Proc. Natl. Acad. Sci. USA
Vol. 78, No. 1, pp. 133–137, January 1981
Biochemistry

Expression of simian virus 40–rat preproinsulin recombinants in monkey kidney cells: Use of preproinsulin RNA processing signals

(RNA splicing/polyadenylylation/gene regulation/proinsulin production)

PETER GRUSS AND GEORGE KHOURY

Laboratory of Molecular Virology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Communicated by Robert M. Chanock, September 17, 1980

ABSTRACT **The complete rat preproinsulin gene *I* was cloned into a simian virus 40 (SV40) vector. Most of the late region of the viral vector, including the SV40 intervening sequences (introns) and all of the major splice junctions, was deleted and replaced by the entire rat insulin gene. The recombinant molecules and a temperature-sensitive helper virus (*tsA28*) were inoculated into monkey kidney cultures. The formation of stable transcripts of the insulin insert was as efficient as the production of late SV40 mRNA. Analysis of these transcripts indicated that the rat preproinsulin gene nucleotide signals involved in RNA splicing and poly(A) addition were used. Examination of the 5′ ends of the mRNAs showed several classes, one of which was the same size as the authentic rat insulinoma mRNA. This suggests that a portion of the transcripts may be initiated or processed faithfully, or both, at their 5′ ends within rat insulin sequences. Significant quantities of a protein identified as rat proinsulin were synthesized. Detection of most of the proinsulin in the tissue culture medium suggests that this protein was secreted.**

Numerous genes have been cloned (see ref. 1), and a comparative analysis of their DNA sequences has helped in deciphering certain nucleotide signals that may be involved in certain steps of gene regulation (2–8). However, characterization of these processing signals requires introduction of the cloned genes into eukaryotic cells. Several approaches have been elaborated to reinsert DNA or RNA into eukaryotic cells. A eukaryotic viral vector [simian virus 40 (SV40)] that is covalently linked to the gene of interest and infected into permissive or nonpermissive cells (9–15) provides an efficient method. Although in this approach the size of the inserted gene is limited, SV40 recombinants offer a tremendous advantage in terms of efficiency of introduction of DNA. Furthermore, in permissive cells there is a significant amplification of transcriptional templates through viral DNA replication. Data obtained from experiments with SV40–eukaryotic recombinant molecules suggest that certain regulatory sequences of the introduced gene, such as the translation initiator AUG, the polyadenylylation signal, and the splice sites, can function normally in the viral genetic environment (13, 14). In this study, we have investigated the regulatory signals for transcription and translation of a rat preproinsulin gene. One of two nonallelic rat insulin genes (rI_1) harbors a single 119-base-pair (bp) intron in the 5′ noncoding region (2, 3). This gene has been inserted into the late region of a SV40 vector, in the orientation of the late viral genes. The results of infection by the recombinant with helper virus are presented.

MATERIALS AND METHODS

Virus Strains and DNA or RNA Preparation. SV40, strain 776, was used for the construction of recombinants. λ Charon 4A–rat preproinsulin *I* (rI_1) recombinant phage were grown and purified (16), and DNA was isolated as described (17). African green monkey kidney (AGMK) cells infected with recombinant virus and *tsA28* (18) as helper virus were maintained for 60–72 hr at 40.5°C, and superhelical DNA was prepared (19). Superhelical DNA was separated from relaxed circles and from small DNA fragments by centrifugation in CsCl/ethidium bromide gradient. The cytoplasmic RNA was harvested 48 hr after infection (20) with 10–20 plaque-forming units of a stock containing *tsA28* and SVL_1-rI_1 recombinant virus.

Nuclease S1 and Exonuclease VII Treatment. The reactions with nuclease S1 and exonuclease VII were performed essentially as described by Berk and Sharp (21, 22). Hybrid products were analyzed by gel electrophoresis as described (23, 24).

Neutral agarose gel electrophoresis was conducted in 1.4% (wt/vol) agarose gels as described (24). Mapping of SVL_1-rI_1 DNA was performed using 4% (wt/vol) polyacrylamide gels (acrylamide:bisacrylamide = 20:1) in 0.04 M Tris, pH 7.8/0.02 M sodium acetate/0.001 M EDTA. For electrophoresis, 7.5% polyacrylamide gels (acrylamide:bisacrylamide = 20:1) were prepared and used as described (25).

