Proc. Natl. Acad. Sci. USA
Vol. 76, No. 10, pp. 5036–5040, October 1979
Biochemistry

Construction and selection of recombinant plasmids containing full-length complementary DNAs corresponding to rat insulins I and II

(mRNA structure/determination of DNA nucleotide sequence/oligonucleotide priming)

SHU JIN CHAN, BARBARA E. NOYES, KAN L. AGARWAL, AND DONALD F. STEINER

Department of Biochemistry, University of Chicago, Chicago, Illinois 60637

Contributed by Donald F. Steiner, July 19, 1979

ABSTRACT **We have used a synthetic deoxydecanucleotide to generate an insulin-specific cDNA probe suitable for selecting transformants that contain nearly full-length cDNAs corresponding to the mRNAs coding for rat insulins I and II. Double-stranded cDNA was synthesized from x-ray-induced rat insulinoma poly(A)-RNA, inserted in pBR322 plasmid DNA by the homopolymeric tailing technique, and cloned in *Escherichia coli* χ1776. Colony hybridization with oligonucleotide-primed cDNA yielded 16 positive clones of which 7 corresponded to rat insulin I mRNA and 9 to rat insulin II mRNA. Restriction endonuclease maps of representative clones of each group indicated that these contained the complete coding sequences, as was confirmed by nucleotide sequence analysis of the 5′ region of the cloned DNA for rat insulin II. Nucleotide sequence analysis also established the amino acid sequence of the prepeptide of rat preproinsulin II. Comparison of the amino acid sequence of the prepeptides of rat preproinsulin I and II shows that three conservative amino acid substitutions have occurred in this region of the molecule.**

Two nonallelic genes are known to code for insulin in the rat (1, 2). Ullrich *et al.* (3) reported the nucleotide sequence for much of the coding region of rat insulin I mRNA on the basis of DNA sequences obtained from cloned double-stranded (ds) cDNA copies of normal rat islet cell poly(A)-RNA. They also reported differences in nucleotide sequence between the two insulin mRNAs in the region coding for the insulin A chain. More recently, Villa-Komaroff *et al.* (4) used cloned ds cDNA made from poly(A)-RNA of x-ray induced islet cell tumors to obtain additional nucleotide sequences for the 5′ end of rat insulin I mRNA, including sequences coding for 21 of the 24 amino acids in the prepeptide segment of preproinsulin. Neither laboratory, however, has reported the complete nucleotide sequence of rat insulin II mRNA, or nucleotide sequences in the 5′-untranslated region of these messengers. In this communication we describe the construction of recombinant plasmids containing inserts representative of nearly full-length copies of both rat insulin mRNA sequences. We have used a synthetic deoxydecanucleotide and poly(A)-RNA from islet cell tumors to generate a cDNA probe specific for sequences in the 5′ region of the insulin mRNA and have used this probe to select transformants containing insulin-related sequences. Restriction maps of the cloned DNA derived from both mRNAs are presented together with the nucleotide sequence for most of the 5′ end of rat insulin II mRNA. The nucleotide sequence also establishes the amino acid sequence of the prepeptide region of preproinsulin II.

MATERIALS AND METHODS

Enzymes. Reverse transcriptase (RNA-dependent DNA polymerase) from avian myeloblastosis virus was obtained from J. W. Beard (Life Sciences, Inc., St. Petersburg, FL). Restriction endonuclease *Hpa* II was generously provided by J. Hines. Other restriction endonucleases were purchased from Bethesda Research Laboratories or New England BioLabs. Calf thymus terminal deoxynucleotidyltransferase was obtained from M. S. Coleman at the University of Kentucky. *Escherichia coli* DNA polymerase I was obtained from Boehringer Mannheim. S1 nuclease was from Sigma.

Preparation of Poly(A)-RNA. An x-ray-induced rat insulinoma (5) originally obtained from W. Chick (Boston) and subsequently maintained in transplant in NEDH strain rats in our laboratory by A. Labrecque was used as the source of insulin mRNA. Total RNA was extracted from frozen tumor tissue as described (6) and was centrifuged through 5.7 M CsCl to remove residual DNA (7). Poly(A)-enriched RNA was recovered after two cycles of oligo(dT)-cellulose chromatography (8). The yield of poly(A)-RNA was about 50 μg/g of tissue.

