The ever-increasing complexities of the exon junction complex
Thomas Ø Tange¹, Ajit Nott² and Melissa J Moore¹,³

Over the past decade many studies have revealed a complex web of interconnections between the numerous steps required for eukaryotic gene expression. One set of interconnections link nuclear pre-mRNA splicing and the subsequent metabolism of the spliced mRNAs. It is now apparent that the means of connection is a set of proteins, collectively called the exon junction complex, which are deposited as a consequence of splicing upstream of mRNA exon–exon junctions.

Introduction
The expression of eukaryotic genes is a process with many individual steps. These include transcription and pre-mRNA processing in the nucleus, mRNA export through the nuclear pore complex, translation in the cytoplasm, and finally decay of the mRNA and protein. Although these processes were traditionally thought of and analyzed as entirely separate events, it has recently become apparent that in vivo they are all highly interconnected [1–3]. One type of interconnection is the ability of an earlier process in a pathway to influence one or more later processes. A prime example of this is pre-mRNA splicing, which can imprint mRNAs with information that affects their subsequent metabolism. Early indications of such mRNA imprinting came from studies of nonsense-mediated mRNA decay (NMD) — the selective degradation of mRNAs containing premature termination codons (PTCs). The observation that, in mammalian cells, this translation-dependent decay pathway requires the presence of a spliceable intron downstream of a PTC [4–6] suggested that the splicing process leaves behind some type of ‘mark’ capable of signaling the positions of exon–exon boundaries to the translation and decay machineries [7–9]. Another early indication that splicing can affect subsequent events in the RNA processing pathway came from experiments in Xenopus oocytes, wherein spliced mRNAs exhibited altered translational yields compared to mRNAs not generated by splicing [10,11].

The simplest way in which splicing might ‘mark’ an mRNA is by altering the complement of associated proteins, which together with the mRNA constitute the messenger ribonucleoprotein particle (mRNP). The first direct evidence for this came from gel shift experiments, where an mRNP generated by splicing in vitro exhibited reduced mobility on native gels compared with an otherwise identical mRNP not generated by splicing [12]. Subsequently, several proteins that specifically associate with spliced RNAs were identified by crosslinking and immunoprecipitation approaches [13,14]. Soon thereafter, it was shown that these splicing-dependent proteins are deposited at a specific position approximately 20–24 nts upstream of exon–exon junctions [15,16]. This set of proteins is now known as the exon junction complex (EJC).

In the four years since its discovery, the EJC has become the focus of intense research by numerous laboratories. In addition to being a key effector of NMD, the EJC also appears to play important roles in subcellular mRNA localization and translational yield. In the following sections, we will review recent progress on EJC structure and assembly, as well as its proposed functions. We also speculate on the origin and evolution of the EJC.

EJC structure, assembly and removal
The EJC is a highly dynamic structure consisting of a few core proteins plus several more peripherally associated factors; most of these peripheral proteins join the complex only transiently, either during EJC assembly or during subsequent mRNA metabolism. The best-characterized core component to date is the Y14:Magoh heterodimer. Y14 and Magoh are both found in spliceosomes that have completed the first chemical step of splicing (lariat formation) [17–19]. After the second chemical step (exon ligation), they remain associated with the mRNA, traveling with it to the cytoplasm [14,20–22]. Because Y14 contains a well-defined RNA recognition motif (RRM),
an early hypothesis was that it might constitute part of the RNA-binding platform upon which the rest of the EJC assembles. However, crystal structures of the Drosophila and human Y14:Magoh heterodimers [23**–25**] revealed an unexpected mode of interaction between Y14 and Magoh that completely masks Y14’s RRM. Thus Y14 is unlikely to bind the spliced mRNA directly, a conclusion supported by Y14’s inability to form crosslinks to highly reactive photoactivatable groups incorporated at the site of EJC deposition [18].

