The Human Nuclear Exosome Targeting Complex Is Loaded onto Newly Synthesized RNA to Direct Early Ribonucleolysis

Graphical Abstract

Highlights
- RBM7 is loaded promiscuously onto newly synthesized RNA
- RBM7/NEXT defines an early nuclear RNA decay pathway
- RBM7/NEXT triggers exosome activity at snoRNA-encoding introns
- RBM7/NEXT assigns many IncRNAs for nuclear degradation

Authors
Michal Lubas, Peter Refsing Andersen, ..., Grzegorz Kudla, Torben Heick Jensen

Correspondence
thj@mb.au.dk

In Brief
The nuclear RNA exosome is directed to many of its nucleoplasmic substrates by the nuclear exosome targeting (NEXT) complex. Lubas et al. now use iCLIP to follow NEXT targeting to RNA via its RNA binding component RBM7. Association of RBM7 with RNAPII-derived RNAs enables degradation upon emergence of unprotected RNA 3′ ends. Thus, the NEXT complex defines an early exosome targeting pathway acting on newly synthesized RNA, including snoRNAs embedded in pre-mRNA introns.

Accession Numbers
GSE63791
GSE63496
GSE52132
GSE48286

Lubas et al., 2015, Cell Reports 10, 178–192
January 13, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2014.12.026
The Human Nuclear Exosome Targeting Complex Is Loaded onto Newly Synthesized RNA to Direct Early Ribonucleolyis

Michal Lubas, Peter Refsing Andersen, Aleks Schein, Andrzej Dziembowski, Grzegorz Kudla, and Torben Heick Jensen

INTRODUCTION

Nuclear RNA metabolism in eukaryotes depends on a large extent on the 3′-5′ exoribonucleytic and endoribonuclcytic activities of the RNA exosome complex (Chlebowski et al., 2013; Schneider and Tollervey, 2013). In vivo exosome activity is regulated by associated enzymatic cofactors and RNA-binding adapters, which facilitate exosome access and recruitment to a multitude of substrates. The best characterized cofactor of the nuclear exosome is the S. cerevisiae Trf4p/5p-Air1p/2p-Mtr4p polyadenylation (TRAMP) complex, containing either one of two oligo(A)-polymerases Trf4p/Trf5p and zinc finger proteins Air1p/Air2p, respectively, as well as the DEAD box RNA helicase Mtr4p (LaCava et al., 2005; Vanácová et al., 2005; Wyers et al., 2005). TRAMP is recruited to substrates by the exosome adaptor- and RNA-binding Nrd1p/Nab3p/Sen1p (NNS) protein complex (Arigo et al., 2006; Thiebaut et al., 2006; Vasiljeva and Buratowski, 2006). While TRAMP predominantly acts to promote exosomal decay by the oligoadenylation and unwinding of RNA targets, the RNA polymerase II (RNAPII)-binding NNS complex targets short sequence motifs in the nascent RNA to trigger RNAPII transcription termination and to couple this process with the handover of substrates with free 3′ ends to TRAMP (Arigo et al., 2006; Carroll et al., 2007; Steinmetz et al., 2001; Thiebaut et al., 2006; Tudek et al., 2014).

Although the principal inner workings of the nuclear exosome machinery appear similar in mammalian cells with the Mtr4p homolog, hMTR4/SKIV2L2, as a central helicase coupling the activity of the exosome to its RNA adapters, the spatial organization and nature of exosome accessory factors have diverged considerably from S. cerevisiae (Lubas et al., 2011; Sloan et al., 2012; Wolin et al., 2012). In mammals, two distinct exosome cofactors have so far been characterized, which are the nucleolar-localized hTRAMP complex (Fasken et al., 2011; Lubas et al., 2011) and the nucleoplasm-specific Nuclear Exosome Targeting (NEXT) complex (Lubas et al., 2011). Sharing hMTR4 with hTRAMP, the otherwise distinct trimeric NEXT complex is further composed of the RNA recognition motif (RRM)-containing RBM7 and the zinc-knuckle ZCCHC8 proteins. Consistent with their different nuclear localization, hTRAMP and NEXT complexes operate on different RNA substrates. While hTRAMP is involved in oligoadenylation of nucleolar rRNAs to facilitate their decay, depletion of NEXT subunits leads to the stabilization of nuclear exosome substrates such as PROMoter uPstream Transcripts (PROMPTs) and 3′-extended products from U1/U2 snRNA- and replication-dependent histone (RDH) genes (Andersen et al., 2013; Lubas et al., 2011; Ntini et al., 2013; Shcherbik et al., 2010).
the NEXT complex reaches its nucleoplasmic RNA targets is a central question as no NNS-homologous complex has been identified in metazoan systems.

A possible clue to NEXT-substrate targeting came recently when a robust interaction between NEXT and the RNA cap-binding complex (CBC) was revealed (Andersen et al., 2013; Hallassi et al., 2013). Specifically, all NEXT components individually copurify the CBC as well as the CBC-associated arsenic-resistance 2 (ARS2) and ZC3H18/NHN1 proteins, forming the so-coined CBC-NEXT (CBCN) protein assembly. This link between the CBC and NEXT provides a direct coupling between the capped 5' ends of RNAIIP transcripts and 3' end processing/degradation processes. A similar connection was reported between the S. cerevisiae NNS and CBC complexes (Vasiljeva and Buratowski, 2006); however, the extent to which this interaction guides the NNS complex to RNAIIP transcription units has not been further explored. Interestingly, a subcomplex of the CBCN assembly, the CBC-ARS2 (CBCA) complex, triggers the termination of PROMPT, U2, and RDI gene read-through transcription, whereas depletion of NEXT components RBM7 and ZCCHC8 shows only modest transcription phenotypes (Andersen et al., 2013; Hallassi et al., 2013). While this provides some mechanistic analogy to the S. cerevisiae system with its coupling of transcription termination and exosomal decay, it still leaves open the mechanism by which NEXT seeks out its targets, including how it organizes RNA binding with summoning of the exosome.