Construction and Cloning of SV40-Preproinsulin Recombinants. DNA fragments were purified by gel electrophoresis and DEAE-cellulose chromatography (26), coprecipitated with a similarly purified *Hpa* II/*Bam*HI SV40 A fragment, and resuspended in 20 μl of ligation buffer. Ligation was performed essentially as described by Maniatis *et al.* (17). The ligated products were used directly for an infectious center assay (27) with helper virus SV40 *tsA28*.

Protein Analysis. AGMK cells were infected with SVL_1-rI_1 recombinant virus (5–10 plaque-forming units per cell). Cells were washed and labeled 40 hr after infection in medium containing 200 μCi of L-[^{35}S]cysteine per ml (465 Ci/mmol; Ci = 3.7×10^{10} becquerels) for 4 hr at 41°C. Lysis was performed by using 2 ml per 150-cm^2 flask of Tris-buffered saline (pH 7.6) containing 1% Nonidet P-40, 2 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol, and 100 μg of *N*-tosylphenylalanine chloromethyl ketone per ml. After immunoprecipitation (28) with anti-bovine insulin antiserum, analysis in NaDodSO$_4$/urea/polyacrylamide gels (29) was followed by fluorography (30). Quantitative radioimmunoassays were performed by Hazelton Laboratories, Vienna, VA.

RESULTS

Construction of an SV40-Rat Preproinsulin Gene *I* Recombinant Molecule. In our studies we used the rat preproinsulin gene *I*, designated rI_1, which is approximately 0.56 kilobase (kb) in length and contains a single 119-bp intervening sequence (intron) prior to the coding sequence (provided by P. Lomedico, A.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: AGMK, African green monkey kidney cells; kb, kilobase(s); SV40, simian virus 40; bp, base pair(s).

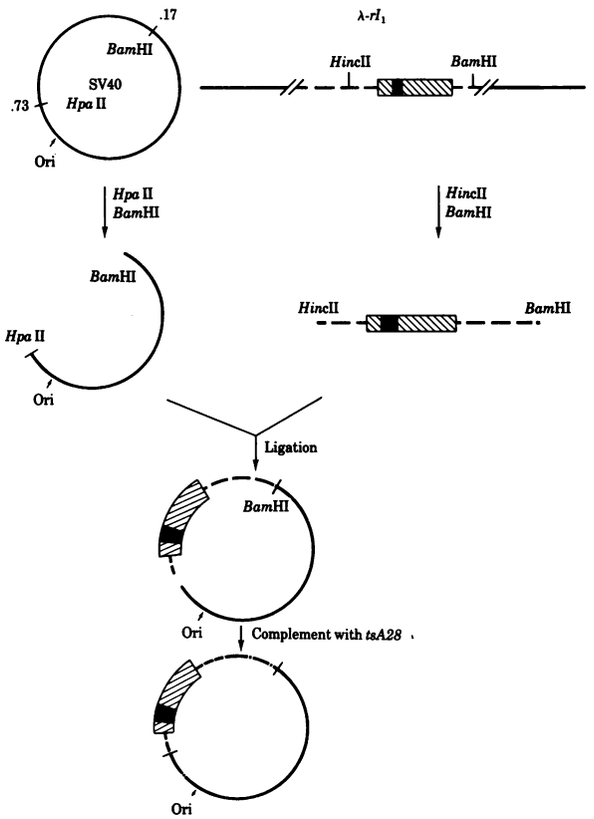

FIG. 1. Schematic representation of the construction of SV40-rat preproinsulin gene *I* (SVL_1-rI_1) recombinants. ———, SV40 sequences; -------, rI_1, flanking sequences. The diagonally striped rectangular box indicates the location of the rI_1 gene. The black square inside the box symbolizes the intron. Ori, functional replication origin.

