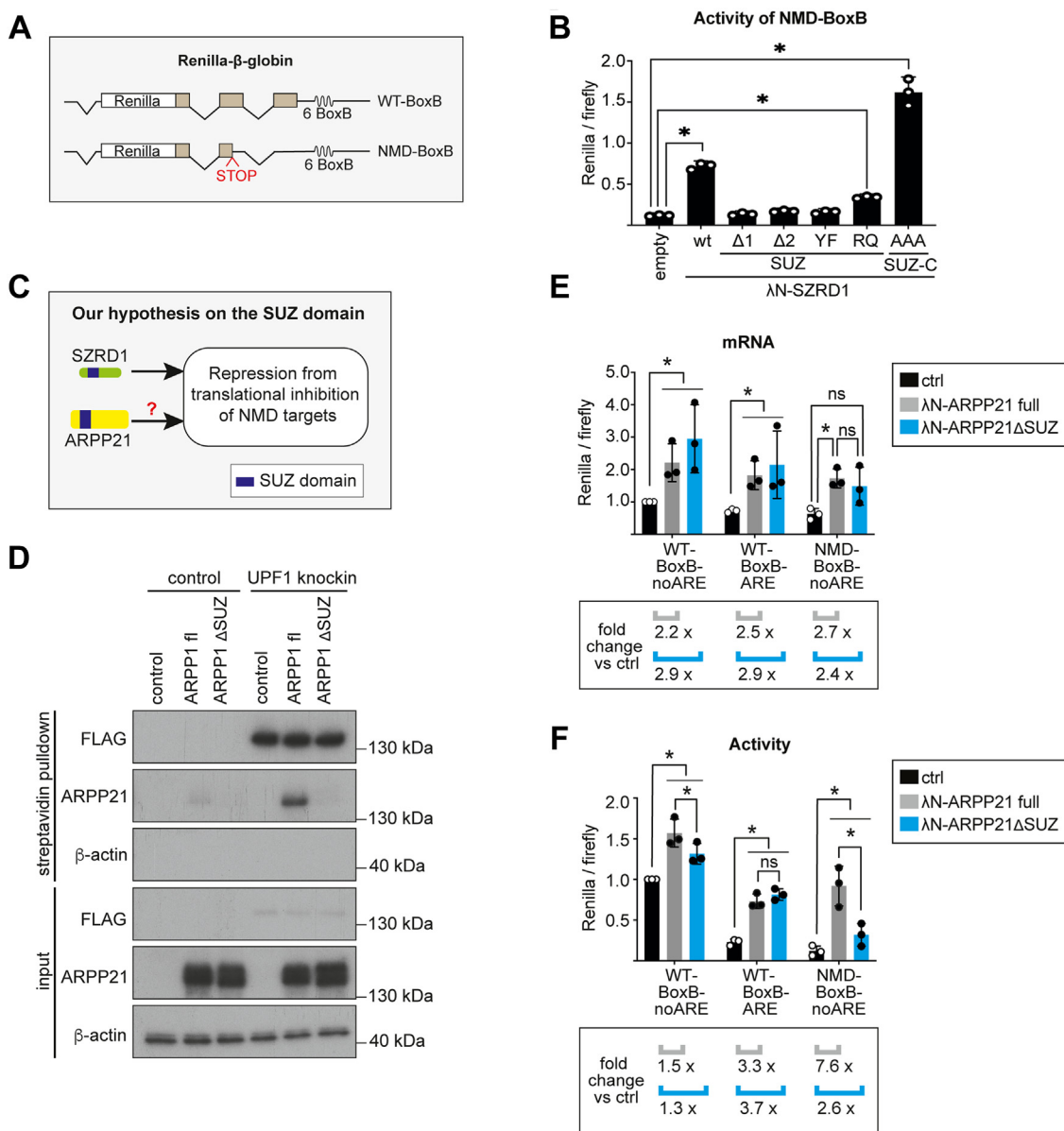


Figure 5. The SUZ domain of SZRD1 and ARPP21 is required to increase protein production from NMD-sensitive reporter constructs. *A*, schematic representation of Renilla reporter gene constructs used in (*B*) and Fig. S3A. *B*, relative luciferase activities of the NMD-BoxB construct in the presence or the absence of λN-tagged long SZRD1 containing or not mutations in the SUZ and SUZ-C domains (Fig. 2, A and B). Values represent the ratio of Renilla to firefly luciferase activities and are means ± SD of three independent experiments, where each condition was performed in triplicates. Asterisks indicate $p < 0.05$ in post hoc testing after two-way ANOVA. *C*, hypothesis about the role of the SUZ domain of SZRD1 and ARPP21. *D*, HEK293 cells with an SFB tag knocked in the UPF1 locus or a control cell line ("control") were transiently transfected with wildtype ARPP21, a deletion mutant lacking part of the SUZ domain, or an empty vector control ("control"). Protein lysates were subjected to affinity purification followed by Western blot analysis with the indicated antibodies. *E* and *F*, relative mRNA levels (*E*) and luciferase activities (*F*) obtained and presented as in panels of the Figure 3, B and C, except that expression constructs for N-terminally λN-tagged ARPP21 (gray bars) or a splice variant of ARPP21 missing a part of the SUZ domain (cyan bars) were transfected. HEK293, human embryonic kidney 293 cell line; NMD, nonsense-mediated decay; SZRD1, SUZ domain-containing protein 1.

indeed interacts with UPF1. To this end, we used expression constructs for wildtype ARPP21 (λN-ARPP21) or for a rare splice variant of ARPP21 that codes for a protein lacking the second half of the SUZ domain (λN-ARPP21ΔSUZ; *Figure 2B*, underlined in black). These constructs were transfected in a cell line where the endogenous UPF1 protein is tagged with a C-terminal SFB tag or parental HEK293 cells. Pulldown of endogenous UPF1 recovered ARPP21 specifically in the knockin cell lines, whereas ARPP21 with a deletion of an SUZ domain was not recovered (Fig. 5D). This demonstrates that

ARPP21 can indeed interact with UPF1 in an SUZ-domain dependent manner. Next, we transfected HEK293T cells with the reporter gene constructs described in Figure 3A, which allow us to recruit λN-tagged proteins to the 3'UTR. When we quantified mRNA levels for Renilla and firefly luciferase by RT-qPCR, we observed that expression of λN-tagged ARPP21 led to a threefold to threefold increase in the ratio of Renilla and firefly mRNA levels for all three reporter gene constructs (Fig. 5E). This effect was independent of the SUZ domain, since wildtype ARPP21 and the variant lacking the SUZ

domain led to comparable increases in the abundance of reporter transcripts.

We next determined luciferase activities as a representation of the protein production from these transcripts. Recruitment of λ N-tagged ARPP21 to the wildtype construct led to a 1.5-fold increase in the ratio between *Renilla* and firefly luciferase activities, similar to what has been observed by Rehfeld *et al.* (56) (Fig. 5F). However, we observed that recruitment of wildtype ARPP21 to the NMD-sensitive transcript led to a much larger increase in reporter activity than in the case of the ARE or wildtype constructs (7.6-fold *versus* 3.3-fold and 1.5-fold, respectively, in Fig. 5F), although the increases in mRNA levels were comparable for all three reporters (Fig. 5E). This was reminiscent of what had been observed for SZRD1 (Fig. 3, B and C) and seemed consistent with a release from translation repression when ARPP21 is recruited. Strikingly, recruitment of the ARPP21 variant lacking the SUZ domain only led to a 2.6-fold increase of *Renilla* luciferase activity from the NMD-sensitive transcript, whereas the wildtype form led to a 7.6-fold increase. This is remarkable, since mRNA levels were similar in the presence of the wildtype and SUZ-deficient ARPP21. Thus, in the absence of the SUZ domain, 2.5-fold less protein is produced from the same amount of transcript. Given that this difference is only observed for the NMD-sensitive construct, our observations indicate that recruitment of ARPP21 and also SZRD1 leads to an SUZ domain-dependent relief from NMD-induced translational repression.

Strong inhibition of NMD upon supraphysiological expression of SZRD1 likely does not represent the physiological function of SZRD1

To evaluate the role of SZRD1 on endogenous transcripts, we overexpressed the long form of SZRD1 in HeLa cells (Fig. 6B). This cell line was chosen since the repertoire of NMD targets has been extensively characterized in this cell line (7). Next, we used RNA-Seq to compare gene expression profiles between cells overexpressing SZRD1 with empty vector control cells (Fig. 6A). Subsequently, we performed a gene set enrichment analysis with the Molecular Signature Database (77) and with sets of genes that had been found to be upregulated when UPF1 or other NMD components were knocked down in the same cell line (7). This revealed a massive enrichment of NMD targets among the transcripts upregulated upon SZRD1 overexpression ($p < 10^{-139}$) (Fig. 6, C and D) suggesting that SZRD1 overexpression can inhibit NMD. We then confirmed these observations by testing the effect of overexpressing the long and short forms of SZRD1 in HEK293 cells. Transcript levels of all four NMD-sensitive transcripts (GAS5, ZFAS1, and splice variants of hnRNP A2B1 and SC-35) (32, 78–80) (Fig. 6, E–H) increased significantly at least threefold upon SZRD1 overexpression. Next, we tested whether knockdown of SZRD1 would lead to reciprocal changes in the abundance of these *bona fide* NMD targets. However, the two different shRNAs targeting SZRD1 levels did not lead to significant reductions in the abundance of these transcripts (Fig. 6, I–L). Given the very high levels of

SZRD1 obtained in the acute overexpression experiments (*i.e.*, more than 20 $\times$ higher than in control cells), this suggested that the observed inhibition of NMD was an artifact because of high SZRD1 protein levels. Of note, NMD targets were also enriched in transcripts upregulated upon ARPP21 overexpression, albeit to a lesser extent (Fig. S6).