Here, we used an individual-nucleotide-resolution UV crosslinking and immunoprecipitation (iCLIP) procedure to achieve a transcriptome-wide view of in vivo RBM7-RNA interactions. Our results demonstrate that RBM7 associates broadly with RNAIIP-derived RNA, thereby enabling transcript degradation upon emergence of unprotected RNA 3' ends. We suggest that the NEXT complex defines an early exosome targeting pathway acting on newly synthesized RNA.

RESULTS

Global Overview of RBM7-RNA Interactions
To achieve a transcriptome-wide view of in vivo NEXT interactions, we performed iCLIP (Figure 1A) (Konig et al., 2011) to map RNA-binding sites of the RRM-containing (Figure S1A) and vertebrate-conserved (Figure S1B) RBM7 protein. Localization and affinity purification (LAP)-tagged RBM7 expressed at near-endogenous levels in HeLa cells (Andersen et al., 2013) crosslinked to RNA with an efficiency comparable to the RNA-binding protein hnRNPK (Figure S1C, lanes 2 and 3). Extensive treatment with either RNase I or RNase TI confirmed that the majority of the 32P-labeled RNA was attached to RBM7-LAP (Figure S1D), although its stable interaction partner, ZCCHC8, was also present in the IP eluates (Figure S1E, lanes 5–8). Notably, hMTR4 was absent after the stringent washing conditions employed, indicating a strong interaction between ZCCHC8 and RBM7. As expected, RBM7 only crosslinked to RNA in the presence of UV irradiation (Figure S1F). RBM7 iCLIP was carried out in two replicates, which were either treated with control luciferase siRNAs or siRNAs targeting the exosome core-component hRRP40 (Figure S1G) to enable visualization of exosome substrates and RBM7-binding targets in the same experimental setup. The nontagged control cell line did not generate detectable PCR products (Figure 1A), implying a low experimental background. For further comparative analysis, we included our previously published RNA sequencing (RNA-seq) datasets of total rRNA-depleted RNA obtained from HeLa cells treated with siRNAs targeting hRRP40, ZCCHC8, or EGFP (control) (Andersen et al., 2013).

To provide a first overview of the data, unique RBM7 iCLIP and corresponding RNA-seq reads from control and hRRP40-depleted samples were distributed into general genomic categories: pre-mRNA intron, long noncoding RNA (lncRNA), repetitive elements, pre-mRNA exon, short ncRNA (miRNA, snoRNA, tRNA, snRNA), unannotated rRNA, and “other” (Figure 1B). Within pre-mRNA, the majority of RBM7 iCLIP tags mapped to introns (37%–43%) and a smaller fraction to exons (8%–11%). The RNA-seq data showed a ~3-fold higher frequency of reads in exons (30%). Such relative enrichment of intronic reads in RBM7 iCLIP over RNA-seq data are in line with the nucleoplasmic localization of RBM7 (Lubas et al., 2011). Similar intron/exon distributions of iCLIP-tags were observed in studies of splicing-related (FUS, U2AF65, and SFSF3/4) (Ånkö et al., 2012; Rogelj et al., 2012; Zarnack et al., 2013), but not cytoplasmic (UPF1), RNA-binding proteins (Zünd et al., 2013). This indication of RBM7 binding, before or during pre-mRNA splicing, was supported by inspection of representative protein-coding genes (see Figure 1C, “red tracks” for the CKS2 and SECISBP2 genes, and further examples below).

Another large fraction of iCLIP tags mapped to a broadly defined collection of IncRNAs (GENCODE v18) (Figure 1B). Unique tag numbers increased upon exosome depletion from 14% (RBM7) to 20% (RBM7/hRRP40), likely reflecting a prolonged association between IncRNA transcripts and the NEXT complex in the absence of exosome-mediated IncRNA decay (Andersson et al., 2014b; Lubas et al., 2011). Consistent with this notion, the CKS2 and SECISBP2 loci also revealed evidence for a physical contact between RBM7 and the transcription start site (TSS)-proximal PROMPTs of these genes (Figure 1C, “blue” tracks), further validated by RNA IP (RIP) of two previously reported PROMPTs (Figure S1H).

Of the remaining genomic categories, “repetitive elements” constituted a noteworthy class of RBM7-bound transcripts of which comprehensive analyses fall beyond the scope of this study. Additionally, notable crosslinking of RBM7 to stable ncRNAs (tRNA, rRNA, snoRNA, and snRNA) was observed. Mature forms of these transcripts are highly abundant and offer binding sites for RNA binding proteins calling the specificity of the interaction into question. Binding to the unprocessed forms of short ncRNAs, specifically snRNA and snoRNA, is addressed in detail later (see below).

Finally, RBM7 iCLIP tags were found to harbor notably fewer nonencoded 3’-terminal adenines than that from CLIP tags of yeast exosome cofactors (~1% versus 6%–30%) (Tuck and Tollervey, 2013; Wlotzka et al., 2011). This likely reflects a requirement in S. cerevisiae for the posttranscriptional addition of 3’ adenosines to promote exosomal activity. In human cells, such a requirement mostly appears to be relevant in nucleoli (Lubas et al., 2011; Preker et al., 2011). Therefore, increased
3' adenylation of RBM7 iCLIP tags from lncRNA, snRNA, and mRNA intron fractions upon hRRP40 depletion (Figure S1I) probably arises from adenylation of nuclear targets incapable of turning over efficiently in the absence of exosome activity.

Taken together, these initial analyses confirmed that RBM7 physically contacts known nuclear RNA exosome substrates, validating the utilized RBM7 iCLIP procedure to reliably describe the detailed in vivo RNA interaction profile of RBM7.