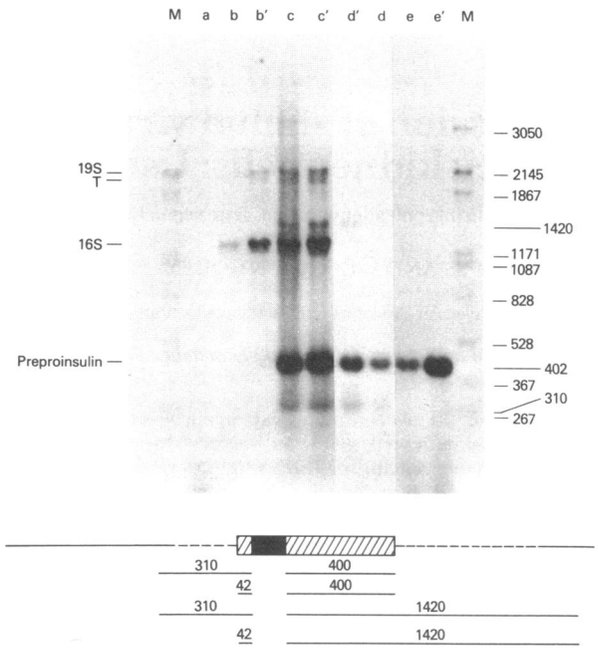

FIG. 2. Nuclease S1 alkaline agarose gel analysis of RNA from SVL_1-rI_1 infection. Purified poly(A)-containing cytoplasmic RNA from SV40- and SVL_1-rI_1-infected cells and polyadenylylated RNA from rat insulinoma cells (see refs. 30 and 33) were used. In all cases, letters without a prime represent the use of RNA from approximately 2×10^6 cells and letters with prime indicate the use of RNA from 5×10^6 cells. SV40 RNA (b and b') and RNA from SVL_1-rI_1-infected cells (containing *tsA28* helper) (c and c') were hybridized to *Taq* I-linearized ^{32}P-labeled *tsA28* and SVL_1-rI_1 DNA. The latter RNA was hybridized to *Taq* I-cleaved, gel purified ^{32}P-labeled SVL_1-rI_1 DNA probe clone (d and d'). This probe was also used for analysis of rat insulinoma RNA (e and e'). Hybridization mixtures were treated with nuclease S1 and subjected to electrophoresis. Tracks a and M: a, the probe contains no added RNA; M, the SV40 restriction fragment markers (with their sizes in nucleotides shown on the right-hand side). The diagrams below the figure represent an interpretation of the RNA mapping data.

Efstratiadis, and W. Gilbert). The entire gene, including the putative regulatory signals at its 5' and 3' ends, is present in the 2.1-kb insert of a clone in λ Charon 4A (λ-rI_1). The DNA of this clone was cleaved with the restriction enzymes *Hin*cII and *Bam*HI (Fig. 1), generating a 1.62-kb segment containing the entire rI_1 genomic sequence, including chromosomal portions flanking the gene. The SV40 *Hpa* II/*Bam*HI A fragment (3138 bp) containing an active early region, a functional replication origin, and the sequences of the extreme 5' and 3' ends of 16S and 19S late viral mRNAs was purified. The gene rI_1, contained within the *Hin*cII/*Bam*HI fragment, was ligated to the SV40 *Hpa* II/*Bam*HI vector. *Bam*HI provides the only cohesive end present in both fragments; thus, ligation of rI_1 can occur only in the sense direction relative to the late viral promotors.

The final product, a linear recombinant DNA molecule (Fig. 1) was inoculated into secondary AGMK cells in an infectious center assay (27). The late functions deleted from the recombinant molecules were complemented by the early SV40 mutant *tsA28*. Although the ligated recombinant DNA molecules were linear, they could be cyclized (presumably by the action of cellular enzymes) prior to encapsidation (31). The presence of insulin-specific sequences in plaques was assayed by hybridization of the ^{32}P-labeled recombinant DNA molecules to DNA blots (32) containing immobilized insulin cDNA (provided by A. Efstratiadis). One of many recombinants with the desired structure (SVL_1-rI_1) was mapped (data not shown). Because the *Hha* I site in SV40 is preserved, no more than two bases were deleted from SV40 during cyclization. Based on the size of the restriction enzyme fragments, very few nucleotides of the chromosomal sequences adjacent to the rI_1 gene at its 5' end were removed. The physical map of the rI_1 gene is similar to the map as published (2, 3).

Mapping Cytoplasmic RNA from SVL_1-rI_1 Recombinants. Polyadenylylated cytoplasmic RNA was prepared 48 hr after infection of AGMK with a viral stock of SVL_1-rI_1 recombinants (containing *tsA28* as a helper virus). Hybrid molecules of RNA and ^{32}P-labeled DNA probes were digested with nuclease S1 followed by alkaline agarose gel electrophoresis (21, 22) to localize introns and to identify 5' and 3' ends of the mRNAs (Fig. 2). In track d and d', two different concentrations of RNA from SVL_1-rI_1-infected AGMK cells were hybridized to a ^{32}P-labeled DNA probe of SVL_1-rI_1 (cleaved at the *Taq* I site; 0.57 SV40 map unit).