We had earlier demonstrated that SZRD1 and ARPP21 can interact with UPF1 (Figs. 2 and 5D). Hence, overexpression of SZRD1 might disturb the stoichiometry of protein complexes in the NMD machinery and thereby disrupt their function.

Endogenous SZRD1 is required to maintain UPF3B protein levels but has minor effects on the abundance of NMD-sensitive transcripts

Next, we wanted to explore the role of SZRD1 at physiological expression levels. To this end, we used shRNAs to knock down SZRD1 or UPF1 in the cell lines HEK293 and U2OS (Fig. 7, A and F). First, we assessed knockdown efficiency and effects on the NMD machinery by Western blot. This revealed a more than 80% reduction in protein levels for UPF1 and SZRD1 (Fig. 7, A, B, F and G). Surprisingly, during these analyses, we found that knockdown of SZRD1 caused a reduction of UPF3B protein levels of 80% in HEK293 and 70% in U2OS cells (Figs. 7, E and J and S7), whereas other NMD components were less affected (Fig. 7, B–D and G–I). Therefore, endogenous SZRD1 seems to be required to maintain cellular expression levels of the NMD-machinery component UPF3B. To investigate whether the change in UPF3B levels was caused by an off-target effect of the shRNAs or by the loss of the short or the long isoform, we re-expressed these isoforms in SZRD1 knockdown cell lines. This led to a recovery of UPF3B levels upon re-expression of the long SZRD1 isoform, whereas expression of the short isoform did not affect UPF3B levels (Fig. S7, C–F). Consistent with this, a specific knockdown of the long SZRD1 isoform was sufficient to reduce cellular UPF3B levels (Fig. S7G), indicating that the long isoform is more important in maintaining UPF3B levels than the short isoform. Of note, in previous experiments, we had noticed that SZRD1 overexpression increased UPF1 protein levels in an SUZ domain-dependent manner (Figs. 1, D and E, 2, C and D and S2, A and B), but this was not the case for the other proteins that interacted with SZRD1 (*i.e.*, STRAP). A more detailed analysis revealed that the abundance of several NMD machinery components (*i.e.*, UPF1, SMG1, UPF2, and UPF3B) was increased upon SZRD1 overexpression, whereas GAPDH protein levels were unaffected (Fig. S7A). Interestingly, the increase of the NMD components was abolished by mutations in the SUZ domain (“ $\Delta 1$ ” and “ $\Delta 2$ ”) (Fig. S7A). Of course, some part of this effect might be caused by the strong overexpression of SZRD1 in these experiments. Yet, taken together, our data demonstrate that endogenous SZRD1 is required to maintain normal UPF3B protein levels, potentially in an SUZ domain-dependent manner.

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

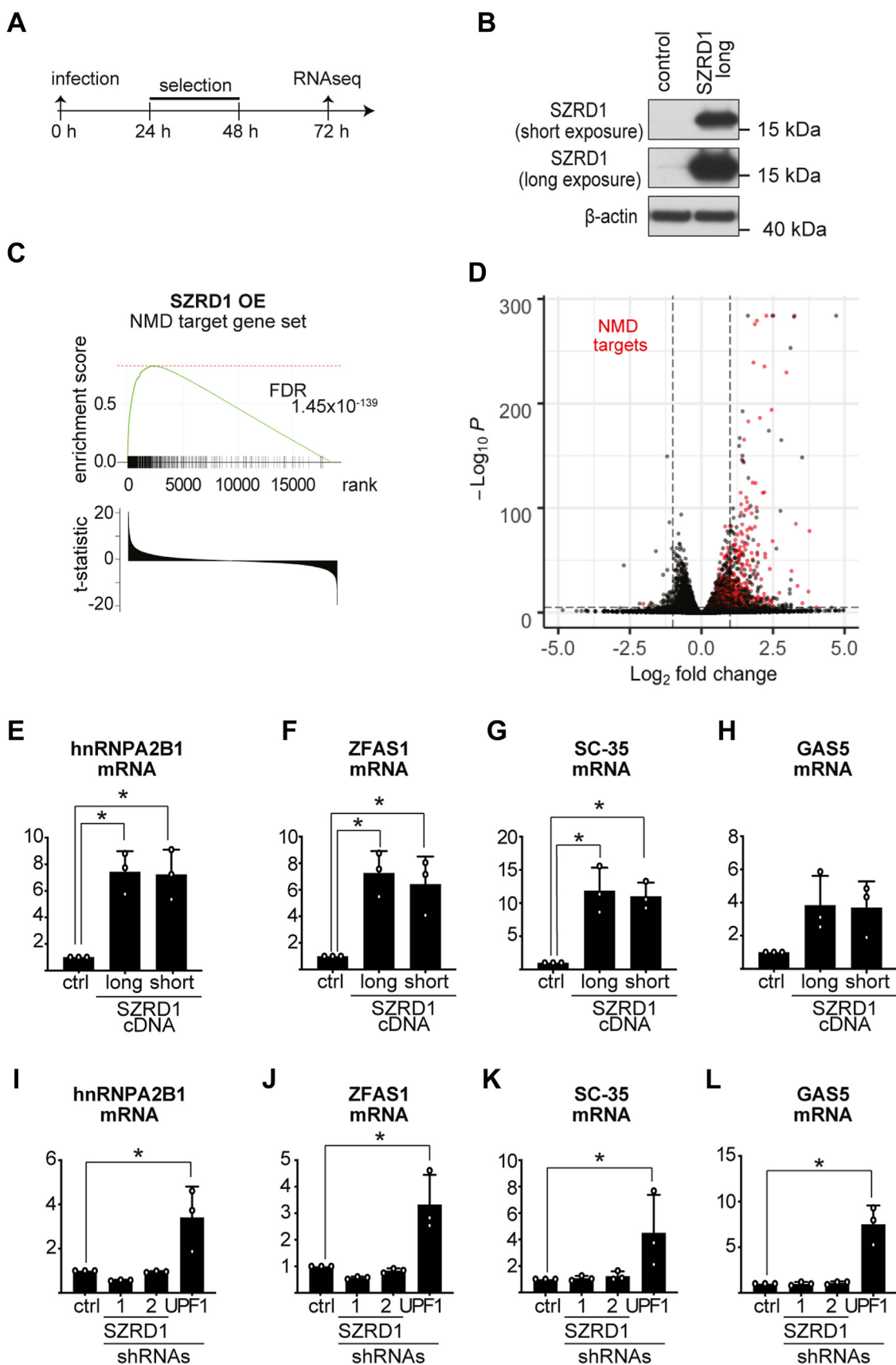


Figure 6. Strong inhibition of NMD upon supraphysiological expression of SZRD1 likely does not represent the physiological function of SZRD1. A, RNA-Seq was performed 48 h after infection of HeLa cells with lentiviruses driving the overexpression of the long form of SZRD1 or an empty expression cassette. B, Western blot analysis of the infected HeLa cell lines expressing SZRD1 or not. C, GSEA with a gene set of NMD targets revealed strong enrichment among transcripts increased upon SZRD1 overexpression. Genes are sorted according to the signed *t*-statistic. Upregulated genes are on the

Reporter gene assays shown in [Figures 3 and 5](#) indicated that recruitment of SZRD1 and ARPP21 to transcripts can prevent some of the translational inhibition occurring during NMD. In addition, shRNA-mediated knockdown revealed that SZRD1 is required to maintain normal protein levels of the NMD machinery component UPF3B ([Fig. 7](#)). This indicated that SZRD1 might have two apparently opposing effects on NMD targets: on the one hand, the reporter assays indicated that SZRD1 might oppose NMD activity by preventing translational inhibition. On the other hand, SZRD1 may serve to maintain NMD function *via* an increase in NMD machinery components such as UPF3B ([Fig. 7O](#)).

We therefore wanted to explore the effect of endogenous SZRD1 on the abundance of mRNAs targeted by NMD. To answer this question, we knocked down SZRD1 in the cell line HeLa. We then used RNA-Seq to compare gene expression profiles between cells expressing SZRD1 shRNA #1 or a nonsilencing control. A gene set enrichment analysis revealed an enrichment of NMD target genes among the genes that were upregulated upon SZRD1 knockdown ([Fig. 7, K–M](#), false discovery rate = 9.7×10^{-12}). Thus, our RNA-Seq analysis indicates that NMD activity on a subset of NMD targets requires normal SZRD1 levels. Yet, while this enrichment was highly significant, changes remained below twofold, indicating that in our experimental system, the reduction of UPF3B caused by SZRD1 knockdown only leads to minor effects on the abundance of NMD-sensitive transcripts ([Fig. 7N](#)).