Recently, a better candidate for an EJC anchoring factor has emerged. This protein, eIF4AIII, is a member of the eIF4A family of RNA helicases and is a nuclear shuttling protein [26**,27**] that forms a stable complex with Y14 and Magoh when co-expressed in E. coli [28**]. Furthermore, eIF4AIII has been shown to be an integral component of EJCs assembled in vitro [26**,27**,29**]. eIF4AIII readily forms crosslinks to spliced mRNA in the area of the EJC, and anti-eIF4AIII antibodies interfere with EJC formation in vitro [26**]. Intriguingly, crystal structures of related helicases indicate that their nucleic acid binding site spans about 8–10 nts [30,31], a length remarkably similar to the size of the EJC footprint on spliced mRNA [15]. Furthermore, eIF4A-type helicases generally interact with RNA in a sequence-independent fashion [31], which correlates with the apparent sequence-independence of EJC deposition [15]. If eIF4AIII does prove to be the EJC anchor, an important question yet to be addressed is how it remains bound so stably to mRNA in contrast to related RNA helicases, which cycle through RNA-bound and RNA-free states coupled to NTP hydrolysis [31].

All of the other known EJC proteins associate more peripherally than Y14, Magoh and eIF4AIII. In the nucleus and/or nuclear extracts, complexed proteins include the splicing co-activators SRm160 and RNPS1, the alternative splicing factor Pinin, the mRNA export factors UAP56, REF/Aly and TAP/NFX1:p15 and the NMD factor Upf3b [14,15,20,21,32,33,34*,35,36]. Although DEK, a nuclear oncoprotein, was originally thought to be EJC-associated, more recent data suggest that DEK is not a genuine EJC component [18]. Immunoprecipitation experiments with the nuclear and cytoplasmic fractions of Xenopus oocytes have indicated that SRm160, RNPS1, REF/Aly and TAP/NFX1:p15 all dissociate from spliced mRNPs during or shortly after mRNA export to the cytoplasm [20–22]. Upf3b might also dissociate, but it is more likely that it is made inaccessible to antibodies by its interaction with Upf2, an NMD factor that joins the EJC in the cytoplasm [20] (see Figure 1).

Once in the cytoplasm, the majority of EJCs are removed during the first or ‘pioneering’ round of translation [37,38]. Immunoprecipitation studies of nuclear and cytoplasmic fractions from mammalian cells have indicated that this first round of translation occurs while mRNAs are still associated with the nuclear cap-binding CBP80:CBP20 heterodimer. In subsequent rounds of translation, CBP80:CBP20 is replaced by the cytoplasmic cap-binding protein, eIF4E, and EJC factors are not detectably associated with eIF4E-bound mRNAs.

**EJC and subcellular mRNA localization**

After mRNAs are generated in the nucleus, they are exported through nuclear pores to the cytoplasm, a process that in metazoans is mediated primarily by the TAP/NXF1:p15 heterodimer [36]. Early experiments in Xenopus oocytes indicated that the presence of an EJC could enhance the efficiency of mRNA export [12,20,39]. As mentioned above, three export factors, UAP56, REF/Aly and TAP/NXF1:p15, have been reported to associate with spliced mRNAs. UAP56 (also known as Sub2 in S. cerevisiae), initially identified as a splicing factor, was recently demonstrated to be essential for mRNA export and is thought to be responsible for recruiting REF/Aly to mRNAs [35,40,41]. REF/Aly can then in turn interact directly with NXF1:p15, a protein which facilitates bulk mRNA export through its interactions with components
of the nuclear pore complex [20,33,39,42,43]. Another link between the EJC and mRNA export was recently suggested to involve the novel EJC factor, Pinin. Expression of antisense Pinin RNA was shown to lead to some accumulation of bulk polyA RNA in the nucleus, indicating a possible role for Pinin in nuclear mRNA export [34].

While a model in which the EJC licenses export of spliced mRNAs is appealing, several recent results suggest that REF/Aly and the EJC are not essential for this process [36]. RNAi-mediated knockdown of several EJC proteins (including REF/Aly), individually or in combination, only partially affected bulk polyA mRNA export in Drosophila cells [44]. Furthermore, two studies reported no defect in mRNA export upon simultaneous knockdown of all three REF/Aly genes in C. elegans [45,46]. These results suggest that although EJC deposition may enhance the export of some spliced mRNAs, it is not the only means by which mRNAs can become associated with export adaptor proteins.

