In vitro processing of the human growth hormone primary transcript

Klaus Hartmuth and Andrea Barta*

Institut für Biochemie, Universität Wien, Währingerstraße 17, A-1090 Wien, Austria

Received June 29, 1987; Revised and Accepted August 13, 1987

ABSTRACT

To study the sequence of events during processing of primary RNA transcripts and to gain more insight into the mechanism of splice site selection, the *in vitro* processing of a 2.5 kb human growth hormone (hGH) pre-mRNA containing four introns and an alternative 3' splice site for intron B was analysed. In order to process the hGH pre-mRNA the preparation of the HeLa cell nuclear extract (1) had to be modified, indicating differences in factor requirement for processing this pre-mRNA. After an unusual long lag phase of one hour splicing intermediates begin to accumulate. Intron A and D are removed with correct ligation of exons 1/2 and 4/5. Most splice sites are used - albeit with variable efficiencies - except the splice sites surrounding exon 3 and the 3' alternative splice site within exon 3; as a consequence "exon skipping" events take place. Using a pre-mRNA containing only intron B neither the 5' nor the 3' splice site is cleaved, indicating that the 3' splice site of intron B is not recognized. The results show that splice sites can differ considerably in their requirement for splicing factors.

INTRODUCTION

The excision of intervening sequences from nuclear pre-mRNAs has to be precise if functional mature RNAs are to be generated. Evidence is accumulating that splicing may contribute to the diversity of gene expression by the use of alternative splice sites to produce more than one mRNA from one gene (2, 3, 4, 5). Instances have been described where this alternative or differential splicing can be highly tissue specific (6, 7, 8, 9, 10). This is indicative of an as yet poorly defined mode of regulation of gene expression.

Recent advances in the biochemistry of splicing depended strongly on the development of *in vitro* splicing systems and it was shown that the process could be divided into two stages

(for review see ref. 11). The first stage is the cleavage of the intron at its 5' end and subsequent joining of its 5'-phosphate through a 2'-5' phosphodiester bond to a nucleotide near the 3' end of the intron. The second stage involves cleavage at the 3' intron/exon junction with concurrent ligation of the two exons, releasing the intron in a lariat form. The structural requirements for the formation of lariats in nuclear pre-mRNAs involves sequences at the 5' and 3' splice site and the polypyrimidine stretch preceeding the 3' splice site (12, 13, 14, 15). Although an invariant UACUAAC sequence within the intron was identified as being crucial for branch point selection in yeast (16, 17) no similar sequence has been found in introns of higher eucaryotes (18, 19). A number of proteins (20, 21, 22, 23) and small ribonucleoprotein particles (snRNPs) have been shown to participate in splicing of nuclear pre-mRNAs (24, 25, 26; for review see ref. 11). Recognition of the 3' part of the intron seems to be an early step in the assembly of an active splice complex (26, 12, 27, 28, 29). However, regulation of selection of different splice sites is as yet poorly understood. This is of particular interest for pre-mRNAs containing large introns, multiple introns or alternative 3' or 5' splice sites for individual introns.

To study these phenomena, we used the pre-mRNA of the human growth hormone (hGH) for in vitro processing in HeLa cell nuclear extracts. hGH is required for postnatal growth and has a variety of effects on the carbohydrate, lipid, mineral and nitrogen metabolism (30). The hGH gene contains four introns and is specifically expressed in the anterior pituitary. In vivo, 10% of the primary transcript uses an alternative 3' splice site of intron B, giving rise to a protein 15 amino acids shorter with a slightly different biological activity (4, 31).

Processing of the hGH pre-mRNA could only be detected by modifying the published procedure for the preparation of the HeLa cell nuclear extract (1). The analyses of the hGH splicing products indicate that the recognition of naturally occurring splice sites can vary considerably, suggesting a difference in the factor requirement of certain splice sites.

MATERIALS AND METHODS

Materials

Restriction enzymes, SP6 RNA polymerase, T-4 DNA ligase and HindIII linkers were from Biolabs (FRG). GpppG, NTPs, dNTPs, polynucleotide kinase and mung bean nuclease were from Pharmacia. Nuclease S1 (24 IU/ml) was from Calbiochem. DNA polymerase (Klenow fragment) was from Staehlin (Switzerland). Reverse transcriptase was provided by J.Beard (Life Sciences, St.Petersburg, FL). α -³²P dNTPs (3000Ci/mmol), α -³²P UTP (440-880Ci/mmol), γ -³²P ATP (5000Ci/mmol) and human placental ribonuclease inhibitor were from Amersham. Formamide was from Sigma and was recrystallized twice at 0°C. All other chemicals were of the highest purity available.