One major and two minor bands were detected. The major band was 400 nucleotides in length and comigrated (Fig. 2, tracks e and e') with rat insulinoma RNA (34) analyzed under the same conditions. This band indicates that the entire coding region of preproinsulin was present in the mRNA made in AGMK cells. It also confirms the use of preproinsulin splicing signals and the use of a 3' mRNA end similar to that of authentic insulinoma mRNA.

A second band (310 nucleotides) was absent from the analysis of rat insulinoma RNA. The likely derivation of this fragment is from the 5' end of this particular class of mRNA. A segment of that size is predicted, if some transcripts of the recombinant use an SV40 promotor (5' end) and the preproinsulin gene splice

points. Thus, one class of hybrid RNA 5′ end is likely to be identical with a SV40 wild-type viral 5′ end. It should be pointed out that the molar quantity of the 310-nucleotide fragment was less than that of the 400-nucleotide fragment. One explanation for this disparity is a heterogeneity within the 5′ ends of insulin-specific RNA molecules. Indeed, below the fragment of 310 nucleotides, at least one more fragment was visible (approximately 260 nucleotides), which appeared to represent a 5′ end of another class of mRNA. Fragments smaller than the 260-nucleotide fragment would not be easily detectable in the gel system used in this experiment (but were examined in another experiment described below).

A third segment, 1420 nucleotides in length (Fig. 2, tracks d and d′), was also absent from rat insulinoma mRNA. The size indicates that this fragment extends from the preproinsulin splice site to the SV40 polyadenylylation site (normally used for 16S and 19S mRNA) rather than to the preproinsulin transcription termination sites. A comparative analysis of sequences preceding the 3′ ends of eukaryotic mRNA (35) indicates that a specific nucleotide sequence (A-A-U-A-A-A) precedes the poly(A) site by 15 to 20 nucleotides. These studies, data obtained with SV40-globin recombinants (36), and SV40 mutants with duplicated signals for polyadenylylation (M. Fitzgerald and T. Shenk, personal communication) show that a certain fraction of molecules are not polyadenylylated at the proximal signal.

In a control experiment (Fig. 2, tracks b and b′) two concentrations of wild-type SV40 RNA were hybridized to a ^{32}P-labeled combined probe containing both *tsA28* helper DNA and SVL_1-rI_1 DNA. These lanes identify bands corresponding to the SV40 large tumor-antigen mRNA, 19S, and 16S late mRNAs, which all migrated differently from the above-mentioned preproinsulin recombinant RNA (Fig. 2, tracks d and d′). This was further clarified by using the combined probe (*tsA28* DNA + SVL_1-rI_1) for nuclease S1 analysis of RNA of SVL_1-rI_1-infected cells (containing *tsA28* helper RNA). Tracks c and c^r (which in effect combine the data from Fig. 2, tracks b and b′ and tracks d and d′) show that the bands derived from SV40-specific helper RNAs were distinct from those related to the recombinant molecules. These tracks furthermore demonstrate that the amount of stable SVL_1-rI_1 mRNA was equal to or greater than the amount of late *tsA28* helper mRNA based on a *tsA28* DNA/SVL_1-rI_1 DNA infection ratio of 1:1.

To confirm and extend the data from the alkaline gel concerning the location of 5′ ends of the transcripts, a nuclease S1 analysis was performed by using a probe (*Hha* I fragment, 370 bp) that hybridized only to the 5′-terminal portion of the preproinsulin gene. This fragment did not extend into SV40 sequences. Nuclease S1-digested heteroduplexes were subsequently electrophoresed in a 7.5% polyacrylamide/urea gel under denaturing conditions. As size markers, a DNA fragment with a known sequence, treated according to the method of Maxam and Gilbert (25), was coelectrophoresed. This marker allowed the precise determination of the number of nucleotides protected by the *Hha* I probe. Since the 3′ end was fixed by the splice junction, variability in length must have been due to differences at the 5′ termini of RNA species. Thus, the length of the fragment determined the location of the 5′ ends.