Deoxydecanucleotide Primer and Synthesis of Insulin-Specific cDNA. The deoxydecanucleotide d(C-C-T-C-C-A-C-C-A-G) was chemically synthesized by using the diester approach (9), and its nucleotide sequence was confirmed by two-dimensional homochromatography (10). The decanucleotide was labeled at the 5′ end by using phage T4 polynucleotide kinase and [γ-^{32}P]ATP (11). Synthesis of ^{32}P-labeled decanucleotide-primed cDNA from rat tumor poly(A)-RNA, isolation of cDNA by polyacrylamide gel electrophoresis, and nucleotide sequence analysis using the procedure of Maxam and Gilbert (12) have been described (11). Hybridization of insulin-specific cDNA to poly(A)-RNA separated by agarose gel electrophoresis and transferred to diazobenzyloxymethyl-cellulose paper (13) was performed as described (11, 14).

Isolation of Plasmid DNA. Closed circular plasmid DNA was isolated from chloramphenicol-amplified cultures as described by Staudenbauer (15). The yield was about 500 μg of plasmid DNA per liter for non-χ1776 strains and 100 μg/liter for χ1776.

Construction of Recombinant Plasmids and Transformation of *E. coli* χ1776. Double-stranded cDNA was synthesized from insulinoma poly(A)-RNA by using reverse transcriptase for first strand synthesis and *E. coli* DNA polymerase I for second strand synthesis. The reverse transcriptase reaction mixture contained: 50 mM Tris·HCl at pH 8.2, 35 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, poly(A)-RNA at 200 μg/ml, $(dT)_{15}$ at 50 μg/ml, 200 μM each of the four unlabeled deoxynucleoside triphosphates, and [^{3}H]dCTP at 200 μCi/ml (27 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels). Reverse transcriptase, 400 units/ml, was added, and the reaction mixture was incubated at 42°C for 30 min. After phenol extraction and ethanol precipitation, RNA was removed by hydrolysis in 0.1 M NaOH for 15 min at 70°C. After neutralization and ethanol precipitation, the second strand was synthesized with *E. coli* DNA polymerase I (200 units/ml) in the presence of [α-^{32}P]dGTP as described by Efstradiatis *et al.* (16). The ds

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.

Abbreviation: ds, double-stranded.

cDNA was purified by gel filtration on Sephadex G-50 (0.6 × 10 cm column; equilibrated in 10 mM Tris·HCl, pH 7.4/0.1 mM EDTA), treated with S1 nuclease (17), and electrophoresed on a 5% polyacrylamide slab gel (18). After autoradiography, radiolabeled material migrating in the region corresponding to 400–600 base pairs was electrophoretically eluted (19). Approximately 50 dAMP residues were added to the ds cDNA by using terminal deoxynucleotidyl transferase as described by Roychoudhury *et al.* (20). Similarly, approximately 50 dTMP residues were added to the 3′ ends of 0.4 μg of *Hin*dIII-digested pBR322 plasmid DNA. After phenol extraction and ethanol precipitation, the DNAs were resuspended in 200 μl of 10 mM Tris·HCl, pH 7.4/140 mM NaCl/0.25 mM EDTA and annealed by heating to 65°C for 5 min, incubating at 42°C for 3 hr, and slow cooling to 4° overnight. The mixture was used directly to transform according to a procedure described by Curtiss *et al.* (21). χ1776 was grown in 50 ml of LB medium (22) containing diaminopimelic acid at 200 μg/ml and thymine at 50 μg/ml to an optical density at 550 nm of 0.3. The cells were collected by centrifugation, washed in 0.5 vol of cold 10 mM NaCl, and resuspended in 0.5 vol of cold $CaCl_2$ buffer containing 75 mM $CaCl_2$, 75 mM NaCl, and 10 mM Tris·HCl, pH 7.4. After 30-min incubation on ice, the cells were collected by centrifugation, resuspended in 400 μl of cold $CaCl_2$ buffer, and mixed with the annealed DNA mixture. The mixture was incubated on ice for 30 min, warmed to 37°C for 2 min, cooled to room temperature for 10 min, and spread on LB agar plates containing diaminopimelic acid at 200 μg/ml, thymine at 50 μg/ml, and ampicillin at 12.5 μg/ml. Transformants were grown for 2 days at 37°C and were transferred with toothpicks to plates containing tetracycline at 25 μg/ml. Because DNA inserted into the *Hin*dIII site of pBR322 is known to inactivate the tetracycline gene (3), transformants sensitive to tetracycline were subjected to colony hybridization (23) with insulin-specific cDNA synthesized from the deoxydecanucleotide primer as described above. All manipulations involving recombinant DNA were carried out in a P3 containment facility in accordance with National Institutes of Health guidelines (24).

