Volume 16 Number 16 1988 Nucleic Acids Research

# Model for tissue specific Calcitonin/CGRP-I RNA processing from *in vitro* experiments

R.A.L.Bovenberg*, G.J.Adema, H.S.Jansz and P.D.Baas

Institute of Molecular Biology and Medical Biotechnology and Laboratory for Physiological Chemistry, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received May 17, 1988; Revised and Accepted July 29, 1988

ABSTRACT

The Calcitonin/CGRP-I (CALC-I) gene is known to be expressed in a tissue specific fashion resulting in the production of Calcitonin mRNA in thyroid C-cells and CGRP-I mRNA in particular nerve cells. The alternative RNA processing reactions include splicing of exons 1,2 and 3 to exon 4 and poly (A) addition at exon 4 (Calcitonin mRNA) or splicing of exons 1, 2 and 3 to exons 5 and 6 and poly (A) addition at exon 6 (CGRP-I mRNA). Using a model precursor RNA containing the exon 3 to exon 5 region of the human CALC-I gene we have investigated the Calcitonin- and CGRP-I mRNA-specific processing reactions in vitro, in nuclear extracts of Hela, PC12 and Ewing-1B cells, respectively. Extracts of PC12- and Ewing-1B cells were expected to perform CGRP mRNA-specific splicing, whereas Calcitonin mRNA specific processing was expected to occur in Hela cell extracts. Surprisingly, CGRP mRNA-specific splicing of exon 3 to exon 5 was the predominant reaction in all three extracts. Significant Calcitonin mRNA- specific splicing of exon 3 to exon 4 only took place upon elimination of the dominant downstream 3' splice site used in CGRP mRNA-specific splicing. This elimination occurs most definitively by cleavage at the Calcitonin mRNA specific poly(A) site at exon 4 which may then be the major regulatory mechanism for tissue-specific expression of the CALC-I gene.

## INTRODUCTION

The rat Calcitonin/CGRP gene was one of the first examples of a cellular gene exhibiting alternative RNA processing in a tissue specific fashion (1,2). Expression of the gene in thyroid C cells results predominantly in the formation of Calcitonin (CT) mRNA containing the exons 1 to 4 with a poly(A) tail at exon 4 (3,4). Exons 1 to 3 are also present in CGRP mRNA which further consists of the CGRP-encoding exon 5 and the non-coding, polyadenylated exon 6. The CT-encoding exon 4 is excluded from CGRP mRNA by splicing of exon 3 to exon 5. CGRP mRNA is the exclusive product of RNA processing in particular cells in the central and peripheral nervous system (1,5). The human Calcitonin/CGRP-I (CALC-I) gene is completely analogous to its rat counterpart in structure and mode of alternative RNA processing (6-12). In man the CGRP-I polypeptide was detected in several neural cells in the brain, in pituitary

gland cells, the spinal cord and in medullary thyroid carcinoma (MTC), a tumor of the thyroid C cell (8, 13-15).

In a previous report we have analyzed processing of CALC-I RNA in vivo, in human MTC (16). The results indicated that RNA processing of the primary transcript leads to the accumulation of a 3.3 kilobases (kb) RNA intermediate which can still be processed into either CT or CGRP-I mRNA. This 3.3 kb RNA is polyadenylated at exon 6 and contains the spliced exons 1,2 and 3, the complete intron 3- exon 4- intron 4 region and the spliced exons 5 and 6. These observations suggest that this RNA is the main target for putative regulatory factors involved in the tissue specific processing patterns. CT mRNA is formed by splicing of intron 3 and poly(A) addition at exon 4. Alternatively, the whole intron 3- exon 4- intron 4 region is spliced from the 3.3 kb RNA yielding CGRP-I mRNA.

To analyze the alternative RNA processing reactions at the molecular level a model CALC-I precursor RNA was engineered which contains all regulatory sequences of the 3.3 kb RNA precursor. This precursor RNA was processed in vitro, in nuclear extracts of Hela, PC12 and Ewing-IB cells (for recent reviews see 17,18). The rat PC12 and human Ewing-IB cell lines are both attractive candidates for studying CGRP-I mRNA-specific processing. Transfection of the rat Calcitonin/CGRP gene into PC12 cells results in the expression of CGRP mRNA (19). In Ewing-IB cells the endogenous CALC-II gene, which is closely related to the CALC-I gene, is expressed into CGRP-II mRNA (20). Surprisingly, CGRP mRNA-specific splicing of exon 3 to exon 5 is strongly favored over CT mRNA- specific splicing of exon 3 to exon 4, not only in PC12 and Ewing-IB extracts but also in nuclear extracts of Hela cells, which are not of neural origin (21). These observations indicate that there are strong differences in strength between the two alternative 3' splice sites and suggest a model for alternative RNA processing which will be presented.

## MATERIALS AND METHODS

### Materials

SP6 RNA polymerase, RNAsin, DNAse I (RNAse free) and AMV Reverse transcriptase were purchased from Promega Biotec. Restriction endonucleases, nucleotides and the cap dinucleotide GpppG were from Pharmacia. Oligonucleotides were synthesized on a Biosearch 8600 DNA synthesizer. [$\gamma$-$^{32}$P]-ATP (3000 Ci/mmol), [$\alpha$-$^{32}$P]-CTP (800 Ci/mmol) and [$\alpha$-$^{32}$P]-dCTP (800 Ci/mmol) were from Amersham.

### DNA templates and in vitro transcription

The BssHII-NarI restricted DNA fragment of the human CALC-I gene containing the 3' half of exon 3, intron3- exon 4- intron4- and 5' half of exon 5 (12) was purified, blunt-ended by T4 DNA polymerase and cloned in SmaI digested

pSP64. Deletions in intron 3 and intron 4 were obtained by oligonucleotide directed mutagenesis after introduction of the cloned insert into M13mp11 (22). Oligonucleotides of a length of 24 nucleotides were synthesized complementary in sequence to the 12 consecutive nucleotides immediate upstream and downstream of the point of deletion. The insert carrying both the intron 3 and intron 4 deletions was reintroduced in pSP64: pSP64CALCI-dI3I4.

