JCI The Journal of Clinical Investigation

Killing the messenger: new insights into nonsense-mediated mRNA decay

Peter H. Byers

J Clin Invest. 2002;109(1):3-6. https://doi.org/10.1172/JCI14841.

Spotlight

Nonsense-mediated mRNA decay (NMD) is one of many quality control mechanisms developed by cells to maintain the metabolic status quo. The cell's objective in this case is to destroy mRNA species that contain premature termination codons (PTCs) so that only full-length proteins are produced. Targeted destruction of proteins that misfold as a result of missense mutations (those that result in substitution for amino acids in the protein) and NMD are ancient and evolutionarily conserved strategies to protect the cell from mutations (or errors in transcription) that could yield truncated, potentially hazardous proteins. Eukaryotes as diverse as yeast, Caenorhabditis elegans, and humans employ a limited and overlapping array of proteins that cooperate to destroy mRNA species harboring PTCs. The site of mRNA destruction and the mechanisms by which the cell recognizes premature, as opposed to the constitutive, termination codons have been the objects of intense scrutiny and continuing debate over the last decade and a half. In that time, seven C. elegans genes have been identified that are essential for NMD (named smg-1 through smg-7). Three of these have yeast homologs and the same three genes have human homologs, although the human repertoire is larger than that in yeast (Table 1) as a result of gene duplication. Some additional genes have been identified in yeast as involved in the NMD pathway, […]

Killing the messenger: new insights into nonsense-mediated mRNA decay

Peter H. Byers

Departments of Pathology and Medicine, University of Washington, Seattle, Washington, USA

Address correspondence to: Peter H. Byers, Department of Pathology, Box 357470, University of Washington, Seattle, Washington 98195-7470, USA.
Phone: (206) 543-4206; Fax: (206) 616-1899;
E-mail: pbyers@u.washington.edu.

J. Clin. Invest. **109**:3–6 (2002). DOI:10.1172/JCI200214841.

Nonsense-mediated mRNA decay (NMD) is one of many quality control mechanisms developed by cells to maintain the metabolic status quo. The cell's objective in this case is to destroy mRNA species that contain premature termination codons (PTCs) so that only full-length proteins are produced. Targeted destruction of proteins that misfold as a result of missense mutations (those that result in substitution for amino acids in the protein) and NMD are ancient and evolutionarily conserved strategies to protect the cell from mutations (or errors in transcription) that could yield truncated, potentially hazardous proteins. Eukaryotes as diverse as yeast, *Caenorhabditis elegans*, and humans employ a limited and overlapping array of proteins that cooperate to destroy mRNA species harboring PTCs.

The site of mRNA destruction and the mechanisms by which the cell recognizes premature, as opposed to the constitutive, termination codons have been the objects of intense scrutiny and continuing debate over the last decade and a half. In that time, seven *C. elegans* genes have been identified that are essential for NMD (named *smg-1* through *smg-7*). Three of these have yeast homologs and the same three genes have human homologs, although the human repertoire is larger than that in yeast (Table 1) as a result of gene duplication. Some additional genes have been identified in yeast as involved in the NMD pathway, but their homologs in mammalian cells and *C. elegans* have not been extensively explored (see Table 1).

Table 1
Nonsense-mediated decay genes and their chromosomal location, putative function, and homologies to yeast and worm genes

Human gene GenBank accession no.[A]	Putative function	*S. cerevisiae* gene GenBank accession no.	*C. elegans* gene GenBank accession no.
SMG-1–like AB061371	Phosphatidylinositol 3-kinase–like	Not identified	*smg-1* AF149821
RENT-1, *hUPF1* XM_051416	RNA helicase, RNA-dependent ATPase, RNA-binding protein	*upf1* YSCUPF1	*smg-2* AF074017
RENT-2, *hUPF2* XM_018031	Acidic 127-kDa protein that shares homology and motifs with eIF4G and a cap-binding protein (CBP80)	*upf2*/*nmd2* SCU14974	*smg-3* AF074017
hUPF3A[B] XM_058893	Basic 45-kDa protein with nuclear localization and nuclear export sequences	*upf3* L41153	*smg-4* CAA94820
No human homolog identified	No structural elements identified in *C. elegans* sequence	No apparent yeast homolog	*smg-5* U64441
No human homolog identified		No apparent yeast homolog	*smg-6* Mapped; sequence not available in WormBase[C]
No human homolog identified		No apparent yeast homolog	*smg-7* Not identified in WormBase
Y14 (RBM8A) AF299118	mRNA export factor has an RNA-binding domain	*yralp* U72633	Not identified in WormBase
ALY/REF AF047002	RNA-binding protein	Not identified	Not identified in WormBase
Not identified in human ESTs	Has a "nudix" domain common in many proteins that interact with other proteins	*nmd1* U31377	Not identified
CGI-07 BC013317	No specific protein motifs found	*nmd3* U31376	Hypothetical protein T25G3.3
Shares homology with many human sequences	Nuclear poly-A–binding protein	*hrp1*/*nab4* U38535/U35737	No specific transcript identified in WormBase

