Accurate *in vitro* splicing of human β -globin RNA

Ryszard Kole and Sherman M. Weissman

Department of Human Genetics, Yale University School of Medicine, New Haven, CT 06510, USA

Received 21 June 1982; Revised and Accepted 19 August 1982

ABSTRACT

Human β -globin RNA transcribed from an exogenous DNA template is spliced *in vitro* by concentrated whole cell extracts from HeLa cells. Using the primer extension technique, we have shown that the small intervening sequence is spliced accurately and that the sequence of the product across the splice junction is identical to that of β -globin mRNA prepared from human reticulocytes.

The efficiency of the splicing reaction is low. The RNA transcript containing both introns and terminated upstream from the polyadenylation site is spliced most efficiently. The transcript which is terminated downstream from the polyadenylation site is not spliced at all. Thalassaemic β -globin RNA which carries an extra splice site in the small intron is also spliced, albeit with a low yield.

INTRODUCTION

Among the steps required for the maturation of most eukaryotic messenger RNAs, the reaction leading to the removal of the introns and the subsequent ligation of the remaining exons (splicing) remains the least understood (for review see 1-3). Elucidation of the molecular mechanism of this process is necessary for better understanding of the regulation of gene expression and of mRNA synthesis. One example of the importance of these studies is the recent observation that the common form of β -⁺thalassaemia is due to an aberration in splicing of human β -globin mRNA (4-8).

Analysis of the biochemical steps involved in splicing requires a suitable *in vitro* system. In two recent reports, evidence was presented that splicing of viral RNA may occur *in vitro* in the whole cell extracts from HeLa (9) and MOPC-315 cells (10). Due to the limitations of the employed assays, it is not clear if the reported *in vitro* splicing reaction is accurate. Since both groups studied the splicing of adenovirus RNA, it is also not known if cellular mRNA can be spliced in these systems. Other investigators (11) were not able to detect any splicing activity in very

similar in vitro systems. We report here that a concentrated whole cell extract from HeLa cells (11- 13) is capable of splicing the small intervening sequence in the RNA transcribed in vitro from the cloned human β -globin gene. The sequence across the splice junction is identical to that of β -globin mRNA prepared from human reticulocytes. Thus the extract is capable of accurate in vitro splicing of cellular mRNA precursors. The structure of the DNA template and/or primary transcript is an important factor in permitting splicing to occur.

MATERIALS AND METHODS

Preparation of DNA.

Plasmids C1, β Pst (14) and C6 (6) were grown and digested with appropriate restriction enzymes by standard procedures. The DNA was used as templates for in vitro transcription reaction. Plasmid C1 was digested with BamHI (template I) or EcoRI (template II). Plasmid β Pst was digested with PstI (template III). Plasmid C6 was digested with EcoRI (template IV).

Preparation of the HeLa cell extract.

Concentrated extracts from HeLa cells were prepared essentially as described with the following modifications. Cells were grown to a density of at most 5×10^5 cells/ml and then collected and lysed as described (12). After lysis of the nuclei the extract was incubated at 35°C for 1 hr (13). All subsequent steps were as described (12). The final pellet was resuspended in a minimal amount of the dialysis buffer (20 mM Hepes, pH 7.9; 100 mM KCl, 12.5 mM $MgCl_2$, 0.1 mM EDTA, 2 mM DTT and 17% glycerol), transferred to the dialysis bag and dialyzed for 6-10 hours. The final protein concentration of the extract was about 20 mg/ml. In our hands, protein concentration of the extract prepared exactly as described by Manley et al., (12) was 7-10 mg/ml.

RNA synthesis.

RNA was transcribed and spliced in vitro in a reaction mixture (400 μ l) containing: 240 μ l of concentrated extract in the dialysis buffer, ATP, CTP, GTP at 0.5mM each, UTP 0.05 mM, 10 μ Ci of α -³²P UTP, 6mM creatine phosphate, and 100-200 μ g/ml of plasmid DNA cut with the appropriate restriction enzyme. Incubation was at 30°C for 60 min. (11). The reaction was stopped by addition of 600 μ l of 7 M urea, 0.35 M NaCl, 10 mM Tris, pH8, 10 mM EDTA and 0.5% SDS. The mixture was extracted twice with phenol:chloroform:isoamyl alcohol (1:1:0.05 vol/vol), once with chloroform

and ethanol precipitated. Before being analyzed by primer extensions, all the reactions were tested on 5% acrylamide, 7M urea gel (15) for correct runoff transcripts.