Construction of plasmids

All cloning steps were essentially carried out as described (32). pAAH5-hGH was constructed by first converting the unique BamHI site at the 5' end of the hGH gene (genomic EcoRI fragment in pBR322; 33, 34) into a HindIII site by treatment with mung bean nuclease and subsequent addition of HindIII linkers (octamer). The resulting 2.1 kb fragment with the small EcoRI/HindIII fragment of pBR322 at its 3' end was then inserted into the unique HindIII site of pAAH5 (35). pAAH5-chGH was constructed by inserting the HindIII fragment containing the hGH cDNA (36) into pAAH5. pH10 and pH12 were constructed by inserting the HindIII fragment of pAAH5-hGH into the SP65 and SP64 vector respectively. pB3 was constructed by inserting the internal PvuII fragment of the hGH gene (Figure 1A) into the SmaI site of SP65. pC10 was constructed by inserting the hGH cDNA (36) into the BamHI site of the pSP65 vector in an inverted orientation relative to the SP6 promoter.

Preparation of RNA substrates

In vivo synthesis of RNA in recombinant yeasts: pAAH5-hGH was transformed (37) into a competent leu⁻ yeast strain (strain WS2-1C). Total RNA was isolated from 1.6 litre log phase cultures (OD₅₆₈ approximately 0.6) by resuspending the pelleted cells in 3ml Tris/HCl pH 6.7, 2%SDS, 2 mM EDTA addition of acid washed glass beads and 3ml equilibrated phenol and vigorous vortexing. The aqueous phase was extracted two

more times with phenol/chloroform before polyA⁺ RNA was selected by two cycles of binding to oligo(dT) cellulose (32) using LiCl instead of NaCl. RNA was stored frozen in water. For in vitro synthesis with SP6 RNA polymerase, DNA templates were linearized by excessive digestion with the appropriate restriction enzymes and transcribed according to published procedures (38, 39). Capping was accomplished by lowering the GTP concentration to 50 μ M and addition of 875 μ M GpppG dinucleotide (40). Radiolabelled RNA was stored at -20°C in 50% ethanol.

Preparation of extracts and splicing reactions

HeLa cell nuclear and cytoplasmic extracts were prepared from HeLa Ohio cells (1) with the following modifications. Volumes refer to 2×10^9 cells starting material; all operations were at 4°C. The pelleted cells were resuspended in 30 ml isotonic buffer (10mM HEPES/KOH pH 7.9, 140mM KCl, 1.5 mM MgCl₂, 0.5 mM DTT, 0.5 mM PMSF, 1 mM EDTA) and disrupted by Dounce homogenization (type A pestle; 150 - 200 strokes; approximately 25 min). Pelleted nuclei were first resuspended in 1.0 ml hypotonic buffer (20 mM HEPES/KOH pH 7.9, 2.5 mM MgCl₂, 1mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 25% glycerol) and stirred for 30 min. Then an equal volume of high salt buffer (hypotonic buffer with 0.88 M NH₄Cl) was slowly added and stirring was continued for an additional 40 min. Subsequent steps were as described (1). For the cytoplasmic S100 extract the cytoplasmic fraction was precipitated with 0.51 g/ml (NH₄)₂SO₄ after the high speed centrifugation and resuspended in 4 ml dialysis buffer. All other steps were exactly as described (1). The protein concentration of the nuclear extract was 7-10 mg/ml.

Splicing incubations all contained 12 mM HEPES/KOH pH 7.9, 60 mM KCl, 0.12 mM DTT, 0.12 mM EDTA, 0.3 mM PMSF, 2.4 mM MgCl₂, 1.3 mM ATP and 20 mM creatine phosphate and were performed at 30°C. Due to the large variability in splicing efficiency among different preparations the following 3 reactions conditions were used to test each new extract: (i) 60% extract (ii) 60% extract, 3% PEG-6000, (iii) 40% extract, 20% S100, 3% polyethylene glycol 6000, further variations proved ineffective. With the yeast polyA⁺ RNA precursor,

containing 1-2% hGH specific pre-mRNA, 1 μ g RNA was incubated in a total volume of 25 μ l. With the SP6 RNAs 50000 to 500000 cpm (Cherenkov) were incubated in 15 μ l and preparative reactions were performed in 200 μ l. Recovery of RNA was as described (41).

Analysis of RNA

DNA fragments for indirect analysis of RNA were labelled either with polynucleotide kinase or DNA polymerase (Klenow fragment) to specific activities ranging between 0.5-2 μ Ci/pmole end and were purified by preparative polyacrylamide gel electrophoresis. Single stranded DNA fragments were prepared as described (32). Nuclease S1 mapping and primer extension analysis of RNA was as described (41), except that 1IU/ml nuclease S1 was used throughout and digestion was at 30°C for 1 hour, unless otherwise indicated in the figures. Chemical sequencing of primer extension products was as described previously (41). ³²P-labelled RNAs were analysed on 8% polyacrylamide/8M urea gels. RNase mapping was performed as in (38) and protected fragments were displayed on a 6% polyacrylamide/8M urea gel.

