

Tissue-specific expression of insulin-like growth factor II mRNAs with distinct 5' untranslated regions

(hypothalamus/liver/RNase H)

JEAN-CLAUDE IRMINGER*, KENNETH M. ROSEN†, RENÉ E. HUMBEL*, AND LYDIA VILLA-KOMAROFF†

*Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland; and †Departments of Neuroscience and Neuropathology, Harvard Medical School, The Children's Hospital, 300 Longwood Avenue, Boston, MA 02115

Communicated by Walter Gilbert, April 6, 1987

ABSTRACT We have used RNA from human hypothalamus as template for the production of cDNAs encoding insulin-like growth factor II (IGF-II). The prohormone coding sequence of brain IGF-II RNA is identical to that found in liver; however, the 5' untranslated sequence of the brain cDNA has no homology to the 5' untranslated sequence of the previously reported liver cDNAs. By using hybridization to specific probes as well as a method based on the properties of RNase H, we found that the human IGF-II gene has at least three exons that encode alternative 5' untranslated regions and that are expressed in a tissue-specific manner. A probe specific to the brain cDNA 5' untranslated region hybridizes to a 6.0-kilobase transcript present in placenta, hypothalamus, adrenal gland, kidney, Wilms tumor, and a pheochromocytoma. The 5' untranslated sequence of the brain cDNA does not hybridize to a 5.3-kilobase transcript found in liver or to a 5.0-kb transcript found in pheochromocytoma. By using RNase H to specifically fragment the IGF-II transcripts into 3' and 5' fragments, we found that the RNAs vary in size due to differences in the 5' end but not the 3' end.

The somatomedins, or insulin-like growth factors (IGFs), are a family of peptides related to insulin by structure and function (for review see refs. 1-5). Two of the somatomedins, IGF-I and IGF-II, are 62% homologous at the amino acid level (6). Despite structural and functional similarities, evidence has begun to accumulate that these two peptides have distinct and different roles. IGF-I mediates the growth effects of growth hormone in childhood and adolescence (7, 8). The role of IGF-II is not known. Levels of IGF-II are less dependent on levels of growth hormone (9) and IGF-II is a less potent mitogen *in vivo* than IGF-I (8). IGF-II can be synthesized by a number of fetal tissues (10) and, in rats, may stimulate fetal growth (11). In humans, IGF-II levels in serum persist into adulthood at 3 or 4 times the level of IGF-I (9). In adult human cerebrospinal fluid and brain, IGF-II, but not IGF-I, was found (12, 13). However, a truncated form of IGF-I as well as IGF-II was recently isolated from fetal and adult human brain (14, 15). Evidence for local synthesis in pituitary and brain (13, 16) suggests a physiological role for IGF-II as a neuroregulator and/or a brain growth factor.

Although several forms of IGF-II that differ in molecular size have been found in human brain and spinal fluid, their exact relationship to serum IGF-II has not been established (12). The major serum form of IGF-II is the classical 7.5-kDa form (6). Two minor variants, postulated to be the result of polymorphic IGF-II alleles, have also been described (17, 18).

Current data indicate that a single gene for IGF-II is located on chromosome 11 (19, 20) adjacent to the insulin gene (21, 22). However, the 5' untranslated region of the human cDNA

clones is different from the 5' sequence reported in the genomic clone (23). To investigate the possibility that different coding regions and/or different 5' untranslated regions may be present on IGF-II mRNAs expressed in different tissues, we have used RNA from human hypothalamus as template for the production of cDNAs encoding IGF-II.