Template DNA pSP64CALCI-dI3I4 was cleaved with EcoRI or BclI, extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and then precipitated with ethanol. Capped transcripts were synthesized in vitro using a cap/priming protocol (23). Standard in vitro transcription reactions contained: 1.5-2 µg template DNA, 40 mM Tris.HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 500 µM ATP, 500 µM UTP, 50 µM GTP, 1 mM GpppG, 50 µM CTP, 20 µCi [$\alpha$-$^{32}$P]-CTP (800 Ci/mmol), 50 U RNasin and 7.5 U SP6 RNA polymerase in a total volume of 50 µl. RNAs of higher specific activity were synthesized when secondary analysis, like debranching, were performed. Transcription reactions were performed for one hour at 40°C . Template DNA was then degraded by RNAse free DNase I. Transcripts were purified by extraction with phenol/chloroform-/isoamylalcohol (25:24:1). After precipitation with ethanol, transcripts were further purified by electrophoresis on 4% (30:1) cross-linked 8.3 M acrylamide-urea gels (24). Full length precursor RNA was subsequently isolated from the gel by electroelution. Transcripts were stored under 75% ethanol at 20°C.

Nuclear extracts and in vitro splicing

Hela cell nuclear extracts were prepared by extraction of nuclei with 0.2, 0.4 and 0.6 M NaCl as described by Dignam (25). Following the same procedure Hela nuclear extracts were also prepared by extraction of nuclei with 0.6 M KCl (26). Nuclear extracts from rat PC12 and human Ewing-IB cell lines were prepared by extraction with 0.4 M NaCl (25). Splicing reactions were essentially performed as described by Krainer (27) and contained, in a total volume of 25 µl: 10 µl nuclear extract In Dignam buffer D, 20 mM creatine phosphate, 0.5 mM ATP, 1-2 mM $MgCl_2$, 2.7% polyvinyl alcohol and 0.2 nM [$^{32}$P]-labelled precursor RNA. Reactions were performed at 30°C. Splicing reactions were stopped by the addition of 175 µl proteinase K solution (50 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.2% SDS, 10 mM EDTA, 5 µg/ml tRNA and, freshly added, 0.25 mg/ml proteinase K). After an incubation of 20 minutes at 30°C the reaction mixtures were extracted once with phenol/chloroform/isoamylalcohol (25:24:1) and then precipitated with ethanol. Splicing reactions were scaled up twenty fold to obtain sufficient material for secondary analysis in debranching and primer extension experiments.

Acrylamide gel analyses of RNA

[$^{32}$P]-RNA from analytical splicing reactions was analyzed routinely on thin 4%, 30:1 cross-linked, 8.3 M acrylamide-urea gels (24). [$^{32}$P]-RNA was detected after electrophoresis by autoradiography on X-ray films. Occasionally, the percentage of acrylamide or the amount of cross-linking was increased to identify putative lariat RNAs. [$^{32}$P]-RNA from preparative splicing experiments was analyzed on similar 1 mm thick acrylamide-urea gels.

Debranching experiments

Presumptive lariat RNAs were purified by acrylamide gel electrophoresis, located by autoradiography, and isolated by electroelution from the gel. Debranching was performed as described (28), using Hela cell S100 fractions. After proteinase K treatment as described above and extraction with phenol/-chloroform/isoamylalcohol (25:24:1) RNA was precipitated with ethanol. Analysis of [$^{32}$P]-RNA was on acrylamide-urea gels as described above.

Primer extension analyses

Primer extension was performed using two synthetic primers: primer I4 (5' CTGATTTGCAGGATGGAG 3') complementary to the 3' end of intron 4 in pre-mRNA,

and primer E5 (5' AGGCTCTCTTCTGGGCAATG 3') complementary to nucleotides 3 to 24 of exon 5. Primers were 5' end-labeled with [$\gamma$-$^{32}$P]-ATP and T4-polynucleotide kinase (24). RNA species of interest were purified by acrylamide-urea gel electrophoresis and isolated as described above. Primer extension was performed essentially as described by Solnick (29). In a slight modification of this procedure primer extension was performed in the presence of 0.25 µM [$\alpha$-$^{32}$P]-dCTP (800 Ci/mmol) for the first 15 minutes at 42°C, followed by a chase of 30 minutes, at the same temperature, with excess dCTP. In some experiments with precursor RNA dI3I4-E as the template RNA, ddTTP was included in the reaction to obtain a partial sequence of adenosine residues in the precursor RNA. After cDNA synthesis, [$^{32}$P]-RNA was degraded by the addition of NaOH to the reaction mixture and incubation for 30 minutes at 55°C. After neutralization the cDNA products were precipitated with ethanol and finally analyzed on 10%, 20:1 cross-linked, 8.3 M acrylamide-urea gels. Detection of [$^{32}$P] cDNA was done by autoradiography.

## RESULTS

### Model precursor RNA

To study the tissue-specific splice site selection in human CALC-I RNA *in vitro*, we first constructed a model precursor RNA appropriate for *in vitro* splicing analyses (Figure 1). To that end the BssHII-NarI fragment, spanning exon 3 to exon 5, of the human CALC-I gene (12) was cloned in SP64 (30). In a second step we deleted approximately 860 nucleotides from intron 3 ( the size and sequence of intron 3 has not been determined completely) and 479 nucleotides from intron 4. The deletions were introduced to enhance the efficiency of the *in vitro* splicing reaction (31,32). They do not affect the sequences which are directly involved in the alternative processing reactions: the 5' splice site (33) of intron 3, the 3' splice site (33) and branch point region (34,35) of intron 3, the sequences necessary for poly(A) addition (36) at exon 4 and the 3' splice site and branch point region preceding exon 5. *In vitro* transcription of EcoRI cleaved template DNA by SP6 RNA polymerase (30) generates a precursor RNA, dI3I4-E, with a length of 1096 nucleotides which contains the 3' 88 nucleotides of exon 3, a 142 nucleotides long truncated form of intron 3, the complete exon 4 sequence (498 nucleotides ), a truncated form of intron 4 (178 nucleotides ) and the 5' 138 nucleotides of exon 5. The presence of vector derived nucleotides in the transcript increases the size of exon 3 and exon 5 to 129 and 149 nucleotides, respectively.