RESULTS

Generation of pre-mRNAs and preparation of extracts

To generate a full length, capped and polyadenylated pre-mRNA of hGH, advantage was taken of the fact that mammalian processing signals are not recognized in yeast cells (16). A 2.1 kb fragment containing the complete transcription unit of the hGH gene (Figure 1A) was cloned into the yeast multicopy vector pAAH-5 under transcriptional control of the yeast ADH regulatory regions (Figure 1B, and Materials and Methods). PolyA⁺ RNA isolated from yeasts transformed with this plasmid contains 1-2% hGH specific transcript. S1 nuclease analysis of this polyA⁺ RNA revealed that neither the polyadenylation signal (unpublished observation) nor the splicing signals of the hGH pre-mRNA are recognized in yeast (see control lanes in Figure 2). This RNA was mainly used in assaying the processing events of the complete hGH pre-mRNA. Radiolabelled pre-mRNAs were synthesized in vitro from SP64 or SP65 derived plasmids where the hGH gene or a fragment of it had been inserted (Figure 1C).

Figure 1: Schematic diagram of the hGH gene and the hGH pre-mRNAs used in this study. **A.** Physical map of the hGH gene cloned into pBR322 as an *EcoRI* fragment (B: *BamHI*, M: *MspI*, P: *PvuII*, F: *HinfI*, R: *EcoRI*, H: *HindIII*, T: *TaqI*). Below this, all restriction fragments used for nuclease S1 mapping and primer extension analysis are indicated, the star (*) showing the relevant ³²P-label. The structure of the 2.5 kb pre-mRNA as isolated from recombinant yeasts (Figure 1B) is shown on the bottom. The lengths of the exons in nucleotides are: 1: 94, 2: 161, 3: 120, 4: 165 and 5: 303. The lengths of the introns in nucleotides are: A: 259, B: 206, B': 251, C: 93 and D: 253 (60). **B.** The yeast multicopy vector pAAH-5 is shown with the hGH gene below inserted into the *HindIII* site. Notation within the hGH gene is as in Figure 1A. With respect to the vector the following notation is valid: B: *BamHI*, H: *HindIII*, E: *EcoRI*, ADH-PRO and ADH-TER: yeast alcoholdehydrogenase I promoter and terminator, respectively (35). **C.** Structure of SP6 generated hGH pre-mRNAs. H12HfI and H12RI are RNAs transcribed from pH12 linearized with *HinfI* and *EcoRI*, respectively. B3HfI is RNA transcribed from pB3 linearized with *HinfI*. Exon 1 of H12HfI and H12RI is 82 nts; exon 2 of B3HfI is 52 nts; exon 3 of H12HfI and B3HfI is 76 nts; the 3' untranslated region of H12RI is 540 nts.

HeLa cell nuclear extract was prepared essentially as described (1) with some modifications (see Materials and Methods). These modifications were found to be essential as the standard protocol yielded extracts which were only active with other eukaryotic pre-mRNAs (first and second leader exon of the major late transcription unit of adenovirus 2) but not with the hGH pre-mRNA (unpublished observations).

The hGH pre-mRNA shows an unusual long lag phase before the onset of processing. Splicing intermediates become visible in the case of the hGH pre-mRNA after one hour incubation time (Figure 4) which is about twice the time as for other eucaryotic pre-mRNAs in this HeLa cell nuclear extract (41, 42). This phenomenon, in combination with the differences in the extract preparation and the sensitivity of the reaction to extract parameters suggest that probably higher concentrations of certain factors might be required for splicing the hGH pre-mRNA.

Frequency of usage of the various splice sites of the complete hGH pre-mRNA

The hGH pre-mRNA, as isolated from recombinant yeast, contains four introns and an alternative 3' splice site of intron B (Figure 1A and 1B). Using the DNA fragments shown in Figure 1A for S1 mapping on processed hGH pre-mRNA (Figure 2(A) and 2(B), lanes S), cleavage at each splice site was assayed. Control hybridizations were done routinely to unprocessed pre-mRNA (lanes C) and yeast tRNA (lanes T).

S1 mapping with the 3' end labelled 772 nt HindIII/HinfI probe gives rise to a 70 nt and a 487 nt protected fragment resulting from cleavage at the 5' splice sites of intron A and B, respectively (Figure 2(A)a, lane S). Lane C' shows control hybridization to 50% of input RNA. The fact that 5' end cleavage of intron B can be detected with this probe indicates that removal of intron A is not a prerequisite for cleavage at the 5' splice site of intron B. The same probe labelled at the 5' end gives rise to a 285 and a 702 nt fragment indicating cleavage at the 5' splice sites of intron B and A, respectively (Figure 2(A)b, lane S). The 446 nt fragment results from cleavage at the 3' splice site of intron A. A 76 nt fragment

expected from processing at the 3' end of intron B (Figure 2(A)b, lane R, control hybridization with mature hGH RNA) could not be detected. Also, a 36 nt fragment that should be visible if the 3' alternative splice site of intron B were used could not be detected using shorter runs and longer autoradiographic exposures (not shown). Nevertheless the 5' end of intron B is frequently used as can be seen in Figure 2(A)c, lane S where a 3' end labelled 440 nt MspI/HinfI probe was used. The protected fragment in this case is the 155 nt band.

