A new era for the RNA world

Conference: RNA 2000

Introduction

The RNA Society was established in 1993 to facilitate sharing and dissemination of experimental results and emerging concepts in RNA research. The Society encompasses RNA research in the broadest sense: from the ribosome to the spliceosome, from RNA viruses to catalytic RNAs. It is a multidisciplinary society representing molecular, evolutionary and structural biology, biochemistry, biomedical sciences, chemistry, genetics and virology, as they relate to questions of the structure and function of RNA and of ribonucleoprotein assemblies. The annual meeting of the RNA Society, which has more than 800 members, was held from May 30 to June 4 in Madison, WI.

The brain and the heart of the ribosome

Understanding the decoding of genetic information and its translation into polypeptides by the ribosome has posed one of the great challenges in molecular biology. Recent years have witnessed a tremendous acceleration in the elucidation of the structure of the ribosome to the point where central insights into ribosome function are becoming apparent.

Unquestionably, highlights of 'RNA 2000' included reports on the high resolution X-ray structures of the 50S and 30S subunits from *Thermus thermophilus* and cryo-EM work on the 70S ribosome of *Escherichia coli*. While the small 30S subunit accommodates the mRNA decoding center (the 'brain of the ribosome' in A. Yonath's words), the 50S subunit harbors the 'heart', the peptidyl transferase center (PTC). T. Steitz's and P. Moore's groups (New Haven, CT) reported structural data for the 50S subunit at 2.4 Å resolution, which reveal the singular most breathtaking finding: the apparent lack of any proteins within at least 20 Å of the PTC (Ban *et al.*, 2000). 'The ribosome is a ribozyme', as P. Nissen (New Haven, CT) summarized it, to emphasize that the catalytic function of peptide bond formation seems to be catalyzed by elements of the rRNA. His presentation, which focused on the description of a complete model of the ~3000 nucleotides of 23S and 5S rRNAs, also identified the 'peptide exit channel', which is lined by the proteins L4 and L22. This report was complemented by A. Yonath's work (Hamburg, Germany) on the 30S subunit at 3.6 Å resolution, which made possible the identification of the decoding center, subunit interface regions and tRNA binding sites (Schluenzen *et al.*, 2000; Figure 1). Cryo-EM mapping of the *E. coli* ribosome by J. Frank

(Albany, NY) and his co-workers provided insights into dynamic conformational rearrangements. A striking example of this appears to be a rotation of the 'sarcin-ricin loop' by 17° compared with its position in the 50S crystal structure (Gabashvili *et al.*, 2000). Clearly the ribosome, the most complex biological machine for which high resolution structural information is currently emerging, has captured the minds of biologists from numerous disciplines. Before long, we should begin to understand the molecular details of peptide bond formation.

Stops on the go

Both in the nucleus and in the cytoplasm, mRNAs form complexes (mRNPs) with RNA binding proteins that affect RNA processing, transport, translation, stability and localization. An emerging theme of the meeting was the realization that nuclear and cytoplasmic mRNA metabolism are intimately linked and that numerous RNA binding proteins fulfill roles both in the nucleus and in the cytoplasm.

An excellent example of this is a process termed 'nonsense-mediated decay' (NMD), which serves to degrade mutant mRNAs with premature translation termination codons (reviewed by Hentze and Kulozik, 1999). Previous work had identified three yeast (*upf 1–3*) and seven *Caenorhabditis elegans* (*smg 1–7*) genes that play specific and essential roles in NMD. The position of a stop codon relative to the 3' most intron had emerged as a critical determinant of whether the stop codon would be interpreted as being 'premature' or 'normal' in mammalian cells. This gave rise to the concept that splicing might deposit a tag at sites of exon-exon junctions. These tags would remain associated with the mRNAs during nucleo-cytoplasmic transport and could be used to mark a stop codon as premature.