Taken together, our data support a model of a dual role of SZRD1 in NMD. First, SZRD1 is required to maintain UPF3B protein levels. As a consequence, SZRD1 can affect wide range of NMD targets albeit with very low efficiency. In contrast, local recruitment to specific mRNAs can prevent the translational inhibition occurring during NMD ([Fig. 7O](#)).

Discussion

Recruitment of SUZ domain-containing proteins SZRD1 and ARPP21 can increase protein production from NMD targets

NMD helps prevent the formation of potentially dangerous proteins that could be formed because of aberrant splicing or to gene mutations. Up to 10% of all transcripts are affected by NMD because of long 3'UTRs, the presence of EJC in the 3'UTR, or upstream ORFs (81). Many of these NMD targets are not “abnormal” transcripts but code for physiologically relevant proteins (9). NMD is best known to induce the degradation of transcripts, but it also inhibits their translation (24, 42, 44, 69, 74). Yet, in some instances, it might be desirable to escape from NMD to increase protein expression, while maintaining a short mRNA half-life. Thus, regulation of translation of some NMD targets might be physiologically relevant.

We found that recruitment of SZRD1 to NMD-sensitive reporter transcripts increases reporter protein production with very little changes in mRNA levels ([Figs. 3 and S3](#)). Interestingly, the discrepancy between the effect on protein levels and mRNA levels was only observed for NMD-sensitive reporters. This suggests that SZRD1 acts, at least in part, by counteracting the translational inhibition observed during NMD. We found that this effect was dependent on the presence of the SUZ domain ([Fig. 5B](#)). While we do not know how SZRD1 achieves this, we excluded that this effect is due to an interaction with components of the EJC ([Fig. S4](#)). Of note, we obtained similar results with two endogenous transcripts targeted by the NMD machinery ([Fig. 4](#)), but we did not succeed to demonstrate direct recruitment of SRZD1 to these transcripts. This leaves open the question whether and how SZRD1 might be recruited to a subset of NMD target transcripts.

To test whether other SUZ domain-containing proteins might have a similar effect, we investigated the effect of ARPP21 recruitment to reporter gene transcripts. This revealed that the function of ARPP21 might consist of two components: an effect on transcripts levels and an effect on transcript translation. In these experiments, we made use of a rare splice variant of ARPP21 lacking part of the SUZ domain. Both forms of ARPP21 led to comparable twofold to threefold increases in RNA levels produced from all three reporter constructs ([Fig. 5E](#)). In addition, both forms of ARPP21 also led to comparable increases in protein production (assessed *via* luciferase activity measurement) from the control reporter construct and from the construct carrying an ARE in the 3'UTR ([Fig. 5F](#)). Strikingly, protein production from the NMD-sensitive reporter construct upon recruitment of wildtype ARPP21 was much stronger increased (7.6-fold) than upon recruitment of the SUZ-deficient ARPP21 (2.6-fold), whereas mRNA levels were almost identical ([Fig. 5E](#)) under both conditions. This indicates that ARPP21 increases translation from the NMD-sensitive reporter construct in a SUZ domain-dependent manner. Given that the loss of the SUZ domain only led to minor changes in protein production from the other reporter constructs, and given that translational inhibition occurs during NMD, our observations are consistent with a model where the SUZ domain in ARPP21 is required for a relief from translational inhibition.

At present, the molecular mechanism of how the SUZ domain is involved in this process remains unclear. Previous studies have reported that the SUZ domain in SZY-20, the *C. elegans* homolog of SZRD1, can bind to single-stranded RNA, but it is unclear whether this would occur in a sequence-specific manner (50). In some of our experiments,

left. D, volcano plot of gene expression changes upon overexpression of SZRD1. Members of the NMD target gene set are highlighted in red. *E–H*, mRNA levels of hnRNP2B1 (*I*), ZFAS1 (*J*), SC-35 (*K*), and GAS5 (*L*) were measured by RT-qPCR in HEK cells expressing the indicating SZRD1 protein (“long” or “short”) or not (“ctrl” = empty vector control). Expression levels were normalized to TBP and β 2-microglobulin mRNA levels. Data represent means $\pm$ SD from three independent experiments and were normalized to the empty vector condition. *I–L*, mRNA levels of hnRNP2B1 (*I*), ZFAS1 (*J*), SC-35 (*K*), and GAS5 (*L*) were measured by RT-qPCR in HEK cells expressing a nonsilencing control shRNA (“ctrl”), two shRNAs (“1” and “2”) targeting SZRD1, or one shRNA targeting UPF1. Expression levels were normalized to TBP and β 2-microglobulin mRNA levels. Data represent means $\pm$ SD from three independent experiments and were normalized to the control shRNA condition. FDR, false discovery rate; GSEA, gene set enrichment analysis; HEK, human embryonic kidney 293 cell line; NMD, nonsense-mediated decay; qPCR, quantitative PCR; SZRD1, SUZ domain-containing protein 1; UPF, up-frameshift.

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

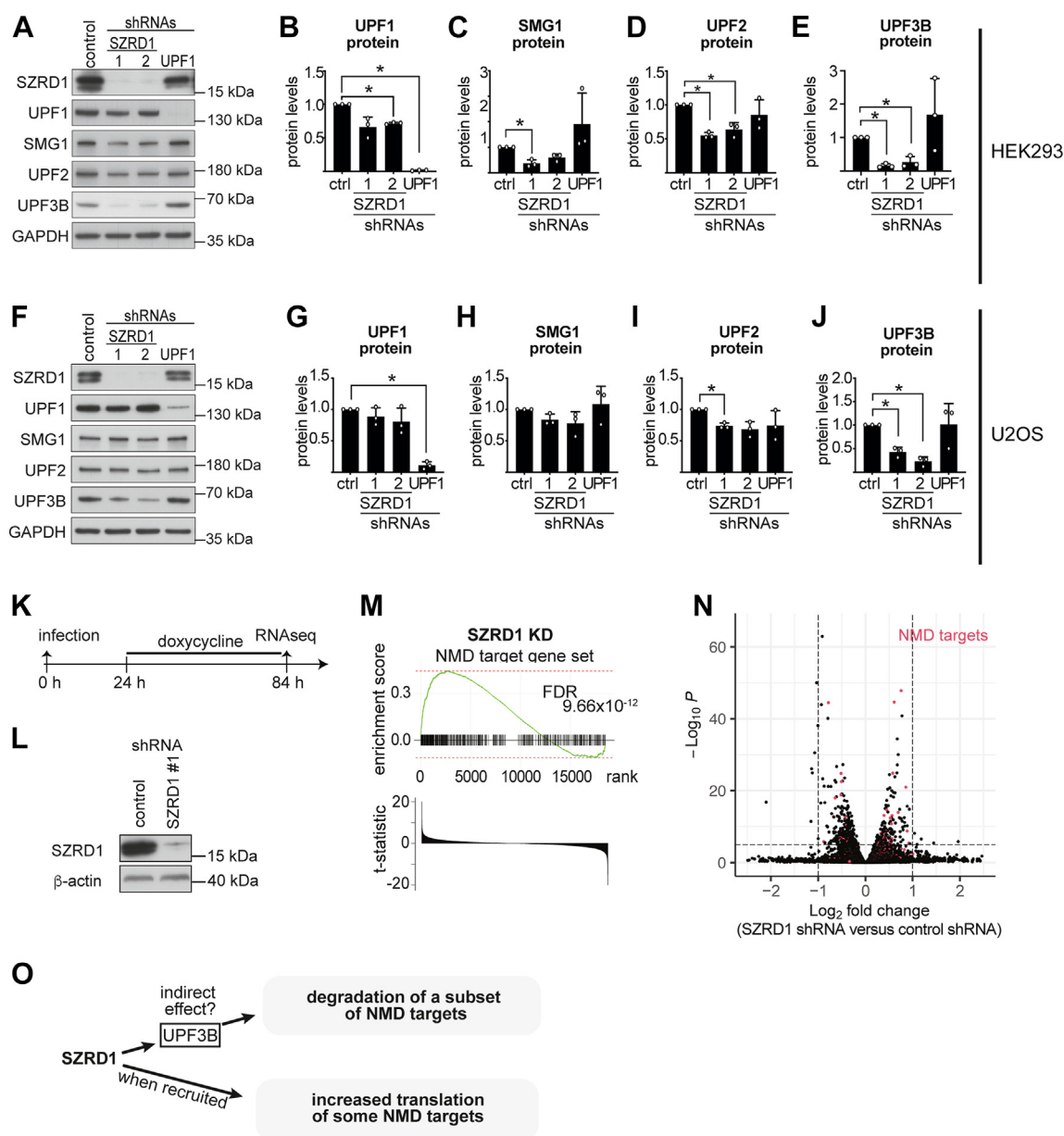


Figure 7. Endogenous SZRD1 is required to maintain UPF3B proteins levels but only has minor effects on mRNA levels of NMD targets. A, Western blot analysis using the indicated antibodies in HEK293 cell lines expressing a nonsilencing shRNA ("control"), two different shRNAs targeting SZRD1 ("1" and "2"), or one shRNA targeting UPF1. B–E, quantification of Western blot signals with the indicated antibodies obtained in three independent experiments using HEK293 cell lines expressing a nonsilencing shRNA ("control"), two different shRNAs targeting SZRD1 ("1" and "2"), or one shRNA targeting UPF1. Signals were normalized to the abundance of GAPDH within each experiment and represent means $\pm$ SD. Post hoc testing after one-way ANOVA was performed in comparison to the control condition. Asterisks indicate $p < 0.05$; not significant comparisons are not indicated. F, Western blot analysis using the indicated antibodies in U2OS cell lines generated as described in A. G–J, quantification of Western blot signals with the indicated antibodies obtained in three independent experiments using U2OS cell lines as described for the B–E. K, experimental setup for the analysis of HeLa cells with doxycycline-inducible expression of an SZRD1 shRNA or a control shRNA. L, Western blot analysis demonstrating knockdown of SZRD1. M, GSEA revealed the enrichment of a gene set of NMD targets (7) among transcripts increased upon SZRD1 knockdown. Genes are sorted according to the signed t -statistic. Upregulated genes are on the left. N, volcano plot of gene expression changes upon SZRD1 knockdown. Members of the NMD target gene set are highlighted in red. O, schematic representation of the potential roles of SZRD1 based on our observations. FDR, false discovery rate; GSEA, gene set enrichment analysis; HEK293, human embryonic kidney 293 cell line; NMD, nonsense-mediated decay; SZRD1, SUZ domain-containing protein 1; UPF, up-frameshift.