In Fig. 3 *Right*, tracks b and c show the analysis of two different concentrations of RNA derived from cells infected with the recombinant virus SVL_1-rI_1. A number of bands representing different 5′ ends are readily discernible. In addition to the 310-nucleotide fragment, which contained a 5′ end in the SV40 sequence at 0.72 map unit (this fragment is not present in Fig. 3 because the probe did not extend into SV40 sequences) two abundant classes of RNA, 260 nucleotides and 240 nucleotides in length, were detected; these 5′ termini were located in rat

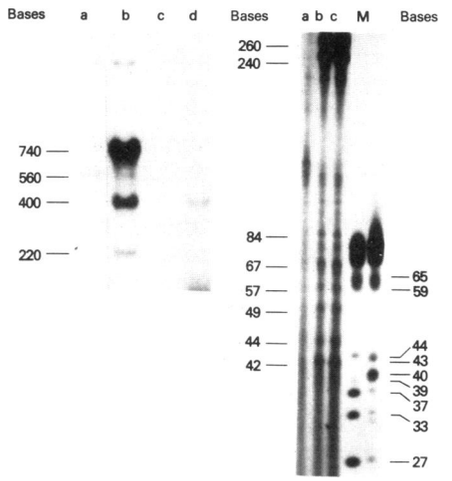

FIG. 3. Exonuclease VII analysis and 5′-end mapping with nuclease S1. (*Left*) Exonuclease VII analysis of purified poly(A)-containing cytoplasmic RNA from rat insulinoma cells (tracks a and d) hybridized with a *Hae* II/*Bam*HI fragment DNA probe representing the entire rat insulin gene in SVL_1-rI_1. Hybrids were treated with nuclease S1 as control (track d) or with exonuclease VII (track a). Poly(A)-containing cytoplasmic RNA from SVL_1-rI_1-infected AGMK was hybridized to the same probe, reacted with exonuclease VII, and coelectrophoresed with the samples from insulinoma cells in a 1.4% alkaline agarose gel (track b). The ^{32}P-labeled DNA probe (track c, control) was a restriction enzyme fragment (*Hae* II/*Bam*HI), 1560 nucleotides in length, with a specific activity of 2×10^6 cpm/μg. (*Right*) A nuclease S1 analysis with a *Hha* I fragment of 370 bp (specific activity, 3×10^6 cpm/μg) as a probe. The gel system was that of Maxam and Gilbert (25). Tracks: a, DNA control with no added RNA; b and c, two different concentrations of RNA from SVL_1-rI_1-infected AGMK cells [cytoplasmic poly(A)-containing RNA from 15×10^6 cells in track b and RNA from 30×10^6 cells in track c]; M, known DNA sequence giving size markers, with the fragment sizes in nucleotide numbers.

sequences. Additionally a series of smaller bands reflected heterogeneity of 5′ RNA ends within the rat insulin sequences. Among these bands, the most abundant was 42 nucleotides. The size of this fragment located the 5′ end at a position that corresponded to the putative 5′ terminus of authentic rat preproinsulin mRNA (2). An analysis of these data is presented in the legend to Fig. 3. To further demonstrate that one class of transcripts has a 5′ end at the same position as authentic rat preproinsulin RNA, we assayed the transcripts after hybridization to a *Hae* II/*Bam*HI probe by using exonuclease VII (Fig. 3 *Left*). Exonuclease VII digests single-stranded ends but unlike nuclease S1 does not cleave internal single-stranded loops. As opposed to a nuclease S1 analysis of authentic rat insulinoma RNA, a size increase (according to the size of the intron plus the 5′ terminal portion) of 162 nucleotides was expected.

The analysis of the transcripts in a 1.4% alkaline agarose gel is shown in Fig. 3 *Left*. In track d is seen the insulinoma RNA that was assayed with nuclease S1. As seen in Fig. 2, a single band 400 nucleotides in length appeared. Track a shows the same RNA analyzed with exonuclease VII. As expected, the size increased by approximately 160 nucleotides. RNA from cells infected with SVL_1-rI_1 analyzed by using exonuclease VII is shown in track b. The most prominent band is 740 nucleotides in length and would appear to represent RNA transcripts with 5′ ends located beyond the *Hae* II site of the probe. A second prominent band migrating at 400 nucleotides most likely represents the body segment resulting from contaminating endonucleases, inadequate hybridization in the 5′ exon, or nicks in the intron of the DNA probe. In this case, a major shortened leader segment of 220 nucleotides and the coding segment of 40 nucleotides is

expected. However, one class of RNA (560 nucleotides) comigrated with the RNA from insulinoma cells. This class appeared to reflect the population bearing 5′ ends similar to those of authentic preproinsulin mRNA. Hybridization of the probe DNA in the absence of RNA revealed no bands (track c). Taken together, the analyses of the SVL_1-rI_1 transcripts indicate that all four available control signals representing 5′ ends and 3′ poly(A) sites in both SV40 and rat insulin were being used, although with different frequencies. A significant heterogeneity of 5′ ends within the rat insulin RNA molecules was observed. The rat insulin splice site appeared to be used in all stable transcripts. Furthermore, one class of terminally processed mRNA was indistinguishable by this analysis from authentic rat insulinoma RNA.

