The EMBO Journal vol.5 no.5 pp.1023-1030, 1986

Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3′ splice site

Alessandra Cellini, Eduard Felder and John J.Rossi

Department of Molecular Genetics, Beckman Research Institute of The City of Hope, Duarte, CA 91010, USA

Communicated by G.Tocchini-Valentini

In the yeast *Saccharomyces cerevisiae* the 5′ and 3′ splice junctions and the internal branch acceptor site (TACTAAC box) are highly conserved intron elements. Analyses of mutants have demonstrated the importance of each of these elements in the splicing process. In the present report we show by three different analytical approaches (splicing-dependent β-galactosidase expression, *in vitro* splicing assays and *in vivo* RNA analyses) that at least two of these elements (the TACTAAC and 3′ splice signals) also have to fulfill certain spacing requirements to allow efficient splicing to occur. In particular, the spacing of the 3′ splice site from the 2′−5′ branch site is a critical factor in determining the efficiency for completion of the final reactions of splicing, intron release and exon−exon joining. Whereas insertions within this region have little or no effect on the first reactions in splicing (the 5′ cleavage and 2′−5′ branch formation), they dramatically affect the efficiency of the final reactions. In contrast, a 15-base deletion between these two sites has no detectable effect on splicing efficiency. We also show that the 5′ cleavage and branch formation can take place, albeit inefficiently, in transcripts in which all of the yeast sequences downstream of the branch site have been replaced by *Escherichia coli* sequences. We conclude from these studies that, in yeast, the 5′ and 3′ splice sites are recognized independently from one another, but always in conjunction with the TACTAAC signal.

Key words: branch site/spacing/splicing/yeast

Introduction

One of the intriguing aspects of the mRNA splicing process is how the positions of three consensus blocks (the 5′ and 3′ splice junctions and the internal branch acceptor site) define an intervening sequence. An attractive model system for addressing this problem is the yeast *Saccharomyces cerevisiae* in which the consensus blocks are comprised of highly conserved sequences. *S. cerevisiae* not only follows the GT/AG rule of higher eukaryotes but, in addition, a larger number of nucleotides appear to be conserved in the intron of its pre-mRNA, including the six nucleotides GTAPyGT at the 5′ intron/exon junction (Langford *et al.*, 1984; Teem *et al.*, 1984) and the seven nucleotides TACTAAC near the 3′ junction (Langford and Gallwitz, 1983; Langford *et al.*, 1984; Pikielny *et al.*, 1983).

These two highly conserved sequences are required for the formation of the 2′−5′ linkage between the last A of the TACTAAC box and the 5′G of the intron during the splicing process (Domdey *et al.*, 1984; Newman *et al.*, 1985; Rodriguez *et al.*, 1984). The resulting branched structure, called a lariat (Grabowski *et al.*, 1984), is an obligatory intermediate before the 3′ intron cleavage and ligation of the exons can occur. The intron is then released as a lariat structure (Domdey *et al.*, 1984; Grabowski *et al.*, 1984; Padgett *et al.*, 1984; Ruskin *et al.*, 1984). Deletions or point mutations inside the 5′ intron boundary or in the TACTAAC box cause inefficient, incorrect or complete loss of splicing activity (Cellini *et al.*, 1986; Gallwitz, 1982; Jacquier *et al.*, 1985; Langford and Gallwitz, 1983; Langford *et al.*, 1984; Newman *et al.*, 1985; Parker and Guthrie, 1985; Pikielny *et al.*, 1983). However, the mere presence of these sequences does not account for efficient splicing. Other factors must play a role in determining the efficiency with which these signals are recognized and used.

For different introns in *S. cerevisiae* the spacing between the TACTAAC box and the 3′ AG varies from six to 53 nucleotides; this variation, as well as variations of the nucleotide sequences in this region, could affect the efficiency of splicing and play a role in the regulation of gene expression. To test this possibility we varied the distance of the yeast actin gene TACTAAC box with respect to the 3′ splice junction. Our results demonstrate that the spatial relationship between the TACTAAC and 3′ splice site is inconsequential to the efficiency of the initial events in splicing, the 5′ exon−intron cleavage and lariat formation. In contrast, the final reactions of splicing, the 3′ intron cleavage and exon−exon joining, are critically dependent upon this spacing.

We demonstrate that in yeast splicing, unlike metazoans, the branch site is rigorously fixed to the TATAAC sequence, and is located within this signal regardless of its distance from the 3′ splice site. In further contrast to metazoan splicing, which requires an intact, intron-encoded polypyrimidine tract adjacent to the 3′ splice site for formation of a splicing complex and subsequent branch formation (Frendewey and Keller, 1985; Reed and Maniatis, 1985; Ruskin and Green, 1985), we observe that a yeast precursor harbouring *Escherichia coli* sequences downstream of the TACTAAC, and lacking an AG, is still a substrate for lariat formation, albeit quite inefficiently. These later results suggest that sequences at the 3′ end of the intron, although not required for branch formation, may facilitate branch site localization.

Results

We have previously reported the construction of a translational fusion plasmid which allows a simple quantitative assay for the efficiency of the splicing process (Larson *et al.*, 1983). A similar plasmid pAHB-i2 (Figure 1) contains a tri-hybrid fusion between the *S. cerevisiae ACT* and *HIS4* genes and a defective *E. coli lacZ* gene. Expression from this vector is dependent upon the actin 5′-flanking sequence for transcriptional and translational start sites and upon proper splicing of the actin intervening sequence (IVS) for in-phase translation and subsequent β-galactoside activity. While precise and efficient splicing of the actin IVS gives a full level of β-galactosidase activity, failure or inaccuracy of splicing prevents functional expression of the β-galactosidase by altering the reading frame, and inefficient splicing causes a decreased level of β-galactosidase (Cellini *et al.*, 1986).