**RBM7 Poises lncRNAs for Exosomal Decay via Initial CBC Contacts**

That RBM7/NEXT recruits the exosome to nuclear substrates has previously been suggested by the direct physical interaction of RBM7/NEXT with the exosome (Lubas et al., 2011), the upregulation of similar substrate classes upon exosome and NEXT depletion (Andersen et al., 2013; Lubas et al., 2011), and the interaction of CBCN with exosome substrates (Andersen et al., 2013). To obtain experimental support that transcripts physically interacting with RBM7 are ultimately degraded by the exosome, we analyzed iCLIP and RNA-seq data for (1) a set of 12153 PROMPT regions, (2) their neighboring genic regions, and (3) 10,014 HeLa cell enhancer RNAs (eRNAs) detectable within both the iCLIP and RNA-seq data as well as in data sets of RNA 5' ends generated by Cap Analysis of Gene Expression (CAGE) (Andersson et al., 2014b; Ntini et al., 2013). Like PROMPTs, eRNAs are capped, short, usually unspliced, and exosome sensitive (Andersson et al., 2014a, 2014b; De Santa et al., 2010; Kim et al., 2010). To address RBM7-binding, we normalized control HeLa RBM7 iCLIP- to RNA-seq-read coverage within 1 kb from the TSSs of the interrogated RNAs. This revealed a higher density of RBM7 on PROMPTs and eRNAs relative to genic transcripts (Figure 2A, left). Less relative binding of RBM7 to genic RNA is likely due to high levels of unbound cytoplasmic mRNA as our iCLIP data later revealed that RBM7 preferentially binds pre-mRNA (Figures 3A and 5B). As expected, PROMPTs and eRNAs were also generally

---

**Figure 1. iCLIP Reveals Direct In Vivo Targets of RBM7**

(A) Schematic overview of the iCLIP procedure. Control ("noTAG") and RBM7-LAP HeLa cells were transfected with luciferase (control [ctrl]) or hRRP40 siRNA and UV irradiated at 254 nm. Crosslinked RNA-protein complexes were purified, subjected to RNase I treatment, and analyzed by SDS-PAGE. Visualized by P32 autoradiography, RBM7-linked RNAs are indicated by a red line, and the black arrow denotes the RBM7-LAP molecular weight; #1 and #2 indicate the two biological replicates. Extracted RNA material was used for preparation of three size fractions of RT-PCR products (high, medium, low) and pooled prior to sequencing.

(B) Distribution charts of unique tags derived from the iCLIP (RBM7 and RBM7/hRRP40) and RNA-seq (ctrl and hRRP40) libraries and mapped to the indicated RNA classes (averages from the two biological repeats).

(C) Genome browser screenshot of RBM7 iCLIP and RNA-seq reads mapped to a representative genomic area encompassing TSS-proximal regions of two protein-coding genes (CKS2 and SECISBP2). Four RBM7 iCLIP (eight tracks) and three RNA-seq (six tracks) libraries are visualized by (–) (blue) and (+) (red) strands. Numbers to the right show the normalized maximum read count value.
stabilized upon hRRP40 (Figure 2A, middle) and ZCCHC8 (Figure 2A, right) depletions. This implies that RBM7 targeting of these transcripts results in their subsequent degradation by the RNA exosome. We next asked whether this correlation between RBM7 binding and exosome sensitivity could also be observed at the level of individual transcripts. To this end, we plotted RBM7 iCLIP enrichment of each transcript (the ratio of RBM7 iCLIP to control RNA-seq signal) against its sensitivity to exosome (hRRP40) or NEXT (ZCCHC8) depletion (the ratio of hRRP40 or ZCCHC8 RNA-seq to control RNA-seq signals) (Figure 2B). This showed that also for individual transcripts, levels of RBM7 binding correlate with sensitivity to hRRP40 and ZCCHC8. Individual examples of PROMPTs (Figure S2A) and eRNAs (Figure S2B) were inspected and confirmed this conclusion.

To attract the exosome to its substrates, NEXT needs first to bind RNA. Our discovery of a CBCN interaction led to the hypothesis that the CBC may constitute one mechanism to recruit NEXT to capped transcripts. To test this idea, we performed RBM7-RIP of four known exosome substrates (three PROMPTs and 3′-extended U1 snRNA) as well as two transcripts not known to be exosome targets (GAPDH mRNA and pre-mRNA) in control cells or in cells depleted of both CBC subunits (CBP80 and CBP20) (Figure S2C). While comparable amounts of RNA were purified with RBM7 from untreated and CBC-depleted cells (data not shown) and while the largely cytoplasmic GAPDH mRNA was unchanged between the two conditions, depletion of the CBC resulted in an ~50% decrease in RBM7-associated GAPDH pre-mRNA and 3′-extended U1 snRNA (Figure 2C). Strikingly, for the tested PROMPTs, CBC depletion resulted in a pronounced decrease in RBM7-bound material (Figure 2C), which was especially prominent taking into account that these transcripts increase in abundance upon CBC depletion (Figure S2D). Taken together, these data suggest that the CBC aids in recruiting the NEXT complex newly synthesized RNAPII transcripts.

RBM7 recruitment to RNA could also be favored by transcript sequence. To analyze this possibility, we determined enriched pentamer sequences of all RBM7 iCLIP crosslinking sites as determined by iCLIP (Figure S3B, left). Moreover, employing published 3′-end sequencing data (3′ TAG-seq) (Ntini et al., 2013), we observed a correlation between the frequency of early 3′ ends and an increased density of RBM7 iCLIP tags in these regions (Figure S3B, right). Thus, despite early 3′-end formation often being suppressed in the sense direction of transcription (Ntini et al., 2013), we suggest that the NEXT complex also here surveys early transcription termination events to target abortive transcripts for exosomal degradation.

We next analyzed RBM7 association with eRNAs, which are most often bidirectionally transcribed from common transcription initiation regions and sensitive to exosome-activity in both directions (Andersson et al., 2014b). Consistent with a role of the NEXT complex in the degradation of eRNAs, ZCCHC8 depletion increased their levels (Figure 3C, top image), and RBM7 iCLIP signal in these regions was prominent in control cells and increased in the absence of hRRP40 (Figure 3C, bottom image; compare with Figure S3C). As in the case of PROMPTs, RBM7 iCLIP tags displayed a narrow distribution around eRNA TSSs consistent with the general similarity between the biogenesis and turnover of these IncRNAs (Andersson et al., 2014b; Ntini et al., 2013).