RESULTS

Synthesis and Characterization of an Insulin-Specific cDNA. In order to efficiently identify clones with large DNA inserts that include the 5′ region of insulin mRNA, we prepared an insulin-specific cDNA probe from rat insulinoma poly(A)-RNA. A decanucleotide primer, d(C-C-T-C-C-A-C-C-A-G), that corresponded in sequence to the region coding for amino acids 11–13 plus the first nucleotide of the codon for residue 14 of the insulin B chain as reported by Ullrich *et al.* (3) was synthesized and labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase. The decanucleotide was then used to prime the synthesis of cDNA from rat insulinoma poly(A)-RNA. The

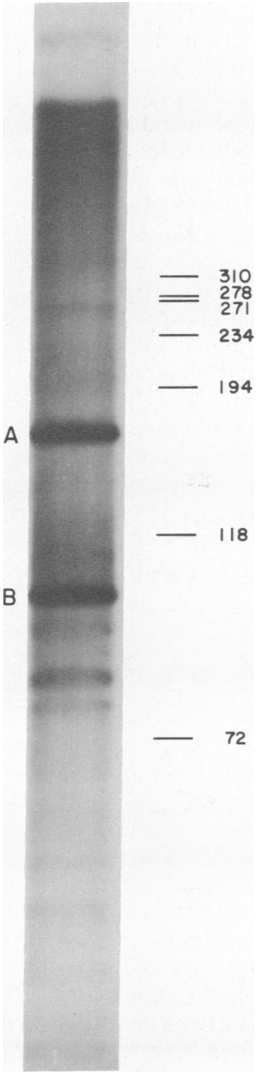

FIG. 1. Electrophoresis of cDNA synthesized from rat insulinoma poly(A)-RNA by using a deoxydecanucleotide primer. cDNA synthesized from 50 μg of poly(A)-RNA by using 1100 pmol of [^{32}P]dC-C-T-C-C-A-C-C-A-G) (2.6 × 10^6 cpm/pmol) as a primer for reverse transcription was electrophoresed on a 12% polyacrylamide/7 M urea slab gel (40 × 24 × 0.3 cm) in 50 mM Tris borate, pH 8.3/1.4 mM EDTA. Bands were visualized by autoradiography. Single-strand phage φX174 DNA digested with restriction endonuclease *Hae* III and end-labeled with [γ-^{32}P]ATP was used for molecular weight markers. The positions of these fragments are marked on the right. Band A cDNA is about 170 nucleotides, and band B cDNA is about 102 nucleotides long.

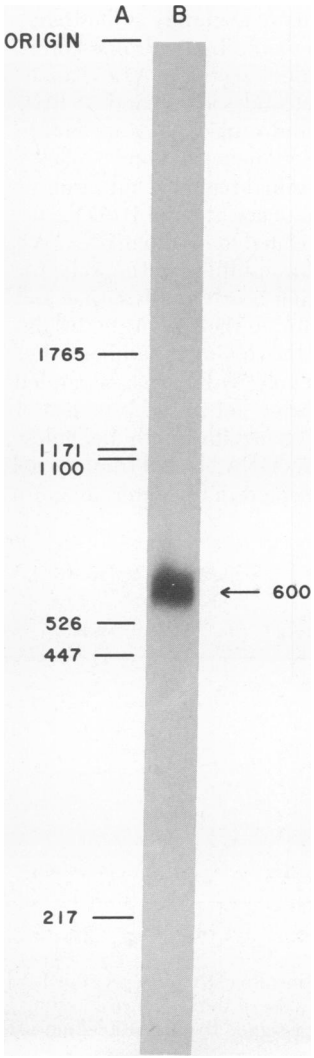

FIG. 2. Hybridization of [^{32}P]cDNA band A to rat insulinoma poly(A)-RNA. RNA, 5.7 μg, was electrophoresed on a 1.5% agarose gel in the presence of 4 mM CH_3HgOH for 15 hrs at 40 V (9). The RNA was transferred to diazobenzyloxymethyl-cellulose paper and hybridized to [^{32}P]cDNA band A extracted from the gel in Fig. 1 (5). Hybridization was for 45 hr at 37°C in 1.4 ml containing 1.4 × 10^5 cpm of cDNA (1.6 × 10^6 cpm/pmol). Autoradiography was for 2 hr with Kodak No Screen film. Simian virus 40 DNA digested with restriction endonuclease *Hin*dIII was electrophoresed on the same gel (lane A) and stained with ethidium bromide for use as molecular weight markers.