MATERIALS AND METHODS

Synthesis and Cloning of cDNA. RNA was prepared using guanidinium thiocyanate CsCl (24, 25). Oligonucleotide (3.8 $\mu\text{g/ml}$) was annealed to poly(A)⁻ RNA (0.38 mg/ml) in 50 mM sodium phosphate, pH 7.2/5 mM EDTA at 90°C for 3 min. KCl was added to 100 mM and the mixture was cooled over 5 min. Double-stranded cDNA was synthesized as described by Gubler and Hoffman. (26), inserted into the *Pst* I site by oligo(dG)-oligo(dC) tailing (27), and used to transform *Escherichia coli* MM294 prepared as described (28). As probes for colony screening, we used oligonucleotides labeled at their 5' end by use of bacteriophage T4 polynucleotide kinase and [γ -³²P]ATP. Filters were prehybridized in 6 \times SSC (0.9 M sodium chloride/90 mM sodium citrate)/1 \times Denhardt's solution (0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin/0.5% NaDodSO₄/sheared and denatured herring sperm DNA (100 $\mu\text{g/ml}$)/0.05% sodium pyrophosphate at 37°C for 2 hr and hybridized in 6 \times SSC/1 \times Denhardt's solution/tRNA (20 $\mu\text{g/ml}$)/0.05% sodium pyrophosphate at 42°C (23-mer) or 37°C (17-mer) for \approx 15 hr. Filters were washed in 6 \times SSC/0.05% sodium pyrophosphate several times at room temperature, at 42°C (23-mer) or 37°C (17-mer) for 1 hr and at a temperature approximately 5°C beyond the "melting" temperature (t_m) for 10 min.

Synthesis of Oligodeoxynucleotides. Oligonucleotides were synthesized in a model 380A DNA synthesizer (Applied Biosystems) and purified by electrophoresis in a 20% polyacrylamide gel.

cDNA Sequencing. Isolated cDNAs were sequenced by either the chemical method (29) or the enzymatic method (30).

Blot Hybridization Analysis of RNA. Poly(A)⁺ RNA (3 μg) was fractionated in 1% agarose/2.2 M formaldehyde gels (31) and transferred to nylon filters (GeneScreenPlus, New England Nuclear). For probing with cDNA, blots were prehybridized in 0.5 M sodium phosphate buffer/1% bovine serum albumin/1 mM EDTA/7% NaDodSO₄ at 65°C for 2 hr and hybridized in the same solution at 65°C for 40 hr. The filters were washed in 40 mM sodium phosphate/5% NaDodSO₄/1 mM EDTA/0.5% bovine serum albumin at room temperature for 5 min, in 40 mM sodium phosphate/1% NaDodSO₄/1 mM EDTA four to seven times at 65°C for 3 min, and in 100 mM sodium phosphate three times at 65°C for 3 min. As probe, we used IGF-II cDNA nick-translated to 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.

Abbreviation: IGF, insulin-like growth factor.

specific activity of 5×10^8 dpm/ μ g, using [α - 32 P]dCTP. The filters were exposed to Fuji RX x-ray films for 12–150 hr.

RNAse H Mapping. Five micrograms of total placental RNA or poly(A)⁺ liver RNA and 20 ng of a specific IGF-II cDNA restriction fragment were dried, dissolved in 25 μ l of hybridization solution [75% (vol/vol) formamide/0.5 M NaCl/0.04 M Pipes, pH 6.4/1 mM EDTA/0.05% NaDodSO₄] and then incubated overnight at 55°C after heating to 90°C for 5 min. The sample was diluted 1:4 and nucleic acids were precipitated by addition of ethanol. Precipitates were washed once with ethanol, dried, and dissolved in RNAse H buffer (20 mM Tris-HCl, pH 7.5/10 mM MgCl₂/100 mM KCl/0.1 mM dithiothreitol/5% sucrose). RNAse H (1 unit) was added and the mixture was incubated for 30 min at 37°C. After digestion, EDTA was added to 10 mM and the sample was extracted with an equal volume of phenol/chloroform (1:1, vol/vol). The aqueous phase was removed and nucleic acids were ethanol-precipitated. The pellet was washed with ethanol, dried, dissolved in sterile distilled H₂O, and electrophoresed in denaturing formaldehyde/agarose gels for blot hybridization analysis as described (32). IGF-II DNA was labeled either by nick-translation to a specific activity of 1–5 $\times 10^8$ cpm/ μ g or by replacement synthesis with phage T4 DNA

polymerase to a specific activity of 3–5 $\times 10^7$ cpm/ μ g (33). After hybridization, filters were washed for 1 hr at 60°C in 0.1 \times SSC/0.1% NaDodSO₄ and then exposed to Kodak XAR-5 film at room temperature or at –70°C with Cronex Lightning Plus intensifying screens.

RESULTS

We extracted RNA from human adult hypothalamus and synthesized two double-stranded cDNA libraries. To enrich for sequences encoding IGF-II, we used IGF-II-specific oligonucleotides to prime cDNA synthesis. One library was constructed using an oligomer that was 21 nucleotides long and complementary to the 3' end of the IGF-II coding sequence (34) (Oligo B, Fig. 1). The second library was constructed using an oligomer that was 23 nucleotides long and complementary to the E peptide of IGF-II (Oligo H, Fig. 1).

