

Isolation of the human insulin-like growth factor genes: Insulin-like growth factor II and insulin genes are contiguous

(gene family/human chromosome 11/diabetes mellitus/Beckwith-Wiedemann syndrome)

GRAEME I. BELL*, DANIELA S. GERHARD†, NOEL M. FONG*, RAY SANCHEZ-PESCADOR*,
AND LESLIE B. RALL*

*Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608; and †Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Donald F. Steiner, June 3, 1985

ABSTRACT Overlapping recombinant clones that encompass the insulin-like growth factor (IGF) I and II genes have been isolated from a human genomic DNA library. Each gene is present once per haploid genome; the IGF-I gene spans >35 kilobase pairs (kbp) and the IGF-II gene is at least 15 kbp. The exon-intron organization of these genes is similar, each having four exons, which is one more than the related insulin gene. Comparison of the restriction endonuclease cleavage maps of the IGF-II and insulin genes, including their flanking regions and hybridization with an IGF-II cDNA probe, revealed that they are adjacent to one another. The IGF-II and insulin genes have the same polarity and are separated by 12.6 kbp of intergenic DNA that includes a dispersed middle repetitive *Alu* sequence. The order of the genes is 5'-insulin-IGF-II-3'.

The insulin-like growth factors (IGFs) and insulin are related polypeptides that have a high degree of sequence homology and exhibit a similar spectrum of biological activities (1-4). They produce rapid metabolic changes and have long-term growth promoting effects as well. However, the relative concentrations required to elicit these responses are different. In general, insulin is more potent in producing short-term metabolic effects, while the IGFs are more potent in promoting growth. Complementary DNAs encoding the precursors to human insulin, IGF-I and IGF-II, have been isolated and sequenced (5-8). The chromosomal location of the human insulin (*INS*), IGF-I (*IGF1*), and IGF-II (*IGF2*) genes has also been determined (refs. 9-15; unpublished data). The human *IGF1* gene has been mapped to the region of bands q22-q24.1 of chromosome 12. The *INS* and *IGF2* genes have both been localized to band p15 of chromosome 11, a region of ≈ 20 cM (20×10^6 base pairs), which also includes the genes for the β -globin complex, parathyroid hormone, and the oncogene *c-Harvey-ras1* (15). In this report, we describe the isolation and partial characterization of the human *IGF1* and *IGF2* genes and demonstrate that the *IGF2* and *INS* genes are contiguous. The gene order is 5'-*INS-IGF2*-3' and they have the same polarity. The 5' end of the *IGF2* gene is within 12.6 kilobase pairs (kbp) of the 3' end of the *INS* gene and the genes are separated by an *Alu* sequence.

METHODS

Isolation of Human *INS*, *IGF1*, and *IGF2* Genes and Flanking Regions. Bacteriophage λ containing DNA segments encoding the *INS*, *IGF1*, and *IGF2* genes were isolated from the partial *Hae* III/*Alu* I fetal human liver DNA library in λ CH4A of Lawn *et al.* (16), which was provided by T. Maniatis. The isolation and characterization of the human

INS gene and flanking regions has been described (17-21). The λ phage, λ hINS-2, was isolated by hybridization with nick-translated (22) human insulin cDNA, pch1 (5). Phage λ hINS-5 and -3 were isolated using the 2.2- and 1.8-kbp *Eco*RI fragments, respectively, of λ hINS-2 as probes. The *IGF1* gene-containing phage were isolated by hybridization (21) with the human IGF-I cDNA, phigf1 (7). λ hIGF1-4 was isolated from a λ EMBL4 (23) library of partial *Sau*3A fragments of DNA from the lymphoblastoid cell line GM1416, which has the karyotype 48,XXXX. The *IGF2* gene-containing phage were isolated by hybridization with the human IGF-II cDNA, phigf2 (7). Phage were plaque-purified and DNA was isolated by standard procedures (24). The order of the *Eco*RI fragments in the human DNA inserts was determined by single and double digests of the phage DNA as well as by digests of restriction fragments eluted from low-melting-point agarose gels (24). The locations of the exons and repeated sequences were determined by restriction mapping, blotting, and hybridization with nick-translated cDNA and human DNA probes, respectively. As phigf1 lacked sequences corresponding to the 5' untranslated region, an oligonucleotide (5' TAATTGGGTTGGAAGACTGC 3') derived from the sequence of a longer IGF-I cDNA reported by Jansen *et al.* (6) was chemically synthesized for use as a probe to locate this region of the *IGF1* gene.