RNA analysis.

RNA synthesized in vitro was analyzed by the primer extension technique (16). As primer, a 5'-³²P labeled (17), 29 nucleotide long fragment of β -globin DNA, located in the second exon between Sau 96I and Sau 3aI sites (18) was used. The primer (about 1 μ g) and the RNA were resuspended in the hybridization buffer (80% formamide, 0.4 M NaCl, 0.01M Pipes, pH 6.4), total volume 100 μ l, and hybridized overnight at 50°C. Before mixing with RNA, the primer was denatured at 100°C for 5' and quick-chilled on ice. After hybridization, the nucleic acids were precipitated with ethanol and redissolved in 200 μ l of buffer containing: 50 mM Tris, pH 8.3, 60 mM NaCl, 6 mM magnesium acetate, 10 mM DTT, 5 mM each dATP, dGTP, dCTP and TTP and 5 units of reverse transcriptase (supplied by Dr. J. Beard, Life Sciences). The reaction was at 41°C for 3hr. RNA in the hybrid was degraded with 0.2N NaOH for 30 min. at 50°C. The mixture was neutralized with HCl. The cDNA was extracted twice with phenol and ethanol precipitated. Digestion with HaeIII was for 2 hr. at 37°C, with 60 units of the enzyme (New England Biolabs) per total amount of the synthesized cDNA. The extension products were analyzed on an 8% sequencing gel (20x40x0.03 cm) (19). For sequencing cDNA was run on a gel 2 mm thick and the proper bands were excised and eluted. Eluted DNA was sequenced according to the procedure of Maxam and Gilbert (17).

Densitometry of autoradiograms was carried out on a Joyce-Loebl microdensitometer model MK-IIICS.

RESULTS

DNA templates.

We have used cloned fragments of the DNA coding for the normal or β^+ -thalassemic human β -globin gene as templates for directing in vitro synthesis of the precursors to β -globin mRNA. The human β -globin gene contains three coding segments and two intervening sequences 130 and 852 nucleotides long, located between codons 30-31 and 104-105, respectively. The total length of the gene from the cap site to the polyadenylation site is about 1600 nucleotides. The nucleotide sequence of the gene and of its flanking sequences is now known (6, 18). Knowledge of the sequence of the

gene and of the mRNA transcribed from it was essential for determining the accuracy of the splicing reaction.

Four templates were used in this study (Fig. 1A). Plasmid C1 containing the 5.2 kb *EcoRI* insert spanning the β -globin gene up to codon

Fig. 1. A. Structure of the human β -globin gene. Coding regions are indicated as black rectangles. Intervening sequences and untranslated regions are indicated as open rectangles. CAP refers to the location of the capped end of mature RNA. Arrows represent the RNA transcribed in vitro from BamHI cut (479nt) and EcoRI cut (1395 nt) plasmid C1 and from PstI cut (2221 nt) plasmid β Pst.

B. Synthesis of cDNA on β -globin RNA transcribed in vitro. The primer (Sau96I- Sau3aI β -globin gene DNA fragment) was hybridized and extended by reverse transcriptase either on unspliced (upper arrow) or on spliced (lower arrow) RNA. The length (in nucleotides) of the full extension products before and after cleavage with HaeIII is indicated.

C. Sequence of the RNA and cDNA from the 5' end of the primer (underlined) to the HaeIII site. Upper sequence unspliced, lower sequence, spliced RNA, respectively. The numbers indicate the positions in unspliced RNA relative to the cap site. The single A over the unspliced sequence represents a point mutation in β -thalassemic β -globin RNA (6). (nt - nucleotide; IVS - intervening sequence.)