PROMPTs and eRNAs are mostly intron-less. To expand our analysis to intron-containing IncRNAs, we investigated RBM7 occupancy over a set of annotated IncRNAs (GENCODE v18) that are expressed in HeLa cells (Andersson et al., 2014a) and divided these into monoexonic and polyexonic transcripts. For both transcript classes, RBM7 signal in PROMPT regions increased upon hRRP40 depletion (Figure 3D). Interestingly, only intron-less IncRNAs displayed a similarly strong hRRP40-sensitivity (Figure 3D, (+) directions of top and bottom images), whereas intron-containing IncRNAs showed a RBM7 iCLIP profile like for all genic transcripts (Figure 3A). Consistent with previous observations (Ntini et al., 2013), this suggests that the presence of introns generally renders RNAs less exosome sensitive.

In summary, RBM7 avidly targets RNAPII-derived transcripts of even short length with enrichment toward RNA 5′ ends, supporting the notion that CBC-aided NEXT recruitment may be universal. While recruitment of the NEXT complex to the investigated IncRNAs most often will lead to exosome-mediated decay, we stress that RBM7 binding per se does not trigger...
Figure 2. Factors Determining RBM7 Recruitment to RNA

(A) Box plots showing RBM7 iCLIP enrichment (log2(RBM7[iCLIP]/RPM/control(RNA-seq) RPM)) (left), hRRP40 (middle), and ZCCHC8 (right) sensitivity (log2(siRNA knockdown(RNA-seq) RPM/control(RNA-seq) RPM)) of the first 1 kb of 12,153 genic RNAs, their associated PROMPTs, and 10,014 eRNA regions as indicated. RPM, reads per million; KD, knockdown. Note a slight overall decrease in genic transcripts upon ZCCHC8 depletion, the nature of which is unknown.

(B) Heat graphs showing RBM7 iCLIP enrichment versus hRRP40 (left) or ZCCHC8 (right) sensitivities calculated as in (A) and for individual transcripts contained in the (A) box plots.

(C) Quantitative RT-PCR analysis of RBM7-RIP'ed RNA purified from CBP80/CBP20-depleted or untreated HeLa cells. PROMPTs (proDNAJB4, proRBM39, and proEXT1), the unstable 3'-extended U1snRNA (U1-ext), GAPDH pre-mRNA, and the mainly cytoplasmic GAPDH mRNA were analyzed as indicated. The graph

(legend continued on next page)
transcript turnover, as exemplified by the robust binding of RBM7 to intron-containing RNA.

**RBM7 Targets 3’-Extended Products from Intron-less RNAPII Genes**

Having observed RBM7 binding to lncRNAs, we asked whether other short RNAPII-derived RNAs would be equally targeted. For this purpose, UsnRNA- and RDH genes were chosen as these produce intron-less RNAPII transcripts of <1,000 nt. Stable 3’-end formation of these transcripts proceeds through cleavage by the endonucleases CPSF73 (RDHs) and CPSF73L (most U snRNAs), residing in processing complexes, which displays the fraction of RBM7-bound RNA following depletion of CBP80/CBP20, normalized to input and RNA levels in control. Mean values were quantified from triplicate RT-PCR experiments of two biological replicates with bars representing SEM.

**Figure 3. RBM7 Binds TSS-Proximal Transcripts**

(A) Density profiles of RBM7 iCLIP reads over regions of 3 kb upstream and downstream of TSSs (left) and EAGs (right) from 12,153 genes. Transcription directions are indicated by arrows. Different samples and their (+) versus (−) strand orientations are indicated.

(B) Density profiles of RNA-seq reads from genomic regions used in (A). To reduce signal from the abundant pool of cytoplasmic RNA, reads mapping to exon annotations from RefSeq and ENSEMBL were filtered before density calculation. Left and right images show (−) and (+) strand reads, respectively.

(C) Density profiles of RNA-seq (top) and RBM7 iCLIP (bottom) reads in a region of 3 kb upstream and downstream of initiation regions of 10,014 eRNAs expressed in the displayed datasets. Data are plotted as in (A).

(D) Density profiles of RBM7 iCLIP reads over regions of 3 kb upstream and downstream of TSSs of monoexonic (left) or polyexonic (right) lncRNAs plotted as in (A).

deal with short transcription units. This is a noteworthy distinction to PROMPTs and eRNAs that appear to utilize at least parts of the mRNA 3’-end processing apparatus placed out of context, a feature that has been suggested to contribute to the unstable nature of these RNAs (Andersen et al., 2013; Andersson et al., 2014b; Hallais et al., 2013; Ntini et al., 2013).

We used RBM7-RIP to assess RBM7-binding to U5 snRNA, which among other snRNAs was enriched in the RBM7 iCLIP data upon exosome inhibition (Figure 4A). The RBM7-RIP assay reliably detected RBM7 targets (Figure 2C), which could also be revealed by northern blotting analysis of the NEXT/exosome-sensitive PROMPT proDNAJB4 (Figure S4A, left), the heterogeneous species of which were precipitated (Figure S4A, right). Such RBM7 binding and NEXT/exosome sensitivity was confirmed by the respective sequence data sets (Figure S4B). Changing to a northern probe targeting U5 snRNA, some specific RBM7 binding to mature U5 could be seen, which was not markedly affected by exosome depletion (Figure 4B, left, lanes 3–6). However, exosome-sensitive RBM7 interaction with longer U5 species was clearly visible (Figure 4B, left, compare lanes 5 and 6) and even more apparent when employing a probe detecting hRRP40-sensitive 3’-extended U5 snRNAs (Figure 4B, right).
RBM7 also interacted with a species of shorter length than mature U5 snRNA of unclear nature (Figure 4B, left, lanes 3–6). We conclude that RBM7 binds to mature U5 snRNA and to its 30-extended products that are likely unprocessed precursors of U5 or transcriptional read-through products (Cuello et al., 1999). As for PROMPTs and eRNAs, RBM7 binding to the 30-extended species suggested its participation in eliciting their efficient trimming/removal by the exosome. This observation was corroborated by similar interactions of RBM7 with exosome-sensitive 30-extended species of U1 snRNA from the RNU1-59P gene (Figures S4C and S4D) and of U4ATAC snRNA (Figures S4E and S4F). Moreover, it was supported at a global level by mapping of RBM7 iCLIP reads to all U1 and U5 snRNA genes (Figure 4C, top and middle), as well as to all snRNA genes (Figure 4C, bottom).