Regardless of its role in mRNA export from the nucleus, it is becoming increasingly clear that the EJC plays an important function in the proper localization of at least one mRNA within the cytoplasm. Although it has not yet been formally proven that splicing results in EJCs being deposited on Drosophila mRNAs, the Drosophila Y14:Magoh and eIF4AIII homologs are essential for proper localization of oskar mRNA during oogenesis [28,47,48]. Localization of oskar to the posterior pole of the developing oocyte is essential for subsequent specification of the germline [49]. A currently open question is whether this apparent role for the EJC in mediating oskar mRNA localization is an exception, or whether it reflects a more general function of the EJC. It might therefore be of interest to investigate whether the EJC has any role in mRNA localization in highly asymmetric somatic cells such as neurons, which are highly dependent on localized translation for synaptic function [50].

**EJC and translation**

As mentioned above, early experiments in Xenopus oocytes showed that splicing can influence mRNA translational yield [10,11]. More recently this has been shown to be true in mammalian cells as well. Quantitative analysis of intron effects on various steps in the process of gene expression revealed that greater amounts of protein are produced per molecule of spliced mRNA than from otherwise identical mRNA molecules not produced by splicing [51,52,53]. Expression profiles of mRNAs with first exons either too short or just long enough to accept an EJC directly implicated the EJC in this phenomenon [54,55]. Further, the EJC proteins RNP51, Y14 and Magoh could all enhance mRNA translational yield when artificially tethered to a reporter mRNA [55].

A clue to the mechanism by which splicing leads to greater protein yields was provided by cytoplasmic polysome analysis [55]. In this study, significantly greater proportions of spliced mRNAs were found associated with polysomes compared with otherwise identical mRNAs transcribed from cDNAs. Moreover, this difference in polysome association could be reproduced by tethering the same EJC proteins that enhanced translational yield above. Although the mechanism is currently unknown, having an EJC appears to enhance the uptake of spliced mRNAs into the translationally active pool. Whether this occurs via EJC-dependent recruitment of translation initiation factors or EJC-dependent targeting of mRNPs to subcellular locations highly active in translation (e.g. the cytoskeleton) is currently unknown.

Surprisingly, another reason why some spliced mRNAs yield greater levels of protein is because they lend enhanced stability to the protein product. Dihydrofolate reductase protein expressed from stably transfected mini-genes was found to have a 2.7-fold longer half-life when expressed from an intron-containing construct than from a cDNA construct [53]. The mechanism by which this occurs, however, is a complete mystery at present.

**EJC and nonsense-mediated decay**

In mammalian cells, EJC deposition is pivotal for the process of NMD. It is now well-established that EJCs provide one means of distinguishing premature stop codons, which are upstream of the last EJC, from natural stop codons, which are downstream of the last EJC [9,56,57]. According to prevailing models, the NMD factor Upf3a/b joins the EJC in the nucleus [33,58,59]. Before or immediately after mRNA export, Upf2 is recruited to the cytoplasmic EJC via interactions with Upf3a/b [20,37,58]. If during the pioneering round of translation a stop codon is encountered before removal of one or more EJCs, eRF1/eRF3-bound Upf1 bridges the terminating ribosome to Upf3 and Upf2 in the downstream EJC(s). The formation of such an activated NMD complex then leads to rapid decay of the PTC-containing mRNA [20,37,58,60–62].

Although NMD and the NMD-specific Upf proteins are conserved in all eukaryotes examined thus far [57], discrimination between PTCs and normal termination codons seems to involve different mechanisms in yeast, flies and mammals. In budding yeast, PTC recognition does not require either introns or splicing [63], and the S. cerevisiae genome contains no apparent orthologs for any of the known EJC core factors (i.e. Y14, Magoh and eIF4AIII). Instead, somewhat loosely defined DSEs (downstream sequence elements), which can destabilize an mRNA when located 3' to a termination codon, apparently serve as the EJC equivalent for NMD in budding yeast [63]. A recent surprise, however, came from RNAi knockdown experiments in flies. Although all of the
known EJC core factors are conserved in *Drosophila*, none of them appear to be essential for NMD in that organism [64**]. Rather, flies seem to be more yeast-like in their mechanism of PTC recognition.