121 was digested either with restriction enzyme BamHI (template I) or EcoRI (template II). These enzymes cleave the β -globin gene within the second and third exons respectively. Therefore, the runoff transcription products from these templates contained only the first intervening sequence for the BamHI cut plasmid and both intervening sequences for EcoRI cut plasmid. Both transcripts terminate upstream from the polyadenylation site. Plasmid β PstI carrying the 5.5 kb PstI insert spanning the entire β -globin gene was digested with Pst I (template III). The Pst I site in the β -globin DNA is located 540 nucleotides downstream from the presumed polyadenylation site (18). In the in vitro system we used transcription by RNA polymerase II is not terminated efficiently (12). Therefore, the runoff transcript from the PstI cut plasmid will carry 540 nucleotides of sequence downstream from the polyadenylation site of the β -globin gene (Fig. 1A and 2).

Plasmid C6 carries an EcoRI insert similar to that in plasmid C1. However, the inserted DNA was obtained from a patient with the β^+ -thalassemic phenotype. It has been recently shown that this DNA carries a point mutation (G-A) located 21 nucleotides upstream from the acceptor site of the small intervening sequence. This single base substitution creates an additional aberrant acceptor site within the first intron (6). Recent reports indicate that aberrant splicing does occur in vivo (4,5,8) and in transfected cells (7, 8). We have used EcoRI cut plasmid C6 as a template (template IV) for in vitro RNA synthesis to determine whether aberrant splicing is also carried out in the in vitro system, and as a control for the demonstration of splicing of normal transcripts.

The templates discussed above allowed us to study the faithfulness of splicing in the in vitro system as well as the effect of the substrate structure on the efficiency of splicing.

Analysis of the RNA synthesized and spliced in vitro.

β -globin mRNA runoff transcripts were synthesized essentially as described (11, 12) with minor modifications (see Materials and Methods). The efficiency of the transcription was assayed by polyacrylamide gel electrophoresis of the purified transcripts (Fig. 2). The mobility of the transcribed globin RNA was as expected for unspliced runoff transcripts. In lanes "EcoRI" and "PstI" there are also additional bands visible which migrated faster than globin RNA. The estimated size of the majority of the bands is not compatible with their being products of normal splicing. The origin of these fragments is under investigation. The concentrations of the DNA template required for optimum transcription and splicing in the

Fig. 2. Analysis of RNA synthesized *in vitro*. RNA was synthesized *in vitro* as described in Materials and Methods from cloned DNA templates. BamHI and EcoRI indicate transcripts obtained when plasmid C1 cut with BamHI and EcoRI respectively served as templates. PstI indicates transcripts obtained with plasmid β Pst cut with PstI as a template. For each template the left lane shows transcript obtained with 100 µg/ml of the DNA, the right lane shows transcripts obtained at 200 µg/ml of the DNA. The indicated sizes (in nucleotides) are for the following RNA markers: 158-5.8S, 171-U1, 196-U2, 216-U3, 300-7S, 555-transcript of Alu family clone (27), 1748-human ϵ -globin RNA transcribed *in vitro* on RI-cut plasmid TEG51B (28). The sizes 479, 1395 and 2221 are for transcripts shown in Lane BamHI, EcoRI and PstI, respectively. The electrophoresis was on a 5% urea-acrylamide gel (40x20x0.2) (15). 1/30 of the total amount of transcribed RNA was analyzed on the gel. The gel was autoradiographed overnight at -70°C with DuPont Lightning plus screen and Fuji RX film.

concentrated extract was found to be 100 to 200 µg/ml. These DNA concentrations are 2-4 fold higher than those used in the standard transcription reaction (11, 12).

RNA synthesized in vitro was then analyzed for spliced products using the primer extension technique (16). The primer selected for the reverse transcriptase reaction hybridized to β -globin mRNA 24 nucleotides downstream from the 3' end of the small intervening sequence. The primer (29 nucleotides long) extended to the 5' end of the unspliced β -globin RNA gave rise to a cDNA product 325 nucleotides long. The same primer, extended on the spliced RNA gave rise to a product 195 nucleotides long. cDNA obtained by primer extension can be digested with HaeIII which cleaves single-stranded DNA. One of the HaeIII sites in the human β -globin gene is located at position 132, 10 nucleotides upstream from the 5' end of the small intron. Cleavage of β -globin cDNA with this enzyme should give fragments 193 and 63 nucleotides long for unspliced and spliced RNA respectively (Fig. 1B).