[A]When the accession numbers are entered at the National Center for Biotechnology Information home site (http://www.ncbi.nlm.nih.gov/) they will provide direct links to nucleotide and protein sequences as well as references to the identification of these genes. [B]*UPF3B* (XM_016123) and *UPF3X* (NM_023010) appear to represent two additional copies of the *UPF3* gene. How they differ in function from the *UPF3A* gene and transcript is not understood. [C]The web address for WormBase, the database for *C. elegans* genome information, is http://www.wormbase.org/. EST, expressed sequence tag.

Cellular recognition of PTCs

Termination codons can arise by mutations within a coding sequence in several different ways. The first and simplest is the introduction of a nonsense mutation, in which a sense codon undergoes a single base pair substitution to yield a TAG, TAA, or TGA termination codon. Second, insertion or deletion of a number of nucleotides not divisible by 3 will create a frame shift in the coding sequence and, on average, yield a new stop codon within 20 downstream codons. Such frame shifts can arise by insertion or deletion during replication, often in regions of repetitive nucleotides. Alternatively, mutations that lead either to the use of cryptic splice sites or to exon-skipping will result in frame shifts in many instances. Finally, intron inclusion, depending on the size of the intron, is very likely to result in either a frame shift or the introduction of a PTC.

In *Saccharomyces cerevisiae*, NMD appears to occur exclusively in the cytoplasm. The site of NMD in mammalian cells remains an open question, with evidence that supports both a nuclear and a cytoplasmic location. Indeed, it may well be that recognition and destruction occur in both locales. But the enduring question remains: How are the interloper PTCs recognized by the cell so that the mRNA can be targeted for destruction? For a long time it has been argued that the cell really has only a single system capable of sensing in-frame stop codons — translation. So, the argument continues, recognition of the abnormal mRNA species must occur in the context of translation and, thus, ought to be localized to the cytoplasm, the universally acknowledged site of translation. As a compromise position, the nuclear pore — the site of exit of the mRNA from the nucleus — might allow recognition and destruction by anchoring ribosomes in the local region and subjecting mRNA to degradation as it leaves the nucleus. Yet, there is clear evidence that at least some mRNA species that house PTCs do not exit the nucleus (1). Two recently published papers that propose that translation of some mRNA species occurs in the nucleus (2, 3) provide one interesting resolution to this dilemma by allowing a "pioneer" round of translation (that is, scanning by a ribosome), which could mark these faulty mRNAs for destruction.

Not all mRNAs that contain PTCs are targeted for destruction. Notably, at least some intronless genes (4) and transfected cDNAs that contain PTCs generally are not subject to NMD, for reasons that are not intuitively obvious at first. Furthermore, not all mRNAs derived from genes that contain introns are unstable if they contain PTCs; they lose stability only when the PTC is located 5′ of the last intron by about 50 or more nucleotides (5, 6). There must be one or more introns downstream from the PTC (7), which must be in the correct reading frame to be active (8). Furthermore, in at least some contexts, NMD requires components of translation, as inhibitors of translation diminish or ablate NMD (9). These observations point to linkage between the splicing machinery and some aspects of the translation (or translational termination) machinery to generate a mechanism that senses reading frames as integral parts of the recognition of PTCs by the cell. The most striking difference in mechanism between yeast NMD and the process in mammalian cells is the relation to splicing. In the yeast *S. cerevisiae,* which has few intron-containing genes, the presence of a downstream element in the 3′-untranslated region is used to "mark" the correct termination codon. This led to models in which a "mark" appears within the processed spliced mRNA in mammalian cells, presumably in the aftermath of splicing.