### In vitro splicing of precursor RNA dI3I4-E

We first analyzed the fate of precursor RNA dI3I4-E in nuclear extracts of Hela cells (expected to perform CT-specific processing) on the one hand, and in nuclear extracts of PC12 and Ewing-IB cells (expected to perform CGRP-I specific processing) on the other hand. Surprisingly, similar results were obtained in the different extracts. The highest efficiency of the splicing

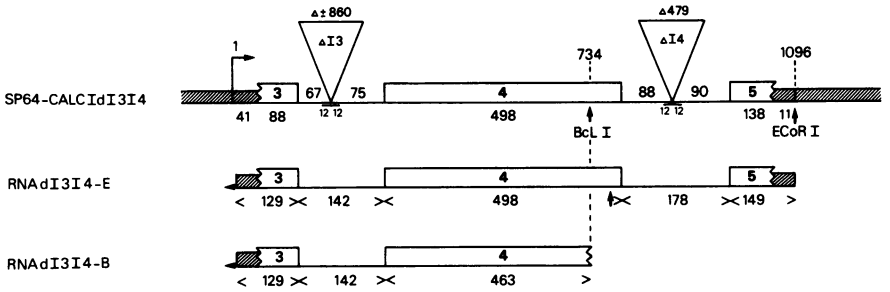



Figure 1. Schematic representation (not drawn to scale) of the structure of SP64CALCI-dI3I4 which contains the genomic CALC-I sequence from the BssHII site in exon 3 to the NarI site in exon 5 with truncated forms of intron 3 and intron 4 (above). CALC-I and vector sequences are represented by open numbered boxes (CALC-I exon sequences), straight lines (CALC-I intron sequences) and dashed lower case boxes (vector sequences) respectively. The size of the different exon-, intron- and vector regions is indicated in nucleotides. The intron 3-(dI3) and intron 4-(dI4) deletions are represented by triangles. The number of nucleotides removed by the deletions and the number of remaining 5' and 3' intron nucleotides in the introns are indicated. The position of the hybridization of the two synthetic deoxyoligonucleotides (both 24-mers) used to generate these deletions is represented by a solid line, located symmetrically around the point deletion. The site of transcription initiation is indicated ([*]) as well as the position of the recognition sequences of the BclI and EcoRI restriction enzymes (arrows). In vitro transcription of EcoRI or BclI cleaved SP64CALCI-dI3I4 template DNA with SP6 RNA polymerase results in capped (◀) run-off transcripts RNA dI3I4-E of 1096 nucleotides (middle) and RNA dI3I4-B of 734 nucleotides (below) respectively. The arrow indicates the position of the exon 4 poly(A) signal sequence in RNA dI3I4-E. The total length of the exon and intron regions in the two precursor RNAs is indicated in nucleotides.

reactions was observed using Hela nuclear extracts. A representative example of a time course experiment is shown (Figure 2A). During the time course two prominent RNA species are generated (RNA 1 and RNA 2 in Figure 2A) which display the characteristic anomalous mobility of lariat RNAs on denaturing acrylamide gels (31,35). The appearance of RNA 1 is preceded by a lag period of 30 minutes (lanes 1-2) which is characteristic for in vitro splicing (27). RNA 1 accumulates after 1 hour of incubation and decreases in concentration with longer incubation times (lanes 3-4). Its kinetics and retarded mobility suggests that RNA 1 is a lariat intermediate in a splicing reaction (34,35). RNA 2 appears only after 1 hour of incubation as a single RNA species (lane 3) and accumulates after three hours of incubation as a series of bands (lane 4). This suggests that RNA 2 is the corresponding free intron lariat of the same splicing reaction (31,35). The conversion of a discrete product lariat

RNA into multiple bands after prolonged incubation times might be due to the action of an 3' exonuclease which removes the linear tail from the lariat molecules. Similar observations were made by others (36). Further analysis (see below) showed that RNA 1 and RNA 2 are the lariat molecules involved in the splicing reaction of exon 3 to exon 5. The spliced product of this reaction, RNA 8, was detected with the expected size and kinetics (lanes 3-4). The accuracy of the splicing reaction was confirmed by a partial sequence spanning the splice junction obtained by primer extension analyses on gel

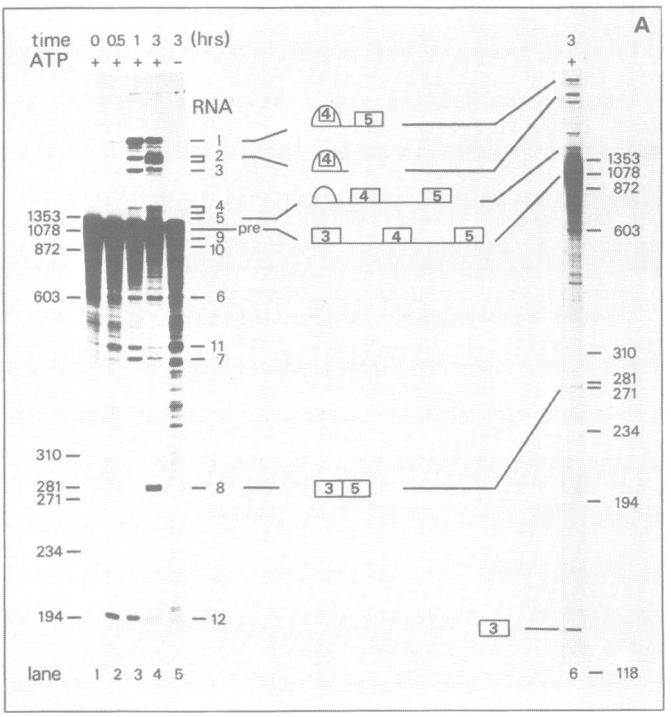
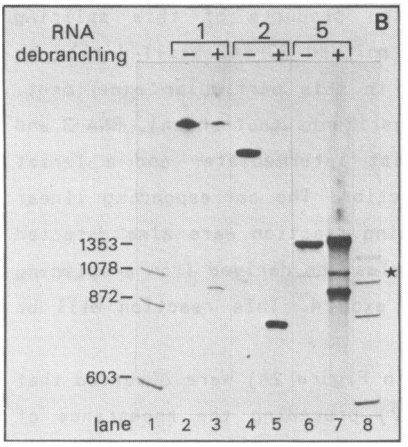