we used λ N-tagged SZRD1 or ARPP21, which are recruited to BoxB sites in the 3'UTR of NMD targets, thereby bypassing a potential requirement of the SUZ domain to bind these proteins to mRNA. Thus, here the SUZ domain was not required for SZRD1 recruitment to reporter transcripts. In this situation, key residues in the SUZ domain were still required to

increase luciferase activity produced from reporters that are targeted by NMD (Fig. 5). This demonstrates that the function of the SUZ domain is not limited to the recruitment of SZRD1 to mRNA but rather that the SUZ domain likely is implicated in the modulation of protein production from transcripts targeted by the NMD machinery.

We showed that SZRD1 interacts with UPF1 in an RNA-independent manner. Furthermore, we showed that an intact SUZ domain is required for this interaction (Figs. 2D and 5D). This is consistent with large-scale immunoprecipitation studies that indicated that ARPP21 (and R3HDM1) might interact with UPF1 and that this interaction is maintained after RNA degradation (supporting information to Refs. (56, 69)). Hence, as for SZRD1, the SUZ domain of ARPP21 and R3HDM1 might be required for their interaction with UPF1.

Considering that UPF1 plays a central role in NMD activity, the interaction between SZRD1 and UPF1 might prevent the interaction of another factor with UPF1 (e.g., by masking its interaction sites). However, mutations in the C-terminal part of the SUZ domain of SZRD1 completely lost their ability to increase luciferase activity from reporters targeted by the NMD but still maintained a certain interaction with UPF1 (Figs. 2D and 5B). This indicates that binding of SZRD1 to UPF1 is not sufficient to increase protein production from NMD-sensitive reporters but that SZRD1 needs to interact with another factor besides UPF1 to exert this effect. Future studies will be required to dissect the mechanistic link between SZRD1 and the NMD machinery.

Regulation of a subset of NMD targets and maintenance of UPF3B expression by SZRD1

NMD has been described to participate in several biological processes, such as cellular stress responses, neuronal differentiation, or development (9, 82, 83) by modulating the expression of specific transcripts. Thus, it is tempting to speculate that regulation of NMD activity is important in some contexts. So far, very little is known about how NMD is regulated. The recruitment of some proteins in the 3'UTR of mRNA has been described to block NMD (34–36). Furthermore, some *cis*-elements in the 3'UTR of mRNA confer some resistance to NMD (37). Here, we add two proteins (ARPP21 and SZRD1) that can increase protein output from transcripts recognized by the NMD machinery, at least in part, by counteracting the translational inhibition occurring during NMD.

These observations are consistent with and expand the observations by Rehfeld *et al.*, who reported that recruitment of ARPP21 can counteract the effect of miR-128. While the authors attributed the effect of ARPP21 mainly to an increase in translation from these transcripts, they also observed an increase in mRNA levels (Fig. 5 in Ref. (56)). This indicates that recruitment of ARPP21 can increase protein production by increasing mRNA levels and translation. Our experiments indicate that the increase in translation might be particularly effective on transcripts that are targeted by NMD.

Recently, two studies provided evidence that orthologs of ARPP21 oppose NMD function. First, Kelliher *et al.* (57) found that the *N. crassa* homolog, PRD2, antagonizes the effect of NMD on the *ck-1a* transcript. Second, Cieřla *et al.* (58) showed that the *S. cerevisiae* homolog of ARPP21, Rbs1, interacts within the 3'UTR of several transcripts that are degraded by the UPF1 homolog. The biological role of these proteins depends on the transcripts that it regulates. For instance, PRD2

plays a role in the circadian rhythm in *N. crassa* by increasing *ck-1a* levels, whereas Rbs1 suppresses assembly defects of polymerase III in a yeast mutant by increasing levels of the protein Rpb10 (57, 58). These publication and our data suggest that ARPP21, as well as its orthologs and paralogs, might play an evolutionary conserved role in modulating NMD.

The situation is more complex in the case of SZRD1. Our data indicate that recruitment to transcripts can alleviate the translational inhibition occurring during NMD, suggesting that it can counteract the effects of NMD. In contrast, knockdown of SZRD1 led to a dramatic reduction of the NMD machinery component UPF3B (Fig. 7, A–J), which might be responsible for the upregulation of a subset of NMD targets (Fig. 7, M and N) (32, 33).

Future mechanistic studies will likely reveal how these seemingly opposing effects are achieved, and in which biological context they might be important. Nevertheless, our data give first insights into the complex role of mammalian SUZ domain-containing proteins in NMD and open the field for future work.

Experimental procedures

Sequences for primers used for cloning are given in Table S1; antibodies are listed in Table S2; plasmids are listed in Table S3; and qPCR primers are listed in Table S4. The nucleotide sequences of the inserts in all plasmids generated in this study were verified by Sanger sequencing.

Generation of expression constructs

The ORF of SZRD1 was amplified using complementary DNA (cDNA) from HCT116 cells as template with the primers *SZRD1long-FW* and *SZRD1-RV* for the long forms of SZRD1 (corresponding to NP_001108072.1 [+CAG] or NP_001258798.1 [–CAG]), and the primers *SZRD1short-FW* and *SZRD1-RV* for the short forms of SZRD1 (corresponding to AAH23988.1 [+CAG] or BAC77397.1 [–CAG]) to generate constructs driving the expression of proteins with a C-terminal SFB tag. The resulting PCR products were digested with the restriction endonucleases *NheI* and *BsrGI* and inserted into the corresponding sites of the lentiviral vector pGTB5002 (Table S3) in frame with an SFB tag (61, 84). Constructs driving the expression of untagged versions of different SZRD1 forms were generated by insertion of PCR products with endogenous stop codons obtained with the reverse primer *SZRD1-STOP-RV* into the same vector. Mutant forms of SZRD1 were obtained by PCR-driven overlap extension (85). The flanking primers were the same used to amplify the wildtype forms, and the internal primers are listed in Table S1.

The ORF of human UPF1 was amplified from cDNA obtained from HCT116 cells using the primers *UPF1-FW* and *UPF1-RV* (corresponding to NP_002902.2). The resulting PCR product was digested and inserted into the vector pGTB5002 using the restriction enzymes *NheI* and *BsrGI*. The ORF of mouse STRAP was amplified from mouse spleen cDNA with the primers *STRAP-FW* and *STRAP-RV* (corresponding to

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

NP_035629.2). The ORF of mouse ARPP21 was amplified from an EST clone (ID: 5686446; GE Healthcare) using the primers *ARPP21-FW* and *ARPP21-RV*, respectively. The resulting ARPP21 PCR product lacked part of the SUZ because of the alternative splicing event and represents a minor splice variant (according to a large-scale RNA-Seq dataset) (59). An expression construct for the most common isoform (corresponding to NP_001348953.1) was generated by inserting the missing exon by Gibson assembly using the gene block gb-ARPP21SUZ. The PCR products obtained for the STRAP and ARPP21 ORFs were digested with NheI and Acc65I and inserted into the NheI and BsrGI sites of the vector pGTB5002.

Vectors used for transient expression of λ N-tagged proteins were based on the pCDNA3.1 vector (Life Technologies). The λ N tag was inserted by ligating the annealing oligos λ N-FW and λ N-RV into the NheI site resulting in the plasmid pJG211. Then, the SZRD1 long ORF was taken from pAH45, using the restriction enzymes NheI–BsrGI and inserted into pJG211 plasmid. The same procedure was followed to insert the mutated forms of SZRD1 that are described previously. ARPP21 was amplified from pMH117 using the primers λ NARPP21-FW and λ NARPP21-RV. For the amplification of the ARPP21 ORF containing the deletion in the SUZ domain, we used the plasmid pOH463 (Table S3) as a template and amplified the ORF with the same primers as for full-length ARPP21. The PCR products were digested by NheI and Acc65I and inserted into corresponding sites in pJG211.