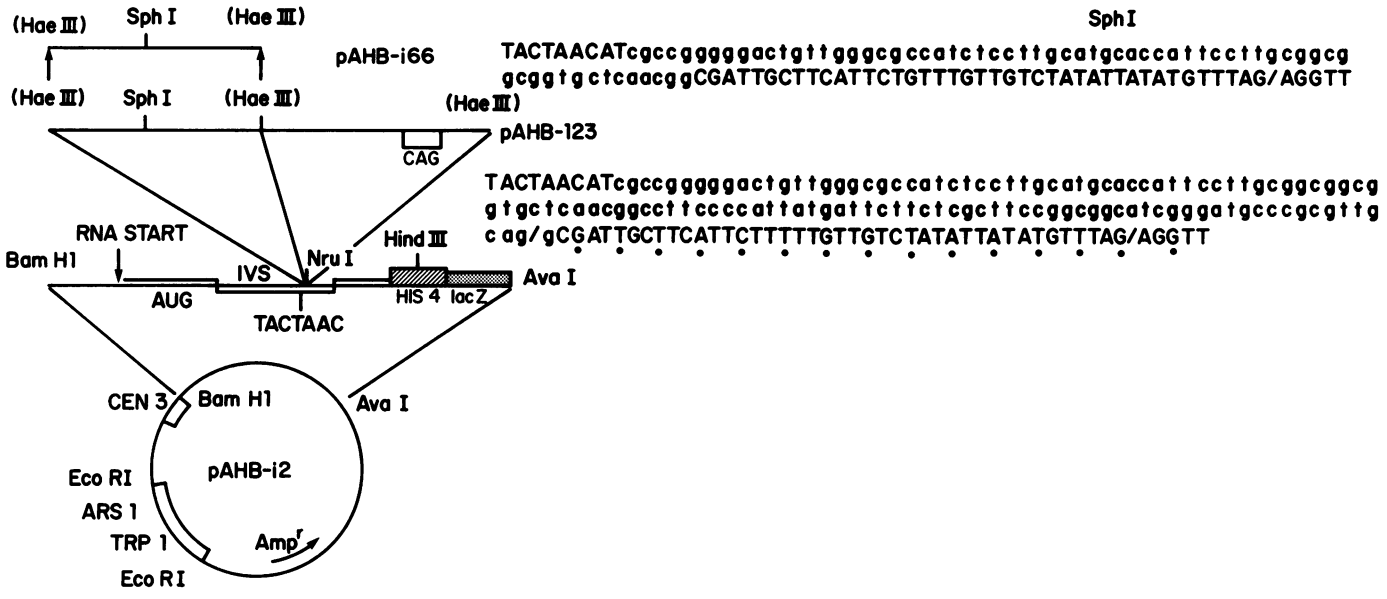

Fig. 1. Sequences of the inserted DNAs between the *ACT* intron TACTAAC and 3′ AG. The DNA sequences from the *ACT* TACTAAC through the 3′ AG are presented. The upper case letters represent *ACT* sequences, whereas the lower case letters depict the pBR322 *Hae*III fragments inserted within the *ACT* intron. The relevant restriction endonuclease sites are indicated. In the flared out diagram above the circle, the *ACT* sequences are represented by the open rectangles. The dots below the pAHB-i123 insertion correspond to the first base of the codons generated if a splice were made from the normal *ACT* 5′ splice site to the nucleotide following the 'cag' introduced in the pBR322 inserts. These codons are translationally in-phase with the downstream *ACT–HIS4–lacZ* sequences.

To construct our insertion mutants, we first created a *Nru*I site between the TACTAAC and AG by filling in a *Cla*I site and rejoining the flush ends (Figure 1). This filled in *Cla*I site created a 2-bp insertion relative to the wild-type ACT intron, as well as creating a unique *Nru*I site. The other insertion mutants were created by inserting *Hae*III fragments from pBR322 into the *Nru*I site of this plasmid, designated pAHB-i2 (see Figure 1). The 64-bp insertion, in the orientation depicted, does not contain any AG. Thus, the closest AG to the TACTAAC sequence is the normal *ACT* AG, which is now spaced 106 nucleotides away from this sequence. The 64- plus 57-bp double insertion does contain an AG at one of the boundaries of the 57-bp fragment. This AG is spaced 122 bases downstream of the TACTAAC in the construction depicted in Figure 1. The potential use of this AG as a 3′ splice signal would result in a fusion message with in-phase translation of the *ACT–HIS–lacZ* sequences.

Each of the above vector constructs was transformed into yeast and assayed for β-galactosidase activity (Table I). The insertion mutant pAHB-i66 resulted in a 70% reduction in activity versus the parental construct, while the pAHB-i123 insertion mutant resulted in a drastic reduction of >95%.

To understand better which steps in the splicing pathway were affected by each of the insertions, primer extension analyses of the fusion transcripts were carried out as described in the legend to Figure 2. Three different primers were used for the 5′ extension analyses. One of these is a 21-base oligonucleotide (A), which bridges the *ACT–HIS* fusion in the 3′ exon and therefore primes from both spliced and unspliced transcripts. The second oligonucleotide used is 24 bases long (B), and is complementary to sequences at the 3′ intron boundary. The third primer utilized is a 23-base oligonucleotide (C) that is specific for *ACT* intron sequences 46 bases downstream of the 5′ splice junction. The yeast host utilized for these studies, strain FC8-24D, contains no *ACT* intron and harbours a deletion of the *HIS4* gene (Parker and Guthrie, 1985). In this strain only products originating from the fusion messages are detected by the above-described primers.

Table I. β-Galactosidase measurements in yeast strain NNY transformed with the fusion vectors

Construct	Percent control
pAHB-i2	100.0[a]
pAHB-i66	27.0[b]
pAHB-i123	2.2[b]
pAHI-Δ15	100.0[c]

[a]The two-base insertion (relative to the wild-type *ACT* intron) in the intron of this vector had no effect on expression of fusion β-galactosidase activity in four independent measurements.
[b]The percentages are relevant to values obtained from pAHB-i2 transformants in yeast strain NNY. These values represent the averages obtained from four independent experiments with duplicate measurements.
[c]The value presented is relative to both pAHB-i2 and the wild-type *ACT* intron fusion vector described in the legend to Figure 5. The percentage represents two independent experiments with duplicate measurements.