Synthesis of Rat Proinsulin in AGMK Cells. Insulin is initially translated as a prehormone, preproinsulin, which undergoes two subsequent processing steps. A 24-amino acid, NH_2-terminal, hydrophobic leader peptide is involved in the transfer of the nascent chain through the microsomal membrane; it is then rapidly cleaved off before packaging of proinsulin into secretory granules (37–39). Conversion of proinsulin to insulin requires the specific removal of the central C peptide, leaving the A and B peptides joined by two disulfide bridges. To determine if the SV40-rat-preproinsulin mRNAs are translated in AGMK cells, a quantitative radioimmune assay was used. Protein extracts from 1×10^6 AGMK cells infected with SVL_1-rI_1 recombinant stock had a level of 53 to 70 microunits (1 unit $\approx$ 48 μg) of immunoreactive insulin. In a 4-hr period (40–44 hr after infection) 530 microunits of immunoreactive insulin accumulated in the tissue culture medium from this same number of cells. Polyacrylamide gel electrophoresis was used to determine the size of the insulin-specific protein(s).

Proteins labeled with [^{35}S]cysteine were extracted from SVL_1-rI_1 infected AGMK cells and were immunoprecipitated with guinea pig anti-bovine insulin antiserum and analyzed on 15% $NaDodSO_4$/urea/polyacrylamide gels (Fig. 4). The immunoprecipitable material (Fig. 4, track b) comigrated approximately with the unlabeled bovine proinsulin marker. This protein was not precipitated with control normal guinea pig antiserum (Fig. 4, track a). The addition of increasing amounts of unlabeled bovine proinsulin (Fig. 4, track c–f) and bovine insulin (not shown) competed, in proportion to the amount used, against the binding of the ^{35}S-labeled protein produced from SVL_1-rI_1 by the antibody. As a control, serum albumin (Fig. 4, track g) did not inhibit this specific binding of ^{35}S-labeled SVL_1-rI_1 protein by the insulin anti-serum. Subsequent immunoprecipitation analysis of the proteins present in the tissue culture medium revealed an insulin-reactive polypeptide which comigrated with the proinsulin marker (data not presented). These experiments suggest that translation and peptide processing of one or more of the SVL_1-rI_1 transcripts results in the production of rat proinsulin that may be secreted.

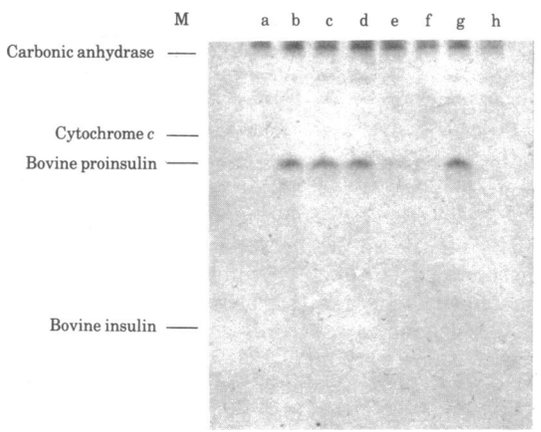

FIG. 4. Insulin-reactive protein analysis of SVL_1-rI_1 infected cells by immunoprecipitation and gel electrophoresis (29). Parallel cultures of secondary AGMK cells were infected with either SV40 as control or SVL_1-rI_1 recombinants. Proteins were labeled and harvested. Aliquots (from 1×10^5 cells each) were immunoprecipitated with 2-μl hamster anti-bovine insulin antiserum (a–g) or normal hamster serum as control (h). No protein band the size of proinsulin appeared in SV40-infected cells (a). Proteins extracted from SVL_1-rI_1-infected cells competed with 0 (b), 1.8 ng (c), 18 ng (d), 360 ng (e), or 3 μg (f) of unlabeled bovine proinsulin for binding to the insulin antiserum. As a control, 3 μg of bovine serum albumin (g) was used instead of bovine proinsulin to rule out nonspecific competition. Unlabeled proteins run in a parallel track (M) were used as size markers and visualized by staining the gel with Coomassie blue.

