Development 105, 183–189 (1989)
Printed in Great Britain © The Company of Biologists Limited 1989

Regulated splicing in early development and stage-specific U snRNPs

IAIN W. MATTAJ and JÖRG HAMM

European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, West Germany

Key words: RNA, splicing mechanism, endogenous gene, U snRNP, ribonucleoprotein, *Xenopus*, mouse.

Introduction

Analysis of the splicing of several different genes has led to the discovery of RNA-processing events that are specific for certain developmental stages. Although some alternative splicing events are clearly regulated by very specific mechanisms, the fact that exogenous (viral) as well as endogenous genes display stage-specific splicing patterns makes it likely that the basic splicing machinery itself differs in composition or activity at different times in development. U snRNPs (small nuclear ribonucleoproteins) are essential components of the complexes in which intron removal takes place. Evidence has been accumulating for several years that U snRNPs, in particular their RNA components the U snRNAs, are expressed in some species in a developmentally regulated manner. In this review, we will discuss this work and its possible significance for stage-specific RNA processing. In addition, we discuss several of the best-understood cases of alternative splicing which are regulated by mechanisms that appear to be more transcript-specific.

The splicing process

A model of the steps and factors involved in intron removal in nuclear extracts of mammalian culture cells is presented in Fig. 1. Since detailed treatment of this topic is not within the scope of this review, readers are referred to Dreyfuss *et al.* (1988) as a source of references to several recent reviews and the original literature. For our purposes, a few aspects of the process are important. The reaction takes place in two steps which are distinguishable *in vitro* but probably concerted *in vivo*. In the first step, the upstream exon–5′ splice site junction is cleaved and the G at the 5′ splice site joined through a 2′–5′ phosphodiester linkage to the A at the branch point to form a lariat-shaped intermediate. In the second step, the 3′ splice site–downstream exon junction is cleaved and the two exons are joined. The intron is released as a lariat product. As indicated in Fig. 1, several *trans*-acting factors takes part in these reactions. These can be divided into two broad classes: the U snRNPs, evolutionarily highly conserved nuclear RNP particles, and protein factors. In higher eukaryotes, the major U snRNPs, most of which are essential for splicing, are highly abundant. They consist of one (U1, U2, U5) or two (U4/U6) RNA molecules, a group of common proteins and, in most cases, proteins unique to individual U snRNPs (Reddy & Busch, 1988; Lührmann, 1988). The protein factors are less well characterized and, at least in some cases, they may even represent U snRNP proteins that have been dissociated from the RNP particle during purification. The snRNPs whose functions in splicing are best understood are U1 and U2. These snRNPs interact directly with the 5′ splice site and the branch point, respectively, in a process which is at least partly dependent on RNA–RNA base pairing. They are thus presumed to be involved in the identification or choice of these two sites. Although U5 snRNP probably interacts with the 3′ end of the intron (Chabot *et al.* 1985), this interaction and the exact function of the U4/U6 snRNP are not well understood.

Alternative splicing

As soon as a primary transcript has an intron, the possibility of producing alternative mature transcripts exists. If an RNA has a single intron, mature mRNAs either with or without the intron can be generated. This could lead to the production of two different protein products from the same primary transcript. Complex transcripts containing multiple introns have the potential for producing many alternative transcripts. This can occur in several ways (Fig. 2). There are many examples of genes whose expression is modified by the choice of alternative splicing pathways. (Again, this topic is outside our scope, see Leff & Rosenfeld, 1986 for a recent review.) In the majority of cases examined, the production of a particular mRNA species from such complex genes is regulated either tissue-specifically or developmentally. In one very well studied case, that of the SV40 T antigen (see below), the production of alternatively spliced products correlates with the utilization of different branch points in a single intron (Noble *et al.* 1987). Unfortunately, the study of other genes has not progressed to the stage where it is possible to say whether or not this is part of a common mechanism utilized in the production of alternatively spliced transcripts.