Transformation with the cDNA inserted into plasmid pBR322 resulted in 10³ colonies, which were screened with ³²P-labeled oligonucleotides A and C (see Fig. 1). Only one of the five colonies that hybridized with both oligonucleotides contained IGF-II sequence. The plasmid, pIGF-II/1, con-

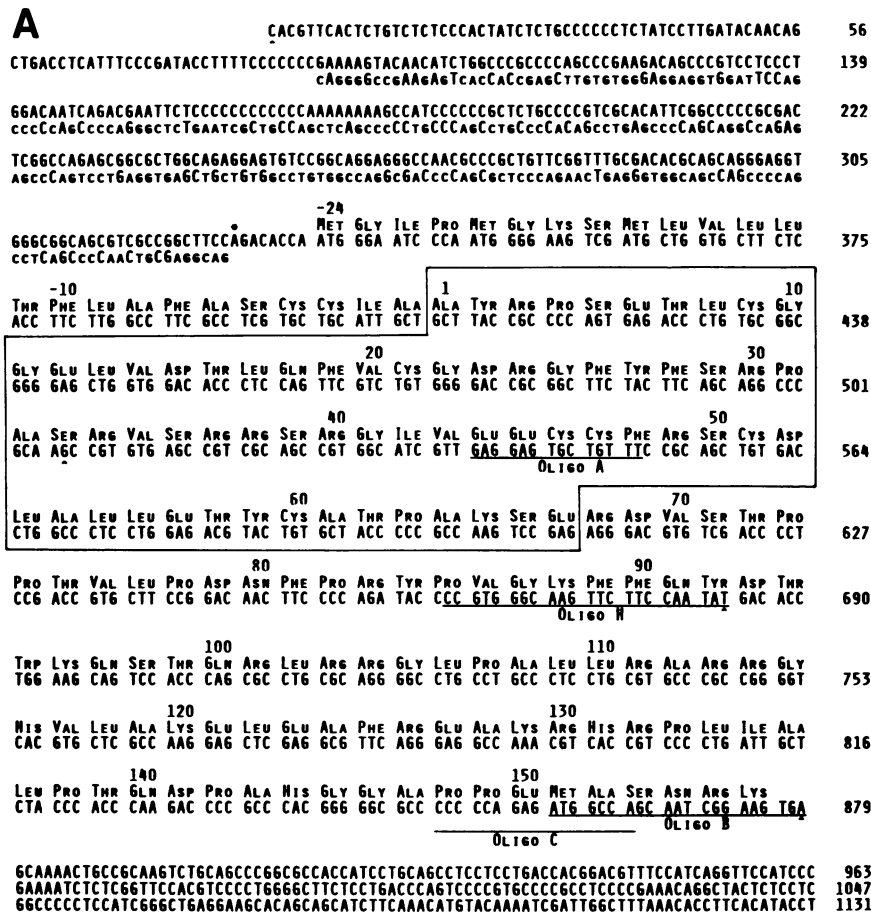
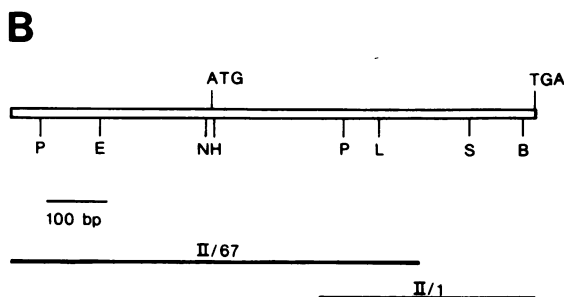


FIG. 1. (A) Sequence of human prepro-IGF-II cDNA and protein. The predicted amino acid sequence of prepro-IGF-II is numbered by designating the first amino acid of IGF-II as 1. The region corresponding to IGF-II is boxed. Clone pIGF-II/67 contains the sequence from nucleotide 1 to 684 and clone pIGF-II/1 contains the sequence from nucleotide 515 to 879. Sequence from nucleotide 880 to 1131 is from ref. 19. The number of the nucleotide at the end of each line is indicated. Oligonucleotides used for priming and screening cDNA libraries are underlined. Homology with other reported cDNA sequences ends in the 5' untranslated sequence at position 329 (indicated by an asterisk). Sequence of the 5' untranslated region from liver cDNA is shown below the sequence of the brain cDNA and is from ref. 34. (B) Composite restriction map of IGF-II cDNAs. P, *Pvu* II; E, *Eco*R I; N, *Nae* I; H, *Hin* I; L, *Sal* I; S, *Sst* I; B, *Bal* I. Positions of the initiator (ATG) and terminator (TGA) codons are indicated. bp, Base pairs.