As controls in the primer extension analyses, FC8-24D was also transformed with *ACT–HIS* fusion vectors in which the TACTAAC sequences have been deleted (Cellini *et al.*, 1986) resulting in accumulation of precursor RNAs. As shown in Figure 2, primer (A), little or no precursor RNA was detected in either the parental pAHB-i2, pAHB-i66 or pAHB-i123 constructs, while significant amounts of precursor accumulated in the TACTAAC deletion mutant (Δ8). Results with this same primer gave some spliced message (band 7, Figure 2) in the pAHB-i66 extensions, but there is a very strong stop at the TACTAAC sequence (band 4, Figure 2, i-66). A similarly strong TACTAAC stop is seen in transcripts derived from the pAHB-i123 construct (band 4, Figure 2, i-123), but in this case there is no extension of the size predicted for a spliced message. That the strong stop observed in both the mutant extensions is due to a branch point was confirmed with the IVS-specific primers (B) and (C). Primer-extended products from both of these oligonucleotides gave very

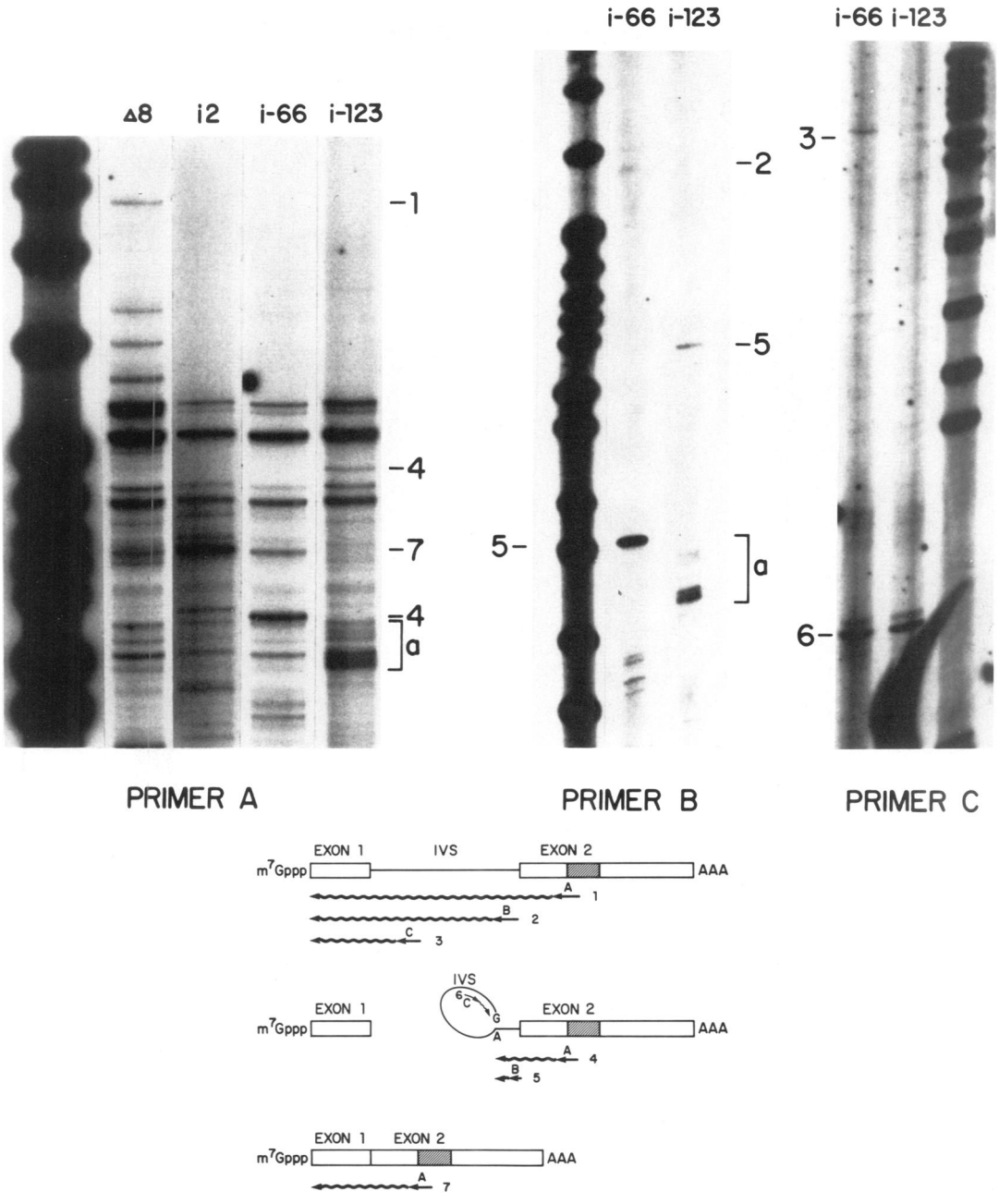

Fig. 2. Primer extension analyses of *in vivo* transcripts from the *ACT* intron insertion mutants. For the DNA sequences of the insertion mutants see Figure 1. The vectors used to transform the yeast host FC8-24D were the high copy number vectors described in Materials and methods. These vectors contain the *ACT* gene transcriptional termination and poly(A) addition signal fused downstream of the *ACT–HIS4* fusion. Total RNAs from the transformed yeast host were used as templates for AMV reverse transcriptase-mediated priming from ^{32}P-labelled synthetic oligonucleotides. The primer-extended products were electrophoresed in an acrylamide–urea gel and the radioactive products autoradiographed. The positions within the transcripts (precursor, 2/3 splicing intermediate, intron and mRNA) to which the synthetic oligonucleotides are complementary are indicated. Each panel depicts an autoradiogram of the primer-extended products obtained with the indicated oligonucleotide primer. The numbers alongside the autoradiograms correspond to the reverse transcripts depicted in the diagram below. The symbol a] brackets a series of reverse transcriptase stops seen only with transcripts from the i-123 insertion mutant, and only with primers A and B. These stops may be due to secondary structure created by the inserted sequences in this mutant. The Δ8-primed reverse transcripts are depicted in the upper panel as a reference for full length, unspliced transcripts. We routinely see several premature stops below the full-length precursor. The other umarked bands in these experiments are probably the consequence of non-specific priming events since similar bands have been observed in transcripts made from RNAs prepared from untransformed cells (data not presented).

prominent stops at the lariat attachment site (Figure 2, band 5, primer B; band 6, primer C).

A confirmation of our interpretations for the *in vivo* results was obtained from an *in vitro* splicing assay (Newman *et al.*, 1985). Transcripts from the i-66 and i-123 mutants were incubated in the splicing extracts as described in the legend to Figure 3. Both the i-66 and i-123 transcripts accumulate 2/3 lariat molecules, with a small amount of intron appearing in the i-66 reactions, but no detectable intron in the i-123 reactions. No message is apparent in the i-66 lanes. This is probably due to the poor efficiency of splicing along with the background of RNA breakdown products in this region of the gel. None the less, the appearance of intron signifies that the splice has been made.