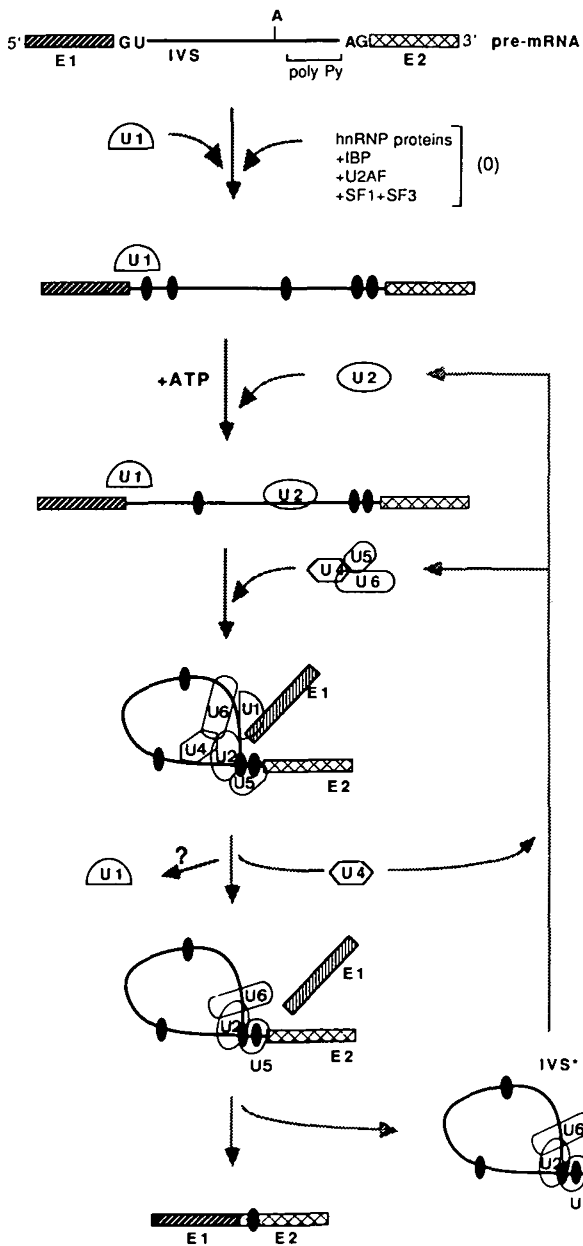

Fig. 1. The splicing pathway of pre-mRNA in mammalian cells. IVS, intervening sequence (intron); *IVS**, intervening sequence in lariat form; *poly py*, polypyrimidine stretch; E, exons; *IBP*, intron-binding protein(s); *U2AF*, U2 snRNP auxiliary factor; *SF1*, *SF3*, splicing factors 1 and 3. The question mark at the exit point of U1 snRNP indicates the uncertainty as to the stage at which U1 snRNP dissociates. (Reproduced with permission from Dreyfuss *et al.* 1988.)

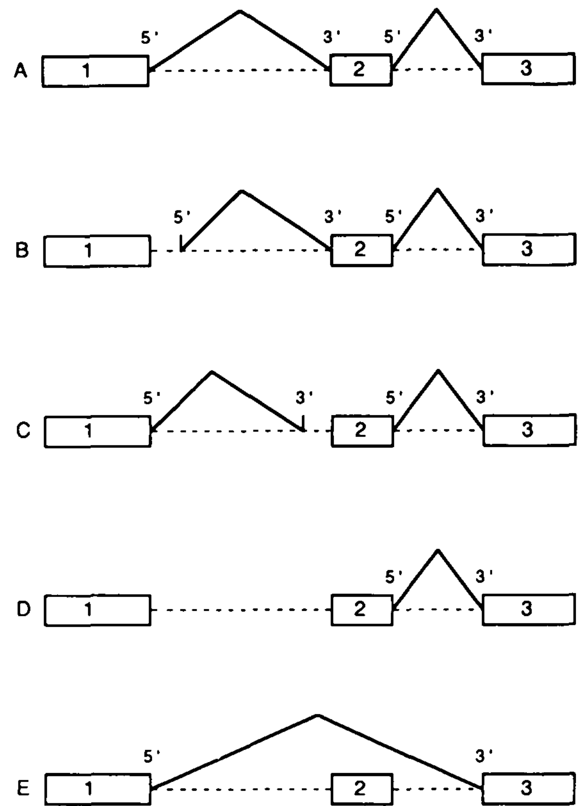

Fig. 2. Some possible ways to produce alternative transcripts from a precursor containing two introns. Boxes numbered 1, 2, 3 are exon sequences and dashed lines intron sequences in example A. Solid lines join splice junctions. (A) Both introns spliced. (B) An alternatve 5′ splice site is used in the first intron. (C) An alternative 3′ splice site is used in the first intron. (D) The first intron is not removed. (E) The 5′ splice site of the first intron is joined directly to the 3′ splice site of the second intron deleting exon 2 (Exon skipping). Depending on the disposition of open reading frames and translational stop codons some or all of the transcripts produced can give rise to different protein products.

Developmentally regulated (alternative) splicing

We will now briefly review a few particular cases of regulated alternative splicing and then describe some developmentally interesting examples. These examples were chosen since their occurrence correlates temporally with the appearance of stage-specific U snRNPs which will be described later in the review.

Sex determination in *Drosophila* is mediated by a hierarchy of regulatory genes whose expression is dependent on the X chromosome: autosome ratio (1:1 in females, 1:2 in males) and whose action brings about terminal sex differentiation (Baker & Belote, 1983; Cline, 1985). The recent molecular cloning of many of the genes involved (Maine *et al.* 1985; Butler *et al.* 1986; Baker & Wolfner, 1988) has led to the discovery that they produce sex-specific transcripts. This was not entirely expected since several of the genes in the hierarchy are not required for male development. Analysis of these sex-specific transcripts is proceeding; currently best studied are the RNA products of a gene called *transformer* (*tra*) whose expression is required for female differentiation. Two transcripts are produced from the *tra* locus, one found in both sexes, the other female specific. Only the female-specific transcript contains a long open reading frame (Boggs *et al.* 1987). In the ubiquitous transcript, the open reading frame is destroyed by the choice of an alternative 3′ splice site in

the first intron. Thus, only the female-specific mRNA is proposed to be capable of giving rise to the *tra* protein product (Boggs *et al.* 1987). The *tra* protein is then proposed to modify, in females, the expression of other genes in the pathway that results in sexual differentiation. The ability to make the female-specific splicing choice in *tra* depends on the expression of the *Sex lethal* (*sxl*) gene, which is situated above *transformer* in the regulatory hierarchy (Nagoshi *et al.* 1988), i.e. *tra* splicing is regulated directly or indirectly by *sex lethal* gene products.