The results of a representative experiment are shown in Fig. 3A. In the control sample (Lane 4) cDNA was synthesized on human reticulocyte RNA. Lane 2+ shows HaeIII cut cDNA derived from the β -globin RNA transcribed from template II. The band 63 nucleotides long is clearly visible, indicating that a fraction of the in vitro synthesized RNA is accurately spliced in this system. A large amount of the RNA remains unspliced leaving the band 193 nucleotides long. As expected, a 63-nucleotide band is not visible in the material not treated with HaeIII (lane 2-). The possible reasons why there is only a faint band of 195 nucleotides (spliced uncut) visible in lane 2- are threefold. First, partial degradation of RNA during workup and reverse transcription may lead to the creation of cDNA molecules with heterogeneous 3' ends. Second, the same result would be obtained if there was inefficient extension by reverse transcriptase. The third reason would be heterogeneity at the 5' ends of the RNA, due to nonspecific initiation by RNA polymerase II. After cutting with HaeIII, all the cDNA molecules which were extended on spliced RNA past the HaeIII site but terminated before the normal 5' end of the β -globin RNA would be scored as spliced products. The factors mentioned above - breakage of RNA, inefficient extension and nonspecific initiations by RNA polymerase II - might be the source of the large number of bands visible on the gel. The fact that the bands have different intensities, and that the same size bands are visible in the extensions of two different transcripts (lanes 1

Fig. 3. Analysis of cDNA synthesized on *in vitro* transcribed β -globin RNA. The primer, labeled at the 5' end was hybridized, extended and the cDNA cleaved with *Hae*III as described in Materials and Methods. The aliquots of the cDNA were analyzed on an 8% urea-acrylamide gel (40x20x0.03 cm). The cDNA was synthesized on β -globin RNA transcribed *in vitro* from: Lane 1, *Pst*I cut plasmid β Pst; Lane 2, *Eco*RI cut plasmid C1, Lane 3, control, without DNA template added, Lane 4 RNA isolated from human reticulocytes. "-" indicates the extension products before and "+" after cleavage with *Hae*III. The sizes of the indicated bands were determined by comparison with cDNA synthesized on RNA isolated from human reticulocytes. 1/8 of the synthesized cDNA was loaded per lane. The gel was autoradiographed for 48 hrs as described in the legend to Fig. 2.

(63 - spliced, cut with *Hae*III; 193 - unspliced, cut with *Hae*III; 195 - spliced, not cut with *Hae*III [see Fig. 1]).

and 2) indicates that the breakage, nonspecific RNA polymerase initiation sites, and/or termination points for reverse transcriptase are distributed in a nonrandom fashion. There do not seem to be any peculiarities in the sequence of the RNA that could explain this phenomenon.

Two additional bands were generated by HaeIII cleavage of cDNA synthesized on RNA transcribed from templates III and II (Fig.3A, lanes 1+ and 2+, indicated by arrows). There was insufficient material in these bands for sequence analysis and their origin is not known.

The control experiment in which no DNA template is added, also exhibited some weak cDNA bands (lane 3). However, none of the bands had the mobility of the β -globin cDNA cut with HaeIII. These bands might originate from the RNA present in the extract. Since the ratio of RNA to primer in the control reaction was low, one can expect some relatively unspecific priming to occur, which then leads to the extended products. A similar result was obtained when RNA obtained from uninfected CV-1 cells was reverse-transcribed with simian virus 40 (SV40) DNA as a primer. The sequence of these extension products was different from that of SV40 (unpublished data).

Lane 1 in Fig. 3 shows cDNA extensions derived from the RNA transcribed from template III. Surprisingly, no spliced product was visible. (The weak band present in this region migrates slightly faster than the 63-nucleotide band. This can be more clearly seen in Fig. 5, lane 1.) The possible reasons for this result are discussed below.