Figure 2. In vitro splicing of uniformly [$^{32}$P]-labeled precursor RNA dI3I4-E. (A) Individual splicing reactions were incubated for 0, 0.5, 1 and 3 hours (lanes 1-6) in the presence (+) or absence (-) of ATP and creatine phosphate in nuclear extracts of Hela cells at 30°C. Samples were prepared for analysis on 4% denaturing acrylamide gels as described in the materials and methods section. After electrophoresis, [$^{32}$P]-RNA was detected by autoradiography. The position of 5'end-labeled HaeIII restriction fragments of øx174 RF DNA in the gel is indicated. RNAs 1 to 12, described in the text, are indicated. RNAs 9 and 10 are obscured in this exposure time by the signal of unreacted precursor RNA which is present in excess. The structure of RNA 1, 2, 5, 8 and precursor (pre)RNA is schematically represented. CALC-I exons 3, 4 and 5 are indicated by open numbered boxes. In an independent experiment (lane 6) the free exon 3 intermediate formed in the E3-E4 and E3-E5 splicing reactions is detected.




(B) Lariat RNAs 1, 2 and 5 were isolated from a splicing reaction on a preparative scale by denaturing acrylamide gel electrophoresis. The purified lariat RNAs (-) (lanes 2, 4, 6) were debranched (+) (lanes 3, 5, 7) in S100 fractions of Hela cells. $^{32}P$ RNA was extracted and analyzed on 4% denaturing acrylamide gels as described above. DNA size markers (lanes 1, 8) are the same as in Figure 2A. The asterix in lane 7 indicates the position of precursor RNA dI3I4-E which is present as a contaminant in the RNA 5 preparation.

purified RNA 8 (not shown). The linear intermediate from this splicing reaction, free exon 3, is run of the gel in this particular experiment but was readily observed in other experiments using shorter running times (lane 6). Splicing was not observed in the absence of ATP (lane 5).

The notion that RNAs 1 and 2 are in lariat configuration was confirmed by the observation that the mobility of these RNAs was specifically and dramatically reduced on gels containing higher percentages of acrylamide (not shown). To characterize the putative lariat RNAs 1 and 2 in more detail debranching experiments (28) were performed (Figure 2B). After debranching RNA 1 has the mobility of a linear RNA species of approximately 980 nucleotides (lane 3). The size of this linear RNA corresponds very well with the expected size for the debranched intermediate lariat from the exon 3 to exon 5 (E3-E5) splicing reaction (954 nucleotides). RNA 2 is debranched into a linear RNA of approximately 800 nucleotides (lane 5), which is in good agreement with the expected size for the debranched product intron lariat from the E3-E5 splice (818 nucleotides).

Additional lariat RNAs, RNAs 3,4 and 5, are indicated in Figure 2A. RNA 5 is only detected as a faint band late in the time-course (lane 4). This RNA was also debranched into a linear RNA of approximately 980 nucleotides (Figure 2B, lane 7) suggesting that RNA 5 is the intermediate lariat in the splicing reaction of exon 3 to exon 4 (E3-E4). This observation is consistent with the notion that E3-E5 and E3-E4 splicing would result in different lariat intermediate RNAs which can be debranched into identical linear RNAs. The lariat formed in E3-E5 splicing consists of a large loop and a small tail whereas a small loop-large tail lariat is formed in E3-E4 splicing. These differences in lariat structure explain for the different mobility of these

RNAs on denaturing acrylamide gels. The two products of this splicing reaction, free intron 3 in lariat form and spliced E3-E4, still linked to intron 4 and exon 5, could not be detected in this particular experiment. They could be detected, however, in other experiments (not shown). RNA 3 and RNA 4 show the typical kinetics of a lariat intermediate- and a lariat product RNA respectively, of a splicing reaction. The corresponding linear intermediate- and product RNAs of this splicing reaction were also detected (RNAs 7 and 6 in Figure 2A). These RNA molecules are derived from a splicing reaction using a cryptic 5' splice site in exon 4. This reaction will be described in detail elsewhere.

Finally, several RNA species (RNAs 9-12 in Figure 2A) were observed that appeared early in the time course, (lane 2) preceding the appearance of lariat RNAs. Although we did not attempt to identify these RNAs in detail several indications were obtained that these RNAs are the result of a 5' exonuclease activity in the nuclear extract. Specific blocks for degradation are presumably caused by splicing factors bound to precursor RNA (37). Interestingly, the most prominent of these RNA species, RNA 12, is only formed under splicing conditions. RNA 12 should correspond to a protected RNA fragment containing exon 5 and approximately 50 nucleotides of intron 4 encompassing the intron 4 3' splice site and branch point region. After incubation under splicing conditions of Ad2 ML pre-mRNA and β-globin pre-mRNA, comparable intron fragments were found to be resistent to degradation in RNAse protection experiments (37, 38-40). This observation suggests that rapid complex formation occurs on the intron 4 3' splice site. Protected fragments corresponding to putative blocks to this exonuclease activity at the intron 3 5' splice site, the intron 3 3' splice site and the cryptic 5' splice site in exon 4 were also detected (RNAs 9-11), albeit at much lower concentrations.

We conclude from the time course and debranching experiments that the 818 nucleotides long intron separating exon 3 from exon 5 is efficiently spliced *in vitro* in nuclear extracts of Hela cells. *In vitro* splicing of the much smaller sized (142 nucleotides long) intron 3 on the other hand is very inefficient.