The above experiments showed that the spacing between the

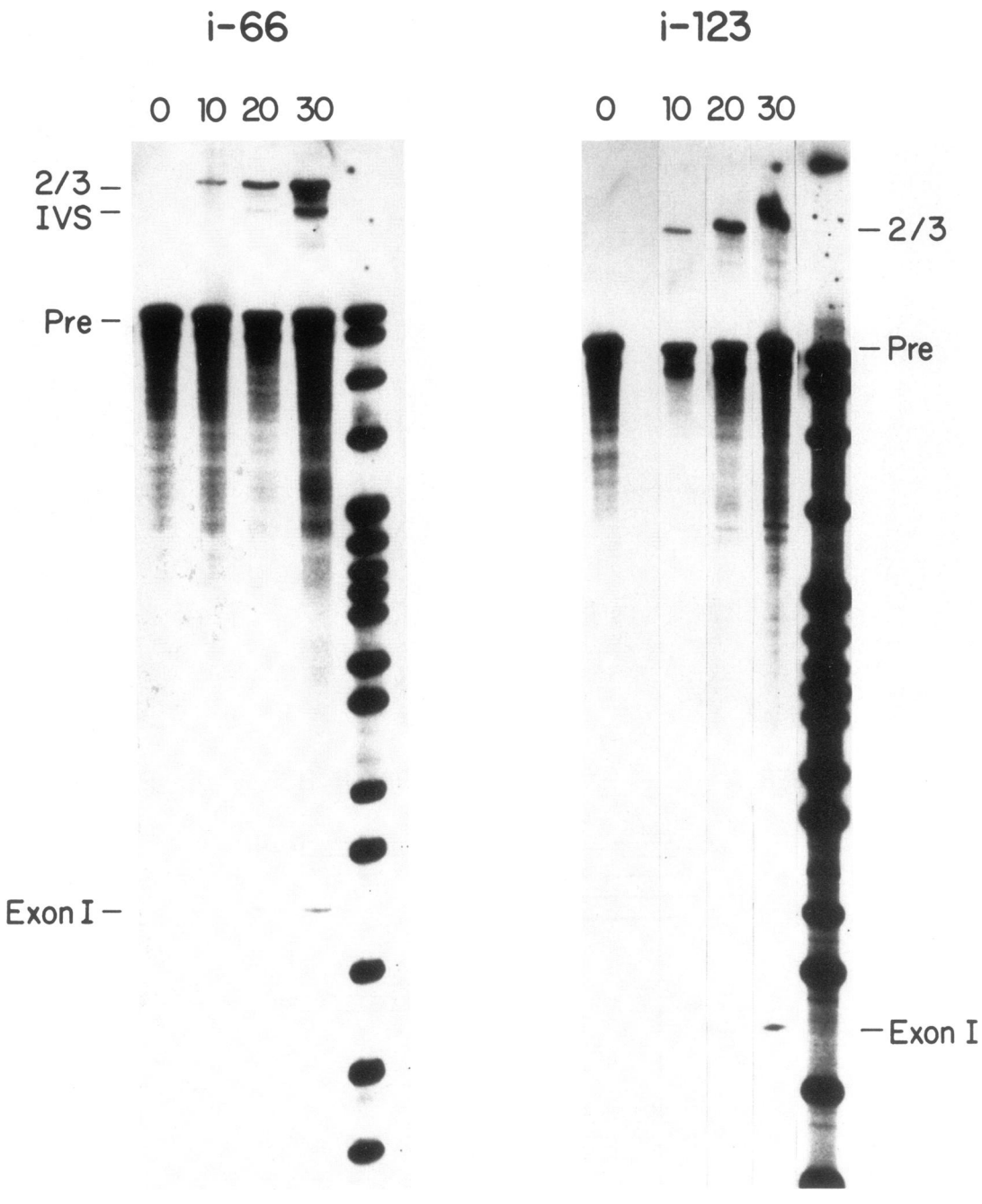

Fig. 3. *In vitro* splicing products derived from the i-66 and i-123 insertion mutant transcripts. Internally ^{32}P-labelled transcripts were incubated with the splicing extracts for the times (in min) indicated. The reactions were terminated and the RNAs extracted twice with a 1:1 mixture of phenol–chloroform. The RNAs were electrophoresed in an 8% polyacrylamide-8 M urea gel (40:1.3 acrylamide to bisacrylamide) and the gels autoradiographed. Under these gel conditions, branch-containing molecules have an anomalous electrophoretic mobility and migrate more slowly than the larger sized molecules (Newman *et al.*, 1985). The rightmost lane in each panel is a *Hpa*II digest of pBR322 used as size marker. 'Pre' stands for uncleaved precursor, 'IVS' for excised IVS and '2/3' for a lariat-containing molecule consisting of the intron and 3′ exon. The position of the excised exon I is also indicated. We have not been able to identify spliced message from the i-66 transcripts due to the heavy background of RNA degradation products in the region of the gel where the message should migrate.

TACTAAC and AG has no marked effect on the efficiency of the initial reactions in the splicing pathway. Each of the transcripts still maintain the *ACT* 3′ intron splice junction sequences, and these may still be necessary for intron recognition. To examine whether or not this was the case, we generated truncated transcripts deficient in the 3′ splice site. This was accomplished by preparing a run-off transcript of *Sph*I-cleaved SP6-i66, which leaves 38 nucleotides downstream of the TACTAAC, but is completely deficient in *ACT* 3′ intron sequences (Figures 1 and 4). This truncated substrate, when incubated in the *in vitro* splicing extracts, gave rise to lariat molecules, but at a greatly reduced efficiency compared with fully intact i-2-derived transcripts (Figure 4).