We then turned to RDH genes. As for tested U snRNA genes, RNAPII transcription at these loci can proceed up to 1 kb past their EAGs (Anamika et al., 2012). Furthermore, the absence of
an intact CBCA complex increases the abundance of RDH mRNAs extending beyond their EAGs, which leads to their efficient targeting by the RNA exosome (Andersen et al., 2013; Gruber et al., 2009). Indeed, RBM7 was amply distributed along the 3′ extensions of three selected RDH transcripts (Figures 4D, S4G, and S4H), an observation that was generalized by RDH metagene analysis (Figure 4E). The RBM7-bound and 3′-extended RDH mRNAs were not evidently targeted by the RNA exosome (Figures 4D and 4E). A similar conclusion could be drawn from the RNA-seq data, which otherwise confirmed a previous notion of exosome-dependent turnover of mature RDH mRNA (Figures 4D, S4G, and S4H; Reis and Campbell, 2007). We suggest that loading of RBM7/NEXT onto RDH transcripts enables rapid targeting by the nuclear exosome, which only occurs upon lowered 3′-end processing efficiency. Separate analysis of mature and 3′-extended RDH RNA levels confirmed that decay of mature histone RNA is largely NEXT-independent (Figure S4I) and likely relies on the cytoplasmic RNA exosome (Reis and Campbell, 2007). A minor increase of, presumably nuclear, 3′-extended histone RNA could be detected upon ZCCHC8 depletion. We suggest this pool arises due to unsuccessful 3′-end processing of these RNAs.

Together, the data add 3′-extended Usn- and RDH-mRNAs to the range of targets bound by RBM7 early in their life cycle.

**RBM7 Targets Pre-mRNA and Accumulates at Intron 3′ Ends**

So far we reported RBM7 interactions with RNAs for which most are ultimately targeted by the RNA exosome. As shown in Figures 2A, 2B, and 3A, this is not always the case, and we therefore analyzed RBM7 distribution on protein coding transcripts in more detail. Given that RBM7 iCLIP tags mapped to both introns and exons (Figures 1B and 1C), we first plotted the data against 10,811 pre-mRNAs sorted by increasing first exon length. While this analysis confirmed the previously noticed bias of RBM7 binding toward RNA 5′ ends, it did not reveal a noticeable distinction between targeting of first exons and subsequent introns (Figure 5A), suggesting that RBM7 establishes RNA contacts prior to splicing. To confirm this possibility, we counted iCLIP tags overlapping unspliced (intron-exon [IE] and exon-intron [EI] borders) and spliced (exon-exon [EE] borders) RNA and also included the RNA-seq data for comparison. The calculated high ratios of IE and EI to EE spanning reads in the iCLIP libraries as compared with the RNA-seq libraries strongly indicate that RBM7-bound RNA is generally unspliced (Figure 5B, left). This was supported by a low fraction of EE spanning iCLIP reads, demonstrating that RBM7 is generally depleted from spliced mRNA (Figure 5B, right). Similarly high unspliced/spliced RNA ratios were reported for the yeast RNA surveillance factors Mtr4p and Trf4p as well as for the CBC (Schneider et al., 2012; Tuck and Tollervey, 2013), which associates early with pre-mRNA. Bearing the avid interaction of RBM7 with PROMPTs and eRNAs in mind, these data considerably strengthen the hypothesis that RBM7 targets newly synthesized RNA.

To gain insight into which tasks RBM7 may exercise when bound to pre-mRNA, we investigated where in intronic regions RBM7 iCLIP tags are positioned. The distribution of iCLIP read densities was plotted in regions of 500 bp upstream and downstream of exon-intron borders, removing first and last introns to prevent possible biases due to RBM7 interaction with RNA termini (Figure 3A). Interestingly, RBM7 binding was generally enriched at the 3′ ends of introns, independent of exon and/or intron length (Figure 5C, right). Accumulation of iCLIP tags in the 3′ ends of upstream exons was also observed, but only for short exons (Figure 5B, left), suggesting that this signal originates from RBM7 accumulation at 3′ ends of preceding introns. Notably, most RBM7 iCLIP reads mapping around 3′ splice sites (SSs) traversed IE borders (data not shown), meaning that they primarily derive from pre-mRNAs as opposed to excised introns. We speculate that this intrinsic 3′-end enrichment of RBM7 may, at least partly, be explained by a RBM7 preference for uridines (Figure 2D).

To investigate what role RBM7 may be playing at introns, we first focused on introns from which miRNAs are cotranscriptionally processed (Morlando et al., 2008). Interrogating RBM7 binding to intronic pre-miRNAs, we noted an enrichment of RBM7 iCLIP tags at pri-miRNA 5′ flanks (Figure 5D). Analysis of a curated pre-miRNA dataset (miRBase) confirmed this finding (Figure 5E). We therefore suggest that the NEXT complex partakes in 3′-5′ exosomal decay of the intronic pre-miRNA 5′ flanks. Thus, although using the NEXT/exosome axis for some purpose of RNA decay, miRNA-containing introns are not likely drivers of the enriched RBM7 signal at intrinsic 3′ ends.

**NEXT Is Targeted to snoRNA-Encoding Introns**

Other prominent human intron-contained ncRNAs are the snoRNAs, the vast majority of which are encoded within introns. Previous efforts revealed that splicing is required for human snoRNA biogenesis (Hirose et al., 2003) and suggested a role for 5′-3′ and 3′-5′ exonucleolysis in their processing (Berndt et al., 2012; Kiss and Filipowicz, 1995). Therefore, we plotted the distribution of RBM7 iCLIP reads over intronic regions flanking snoRNAs and found an increased signal at these 3′ ends, which was enriched upon hRRP40 depletion (Figure 6A). In support of a role of the NEXT/exosome complexes on these introns subsequent to pre-mRNA splicing, hRRP40 depletion led to an increased number of genuine intronic reads as compared with those spanning IE borders (Figure S5A). Inspection of individual snoRNA host genes supported this observation (Figures 6B and S5B–S5D) and further revealed that RBM7 binding, in the absence of the exosome, was enriched toward the very 3′ ends of introns likely representing the 3′ ends of snoRNA precursors.