DNA Sequencing. The sequence of the exons and adjacent introns of the *IGF1* and *IGF2* genes was determined by using the dideoxy method (25) after subcloning appropriate restriction fragments into the vectors M13 mp10 and mp11 (26). The positions of the exon-intron junctions were assigned by comparing the cDNA and gene sequences and using the "GT-AG" rule (27).

Southern Blot Analysis. One microgram of recombinant phage DNA was digested with *Eco*RI and the fragments were separated by electrophoresis in a horizontal 1% agarose gel. The denatured DNA fragments were transferred to a nitrocellulose filter (28) and hybridized with a 32 P-labeled (22) 833-bp *Pst* I fragment of phigf2 that contains the 5' untranslated and protein coding regions and 21 bp of 3' untranslated sequence. The filter was preannealed and hybridized at 42°C in a solution of 5 \times NaCl/Cit (1 \times NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate)/50% (vol/vol) formamide/50 mM sodium phosphate, pH 6.8/5 \times Denhardt's solution (1 \times Denhardt's solution is 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/250 μ g of sonicated and denatured salmon testes DNA per ml/0.1% NaDodSO₄. The 32 P-labeled probe (specific activity, 2×10^8 cpm/ μ g) was present at 2.5×10^5 cpm/ml. Before autoradiography, the filter was washed at room temperature for 30 min in 0.1 \times NaCl/Cit/0.1% NaDodSO₄ and then at 50°C for 60 min.

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: kbp, kilobase pair(s); IGF, insulin-like growth factor.

Isolation of RNA and RNA Blot Analysis. RNA was isolated from adult human liver and kidney by using the guanidinium thiocyanate/LiCl procedure (29). Poly(A)⁺ RNA was prepared by chromatography on an oligo(dT) cellulose column (30). Ten micrograms of poly(A)⁺ RNA was denatured with glyoxal (31, 32) and, after electrophoresis through a 1% agarose gel, was transferred to a nitrocellulose membrane (32). ³²P-labeled fragments of a *Hind*III digest of λ DNA were included in an adjacent lane to serve as size markers. The nitrocellulose filter was hybridized (21) with ³²P-labeled cDNA inserts from the plasmids phigf1 and phigf2.

RESULTS

Isolation and Exon–Intron Organization of the Human *IGF1* and *IGF2* Genes. The *IGF1* and *IGF2* genes were isolated as overlapping DNA fragments (Figs. 1 and 2). The locations of the exons in each gene were determined by restriction mapping, and then each exon was sequenced. The *IGF1* gene encompasses >35 kbp and there is a gap of unknown size between exons 2 and 3. Two different λ phage libraries and one cosmid library were screened with the IGF-I cDNA probe and none contained a DNA segment that overlapped the inserts in λhIGF1-3 and -4. The *IGF2* gene, which is at least 15 kbp, is smaller than the *IGF1* gene. As the transcriptional start and polyadenylation sites have not been rigorously determined for either gene, the extent of exons 1 and/or 4 in both cases is probably greater than indicated (Figs. 1 and 2). All the *Eco*RI fragments observed in a Southern blot of human genomic DNA that hybridize with IGF-I or -II cDNA probes are accounted for in the linkage map of these genes (data not shown). Thus, *IGF1* and *IGF2* are single-copy genes in humans.