*Localization of the branch point in the splicing reaction of exon 3 to exon 5*

To localize the branch point in intron 4 used in the splicing of exon 3 to exon 5, precursor RNA dI3I4-E was spliced in Hela nuclear extracts on a preparative scale. The two lariat RNAs generated in this splicing reaction (RNAs 1 and 2 in Figure 2A) were isolated. Primer I4, a synthetic 18-mer, was

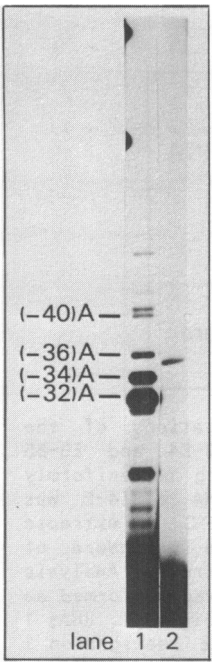



Figure 3. Determination of the nucleotide at the branch in the splicing reaction of exon 3 to exon 5. The 18 nucleotides long synthetic primer I4 (see below) was used to prime reverse transcription on gel purified RNA 2, the product lariat of the E3-E5 splicing reaction (lane 2). The position of the adenosine residues in the intron 4 sequence in the vicinity of the branch point was deduced from the specific stops occurring when the same I4 primer was used to prime reverse transcription on gel purified precursor RNA dI3I4-E in the presence of ddTTP (lane 1). The complete nucleotide sequence from the adenosine residue at -40 to the AG dinucleotide at the 3' splice site in intron 4 is:

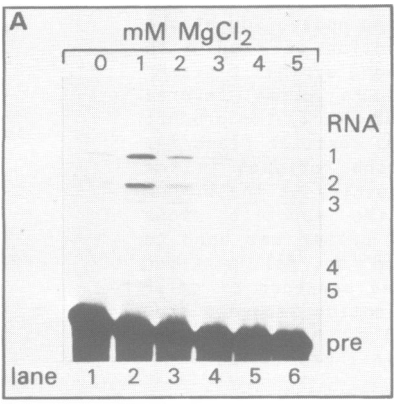
5' ACTCA*CAGATCTTCTCTTCTTTCTCCATCCTGCAAATCAG 3'.

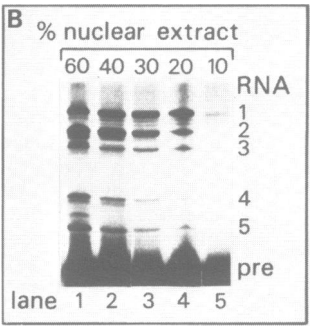
The nucleotide at the branch is indicated with an asterix in this sequence. The intron 4 sequence complementary to the I4 primer is underlined. Extra bands are visible in lane 1 which do not correspond to stops at adenosine residues in the intron 4 sequence. They are caused by contamination of the I4 primer with an intron 3 specific primer (I3). The presence of primer I3 did not affect the branch point determination.

hybridized to the 3' end of the free intron lariat (RNA 2) and extended by transcription by reverse transcriptase. It is known that primer extension on branched RNA systematically stops at the residue preceding the branch point (35). A strong stop for reverse transcription was observed (Figure 3, lane 2). In lane 1 the I4 primer was extended on precursor RNA dI3I4-E in the presence of ddTTP to identify the adenosine residues in intron 4. Reverse transcriptase terminates the elongation reaction on RNA 2 at the nucleotide preceding the adenosine residue at position -36 (lane 2). This localizes the branch point to the adenosine residue 36 nucleotides upstream of the intron 4 3' splice site. In an independent experiment we determined the nucleotide at the branch of the lariat intermediate RNA (RNA 1), which still contains exon 5, using an exon 5 primer for reverse transcription. This experiment confirmed the branch point at A -36 in the sequence CUCAC (not shown).

Effect of variations of the reaction conditions on Calcitonin-mRNA specific E3-E4 and CGRP-mRNA specific E3-E5 splicing

Several variations in splicing conditions were studied to address the question whether the predominance of splicing of exon 3 to exon 5 over splicing of exon 3 to exon 4 could be influenced. The ratio of the two lariat



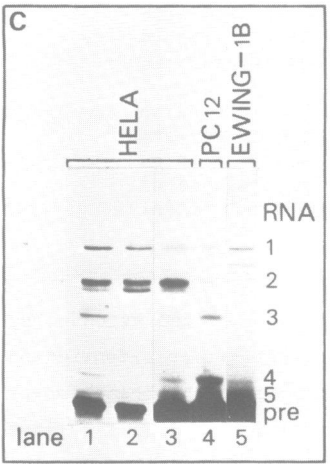




Figure 4. Effect of variations of the reaction conditions on E3-E4 and E3-E5 splicing. In vitro splicing of uniformly [$^{32}$P]-labeled precursor RNA dI3I4-E was done for two hours at 30°C in extracts prepared according to the procedure of Dignam unless indicated otherwise. Analysis of the splicing reactions was performed as described in the legend to Figure 2. RNAs 1 to 5 correspond to the RNAs designated 1 to 5 in Figure 2A. The precursor RNA dI3I4-E is also indicated (pre). (A) In vitro splicing in a Hela nuclear extract with varying amounts of exogenous $MgCl_2$ added to the reaction. The final concentration of $Mg^{2+}$ (mM) in the reaction is indicated above the lanes (lane 1-6). (B) In vitro splicing in a Hela nuclear extract using several dilutions of this extract in Dignam buffer D. The amount of undiluted extract added to the reactions was so varied from 60% to 10% of the total reaction volume (lanes 1-5). (C) In vitro splicing in nuclear extracts prepared from two independently maintained Hela cell lines (lanes 1 and 2), the rat PC12 cell line (lane 4) and the human Ewing-IB cell line (lane 5). These extracts were all prepared by extraction of the nuclei with 0.4 M NaCl (25). In vitro splicing in a Hela nuclear extract prepared by extraction of the nuclei with 0.6 M KCl (26) is also shown (lane 3). Different exposure times of the same autoradiograph are shown in lanes 1-5 because the efficiency of the splicing reactions differed between the different extracts.

intermediate RNAs, RNA 1 and RNA 5 for E3-E5 and E3-E4 splicing respectively, was used to determine the ratio of E3-E5 to E3-E4 splicing. Firstly, RNAs 1 and 5 are diagnostic for splicing of exon 3 to exon 5 and exon 4 respectively. Secondly, they are easily monitored because they can be separated from each other, from precursor RNA and RNA breakdown products. Finally, when uniformly labeled precursor RNA is used the relative intensities of the two