A molecular model for cytoplasmic NMD

The emerging picture by which these PTC-containing transcripts are recognized is proving increasingly complex and interesting. Although textbooks often depict precursor mRNA molecules and mRNA molecules as lonely travelers in a complex nucleus, this image could hardly be farther from the truth. From the moment of transcriptional initiation, the mRNAs in the nucleus are in the company of numerous proteins. Soon after transcription is completed, the 5′ end of the nascent mRNA is modified ("capped"), and a protein complex associates with the newly capped end of the molecule. In the cytoplasm, translational initiation and decapping reactions compete for the site, to determine whether protein synthesis or mRNA degradation carries the day. Mature mRNA species in the nucleus are also protected from decapping, although the mechanism for this protection is likely to be distinct from that which occurs in the cytoplasm. If, as is generally assumed, translational initiation occurs solely in the cytoplasm, or even if a single round of pioneer translation occurs within this compartment, it is unlikely that initiation complexes would be sufficiently abundant to protect the 5′ end of mRNAs from degradation.

The next set of events for most transcripts involves splicing, the removal of intervening sequences (introns) from between the coding domains (exons). Intron removal is orderly but not processive; introns are removed from a given pre-mRNA species in a characteristic but not invariant order, which, in large genes, does not correspond to simple progression from 5′ to 3′ (although, in general the 5′-end introns are processed prior to those at the 3′ end). Distinct small nuclear ribonuclear proteins recognize the branch site, the 3′ end of the intron, and the 5′ end of the intron. The resulting multisubunit RNA-protein complex, the spliceosome, facilitates cleavage at the 5′ end of the intron and formation of a lariat at a branch point close to the 3′ end of the intron, after which it cleaves the 3′ end of the intron and ligates the two exon ends.

During splicing, a protein complex is deposited about 20–24 nucleotides upstream of the splice site. Proteins so far identified in this splice junction complex include Y14 (an RNA-binding protein) (10), Aly/Ref (an RNA-binding and export factor) (11), RNPS1 (an RNA-binding protein previously implicated in splicing) (12), SRm160 (a protein that associates with the splicing complex but that does not bind RNA) (13, 14), DEK (a 45-kDa phosphoprotein that binds SRm160 and is part of the spliceosome complex) (15), and magoh (which binds to Y14 and TAP, a protein involved in mRNA export to the cytoplasm) (16). To date, none of these proteins seems to be essential for NMD in yeast.

Both Y14 and Aly/Ref bind only spliced mRNA species; they do not bind to unspliced, intron-containing mRNA or to intronless mRNA species. Aly/Ref is associated initially with the spliceosome, but following intron removal, it translocates on the mRNA to the site of splicing complex formation, upstream of the former intron-exon boundary. Y14 does not appear to be part of the spliceosome and thus probably depends on Aly/Ref for positioning on the spliced mRNA. The other components of the postsplicing marker complex bind one or more of these proteins and are transported out of the nucleus with the mature mRNA.

According to recently emerging models (14, 17), the formation of this marker complex represents a key step in NMD, because it links the intranuclear process of splicing to translation, a predominantly (if not exclusively: see below) cytoplasmic event. The protein product of the *hUPF3* gene binds to Y14 protein in the nucleus on the spliced mRNA. The other two UPF gene products, hUPF2p and hUPF1p, are thought to reside at the periphery of the nucleus and in the cytoplasm, respectively (18). The hUPF2p appears to provide a bridge between hUPF3p and hUPF1p in the complex. The hUPF1p protein appears to be one of the key structural and functional elements to nonsense-containing mRNA degradation, providing links between the exon-exon boundary marks, the translation termination complex, and the mRNA cap complex (19).

In the cytoplasm, these proteins are stripped from the mRNA with the first passage of a ribosome during translation. If, however, translation terminates at a PTC, hUPF1p present at the proximal exon-exon boundary can interact with proteins in the translational termination complex and can then interact with DCP2p, thereby activating the decapping protein DCP1p. Once the cap is removed, the mRNA is rapidly degraded by the action of constitutively active intracellular 5′→3′ exonucleases. Because these complexes do not bind 3′ to the constitutive termination codon in mammalian mRNAs, those mRNA species that lack a PTC are protected from NMD.