Generation of lentiviral shRNA constructs

To express shRNAs, we used the plasmid pLVXpuro (Clontech), in which we inserted an enhanced GFP expression cassette followed by an shRNA expression cassette that closely resembles the improved miR-30 expression scaffold miRE (86) to generate the plasmid pJG97. shRNA sequences were identified using the splashRNA tool (87). All shRNA cassettes were amplified from synthetic oligonucleotides (IDT) using the primers *shRNA-FW* and *shRNA-RV* and inserted into the plasmid pJG97 via the XhoI and EcoRI sites. A nonsilencing control shRNA was shuttled from the vector pGIPZ-negative (Open Biosystems). The sequences of oligonucleotide templates are given in Table S1. With regard to SZRD1 shRNAs, #1 and #2 target the 3'UTR and therefore reduce all SZRD1 isoforms. In contrast, shRNAs #3 to #6 target exon 2, which is present only in the long isoforms of SZRD1. To generate inducible shRNA expression constructs, shRNAs were amplified from the synthetic oligonucleotide corresponding to shRNA #1 (CAGTGCCAGCAATAACAGT, Table S1) or from the vector pGIPZ-negative with the primers *miR30PCRX-hoIF_s* and *miR30PCREcoRIF_rev* and inserted into the XhoI and EcoRI sites of the plasmid pTRIPZ-empty (Open Biosystems).

Generation of reporter gene constructs

A reporter construct driving expression of a *Renilla* luciferase gene fused on its C terminus with part of the wildtype

human beta-globin sequence (*Renilla*-HBB wt) or a version containing a premature stop codon (*Renilla*-Hbb NS39) was a kind gift from Jana Loeber and Andreas Kulozik (71). Reporter constructs that allow recruitment of λ N-tagged proteins contain destabilizing elements, and/or miRNA-binding sites were produced similar to what is described by Gehring *et al.* (70). Briefly, a synthetic linker (BoxB adaptor) was inserted using the NotI site. A total of six BoxB sites were inserted by three consecutive insertions of the *BoxB genblock* (containing two BoxB sites) into a single MluI site. In addition, a poly A/U-rich site sequence (ARE) (*ARE-MluI*) or its control (*AREctrl-MluI*) were inserted downstream of the BoxB sites using the MluI restriction enzyme. To assess the effects of miRNAs, four complementary or seed-mutated let-7a-binding sites were inserted in the same location as the ARE sequence. These sequences were obtained by PCR from a plasmid kindly provided by Thomas Michiels from UCLouvain (unpublished data) using the primers *pADC-miRs-FW* and *pADC-miRs-RV*.

Insertion of an SFB tag into the UPF1, eIF4A3, and RNPS1 loci (“knock-in”)

To generate CRISPR–Cas9 constructs cutting immediately downstream of the ORFs, we inserted annealed primer pairs *UPF1-guide*, *RNPS1-guide*, or *eIF4A3-guide* into the BbsI site of the vector pX330 (Addgene #42230, a kind gift from Feng Zheng, MIT) (88).

To select for correct insertions, we generated a cassette consisting of an SFB-tag sequence, a self-cleaving P2A site, and a puromycin-resistant gene, similar to what has been described by Sheridan and Bentley (89), into the pGolden-adenovirus-associated virus (AAV) vector (Addgene #51424, a kind gift from Younlgun Luo) (90) generating the vector pJG180. Homology arms containing the sequences immediately upstream and downstream of the stop codons were amplified from genomic DNA of the U2OS cell line with the primer pairs indicated in Table S1 (*UPF1HL-FW* and *UPF1HL-RV* for the left homology arm of UPF1; *UPF1HR-FW* and *UPF1HR-RV* for the right homology arm of UPF1; *RNPS1HL-FW* and *RNPS1HL-RV* for the left homology arm of RNPS1; *RNPS1HR-FW* and *RNPS1HR-RV* for the right homology arm of RNPS1; *eIF4A3HL-FW* and *eIF4A3HL-RV* for the left homology arm of eIF4A3; and *eIF4A3HR-FW* and *eIF4A3HR-RV* for the right homology arm of eIF4A3). The selection cassette was amplified with the primers *SFB-P2A-puro-FW* and *SFB-P2A-puro-RV* using the vector pJG180 as a template. Subsequently, homology arms and selection cassettes were fused and inserted into the vector pGolden-AAV using Gibson assembly.

To generate helper-free recombinant AAV, we transiently transfected the aforementioned donor vectors together with the plasmids pHelper and pAAV-RC into HEK293 cells (Agilent; catalog no.: 240071) using calcium phosphate precipitation. After 6 h, the medium was changed, and 72 h later, medium and cells were recovered and viruses were liberated by three freeze–thaw cycles.

HEK293 cells were transfected with CRISPR–Cas9 constructs at 70% confluence in 6-well plates using Lipofectamine 2000 (2 µg of DNA for 6 µl of Lipofectamine 2000). After 6 h, the medium was changed, and half of the AAV supernatant was added. About 72 h later, cells were plated in two 10 cm dishes with 0.5 µg/ml of puromycin. Single-cell clones were isolated by limiting dilution. Insertion of the SFB tag was assessed by Western blot and by sequencing of the insertion site (Fig. S4, A and B).

Cell culture and lentiviral transduction

HEK293, HEK293-T, U2OS, and HeLa cells were cultured in Dulbecco's modified Eagle's medium containing 4.5 g/l D-glucose, 10% fetal calf serum, 2 mM ultraglutamine I (Lonza), and 100 U/ml penicillin–streptomycin (Lonza) at 37 °C and 5% CO₂. Jurkat cells were cultured in RPMI1640 medium (Lonza; catalog no.: 12-702F), 10% fetal calf serum, 2 mM ultraglutamine I, and 100 U/ml penicillin–streptomycin at 37 °C and 5% CO₂. Jurkat cells were maintained between 0.2 and 1.5 million cells/ml. To inhibit proteasomal degradation, cells were treated with 10 µM MG132 (Millipore; catalog no.: 133407-82-6) for the indicated times.

To generate recombinant lentiviruses used for the stable overexpression or knockdown of proteins, we transiently transfected HEK293T cells with lentiviral vectors and packaging plasmids using calcium phosphate precipitation (91). Briefly, HEK293T cells were transfected at 80% confluence in 6 cm dishes with 8.4 µg of psPAX2 (Addgene #12260) and 4.2 µg of pMD2.G (Addgene #12259), both kind gifts from Didier Trono (University of Geneva), together with 8.4 µg of the lentiviral vectors described previously. Medium was changed 6 h after transfection, and 24 h later, tissue culture supernatants containing viruses were harvested, filtered through a 0.4 µm filter, and diluted to infect target cells in the presence of polybrene (8 µg/ml; Sigma). After 24 h, cells were subcultured, and infected cells were selected with 2 µg/ml of puromycin for 2 days or until all cells on a noninfected control plate were dead. In cases where two lentiviruses were used to infect cells, vectors were generated containing a hygromycin resistance cassette but were otherwise identical. This allowed a parallel selection with 300 µg/ml of hygromycin and 2 µg/ml puromycin.

Affinity purification via streptavidin-coated beads

Cells growing in 10 cm plates were washed twice with ice-cold PBS and harvested in 300 µl NETN lysis buffer (150 mM NaCl, 1 mM EDTA, 20 mM Tris–HCl [pH 8.0], and 1% NP-40) containing phosphatase inhibitors (10 mM sodium fluoride, 1 mM sodium orthovanadate [pH 9.0], 2 mM sodium pyrophosphate, and 2 mM β-glycerophosphate) and protease inhibitors (cOmplete; Roche). Lysates were incubated on ice for 10 min and briefly sonicated to shear genomic DNA. Subsequently, lysates were clarified by centrifugation at 27,000g for 30 min at 4 °C, and the supernatants were recovered. We measured protein concentrations with the bicinchoninic acid assay and incubated equal amounts of proteins with

streptavidin-coupled Sepharose beads (GE Healthcare; catalog no.: 17-5113-01) or magnetic beads (Thermo Fisher; catalog no.: 88817; Fig. 3, C–F), which had been pre-equilibrated with NETN buffer. Samples were incubated on a rotating wheel for at least 1 h at 4 °C. Approximately 10 µl of beads were used per milligram of protein. Where indicated, 5 µl of RNase A (Thermo Scientific; EN0531) and 5 µl of RNase T1 (Thermo Scientific; EN0541) per milliliter of NETN buffer were added during the pulldown.

Beads were washed three times by adding 750 µl of NETN buffer and separated from the supernatant by centrifugation at 500g for 3 min at 4 °C or by exposure to a strong magnet for 30 s. Retained proteins were removed with denaturing loading dye and heated for 5 min at 85 °C. As an input control, a fixed fraction of the amount used in the pulldown was processed in parallel, representing between 5% and 12% as specified in the figure legends. Initial experiments were performed to identify interacting partners of SZRD1 followed by a tandem affinity purification protocol (61), where proteins retained on the streptavidin beads were eluted with 1 mg/ml biotin and loaded onto S-protein-coated agarose beads (Novagen; catalog no.: 69704). After elution with denaturing loading dye and gel electrophoresis, gels were stained with Coomassie brilliant blue G250 (PageBlue, Thermo Scientific), bands were excised and proteins were identified after in-gel digestion with trypsin using an LTQ mass spectrometer (92).