RNA-seq data provided further support that the observed RBM7 binding can elicit exosomal decay/processings, in that accumulation of reads spanning the 3′-extended snoRNAs increased upon both ZCCHC8 and hRRP40 depletion (Figure 6B). This implies that NEXT/exosome controls the metabolism of intronic snoRNA through binding of RBM7 to emerging 3′ ends. Given the otherwise uniform distribution of RBM7 along pre-mRNA, this suggests a delay between early binding of RBM7 and 3′-5′ intron decay by the RNA exosome. We do not observe a general increase of intrinsic RBM7 signal or intron RNA accumulation upon hRRP40 depletion, which is most likely due to rapid 5′-3′ decay/processings of introns by the exoribonuclease XRN2 (Valen et al., 2011).
To examine the role of NEXT/exosome complexes in snoRNA metabolism, we depleted cells of either the core exosome subunit hRRP40 or one of its two catalytic subunits hRRP6 and hDIS3. Northern probes targeting the 3'-extended regions of SNORD83A, SNORD43, SNORA6, and SNORA62 revealed an accumulation of 3'-lengthened transcripts in all knockdown experiments (Figure 6C, top). hRRP40 depletion yielded the strongest phenotype, in line with the reported functional redundancy of hDIS3 and hRRP6 (Preker et al., 2008). Despite effective protein depletions (Figure S5E), we detected no major differences in levels of mature snoRNAs (Figure 6C, bottom). The interaction of RBM7 with 3'-extended snoRNAs was verified by employing the RBM7-RIP approach, combined with depletion of hRRP40. For all four tested RNAs, SNORD43, SNORD83A, SNORA73A, and SNORA73B, species corresponding to snoRNAs matured at their 5' ends and extended at their 3' ends, down to the host intronic 3'ss, were detected in a hRRP40 depletion-dependent manner (Figure 6D). IP of the

Figure 5. RBM7 Binds Pre-mRNA and Accumulates at Intronic 3’ Ends
(A) RBM7 iCLIP reads mapped along pre-mRNAs using their annotated 5’ ends as anchoring sites (vertical purple line). Displayed transcripts were ordered by ascending first exon length with the EI boundary colored purple. Read density over all analyzed RNAs is shown at the bottom. (B) (Left) Bar plots showing the fraction of mRNA/pre-mRNA mappers of iCLIP and RNA-seq reads, from the indicated libraries, spanning IE or EI relative to EE junctions. (Right) Same as left, but displaying reads spanning EE relative to the sum of IE, EI, and EE junctions (average of the two iCLIP replicates). (C) Density profiles of RBM7 iCLIP reads over a region spanning 0.5 kb upstream and downstream of EI (left) or IE (right) boundaries. Reads were assembled from nonterminal introns flanked by exons of different lengths as indicated by color coding. (D) Genome browser screenshot of RBM7 iCLIP and RNA-seq reads mapped to representative introns encoding the indicated pre-miRNAs (+) strand only. The nomenclature is as in Figure 1C. (E) Density plots of iCLIP read clusters over regions spanning 0.5 kb upstream and downstream of pre-mRNA borders. The nomenclature is as in Figure 4E.

Figure 6. RBM7 Targets the RNA Exosome to SnoRNA-Encoding Introns

(A) Density profiles of RBM7 iCLIP reads from control and hRRP40-depleted libraries across snoRNA-encoding introns, plotted as in Figure 4C.

(B) Genome browser screenshot of RBM7 iCLIP and RNA-seq reads mapped to the RPL3 locus, hosting intronic snoRNAs ([−] strand only). Red arrows and bars indicate 3' SSs and snoRNA 3' extensions, respectively. Green arrow indicates the mature snoRNA 3' end.

(C) Northern blotting analysis of total RNA from siRNA-depleted samples, as indicated. The top and bottom images show membranes probed against the 3'-extended and mature species of the noted snoRNAs, respectively. 3' SS denotes the position of the 3' SS of the snoRNA host intron.

(D) Northern blotting analysis (left) of RBM7-RIP'ed RNAs from control- or hRRP40-depleted samples. Both total (noTag) and RBM7-bound (RBM7-LAP) samples were probed against SNORD43, SNORD83A, and SNORA73A/B, as indicated. M, molecular weight markers. Western blotting analysis (right) of protein extracts from cells treated with siRNA against hRRP40 or luciferase (control). β-actin served as a loading control. Northern blots showing mature snoRNA levels of input samples are in Figure S5F.

(E) Global effects of exosome and NEXT complex component depletion on mature snoRNA/snRNA and their flanking regions. Standard RNA-seq (RNA-seq, top) and RNA-seq libraries targeting mature snoRNA and snRNA (short RNA-seq, bottom) were employed. Y axes show log2-scaled mean reads per 10 million mapped reads (RP10M) fold changes in the indicated samples relative to their GFP siRNA control. C/D- and H/ACA box intronic snoRNAs as well as all snRNAs are plotted with separate mean RP10M ratios for the mature ncRNA (blue), the upstream (orange), and downstream (green) flanking regions. 5' and 3' flanks of intronic snoRNA were taken as the respective intronic regions of the host intron. 5' and 3' flanks of snRNAs were the immediate upstream and downstream 500 bp of each gene.
mature snoRNA counterparts was subjected to some background and did not yield conclusive results (Figures 6D and S6F). Still, we suggest that RBM7 recruits the exosome to 3′-extended snoRNA transcripts and that exosome activity ensues subsequent to snoRNA 5′ processing.