Two other related examples are the regulated splicing of transcripts of the *Drosophila* gene encoding suppressor of white apricot, $su(w^a)$ and of the *Xenopus* LI ribosomal protein gene. In both cases, accumulation of the protein products results in a block of RNA processing steps, possibly by a feedback mechanism.

The *Xenopus* LI gene has nine introns (Bozzoni *et al.* 1984). In circumstances when excess LI protein is produced, such as after microinjection of the gene into *Xenopus* oocytes, the accumulated free LI causes the removal of two of these introns to be inhibited (Bozzoni *et al.* 1984). The resulting, partially processed, precursor RNAs are destabilized by a specific endonuclease which recognizes the unspliced introns (Caffarelli *et al.* 1987).

Three $su(w^a)$ transcripts accumulate at specific developmental stages (Chou *et al.* 1987). The two longer transcripts retain one or two of the seven introns found in this gene. The presence of these introns results in the closure of the functional $su(w^a)$ open reading frame (Zachar *et al.* 1987). From genetic evidence it appears that the $su(w^a)$ protein product is required for the block to splicing. This suggests that, as in the case of protein LI, intron retention is part of an autoregulatory mechanism that controls the level of the $su(w^a)$ protein. Although the mechanism of splicing regulation is not understood in any of the above examples, it is perhaps suggestive that the putative $su(w^a)$ protein sequence displays similarity to part of the U1 snRNP specific 70K protein (Bingham *et al.* 1988; Theissen *et al.* 1986).

The three cases of alternative splicing mentioned so far all appear to utilize specific regulatory mechanisms. We will now discuss two cases of germ cell and early embryo-specific processing reactions. The examples chosen are the splicing of transcripts of the viral SV40 T antigen and of the *Drosophila* P element transposon. In the former case, splicing is cell-type- and perhaps also differentiation state-specific and in the latter it is germline-specific. These examples are particularly interesting since both transcripts are required for an essentially parasitic lifestyle. This increases the likelihood that the alternative RNA processing observed occurs as a result of differences in the basic splicing machinery of the host cell rather than by transcript-specific regulatory mechanisms.

P elements are a family of transposable elements found in *Drosophila*, which vary in size. Transposition of P elements is frequent in the germ line, but has not been detected in somatic cells (Engels, 1983). Autonomous P elements, capable of transcription in the absence of other P elements, are 2·9 kb long and encode transposase, an activity required for transposition (O'Hare & Rubin, 1983; Engels, 1984). Due to the scarcity of germ cell P element transcripts, it is not feasible to study them directly. However, in an elegant series of experiments Laski *et al.* (1986) provided strong evidence for the fact that transposase production required a germline-specific splicing event. Two introns are removed from P elements in all tissues but Laski *et al.* showed that the removal of a third intron was required before a functional transposase product could be produced. This intron is not normally removed in somatic tissues. However, when mutant P elements were constructed from which the putative third intron was removed these elements transposed actively in somatic cells. Conversely, mutations introduced into the putative splice junctions of the third intron prevented transposition from occurring even in germ cells. These data provided strong circumstantial evidence that this intron must be excised specifically in the germ line to allow the expression of a functional transposase and subsequent P element transposition. The mechanism of regulation of splicing of the third intron is not known, but could either result from specific inhibition of splicing in somatic cells or specific activation in germ cells.

The other example of a stage-specific alternatively processed precursor mRNA, the mRNA encoding the SV40 T-antigen, has been analysed in great detail at the molecular level. Processing of the T-antigen precursor mRNA can result in small t mRNA or large T mRNA. In large T mRNA about 300 nt of exon sequence present in small t mRNA is missing because of the joining of alternative 5′ splice sites to the same 3′ splice site (Tooze, 1981). The resulting t mRNA has much less protein-coding capacity than T mRNA because a translational stop codon is generated (Fig. 3). The ratio of T/t mRNA observed is dependent on the cell type in which the splicing takes place. In HeLa cells, the ratio of large T to small t is about 5 to 1 whereas in oocytes or in 293 cells (an Adenovirus transformed human cell line) the ratio is the inverse (Fu & Manley, 1987). When processing is analysed in HeLa nuclear extracts, the large T splice site is used almost exclusively (Noble *et al.* 1986). The results of studies addressing the reasons for the observed ratios suggest that *trans*-acting factors are involved.