In addition to evaluating the effects of insertions on messenger splicing, we have also constructed a 15-nucleotide deletion between the TACTAAC and AG (Figure 5). The 15-base deletion was created by synthetic DNA-mediated site-specific mutagenesis, and the sequences deleted are depicted in Figure 5. The

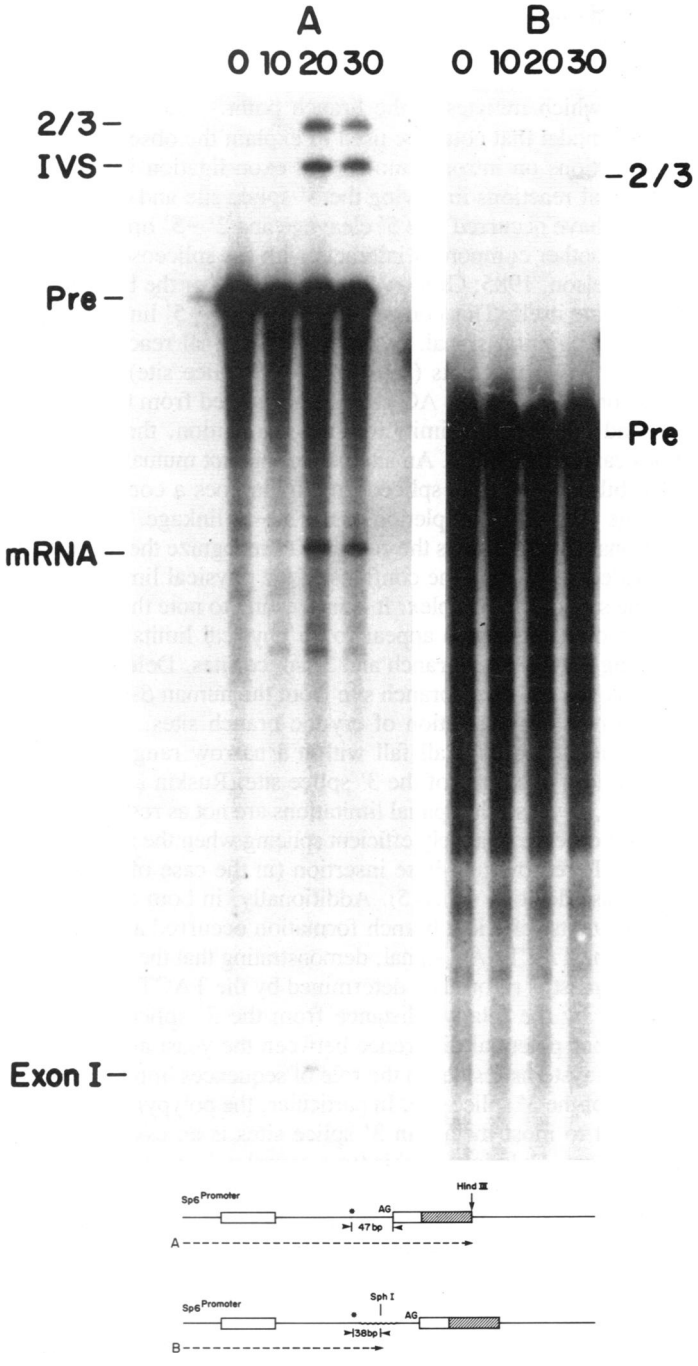

Fig. 4. *In vitro* splicing products from intact and truncated *ACT* transcripts. The conditions used for splicing, electrophoresis and the symbols are the same as presented in Figure 3. The diagrams below the autoradiograms illustrate the *in vitro* produced transcripts which were incubated in the splicing extracts. The open rectangles depict *ACT* exon sequences and the shaded rectangles *HIS4*-coding sequences. The intron is represented as a solid line, the insertion with a wavy line. The * marks the position of the TACTAAC box. Transcript A was produced from an SP65 vector harbouring exon and intron sequences derived from pAHB-i2. Transcript B was prepared from an SP65 vector harbouring exon and intron sequences from pAHB-i66 cleaved with *Sph*I prior to the *in vitro* transcription reactions. To be sure that no full-length transcript was present in the splicing reactions involving the truncated precursor, the following measures were taken. The vector was digested exhaustively with *Sph*I, and the completeness of digestion verified by analytical gel electrophoresis prior to its use as an *in vitro* transcription template. The radioactively labelled transcript prepared from this cleaved vector was purified after sizing it alongside full-length transcripts using denaturing polyacrylamide–urea gel electrophoresis.

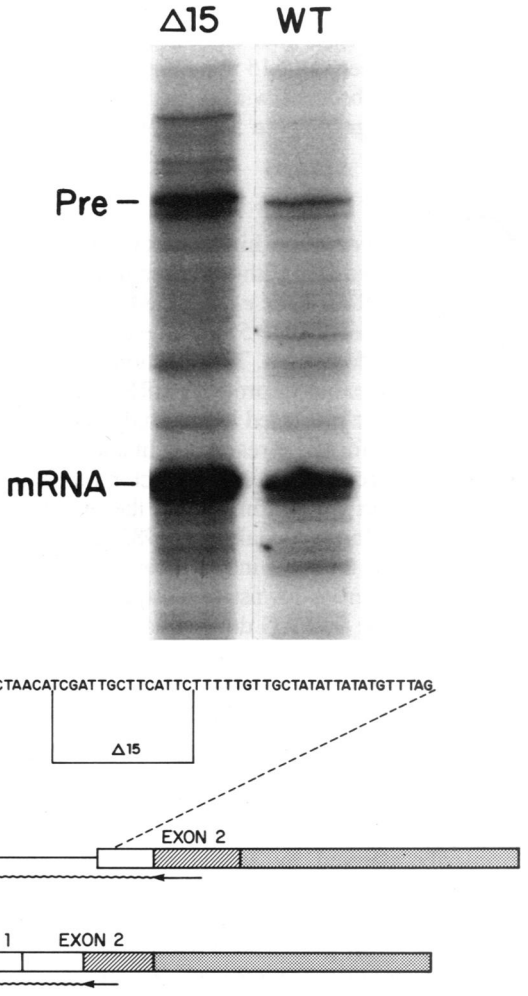

Fig. 5. Primer extension products obtained from yeast strain FC8-24D transformed with either a wild-type *ACT* intron containing *ACT–HIS4–lacZ* fusion vector or the Δ15 mutant. The *ACT–HIS4* fusion in these constructs is different from that contained in the pAHB vectors. The vectors depicted here have a different, and shorter segment of *HIS4*, but contain 200 more bases of *ACT* exon 2 sequences. In addition, the transcripts depicted in Figure 2 derive from vectors in which the *ACT* transcriptional termination signals and poly(A) addition sites have been joined to the *ACT–HIS4* coding sequences. The vectors described in this figure are *ACT–HIS4–lacZ* fusion vectors, and do not contain yeast transcriptional termination and poly(A) addition signals. These differences are probably responsible for the detectable amount of unspliced precursor observed in the primer extensions depicted above. The primer used in these experiments is complementary to *HIS4* sequences just downstream of the *ACT–HIS4* fusion point. The open rectangles depict *ACT* exon sequences, the rectangles with slashes the *HIS4*-coding sequences and the rectangles filled in with stipples depict the *lacZ* sequences. The *ACT* intron sequence is represented by a solid line. The nucleotides which have been deleted in the Δ15 mutant are indicated. The primer extension reactions were carried out on total RNAs as described in Materials and methods and the legend to Figure 2. Following autoradiography, the ratio of precursor to message was determined by densitometry using an integrated chart recorder. The ratios of precursor:message were 0.39 for Δ15 and 0.36 for the wild-type (WT) *ACT* intron construct.

mutant intron was incorporated into an *ACT–HIS–lacZ* fusion vector, and assayed for β-galactosidase activity (Table I). The deletion mutation did not result in any reduction of β-galactosidase versus the control construct which contained a wild-type *ACT* intron.