Several groups have now identified proteins that associate with mRNAs during splicing and that distinguish post-spliced from unspliced mRNPs. These include Y14 (G. Dreyfuss, Philadelphia, PA), REF/Aly, SRm160 and DEK (B. Blencowe, Toronto, Canada) (Kataoka *et al.*, 2000; McGarvey *et al.*, 2000). UV-crosslinking and footprinting approaches had been used to identify these proteins, components of a 340 kDa complex that binds ~20–24 nucleotides upstream of a splice junction in a splicing-dependent, but sequence-independent way (H. Le Hir, L.E. Maquat and M.J. Moore, submitted; Le Hir *et al.*, 2000 Figure 2). It is not yet clear how such a large complex can bind with high affinity to a relatively short stretch of RNA in a sequence-independent manner. However, the complex meets

Fig. 1. Two sides of the crystal structure of the small ribosomal subunit from the bacterium *Thermus thermophilus*. The small subunit contains 19 proteins, a small polypeptide of 26 amino acids and an RNA chain, called 16S, of 1518 nucleotides. The structure shown in this figure covers 1457 nucleotides and 19 proteins. RNA is shown in gold, using a ribbon backbone and simple lines for base pairs. The red colored helical segments and loops are the proteins. The structure shows that the decoding center, which positions mRNA and three tRNAs, is constructed entirely of RNA. (A) The view from outside the ribosome (solvent face). (B) The view from the 50S subunit (subunit interface). In both panels, the head is at the top of the drawing and the foot at the bottom (Schlueder et al., 2000).

DEK, REF, SRm160 and RNPS1) may influence the transport, translation or stability of spliced mRNAs.

NMD regulation by phosphorylation/dephosphorylation cycles was addressed in a presentation from the P. Anderson (Madison, WI) laboratory. SMG-2, the *C. elegans* homolog of yeast Upf1p, is a phosphoprotein whose phosphorylated form accumulates in *smg-5*, *smg-6* and *smg-7* mutants, suggesting that the function of these three proteins is required for dephosphorylation. In fact, *smg-5* appears to encode a regulatory subunit of protein phosphatase 2A. On the other hand, SMG-2 accumulates in the dephosphorylated state in *smg-1*, *smg-3* and *smg-4* mutants, suggesting that these factors are required for phosphorylation, potentially by the SMG-1 kinase. While the biochemical roles of phosphorylation/dephosphorylation of SMG-2 remain to be uncovered, the identification of SMG-3/Upf2p and SMG-4/Upf3p homologs in human cells, together with the finding that the human homolog to SMG-2 is a phosphoprotein (laboratories of H. Dietz, Baltimore, MD; L. Maquat, Rochester, NY; and J. Steitz, New Haven, CT), highlights the conservation of these proteins and implies that other components of the NMD pathway are likely to be conserved as well.

Not all stop codons located within the open reading frame of eukaryotic mRNAs trigger NMD. Some UGA codons fulfill the specific function of encoding the amino acid selenocysteine. For a UGA codon to be decoded in this way rather than specifying translation termination a selenocysteine insertion sequence (SECIS) must be present in the 3' untranslated region of selenoprotein mRNAs. The protein SBP2 had recently been found specifically to bind the SECIS and to be important for selenocysteine incorporation. M. Berry (Boston, MA) now reported the identification of a dedicated translation elongation factor, termed EF_{sec}, which specifically binds aminoacylated selenocysteinyl tRNA and bridges to the SECIS via SBP2. The function of the eubacterial protein selB, which binds to the bacterial SECIS and serves as a specific elongation factor, has thus been split and taken over by SBP2 and EF_{sec} in eukaryotes. This may aid the decoding of multiple UGA codons from the 3' UTR of eukaryotic selenoprotein mRNAs.

Role of 3' UTRs in translational regulation: the last will be first

Until recently, the 3' UTR was considered to be a relatively inert appendage of most mRNAs. Now it has become clear that it represents a major relay station for regulation. Not only does it serve to specify selenocysteine insertion but it also plays a major role in determining mRNA localization, mRNA stability and translation. At first glance, the latter may seem surprising, because ribosome assembly occurs at the 5' end of mRNAs. Surprisingly, the tethering of several different translation factors to the 3' UTR stimulates the expression of the respective mRNA (laboratories of M. Wickens, Madison, WI; and S. Fields, University of Washington, Seattle, WA). Might the 3' UTR serve as a pre-assembly site for translation initiation? More conservatively, several reports suggest that the 5' and 3' ends of mRNAs must functionally, and in some cases physically, interact. For example, the translational repressor protein sex-lethal (SXL) must be bound to specific sites of both the 5' and the 3' UTR of *msl-2* mRNA to repress the synthesis of MSL2 protein in female flies, an event that is biologic-

one prediction for the hypothetical 'splicing tag' for NMD, suggesting that this complex or some of its components (Y14,