First, different branchpoints are used with different efficiencies for large T splicing in different cell types. In HeLa cells the major branchpoint is located at position -19, relative to the 3′ splice site, and minor branchpoints at -22, -28 and -31/32 are also used. In contrast, in oocytes or in 293 cells, the branch site at -31/32 is predominantly used in the production of large T mRNA. For the splicing of small t mRNA, a single branch point at position -18 is utilized (Fig. 3B; Noble *et al.* 1987). Since U2 snRNP is likely to interact with the branch site (Black *et al.* 1985; Parker *et al.* 1987), it is possible that U2 snRNP structure or factors required for the binding of U2 snRNPs to the branch site (reviewed in Dreyfuss *et al.* 1988) might vary in differ-

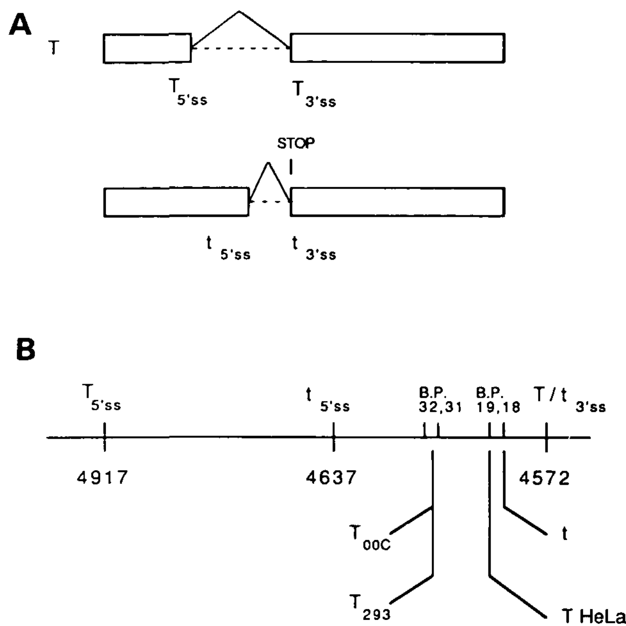

Fig. 3. (A) A diagrammatic representation of the splicing events that generate large T and small t antigen transcripts. The small t open reading frame termination codon (STOP) is indicated. (B) A more detailed diagram of the large T/small t intron. The 5′ and 3′ splice sites are shown as are the major branch points (B.P.) used in the production of t or T mRNAs in various cell types (see text). The numbers at the bottom indicate the positions of the splice sites in the SV40 genome (Tooze, 1981).

ent cell types.

Second, small t intron size (66 nt) is below the minimum length found to be necessary for efficient splicing (75–80 nt) (Wieringa *et al.* 1984). This minimum intron size is thought to reflect the spatial requirements of factors that bind to the intron. Consistent with this model is the fact that the utilization of the small t 5′ splice site is greatly increased (~15 fold) in all cell types by expanding the small t intron to 77 or more nucleotides (Fu & Manley, 1987). It is therefore possible that splicing of the very short small intron requires a *trans*-acting factor that is more abundant in oocytes and 293 cells than in HeLa cells. The observation that only very little small t mRNA is generated in *in vitro* splicing reactions independently of whether HeLa or 293 nuclear extracts are used (Noble *et al.* 1987) could be explained by the loss of the required factor during extract preparation.

Third, the small t 5′ splice site has a better match to the consensus 5′ splice site sequence than the large T 5′ splice site (8/9 and 6/9, respectively). As a consequence, binding of U1 snRNPs to the small t splice site is favoured. When small and large T splice sites were inserted at identical positions into precursor RNAs the small t 5′ splice sites was chosen three times more frequently than large T (Fu & Manley, 1987). Gradually increasing the match of the large T splice site to the consensus sequence reduced this difference (Zhuang *et al.* 1987). Differential affinity for U1 snRNP is perhaps also the cause of the relative insensitivity of small t compared to large T splicing to inhibition by anti-U1 snRNP antibodies (Fradin *et al.* 1984). Differences in the concentration or composition of U1 snRNPs in different cell types could therefore affect which 5′ splice site is chosen. More generally, all the molecular differences that correlate with changes in the alternative splicing of SV40 T-antigen precursor mRNA in different cell types might be explained by parallel changes in the quantity or quality of *trans*-acting factors. As described above, the U snRNPs are predominant among those *trans*-acting factors. We will now discuss the production of developmental variants of U snRNAs and U snRNPs.