To date, only a small number of mRNAs have been examined for binding of complexes, and it seems likely that NMD pathways for various mRNA species will prove idiosyncratic. For instance, special arrangements may be needed for mRNA quality control in the small number of genes that contain introns within their 3′ untranslated regions. Moreover, given the sheer number of splicing events that occur in the nucleus, it seems too much to ask that every exon-exon boundary be marked with a splicing complex. There is, as yet, no evidence for an upper limit to the distance between the splicing complex and the upstream PTC, so the system could probably operate efficiently as long as some exon-exon boundaries near the 3′ end of the gene receive a mark. Finally, selection for efficient NMD may have varied between different gene products, and there may be some for which a degree of variation at the C-terminal end of the protein is permissible or even advantageous. The extent to which cells tolerate the accumulation of a given PTC-bearing mRNA varies greatly and cannot be predicted in advance, a point to which I return later.

Intranuclear translation and NMD

While the above model provides a satisfying explanation for cytoplasmic decay — with some caveats, as described — it fails to explain nuclear NMD because of the absence of a candidate for frame reading in the nucleus. It is at this point that the two recent studies that suggest that nuclear frame determination may derive from early rounds of protein synthesis (2, 3) provide some relief. Iborra and colleagues (2) studied permeabilized cells and showed that nascent peptide can incorporate labeled amino acids, that this incorporation was reduced or ablated by protein synthesis inhibitors cycloheximide and puromycin but not chloramphenicol (an inhibitor of bacterial protein synthesis), and that labeled peptides can be detected in the nucleus by light and electron microscopy. Isolated nuclei appeared to have similar capacity. Although these studies used permeabilized cells and nuclei, the pore size of the nuclei seemed sufficient to exclude newly synthesized proteins from the cytoplasm. Incorporation depended on the activity of polymerase II, suggesting that transcription and translation of at least some transcripts are directly coupled.

The suggestion that pioneer translation occurs in the nucleus — at levels sufficient to account for efficient intranuclear NMD — remains highly unorthodox. Perhaps the greatest resistance to this model stems from the recognition that it would require that not only ribosomes, but also a long list of other accessory macromolecules, be available within the nucleus. These include initiation, elongation, and — especially — termination factors, none of which would be expected to be there in high abundance. The recent work of Ishigaki and colleagues (3) helps address this concern by showing that nuclear mRNA species are associated with factors normally found only in the translation complexes in the cytoplasm.

These findings are consistent with the proposal that reading frame scanning is done by ribosomes assembled in the nucleus and can initiate translation on nuclear mRNAs. To be effective, however, such an association would have to occur in conjunction with splicing, as unspliced mRNAs would normally be expected to contain PTCs. Missing in this model is the nature of communication to the decapping proteins and the 5′→3′ exonucleases of the stalled transit of the ribosome; the activity of the hUPF1p protein (which is also thought to be at least predominantly cytoplasmic) would clearly be required for nuclear NMD to proceed through the pathway described above.

The various locations in which NMD has been proposed to occur may prove difficult to distinguish on kinetic grounds. Thus, if hUPF1p or some equivalent is deposited on the marked mRNAs during transit through the nuclear membrane, rapid juxtanuclear mRNA degradation or intranuclear degradation would each prevent the target mRNAs from appearing in the cytoplasm. Interestingly, certain mRNAs that carry splice site alterations that would produce PTCs in fact accu-

mulate within the nucleus, trapped within a splicing complex. Whether these mRNAs are subject to destruction by the same NMD pathway as other, fully processed mRNAs remains to be studied (20), but it would not be surprising to learn that cells employ more than one strategy to recognize and destroy PTC-bearing mRNAs.