DISCUSSION

In this study, we have employed SV40 as a vector for introducing the rat insulin gene, rI_1 (2, 3) into AGMK cells. The exclusive accumulation of proinsulin suggests that the AGMK cells have enzymes capable of removing the hydrophobic leader peptide from rat preproinsulin but not the internal C peptide from proinsulin. In pancreatic β cells, the hydrophobic leader appears to be cleaved off rapidly, because a prehormone can hardly be detected *in vivo* (40, 41). The lack of further maturation from proinsulin to insulin in AGMK cells may be due to an absence of the specific proteases involved in the final processing step. In the tissue culture medium, significant amounts of protein immunoreactive with insulin antisera and comigrating with a proinsulin marker in $NaDodSO_4$/polyacrylamide gels could be attributed either to the secretion of insulin-reactive polypeptides or to the leakiness of cellular membranes caused by the infecting virus. Because the cell cultures at the time of analysis (44 hr after inoculation) are largely intact, the former possibility seems more likely.

Our primary goal in constructing SV40 recombinant molecules is in their application to the study of gene regulation. Three critical genetic components required for the generation of stable mRNAs include (*i*) the polyadenylylation site at the 3′ end of RNA molecules, (*ii*) the splicing site(s), and (*iii*) the promoter and capping sites at the 5′ ends of the transcript.

Within 10 to 20 nucleotides proximal to the 3′-terminal poly(A) tail of eukaryotic mRNA molecules, one usually finds the hexanucleotide sequence 5′-A-A-U-A-A-A-3′ (35). This sequence, thought to act as a polyadenylylation signal, is present in both the rat insulin segment of SVL_1-rI_1 and in the more distal SV40 segment at the normal late poly(A) site (about 0.17 SV40 map unit). Whether the polyadenylylation event is preceded by transcriptional termination or RNA processing is not known. Based on the finding of two sets of RNA molecules that appear to have 3′ ends near both of these signals, we presume that both polyadenylylation sites can be used. The marked difference in the abundance of these two sets of molecules (about 20:1) appears to reflect a preference for the proximal poly(A) site. A similar finding has been made with SV40–mouse β-globin recombinants (14, 15). Whether this relative abundance reflects the distance between the two poly(A) signals is unknown. All of the major late SV40 splice sites have been deleted from SVL_1-rI_1. The only remaining splice junction situated within rat insulin sequences appears to have been used in the formation of all of the stable recombinant RNA species. The relationship between splicing and the generation of certain stable eukaryotic mRNAs has been considered in detail (23, 24, 36, 42).

An analysis of the rat insulin-containing transcripts found in

AGMK cells indicates that several classes of 5′ ends as well as two 3′ ends are present among the stable, polyadenylylated, cytoplasmic RNAs. One of these 5′ ends maps within SV40 sequences at a site analogous to that found on late SV40 transcripts. Another set of 5′ ends appears to map within the rat genomic sequences, proximal to the coding sequences. Although we cannot distinguish between transcription initiation and processing in the generation of these 5′ ends, it should be pointed out that the size of one transcript that has its 5′ end in insulin gene sequences is similar to "rat insulinoma" RNA (see Figs. 2 and 3). Thus, it appears that the rat insulin "promoter" or 5′ processing signals, or both, are functioning in the recombinant molecules. Based on their studies with SV40 recombinant molecules, a similar conclusion has been reached by other investigators (Chi Nguyen-Huu, S. Clark, P. Berg, and H. Goodman; D. Hamer, M. Kaehler, and P. Leder, personal communications). If this is the case, we may be able to analyze the sequences that determine the formation of the 5′ ends of eukaryotic mRNAs.

We thank Drs. P. Lomedico, A. Efstratiadis, and W. Gilbert for generously providing the genomic clones of rat preproinsulin DNA; Dr. A. Gazdar for rat insulinoma cells; Dr. A. Notkins for anti-bovine insulin antiserum; M. König for expert technical assistance; Drs. G. Jay and N. Goldman for advice in protein analysis; and Drs. C.-J. Lai, R. Dhar, S. Segal, and J. Alwine for helpful comments.