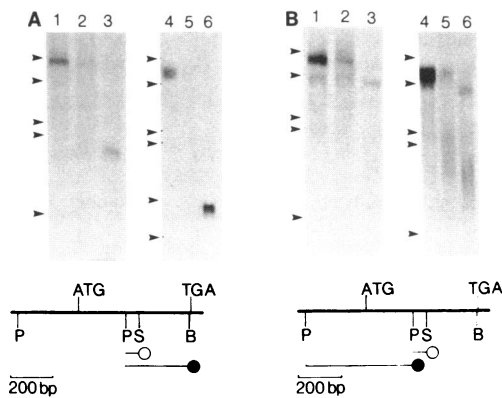


FIG. 4. RNase H mapping. Five micrograms of placental total RNA or liver poly(A)⁺ RNA was annealed to a *Pvu* II-*Sal* I restriction fragment and digested with RNase H, and (after electrophoresis) hybridized to a probe specific for a region either 5' or 3' to the annealing fragment. Lanes 1 and 4: RNA without treatment. Lanes 2 and 5: RNA mock-annealed without restriction fragment and digested with RNase H. Lanes 3 and 6: RNA annealed with fragment and digested with RNase H. (A) Hybridization with a 5'-specific probe. The schematic below shows location of annealing fragment (○) and probe (●) against cDNA sequence. P, *Pvu* II; S, *Sal* I; B, *Bal* I. (B) Hybridization with 3'-specific probe. Schematic shows location of annealing fragment (○) and probe (●). Arrowheads show positions of size markers, which are, in decreasing order, 6682, 4361, 2322, 2027, and 564 bases long.

these RNAs is ≈4150 bases beyond the UGA termination codon.

DISCUSSION

Previous reports indicate that there is a single gene on chromosome 11 encoding IGF-II. The sequences and restriction maps reported (23, 34) are in agreement for the exons that encode the prohormone. The 5' untranslated regions, how-

ever, differ completely. Bell *et al.* (34) and Jansen *et al.* (35) reported the sequence of the 5' untranslated region found on cDNA clones isolated from a human liver cDNA library. An exon containing the liver 5' untranslated region was found to be ≈11 kb from the exon encoding the amino-terminal end of the prohormone. Recently, de Pagter-Holthuizen *et al.* (18) reported that two exons encode the 5' untranslated region reported by Bell *et al.* (34). Dull *et al.* (23) isolated cDNA encoding rat IGF-II from a rat liver cell line and used this cDNA to isolate a human genomic sequence. They identified an exon containing a 5' untranslated region, 2.6 kb from the exon encoding the amino-terminal end, that was ≈80% homologous to the rat 5' sequence.

Soares *et al.* (40) also isolated cDNA clones from the same rat liver cell line and found a 5' untranslated region homologous to the sequence reported by Bell *et al.* (34). We have isolated cDNAs encoding IGF-II from a human brain cDNA library and found a 5' untranslated sequence that corresponds exactly to the genomic sequence reported by Dull *et al.* (23). This sequence is present in a 6-kb IGF-II mRNA from hypothalamus, placenta, adrenal gland, kidney, Wilms tumor, and pheochromocytoma but is not present either in the 5.3-kb IGF-II mRNA present in liver or in an additional 5-kb transcript present in a pheochromocytoma. Thus, there are at least three exons that encode two alternative 5' untranslated regions of IGF-II mRNA and are expressed in a tissue-specific manner. A composite map that reconciles the differences in the previous reports and indicates the sizes of the exons encoding the 5' and 3' regions of the IGF-II mRNAs is shown in Fig. 5. Use of these exons appears to be tissue-specific in normal tissues. Placenta, adrenal gland, and brain all express a 6.0-kb transcript that contains a 5' untranslated region encoded by exon 3. Liver expresses a 5.3-kb transcript that contains a 5' untranslated region encoded by exons 1 and 2. These two transcripts could arise by at least two separate mechanisms, which are diagrammed in Fig. 5. The first possibility is that the IGF-II gene, like the amylase gene (41), has two separate promoters, one upstream of exon 1 and one upstream of exon 3. In this case, the 5' ends of the two RNAs