Implications for understanding the effects of mutations

Regardless of the site or mechanism of destruction of the nonsense codon–containing mRNAs, the existence of efficient NMD in eukaryotic cells is a matter of considerable importance in the analysis of naturally occurring mutations. The casual assumption that a PTC-containing allele identified during a mutation search encodes a truncated protein (or, in the case of a frame shift, a protein with novel properties) is not justified in general. The fate of any such gene product must be scrutinized carefully. In such cases, proof requires demonstration that the mRNA is stable, that the protein is synthesized, and that the protein disturbs function. For splice site mutations, it may well require that the true products of the mRNA be identified. Often, multiple mRNA species can result from the same mutation, some of which may be unstable, while others may yield in-frame, stable mRNAs that indeed encode potentially deleterious proteins (1). These outcomes cannot be currently predicted on the basis of sequence alone and require direct demonstration.

A vital consideration in this regard is the experimental use of expressed cDNAs derived from full-length mRNA to demonstrate that the protein synthesized runs amok in the cell. Because NMD is such a powerful surveillance strategy, honed by evolutionary forces to recognize and destroy PTC-containing mRNAs, very few of these mRNAs survive to tell their story. Thus use of cDNAs, which sidesteps these mechanisms and leaves the cell defenseless against PTCs in this nonbiological context**,** provides anything but a faithful rendition of the effects of mutations that result in PTCs. Unless the PTC-bearing mRNA can be shown to be present in the cytoplasm in amounts likely to produce enough abnormal protein to interfere with cell function, an alternative strategy to define molecular pathogenesis should be pursued. Perhaps if we can convince ourselves to abandon the term "truncated protein" in most of these situations, our perception of the effects of these mutations will be rectified.

1. Schwarze, U., Starman, B.J., and Byers, P.H. 1999. Redefinition of exon 7 in the COL1A1 gene of type I collagen by an intron 8 splice-donor-site mutation in a form of osteogenesis imperfecta: influence of intron splice order on outcome of splice-site mutation. *Am. J. Hum. Genet.* **65**:336–344.
2. Iborra, F.J., Jackson, D.A., and Cook, P.R. 2001. Coupled transcription and translation within nuclei of mammalian cells. *Science.* **293**:1139–1142.
3. Ishigaki, Y., Li, X., Serin, G., and Maquat, L.E. 2001. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell.* **106**:607–617.
4. Maquat, L.E., and Li, X. 2001. Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay. *RNA.* **7**:445–456.
5. Cheng, J., Belgrader, P., Zhou, X., and Maquat, L.E. 1994. Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. *Mol. Cell. Biol.* **14**:6317–6325.
6. Carter, M.S., Li, S., and Wilkinson, M.F. 1996. A splicing-dependent regulatory mechanism that detects translation signals. *EMBO J.* **15**:5965–5975.
7. 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.
8. Belgrader, P., Cheng, J., and Maquat, L.E. 1993. Evidence to implicate translation by ribosomes in the mechanism by which nonsense codons reduce the nuclear level of human triosephosphate isomerase mRNA. *Proc. Natl. Acad. Sci. USA.* **90**:482–486.
9. Carter, M.S., et al. 1995. A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *J. Biol. Chem.* **270**:28995–29003.
10. Kataoka, N., et al. 2000. Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol. Cell.* **6**:673–682.
11. Zhou, Z., et al. 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature.* **407**:401–405.
12. Mayeda, A., et al. 1999. Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. *EMBO J.* **18**:4560–4570.
13. Blencowe, B.J., Issner, R., Nickerson, J.A., and Sharp, P.A. 1998. A coactivator of pre-mRNA splicing. *Genes Dev.* **12**:996–1009.
14. Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**:6860–6869.
15. McGarvey, T., et al. 2000. The acute myeloid leukemia-associated protein, DEK, forms a splicing-dependent interaction with exon-product complexes. *J. Cell Biol.* **150**:309–320.
16. Kataoka, N., Diem, M.D., Kim, V.N., Yong, J., and Dreyfuss, G. 2001. Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. *EMBO J.* **20**:6424–6433.
17. Kim, V.N., Kataoka, N., and Dreyfuss, G. 2001. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. *Science.* **293**:1832–1836.
18. Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. 2000. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell.* **103**:1121–1131.
19. Bhattacharya, A., et al. 2000. Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. *RNA.* **6**:1226–1235.
20. Johnson, C., et al. 2000. Tracking COL1A1 RNA in osteogenesis imperfecta: splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. *J. Cell Biol.* **150**:417–432.