Fig. 2. SRm160, DEK, Y14, REFs and RNPS1 are components of a large protein complex that preferentially associates with spliced mRNAs (H. Le Hir, L.E. Maquat and M.J. Moore, submitted). These proteins may recruit factors implicated in mRNA export. The precise order of association of these proteins with the mRNA is unknown. In the cytoplasmic side of the NPC, this protein complex, or some of its constituents, may provide clues to define stop codons (red asterisks) as premature (1) or normal (2). Even though direct evidence for a role of some of these proteins in RNP export or NMD has yet to be demonstrated, such evidence is expected to be forthcoming.

ally critical to enact dosage compensation (M. Hentze, Heidelberg, Germany). Also, in the barley yellow dwarf virus, a 3' UTR translational enhancer sequence is critical for the translation of the uncapped and non-polyadenylated viral RNA. Apparently, 5' and 3' UTR sequences must hydrogen bond with each other to allow the enhancer to function (W.A. Miller, Ames, IA).

While many 3' UTR functions are mediated by binding proteins such as EF_{Sec} or SXL, unusual examples of target mRNA regulation like that by lin-4 antisense RNA in *C. elegans* have long been known. At this meeting, evidence was presented that lin-4 RNA regulates specifically the translation of the lin-14 and lin-28 mRNAs, and that this occurs at the level of translation elongation or termination (E. Moss, Philadelphia, PA). Moreover, the 21 nt RNA let-7 was reported as another example of a short antisense RNA that controls the expression of a target mRNA (lin-41). Intriguingly, homologs of let-7 were found in *Drosophila* and even humans, suggesting that this mode of regulation may be more widespread than previously anticipated (G. Ruvkun, Boston, MA).

Splicing: how many more factors to go?

Traditionally, a significant fraction of the RNA meeting is devoted to nuclear pre-mRNA splicing. Several groups reported on the identification of new mammalian and yeast factors

involved in splicing. With the purification of complexes from yeast that may represent pre-assembled apo-spliceosomes that lack only a pre-mRNA substrate and the identification of its constituents (J. Abelson, CA), one might expect that most of the yeast splicing factors have now been identified. However, much work is still required to confirm the participation of the newly identified factors in pre-mRNA splicing and to pinpoint their exact functions.

A major function of splicing factors is to recognize the degenerate sequences located at the splice sites and branchpoints of all introns. Previous studies have shown that recurrent recognition of these sequences ensures splicing accuracy, even if the target sequences are only partially conserved. This idea was further strengthened by the work of several groups on branchpoint recognition. Early recognition of the branchpoint is initiated by the SF1/BBP protein. This factor, together with U2AF and other proteins, mediates early recognition of the intron 3' end (A. Krämer, Geneva, Switzerland; J. Valcarcel, Heidelberg, Germany). Interestingly, studies in *C. elegans* and yeast have now revealed that this recognition not only contributes to splicing complex assembly but is also essential *in vivo* to prevent pre-mRNA leakage into the cytoplasm (T. Blumenthal, Denver, CO; B. Séraphin, Heidelberg, Germany). In yeast and mammalian systems, branchpoint recognition also occurs through base-pairing of the U2 snRNA with the pre-mRNA. Surprisingly, analyses performed in trypanosomes indicate that this feature is not conserved in this organism (A. Bindereif, Giessen, Germany). To add to this complexity, the groups of C. Query (Cambridge, MA) and R. Lührmann (Goettingen, Germany) reported the purification of a new protein that interacts with the branchpoint in the human system, and the characterization of a related protein from yeast. Interestingly, the human protein is shared by the two splicing machineries that co-exist in human cells, even though the corresponding branchpoint sequences are different. While no definitive mechanism for branchpoint recognition emerges from these studies, it is clear that this event requires the sequential action of several factors to achieve high accuracy.