Biosynthesis of U snRNPs during oogenesis and embryogenesis in *Xenopus*

Since the bulk of the experimentation relating to stage-specific variation in U snRNP composition has been carried out with *Xenopus*, we will begin with some background information on U snRNP biogenesis during the early development of this organism. This subject has previously been reviewed more extensively (Mattaj *et al.* 1985). U snRNPs are composed of one or two RNAs and up to ten different proteins. In somatic cells, roughly equivalent amounts of protein and RNA components appear to accumulate resulting in the presence in the nucleus of between 2×10^5 and 10^6 copies of each of the major U snRNPs. In different somatic cells examined, U1 is always the most abundant U snRNP and the other major U snRNPs are present in different cell types at similar relative ratios (see Reddy & Busch, 1988 and Lührmann, 1988 for recent reviews of U snRNAs and U snRNP proteins). During early *Xenopus* development, U snRNP biosynthesis is different. U snRNAs and U snRNP proteins accumulate non-coordinately in oogenesis, resulting in the accumulation of a relatively low concentration of nuclear U snRNPs and a large pool of excess, free, U snRNP protein (De Robertis *et al.* 1982; Zeller *et al.* 1983; Forbes *et al.* 1983; Fritz *et al.* 1984). Free U snRNP proteins are cytoplasmic, and can only enter the nucleus after binding U snRNA to form a U snRNP (De Robertis *et al.* 1982; Zeller *et al.* 1983; Mattaj & De Robertis, 1985). Following the midblastula transition, at the 4000- to 8000-cell stage, very active U snRNA gene transcription takes place (Newport & Kirschner, 1982; Forbes *et al.* 1983; Fritz *et al.* 1984). Within a matter of a few hours sufficient U snRNA is synthesized to combine with the maternal store of U snRNP proteins. The complete U snRNPs then translocate to the nucleus (Zeller *et al.* 1983) where they are required for the processing of zygotic transcripts. For the purposes of later discussion, it should be noted that the low concentration of U snRNPs in the oocyte nucleus (5–10 % of somatic levels being a maximal estimate; Forbes *et al.* 1983; Fritz *et al.* 1984) is perfectly capable of supporting normal (somatic-type) splicing of some exogenous transcripts (e.g. Green *et al.* 1983) but not of others (see above).

This pattern of developmental regulation is not universal. For example, in several sea urchin species maternal U snRNPs do not appear to be reutilized in the embryo, since they remain in the cytoplasm during early development (Nash *et al.* 1987). The nuclei of gastrula stages contain U snRNAs, but they are the products of zygotic, not maternal, transcription, indicating that embryonic nuclei contain only U snRNPs assembled on newly synthesized U snRNAs.

In *Xenopus*, as in most other eukaryotes studied, U snRNA genes are encoded by multigene families (Mattaj *et al.* 1985). These families include members with minor variations in sequence. Using these variations as markers, it has been possible to show that different U1 and U4 genes are expressed at different times in oogenesis and embryogenesis (Forbes *et al.* 1984; Lund & Dahlberg, 1987). Different U1 and U4 variants expressed during adult life, during early embryogenesis, or during tadpole stages have been identified and sequenced. Interestingly, although no variation in the sequence of U2 has been found, the modification state of U2 RNA changes during embryonic life (Lund & Dahlberg, 1987). Of several adult tissues tested only ovary and testis contained detectable levels of embryonic U1 and U4 variants. The presence of embryonic U snRNAs may therefore be related to the identification (Fritz *et al.* 1984) of two proteins recognized by anti-Sm antisera which were found in ovary and testes, but not in a variety of other *Xenopus* tissues. Anti-Sm antisera recognize several U snRNP proteins (Lerner & Steitz, 1979). The germline-specific proteins described by Fritz *et al.* have not, however, yet been shown to be U snRNP-associated. A more detailed study of this point might well be rewarding since McAllister *et al.* (1988) have recently identified a tissue-specific U snRNP protein in another context. They demonstrated the existence in rat culture cell lines and tissues of a neural cell-specific U snRNP protein, designated N, which is recognized by some anti-Sm antisera, and which they speculate may have a role in neural-specific splicing pathways. It is not clear whether N is an additional U snRNP polypeptide or whether it substitutes for one of the polypeptides found in other tissues.

A preliminary analysis of U snRNPs in various *Xenopus* tissues has been made and it has been found that there are indeed differences between the U snRNPs found in germ cells and in somatic cells. Both the relative abundance and apparently also the composition of the major U snRNPs is distinctly different in germline and somatic cells (our unpublished data). Thus the changes in U snRNAs (described below) may reflect more general alterations in the splicing machinery which are stage- or tissue-specific.