Processing around introns C and D was investigated with a single stranded 3' end labelled 446 nt HinfI/HinfI probe Figure 2(A)d, lane S). Although a variety of different hybridization conditions were tried (data not shown) the background remained high. One striking feature is that processing of the 5' end of

Figure 2: Nuclease S1 mapping of 5' and 3' splice sites of introns A to D of the hGH pre-mRNA processed *in vitro*. **Lanes M:** 32 P-labelled *Msp*I cleaved pBR322 DNA; **lanes S:** hybridization to processed RNA; **lanes C:** hybridization to unprocessed RNA; **lanes T:** hybridization to yeast tRNA; **Lane C':** hybridization to unprocessed RNA (50% of input RNA); **lane R:** hybridization to mature hGH RNA; **lane F:** fragment used in the analysis. Diagrams of the nuclease S1 protected fragments are depicted below the autoradiograms along with the relevant details of the pre-mRNA. (A). Analysis of processing of introns A and B. (B). Analysis of processing of introns C and D. C. Schematic outline of the frequency of utilization of all the splice sites mapped and the cleavage at the polyadenylation site (not shown). The size of the arrows represents a semiquantitative estimate of splice site usage, the largest being 60-70% and the smallest 1-5%, when processed RNA is compared to unprocessed RNA.

intron C could not be detected. The 46 nt fragment expected is readily seen when mature hGH RNA was used for hybridization (Figure 2(A)d, lane R). Cleavage at the 3' end of intron C is evidenced by the fragment migrating at 139 nt. The third protected fragment at 304 nt indicates that splicing at the 5' end of intron D has occurred, albeit with low efficiency (see below).

To delineate further these events, a 5' end labelled 443 nt double stranded HinfI/HinfI fragment was employed for hybridization (Figure 2(A)e, lane S). The 307 nt protected fragment results from cleavage at the 3' end of intron C. The 142 nt protected fragment shows again correct cleavage at the 5' end of intron D.

Finally, processed RNA was assayed with a 5' end labelled double stranded 1386 nt HindIII/TaqI fragment (Figure 2(A)f). A 61 nt protected fragment, resulting from correct cleavage at the 3' end of intron D is readily seen in Figure 2(A)f, lane S (compare also with lane R, hybridization to mature hGH RNA). The 476 nt fragment again corresponds to cleavage at the 3' end of intron C and the 901 nt fragment to cleavage at the 5' end of intron B. Longer exposures reveal a faint band at 314 nt, corresponding to cleavage at the 5' end of intron D (not shown). Again, no fragment indicating cleavage around exon 3 could be detected in this assay.

From a combined analysis of Figure 2(A) and 2(B) a map with the approximate frequencies of usage of the various splice sites could be established (Figure 2(C)). The size of the arrows represents a semiquantitative estimate of the utilisation frequency of the splice sites. The largest arrows represent 50-70% and the smallest 1-5% processing efficiency comparing processed RNA with unprocessed RNA remaining after the reaction. Figure 2(C) also indicate efficient 3' end processing and polyadenylation of this pre-mRNA in the HeLa cell extract (Hartmuth and Barta, manuscript in preparation). Whilst processing events around exons 1,2,4,5 are variably efficient, neither the splice sites around exon 3 nor the alternative splice site within exon 3 are recognized.

Figure 3: Correct excision of Intron A and Intron D. Lanes are as in Figure 2. **A.** RNase mapping of processed ^{32}P -H12RI RNA. Exon 1 of the pre-mRNA is homologous to that of the anti-mRNA for only 36 nt (34, 36). **B.** Combined primer extension and nuclease S1 mapping analysis of intron D processing. A 83 bp $3'$ - ^{32}P TagI/TagI primer was used to prepare cDNAs from unprocessed (lane C) or processed (lane S) RNA. These cDNAs were then hybridized to mature hGH RNA and digested nuclease S1. Thick bars in the diagram represent the lengths of the S1 protected fragments.

Intron A and D are correctly removed

The S1 analyses (Figure 2(C)) demonstrated that both splice sites of intron A, and D were cleaved, implying that these introns were processed correctly. To show the correct

excision of intron A, processed ^{32}P -H12RI RNA was analysed by the RNase mapping procedure (38) with an unlabelled anti-mRNA probe transcribed from pC10 linearized with HindIII. Ligation of exon 1 to exon 2 results in a protected RNA fragment migrating at 197 nt (Figure 3A). Overexposure of the gel revealed another RNA fragment migrating at the position expected for correct ligation of exon 4 to exon 5 (not shown).