Sequencing of the extension products.

To further determine the accuracy of the in vitro splicing reaction, we have sequenced the 63 nucleotide long DNA fragment discussed above (Fig. 4). After isolation from the gel the DNA was sequenced by the procedure of Maxam and Gilbert (17). Since the amount of the radioactivity was very small (about 1000 cpm) only two sequencing reactions, G and C, were carried out. The reasons why these two reactions were chosen are the following. The C reaction was essential to establish the accuracy of splicing since C is the last nucleotide in the coding sequence at the 3' end of the small intron (Fig. 1C, cDNA sequence). There are four C's and four G's between the HaeIII site in the coding sequence and the splice junction which are diagnostic for the spliced material. A and T reactions are less informative since there is only one A and one T between the HaeIII site and the splice junction.

The reaction products were electrophoresed through an 80 cm gel to

Fig. 4. Sequencing of the β -globin cDNA. The 63 nucleotide long fragments (see Fig. 3 and 6) of cDNA synthesized on β -globin RNA either isolated from human reticulocytes (panel B1) or transcribed *in vitro* on EcoRI cut plasmid C1 (panel B2) were sequenced and analyzed on 8% urea-acrylamide gels (80x20x0.03 cm). Only G and C reactions are shown (see text). Panels A and C show sequence of spliced and unspliced RNA respectively.

obtain maximum separation and distinct patterns of the products near the 3' terminus of the cDNA. Fig. 4B, lane 2, shows the sequence of the 63-nucleotide long cDNA fragment derived from the *in vitro* transcribed RNA, proximal to the HaeIII site. Lane 1 shows the sequence of the similar cDNA fragment which was synthesized on the β -globin mRNA isolated from human reticulocytes. Both sequences were identical and complementary to the

sequence of the spliced β -globin mRNA (Fig. 4A) and different from that of the unspliced precursor (Fig. 4C). The sequence corroborated the findings of the primer extension experiment shown in Fig. 3 and further demonstrated that the small intervening sequence in human β -globin RNA was faithfully and accurately spliced in vitro.

The band 193 nucleotides long was also sequenced and was complementary to the unspliced β -globin RNA (not shown).

Efficiency of the in vitro splicing reaction.

We have attempted to quantitate the yields of the spliced RNA only for the RNA transcribed from template II. This RNA was spliced most efficiently (Fig. 5 and paragraph below). The efficiency of the splicing reaction could be quantitated only approximately by the densitometry of autoradiograms of primer extension gels from four different experiments. The yield of the spliced product (63 nucleotide band) varied from 2-10%. The yield was calculated as a ratio of the area under the peak of 63 nucleotide band to the total area under the peaks contained between the bands 63 and 193 nucleotides long. The bands longer than 193 were due to the material which was not digested by HaeIII. Some bands shorter than 63 nucleotides were due to contaminants (Fig. 3, lane 3); bands closer to the primer were obscured by strong radioactivity from the unhybridized primer. Therefore, bands longer than 193 and shorter than 63 nucleotides were not scored. The upper yield value may be an overestimate, since, in some gels the band 193 nucleotides long (HaeIII-cleaved cDNA synthesized on unspliced RNA) was overexposed, probably beyond the linear response of the film. On the other hand, some omitted short extension products (shorter than 63 nucleotides) could contribute to the amount of both spliced and unspliced RNA scored.

Effect of RNA structure on in vitro splicing.

In the experiment shown in Fig. 5, RNA was transcribed from three different templates: PstI cut plasmid β Pst, EcoRI cut plasmid C1, and EcoRI cut plasmid C6 (templates III, II and IV, respectively). cDNA was then synthesized, cut with HaeIII and analyzed on the gel.