Certain intron mutations do not affect the steady-state level of mRNA, but do result in an accumulation of unspliced precursor RNA (Pikielny and Rosbash, 1985). To examine whether or not

the Δ15 mutant might create such an effect, *in vivo* transcripts were analyzed by primer extension analyses (Figure 5). No differences between the Δ15 mutant and wild-type *ACT* intron constructs were observed with respect to the ratio of precursor to spliced message (Figure 5 legend).

Discussion

In yeast, as well as in higher eukaryotes, pre-mRNA splicing occurs via a two-step pathway (Domdey *et al.*, 1984; Grabowski *et al.*, 1984; Rodriguez *et al.*, 1984; Ruskin *et al.*, 1984). The first step consists of the 5′ intron−exon cleavage and concommitant branch formation. This reaction is followed by a concerted joining of the exons to produce mRNA, and intron release. The results we have presented in this study elaborate on the substrate requirements for these reactions. In addition to the stringent sequence requirements, a critical factor for recognition of the 3′ splice junction is its promixity to the branch site. Prior to our studies, Langford and Gallwitz (1983) deleted the *ACT* 3′ splice site and fused the *ACT* intron 5′ and TACTAAC sequences to a series of different downstream sequences. Their results demonstrated that the first AG downstream of the TACTAAC is used for 3′ splice site recognition provided it is a critical minimum distance from the TACTAAC. Their data also suggested that increased distances from the TACTAAC to the AG diminished the efficiency of splicing. At the time these studies were presented, it was not feasible to elucidate which step in the splicing pathway was primarily affected by the spatial alterations. Our results strongly suggest that the final exon−exon joining reaction is primarily affected by spatial alterations between these two regions.

We have clearly shown that 66- and 123-base insertions between the actin gene TACTAAC and AG splice signals have marked consequences on the splicing efficiency of the fusion transcripts. Moreover, our results suggest that the insertions have little or no effect on the first step in the splicing pathway, but alter markedly the efficiency of the exon−exon joining reaction. The 66-base insertion does not contain an AG dinucleotide, and therefore preserves the normal *ACT* 3′ splice signal. In contrast, the 123-base insertion contains a single AG which is a potential 3′ splice signal. An mRNA resulting from use of this AG would still be translationally in-phase with the normal *ACT−HIS* and *lac*Z sequences (see Figure 1). Since the actual distances from the branch site in the TACTAAC to the first AG differ by only 15 nucleotides in these two mutants, it is surprising that they differ so much in their relative splicing efficiencies (Table I, Figures 2 and 3). It is possible that the context of sequences surrounding the i-123 AG are not favourably recognized by the splicing apparatus and hence result in poor splicing efficiency. An equally likely alternative is that this distance simply exceeds some physical constraint upon the splicing complex making it no longer capable of simultaneous recognition of the branch site and 3′ AG. The residual 3% β-galactosidase activity remaining in this mutant could be due to an inefficient splice using this introduced AG signal or the normal AG signal. Since we have been unable to detect an mRNA from this mutant, we cannot distinguish between the above possibilities.

Klinz and Gallwitz (1985) have reported spatial constraints between the message cap and intron, as well as between the 5′ and TACTAAC signals. Based upon their results, they have suggested that the messenger splicing process in yeast may be initiated by a 5′ to 3′ scan of the precursor RNAs. The spatial constraints which they have observed are measured in hundreds of base pairs, whereas ours are measured in tens of base pairs (the difference between the branch site and nearest AG in i-66 versus i-123) for the final splicing reactions. It therefore seems unlikely that recognition of the 3′ splice site in yeast involves a scan which initiates at the branch point.

One model that could be used to explain the observed effects of insertions on intron removal and exon ligation is that, after the initial reactions involving the 5′ splice site and TACTAAC signals have occurred (the 5′ cleavage and 2′−5′ branch formation), another component interacts with the spliceosome (Brody and Abelson, 1985; Grabowski *et al.*, 1985) or the branched intermediate itself. This component uses the 2′−5′ linkage as part of its recognition signal. In order for the final reaction to proceed, the two reactants (5′ exon and 3′ splice site) must be in close proximity. If the AG signal is separated from the complex beyond the physical limits for this recognition, the final reactions cannot take place. An alternative, but not mutually exclusive possibility, is that the spliceosome undergoes a conformational change following completion of the 2′−5′ linkage. The conformational change allows the complex to recognize the nearest AG provided it is within the confines of the physical limits imposed by the size of the complex. It is interesting to note that, in higher eukaryotes, there also appear to be physical limitations on the spacing between the branch and 3′ splice sites. Deletions which remove the authentic branch site from the human β-globin large intron result in activation of cryptic branch sites. These cryptically activated sites all fall within a narrow range of 22−37 nucleotides upstream of the 3′ splice site (Ruskin *et al.*, 1985). Clearly, in yeast, the spatial limitations are not as restricted since we still observe relatively efficient splicing when the normal spacing is altered by a 66-base insertion (in the case of i-66) or by a 15-base deletion (in Δ15). Additionally, in both of our insertion mutants, efficient branch formation occurred at the normal site in the TACTAAC signal, demonstrating that the branch position in yeast is rigorously determined by the TACTAAC signal, and not by the relative distance from the 3′ splice signal.