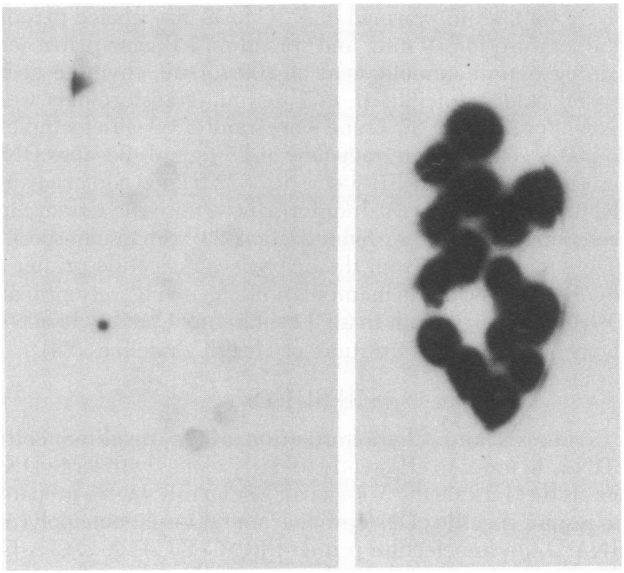

FIG. 3. Colony hybridization of transformed clones with cDNA band A or band B. Twenty-one clones that gave positive reactions when hybridized with total cDNA synthesized from decanucleotide-primed insulinoma poly(A)-RNA were immobilized on Millipore nitrocellulose filters and hybridized (23) with insulin-specific cDNA band A (*Right*) or non-insulin-specific band B (*Left*) purified from the polyacrylamide gel (Fig. 1).

terminally labeled cDNA products were separated by gel electrophoresis and visualized by autoradiography as illustrated in Fig. 1. Two prominent bands were observed, one corresponding to a species 170 nucleotides long (band A) and a second, equally intense, band 102 nucleotides long (band B). Both bands were eluted from the gel and subjected to nucleotide sequence analysis. The nucleotide sequence of band A cDNA corresponded to the known 5′-terminal region of rat insulin I mRNA (3), whereas the partial sequence of band B cDNA indicated that this product was not related to insulin mRNA. We therefore used band A cDNA as an insulin-specific probe for identifying transformants containing nearly full-length copies of insulin mRNA and for determining the size of the insulin mRNA in the preparation used for the cloning experiments. For size determination, rat insulinoma poly(A)-RNA was separated by electrophoresis on a 1.5% agarose gel in the presence of CH_3HgOH, transferred to diazobenzyloxymethyl-cellulose paper, and hybridized with band A cDNA isolated from the gel shown in Fig. 1. The results, presented in the autoradiogram of Fig. 2, show that band A cDNA hybridizes to a single size class of poly(A)-RNA about 600 nucleotides in length. This result agrees with the previous size estimate of ^{125}I-labeled rat insulin mRNA (6).

Identification and Characterization of Insulin Clones. Recombinant plasmids were constructed from insulinoma poly(A)-RNA by following a protocol similar to that used by Maniatis *et al.* (25) and others (26, 27) to clone globin cDNA except that we fractionated the ds cDNA on a 5% polyacrylamide slab gel after S1 nuclease digestion to further enrich for material containing complete insulin cDNA sequences. The ds cDNA that migrated in the region corresponding to 400–600 nucleotide base pairs in length was isolated by electrophoretic elution, and approximately 50 dAMP residues were added to the 3′ ends with terminal transferase. Similarly, *Hin*dIII-treated pBR322 DNA was tailed with approximately 50 dTMP nucleotides. The tailed plasmid and ds cDNA were annealed together and used to transform *E. coli* χ1776. Colonies containing recombinant plasmid DNA were identified on the basis of their resistance to ampicillin and sensitivity to tetracycline. From 40 ng of ds cDNA, 382 ampicillin-resistant tetracycline-sensitive transformants were obtained. Those transformants that contained insulin cDNA inserts were identified in two steps by using the *in situ* hybridization procedure described by Grünstein and Hogness (23). In the first step total cDNA synthesized from the decanucleotide primer was used as a probe and 21 positive clones were selected. These colonies were then screened with purified cDNA band A or band B isolated after gel electrophoresis. With the insulin-specific cDNA band A, 16 of the 21 clones gave positive reactions of about equal intensity. No positive colonies were obtained with band B (Fig. 3). Plasmid DNA was isolated from each of the 16 positive clones and analyzed by restriction endonuclease digestion and gel electrophoresis. The results (data not shown) indicated that all of the plasmids contained inserts of 400–700 base pairs, including the homopolymeric dA tails. Moreover, the restriction patterns obtained after digestion with endonuclease *Hae* III indicated that 7 of the clones contain rat insulin I DNA sequences while the remaining 9 clones contain insulin II DNA sequences.