RNA transcribed in vitro from template II was spliced with the highest efficiency. The RNA transcribed from template III was not spliced at all in the experiments shown (Fig. 5, lane 1). In other experiments, traces of the spliced product could be detected. RNA transcribed from template I, i.e., terminated shortly downstream from the small intervening sequence (see Fig. 1A), was spliced more efficiently than the PstI runoff RNA from

Fig. 5. Analysis of cDNA was carried out as described in Materials and Methods and in Fig. 3. All cDNAs shown here were cleaved with *Hae*III and were synthesized on β -globin RNA transcribed *in vitro* from: Lane 1, *Pst*I cleaved plasmid β Pst; lane 3, *Eco*RI cleaved plasmid C1; lane 4, *Eco*RI cleaved plasmid C6. Lane 2 and 5 show cDNA derived from reticulocyte RNA and from β -thalassemic RNA isolated from transfected CV1 cells respectively (see text). The numbers 63, 82 and 193 indicate the position of *Hae*III cut cDNA corresponding to the normal splice, the new splice in β thalassemia, and unspliced RNA. The band at 193 due to unspliced RNA has the same mobility on this gel as the band 195 nucleotides long representing uncut extension products on full-length mRNA. The amount of *Hae*III restriction endonuclease that could be used without excessive phosphatase activity often failed to give complete cutting of the single-stranded cDNA. This accounts for the band at "195" in reticulocyte RNA.

template III but less than RNA transcribed from the EcoRI cut template II (not shown). As a control, RNA from normal reticulocytes was analyzed by reverse transcription and HaeIII digestion of the single-stranded cDNA (Fig. 5, lane 2). The principal product was the 63-nucleotide long band whose sequence corresponded to cut cDNA from spliced RNA. The band located at approximately 195 residues was due to the fact that in this sample the hydrolysis by HaeIII was not complete. This band was sequenced and confirmed to be uncut cDNA from spliced RNA (195 residues), rather than cut cDNA from unspliced RNA (193 residues).

Lane 5 shows cDNA synthesized on thalassemic β -globin RNA. This RNA was obtained from CV-1 cells transfected with DNA obtained from the recombinant plasmid pVB SVC6 (8). This construct contained the BamHI fragment of plasmid C6 (spanning the small intervening sequence of thalassemic β -globin gene) cloned into the SV40-pBR 322 vector. It has been shown (8) that the thalassemic β -globin gene was expressed and the RNA accurately spliced in vivo in this system. In fact, the cDNA pattern (lane 5) shows the presence of bands 63 and 82 nucleotides long that are products derived from RNA with normal and aberrant splices, respectively. Unspliced transcript was present in amounts similar to that seen when the normal gene was used as a template. However, the cDNA products corresponding to the two forms of spliced RNA were present in much lower amounts than the single product derived from the normal gene.

In vitro spliced RNA, transcribed from template IV (lane 4) shows a virtually identical pattern with both bands visible, but the 63-nucleotide band is markedly reduced compared with that from the normal gene. The abnormally spliced band must be derived from the transcript of the added DNA since HeLa cells would not be expected to contain the β^+ -thalassemia gene. This provides a further control demonstrating that the 63-nucleotide long band seen with the normal template was derived from in vitro spliced RNA. Preliminary sequencing seems to indicate that the band 82 nucleotide long has the sequence expected for the aberrantly spliced material. Thus, it seems likely that the splicing enzymes present in the extract were able to splice in vitro the thalassemic RNA in both of the modes of splicing that occurs in vivo.

DISCUSSION

We have presented evidence that a small intervening sequence from human β -globin RNA was accurately spliced in vitro by concentrated crude extracts prepared from HeLa cells.

Three factors probably contributed to obtaining the splicing reaction in vitro. First, the protein concentration of the extract and the concentration of the DNA template is high which in turn presumably leads to a higher concentration of the RNA precursor - the substrate for the splicing enzyme.

Second, the choice of the small intervening sequence of the human β -globin gene may be important. This intervening sequence is relatively short and it can be expected that any type of intermediate or nonspecific reaction which would lead to an incorrectly spliced product would be minimized.

Third, analysis of the products by the primer extensions is very sensitive and is the method of choice to obtain unambiguous evidence that cleavage and ligation of the RNA in the splicing reaction is accurate.