**What is the evolutionary origin of the EJC?**

The finding that EJC core proteins are present in flies but not required for NMD provides new fodder for speculation regarding the evolutionary origin of the EJC. A model we favor is that the EJC originally arose as a means to promote selective translation of spliced mRNAs. If, in early eukaryotes, pre-mRNA splicing was a relatively inefficient process, there would have been a strong selective pressure to tag RNAs as spliced so that only these RNAs would be efficiently translated. Additionally, the EJC may have helped to retain these spliced mRNAs on a solid support (e.g. the cytoskeleton) where they could be readily shunted into the cytoskeleton-bound translation machinery. Thus, the present functions of the EJC in mRNA localization and regulated translation during *Drosophila* oogenesis may reflect these primordial roles.

In the mammalian lineage, the abundance of introns in pre-mRNAs, coupled with poor conservation of splice site consensus sequences, has led to an explosion of alternative splicing. Hand-in-hand with productive alternative splicing, however, comes nonproductive mis-splicing, defined here as alternative splicing events leading to mRNAs containing prematurely truncated open reading frames. The prevalence of mis-spliced mRNAs in the mammalian lineage may therefore have provided an impetus to also employ the EJC as a signal for premature termination codon recognition. This could have been accomplished by simply adding to one of the EJC core factors a domain that recruits the Upf proteins. In this scenario, the core EJC factors may be important for NMD in mammalian cells not because they are directly involved in NMD, but rather because they are essential for the integrity of the EJC structure which, in turn, serves as a platform for recruitment of the Upf factors.

**Conclusions**

The discovery and characterization of the EJC over the past few years have added significantly to our understanding of the mechanisms by which pre-mRNA splicing can influence subsequent steps in gene expression. It has also opened an important new window into the complexities and dynamics of mRNP formation and metabolism. However, much remains to be done. Even though numerous EJC proteins have been identified, the current list is almost certainly incomplete. Determination of the 3D structure of the EJC will undoubtedly lend insight into its means of assembly, its mechanism of RNA binding and its evolutionary origin. Further functional analysis, particularly with regard to the roles of the EJC in mRNA localization and translational efficiency, are sure not only to provide additional surprises, but also to increase our understanding of the advantages of allowing the synthetic histories of mRNAs to dictate their subsequent metabolism.

**References and recommended reading**

Papers of particular interest, published within the annual period of review, have been highlighted as:
- of special interest
- of outstanding interest

The ever-increasing complexities of the exon junction complex
Tange, Nott and Moore 283


The crystal structures of Drosophila and human Y14:Magoh heterodimers are reported in three independent studies (this study and [24**,25**]). All three structures reveal an unusual mode of interaction between Y14 and Magoh that completely buries the Y14 RRM, thus making it unavailable for binding RNA when dimerized with Magoh. Additionally, Fribourg et al. demonstrate that, in mammalian cells, tethering Magoh to a reporter mRNA downstream of a termination codon can triggers NMD and that the ability to dimerize with Y14 is crucial for tethered NMD.


26. Shibuya T, Tange TO, Sonenberg N, Moore MJ: eIF4AIII binds

In this paper it is demonstrated that eIF4AIII is a novel EJC component that interacts with both Y14:Magoh and the spliced mRNA. This, coupled with antibody inhibition data, suggests that eIF4AIII could be the anchor protein upon which the remainder of the EJC assembles. RNAi knockdown of eIF4AIII leads to an NMD defect in mammalian cells, and eIF4AIII is shown to be a nucleo-cytoplasmic shuttling protein.


This paper reports that eIF4AIII, but not eIF4AI/II, is essential for NMD in mammalian cells and that eIF4AIII is a novel EJC component. Additionally, eIF4AIII is found to be a nucleo-cytoplasmic shuttling protein.