The exact removal of intron D is shown in another experiment: a 3' end labelled 83 bp TaqI/TaqI primer was used to prepare cDNAs from processed yeast pre-mRNA. The cDNA pool was then used as a hybridization probe for mature hGH RNA (synthesized in vivo from recombinant yeast, see Materials and Methods). Hybrids were digested with S1 nuclease and analysed on a denaturing polyacrylamide gel (Figure 3B). The 142 nt fragment (lanes C and S) corresponds to protected primer elongation up to the 5' end of exon 5, indicating unprocessed RNA with respect to intron D. The protected fragment at 307 nt (lane S) corresponds to correctly ligated exons 4/5. Further ligations upstream of exon 4 could not be detected. The fact that the 307 nt band is very faint indicates that exon 4/5 ligation is not very frequent, which is consistent with the low frequency of usage of the 5' splice site of intron D (Figure 2(C)) as determined by nuclease S1 mapping (Figure 2(C)).

The normal and the alternative 3' splice sites of intron B are not recognized in vitro

It was of particular interest to determine if and to what extent the alternative splice site within exon 3 was utilized in this in vitro assay. From the S1 mapping data discussed above it was clear that neither splice site around exon 3 is recognized nor is there any indication of usage of the alternative splice site within exon 3 (Figure 2(C)). As these data were derived from processing of the complete 2.5 kb hGH pre-mRNA, it was possible that these splice sites would be recognized using shorter pre-mRNAs. Thus the radiolabelled pre-mRNAs H12HfI and B3HfI (Figure 1C) were synthesized and incubated for various times in a HeLa cell nuclear extract. Splicing products were directly analysed on a 8% polyacrylamide/8M urea gel (Figure 4). H12HfI pre-mRNA contains

Figure 4: The normal and alternative 3' splice sites of intron B are not recognized in *in vitro* processing. SP6 generated B3HfI (A) and H12RI (B) pre-mRNAs were processed for various times and analysed on a 8% polyacrylamide/8 M urea gel. The precursor, intermediates and products produced by processing the H12RI pre-mRNA are indicated on the right, whereas the B3HfI pre-mRNA is indicated on the left. Lane M: 32 P labelled MspI cleaved pBR322 DNA.

intron A and intron B and 76 nucleotides of exon 3. Splicing intermediates and products were identified by debranching reactions and RNase A mapping following hybridisation to cmRNA (not shown). As can be seen in Figure 4B, only intron A is removed. This interpretation is consistent with the appearance of the 82 nt exon 1 and a 522 nt processed RNA, comprised of exon 1, exon 2, intron B and exon 3. The large exon lariat intermediate hardly migrates into the gel. Note the long lag

phase of one hour in this processing experiment, which is characteristic for splicing of the hGH pre-mRNA. B3HfI pre-mRNA containing only intron B is not processed at all (Figure 4A). The fact that this pre-mRNA is rapidly degraded suggests that it is not protected by assembly with splicing factors (see Discussion). On both the B3HfI and H12HfI pre-mRNAs the 5' splice site of intron B is not utilized. Nevertheless efficient cleavage at this site was demonstrated using the complete hGH pre-mRNA which possesses several additional 3' splice sites downstream of exon 3 (Figure 2(C)). The inability of this in vitro splicing system to process intron B thus appears to result from a failure to recognize the 3' splice site of intron B. In addition, the alternative 3' splice site within intron B is not recognized.

Exon skipping events in the in vitro processing of the complete hGH pre-mRNA

One exon skipping event leading to the joining of exon 1 and exon 5 could be verified by using a 5' end labelled double stranded 81 nt TaqI/TaqI primer (Figure 1A) for reverse transcription on processed H12RI RNA (Figure 1C). The cDNA corresponding in length to the exon 1/exon 5 product was isolated from a 8% polyacrylamide/8 M urea gel and chemically sequenced. The sequencing data (Figure 5) demonstrate the exact joining of exon 1 to exon 5 (34).

The summary of the S1 mapping data depicted in Figure 2(C) shows that the frequency of usage of the various splice sites is suggestive of several other exon skipping events. Interestingly, the 5' splice site of intron B is cleaved with a relatively high frequency although the 3' splice site of this intron is not cleaved (Figure 2(A)b) and not involved in lariat formation (Figure 4A and 4B). Since it has been demonstrated that cleavage at the 5' splice site requires the presence of a 3' splice site (12, 14), lariats including intron B can only be formed with the 3' end of intron C or D. That this might be the case is strengthened by the fact that the 5' splice site of intron C is not cleaved at all so that exon 4 can either be joined to exon 2 or exon 1. For intron D we observe much more cleavage at the 3' splice site than at the 5' splice site,

Figure 5: Exon 1 and exon 5 are spliced via "exon skipping". The SP6 generated H12RI pre-mRNA (Figure 1C) was incubated in a preparative reaction for 6 hours. A 81 bp 5'-³²P *TaqI*/*TaqI* primer homologous to exon 5 sequences was used to prime reverse transcription. A 224 nt cDNA product was isolated and sequenced (41). **Lane M:** ³²P labelled *MspI* cleaved pBR322 DNA; **lane F:** isolated primer extension product; **G, G/A, C/T and C:** respective sequencing reactions. Next to the autoradiogram the sequence of the cDNA and the deduced sequence of the RNA is shown. The filled arrowhead points to the junction of exon 1 and exon 5.

indicating that besides being joined to exon 1 (see above), exon 5 might also be joined to exon 2.