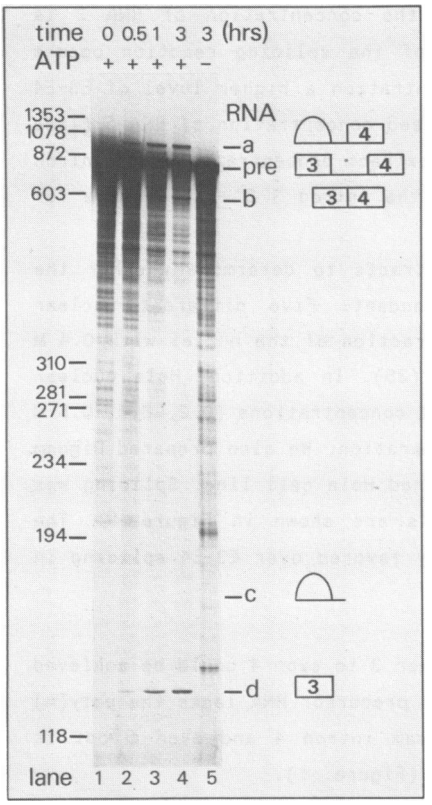



Figure 5. In vitro splicing of uniformly [$^{32}$P]-labeled precursor RNA dI3I4-B in Hela nuclear extracts for the times indicated above the lanes (in hours) in the presence (+) (lanes 1-4) or absence (-) (lane 5) of ATP and creatine phosphate. The position of [$^{32}$P] 5'end-labeled HaeIII restriction fragments of øx174 RF DNA is indicated. RNAs labeled a to d correspond to the RNAs a to d mentioned in the text. A schematic representation of their structure is shown. Open, numbered boxes indicate CALC-I exon 3 and 4 respectively. The position of the precursor RNA in the gel is also indicated (pre).

signals can be directly compared as a consequence of their uniform length and sequence.

Variation of the free $Mg^{2+}$ concentration in the reaction from 0-5 mM (Figure 4A) did not result in a significant change of the ratio of E3-E5 to E3-E4 splicing. Both E3-E5 and E3-E4 splicing proceed optimally between 1-2 mM $Mg^{2+}$ (lanes 2-3) and with higher levels of $Mg^{2+}$ both reactions gradually extinguish (lanes 4-6).

The concentration of nuclear extract in the reaction did affect the ratio of the two reactions (Figure 4B). At a nuclear extract concentration of 10% only the E3-E5 splice is detected (lane 5). The level of E3-E5 splicing is drastically increased when the nuclear extract concentration in the reaction is increased to 20% (lane 4). The E3-E4 splice can now be detected as well. Further increase of the nuclear extract concentration to 40% shows that the level of E3-E5 splicing does not change significantly whereas E3-E4 splicing is slightly increased (lanes 2-3). At 60% nuclear extract the level of RNA 1

does not further increase. Concomitantly, the concentration of RNA 2 is increased indicating that the second step of the splicing reaction occurs more efficiently. Also at 60% extract concentration a higher level of E3-E4 splicing is obtained reflected by the increased concentration of RNA 5 (lane 1). This observation shows that at limiting extract concentrations the intron 4 3' splice site efficiently competes with the intron 3 3' splice site for splicing factors.

We finally tested different nuclear extracts to determine whether the preferred E3-E5 splicing was extract dependent. Five different nuclear extracts of Hela cells were prepared by extraction of the nuclei with 0.4 M NaCl according to the procedure of Dignam (25). In addition, Hela nuclear extracts were prepared employing various salt concentrations (0.2, 0.4, 0.6 M NaCl, 0.6 M KCl) during nuclear extract preparation. We also prepared Dignam nuclear extracts of an independently maintained Hela cell line. Splicing was assayed as described above. Typical results are shown in Figure 4C. The results show that E3-E5 splicing is strongly favored over E3-E4 splicing in all nuclear extracts tested sofar.

In vitro splicing of precursor RNA dI3I4-B

To determine whether efficient splicing of exon 3 to exon 4 could be achieved in vitro we used precursor RNA dI3I4-B. This precursor RNA lacks the poly(A) addition signal of exon 4 and the downstream intron 4 and exon 5 but is otherwise identical to precursor RNA dI3I4-E (Figure. 1).

In a time course experiment with uniformly [$^{32}$P]-labeled dI3I4-B RNA (Figure 5) both intermediates of the exon 3 to exon 4 splicing reaction appear after a lag period of 30 minutes: a linear RNA (RNA d) with the expected size of free exon 3 and a lariat RNA (RNA a) with a slightly lower mobility than precursor RNA which is indicative for a small loop-large tail structure in this gel system (see above, Figure 2A). The products of the splicing reaction, RNA b (exon 3-exon 4) and RNA c (free lariat intron 3 ) accumulate after three hours of incubation. Similar results were obtained (data not shown) with several other substrate RNAs containing exon 3 and different truncated portions of exon 4 and a substrate RNA containing a deletion of the central portion of exon 4. The efficiency of E3-E4 splicing was not influenced either by the presence of the exon 4 poly (A) signal and the 5' portion of intron 4. From these experiments we conclude that exon 4 sequences do not play an important role in splicing of exon 3 to exon 4.

The results show that coupling between exon 3 and exon 4 is a poor reaction in vitro, which takes place more efficiently in Hela nuclear extracts using a RNA substrate lacking the 3' splice site of exon 5.

## DISCUSSION

### CGRP specific splicing in vitro

Using a model precursor RNA (dI3I4-E) and nuclear extracts from different cell lines we have investigated the alternative RNA processing reactions which occur during the expression of the human CALC-I gene. Splicing of this precursor RNA was observed in different Hela nuclear extracts, a PC12 and an Ewing-1B nuclear extract. CGRP mRNA specific splicing of exon 3 to exon 5 was the predominant reaction in all extracts tested. Only a small amount of CT mRNA-specific splicing of exon 3 to exon 4 could be detected. The amount of CT mRNA-specific splicing increased at higher nuclear extract concentrations but never approximated CGRP mRNA-specific splicing. Variation of other reaction parameters (i.e.: temperature, monovalent cation concentration, preincubation of the extracts) did not change significantly the ratio of CGRP to CT-mRNA specific splicing.