Another potential difference between the yeast and metazoan splicing systems resides in the role of sequences immediately upstream of the 3′ splice site. In particular, the polypyrimidine tract adjacent to most metazoan 3′ splice sites is an essential intron component. Deletions of this tract completely abolish lariat formation, and prevent spliceosome assembly (Frendewey and Keller, 1985; Reed and Maniatis, 1985; Ruskin and Green, 1985). In contrast, Rymond and Rosbash (1985) have demonstrated relatively efficient lariat formation *in vitro* in an RP51 transcript in which ~18 nucleotides upstream of, and including, the 3′ splice site were deleted. Our results are qualitatively similar to those of Rymond and Rosbash in that the truncated precursor was still capable of undergoing branch formation (Figure 4). Two important differences between our and the above experiments should be noted. First, the precursor which we constructed contained only four of the naturally occurring yeast nucleotides downstream of the TACTAAC, the remainder being replaced by *E. coli* sequences. In essence, our transcript had no remnants of the pyrimidine-rich sequence normally present in the actin intron. The RP51 precursor cited above still contained a considerable percentage of the endogenous sequences downstream of the TACTAAC. Secondly, the efficiency with which branch formation occurred in our precursor was very poor compared with the non-truncated controls (Figures 3 and 4). Again, this is in contrast to the efficient branch formation in the truncated RP51 precursor (Rymond and Rosbash, 1985). The differences in branch formation efficiency cannot be attributed to the the length

of sequences downstream of the TACTAAC signal, since these lengths were approximately the same in both cases. Therefore, the sequence composition must be important.

One plausible explanation for the observed differences in lariat formation efficiencies between the *ACT* and RP51 truncated precursors is that the endogenous, remaining sequences downstream of the TACTAAC signal in the RP51 precursor somehow still contribute to branch site localization, possibly via interaction with a yeast analogue of the metazoan U5 snRNP (Chabot *et al.*, 1985). Whereas the polypyrimidine tract sequences in metazoans appear to be essential for splicing complex formation and choice of branch site, in yeast the analogous sequences may only serve to facilitate primary interactions of the splicing apparatus with the precursor.

The above cited differences in splice signal requirements between yeast and the metazoans perhaps reflect a divergence in the pathways of evolution of the messenger splicing mechanism. In higher eukaryotes, most messenger encoding genes contain multiple intervening sequences. In some instances alternative splice sites are used in the same transcript. Recognition of the 3′ splice signal before branch site localization can be viewed as a mechanism for avoiding splice site skipping or, contrarily, for ensuring that an alternative splice site is utilized. In yeast, with the exception of the *MATa* gene (Miller, 1984), all of the known interrupted genes contain a single intron and alternative splicing has not been demonstrated. The highly conserved 5′ splice and TACTAAC signals provide the primary intron recognition elements. Choice of a 3′ splice site appears to follow branch formation at the TACTAAC signal (Langford and Gallwitz, 1983; Rymond and Rosbash, 1985; this study). These observations along with the fact that TACTAAC mutants in yeast fail to undergo the 5′ cleavage and lariat formation reactions (Cellini *et al.*, 1986; Newman *et al.*, 1985; Parker and Guthrie, 1985), suggest somewhat different mechanisms for recognition and removal of introns in yeast as compared with higher eukaryotes.

Materials and methods

Enzymes and biochemicals

Restriction endonucleases, T4 DNA ligase, polynucleotide kinase, DNA polymerase I (Klenow fragment) were purchased from Bethesda Research Laboratories. ^{32}P-Labelled nucleotides were purchased from New England Nuclear. AMV reverse transcriptase was obtained from Life Sciences, Inc. The *in vitro* transcription vector SP65 and SP6 RNA polymerase were purchased from Promega Biotech. Synthetic deoxyribooligonucleotides A (Figure 2) and the mutator oligonucleotide were synthesized on an automated DNA synthesizer from Systec, Inc.

Oligonucleotide B (Figure 2) was a gift from J.Abelson, and oligonucleotide C (Figure 2) was a gift from C.Guthrie.

Strains

E. coli strain MC1061 (Casadaban and Cohen, 1980) was used for routine manipulations and was provided by M.Casadaban. *E. coli* strain JM101 (Messing, 1983) was used for M13 phage transfections. The yeast strains used in this study were NNY [*MATα,trp1,gal2,gal10,ura3-52* (J.Wallis)], and FC8-24D [*MATa, trp1,leu2,his4-401,ura3-52,HOL,ACT ΔIVS* (Parker and Guthrie, 1985)]. Yeast strain EJ101 [a protease-deficient strain (Newman *et al.*, 1985)] was used for preparation of the splicing extracts.

Vectors

The insertion mutant pAHB-i66 was constructed by inserting the pBR322 *Hae*III 64-bp fragment into the *Nru*I site of pAHB-i2. The mutant pAHB-i123 was constructed from a double insertion of the pBR322 *Hae*III 64-bp and 57-bp fragments into this same *Nru*I site. The sequences of the inserts and flanking regions were verified by the chemical DNA sequencing procedure of Maxam and Gilbert (1977). The construction of pAHB-i2 and pAHB-Δ8 are described elsewhere (Cellini *et al.*, 1986). Each of these vectors contains a tri-hybrid fusion of yeast *ACT−HIS4* and *E. coli lacZ* sequences. The vectors also contain the *TRP1* gene and *ARS1* element of yeast (Tschumper and Carbon, 1980), as well as a yeast centromere 3-containing fragment (Clarke and Carbon, 1980).

For RNA analyses, high copy number vectors containing the *ACT−HIS4* fusion, *ACT* transcriptional termination region and the *TRP1−ARS1* sequences were constructed as follows. A 1-kb *Bam*HI−*Hin*dIII fragment harbouring the *ACT* 5′ regulatory region, intron and *ACT−HIS4* fused coding sequences was inserted into the *Bam*HI and *Hin*dIII sites of pBR327. A 1.6-kb *Hin*dIII−*Eco*RI fragment from the *ACT*-coding and 3′-untranslated sequences (Gallwitz and Sures, 1980; Ng and Abelson, 1980) was then inserted into the *Hin*dIII−*Eco*RI sites of the above plasmid. Finally, a 1.4-kb *Eco*RI fragment containing the *TRP1−ARS* sequences was inserted into the unique *Eco*RI site of the second construct.

In vitro transcripts were prepared from vectors harbouring fragments of the *ACT−HIS4* fusion which contains the intron and intron variants. These were constructed by inserting a 600-bp *Alu*I fragment into the *Sma*I site of SP65 (Promega Biotech). Prior to transcription, the vectors were linearized at the vector *Hin*dIII site immediately downstream of the insert. Internally labelled splicing substrates were prepared using conditions recommended by the supplier. Other manipulations were according to Newman *et al.* (1985).