DISCUSSION

The first puzzling observation in the *in vitro* processing of the hGH pre-mRNA was that nuclear extracts prepared according to standard procedures were able to process other eukarotic pre-mRNAs (the major late transcription unit of

adenovirus 2, unpublished observation) but not the hGH pre-mRNA. By changing the lysis conditions of the HeLa cells from hypotonic to isotonic, nuclear extracts competent for splicing the hGH pre-mRNA were obtained. But even employing this modification for the extract preparation, considerable variations in the splicing efficiencies using the hGH pre-mRNA were noted. The observation that the isotonic lysis of the HeLa cells as well as having to add back S100 extract, if necessary, stimulates processing of hGH pre-mRNA, suggests that a factor either easily escapes from the nuclei during the isolation procedure or that a factor resides differentially in the cytoplasmic or nuclear compartment, possibly depending on the exact growth conditions of the HeLa cells. The difference in the *in vitro* splicing conditions for the hGH pre-mRNA could be explained by assuming that a common splicing factor has a very low affinity for this pre-mRNA. However, inspection of the primary sequence of the 5' and 3' splice sites of the hGH introns (Table 1), reveals no significant differences from known consensus sequences.

Splicing of the hGH pre-mRNA is initiated not before one hour incubation time which is about twice the time needed for other eukaryotic pre-mRNAs in these extracts (41, 42). As the lag phase has been correlated with the time required to assemble a sufficient amount of splicing complex to allow detection of the splicing products (12), we assume that the hGH pre-mRNA requires a longer incubation time to assemble into a

TABLE 1. Sequences bordering the 5' and 3' splice sites of hGH pre-mRNA introns A to D.

	5' exon:	intron	:3' exon
IVS A	CAG:GUAAGC	. . . CCCUCUGGUUUCUCCCCAG:GC	
IVS B	UUU:GUAAGC	. . . UGGGCGGUCCUUCUCCUAG:GA	
IVS B ^a	UUU:GUAAGC	. . . GAAGUAUUCAUUCCUGCAG:AA	
IVS C	UCC:GUGAGU	. . . GCCCGUCCUCCCCUGCAG:AA	
IVS D	GGG:GUGGGG	. . . CUCUCCUUCACUUUGCAG:AG	
animal consensus	^A CAG:GU ^A AGU	. . . Y (11)	N ^C AG:GG

^a B' denotes the intron if the alternative 3' splice site of intron B is used.

functional spliceosome. This also might be due to a lower affinity of this pre-mRNA to one of the splicing factors. The processing efficiency of the hGH pre-mRNA is however comparable to that of other pre-mRNAs tested.

As our primary interest was the in vitro splicing pattern of a long pre-mRNA containing multiple introns and an alternative splice site, we initially analysed all potential cleavage sites by nuclease S1 mapping. The data clearly show that for the complete hGH pre-mRNA most splice sites are cleaved - albeit with different efficiencies - except the 3' splice site of intron B and the 5' splice site of intron C. Since the correct removal of intron A and intron D could be demonstrated, the data show that the central portion of the hGH pre-mRNA is not recognized in vitro. That this is not due to the length of this pre-mRNA was shown by the experiment in which a pre-mRNA containing only intron B was incubated in the HeLa cell nuclear extract (Figure 4A). Although the 5' splice site of intron B is cleaved using the complete pre-mRNA, neither the 5' nor the 3' splice site of intron B is used in this experiment. This shows that the inability to process intron B is to be found in structural features around the 3' splice site of intron B. Interestingly, this pre-mRNA is degraded at a much faster rate than functional pre-mRNAs, suggesting that the pre-mRNA is not protected by the assembly with splicing factors. This was further confirmed by employing native gel electrophoreses to analyse complex formation; no splicing complexes could be detected with the B3HfI pre-mRNA (unpublished observations). Several studies suggest that an early event in pre-mRNA splicing is the association of a factor, possibly U5, with the 3' end of the intron (12, 26, 27, 28, 29) before assembly of the splicing complex (for review see ref. 11). Therefore, it seems likely that this early step cannot take place with intron B. At present there is no explanation for these results as no obvious peculiarities in the sequence of the 3' end of intron B can be detected (see Table 1).