Reporter gene assays

Cells were plated 1 day before transfection in 24-well plates at 75,000 cells per well for HEK293T cells and U2OS cells and 35,000 cells per well for HeLa cells. Transfections were performed using a total of 200 ng of DNA per well with 1 µl of Lipofectamine 2000 following the manufacturer's protocol. The total amount of DNA consisted of 50 ng of the firefly luciferase control plasmid pCLneoFirefly, 50 ng of the *Renilla* luciferase constructs derived from the plasmid pCLneoRenilla (wt or NS39, Fig. 3), and 100 ng of the constructs driving the indicated effector proteins or empty vector controls. When overexpressing tethered SZRD1 or ARPP21 with a λN tag, we used 10 ng of the pDNA3.1 constructs driving expression of these proteins and 90 ng of the empty vector (pJG211). In the case of λN-tagged R3HDM1, we used 100 ng of the construct since R3HDM1 overexpression was undetectable at lower levels. After 24 h, cells were harvested and analyzed using a commercial dual luciferase kit (Dual-luciferase; Promega, catalog no.: E1960) with the help of a luminescence plate reader (GloMax Discover; Promega). To assess relative expression levels of endogenous proteins and overexpressed proteins, luciferase reporter assay lysates were diluted two times with water, heated after addition of 1/6 volume of 6× concentrated denaturing loading buffer (0.415 M of SDS; 0.9 mM of bromophenol blue; 60 mM of Tris [pH 6.8]; 47% of glycerol, and 0.6 M of DTT) (93) for 10 min at 72 °C, and analyzed by Western blot as described later.

When luciferase mRNA levels were analyzed, a set of parallel transfections was performed in 12-well plates using two

Effects of SZRD1 and ARPP21 on nonsense-mediated decay

times more of each component for the transfection. RNA purification, RT, and qPCR were performed as described later.

When miRNA mimics were cotransfected in reporter assays, 35,000 HeLa cells per well of a 24-well plate were plated 1 day before transfection. Transfection conditions were similar to the ones described previously (*i.e.*, 50 ng pGL3-control plasmid with or without SZRD1 3'UTR, 50 ng of an SV40 promoter-driven *Renilla* luciferase construct [Promega], and 100 ng of an empty expression vector [pJG211] to reach 200 ng) except for the addition of the indicated mimics at a final concentration of 3.3 nM.

RNA extraction, RT, and qPCR

RNA extraction, RT-PCR, and qPCR were performed as described (94). In brief, cells were harvested with TRIzol (Life Technologies; catalog no.: 15596018) and RNA was extracted according to the manufacturer's recommendations. Subsequently, 1.2 µg total RNA was reverse transcribed with RevertAid reverse transcriptase (Fermentas) and random hexamer primers in a total volume of 20 µl (94). The resulting cDNAs were diluted 10-fold, and 5 µl were used to perform a qPCR with a commercial SYBR Green Master Mix (Takyon mix; Eurogentec) according to the protocol given by the manufacturer. Results were analyzed using the delta-delta Ct method (95). Primers used for qPCR are listed in Table S4, and amplicons comprise exon-exon junctions unless indicated otherwise.

To quantify miR-128, the hairpin-based TaqMan MicroRNA assays from "Applied Biosystems" (kit: P02840360) and the corresponding RT kit were used.

Western blot analysis

Western blots were performed as described previously (92). In brief, cells were harvested using radioimmunoprecipitation assay buffer or NETN buffer as described previously. Cell lysates were sonicated and clarified by centrifugation for 30 min at 27,000g at 4 °C. Protein concentrations were determined using the bicinchoninic acid assay. Equal protein amounts were loaded on 10% or 12% Bis-Tris polyacrylamide gels. Electrophoresis was performed in Mops buffer (50 mM Mops, 50 mM Tris, 1 mM EDTA [pH 8.0], and 0.1% of SDS). Proteins were transferred onto polyvinylidene difluoride membranes (Immobilon P; Millipore) using a tank transfer system (Bio-Rad), and membranes were blocked with 5% nonfat dry milk (Regilait) in T-TBS buffer (0.05% Tween-20 in 20 mM Tris base [pH 7.2], and 150 mM NaCl). Incubation with the primary antibody was performed using T-TBS buffer containing 2% bovine serum albumin (Fisher Scientific) overnight at 4 °C using the concentrations indicated in Table S2. Horseradish peroxidase-coupled secondary antibodies against mouse (catalog no.: A5278; Sigma; 1/10,000 dilution) or rabbit (catalog no.: NA934V; GE Healthcare; 1/20,000 dilution) immunoglobulins were diluted in 5% nonfat dry milk T-TBS solution and incubated for 1 h at room temperature. Washing steps after first and secondary antibodies were performed with T-TBS solution three times for

5 min at room temperature. Signals were detected using the chemiluminescence reagent (WBKLS0500; Millipore) and autoradiography films (Fuji X-ray film) or a digital image acquisition system (catalog no.: LAS4000; GE Healthcare). Quantification was performed on nonsaturated images and, where indicated, by comparison with dilution series of lysates.

RNA-Seq analysis

For overexpression experiments, HeLa cells were infected with a lentivirus driving expression of the long SZRD1 isoform (with the addition of a serine residue, pAH45), human ARPP21, or an empty control lentivirus. After 24 h, cells were divided into medium containing 2 µg/ml puromycin. The medium was changed to medium without antibiotic 24 h later. Another 24 h later, two plates were harvested with TRIzol for RNA extraction and NETN lysis buffer for protein extraction. For knockdown experiments, we transduced HeLa cells with a recombinant lentivirus driving doxycyclin-inducible expression of SZRD1 shRNA #1 from the backbone pTRIPZ—empty (Open Biosystems). As a negative control, we used constructs expressing a nonsilencing control shRNA (pTRIPZ-negative; Open Biosystems). Expression of shRNAs was induced by addition of doxycyclin for 60 h before collection of cells in TRIzol.

After RNA extraction, 15 µg of RNA were incubated with 1 µl of DNase (RQ1 RNase-Free DNase; catalog no.: M6101; Promega) for 30 min at 37 °C with the indicated buffer, and RNA was cleaned up using the Qiagen RNeasy kit (Qiagen; catalog no.: 74104). Purity and concentrations were determined with an Agilent 2100 bioanalyzer, and 1 µg was sent for RNA-Seq analysis (after polyA selection) to Beckman Coulter Genomics.

The quality of the raw reads was assessed with FastQC (version 0.11.8) (96). The reads were trimmed of the universal adaptor with Trimmomatic (version 0.36) (97) and aligned using hisat2 (version 2.1.0) (98) to the Genome Reference Consortium Human Build 38 reference assembly (99). The resulting aligned reads were sorted and indexed using Samtools (version 1.9) (100), and counts were then generated using htseq-count (version 0.11.1) (101). The RNA-Seq count data were subsequently analyzed using the Bioconductor package (102) DESeq2 (version 1.26.0) (103), and volcano plots were generated with the EnhancedVolcano R package (<https://github.com/kevinblighe/EnhancedVolcano>). Transcripts were sorted according to the *t*-statistic, and a gene set enrichment analysis was performed using the fgsea package (104) using a list of human NMD taken from the study (7) as well as the molecular signature database (77) (Tables S5–S7).

Data availability

RNA-Seq data are available *via* ArrayExpress under the accession number E-MTAB-9627.

Supporting information—This article contains supporting information (7).

Acknowledgments—We thank Andreas Kulozik and Jana Loeber (University of Heidelberg, Heidelberg, Germany), Feng Zheng (MIT, Cambridge, MA, USA), Younglun Luo (Aarhus University, Aarhus, Denmark), and Didier Trono (University of Geneva, Geneva, Switzerland) for plasmids, which were in part obtained *via* Addgene as specified in the [Experimental procedures](#) section. We thank Emile Van Schaftingen for continuous support.

Author contributions—M. H., M. B., A. H., I. G., L. G., and G. T. B. conceptualization; T. K. and D. V. methodology; M. H., M. B., A. H., I. G., J. G., T. K., D. V., and G. T. B. investigation; M. H. and G. T. B. writing—original draft; M. B., A. H., I. G., J. G., T. K., L. G., and D. V. writing—review & editing; L. G. and G. T. B. supervision; G. T. B. project administration; G. T. B. funding acquisition.

Funding and additional information—This work was supported by the Fonds National de la Recherche Scientifique (FRIA; to M. H.; MIS, CDR, and WELBIO; to G. T. B.), Université Catholique de Louvain (FSR; to G. T. B.), the European Research Council (#771704 NoMePaCa; to G. T. B.), Fonds Maisin (to G. T. B.), Fondation Contre Le Cancer (to G. T. B.), and Televie (to M. B.). Funding for open access charge: Fonds National de la Recherche Scientifique and Fonds Maisin.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AAV, adeno-associated virus; C1QBP, complement component 1Q subcomponent-binding protein; cDNA, complementary DNA; eIF4A3, eukaryotic translation initiation factor 4A3; EJC, exon junction complex; eRF3, eukaryotic release factor 3; HEK293, human embryonic kidney 293 cell line; NETN, NaCl, EDTA, Tris-HCl, NP-40; NMD, nonsense-mediated decay; PABPC1, polyA-binding protein cytoplasmic 1; qPCR, quantitative PCR; R3HDM, R3H domain-containing protein; STRAP, serine/threonine kinase receptor-associated protein; SZRD1, SUZ domain-containing protein 1; SZY-20, suppressor of zyg-1 protein 20; T-TBS, Tween-20 in 20 mM Tris base; UPF, up-frameshift.