The exon–intron organization of the *IGF1* and *IGF2* genes is similar (Figs. 1 and 2; Table 1) and each has four exons, which is one more than the related *INS* gene (17) whose exons correspond to 1–3 of the *IGF* genes. Although the introns in the IGF genes are in similar positions, those in the *IGF1* gene are much larger. The first and second introns of the *IGF1* and *IGF2* genes are located in the 5' untranslated region and near the end of the B domain, respectively, and when the sequences are aligned for maximum homology they are in identical positions. The third intron, which is in the COOH-terminal E domain of each precursor is displaced by 15 bp in the two genes. The introns in the human *INS* gene (17) are in similar but not identical positions to the first two of the *IGF* genes; the first is in the 5' untranslated region 17 bp before the start of translation and the second is in the region of the C peptide and thus has shifted slightly during the evolution of the *INS* and *IGF* genes. All of the published IGF-I and -II

cDNA sequence (6–8) is present in exons 1–4 of these genes. In addition, with one exception, the cDNA and gene sequences are identical. There is a silent change in the codon for glycine-42 of IGF-I, which is GGC in the gene and GGT in the cDNA sequence. Dull *et al.* (33) deduced the organization of the human *IGF2* gene based on its nucleotide sequence homology with a rat IGF-II cDNA. The positions of exons 2–4 are the same in their study and in this report. However, their location for exon 1 does not agree with our assignment, which is based on the comparison of the sequence of the gene with that of the 5' untranslated region of a human IGF-II cDNA. Ullrich *et al.* (34) have also isolated a segment of the human IGF-I gene that encodes exon 2.

Heterogeneity of IGF-I and IGF-II mRNA. Hybridization of human liver IGF-I and -II cDNAs to RNA blots of poly(A)⁺ RNA from adult human liver and kidney revealed multiple forms for each mRNA and also indicated that there is tissue-specific processing of IGF-II mRNA (Fig. 3). The insert from the IGF-I cDNA, phigf1, which is 607 bp exclusive of a poly(dA) tract of 18 bp, hybridized to liver mRNAs of 900, 5300, and 7700 bases and to kidney mRNAs of 5300 and 7700 bases (the 900-base transcript may also be in the kidney RNA preparations; however, because of its low abundance, more sensitive hybridization procedures would be required to detect it). The insert from the IGF-II cDNA, phigf2, which is 1046 bp and lacks a poly(dA) tract, hybridized to a liver mRNA of 5300 bases and to kidney mRNAs of 4900 and 6000 bases. The large size of the major IGF-I and -II transcripts is surprising because preproIGF-I and -II are encoded by 390 and 540 bases, respectively. It is unknown if the additional sequence present in these large mRNAs is 5' or 3' untranslated. Consequently, exons 1 and/or 4 of the *IGF1* and *IGF2* genes are most likely larger than indicated in Figs. 1 and 2.

Linkage of the Human *IGF2* and *INS* Genes. Linkage of the *IGF2* and *INS* genes was first suggested because of hybridization of λhIGF2-1 DNA and the 8.6-kbp *Eco*RI fragment of λhINS-3 to restriction fragments of similar size when they were used as probes in a search for polymorphic restriction sites. Inspection of the inserts in λhIGF2-1 and λhINS-3 revealed the presence of three *Eco*RI fragments of similar size and order (Fig. 2). The linkage of the *IGF2* and *INS* genes was confirmed by hybridization of ³²P-labeled phigf2 insert to *Eco*RI fragments from each of the λhIGF2 clones as well as from λhINS-3 (Fig. 4). The map of the *INS* and *IGF2* genes spans ≈60 kbp (Fig. 2) and the *INS* and *IGF2* genes have the same polarity. Although the boundaries of the *INS* gene have been determined (17), those of the *IGF2* gene are unknown and the 5' end of the *IGF2* gene may be nearer the *INS* gene than indicated. The data show, however, that the 5'-end of