As the hGH-gene is expressed only in highly specialized cells of the anterior pituitary, one might argue that a special

factor from these cells is needed for splicing, thus controlling gene expression. But this is unlikely, considering the many in vivo experiments, in which the hGH gene was transfected into different cell lines (43) or was used as a reporter gene for a variety of promoters (44, 45, 46) and hence spliced in these cells.

It is interesting that both splice sites which are not recognized in vitro border exon 3. It was noted earlier (47) that exon 3 might have a different evolutionary origin than the other exons of the hGH gene. Since the protein sequences within exon 2, 4, and 5 show internal homology, it has been proposed that these portions of the gene arose by gene duplications of a smaller precursor gene segment (47). In another separate evolutionary event the primordial gene of the hGH acquired exon 3. We only can speculate, as to whether the observed in vitro splicing abnormality of this exon is in any way connected to the different evolutionary origin proposed for exon 3.

The experiments also demonstrate that the alternative 3' splice site of intron B, located within exon 3, is not recognized in the HeLa cell extract. There are ample indications that alternative splice sites are in fact precisely regulated, most probably involving special trans acting factors (48, 7, 9, 49). When tested in an in vitro splicing system neither the alternative splice site of the SV40 (50) nor of the influenza virus (49) were recognized. Recently, Van Santen and Spritz (51) have shown the usage of the alternative 5' splice site of the SV40 early pre-mRNA producing the small t antigen mRNA in an in vitro splicing system, however the efficiency was very low. These data indicate that special conditions might be necessary to mimic an alternative splicing pattern. This explanation may also apply to the alternative splice site of intron B of the hGH pre-mRNA.

On the other hand the biological importance of the variant hGH which is produced by use of the alternative splice site of intron B is not yet established, nor is anything known about the regulation concerning this alternative splice site. Therefore the fact that in vivo 10% uses the alternative splice

site, might be just a statistical splicing error due to the poor performance of the normal splice site of intron B.

Our data also demonstrate the exact ligation of exon 1 and exon 5 and there is additional evidence for other "exon skipping" events. "Exon skipping" events in in vitro splicing extracts have been demonstrated mainly when either exons were shortend (52) or when long hair pin loops involving exon sequences were constructed (53). Using a RNA folding program, based on energy minimization (54), no such extreme hair pins could be detected. Therefore, we assume at present that the exon skipping events are in vitro artefacts, due to the inability of the splicing system to recognize the splice sites surrounding exon 3 and have no function in vivo.

The experiments presented in this paper show that considerable variations in splice site recognition within one pre-mRNA can be observed. In particular, a natural alternative splice site in exon 3 of the hGH pre-mRNA is not recognized under standard conditions. The establishment of a nuclear extract recognizing all catagories of splice sites and the knowledge of the structural features which cause the differential requirement for factors is a prerequisite if these questions are to be answered.

ACKNOWLEDGEMENTS

We would like to thank Drs. T. Skern and Z. Rattler for invaluable discussions and enlightening comments. We also thank D. Thompson and L. Frasel for excellent technical assistance. Plasmid pAAH5 was a kind gift of G. Ammerer; the yeast strain WS2-1C was a gift of H. Ruis. This work was supported by a grant from the Österreichischer Fonds zur Förderung der wissenschaftlichen Forschung to A.B.

*To whom correspondence should be addressed

REFERENCES

1. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475-1489.
2. Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., and Evans, R.M. (1982). *Nature* **298**, 240-244.
3. Behlke, M.A. and Loh, D.Y. (1986) *Nature* **322**, 379-382.