References

1. Losson, R., and Lacroute, F. (1979) Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5134–5137
2. Leeds, P., Peltz, S. W., Jacobson, A., and Culbertson, M. R. (1991) The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes Dev.* **5**, 2303–2314
3. Chang, J. C., Temple, G. F., Trecartin, R. F., and Kan, Y. W. (1979) Suppression of the nonsense mutation in homozygous beta 0 thalassemia. *Nature* **281**, 602–603
4. Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A., and Ross, J. (1981) Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. *Cell* **27**, 543–553
5. Buhler, M., Steiner, S., Mohn, F., Paillusson, A., and Muhlemann, O. (2006) EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* **13**, 462–464
6. Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z., and Muhlemann, O. (2008) Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol.* **6**, e92
7. Colombo, M., Karousis, E. D., Bourquin, J., Bruggmann, R., and Muhlemann, O. (2017) Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA* **23**, 189–201
8. Karousis, E. D., Nasif, S., and Muhlemann, O. (2016) Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley interdisciplinary reviews. RNA* **7**, 661–682
9. Nasif, S., Contu, L., and Muhlemann, O. (2018) Beyond quality control: the role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. *Semin. Cell Dev. Biol.* **75**, 78–87
10. Kurosaki, T., Popp, M. W., and Maquat, L. E. (2019) Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.* **20**, 406–420
11. Le Hir, H., Sauliere, J., and Wang, Z. (2016) The exon junction complex as a node of post-transcriptional networks. *Nat. Rev. Mol. Cell Biol.* **17**, 41–54
12. Lejeune, F., Ishigaki, Y., Li, X., and Maquat, L. E. (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.* **21**, 3536–3545
13. Schlautmann, L. P., and Gehring, N. H. (2020) A day in the life of the exon junction complex. *Biomolecules* **10**, 866
14. Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., *et al.* (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
15. Hug, N., and Caceres, J. F. (2014) The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell Rep.* **8**, 1845–1856
16. Melero, R., Hug, N., Lopez-Perrote, A., Yamashita, A., Caceres, J. F., and Llorca, O. (2016) The RNA helicase DHX34 functions as a scaffold for SMG1-mediated UPF1 phosphorylation. *Nat. Commun.* **7**, 10585
17. Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008) NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* **15**, 85–93
18. Lopez-Perrote, A., Castano, R., Melero, R., Zamarró, T., Kurosawa, H., Ohnishi, T., *et al.* (2016) Human nonsense-mediated mRNA decay factor UPF2 interacts directly with eRF3 and the SURF complex. *Nucleic Acids Res.* **44**, 1909–1923
19. Boehm, V., Haberman, N., Ottens, F., Ule, J., and Gehring, N. H. (2014) 3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. *Cell Rep.* **9**, 555–568
20. Metzger, S., Herzog, V. A., Ruepp, M. D., and Muhlemann, O. (2013) Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. *RNA* **19**, 1432–1448
21. Karousis, E. D., Gutzler, L. A., Annibaldis, G., Dreos, R., and Muhlemann, O. (2020) Human NMD ensues independently of stable ribosome stalling. *Nat. Commun.* **11**, 4134
22. Plant, E. P., Wang, P., Jacobs, J. L., and Dinman, J. D. (2004) A programmed -1 ribosomal frameshift signal can function as a cis-acting mRNA destabilizing element. *Nucleic Acids Res.* **32**, 784–790
23. Hoek, T. A., Khuperkar, D., Lindeboom, R. G. H., Sonneveld, S., Verhagen, B. M. P., Boersma, S., *et al.* (2019) Single-molecule imaging uncovers rules governing nonsense-mediated mRNA decay. *Mol. Cell* **75**, 324–339.e311
24. Isken, O., Kim, Y. K., Hosoda, N., Mayeur, G. L., Hershey, J. W., and Maquat, L. E. (2008) Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell* **133**, 314–327
25. Miller, J. E., and Reese, J. C. (2012) Ccr4-Not complex: the control freak of eukaryotic cells. *Crit. Rev. Biochem. Mol. Biol.* **47**, 315–333
26. Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C. T., *et al.* (2014) A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Mol. Cell* **54**, 737–750
27. Collier, J., and Parker, R. (2005) General translational repression by activators of mRNA decapping. *Cell* **122**, 875–886
28. Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., *et al.* (2014) Structural and biochemical insights to the role of

- the CCR4-NOT complex and DDX6 ATPase in microRNA repression. *Mol. Cell* **54**, 751–765
29. Jonas, S., and Izaurralde, E. (2015) Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **16**, 421–433
30. Ferraiuolo, M. A., Basak, S., Dostie, J., Murray, E. L., Schoenberg, D. R., and Sonenberg, N. (2005) A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *J. Cell Biol.* **170**, 913–924
31. Zhang, Z., Zhou, L., Hu, L., Zhu, Y., Xu, H., Liu, Y., *et al.* (2010) Nonsense-mediated decay targets have multiple sequence-related features that can inhibit translation. *Mol. Syst. Biol.* **6**, 442
32. Gehring, N. H., Kunz, J. B., Neu-Yilik, G., Breit, S., Viegas, M. H., Hentze, M. W., *et al.* (2005) Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol. Cell* **20**, 65–75
33. Chan, W. K., Huang, L., Gudikote, J. P., Chang, Y. F., Imam, J. S., MacLean, J. A., 2nd, *et al.* (2007) An alternative branch of the nonsense-mediated decay pathway. *EMBO J.* **26**, 1820–1830
34. Ge, Z., Quek, B. L., Beemon, K. L., and Hogg, J. R. (2016) Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. *Elife* **5**, e11155
35. Kishor, A., Ge, Z., and Hogg, J. R. (2019) hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. *EMBO J.* **38**, e99128
36. Fritz, S. E., Ranganathan, S., Wang, C. D., and Hogg, J. R. (2020) The RNA-binding protein PTBP1 promotes ATPase-dependent dissociation of the RNA helicase UPF1 to protect transcripts from nonsense-mediated mRNA decay. *J. Biol. Chem.* **295**, 11613–11625
37. Toma, K. G., Rebbapragada, I., Durand, S., and Lykke-Andersen, J. (2015) Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay. *RNA* **21**, 887–897
38. Yepiskoposyan, H., Aeschmann, F., Nilsson, D., Okoniewski, M., and Muhlemann, O. (2011) Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* **17**, 2108–2118
39. Huang, L., Lou, C. H., Chan, W., Shum, E. Y., Shao, A., Stone, E., *et al.* (2011) RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol. Cell* **43**, 950–961
40. Longman, D., Hug, N., Keith, M., Anastasaki, C., Patton, E. E., Grimes, G., *et al.* (2013) DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and *Caenorhabditis elegans*. *Nucleic Acids Res.* **41**, 8319–8331
41. Joncourt, R., Eberle, A. B., Rufener, S. C., and Muhlemann, O. (2014) Eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay by two genetically separable mechanisms. *PLoS One* **9**, e104391
42. Fatscher, T., Boehm, V., Weiche, B., and Gehring, N. H. (2014) The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. *RNA* **20**, 1579–1592
43. Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007) A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J.* **26**, 1591–1601
44. Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W., and Kulozik, A. E. (2008) Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* **27**, 736–747
45. Kervestin, S., Li, C., Buckingham, R., and Jacobson, A. (2012) Testing the faux-UTR model for NMD: analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p. *Biochimie* **94**, 1560–1571
46. Silva, A. L., Ribeiro, P., Inacio, A., Liebhauer, S. A., and Romao, L. (2008) Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. *RNA* **14**, 563–576
47. Matsuda, A., Suzuki, Y., Honda, G., Muramatsu, S., Matsuzaki, O., Nagano, Y., *et al.* (2003) Large-scale identification and characterization of human genes that activate NF- κ B and MAPK signaling pathways. *Oncogene* **22**, 3307–3318
48. Zhao, N., Zhang, G., He, M., Huang, H., Cao, L., Yin, A., *et al.* (2017) SZRD1 is a novel protein that functions as a potential tumor suppressor in cervical cancer. *J. Cancer* **8**, 2132–2141
49. Geng, Y. B., Pan, C. C., Xu, C., Zuo, P. C., Wang, Y., Li, X. O., *et al.* (2020) Long non-coding RNA LINC00346 regulates proliferation and apoptosis by targeting miR-128-3p/SZRD1 axis in glioma. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 9581–9590
50. Song, M. H., Aravind, L., Muller-Reichert, T., and O'Connell, K. F. (2008) The conserved protein SZY-20 opposes the Plk4-related kinase ZYG-1 to limit centrosome size. *Dev. Cell* **15**, 901–912
51. Carissimi, C., Baccon, J., Straccia, M., Chiarella, P., Maiolica, A., Sawyer, A., *et al.* (2005) Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett.* **579**, 2348–2354
52. Hunt, S. L., Hsuan, J. J., Totty, N., and Jackson, R. J. (1999) unr, a cellular cytoplasmic RNA-binding protein with five cold-shock domains, is required for internal initiation of translation of human rhinovirus RNA. *Genes Dev.* **13**, 437–448
53. Vukmirovic, M., Manojlovic, Z., and Stefanovic, B. (2013) Serine-threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. *Mol. Cell Biol.* **33**, 3893–3906
54. Grishin, N. V. (1998) The R3H motif: a domain that binds single-stranded nucleic acids. *Trends Biochem. Sci.* **23**, 329–330
55. He, G. J., Zhang, A., Liu, W. F., and Yan, Y. B. (2013) Distinct roles of the R3H and RRM domains in poly(A)-specific ribonuclease structural integrity and catalysis. *Biochim. Biophys. Acta* **1834**, 1089–1098
56. Rehfeld, F., Maticzka, D., Grosser, S., Knauff, P., Eravci, M., Vida, I., *et al.* (2018) The RNA-binding protein ARPP21 controls dendritic branching by functionally opposing the miRNA it hosts. *Nat. Commun.* **9**, 1235
57. Kelliher, C. M., Lambregts, R., Xiang, Q., Baker, C. L., Loros, J. J., and Dunlap, J. C. (2020) PRD-2 directly regulates casein kinase I and counteracts nonsense-mediated decay in the *Neurospora* circadian clock. *Elife* **9**, e64007
58. Ciesla, M., Turowski, T. W., Nowotny, M., Tollervey, D., and Boguta, M. (2020) The expression of Rpb10, a small subunit common to RNA polymerases, is modulated by the R3H domain-containing Rbs1 protein and the Upf1 helicase. *Nucleic Acids Res.* **48**, 12252–12268
59. Consortium, G. T. (2013) The genotype-tissue expression (GTEx) project. *Nat. Genet.* **45**, 580–585
60. Geng, L., Huntoon, C. J., and Karnitz, L. M. (2010) RAD18-mediated ubiquitination of PCNA activates the Fanconi anemia DNA repair network. *J. Cell Biol.* **191**, 249–257
61. Kim, H., Chen, J., and Yu, X. (2007) Ubiquitin-binding protein RAP80 mediates BRCA1-dependent DNA damage response. *Science* **316**, 1202–1205
62. Sato, H., and Maquat, L. E. (2009) Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. *Genes Dev.* **23**, 2537–2550
63. Bai, Y., Wang, W., Li, S., Zhan, J., Li, H., Zhao, M., *et al.* (2019) C1QBP promotes homologous recombination by Stabilizing MRE11 and controlling the assembly and activation of MRE11/RAD50/NBS1 complex. *Mol. Cell* **75**, 1299–1314.e1296
64. McGee, A. M., Douglas, D. L., Liang, Y., Hyder, S. M., and Baines, C. P. (2011) The mitochondrial protein C1qbp promotes cell proliferation, migration and resistance to cell death. *Cell Cycle* **10**, 4119–4127
65. Petersen-Mahrt, S. K., Estmer, C., Ohrmalm, C., Matthews, D. A., Russell, W. C., and Akusjarvi, G. (1999) The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J.* **18**, 1014–1024
66. Yoshikawa, H., Ishikawa, H., Izumikawa, K., Miura, Y., Hayano, T., Isobe, T., *et al.* (2015) Human nucleolar protein Nop52 (RRP1/NNP-1) is involved in site 2 cleavage in internal transcribed spacer 1 of pre-rRNAs at early stages of ribosome biogenesis. *Nucleic Acids Res.* **43**, 5524–5536
67. Zund, D., Gruber, A. R., Zavolan, M., and Muhlemann, O. (2013) Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nat. Struct. Mol. Biol.* **20**, 936–943
68. Hurt, J. A., Robertson, A. D., and Burge, C. B. (2013) Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res.* **23**, 1636–1650