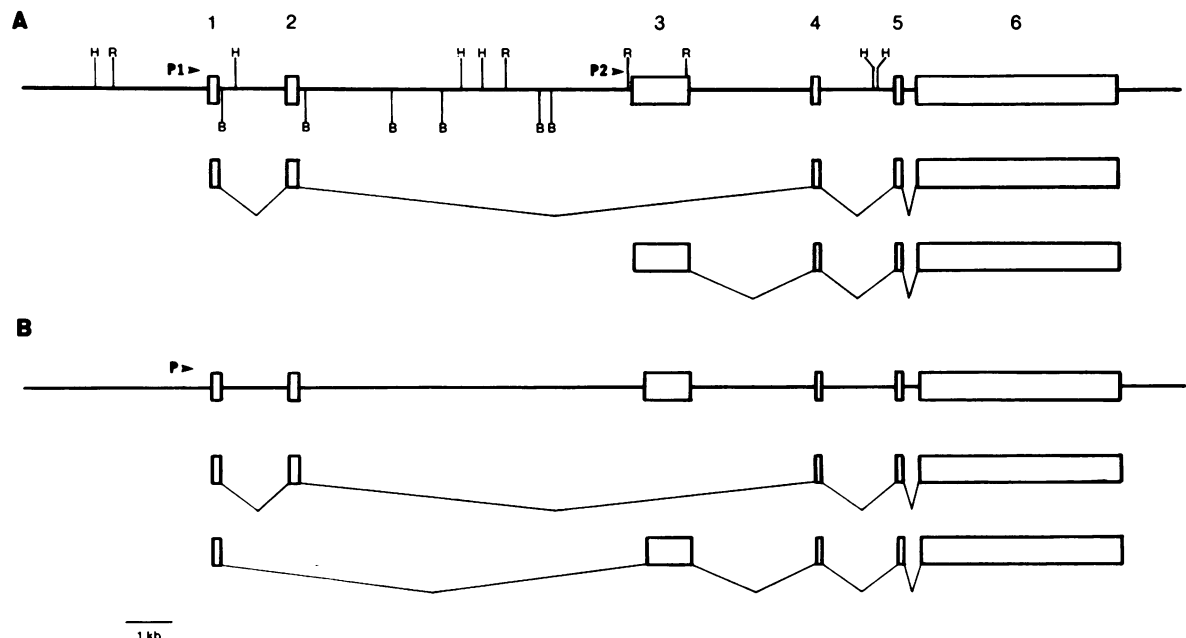


FIG. 5. Genomic organization of the human IGF-II gene. Two alternative mechanisms for generating IGF-II transcripts of 6.0 and 5.3 kb. Blot hybridization of human genomic DNA with a cDNA probe are completely consistent with this map (refs. 18, 23, 34, and data not shown). (A) The gene and transcripts produced utilizing two promoters (P1 and P2). Exon sizes: 1, 220 bp; 2, 242 bp; 3, 1280 bp; 4, 163 bp; 5, 149 bp; 6, 4150 bp. Restriction sites: R, *Eco*RI; H, *Hind*III; B, *Bam*HI. Restriction sites and exon sizes were compiled from refs. 18, 23, 34, and data presented in this paper. (B) The gene and transcripts produced utilizing one promoter (P). Exons are the same size as in A except for exon 3, which is ≈1080 bp.

would be different up to the cap site and exon 3 would be 1280 bp long. The second possibility is that the IGF-II gene has a single promoter upstream of exon 1 and the two transcripts are generated by differential splicing. In this case, the liver-specific transcripts would contain exons 1, 2, 4, 5, and 6; the 6-kb specific transcript would contain exons 1, 3, 4, 5, and 6; and both types of transcripts would have a common 5' sequence near the cap site. If this is the case, exon 3 would be 1080 bp long. Nuclease S1 protection experiments using fragments of genomic clones or hybridization of cDNA complementary to exon 1 would distinguish between these two possibilities. The 5-kb transcript in the pheochromocytoma could be explained by a transcript containing exons 1, 4, 5, and 6. Interestingly, pheochromocytomas contain high amounts of immunoreactive IGF-II, whereas Wilms tumors contain the same level of immunoreactive IGF-II as normal tissue despite increased levels of IGF-II RNA (42).