Finally, we performed an RNA-seq-based quantification of human snoRNAs and snRNAs, including their flanking regions. As our original datasets were biased toward longer transcripts (>150 nt) and therefore did not detect most mature snoRNAs, we prepared libraries enriched for shorter species (<200 nt, called “short RNA-seq”) from cells depleted of hRRP40 or all three catalytic subunits of the RNA exosome (hDIS3/hRRP6/hDIS3L). In the original RNA-seq libraries a pronounced increase in tags mapping to a region covering the gene body as well as the 3′ flanks of both the H/ACA and C/D box snoRNAs was evident upon single ZCCHC8 and hRRP40 depletions as well as in their combined depletion (Figure 6E, top). However, while fragments upon single ZCCHC8 and hRRP40 depletions as well as in their codepletion of hDIS3, hDIS3L, and hRRP6 did not alter levels of mature snoRNAs and snRNAs (Figure 6E, bottom). Although this could indicate that the RNA exosome is not involved in intronic snoRNA biogenesis, we note that residual exosome complexes, long half-lives of mature snoRNA, and possible compensatory mechanisms that ensure stable snoRNA levels may mask such involvement (see below and Discussion). In any event, we conclude that targeting of NEXT/exosome to 3′-extended snoRNAs/snRNAs is general.

**RBM7 Targets Newly Synthesized RNA**

All analyses so far are compatible with early binding of RBM7 during RNA production. To directly assess whether NEXT acts at the level of newly synthesized RNA, we utilized Northern blotting analysis to monitor 3′-extended snoRNA transcript purified by 5-bromouridine (BrU)-IP. Following the efficient depletion of hRRP40 and ZCCHC8 (Figures S6A and S6B) and a 1-hour BrU labeling pulse, a strong accumulation of BrU-IP′ed 3′-extended SNORA73A and SNORA73B was apparent (Figure 7A, compare lanes 3 and 4 with lane 2). Depletions of hDIS3 and hRRP6 showed less accumulation probably because their codepletion was not highly efficient (Figures 7B and S6B). Interestingly, depletion of ZCCHC8 resulted in RNA levels comparable (3′-extended SNORA73A) or even higher (3′-extended SNORA73B) than after hRRP40 depletion. This is counter to the effects of these knockdowns on steady-state RNA levels (Andersen et al., 2013; Lubas et al., 2011) and may reflect a dedicated utility of NEXT at the level of newly synthesized RNA, whereas the exosome also employs a cofactor(s) that acts posttranscriptionally (see Discussion). An impact of RBM7 on newly synthesized RNA was directly revealed by extending the BrU-IP samples to include knockdowns of RBM7 and its combination with ZCCHC8. Assaying intronic snoRNAs and U2 snRNA showed that depletion of RBM7 led to the accumulation of BrU-labeled 3′-extended transcripts at an extent similar to depletion of ZCCHC8 and hRRP40 (Figures 7B and S6C–S6I).

We also probed for U6 snRNA and observed faster migrating species upon depletion of hDIS3+hRRP6 or hRRP40 (Figure 7C, top, lanes 6 and 7). These exosome substrates likely represent...
U6 decay products that are present in the BrU-IPs due to TUTase-mediated incorporation of BrU at their 3’ ends (Trippé et al., 2006). Steady-state RNA analyses demonstrated that these intermediates represent a minor portion of the total U6 snRNA pool (Figure 7C, bottom). Importantly, however, neither individual depletion of ZCCHC8 or RBM7 nor their combined depletion allowed the detection of U6 decay intermediates. This suggests that NEXT is not involved in this, likely posttranscriptional, exosome targeting to U6 snRNA. In summary, the similar phenotypes of NEXT complex and RNA exosome inhibition at the level of BrU-labeled RNA further support the idea that the NEXT complex functions in specifying exosome substrates among newly synthesized transcripts.

**DISCUSSION**

The variety of substrates of the human nuclear exosome begs the question of how these different RNAs are targeted. As human Nrd1p/Nab3p homologs have not been reported, the exosome-targeting mechanism remains to be elucidated. For now we suggest that these human RNAs are targeted by the exosome via their deposited RBM7 protein.

Given that splicing is often cotranscriptional (Brugiolo et al., 2013), the proposed targeting of newly synthesized RNA by RBM7 is further revealed by its preferential association with pre-mRNA relative to mRNA (Figures 5A and 5B). Furthermore, the general binding of RBM7 along pre-mRNAs with a clear enrichment for the cap-proximal ~1,000 nts (Figure 5A) highlights a couple of interesting features of RBM7 biology. First, the presence of RBM7 on RNA does not inevitably result in exosome degradation. Instead, it appears that RBM7 associates rather generally with RNA and only if a 3’ end becomes available in the vicinity will 3’-5’ exonucleolysis by the exosome follow. PROMPTs, eRNAs, and 3’-extended RNAs represent nuclear exosome targets where RBM7 association, NEXT assembly, and the occurrence of a naked 3’ end is happening within a short time frame. Second, the enrichment of RBM7 at 5’ ends of pre-mRNA and our RBM7 RIP experiment combined with CBC depletion (Figure 2C) suggests that one contribution to recruiting the protein to early RNAII-transcripts is via its interaction with the CBC (Andersen et al., 2013; Lubas et al., 2011). An early recruitment of RBM7 is further supported by its recently reported localization in both chromatin and nucleoplasmic cell fractions (Blasius et al., 2014). In other cases, as exemplified by intron binding, RBM7/NEXT could also intersect with other RNA processing events, such as splicing.

**RBM7 Marks Newly Synthesized RNA for Exosome Degradation**

Our data illustrate several examples of physical interactions between RBM7 and known nuclear substrates of the RNA exosome. The enrichment of RBM7 is pronounced in proximity to cellular TSSs that direct the synthesis of PROMPTs and eRNAs, transcript species that also accumulate upon depletion of the NEXT components ZCCHC8 and hMTR4 as well as exosome core components (Figures 2 and 3). The binding of RBM7 to intron-less IncRNA and the increase in RBM7 iCLIP signal upon hRRP40 depletion (Figure 3D) further add to the suggestion that NEXT recruits the exosome to degrade RNA. RBM7 is also found associated with relatively short and exosome-sensitive 3’-extended RNAs arising from snRNA and RDH genes (Andersen et al., 2013). RBM7 binding is detected in the mature snRNAs, but considering their much higher abundance, this interaction is likely taking place very transiently or to a minor pool of the total mature snRNA. Altogether, the data clearly confirm that the investigated prominent classes of exosome substrates in human nuclei all associate with RBM7.