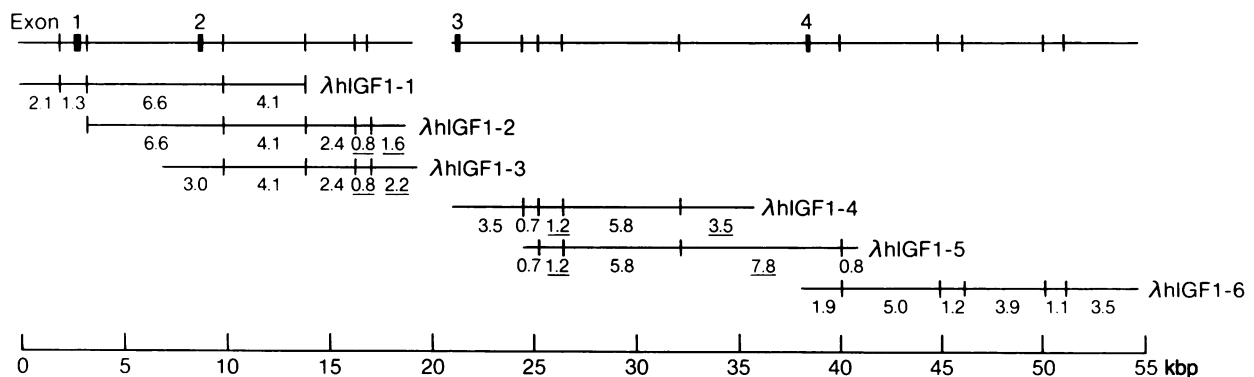


Fig. 1. Map of the human *IGF1* gene. The structures of six λ phage containing overlapping segments of DNA are indicated. The natural *Eco*RI sites are indicated by vertical lines. The sizes (in kbp) of the *Eco*RI fragments in each cloned DNA segment are given and those that hybridize with ³²P-labeled human DNA are underlined. The positions of the exons are noted by the filled boxes. The left-hand 3.5-kbp *Eco*RI fragment of λhIGF1-4 hybridizes to a genomic 4.2-kbp *Eco*RI fragment.

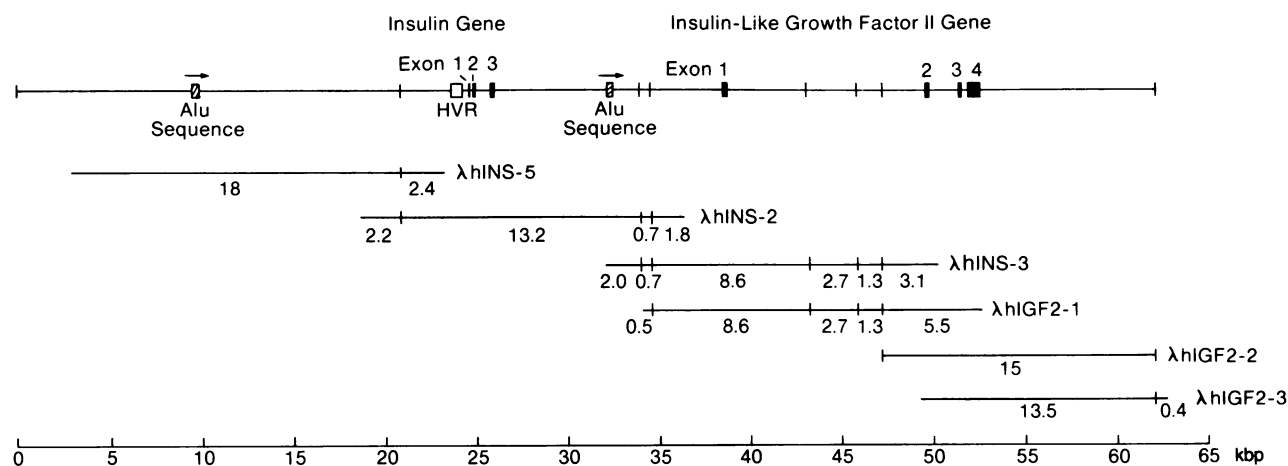


FIG. 2. Map of the human IGF-II and insulin gene region. The structures of six λ phage containing overlapping DNA segments are indicated. The natural *Eco*RI sites are indicated by vertical lines. The sizes (in kbp) of the *Eco*RI fragments in each cloned DNA segment are given. The positions of the exons are noted by the filled boxes, and the hypervariable region (HVR) flanking the insulin gene is shown by an open box. The positions and polarity of the two *Alu* sequences, which are the only repetitive sequences in this region, are indicated. The size of the \approx 22-kbp left-hand *Eco*RI fragment in the map was determined by hybridization of the 2.2-kbp fragment of λ HINS-2 to an *Eco*RI digest of human genomic DNA. The order of the 2.7- and 1.3-kbp fragments of λ HINS-3 were inverted in a previous report (20).