Fig. 4 shows the restriction maps of the cDNA inserts in two clones, pRI-7 and pRI-11, which were studied in more detail. The cDNA insert in pRI-7, including the homopolymeric dA tails, is 540 base pairs in length. Restriction endonuclease analysis of the inserted cDNA showed the presence of a *Hae* III site at position −38 and a *Hin*f site at position 261, indicating that this clone contains the entire coding sequence for rat preproinsulin 1. Clone pRI-11 contained a larger cDNA insert

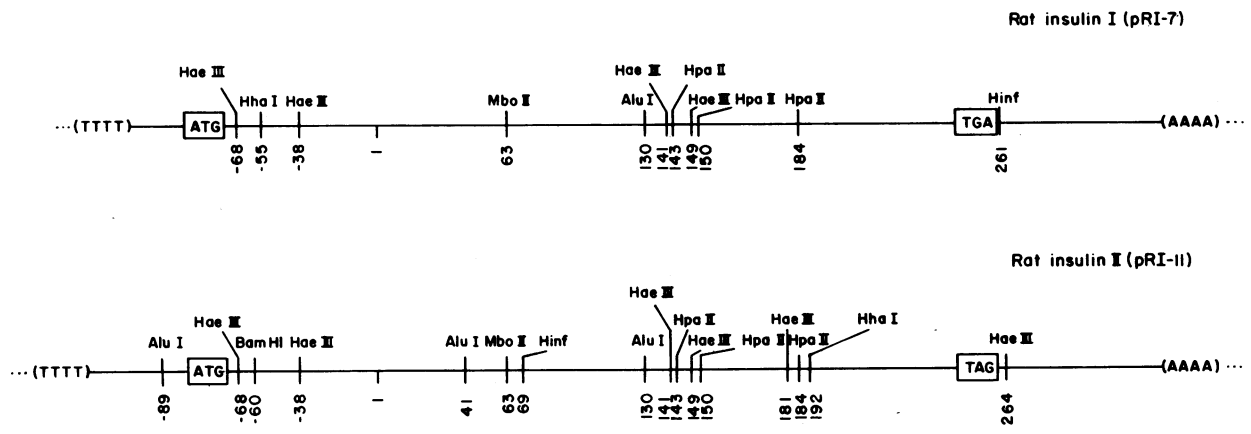

FIG. 4. Restriction maps of the insulin cDNA inserts in pRI-7 and pRI-11. Nucleotides are numbered as suggested by Villa-Komaroff *et al.* (4); i.e., position 1 indicates the first base of the sequence encoding proinsulin and nucleotides in the 5′ direction from this position are identified by negative numbers. The number assigned to each restriction site indicates the 5′-terminal nucleotide generated by cleavage at the message strand.