Bypassing essential RNA helicases

A recurring theme at the meeting was the observation that putative RNA helicases required for splicing under normal conditions may become dispensable if a second splicing factor is defective. An early event during splicing complex formation involves binding of the U1 snRNA to the 5' splice site. Transition from these early complexes to a full spliceosome involves displacement of the U1 snRNP bound to the 5' splice site. In yeast, this has been shown to require the Prp28, an essential protein that belongs to the DexH/D family of putative RNA helicases. Surprisingly, PRP28 was found to be dispensable in a strain carrying certain mutations of U1C (a U1 snRNP subunit) that destabilize the U1 snRNA-pre-mRNA interaction (T.-H. Chang, Columbus, OH). Similarly, the Sub2 helicase, which is essential for early steps of spliceosome assembly (C. Guthrie, UCSF, San Francisco, CA; R. Green, Baltimore, MD; D. Libri, Gif-sur-Yvette, France), was shown to be dispensable in a strain lacking the *MUD2* gene encoding the yeast U2AF65 homolog (C. Guthrie). In the mammalian system, ATP hydrolysis, which normally occurs during association of U2 snRNP with the

branchpoint, can be bypassed when an artificial substrate devoid of sequence upstream of the branchpoint is used. As the putative RNA helicase Prp5 is likely to mediate ATP hydrolysis at this step of spliceosome assembly, this result further suggests that an essential RNA helicase may become dispensable if a specific constitutive region of the pre-mRNA substrate is altered (C. Query).

Structure of the spliceosome: factor by factor

Although the structure of the spliceosome is lagging behind that of the ribosome, remarkable progress has been made in determining the structure of some of its individual components. The structure of a 15.5 kDa U4 snRNP protein bound to its cognate RNA binding site was reported (R. Lührmann). Interestingly, this protein contains a conserved domain that is also found in yeast ribosomal protein L32. More unexpectedly, the two proteins bind related, though not identical, RNA structures. L32 binds not only to ribosomal RNA but also to its own pre-mRNA and, through this latter interaction, regulates pre-mRNA splicing (J. Warner, New York, NY). Surprisingly, the 15.5 kDa U4 snRNP protein was also reported at the meeting to be a subunit of the box C/D snoRNPs, which are involved in rRNA 2'-O-methylation (R. Lührmann; C. Branlant, Nancy, France; M. Rosbash, Waltham, MA). This finding reveals an unexpected link between pre-mRNA splicing and ribosomal RNA modification.

A class of proteins involved in various cellular processes including splicing are called Sm proteins. These proteins form the core structure of several snRNPs and have been extensively characterized. Cross-linking analyses revealed that several Sm proteins contact specific bases in snRNAs through a conserved portion of the Sm domain (R. Lührmann). Interestingly, when these proteins form part of the U1 snRNP, some of the C-terminal extensions contact the pre-mRNA, thereby favoring early splicing complex assembly (M. Rosbash). In addition to canonical Sm proteins, eukaryotic cells have also been shown to contain a set of Sm-like proteins that associates with two distinct heptameric complexes, one of which interacts with the U6 snRNA, and the other was reported to be involved in specific steps of mRNA degradation (J. Beggs, Edinburgh, UK; R. Parker, Tucson, AZ; B. Séraphin). Sm proteins are also found associated with the U7 snRNP, which is involved in histone mRNA 3' end processing. Surprisingly, however, in this complex one of the canonical Sm proteins is substituted by a specific U7 Sm protein. This probably contributes to the recognition of a slightly different cognate sequence present in the U7 snRNA (D. Schümperli, Bern, Switzerland). Sm-related proteins have also been identified in archaea. Structural analysis by crystallography (D. Suck, Heidelberg, Germany) and electron microscopy (R. Lührmann) has demonstrated that these proteins are able to form homo-heptameric structures that interact with RNA *in vitro* and *in vivo*. These prokaryotic proteins may provide clues regarding the origin of eukaryotic Sm and Sm-like proteins.

Splicing regulation and feedback loops

Although splicing regulation is not used frequently in yeast, splicing of some transcripts is limited to meiosis. Characteriza-

tion of the corresponding introns has revealed the presence of a specific splicing enhancer located between the 5' splice site and the branchpoint (M. Ares, Santa Cruz, CA). A meiosis-specific regulator (Mer1) and a U1 snRNP protein (Nam8) are specifically required for these regulated splicing events. While the role of Nam8 in pre-mRNA splicing has been well characterized, no functional human counterpart to this protein was previously known. A recent study has, however, revealed that the related human Tia-1 protein functions like Nam8 in higher organisms, facilitating U1 snRNP binding through interaction with intron sequences downstream of the 5' splice site (J. Valcarcel).