Differential expression of variant U snRNAs during mammalian development

The expression of embryonic U1 and U4 variants in *Xenopus* occurs after midblastula and in previtellogenic oocytes, times when an unusually large amount of U snRNA synthesis is required. It is therefore possible to argue that embryonic U snRNA genes, which are present in many more copies than the adult genes, are expressed in response to heavy demand for U snRNA. A more interesting possibility is that the variants might make functionally different U snRNPs which are specifically required at certain times in development. In order to probe further the significance of variant U1 snRNAs Lund, Dahlberg and their colleagues analysed U1 snRNA expression in different mammalian species, mainly mouse and human. The differences between mammalian and amphibian oogenesis and embryogenesis are such that no periods equivalent to those during which there is a high U snRNA synthetic requirement in frogs are discernable in mammals. Nevertheless, stage-specific U1 variants were also found when U1 synthesis in the mouse was examined (Lund *et al.* 1985). This supports the suggestion that U1 variants might have functional significance related to their particular sequences. In the mouse, two major classes of U1 snRNA were found. The first class was expressed ubiquitously while the second, analogous to *Xenopus* embryonic U1, was only found during early embryonic stages, in germ cells and in a few tissues (e.g. spleen, thymus) in which a significant number of stem cells capable of further differentiation were present. When U1 expression was examined in various cultured mouse cell lines (Lund *et al.* 1985), it was found that the type of U1 snRNA accumulated reflected the source of the cells. Cell lines of embryonal origin and transformed cell lines expressed high levels of embryonic U1, while untransformed lines and highly differentiated teratocarcinoma cells expressed little, if any, embryonic U1. These data would support the idea of a significant role for variant U1 snRNAs. However, data obtained from the examination of U1 expression in human cell lines (Lund, 1988) have indicated that there is no obligatory association between embryonic development and particular U snRNA variants. Although human cell lines contain several U1 variants, there is not a class whose pattern of expression corresponds to the mouse or *Xenopus* embryonic U1 snRNAs (Lund, 1988). Intriguingly, in all three species examined a single U1 residue, A70, is found to be ribose-methylated in adult tissues or differentiated cells but not in embryonal tissues or cells. None of the U1 variants so far sequenced, with the possible exception of the human a-2 variant in which G37 is changed to U (Lund, 1988) are altered in positions required for the assembly of U1 snRNP (Hamm *et al.* 1987, 88). This indicates that all the different U1s are capable of associating with the standard complement of U1 snRNP proteins, but does not rule out the possibility that some variant U1s are associated with additional, specific, proteins.

The pattern of expression of *Xenopus* embryonic U1 snRNAs has been shown to be accurately reproduced when cloned genes are microinjected into oocytes or eggs (Lund *et al.* 1987). Also, a mouse embryonic U1 gene has been shown to be accurately expressed in a transgenic mouse line (E. Lund, personal communication). These results indicate that all the information

required for correct developmental expression of U1 variants is encoded in short DNA segments surrounding the genes and does not require the tandem or clustered arrangement of genes observed in most species.

Prospects

If it is true that some examples of alternative splicing in different cell types are due to the composition and/or concentration of U snRNPs in these cell types, experimental alteration of U snRNP composition might result in a consequent alteration in splicing patterns. The *Xenopus* oocyte or embryo seem to be ideal experimental systems for testing this hypothesis. Purified U snRNPs from, for example, HeLa cells can be introduced into the oocyte by microinjection. Their effect on splicing of a suitable precursor, like the SV40 T antigen, can then be tested. Of particular interest for such experiments is the recent demonstration (Pan & Prives, 1988) that the endogenous oocyte U1 and U2 snRNPs can be destroyed by injection of a DNA oligonucleotide complementary to U1 or U2 snRNA. The resulting DNA-RNA hybrids are substrates for the oocyte RNase H. After a short time, endogenous deoxyribonucleases destroy the remaining, excess, oligonucleotide. Later, the U snRNA can be reintroduced into the oocyte by microinjection of purified U1 or U2 snRNA. The RNA assembles with the stored snRNP proteins (see above) and splicing is restored. If these results can be generalized to other snRNAs then a chosen endogenous U snRNA can be replaced by other sequence variants or be reduced or augmented in concentration. What is now required are a series of suitable substrate transcripts on which the effect of these alterations in the U snRNPs can be tested. A very similar approach was pioneered in the work of Birnstiel and his collaborators with U7 snRNA (Birnstiel & Schaufele, 1988). This RNA was discovered due to its ability to complement the incapacity of *Xenopus* oocytes to process correctly the 3′ ends of sea urchin histone H3 mRNA. In principle any cell-type-specific component of the splicing machinery might be identified in this way.

Conclusions

Several examples of regulated alternative splicing have been discussed, indicating that specific mechanisms for controlling alternative RNA processing exist. We have also described some cases of differential splicing that are germ-cell- or differentiation-state-specific. In *Xenopus* and mouse, the differential production of specific variant U snRNAs and U snRNPs correlates with these differential splicing events. This correlation suggests that some examples of alternative RNA processing may result from the presence of tissue- or stage-specific components of the splicing machinery rather than from transcript-specific mechanisms. Direct experimental evidence for or against this hypothesis is lacking at present. The circumstantial evidence should, however, be sufficiently interesting to encourage further research in this area.

We thank Graham Tebb, Huw Parry, Daniel Scherly, Angus Lamond, David Tollervey, Elsebet Lund and James Dahlberg for their comments on the manuscript.

References

BAKER, B. S. & BELOTE, J. M. (1983). Sex determination and dosage compensation in *Drosophila melanogaster*. *A. Rev. Genet.* **17**, 345–397.

BAKER, B. S. & WOLFNER, M. F. (1988). A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. *Genes Dev.* **2**, 477–489.