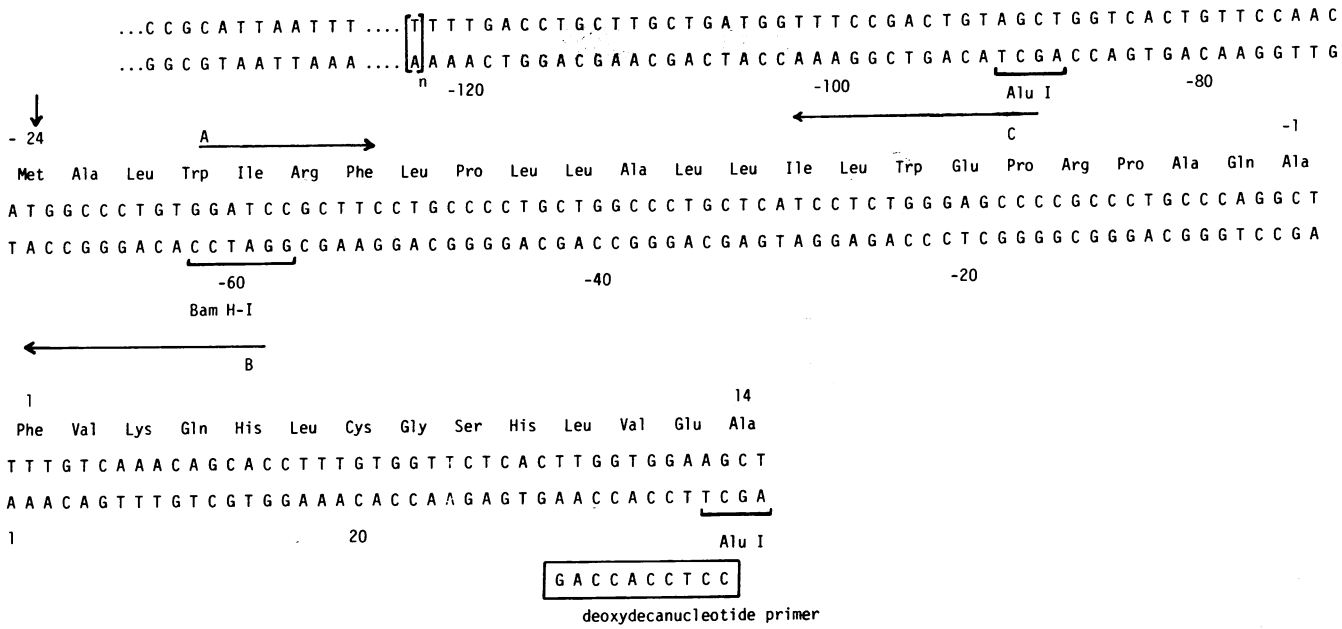

FIG. 5. Partial nucleotide sequence of the insertion in clone pRI-11. The nucleotide sequence corresponding to the 5′ end of rat insulin II mRNA was determined from 5′-labeled restriction fragments isolated from the clone pRI-11. Fragments A (129 nucleotides) and B (260 nucleotides) were isolated after digestion with *Bam*HI, labeling, and secondary cleavage with *Hinf* and *Hpa* II. Fragment C (243 base pairs) was isolated after *Hpa* II and *Alu* I digestion, and the sequences of the separated strands were determined to verify the data obtained from fragment B. The A/T junction between plasmid pBR322 and insulin cDNA is about 85 base pairs in length. The amino acid sequence for preproinsulin II is shown above the DNA sequence. The presence of serine at position 9 confirms the sequence as that of preproinsulin II. A vertical arrow indicates the initiator methionine. Negative numbers correspond to amino acids in the prepeptide, and positive numbers refer to the amino-terminal residues of the B chain. The deoxydecanucleotide used as a primer for synthesis of insulin I-specific cDNA is shown in the box.

of about 600 base pairs (including homopolymeric tails) and had a restriction map consistent with rat insulin II cDNA. As illustrated in Fig. 4, the cDNA insert in pRI-11 contains different cleavage sites for *Hha* I and *Hinf* and additional sites for *Alu* I and *Hae* III relative to pRI-7. The insert in pRI-11 also contains a single cleavage site for *Bam*HI near the beginning of the coding sequence.

For nucleotide sequence analysis, plasmid pRI-11 was cleaved with *Bam*HI and the resulting fragments were labeled with [γ-^{32}P]ATP by using polynucleotide kinase. The labeled fragments were further digested with *Hinf* and *Hpa* II. Fragments A (129 nucleotides) and B (260 nucleotides) were isolated and their nucleotide sequences were determined. Fig. 5 shows the nucleotide sequences of fragments A and B. To further confirm the sequence of fragment B, fragment C (243 base pairs) was isolated after *Hpa* II and *Alu* I digestion and the strands were separated and their nucleotide sequences were determined. Thus the 5′-terminal nucleotide sequence shown in Fig. 5 was confirmed by studying two different restriction fragments.

DISCUSSION

We previously showed that when poly(A)-RNA isolated from an x-ray-induced rat insulinoma is translated in a wheat germ cell-free system, approximately 25% of the synthesized protein can be immunoprecipitated with anti-insulin sera (28). However, the mass of insulin mRNA present in these preparations appears to be substantially less. Restriction endonuclease *Hae* III analysis of the ds cDNA prepared from total poly(A)-RNA indicates that only 1–2% of the cDNA contains sequences related to insulin. The reason for this discrepancy is unclear, although a similar observation has been reported by Villa-Komaroff *et al.* (4). This finding, and the known difficulties in obtaining bacterial clones containing full-length insulin cDNA inserts, necessitated the development of a more efficient method for selection of the small number of clones expected to have longer inserts. Our approach was to construct a highly specific cDNA probe corresponding in sequence to the 5′ region of rat insulin mRNA by using a specific deoxydecanucleotide as a primer for the reverse transcription of total rat insulinoma poly(A)-RNA. When the transcribed insulin cDNA was isolated by polyacrylamide gel electrophoresis and used in colony hybridization experiments, only those colonies containing nearly full-length copies of the mRNA were detected. Clones containing shorter inserts or sequences unrelated to insulin were not detected. Because the sequence of the decanucleotide corresponds to a region of amino acid sequence in the insulin chain that is highly conserved in several species (29), it may prove useful in studies with other insulins as well.