The possibility that HeLa cells express trace amounts of β -globin RNA can be excluded since no spliced product (band 63 nucleotides long) is visible in the control sample without added DNA template and in the transcript derived from PstI cut plasmid β Pst (template III). The fact that in thalassemic RNA there are lower amounts of the spliced product also supports this argument. The sequence around the β -globin gene is known (unpublished data). There are no sequences which are homologous to the β -globin gene within the length of the DNA templates used for transcription. This excludes the possibility that the spliced RNA might be transcribed from the "processed gene" (20) or a pseudogene which lost its intervening sequence.

The RNA which was spliced most efficiently (derived from EcoRI cut template II) was terminated upstream from the normal polyadenylation site. It is therefore very likely that the in vitro transcripts were spliced without being polyadenylated.

The fact that RNA transcribed from template II was spliced much more efficiently than the longer transcript (from template III) is rather striking. It could be expected that RNA containing the whole third exon, a polyadenylation site and sequences farther downstream would provide a better substrate for the splicing enzyme. One explanation for our results is that downstream sequences inhibit splicing, either directly, or by preventing precursor RNA from attaining its proper in vivo conformation.

This hypothesis would further imply that cleavage of mRNA precursors at the polyadenylation site might facilitate splicing, even in the absence of polyadenylation, reconciling the data of Zeevi et al. and Manley et

al.(21, 22), who have shown that adenovirus RNA can be spliced in vivo and in isolated nuclei without being polyadenylated, and the fact that cells may contain intranuclear polyadenylated unspliced RNA. A role for downstream sequences in inhibiting upstream splicing could also account for the observation of "coupled splices" in late SV40 RNA in which an intraleader splice acceptor apparently only functions in RNA in which there is a downstream 16S splice (23).

The efficiency of splicing in our system was rather low. Goldenberg and Raskas reported that 80% of the precursor RNA was processed in vitro (10), but did not sequence their products. In addition to the use of extract from a different cell line, these authors used as a substrate a 28S RNA isolated from cultures early in adenovirus infection. This RNA is polyadenylated and appears to be a precursor to one of the early adenovirus RNAs. It will be interesting to see if polyadenylation of the precursors to β -globin mRNA increases the efficiency of splicing in the HeLa cell extract.

Recent reports indicate that in vitro splicing of adenovirus RNA precursors (9) could not be consistently reproduced (24). It should be stressed that three different extract preparations obtained from HeLa cells from two different sources have been used in the experiments described above with consistent results.

When β -globin RNA transcribed from template II was purified by phenol extraction and then added to the extract under standard conditions, after a 15-min incubation, most of the RNA was unspecifically degraded. Transcripts which had not been phenol-extracted were stable in the extract for 1 to 2 hr (unpublished data). This result suggests that RNA transcribed in vitro is quickly converted into a ribonucleoprotein complex. It is possible that the formation and structure of this complex affects the splicing reaction. In this study we did not attempt to analyze the splicing of the large intron in β -globin RNA. Work on this question is in progress.

In erythroid cells and in recombinant SV40 genomes containing the small globin intervening sequence (8), but not in transfection studies with the whole globin gene (7), cytoplasmic RNA production from the β + β -thalassemic globin gene is substantially reduced, compared with the normal gene RNA. This could be a consequence of competition between degradation and splicing of the globin precursor, or excessive degradation of the abnormally spliced precursor prior to its reaching the cytoplasm.

The present results indicate there is a primary defect in the rate at which the thalassemic RNA can be spliced, even when the precursor is stable, and favor the interpretation that successful completion of splicing is directly related to the appearance of cytoplasmic RNA.

Another point of interest is the implied role in splicing of ribonucleoproteins containing small nuclear RNAs (snRNPs) (25). Obviously, in vitro splicing systems should facilitate studies of this question. Fractionation of the extract and/or supplementing it with snRNPs should increase efficiency of splicing. Conversely, addition of antibodies directed against snRNPs might inhibit or arrest the splicing reaction (26).

The results presented above indicate that whole cell extract from HeLa cells can serve as a good starting material for the study of the molecular mechanism of the splicing reaction, and that the structure of the primary transcript is critical for demonstrating in vitro splicing.