the *IGF2* gene is within 12.6 kbp of the 3' end of the *INS* gene. The only other genetic elements that we have detected in this 60-kbp segment are two dispersed middle repetitive *Alu* sequences (ref. 18; unpublished observations) that have the same polarity [the poly(dA) tract is at the 3' end (35)] as the *INS* and *IGF2* genes. Although in this instance these *Alu* sequences could represent gene boundaries, this is not the only function of *Alu* sequences, as in other genes they are present within introns as well as in the 3' untranslated portion of the mRNA (35–37).

DISCUSSION

The insulin family includes the insulin, relaxin, and IGF genes (38, 39). Relaxin is synthesized by the ovary and promotes the remodeling of the female reproductive tract that occurs prior to birth (40). Whereas the IGFs have to date only been described from mammals, insulin is probably present in

all vertebrates (41) and relaxin has been isolated from both fish and mammals (40, 42). In humans, *INS*, *IGF1*, and *IGF2* are single-copy genes, whereas there are two nonallelic relaxin (*REL*) genes (43). Interestingly, in mice, rats, and some fish, there are two nonallelic insulin genes (41). Furthermore, in humans the *INS* and *IGF2* genes are closely linked on chromosome 11 (9–11, 13–15), but the *IGF1* and *REL* genes are on different chromosomes (12 and 9, respectively) (10–12, 43). Comparison of the amino acid sequence homology between members of the insulin family indicates that the lines that gave rise to insulin and relaxin diverged first, and then the IGFs diverged from the insulin line (38). The ancestral gene was probably like the insulin gene and had three exons: exon 1, encoding the 5' untranslated region; exon 2, the signal peptide, a B chain-like-peptide and part of a C peptide-like peptide; and exon 3, the remainder of the C peptide-like peptide and an A chain-like peptide and 3'

Table 1. Exon–intron organization of the human IGF genes

Exon	Exon size, bp	5' splice donor	3' splice acceptor	Intron size, kbp
IGF-I				
1	>72	5' -UT CTTGAAGgtaaattttcttgggcagGTGAAG	-25 Met ATG	6
2	157	Tyr Phe A TAT TTC A gtaagt...ttgattgtgcag	27 sn Lys Pro AC AAG CCC	>10
3	182	Thr Gln Lys ACC CAG AAG gtaagc...ctttatttatag	87 Glu Val His GAA GTA CAT	17
4	307			
IGF-II				
1	>240	5' -UT CACAGAGgtgagt...tcccgccccagACACCA	-24 Met ATG	11
2	163	Tyr Phe S TAC TTC A gtaagt...tgggacttccag	30 er Arg Pro GC AGG CCC	1.7
3	149	Val Leu Pro GTG CTT CCG gtgagg...atcccctcccag	79 Asp Asn Phe GAC AAC TTC	0.3
4	>490			

The position at which the intron interrupts the IGF-I and -II mRNA (6, 7) is indicated. Exon sequences are in capital letters; intron sequences are in lowercase letters. Each exon was sequenced in its entirety. The sizes of the introns were estimated from restriction mapping. UT, untranslated region.