In five independent experiments using three different preparations of rat insulinoma RNA, band A cDNA (as shown in Fig. 1) was consistently found to be 170–180 nucleotides long. In each case sequence analysis indicated that the product is derived from rat insulin I mRNA. However, band A was used successfully to identify both rat I and II DNA sequences. Of 382 colonies tested, 7 clones were found to contain nearly full-length inserts corresponding to rat insulin I and 9 clones contained long rat insulin II inserts.

The length of band A DNA suggests that this cDNA may represent a complete copy of the mRNA from the priming site through the 5′ end of the molecule. Taken together with nucleotide sequence data previously published (3, 4), this indicates that the length of rat insulin I mRNA is 455 nucleotides, exclusive of the 3′ poly(A) tail of about 150 nucleotides (ref. 6; Fig. 2 of the present paper). Of the 455-nucleotide segment, 333 nucleotides code for preproinsulin, including the initiator methionine and termination codons, and there are 54 nucleotides in the untranslated 3′ end (3). Thus, we estimate that the untranslated 5′ end of the mRNA is about 68 nucleotides long. By restriction mapping and partial nucleotide sequence analysis we have shown that clone pRI-11 contains the complete coding sequence for preproinsulin II in addition to 47 nucleotides of

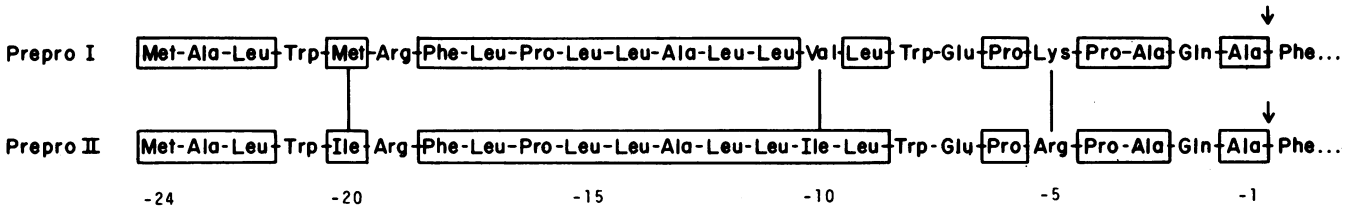

FIG. 6. Amino acid sequences of the pre-peptide segments of rat preproinsulin I and II. The sequence for rat preproinsulin I is derived from nucleotide sequence data (3, 4) and radiosequences of labeled preproinsulin (28). The sequence for preproinsulin II is derived from the nucleotide sequence of pRI-11. Boxes indicate residues that were identified by radiosequencing. The identification of the initiator methionine at position 24 was confirmed by specific labeling with formyl-[^{35}S]Met-tRNA$_f^{Met}$ (unpublished observations).

the untranslated 5′ region. Assuming that rat insulin I and II mRNAs are approximately the same size, then pRI-11 is only about 20 nucleotides short of being a full-length copy.

Partial nucleotide sequence analysis of the insert in pRI-11 also established the complete amino acid sequence of the prepeptide for rat preproinsulin II. A comparison of the amino acid sequences of the two rat prepeptides (Fig. 6) reveals only three conservative amino acid substitutions. These results do not support our previous assignment of two phenylalanine residues at positions −15 and −16 in the prepeptide of rat preproinsulin II (28). However, in further radiosequencing studies (unpublished data) we have confirmed these assignments in the cell-free translation product of mRNA isolated from Sprague–Dawley islets, while sequence studies on the tumor insulin mRNA cell-free translation product identified phenylalanines only at positions −18 and +1, in agreement with our nucleotide sequence data. Possible strain differences among rat insulin mRNAs might account for this discrepancy.