BINGHAM, P. M., CHOU, T.-B., MIMS, I. & ZACHAR, Z. (1988). On/off regulation of gene expression at the level of splicing. *TIG* **4**, 134–138.

BIRNSTIEL, M. L. & SCHAUFELE, F. J. (1988). Structure and function of minor snRNPs. In *Structure and Function of Major and Minor Small Nuclear ribonucleoprotein Particles* (ed. M. L. Birnstiel), pp. 155–182. Berlin: Springer Verlag.

BLACK, D. L., CHABOT, B. & STEITZ, J. A. (1985). U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* **42**, 737–750.

BOGGS, R. T., GREGOR, P., IDRISS, S., BELOTE, J. M. & MCKEOWN, M. (1987). Regulation of sexual differentiation in *Drosophila melanogaster* via alternative splicing of RNA from the transformer gene. *Cell* **50**, 739–747.

BOZZONI, I., FRAGAPANE, P., ANNESI, F., PIERANDREI-AMALDI, P., AMALDI, F. & BECCARI, E. (1984). Expression of two *Xenopus laevis* ribosomal protein genes in injected frog oocytes. A specific splicing block interferes with the L1 RNA maturation. *J. molec. Biol.* **180**, 987–1005.

BUTLER, B., PIROTTA, V., IRMINGER-FINGER, I. & NOETHIGER, R. (1986). The sex-determining gene *tra* of *Drosophila*: molecular cloning and transformation studies. *EMBO J.* **5**, 3607–3613.

CAFFARELLI, E., FRAGAPANE, P., GEHRING, C. & BOZZONI, I. (1987). The accumulation of mature RNA for the *Xenopus laevis* ribosomal protein L1 is controlled at the level of splicing and turnover of the precursor RNA. *EMBO J.* **6**, 3493–3498.

CHABOT, B., BLACK, D. K., LEMASTER, D. M. & STEITZ, J. (1985). The 3′ splice site in pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science* **230**, 1344–1347.

CHOU, T.-B., ZACHAR, Z. & BINGHAM, P. M. (1987). Developmental expression of a regulatory gene is programmed at the level of splicing. *EMBO J.* **6**, 4095–4104.

CLINE, T. W. (1985). Primary events in the determination of sex in *Drosophila melanogaster*. In *Origin and Evolution of Sex* (ed. H. O. Halvorson & A. Monroy), pp. 301–327. New York: Alan R. Liss.

DE ROBERTIS, E. M., LIENHARD, S. & PARISOT, R. F. (1982). Intracellular transport of microinjected 5S and small nuclear RNAs. *Nature, Lond.* **295**, 572–577.

DREYFUSS, G., PHILIPSON, L. & MATTAJ, I. W. (1988). Ribonucleoprotein particles in cellular processes. *J. Cell Biol.* **106**, 1419–1425.

ENGELS, W. R. (1983). The P family of transposable elements in Drosophila. *A. Rev. Genet.* **17**, 315–344.

ENGELS, W. R. (1984). A trans-acting product needed for P factor transposition in Drosophila. *Science* **226**, 1194–1196.

FORBES, D. J., KIRSCHNER, M. W., CAPUT, D., DAHLBERG, J. E. & LUND, E. (1984). Differential expression of multiple U1 small nuclear RNAs in oocytes and embryos of Xenopus laevis. *Cell* **38**, 681–689.

FORBES, D. J., KORNBERG, T. B. & KIRSCHNER, M. W. (1983). Small nuclear RNA transcription and ribonucleoprotein assembly in early *Xenopus* development. *J. Cell Biol.* **97**, 62–72.

FRADIN, A., JOVE, R., HEMENWAY, C., KEISER, H. D., MANLEY, J. M. & PRIVES, C. (1984). Splicing pathways of SV40 mRNAs in

Xenopus laevis oocytes differ in their requirements for snRNPs. *Cell* **37**, 927–936.

Fritz, A., Parisot, R., Newmeyer, D. & De Robertis, E. M. (1984). Small nuclear U-ribonucleoproteins in *Xenopus laevis* development. Uncoupled accumulation of the protein and RNA components. *J. molec. Biol.* **178**, 273–285.

Fu, X. Y. & Manley, J. L. (1987). Factors influencing alternative splice site utilization *in vivo*. *Mol. Cell. Biol.* **7**, 738–748.

Green, M. R., Maniatis, T. & Melton, D. A. (1983). Human beta-globin pre-mRNA synthesized *in vitro* is accurately spliced in Xenopus oocyte nuclei. *Cell* **32**, 681–694.

Hamm, J., Kazmaier, M. & Mattaj, I. W. (1987). *In vitro* assembly of U1 snRNPs. *EMBO J.* **6**, 3479–3485.

Hamm, J., van Santen, V. L., Spritz, R. A. & Mattaj, I. W. (1988). Loop I of U1 snRNA is the only essential RNA sequence for binding of specific U1 snRNP proteins. *Mol. Cell. Biol.* **8**, 4787–4791.