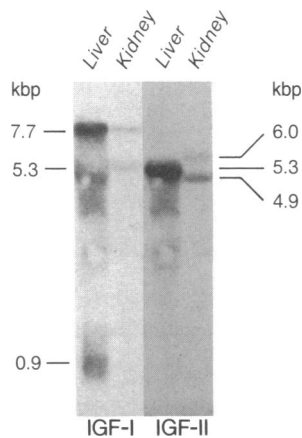


FIG. 3. Hybridization of ³²P-labeled human IGF-I and -II cDNAs to human liver and kidney poly(A)⁺ RNA. The sizes of the hybridizing transcripts are indicated. Each lane contained 10 μg of liver or kidney poly(A)⁺ RNA.

untranslated region. The IGF genes acquired exon 4, which encodes the COOH-terminal or E domain of the precursor, after they diverged from the insulin line. Moreover, because of the absence of homology between exon 4 of the *IGF1* and *IGF2* genes, this exon was probably acquired independently by these genes. Since both IGF-I and -II possess 47% identity with insulin (38), it is possible that the primordial *IGF* genes arose at the same time. In the case of the IGF-I gene, there was a duplication and simultaneous translocation to another chromosome, whereas the IGF-II gene remained closely linked with the gene from which it arose. Subsequently, the *IGF* genes acquired a fourth exon.

IGF-I and -II mRNA is synthesized by many, if not all tissues, including the liver, heart, lung, kidney, pancreas,

spleen, small intestine, colon, brain, and pituitary, although its abundance in each tissue varies (unpublished results). In contrast, insulin mRNA is only synthesized by the beta cells of the endocrine pancreas. Thus, even though they are contiguous and within 12.6 kbp of one another, the changes in chromatin structure that must occur to facilitate transcription of the IGF-II gene do not result in activation of the adjacent insulin gene. Indeed, transcription of the insulin gene may require factors that are present only in the beta cell (44). It is unknown whether expression of the IGF-II gene is constitutive or if it also requires specific transcriptional factors.

The 5' flanking region of the human insulin gene contains a hypervariable region (18–21) (Fig. 2) that is composed of variable numbers of a tandemly repeating 14- to 15-bp sequence (19). The high polymorphic content of this region makes it a useful genetic marker for closely linked genes. As the hypervariable region is within 15 kbp of the *IGF2* gene, it may be useful in genetic studies of fetal growth disorders, which may be a consequence of abnormal expression of the *IGF2* gene. For example, Beckwith–Wiedemann or Exomphalos–Macroglossia–Gigantism syndrome has been associated with trisomy of the p15 region of chromosome 11 (45, 46) and thus may be due to increased levels of IGF-II in fetal tissues. Various alleles of the insulin gene hypervariable region have been found to occur more frequently in Caucasians affected with insulin-dependent diabetes mellitus (20), atherosclerosis (47), or diabetic hypertriglyceridemia (48), than in appropriate control groups. Its association with non-insulin-dependent diabetes mellitus has also been reported (49) but has not been confirmed when larger groups of racially separated individuals were examined (20, 50, 51). The association of various features of this polymorphism with three different disorders suggests that the hypervariable region is in linkage disequilibrium with an unidentified predisposing gene(s). The *IGF2* gene is not known to be involved in the etiology of insulin-dependent diabetes mellitus, atherosclerosis, or diabetic hypertriglyceridemia. However, given the apparent importance of IGF-II for the growth and maintenance of cells (3, 4), and the occurrence of variant IGF-II polypeptides (8) to which cells may respond in an abnormal fashion, the role of the *IGF2* gene in the etiology of each of these disorders should be examined.

We thank Ms. Diana Wyles for her patience and assistance in the preparation of this manuscript, Ms. Michelle Stempien for her expert technical assistance, and Drs. James Merryweather and Martha Truett for providing the oligonucleotide probe and λEMBL4 human DNA library, respectively. D.S.G. would like to thank Dr. David E. Housman for his encouragement and support. D.S.G. was supported by Damon Runyon–Walter Winchell Cancer Fund Fellowship DRG-665, and by Grants GM 27882 and CA 17575 to D. E. Housman.