Bacterial and yeast transformations

E. coli cells were transformed by the method of Kushner (1978). *S. cerevisiae* strains were transformed by the lithium acetate procedure (Ito *et al.*, 1983).

β-Galactosidase assays

Strains carrying the *lac*Z fusion plasmids were grown to mid to late log phase (OD 600 of 0.8−1.5) in minimal selective media. For the assays, 0.1 or 0.5 ml of cells were pelleted in a microfuge, resuspended in 1.0 ml of Z buffer (Miller, 1972) and assayed as described by Larson *et al.* (1983). For each sample, duplicate measurements were carried out and the average value calculated.

In vivo and in vitro RNA analyses

Preparation of *in vivo* synthesized RNAs and primer extension analyses were performed according to Domdey *et al.* (1984). *In vitro* splicing extracts were prepared and assays carried out according to the procedures outlined by Newman *et al.* (1985).

In vitro site-directed mutagenesis

The Δ15 mutant was constructed using a 30-base synthetic oligonucleotide to prime repair replication of an *ACT−HIS4*-containing M13mp9 (Messing, 1983) single-stranded template. The general procedures used have been outlined by Zoller and Smith (1983). Phage plaques were screened using the nitrocellulose filter adsorption procedure of Benton and Davis (1977), and screened with the ^{32}P-labelled mutator oligonucleotide according to the methodology described by Zoller and Smith (1983). Once the mutant was identified and characterized by dideoxy sequence analysis (Sanger *et al.*, 1977), a fragment harbouring the mutation was subcloned into an *ACT−HIS4−lacZ* fusion vector for expression in yeast.

Acknowledgements

The authors thank John Abelson and his laboratory colleagues for providing us with the protocol for preparation of *in vitro* splicing extracts prior to publication, and Garry Larson for construction of the yeast centromere 3-containing vectors. We also wish to acknowledge the expert technical assistance of Deanna Lundin. This research was supported by P.H.S. grant GM30134 and Beckman Institute Research Grant to J.J.R. E.F. was supported by a post-doctoral fellowship from the Swiss National Foundation.

References

Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180−182.
Brody,E. and Abelson,J. (1985) *Science*, **228**, 963−967.
Casadaban,M.J. and Cohen,S. (1980) *J. Bacteriol.*, **143**, 971−980.
Cellini,A., Parker,R., McMahon,J., Guthrie,C. and Rossi,J. (1986) *Mol. Cell Biol.*, in press.
Chabot,B., Black,D.L., LeMaster,D.M. and Steitz,J.A. (1985) *Science*, **230**, 1344−1349.
Clarke,L. and Carbon,J. (1980) *Nature*, **287**, 504−509.
Domdey,H., Apostol,B., Lin,R.J., Newman,A., Brody,E. and Abelson,J. (1984) *Cell*, **39**, 611−621.
Frendewey,D. and Keller,W. (1985) *Cell*, **42**, 355−367.
Gallwitz,D. (1982) *Proc. Natl. Acad. Sci. USA*, **70**, 3493−3497.
Gallwitz,D. and Sures,I. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2546−2550.
Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **37**, 415−427.
Grabowski,P.J., Seiler,S.R. and Sharp,P.A. (1985) *Cell*, **42**, 345−353.
Ito,H., Fukuda,Y., Murata,K. and Kimura,A. (1983) *J. Bacteriol.*, **153**, 163−168.
Jacquier,A., Rodriguez,J.R. and Rosbash,M. (1985) *Cell*, **43**, 423−430.
Klinz,F.-J. and Gallwitz,D. (1985) *Nucleic Acids Res.*, **13**, 3791−3804.
Kushner,S.R. (1978) In Boyer,H.W. and Nicosia,S. (eds), *Proceedings of the International Symposium on Genetic Engineering*. Elsevier/North Holland Biomedical Press, pp. 17−23.
Langford,C.J. and Gallwitz,D. (1983) *Cell*, **33**, 519−527.

Langford,C.J., Klinz,F.J., Donath,C. and Gallwitz,D. (1984) *Cell*, **36**, 645–653.
Larson,G.P., Itakura,K., Itoh,H. and Rossi,J.J. (1983) *Gene*, **22**, 31–39.
Maxam,A. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Miller,A.M. (1984) *EMBO J.*, **3**, 1061–1065.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, NY.
Newman,A., Lin,R.J., Cheng,S.C. and Abelson,J.A. (1985) *Cell*, **42**, 335–344.
Ng,R. and Abelson,J. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3912–3916.
Padgett,R.A., Konarska,M., Grabowski,P.J., Hardy,S.F. and Sharp,P.A. (1984) *Science*, **225**, 898–903.
Parker,R. and Guthrie,C. (1985) *Cell*, **41**, 107–118.
Pikielny,C.W. and Rosbash,M. (1985) *Cell*, **41**, 119–126.
Pikielny,C.W., Teem,J.L. and Rosbash,M. (1983) *Cell*, **34**, 395–402.
Reed,R. and Maniatis,T. (1985) *Cell*, **41**, 95–105.
Rodriguez,J.R., Pikielny,C.W. and Rosbash,M. (1984) *Cell*, **39**, 603–610.
Ruskin,B. and Green,M.R. (1985) *Nature*, **317**, 732–734.
Ruskin,B., Krainer,A.R., Maniatis,T. and Green,M.R. (1984) *Cell*, **38**, 317–331.
Ruskin,B., Greene,J.M. and Green,M.R. (1985) *Cell*, **41**, 833–844.
Rymond,B.C. and Rosbash,M. (1985) *Nature*, **317**, 735–737.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Teem,J.L., Abovich,N., Kaufer,N.F., Schwindinger,W.F., Warner,J.R., Levy, A., Woolford,J., Leer,R.J., Van Raamsdonk-Duin,M.M., Mager,W.H., Planta,R.J., Schultz,L., Friesen,J.A. Fried,H. and Rosbash,M. (1984) *Nucleic Acids Res.*, **12**, 8295–8312.
Tschumper,G. and Carbon,J. (1980) *Gene*, **10**, 157–166.
Zoller,M.J. and Smith,M. (1983) *Methods Enzymol.*, **100**, 468–500.

Received on 18 December 1985; revised on 20 February 1986