The extent of the 5'- and 3'-most exons has been uncertain because none of the reported cDNAs are full-length. We exploited the properties of RNase H to define the number of bases 5' and 3' of the coding region. Since the 3' fragment produced using liver RNA is the same size as the 3' fragment produced using placenta RNA, it appears that the 3' end of the IGF-II RNAs contain the same sequence, although this sequence, shown as a single exon of 4380 bp in Fig. 5, might be broken up into more than one exon. However, small differences (50–100 bases) in the sizes of the 3' end would not have been detected in our assay. The 5' end of the transcripts is still somewhat ambiguous. Comparing the size of the 5' region as determined by the RNase H experiments to the sizes of exons 1 and 2 reported by de Pagter-Holthuisen *et al.* (18), about 100 nucleotides of 5' sequence present in the liver mRNA remain to be accounted for. It is possible that an additional 5' exon remains to be found.

The function of IGF-II in brain is not clear. Some evidence points to a role in brain growth and development of certain neurons and glial cells (43–46). The postulated local synthesis of IGF-II in brain (13) is confirmed by our finding of IGF-II mRNA in hypothalamus. Since the coding region of the transcript expressed in hypothalamus is identical with the one coding for serum IGF-II, at least some of the 7.5-kDa form of brain IGF-II is identical with serum IGF-II.

We are very grateful to Dr. Rolf Gaillard (University of Geneva) for providing us with human hypothalamus and to Dr. E. Schoenle for providing us with Wilms tumor tissue. We wish to thank Dr. R. Jaussi and Ed Lamperti for advice and many helpful discussions and I. Irminger for critical reading of the manuscript. The work was supported by Basic Research Grant I-946 from the March of Dimes Birth Defects Foundation, a grant from the Association for Retarded Citizens of the United States, National Institutes of Health Mental Retardation Research Center Program Grant HD18655, and Swiss National Science Foundation Grant 3.328.82. K.M.R. was supported by Developmental Neurology Training Grant 5T32NS07264-03 from the National Institutes of Health.