Alternative splicing also affects pre-mRNAs that encode splicing factors and RNA binding proteins, thereby generating related isoforms with different functions. Three examples of regulated alternative splicing events were reported. These involved the creation of a feedback loop by the product of the corresponding mRNA, thus autoregulating the production of the corresponding protein. This is the case for hnRNP A1 (B. Chabot, Quebec, Canada), the splicing factor SC35 (J. Soret, Montpellier, France) and the splicing regulator Htra2 (S. Stamm, Martinsried, Germany). These autoregulatory loops probably help to maintain cellular homeostasis by keeping the concentration of the corresponding factors essentially constant.

Extensive alternative splicing occurs in the brains of mammals. Identification of specific splicing regulators affecting these events is underway. Among the known regulators, the Nova-1 protein affects splicing of a glycine receptor and other neuronal pre-mRNAs (J. Darnell, New York, NY). mRNAs targeted by this protein were identified by phenotypic analysis of transgenic mice deficient for the Nova-1 protein, and by the identification of potential Nova-1 binding sites. These *in vivo* studies were nicely complemented by structural analyses revealing the interaction mode of the Nova-1 KH-domain with its cognate binding site (S. Burley, New York, NY). As for constitutive splicing, protein phosphorylation appears to contribute to splicing regulation in the brain. A striking example is provided by the Slo potassium channels, changes in whose splicing patterns are induced by calcium signals. This regulation appears to be mediated by a specific calmodulin-dependent protein kinase that links specific changes in extracellular calcium levels with the alteration of endogenous transcripts (D. Black, Los Angeles, CA).

A myriad of non-coding small RNAs in the brain

Searching for factors involved in neuron-specific alternative splicing, the group of P. Grabowski made the surprising observation that a small RNA associates specifically with a pre-mRNA in a manner consistent with a role for this molecule in splicing regulation. This interaction occurs in a conserved region of the pre-mRNA and is blocked by the protein splicing regulator PTB. While further work will be required to prove a competition model between these molecules in splicing regulation, the identification of hundreds of small RNA species in the mouse brain (A. Hüttenhofer, Münster, Germany) suggests that a role of non-coding small RNAs in various neuronal functions should be seriously considered.

Perspectives

The RNA 2000 meeting was indubitably exceptional. Without question, the highlight can be summarized in a single sentence: the ribosome is a ribozyme. The structure of the ribosome not only opens new avenues in understanding protein synthesis, but also offers a real feast for structural biologists to analyze new ways in which RNAs can fold, interact with proteins and be catalytically active. Bioinformatics has made a strong entry into the RNA field, and it seems to be a safe prediction that this discipline will engage into an even closer symbiosis with RNA biologists. Finally, today's 'RNA World' really presents itself as an 'RNP World', with fertilizing winds blowing across the entire RNA landscape from RNA processing to translation and with a growing appreciation of the biology and regulation of 'RNP remodeling'. The next meeting of the RNA Society, in May 2001 in Banff, is eagerly awaited by many.

Acknowledgements

We apologize to the many participants whose work is not discussed due to space limitations. We wish to thank the many colleagues who helped in the preparation of this report by providing us with unpublished information and critical comments. We also thank I.W. Mattaj and the anonymous referees for constructive comments. We are also grateful to all members of the RNA Society for making this stimulating meeting possible every year.

References

Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, **289**, 905–920.

- Gabashvili, I.S., Agrawal, R.K., Spahn, C.M., Grassucci, R.A., Svergun, D.I., Frank, J. and Penczek, P. (2000) Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. *Cell*, **100**, 537–549.
- Hentze, M.W., and Kulozik, A.E. (1999) A perfect message: RNA surveillance and nonsense-mediated decay. *Cell*, **96**, 307–310.
- Kataoka, N., Yong, J., Kim, V.N., Velazquez, F., Perkinson, R.A., Wang, F. and Dreyfuss, G. (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.
- Le Hir, H., Moore, M.J. and Maquat, L.E. (2000) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon–exon junctions. *Genes Dev.*, **14**, 1098–1108.
- 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.
- Schlutzen, F. *et al.* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, **102**, 615–623.

Matthias W. Hentze, Elisa Izaurralde⁺ and Bertrand Séraphin

European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69012 Heidelberg, Germany

Received September 4, 2000; revised October 2, 2000;
accepted October 4, 2000

⁺Corresponding author. Tel: +49 6221 387 389; Fax: +49 6221 387 518;
E-mail: izaurralde@embl-heidelberg.de

DOI: 10.1093/embo-reports/kvd100