ACKNOWLEDGMENTS

We thank Dr. Y. Fukumaki for the thalassemic RNA, Dr. E. Benz for human reticulocyte RNA and Dr. P. Ghosh for help in initial primer extension experiments. This investigation was supported by grant number CA-16038 from the National Cancer Institute, DHHS, and by a National Institutes of Health postdoctoral fellowship to R. K.

REFERENCES

1. Abelson, J. (1979) *Ann. Rev. Biochem.* 48, 1035-1069.
2. Breathnach, R. and Chambon, P. (1981) *Ann Rev. Biochem* 50, 349-383.
3. Sharp, P. A. (1981) *Cell* 23, 643-646.
4. Maquat, L. E., Kinniburgh, A. J., Beech, L. R., Honig, G. R., Lazerson, J., Ershler, W. B. and Ross, J. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4287-4291.
5. Kantor, J. A., Turner, P. H. and Nienhuis, A. W. (1980) *Cell* 21, 149-157.
6. Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., deRiel, J. K., Manley, J. L., Gefter, M. L., Forget, B. G. and Weissman, S. M. (1981) *Proc. Natl. Acad. Sci. USA* 78, 2455-2459.
7. Busslinger, M., Moschonas, N., and Flavell, R. A. (1981) *Cell* 27, 289-298.
8. Fukumaki, Y., Ghosh, P. K., Benz, E. J., Reddy, V. B., Lebowitz, P., Forget, B. G. and Weissman, S. M. (1982) *Cell* 28, 585-593.
9. Weingartner, B. and Keller, W. (1981) *Proc. Natl. Acad. Sci. USA* 78, 4092-4096.
10. Goldenberg, C. J. and Raskas, H. J. (1981) *Proc. Natl. Acad. Sci.* 78, 5430-5434.
11. Handa, H., Kaufman, R. J., Manley, J. L., Gefter, M. L., Sharp, P. A. (1981) *J. Biol. Chem.* 256, 478-482.
12. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. and Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* 77, 3855-3859.

13. Sugden, B., and Keller, W. (1973) *J. Biol. Chem.* 248, 3777-3788.
14. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. and Maniatis, T. (1978) *Cell* 15, 1157-1174.
15. Donis-Keller, H. (1979) *Nucl. Acids Res.* 7, 179-188.
16. Ghosh, P. K., Reddy, V. B., Piatak, M., Lebowitz, P. and Weissman, S. M. (1980) *Meth. in Enzymol.* 65, 580-595.
17. Maxam, A. M. and Gilbert, W. (1980) *Meth. in Enzymol.* 65, 499-560.
18. Lawn, R. M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) *Cell* 21, 647-651.
19. Sanger, F. and Coulson, A. R. (1978) *FEBS Lett.* 87, 107 (1978).
20. Hollis, G. F., Hieter, P. A., McBride, O. W., Swan, D. and Leder, P. (1982) *Nature* 296, 321-325.
21. Zeevi, M., Nevins, J. R. and Darnell, J. E. (1981) *Cell* 26, 39-46.
22. Manley, J. L., Sharp, A. and Gefter, M. L. (1979). *J. Mol. Biol.* 135, 171-197.
23. Piatak, M., Subramanian, K. N., Roy, P. and Weissman, S. M. (1981) *J. Mol. Biol.* 153, 589-614.
24. Crefeld, T., Hernandez, N., Sass, K., Sergeant, A., Solnick, D., Weingartner, B. and Keller, W. (1982) in *RNA Processing*, p. 136, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
25. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. and Steitz, J. A. (1980) *Nature (London)* 283, 220-224.
26. Yang, V. W., Lerner, M. R., Steitz, J. A. and Flint, S. J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1371-1375.
27. Duncan, C. H., Jagadeeswaran, P., Wang, R. R. C. and Weissman, S.M. (1981) *Gene* 13, 185-196.
28. Kaufman, R. E., Kretschmer, P. J., Adams, Y. W., Coon, H. C., Anderson, W. F. and Nienhuis, A. W. (1980) *Proc. Natl. Acad. Sci. USA* 77, 4229-4233.