4. DeNoto, F.M., Moore, D.D. and Goodman, H.M. (1981) Nucl. Acids Res. 9, 3719-3730.
5. Tamkun, J.W., Schwarzbauer, J.E. and Hynes, R.O. (1984) Proc. Natl. Acad. Sci. USA 81, 5140-5144.
6. Breitbart, R.E., Nguyen, H.T., Medford, R.M., Destree, A.T., Mahdavi, V. and Nadal-Ginard, B. (1985) Cell 41, 67-82.
7. Laski, F.A., Rio, D.C. and Rubin, G.M. (1986) Cell 44, 7-19.
8. Medford, R.M., Nguyen, H.T., Destree, A.T., Summers, E. and Nadal-Ginard, B. (1984) Cell 38, 409-421.
9. Nawa, H., Kotani, H. and Nakanishi, S. (1984) Nature 312, 729-734.
10. Rozek C.E. and Davidson, N. (1986) Proc. Natl. Acad. Sci. USA 83, 2128-2132.
11. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Ann. Rev. Biochem. 55, 119-150.
12. Frendewey, D. and Keller, W. (1985) Cell 42, 355-367.
13. Mount S.M. (1982) Nucleic Acids Res. 10, 459-472.
14. Reed, R. and Maniatis, T. (1985) Cell 41, 95-105.
15. Ruskin, B. and Green, M.R. (1985) Nature 317, 732-734.
16. Langford, C.J. and Gallwitz, D. (1983) Cell 33, 519-527.
17. Pikielny, C.W., Teem, J.L. and Rosbash, M. (1983) Cell 34, 395-403.
18. Padgett, R.A., Konarska, M.M., Aebi, M., Horning, H., Weissmann, C. and Sharp, P.A. (1985) Proc. Natl. Acad. Sci. USA 82, 8349-8353.
19. Ruskin, B., Greene, J.M. and Green, M.R. (1985) Cell 41, 833-844.
20. Choi, Y.D., Grabowski, P.J., Sharp, P.A. and Dreyfuss, G. (1986) Science 231, 1534-1539.
21. Krainer, A.R. and Maniatis, T. (1985) Cell 42, 725-736.
22. Krämer, A. and Keller, W. (1985) EMBO J. 4, 3571-3581.
23. Mayrand, S.H., Pedersen, N. and Pedersen, T. (1986) Proc. Natl. Acad. Sci. USA 83, 3718-3722.
24. Berget, S.M. and Robberson, B.L. (1986) Cell 46, 691-696.
25. Black D.L. and Steitz, J.A. (1986) Cell 46, 697-704.
26. Chabot, B., Black, D.L., LeMaster, D.M. and Steitz, J.A. (1985) Science 230, 1344-1349.
27. Gerke, V. and Steitz, J.A. (1986) Cell 47, 973-984.
28. Ruskin, B. and Green, M.R. (1985) Cell 43, 131-142.
29. Tazi, J., Alibert, C., Tamsamani, J., Reveillaud, I., Cathala, G., Brunel, C. and Jeanteur, P. (1986) Cell 47, 755-766.
30. Martin, J.B. (1973) N. Engl. J. Med. 288, 1384-1393.
31. Miller, W.L. and Eberhardt, N.L. (1983) Endocrine Reviews, 4, 97-130.
32. Maniatis T., Fritsch, E. and Sambrook, J. (1982) Molecular cloning: A Laboratory Manual (Cold Spring Harbour Laboratory, Cold Spring Harbour, NY).
33. Hirt, H., Kimelman, J., Birnbaum, M., Chen, E., Seeburg, P.H., Eberhardt, N.L. and Barta, A. (1987) DNA 6, 59-70.
34. Seeburg P.H. (1982) DNA 1, 239-249.
35. Ammerer, G. (1983) Methods Enzymol. 101, 192-201.

-
36. Martial, J.A., Hallewell, R.A., Baxter, J.D. and Goodman, H.M. (1979) *Science* **205**, 602-607.
 37. Beggs, J.D. (1978) *Nature* **275**, 104-109.
 38. Melton, D.A., Kriek, P.A., Rebaqliati, M.R., Maniatis, T., Zinn, K. and Green M.R. (1984) *Nucl. Acids Res.* **12**, 7035-7056.
 39. Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) *Cell* **38**, 317-331.
 40. Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984) *Cell* **38**, 731-736.
 41. Hartmuth, K. and Barta, A. (1986) *Nucl. Acids Res.* **14**, 7513-7528.
 42. Patzelt, E., Thalmann, E., Hartmuth, K., Blaas, D. and Kuechler, E. (1987) *Nucl. Acids Res.* **15**, 1387-1399.
 43. Robins, D.M., Paek, I., Seeburg, P.H. and Axel, R. (1982) *Cell* **29**, 623-631.
 44. Morello, D., Moore, G., Salmon, A.M., Yaniv, M. and Babinet, C. (1986) *EMBO J.* **5**, 1877-1883.
 45. Ornitz, D.M., Palmiter, R.D., Hammer, R.E., Brinster, R.L., Swift, G.H. and MacDonald, R.J. (1985) *Nature* **313**, 600-602.
 46. Palmiter, R.D., Norsted, G., Hammer, R.E. and Brinster, R.L. (1983) *Science* **222**, 809-814.
 47. Barta, A., Richards, R.I., Baxter, J.D. and Shine, J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4867-4871.
 48. Feinberg, M.B., Jarret, R.F., Aldovini, A., Gallo, R.C. and Wong-Staal, F. (1986) *Cell* **46**, 807-817.
 49. Plotch, S.J. and Krug, R.M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5444-5448.
 50. Noble, J.C.S., Prives, C. and Manley, J.L. (1986) *Nucl. Acids Res.* **14**, 1219-1235.
 51. Van Santen, V.L. and Spritz, R.A. (1986) *Nucleic Acids Res.* **14**, 9911-9926.
 52. Reed, R. and Maniatis, T. (1986) *Cell* **46**, 681-690.
 53. Solnick, D. (1985) *Cell* **43**, 667-676.
 54. Williams, A.L. and Tinoco Jr., J. (1986) *Nucleic Acids Res.* **14**, 299-315.