Drosophila eIF4AII is shown to interact with both Barentsz and Magoh:Y14 as part of a complex required for oskar mRNA localization during oogenesis. Further, RNAi knockdown of either eIF4AII or MLN51 (the human Barentsz homolog) in HeLa cells leads to an NMD defect. Thus, all four proteins are part of a conserved mRNA localization complex in Drosophila and are essential for NMD, probably as part of the EJC, in mammalian cells.


This paper reports that eIF4AII associates with Y14 and Magoh in vivo, and is part of the in vitro-derived EJC.


34. Li C, Lin R, Lai MC, Ouyang P, Tarn WY: Nuclear Pnn/DR5 protein

- binds to spliced mRNPs and participates in mRNA processing and export via interaction with RNP51. Mol Cell Biol 2003, 23:7363-7376.

This paper demonstrates that Pinin is a nuclear protein and novel EJC component. Expression of antisense Pinin RNA yields partial nuclear accumulation of polyA RNA, indicating a possible mRNA export defect.


44. Gatfield D, Izaurralde E: REF1/Aly and the additional exon


This report shows that RNAi knockdown in Drosophila cells of the EJC protein REF/Aly as well as several other EJC factors yields no significant nuclear accumulation of polyA RNA, suggesting that these proteins are not essential for mRNA export. However, codepletion of RNP51 and REF/Aly or all EJC proteins at the same time does lead to some nuclear RNA accumulation, indicating that these factors can enhance mRNA export.

45. Longman D, Johnstone IL, Caceres JF: The Ref/Aly proteins are

This study uses RNAi in C. elegans to demonstrate that all known REF isoforms are dispensable for mRNA export and normal development, whereas NXF1/Tap is essential.


www.sciencedirect.com
Current Opinion in Cell Biology 2004, 16:279-284


52. Nott A, Meislin SH, Moore MJ: A quantitative analysis of intron effects on mammalian gene expression. RNA 2003, 9:607-617. This study and [51] report quantitative analyses of the effects of splicing on various steps in gene expression and suggest that, in addition to stimulating mRNA production, splicing enhances the amount of protein produced per mRNA molecule.


54. Wiegand HL, Lu S, Cullen BR: Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci USA 2003, 100:11327-11332. This study demonstrates that deposition of an EJC is important for the enhanced translation of spliced mRNA in mammalian cells using short first-exon β-globin constructs. It additionally shows that tethering RNPS1 or SRm160 to intronless β-globin mRNA results in efficient translation of the mRNA. Other experiments clearly demonstrate that the EJC-dependent enhancement in mRNA translation is an effect distinct from EJC-dependent mRNA export.

55. Nott A, Le Hir H, Moore MJ: Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev 2004, 18:210-222. This study shows that the splicing-dependent enhancement of mRNA translation is due to EJC deposition and, in mammalian cells, correlates with a greater percentage of spliced mRNAs being associated with polysomes. Tethering of the EJC proteins Y14, Magoh and RNPS1 or of the Upf proteins Upf1, Upf2 and Upf3b leads to enhanced translation of intronless mRNAs and also correlates with enhanced polysome association.


59. Gehring NH, Neu-Yilik G, Schell T, Hentze MW, Kulozik AE: Y14 and hUpf3b form an NMD-activating complex. Mol Cell 2003, 11:939-949. Using the iN - boxB tethering system, more robust Y14-tethered NMD than was previously reported by Lykke-Andersen et al. [58] is demonstrated. Furthermore, Y14 and a conserved amino acid stretch in Upf3b that interacts directly or indirectly with Y14 is essential for Upf3b-tethered NMD.


64. Gatfield D, Unterholzner L, Ciccarelli FD, Bork P, Izaurrelde E: Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. EMBO J 2003, 22:3960-3970. In this study, it is found that NMD occurs independently of splicing in Drosophila and that RNAi knockdown of several EJC proteins fails to abrogate NMD. Thus, Drosophila NMD seems to be more yeast-like than mammal-like in its mode of PTC recognition.