69. Flury, V., Restuccia, U., Bachi, A., and Muhlemann, O. (2014) Characterization of phosphorylation- and RNA-dependent UPF1 Interactors by quantitative proteomics. *J. proteome Res.* **13**, 3038–3053
70. Gehring, N. H., Hentze, M. W., and Kulozik, A. E. (2008) Tethering assays to investigate nonsense-mediated mRNA decay activating proteins. *Methods Enzymol.* **448**, 467–482
71. Boelz, S., Neu-Yilik, G., Gehring, N. H., Hentze, M. W., and Kulozik, A. E. (2006) A chemiluminescence-based reporter system to monitor nonsense-mediated mRNA decay. *Biochem. Biophys. Res. Commun.* **349**, 186–191
72. Zhang, J., Sun, X., Qian, Y., and Maquat, L. E. (1998) Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. *RNA* **4**, 801–815
73. Shaw, G., and Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46**, 659–667
74. Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L. E. (2010) UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. *Mol. Cell* **39**, 396–409
75. Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., et al. (2005) Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573–1576
76. Sieber, J., Hauer, C., Bhuvanagiri, M., Leicht, S., Krijgsvel, J., Neu-Yilik, G., et al. (2016) Proteomic analysis reveals branch-specific regulation of the Unfolded protein response by nonsense-mediated mRNA decay. *Mol. Cell. Proteomics* **15**, 1584–1597
77. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550
78. Gerbracht, J. V., Boehm, V., Britto-Borges, T., Kallabis, S., Wiederstein, J. L., Ciriello, S., et al. (2020) CASC3 promotes transcriptome-wide activation of nonsense-mediated decay by the exon junction complex. *Nucleic Acids Res.* **48**, 8626–8644
79. McGlincy, N. J., Tan, L. Y., Paul, N., Zavolan, M., Lilley, K. S., and Smith, C. W. (2010) Expression proteomics of UPF1 knockdown in HeLa cells reveals autoregulation of hnRNP A2/B1 mediated by alternative splicing resulting in nonsense-mediated mRNA decay. *BMC Genomics* **11**, 565
80. Tani, H., Torimura, M., and Akimitsu, N. (2013) The RNA degradation pathway regulates the function of GAS5 a non-coding RNA in mammalian cells. *PLoS One* **8**, e55684
81. Kurosaki, T., Hoque, M., and Maquat, L. E. (2018) Identifying cellular nonsense-mediated mRNA decay (NMD) targets: immunoprecipitation of phosphorylated UPF1 followed by RNA sequencing (p-UPF1 RIP-seq). *Methods Mol. Biol.* **1720**, 175–186
82. Goetz, A. E., and Wilkinson, M. (2017) Stress and the nonsense-mediated RNA decay pathway. *Cell. Mol. Life Sci.* **74**, 3509–3531
83. Han, X., Wei, Y., Wang, H., Wang, F., Ju, Z., and Li, T. (2018) Nonsense-mediated mRNA decay: a 'nonsense' pathway makes sense in stem cell biology. *Nucleic Acids Res.* **46**, 1038–1051
84. Gerin, I., Ury, B., Breloy, I., Bouchet-Seraphin, C., Bolsee, J., Halbout, M., et al. (2016) ISPD produces CDP-ribitol used by FKTN and FKR to transfer ribitol phosphate onto alpha-dystroglycan. *Nat. Commun.* **7**, 11534
85. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* **2**, 924–932
86. Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., et al. (2013) An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713
87. Pelossof, R., Fairchild, L., Huang, C. H., Widmer, C., Sreedharan, V. T., Sinha, N., et al. (2017) Prediction of potent shRNAs with a sequential classification algorithm. *Nat. Biotechnol.* **35**, 350–353
88. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823
89. Sheridan, R. M., and Bentley, D. L. (2016) Selectable one-step PCR-mediated integration of a degron for rapid depletion of endogenous human proteins. *Biotechniques* **60**, 69–74
90. Luo, Y., Lin, L., Bolund, L., and Sorensen, C. B. (2014) Efficient construction of rAAV-based gene targeting vectors by Golden Gate cloning. *Biotechniques* **56**, 263–268
91. Jordan, M., Schallhorn, A., and Wurm, F. M. (1996) Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res.* **24**, 596–601
92. Collard, F., Baldin, F., Gerin, I., Bolsee, J., Noel, G., Graff, J., et al. (2016) A conserved phosphatase destroys toxic glycolytic side products in mammals and yeast. *Nat. Chem. Biol.* **12**, 601–607
93. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
94. Gerin, I., Noel, G., Bolsee, J., Haumont, O., Van Schaftingen, E., and Bommer, G. T. (2014) Identification of TP53-induced glycolysis and apoptosis regulator (TIGAR) as the phosphoglycolate-independent 2,3-bisphosphoglycerate phosphatase. *Biochem. J.* **458**, 439–448
95. Livak, K. J., and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408
96. Steven W, W., and Simon, A. (2018) FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Res.* **7**, 1338
97. Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120
98. Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915
99. Schneider, V. A., Graves-Lindsay, T., Howe, K., Bouk, N., Chen, H. C., Kitts, P. A., et al. (2017) Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res.* **27**, 849–864
100. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The sequence alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079
101. Anders, S., Pyl, P. T., and Huber, W. (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169
102. Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., et al. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121
103. Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550
104. [preprint] Korotkevich, G., Sukhov, V., and Sergushichev, A. (2019) Fast gene set enrichment analysis. *bioRxiv*. <https://doi.org/10.1101/060012>