The branch point used in CGRP mRNA splicing was mapped to an adenosine residue, 36 nucleotides upstream of the 3' splice site. The flanking sequence CACUCAC resembles the highly conserved yeast branch point sequence UACUAAC (41) where branching occurs on the underlined A residue. In mammalian pre-mRNA introns there is only a loosely defined branch point consensus sequence PyXPyUPuAPy (42-44). The overall sequence of the intron 4 branch point fits well to this consensus sequence. It is striking, however, that the purine residue which precedes the branch point in most introns is not conserved in the intron 4 branch point sequence. Two other typical features of branch points are also displayed by the intron 4 branch point: its position, which is in agreement with the observed constraint on the distance of branch points relative to the AG dinucleotide of the 3' splice site (44,45), and the fact that branching occurs on an adenosine residue, which is a common feature of the authentic branch points mapped sofar (with the exception of the first intron of the human growth hormone gene, (46))

### Selection of 3' splice sites in CALC-I RNA dI3I4-E

We conclude from the results presented in this paper that there is a strong selection during spliceosome formation (47,48) for the intron 4 3' splice site in precursor RNA which contains both the intron 3 3' splice site and the intron 4 3' splice site. The prevalence for the intron 4 3' splice site was not affected by the variation of reaction parameters or by the usage of nuclear extracts of different cell origins. The ratio of CGRP to CT mRNA-specific splicing seems to be determined by the intrinsic properties of the intron 3 and intron 4 3' splice sites. This conclusion is at variance with the notion that a neuron specific splice commitment factor is necessary for

splicing of exon 3 to exon 5 (19,49). More complex explanations, such as the possibility that (a) factor(s) which are involved in the inhibition of CGRP mRNA-specific splicing or the regulation of the poly(A) reaction at exon 4 are lost during the preparation of nuclear extracts, could also account for the observed reaction patterns. The deletions introduced in intron 3 and intron 4 do affect the efficiency of splicing but not the ratio of Calcitonin- to CGRP-I mRNA specific splicing (R.A.L. Bovenberg, in preparation). The effect of the absence of exon 1, exon 2 and exon 6 sequences on Calcitonin/-CGRP mRNA-specific splicing in vitro is not clear yet.

Splicing of exon 3 to exon 4 occurs more efficiently at higher nuclear extract concentrations and occurs when intron 4 and exon 5 are artificially removed from the precursor RNA. These observations further support the idea that the intron 3 3' splice site is not easily recognized in vitro. The sequence of the intron 3 3' splice site GUGUUUUCCCUGCAG, is in agreement with the consensus sequence for 3' splice sites (33) although the tract of pyrimidine residues is rather short, especially when compared with the tract of pyrimidine residues in the intron 4 3' splice site: ACAGAUCUUCUCUUCUUUCUC-CAUCCUGCAAAUCAG (the underlined A denotes the branch point). The significance of this observation is not clear at present. It is interesting to note however that one of the earliest steps in 3' splice site complex formation seems to be the recognition of the branch point and the polypyrimidine tract by protein factors (50-52). It is tempting to speculate that the length of the polypyrimidine tract influences the efficiency of binding of these factors and hence influences the efficiency of spliceosome formation.

Furthermore, the notion that splicing of exon 3 to exon 4 is poor in vitro is further substantiated by the observation that the major branch point used in this reaction is an uracil residue (G.J. Adema, in preparation). Branching on uracil is unusual and it is known that branching on uracil lowers the efficiency of the in vitro splicing reaction (45).

Model for alternative CALC-I RNA processing

The following model for alternative processing of CALC-I pre-mRNA in vitro can be envisaged from the results presented in this paper. In precursor RNA containing exon 3- intron 3- exon 4- intron 4- exon 5 both the intron 3 and intron 4 3' splice site can be recognized by splicing factors albeit with different efficiencies: the intron 4 3' splice site is a high-affinity splice site, the intron 3 3' splice site is a low-affinity splice site. The high affinity intron 4 3' splice site competes efficiently, but not entirely for splicing with the low-affinity intron 3 3' splice site which results in a

high ratio of CGRP mRNA-specific versus CT mRNA-specific splicing. To change this ratio in favor of CT mRNA specific splicing it is necessary to inactivate the intron 4 3' splice site. This could be achieved most definitively by the poly(A) reaction at exon 4 which removes the intron 4 3' splice site from the transcript.

A comparable splicing/poly(A) cleavage competition model has been proposed for the switch in immunoglobulin μ heavy chain RNA processing from membrane bound (μm) to secreted form (μs) mRNA during B-cell development (53).

ACKNOWLEDGEMENTS

We like to acknowledge Dr. C.J.M. Lips (Department of Internal Medicine, University Hospital, Utrecht, NL) for stimulating discussions, Dr. P.T.G. Sillekens (University of Nijmegen, NL) for helpful advice on initial in vitro splicing experiments, W.P.M. van de Meerendonk and B. Machielse for help in initial experiments and discussions, Mrs. T.C. Moen for technical assistance and R. Schiphof, Mrs. M.H. Schmidt-Michels and Mrs. A. Koonen for culturing of different cell lines.

The research was in part supported by the Netherlands Organization for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