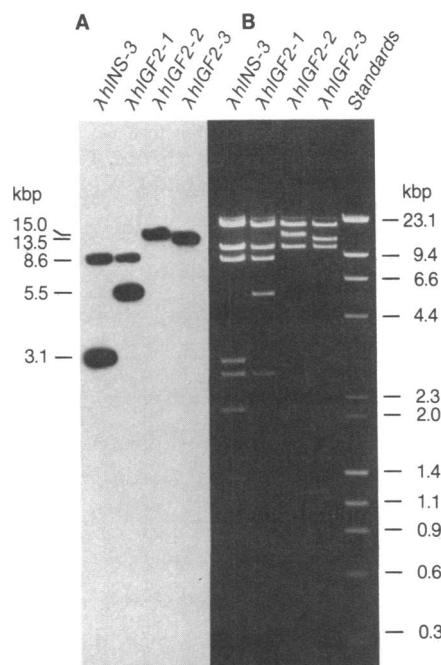


FIG. 4. The human IGF-II and insulin genes are contiguous. (A) Hybridization of the *Pst* I insert of the human IGF-II cDNA, phig2, to *Eco*RI digests of λ clones containing the 3' flanking region of the insulin gene and the *IGF2* gene. The sizes of the hybridizing fragments are indicated. (B) Photograph of the ethidium bromide-stained gel that was analyzed in A, showing the pattern of *Eco*RI fragments present in each phage. The standards are *Hind*III and *Hae* III fragments of λ and φX174 DNA, respectively.

- Rinderknecht, E. & Humbel, R. E. (1978) *J. Biol. Chem.* **253**, 2769–2776.
- Rinderknecht, E. & Humbel, R. E. (1978) *FEBS Lett.* **89**, 283–286.
- Clemmons, D. R. & Van Wyk, J. J. (1981) in *Tissue Growth Factors*, ed. Baserga, R. (Springer, Berlin), pp. 161–208.
- Hall, K. & Sara, V. R. (1983) *Vitam. Horm. (N.Y.)* **40**, 175–233.
- Bell, G. I., Swain, W. F., Pictet, R., Cordell, B., Goodman, H. M. & Rutter, W. J. (1979) *Nature (London)* **282**, 525–527.
- Jansen, M., van Schaik, F. M. A., Ricker, A. T., Bullock, B., Woods, D. E., Gabbay, K. H., Nussbaum, A. L., Susenbach, J. S. & Van den Brande, J. L. (1983) *Nature (London)* **306**, 609–611.
- Bell, G. I., Merryweather, J. P., Sanchez-Pescador, R., Stempien, M. M., Priestley, L., Scott, J. & Rall, L. B. (1984) *Nature (London)* **310**, 775–777.