1. Abelson, J. & Butz, D., eds. (1980) *Science* **209**, No. 4463.
2. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. & Tizard, R. (1979) *Cell* **18**, 545–558.
3. Cordell, B., Bell, G., Tischer, E., DeNoto, F. M., Ullrich, A., Pictet, R., Rutter, W. J. & Goodman, H. M. (1979) *Cell* **18**, 533–543.
4. Konkel, D. A., Tilghman, S. M. & Leder, P. (1978) *Cell* **15**, 1125–1312.
5. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
6. Seif, I., Khoury, G. & Dhar, R. (1979) *Nucleic Acids Res.* **6**, 3387–3398.
7. Dawid, I. B. & Wahli, W. (1979) *Dev. Biol.* **69**, 305–328.
8. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–142.
9. Goff, S. P. & Berg, P. (1976) *Cell* **9**, 695–705.
10. Ganem, D., Nussbaum, A. L., Davoli, D. & Fareed, G. C. (1976) *Cell* **7**, 349–359.
11. Hamer, D. H., Davoli, D., Thomas, C. A., Jr. & Fareed, G. C. (1977) *J. Mol. Biol.* **112**, 155–182.
12. Upcroft, P., Skolnik, H., Upcroft, J. A., Solomon, D., Khoury, G., Hamer, D. H. & Fareed, G. C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2117-2121.
13. Mulligan, R. C., Howard, H. & Berg, P. (1979) *Nature (London)* **277**, 108–114.
14. Hamer, D. H. & Leder, P. (1979) *Nature (London)* **281**, 35–40.
15. Hamer, D. H., Smith, K. D., Boyer, S. H. & Leder, P. (1979) *Cell* **17**, 725–735.
16. Blattner, I. R., Williams, B. G., Blechl, A. E., Thompson, K. D., Faber, H. E., Furlong, L. A., Grunwald, D. J., Kiefer, D. O., Moore, D. D., Schumm, J. W., Sheldon, E. L. & Smithies, O. (1977) *Science* **196**, 161.
17. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Coonnel, C., Quon, D., Sim, G. U. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
18. Tegtmeyer, P. & Ozer, H. L. (1971) *J. Virol.* **8**, 516–524.
19. Hirt, B. (1967) *J. Mol. Biol.* **26**, 365–369.
20. Penman, S. (1966) *J. Mol. Biol.* **17**, 117–130.
21. Berk, A. J. & Sharp, P. A. (1977) *Cell* **12**, 721–732.
22. Berk, A. J. & Sharp, P. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1274–1278.
23. Lai, C.-J. & Khoury, G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 71–75.
24. Gruss, P., Lai, C.-J., Dhar, R. & Khoury, G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4317–4321.
25. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
26. Zain, B. S. & Roberts, R. J. (1979) *J. Mol. Biol.* **131**, 341–352.
27. Brockman, W. W. & Nathans, D. (1974) *Proc. Natl. Acad. Sci. USA* **72**, 942–946.
28. Simmons, D. T., Takemoto, K. K. & Martin, M. A. (1977) *J. Virol.* **24**, 319–325.
29. Gerner, E. W. & Humphrey, R. M. (1973) *Biochim. Biophys. Acta* **331**, 117.
30. Laskey, R. A. & Mills, A. D. (1975) *Eur. J. Biochem.* **56**, 335–341.
31. Lai, C.-J. & Nathans, D. (1974) *J. Mol. Biol.* **89**, 170–193.
32. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
33. Duguid, J. R., Steiner, D. F. & Chick, W. L. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3539–3543.
34. Chick, W. L., Warren, S., Chute, R. N., Like, A. A., Lauris, V. & Kitchen, K. C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 628–632.
35. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
36. Hamer, D. H. & Leder, P. (1979) *Cell* **18**, 1299–1302.
37. Blobel, G. & Sabatini, D. D. (1971) *Biomembranes* **2**, 193–195.
38. Milstein, C., Brownlee, G. G. & Harrison, T. M. (1972) *Nature (London) New Biol.* **239**, 117–120.
39. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol.* **67**, 835.
40. Tager, H. S. & Steiner, D. F. (1974) *Annu. Rev. Biochem* **43**, 509–538.
41. Lomedico, P. T., Chan, S. J., Steiner, D. F. & Saunders, G. F. (1977) *J. Biol. Chem.* **252**, 7971–7978.
42. Khoury, G., Gruss, P., Dhar, R. & Lai, C.-J. (1979). *Cell* **18**, 85–92.