Since this work was completed we learned that Efstratiadis and coworkers (30) have identified and determined the nucleotide sequence of the portions of bacterial clones that contain the chromosomal genes coding for rat insulins I and II. The nucleotide sequence of the mRNA that we derived from clone pRI-11 agrees with the data from the chromosomal gene sequence for rat insulin II, in the region coding for the prepeptide and for the first 15 nucleotides of the adjacent untranslated 5′ region. However, the sequence of the remaining 32 nucleotides corresponding to the untranslated 5′ region of the mRNA, (positions −86 to −117 in the pRI-11 segment shown in Fig. 5) are inverted and complementary to the corresponding region on the chromosomal gene, which also contained a large intervening sequence in this region. The reason for this unusual discrepancy is unclear. One possibility is that a DNA sequence inversion has occurred in the rat insulinoma from which we extracted the mRNA. Alternatively, we cannot exclude the possibility that an error has occurred in the cloning procedure.

We thank Argiris Efstratiadis and Peter Lomedico for helpful discussions. We thank Lise McKean for expert assistance in preparing this manuscript. This work was supported in part by U.S. Public Health Service Grants AM-13914 and AM-21901, Diabetes Center Grant AM-20295, and American Cancer Society Grant (IN-41-Q). K.L.A. is a Research Career Development Awardee of the U.S. Public Health Service (GM-00224) and B.E.N. is a Research Development Awardee of the American Diabetes Association.

1. Smith, L. F. (1966) *Am. J. Med.* **40**, 662–666.
2. Clark, J. L. & Steiner, D. F. (1969) *Proc. Natl. Acad. Sci. USA* **62**, 278–285.
3. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. & Goodman, H. M. (1977) *Science* **196**, 1313–1315.
4. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727–3731.
5. Chick, W. L., Warren, S., Chute, R. N., Like, A., Lauris, V. & Kitchen, K. C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 628–633.
6. Duguid, J. R., Steiner, D. F. & Chick, W. L. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3539–3543.
7. Glisin, V., Crkvenjakov, R. & Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
8. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
9. Agarwal, K. L., Yamazaki, A., Cashion, P. & Khorana, H. (1972) *Angew. Chem. Int. Ed. Engl.* **11**, 451–498.
10. Jay, E., Bambara, R., Padmanabhan, R. & Wu, R. (1974) *Nucleic Acids Res.* **1**, 331–353.
11. Noyes, B., Mevarech, M., Stein, R. & Agarwal, K. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1770–1774.
12. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
13. Allwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.
14. Mevarech, M., Noyes, B. E. & Agarwal, K. L. (1979) *J. Biol. Chem.* **254**, 7472–7475.
15. Staudenbauer, W. L. (1976) *Mol. Gen. Genet.* **145**, 273–280.
16. Efstratiadis, A., Kafatos, F. C., Maxam, A. & Maniatis, T. (1976) *Cell* **7**, 279–288.
17. Shenk, T. E., Rhodes, C., Rigby, P. W. J. & Berg, P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 989–993.
18. Maniatis, T., Jeffery, A. & van de Sande, H. (1975) *Biochemistry* **14**, 3787–3794.
19. McDonell, M. W., Simon, M. N. & Studier, F. W. (1977) *J. Mol. Biol.* **110**, 119–146.
20. Roychoudhury, R., Jay, E. & Wu, R. (1976) *Nucleic Acids Res.* **3**, 101–116.
21. Curtiss, R., Pereira, D. A., Hsu, J. C., Hull, S. C., Clark, J. E., Maturin, L. J., Sr., Goldschmidt, R., Moody, R., Inoue, M. & Alexander, L. (1977) in *Recombinant Molecules: Impact on Science and Society*, eds. Beers, R. F., Jr. & Basset, E. G. (Raven, New York), pp. 45–56.
22. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 433.
23. Grünstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
24. Federal Register (1978) *Fed. Reg.* **43**, 60080–60105.
25. Maniatis, T., Kee, S. G., Efstratiadis, A. & Kafatos, F. C. (1976) *Cell* **8**, 163–182.
26. Higuchi, R., Paddock, G. V., Wall, R. & Salser, W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3146–3150.
27. Humphries, P., Old, R., Coggins, L. W., McShane, T., Watson, C. & Paul, J. (1978) *Nucleic Acids Res.* **5**, 905–924.
28. Chan, S. J., Keim, P, & Steiner, D. F. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1964–1968.
29. Peterson, J. D., Steiner, D. F., Emdin, E. O. & Falkmer, S. (1975) *J. Biol. Chem.* **250**, 5183–5191.
30. Efstratiadis, A., Broome, S., Chick, W. L., Gilbert W., Villa-Komaroff, L., Lomedico, P., Naber, S. P., Perler, F. & Tizard, R. (1979) *J. Supramol. Struct. Suppl.* **3**, 39.