In addition to its binding to PROMPTs, RBM7 also associates with RNAs transcribed in the sense direction of genic TSSs. A part of the TSS-proximal signal presumably originates from RBM7-association with short ~100–1,000 nt RNAs derived from promoter-proximal transcription termination events (Ntini et al., 2013) and constituting substrates of the exosome. In S. cerevisiae, Mtr4p was reported to bind to similar promoter-proximal sense transcripts (Tuck and Tollervey, 2013). Whether this reflects a conserved difficulty for RNAII to achieve sufficient elongation capacity during the first hundred base pairs of transcription from the TSS or whether the production of such promoter-proximal RNAs is part of a conserved regulatory mechanism remains to be elucidated. For now we suggest that these human RNAs are targeted by the exosome via their deposited RBM7 protein.

The NEXT Complex Targets 3’-Extended snoRNAs

RBM7 occupies introns, particularly at their 3’ ends. Most prominently, upon exosome/NEXT depletion, snoRNA-hosting introns display a marked stabilization of, and an increased RBM7 binding to, the 3’-intronic region flanking the snoRNA (Figure 6). In these cases, full intron decay by XRN2 is likely blocked by structures of early-formed pre-snoRNP (Richard et al., 2006). This suggests that the 5’ and 3’ ends of snoRNA processing/decay operate independently. It also underscores a role of NEXT in targeting the exosome to these 3’-extended substrates. Given that mature snoRNA levels are orders of magnitudes higher than their 3’-extended counterparts, RBM7 must bind the latter with high affinity, directly involving NEXT complex function in the metabolism of these transcripts.

We find that depletion of exosome/NEXT components does not markedly affect steady state levels of mature snoRNA (Figures 6C and 6E) and only to some extent in BrU-IP’ed samples (Figure 7C). This could suggest that the RNA exosome does not play a major role in the biogenesis of human intronic snoRNA. However, because of the long half-lives of mature snoRNA and
possible redundant processing pathways, further investigations are needed. While the observed NEXT-mediated exosome targeting of 3′-extended snoRNAs may result in the production of mature species, it is also possible that NEXT facilitates their complete degradation. A conceptually similar observation was reported for lncRNA precursors in *S. cerevisiae* (Gudipati et al., 2012). In this way, the complete turnover of apparent precursor molecules may either represent a mechanism of snoRNP quality control, in which improperly assembled snoRNP particles are degraded, or the simple removal of “surplus” ncRNA material.

**Temporal Separation of Human Nuclear RNA Decay Pathways**

Strikingly, our metabolic labeling experiments revealed an accumulation of 3′-extended newly synthesized snoRNA species in cells depleted of ZCCHC8, RBM7, and hRRP40 (Figures 7A and 7B). In contrast, steady-state analysis of snoRNAs (Figure 6E) and other exosome substrates (Andersen et al., 2013; Lubas et al., 2011) results in less pronounced phenotypes upon NEXT depletion compared with RRP40-depletion. This supports our model that NEXT specifies early exosome substrates like PROMPTs, eRNAs, and 3′-extended snoRNAs. It also argues that other exosome-dependent decay pathway(s) may take over in the absence of NEXT or when NEXT is saturated. One such pathway is specified by the nuclear poly(A)-binding protein, PABPN1, which functionally connects the RNA exosome to substrates that are already spliced and polyadenylated (Beaulieu et al., 2012). This, and other possible nuclear turnover pathways, likely constitute separate, but complementary check points, enabling continued RNA quality control en route from RNA transcription to nuclear export. A major future challenge will be to decompose specific transcript half-lives into their different kinetic components and turnover mechanisms.

**EXPERIMENTAL PROCEDURES**

A more detailed description of the Materials and Methods is provided in the Supplemental Experimental Procedures.

**Cell Culture and RNAi**

HeLa cells were seeded in Dulbecco’s modified Eagle’s medium and transfected using Lipofectamine 2000 (Invitrogen) and siRNA in a final concentration of 15 nM. Knockdown efficiency was validated by western blotting and RNA isolated for analyses using TRIzol (Life Technologies).

**RNA Abundance Analyses**

The abundance of individual transcripts was measured in RNA samples either by Northern blotting analysis, using standard protocol or globally by RNA-seq. RNA libraries were prepared using rRNA-depleted total RNA (Ribozero, Epicenter) and the NEBNext Small RNA Library Kit (NEB). For detection of newly synthesized RNA, metabolic labeling and purification of newly synthesized transcripts were done as described in (Paulsen et al., 2013) with modifications described in the Supplemental Experimental Procedures.

**iCLIP and RBM7-RIP**

iCLIP and RBM7-RIP experiments were performed using HeLa cells stably expressing a C-terminally LAP-tagged RBM7 variant, including a GFP moiety. Cells were irradiated with UV light at 254 nm, and purification of RBM7-bound RNA was done as described in the original iCLIP protocol (König et al., 2010) with modifications noted in the Supplemental Experimental Procedures.

**ACCESSION NUMBERS**

The data associated with the manuscript have been deposited to the NCBI Gene Expression Omnibus and are available under accession numbers GSE63791 (iCLIP) and GSE63496 (short RNA-seq). The previously published RNA-seq, CAGE, and 3′ TAG datasets used in this study were downloaded from accession numbers GSE52132 and GSE48286.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.12.026.

**AUTHOR CONTRIBUTIONS**


**ACKNOWLEDGMENTS**

We thank Joanna Kufel, Edouard Bertrand, Sören Lykke-Andersen, and Christian K. Damgaard for critical comments on the manuscript; Dorthe C. Riishøj and Claudia Scheffler for technical assistance; and Robin Andersson and Thomas Hansen for help with RNA annotations. This work was supported by the Danish National Research Foundation (grant DNRF58), the Danish Cancer Society and the Lundbeck and Novo Nordisk Foundations (to T.H.J.), the Polish National Science Center (2011/02/A/00001 Maestor grant to A.D.) and the Welcome Trust (grant 097383) and the Medical Research Council (to G.K.). M.L. was supported by a Boehringer-Ingelheim PhD fellowship and a START fellowship from the Foundation for Polish Science.

Received: June 24, 2014
Revised: October 29, 2014
Accepted: December 11, 2014
Published: January 8, 2015

**REFERENCES**