*To whom correspondence should be addressed

REFERENCES

1. Rosenfeld, M.G., Mermod, J.J., Amara, S.G., Swanson, L.W., Sawchenko, P.E., River, J., Vale, W.W. and Evans, R.M. (1983) Nature, 304, 129-135
2. Leff, S.E. and Rosenfeld, M.G. (1986) Ann. Rev. Biochem., 55, 1091-1117.
3. Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., and Evans, R.M. (1982) Nature 298, 240-244.
4. Sabate, M.I., Stolarsky, L.S., Polak, J.M., Bloom, S.R., Varndell, I.M., Ghatei, M.A., Evans, R.M. and Rosenfeld, M.G. (1985) J. Biol. Chem., 260, 2589-2592.
5. Amara, S.G., Evans, R.M. and Rosenfeld, M.G. (1984) Mol. Cell. Biol. 4, 2151-2160.
6. Craig, R.K., Hall, L., Edbrooke, M.R., Allison, J. and MacIntyre, I. (1982) Nature, 295, 345-347.
7. Edbrooke, M.R., Parker, D., McVey, J.H., Riley, J.H., Sorenson, G.D., Pettengill, O.S. and Craig, R.K. (1985) EMBO J., 4, 715-724.
8. Jonas, V., Lin, C.R., Kawashima, E., Semon, D., Swanson, L.W., Mermod, J.J., Evans, R.M. and Rosenfeld, M.G. (1985) Proc. Natl. Acad. Sci. USA, 82, 1994-1998.
9. LeMoullec, J.M., Jullienne, A., Chenais, J., Lasmoles, F., Guliana, J.M., Milhaud, G. and Moukhtar, M.S. (1984) FEBS Lett., 167, 93-97.
10. Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., van de Ven, W.J.M., Jansz, H.S. and Lips, C.J.M. (1984) J. Clin. Endocrinol. Metab., 59, 358-360.
11. Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., Cremers, A.F.M., Jansz.

H.S. and Lips, C.J.M. (1985) In Pecile,A (ed), Proc. Int. Symp. Calcitonin, Elsevier Sci. Publ. BV, Amsterdam, N.Y. 23-31.
12. Steenbergh, P.H., Höppener, J.W.M., Zandberg, J., Visser, A., Lips, C.J.M. and Jansz, H.S. (1986) FEBS Lett., 209, 97-103.
13. Morris, H.R., Panico, M., Etienne, T., Tippins, J., Girgis, S.I. and MacIntyre, I. (1984) Nature, 308, 746-748.
14. Petermann, J.B., Born, W., Chang, J.Y and Fischer, J.A. (1987) J. Biol. Chem., 262, 542-545.
15. Tschopp, F.A., Henke, H., Petermann, J.B., Tobler, P.H., Janszer, R., Hokpelt, T., Lundberg, J.M., Cuello, C. and Fischer, J.A. (1985) Proc. Natl. Acad. Sci. USA, 82, 248-251.
16. Bovenberg, R.A.L., Van de Meerendonk, W.P.M., Baas, P.D., Steenbergh, P., Lips, C.J.M. and Jansz, H.S. (1986) Nucl. Acids Res., 14, 8785-8803.
17. Maniatis, T. and Reed, R. (1987) Nature, 325, 673-678.
18. Sharp, P.A. (1987) Science, 235, 766-771.
19. Leff, S.E., Evans, R.M. and Rosenfeld, M.G. (1987) Cell, 48, 517-524.
20. Höppener, J.W.M., Steenbergh, P.H., Slebos, R.J.C., Visser, A., Lips, C.J.M., Jansz, H.S., Bechet, J.M., Lenoir, G.M., Born, W., HallerBrem, S., Petermann, J.B. and Fischer, J.A. (1987) J. Clin. Endocrinol. Metab., 64, 809-817.
21. Gey, G.O., Coffman, W.D., and Kubicek, M.T. (1952) Cancer Res., 12, 264.
22. Strauss, M., Streuli, C.H. and Griffin, B.E. (1986) Gene, 49, 331-340.
23. Konarska, M.M., Padgett, R.A. and Sharp, P.A. (1984) Cell, 38, 731-736.
24. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol., 65, 499-560.
25. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucl. Acids Res., 11, 1475-1489.
26. Heintz, N. and Roeder, R.G. (1984) Proc. Natl. Acad. Sci. USA, 81, 2713--2717.
27. Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984) Cell, 36, 993-1005.
28. Ruskin, B. and Green, M.R. (1985) Science, 229, 135-140.
29. Solnick, D. (1985) Cell, 43, 667-676.
30. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Nucl. Acids Res., 12, 7035-7056.
31. Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984) Cell, 37, 415-427.
32. Kedes, D.H. and Steitz, J.A. (1987) Proc. Natl. Acad. Sci. USA, 84, 7928-7932.
33. Mount, S.M. (1982) Nucl. Acids Res., 10, 459-472.
34. Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984) Science, 225, 898-903.
35. Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984) Cell, 38, 317-331.
36. McLauchlan, J., Gaffney, D., Whitton, J.L. and Clements, J.B. (1985) Nucl. Acids Res., 13, 1347-1368.
37. Ruskin, B. and Green, M.R. (1985) Cell, 43, 131-142.
38. Schmitt, P., Gattoni, R., Keohavong, P. and Stevenin, K. (1987) Cell, 50, 31-39.
39. Chabot, B. and Steitz, J.A. (1987) Mol. Cell. Biol., 7, 281-293.
40. Krämer, A. (1987) J. Mol. Biol., 196, 559-573.
41. Langford, C.J., Klin, F.J., Donath, C. and Gallwitz, C. (1984) Cell, 36, 645-653.
42. Keller, E.B. and Noon, W.A. (1984) Proc. Natl. Acad. Sci. USA, 81, 7417--7420.
43. Reed, R. and Maniatis, T. (1985) Cell, 41, 95-105.
44. Ruskin, B., Greene, J.M., Green, M.R. (1985) Cell, 41, 833-844.

45. Hornig, H., Aebi, M. and Weismann, C. (1986) Nature, 324, 589-591.
46. Hartmuth, K. and Barta, A. (1988) Mol. Cell. Biol., 8, 2011-2020.
47. Frendewey, D. and Keller, W. (1985) Cell, 42, 355-367.
48. Grabowski, P.J., Seiler, S.R. and Sharp, P.A. (1985) Cell, 42, 345-353.
49. Crenshaw III, E.B., Russo, A.F., Swanson, L.W. and Rosenfeld, M.G. (1987) Cell, 49, 389-398.
50. Gerke, V. and Steitz, J.A. (1986) Cell, 47, 973-984.
51. Ruskin, B., Zamore, P.D. and Green, M.R. (1988) Cell, 52, 207-219.
52. Tazi, J., Alibert, C., Temsamani, J. Reveillaud, I. Cathala, G., Brunei, C. and Jeanteur, P. (1986) Cell, 47, 755-766.
53. Peterson, M.L. and Perry, R.P. (1986) Proc. Natl. Acad. Sci. USA, 83, 8883-8887.