Laski, F. A., Rio, D. C. & Rubin, G. M. (1986). Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. *Cell* **44**, 7–19.

Leff, S. E. & Rosenfeld, M. G. (1986). Complex transcriptional units: diversity in gene expression by alternative RNA processing. *A. Rev. Biochem.* **55**, 1091–1117.

Lerner, M. R. & Steitz, J. A. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5495–5499.

Lund, E. & Dahlberg, J. E. (1987). Differential accumulation of U1 and U4 small nuclear RNAs during *Xenopus* development. *Genes Dev.* **1**, 39–46.

Lund, E. (1988). Heterogeneity of human U1 snRNAs. *Nucl. Acids Res.* **16**, 5813–5826.

Lund, E., Bostock, C. J. & Dahlberg, J. E. (1987). The transcription of Xenopus laevis embryonic U1 snRNA genes changes when oocytes mature into eggs. *Genes Dev.* **1**, 47–56.

Lund, E., Kahan, B. & Dahlberg, J. E. (1985). Differential control of U1 small nuclear RNA expression during mouse development. *Science* **229**, 1271–1274.

Lührmann, R. (1988). snRNP Proteins. In *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (ed. M. L. Birnstiel), pp. 71–99. Berlin: Springer Verlag.

Maine, E. M., Salz, H. K., Schedl, P. & Cline, T. W. (1985). Sex-lethal, a link between sex determination and sexual differentiation in Drosophila melanogaster. *Cold Spring Harbor Symp. quant. Biol.* **50**, 595–604.

Mattaj, I. W. & De Robertis, E. M. (1985). Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell* **40**, 111–118.

Mattaj, I. W., Zeller, R., Carrasco, A. E., Jamrich, M., Lienhard, S. & De Robertis, E. M. (1985). In *Oxford Survey on Eukaryotic Genes*, vol. 2 (ed. N. Maclean), pp. 121–140. Oxford University Press.

McAllister, G., Amara, S. G. & Lerner, M. R. (1988). Tissue-specific expression and cDNA cloning of small nuclear ribonucleoprotein-associated polypeptide N. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5296–5300.

Nagoshi, R. N., McKeown, M., Burtis, K. C., Belote, J. M. & Baker, B. S. (1988). The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster. *Cell* **53**, 229–236.

Nash, M. A., Kozak, S. E., Angerer, L. M., Angerer, R. C., Schatten, H., Schatten, G. & Marzluff, W. F. (1987). Sea urchin maternal and embryonic U1 RNAs are spatially segregated in early embryos. *J. Cell Biol.* **104**, 1133–1142.

Newport, J. & Kirschner, M. (1982). A major developmental transition in early Xenopus embryos. I. Characterization and timing of cellular changes at the midblastula stage. *Cell* **30**, 675–686.

Noble, J. C. S., Pan, Z. Q., Prives, C. & Manley, J. L. (1987). Splicing of SV40 early pre-mRNA to large T and small t mRNAs utilises different patterns of lariat branch sites. *Cell* **50**, 227–236.

Noble, J. C. S., Prives, C. & Manley, J. L. (1986). *In vitro* splicing of simian virus 40 early pre-mRNA. *Nucl. Acids Res.* **14**, 1219–1235.

O'Hare, K. & Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. *Cell* **34**, 25–35.

Pan, Z.-Q. & Prives, C. (1988). Assembly of functional U1 and U2 human–amphibian hybrid snRNPs in *Xenopus laevis* oocytes. *Science* **241**, 1328–1331.

Parker, R., Siliciano, P. & Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**, 220–239.

Reddy, R. & Busch, H. (1988). Small nuclear RNAs: RNA sequences, structure, and modifications. In *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (ed. M. L. Birnstiel), pp. 1–37. Berlin: Springer Verlag.

Theissen, H., Etzerodt, M., Reuter, R., Schneider, C., Lottspeich, F., Argos, P., Lührmann, R. & Philipson, L. (1986). Cloning of the human cDNA for the U1 RNA-associated 70K protein. *EMBO J.* **5**, 3209–3217.

Tooze, J. (1981). *DNA Tumor Viruses*, 2nd ed. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory.

Wieringa, B., Hofer, E. & Weissmann, C. (1984). A minimal intron length is required for splicing the large rabbit beta-globin intron. *Cell* **37**, 915–925.

Zachar, Z., Chou, T.-B. & Bingham, P. M. (1987). Evidence that a regulatory gene autoregulates splicing of its transcript. *EMBO J.* **6**, 4105–4111.

Zeller, R., Nyffenegger, T. & De Robertis, E. M. (1983). Nucleocytoplasmic distribution of snRNPs and stockpiled snRNA-binding proteins during oogenesis and early development in Xenopus laevis. *Cell* **32**, 425–434.

Zhuang, Y., Leung, H. & Weiner, A. M. (1987). The natural 5′ splice sites of SV40 large T antigen can be improved by increasing the base complementarity to U1 RNA. *Mol. cell. Biol.* **7**, 3018–3020.

(*Accepted 10 November 1988*)