8. Jansen, M., van Schaik, F. M. A., van Tol, H., Van den Brande, J. L. & Sussenbach, J. S. (1985) *FEBS Lett.* **179**, 243–246.
9. Owerbach, D., Bell, G. I., Rutter, W. J. & Shows, T. B. (1980) *Nature (London)* **286**, 82–84.
10. Tricoli, J. V., Rall, L. B., Scott, J., Bell, G. I. & Shows, T. B. (1984) *Nature (London)* **310**, 784–786.
11. Brissenden, J. E., Ullrich, A. & Franke, U. (1984) *Nature (London)* **310**, 781–784.
12. Hoppener, J. W. M., de Pagter-Holthuisen, P., Geurts van Kessel, A. H. M., Jansen, M., Kittur, S. D., Antonarakis, S. E., Lips, C. J. M. & Sussenbach, J. S. (1985) *Hum. Genet.* **69**, 157–160.
13. de Pagter-Holthuisen, P., Hoppener, J. W. M., Jansen, M., Geurts van Kessel, A. H. M., van Ommen, G. J. B. & Sussenbach, J. S. (1985) *Hum. Genet.* **69**, 170–173.
14. Harper, M. E., Ullrich, A. & Saunders, G. F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4458–4460.
15. Zabel, B. U., Kronenberg, H. M., Bell, G. I. & Shows, T. B. (1985) *Cytogenet. Cell Genet.*, in press.
16. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
17. Bell, G. I., Pictet, R. L., Rutter, W. J., Cordell, B., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **284**, 26–32.
18. Bell, G. I., Pictet, R. & Rutter, W. J. (1980) *Nucleic Acids Res.* **8**, 4091–4109.
19. Bell, G. I., Selby, M. J. & Rutter, W. J. (1982) *Nature (London)* **295**, 31–35.
20. Bell, G. I., Horita, S. & Karam, J. H. (1984) *Diabetes* **33**, 176–183.
21. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5759–5763.
22. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
23. Karn, J., Brenner, S. & Barnett, L. (1983) *Methods Enzymol.* **101**, 3–19.
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
25. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143**, 161–178.
26. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
27. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4853–4857.
28. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
29. Cathala, P., Savouret, J., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) *DNA* **2**, 329–335.
30. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
31. Carmichael, G. C. & McMaster, G. K. (1980) *Methods Enzymol.* **65**, 380–391.
32. Thomas, P. S. (1983) *Methods Enzymol.* **100**, 255–266.
33. Dull, T. J., Gray, A., Hayflick, J. S. & Ullrich, A. (1984) *Nature (London)* **310**, 777–781.
34. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. (1984) *EMBO J.* **3**, 361–364.
35. Jelinek, W. R. & Schmid, C. W. (1982) *Annu. Rev. Biochem.* **51**, 813–844.
36. Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. & Russell, D. W. (1984) *Cell* **39**, 27–38.
37. Lehrman, M. A., Schneider, W. J., Sudhoff, T. C., Brown, M. S., Goldstein, J. L. & Russell, D. W. (1985) *Science* **227**, 140–146.
38. Dayhoff, M. L., ed. (1978) *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, DC), Vol. 5, Suppl. 3, pp. 150–151.
39. Blundell, T. L. & Humbel, R. E. (1980) *Nature (London)* **287**, 781–787.
40. Bryant-Greenwood, G. D. (1982) *Endocr. Rev.* **3**, 62–90.
41. Blundell, T. L. & Wood, S. P. (1975) *Nature (London)* **257**, 197–203.
42. Hudson, P., Haley, J., John, M., Cronk, M., Crawford, R., Haralambidis, J., Tregar, G., Shine, J. & Niall, H. (1983) *Nature (London)* **301**, 628–631.
43. Crawford, R. J., Hudson, P., Shine, J., Niall, H. D., Eddy, R. L. & Shows, T. B. (1984) *EMBO J.* **3**, 2341–2345.
44. Walker, M. D., Edlund, T., Boulet, A. M. & Rutter, W. J. (1983) *Nature (London)* **306**, 557–561.
45. Waziri, M., Patil, S. R., Hanson, J. W. & Bartley, J. A. (1983) *J. Pediatr.* **102**, 873–876.
46. Turieau, C., deGrouchy, J., Chavin-Colin, F., Martelli, H., Voyer, M. & Charias, R. (1984) *Hum. Genet.* **67**, 219–221.
47. Mandrup-Poulsen, T., Mortensen, S. A., Meinertz, H., Owerbach, D., Johansen, K., Sorensen, H. & Nerup, J. (1984) *Lancet* **i**, 250–253.
48. Jowett, N. I., Williams, L. G., Hitman, G. A. & Galton, D. J. (1984) *Br. Med. J.* **288**, 96–99.
49. Rotwein, P. S., Chirgwin, J., Province, M., Knowler, W. C., Pettitt, D. J., Cordell, B., Goodman, H. M. & Permutt, M. A. (1983) *N. Engl. J. Med.* **308**, 65–71.
50. Permutt, M. A., Andreone, T., Chirgwin, J., Elbein, S., Knowler, W. & Rotwein, P. (1984) in *Endocrinology, Proceedings of the 7th International Congress*, eds. Labrie, F. & Proulx, L. (Excerpta Medica, Amsterdam), pp. 245–248.
51. Permutt, M. A., Rotwein, P., Andreone, T., Ward, W. K. & Porte, D. (1985) *Diabetes* **34**, 311–314.