- Blundell, T. L., Bedarkar, S. & Humbel, R. E. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **42**, 2592–2597.
- Hall, K. & Sara, V. R. (1983) *Vitam. Horm.*, (N.Y.) **40**, 175–233.
- Humbel, R. E. (1984) in *Hormonal Proteins and Peptides*, ed. Li, C. H. (Academic, New York), Vol. 12, pp. 57–79.
- Nissley, P. S. & Rechler, M. M. (1984) in *Hormonal Proteins and Peptides*, ed. Li, C. H. (Academic, New York), Vol. 12, pp. 127–203.
- Froesch, E. R., Schmid, C., Schwander, J. & Zapf, J. (1985) *Annu. Rev. Physiol.* **47**, 443–467.
- Rinderknecht, E. & Humbel, R. E. (1978) *FEBS Lett.* **89**, 283–286.
- Daughaday, W. H., Hall, K., Ruben, M. S., Salmon, W. D., Vanden Brande, J. L. & Van Wyk, J. J. (1972) *Nature (London)* **235**, 107.
- Schoenle, E., Zapf, J., Humbel, R. E. & Froesch, L. M. (1982) *Nature (London)* **296**, 252–253.
- Zapf, J., Walter, H. & Froesch, E. R. (1981) *J. Clin. Invest.* **68**, 1321–1330.
- D'Ercole, A. J., Applewhite, G. T. & Underwood, L. E. (1980) *Dev. Biol.* **75**, 315–328.
- Adams, S. O., Nissley, S. P., Handwerger, S. & Rechler, M. M. (1983) *Nature (London)* **302**, 150–153.
- Haselbacher, G. & Humbel, R. (1982) *Endocrinology* **110**, 1822–1824.
- Haselbacher, G., Schwab, M. E., Pasi, A. & Humbel, R. E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2153–2157.
- Sara, V. R., Carlsson-Skwirut, C., Andersson, C., Hall, E., Sjöhoegren, B., Holmgren, A. & Joernvall, H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4904–4907.
- Carlsson-Skwirut, C., Joernvall, H., Holmgren, A., Andersson, C., Bergman, T., Lundquist, G., Sjöhoegren, B. & Sara, V. R. (1986) *FEBS Lett.* **201**, 46–50.
- Binoux, M., Hossenlopp, P., Lassarre, C. & Hardoin, N. (1981) *FEBS Lett.* **24**, 178–184.
- Zumstein, P. P., Luthi, C. & Humbel, R. E. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3169–3172.
- de Pagter-Holthuisen, P., van Schaik, F. M. A., Verduijn, G. M., van Ommen, G. J. B., Bouma, B. M., Jansen, M. & Sussenbach, J. S. (1986) *FEBS Lett.* **195**, 179–184.
- Brissenden, J. E., Ullrich, A. & Franke, U. (1984) *Nature (London)* **310**, 781–783.
- Tricoli, J. V., Rall, L. B., Scott, J., Bell, G. I. & Shows, T. B. (1984) *Nature (London)* **310**, 784–786.
- Bell, G. I., Gerhard, D. S., Fong, N. M., Sanchez-Pescador, R. & Rall, L. B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6450–6454.
- Kittur, S. D., Hoppener, J. W. N., Antonarakis, S. E., Daniels, J. D. J., Meyers, D. A., Maestri, N. E., Jansen, M., Korneluk, R. G., Nelkin, B. D. & Kazazian, H. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5064–5067.
- Dull, I. J., Gray, A., Hayflick, J. S. & Ullrich, A. (1984) *Nature (London)* **310**, 775–777.
- Chirgwin, J. M., Przybyla, A. E. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5298.
- Glisin, V., Crkvenjakov, R. & Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
- Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 263–269.
- Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tozard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727–3731.
- Hanahan, D. (1983) *J. Mol. Biol.* **155**, 557–568.
- Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, J. H. & Roe, B. (1980) *J. Mol. Biol.* **143**, 161–178.
- Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16**, 4743–4751.
- Dobner, P. R., Kislaukis, E., Wentworth, B. M. & Villa-Komaroff, L. (1987) *Nucleic Acids. Res.* **15**, 199–218.
- O'Farrell, P. H., Kutter, E. & Nakanishi, M. (1980) *Mol. Gen. Genet.* **179**, 421–435.
- Bell, G. I., Merryweather, J. P., Sanchez-Pescador, R., Stempien, M. M., Priestly, L., Scott, J. & Rall, L. B. (1984) *Nature (London)* **310**, 775–777.
- Jansen, M., van Shaik, F. M. A., van Tol, H., Vanden Brande, J. L. & Sussenbach, J. S. (1985) *FEBS Lett.* **179**, 243–246.
- Reeve, R. E., Eccles, M. R., Wilkins, R. J., Bell, G. I. & Millow, L. J. (1985) *Nature (London)* **317**, 258–260.
- Scott, J., Cowell, J., Robertson, M. E., Priestly, L. M., Wadey, R., Hopkins, B., Pritchard, J., Bell, G. I., Rall, L. B., Graham, C. F. & Knott, T. J. (1985) *Nature (London)* **317**, 260–262.
- Spector, D. H., Villa-Komaroff, L. & Baltimore, D. (1975) *Cell* **6**, 41–44.
- Vournakis, J. N., Efstratiadis, A. & Kafatos, F. C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2959–2963.
- Soares, M. B., Ishii, D. N. & Efstratiadis, A. E. (1985) *Nucleic Acids Res.* **13**, 1119–1134.
- Schibler, U., Hagenbuchle, O., Wellauer, P. K. & Pittet, A. C. (1983) *Cell* **33**, 501–508.
- Haselbacher, G. K., Irminger, J.-C., Zapf, J., Ziegler, W. H. & Humbel, R. E. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1104–1106.
- Bothwell, M. (1982) *J. Neurosci. Res.* **8**, 225–231.
- Lenoir, D. & Honegger, P. (1983) *Dev. Brain Res.* **7**, 205–213.
- Recio-Pinto, E. & Ishii, D. N. (1984) *Brain Res.* **302**, 323–334.
- Schoenle, E., Haselbacher, G., Briner, J., Janzer, R. C., Gammeltuft, S., Humbel, R. E. & Prader, A. (1986) *J. Pediatr.* **108**, 737–740.