Studies on mutant human insulin genes: Identification and sequence analysis of a gene encoding [Ser^{B24}]insulin

(hyperinsulinemia/diabetes mellitus/autosomal dominant/Mbo II cleavage defect)

MASAKAZU HANEDA, SHU J. CHAN, SIMON C. M. KWOK, ARTHUR H. RUBENSTEIN, AND DONALD F. STEINER

Departments of Biochemistry and Medicine, The University of Chicago, Chicago, Illinois 60637

Contributed by Donald F. Steiner, June 27, 1983

ABSTRACT Both alleles of the insulin gene of a patient with mild diabetes [maturity-onset-diabetes-of-the-young (MODY)-type syndrome] associated with hyperinsulinemia have been cloned, and the sequences have been determined. One allele contained a mutation (single nucleotide transition) in the coding sequence for the B chain at position 24 (TTC \rightarrow TCC), resulting in the loss of a restriction enzyme (*Mbo* II) cleavage site in the gene. This mutation results in the substitution of serine for phenylalanine in a critically important region of the insulin molecule that is intimately involved in receptor binding. Both insulin alleles were of the α type and, aside from a single nucleotide deletion in the 5' region of the normal allele, their sequences were identical to those previously determined.

Diabetes is a heterogeneous group of metabolic disorders characterized by hyperglycemia and a relative or absolute deficiency of insulin. A strong genetic component is particularly evident in the more common noninsulin-dependent or maturity-onset form of the disease (type II diabetes), but the nature of the underlying genetic factor or factors responsible for the disorder remains obscure. Analysis of the inheritance pattern in affected families is complicated by the relatively late onset of the disease in many individuals and the probable existence of secondary contributing factors, which may be either acquired or genetic (1, 2). Despite these complexities, it is important to establish whether defects residing in or near the insulin gene may contribute to the relative insulin deficiency in this group of patients.

We have identified three individuals with mild diabetes associated with increased plasma insulin levels and have shown that these patients all have abnormal circulating insulin molecules that can be distinguished from each other and from normal human insulin (3). These abnormal insulins exhibited both impaired biological activity and reduced rates of clearance from the circulation (4-6). In one of these cases, insulin isolated from the patient's pancreas consisted of roughly equal proportions of normal insulin and an insulin of low biological potency, having a leucine-for-phenylalanine substitution at position 25 in the B chain (3, 7). This substitution resulted in the loss of a normal restriction enzyme cleavage site in one allele of the patient's insulin gene in the coding region for the B chain at the phenylalanine-phenylalanine sequence at positions B24 and B25 (8). The codons for these two amino acids contain a pentanucleotide recognition sequence for the enzyme Mbo II. Cloning and nucleotide sequence analysis of the insulin gene of this patient resulted in the identification of a single nucleotide transversion in the codon for phenylalanine-B25, changing it to a codon specifying leucine (9).

We have now studied in greater detail another affected individual (J.F.) and her family. This patient had fasting hyperinsulinemia without any resistance to exogenously administered insulin (5). Similar fasting hyperinsulinemia was demonstrated in five additional family members of both sexes over three successive generations, indicating an autosomal dominant pattern of inheritance. Insulin isolated from the patient's serum by immunoaffinity chromatography had markedly reduced biological activity as compared to its immunoreactivity (5). HPLC analysis of the serum insulin revealed a major component that was less hydrophobic than native human insulin and only small amounts of the normal hormone (3). Furthermore, restriction endonuclease cleavage analysis of genomic DNA isolated from the patient's leukocytes revealed the apparent loss in one allele of the insulin gene of the same Mbo II recognition site that had been found to be defective in the first patient (3, 5). However, this abnormal insulin could be distinguished clearly by HPLC from the [Leu^{B25}]insulin found in the first case (3). In order to define the nature of this mutation, we cloned and determined the nucleotide sequences of both alleles of the insulin gene of this patient. The results demonstrate that a point mutation has occurred in the codon for position B24 resulting in the substitution of serine for phenylalanine at this site.

MATERIALS AND METHODS

Preparation of Genomic DNA. Genomic DNA was prepared from the leukocytes from 20–50 ml of blood as described (8). In brief, the leukocytes were digested with proteinase K (100 μ g/ml) in the presence of 100 mM EDTA, pH 8.0/0.5% NaDodSO₄ at 55°C overnight and were extracted three times with phenol/chloroform/isoamyl alcohol, 25:24:1 (vol/vol), and twice with chloroform; the DNA was dialyzed extensively against 10 mM Tris·HCl, pH 7.5/1 mM EDTA. DNA was then digested with ribonuclease A (100 μ g/ml) at 37°C for 2 hr and with proteinase K at 55°C for 2 hr, extracted twice with the phenol solution and once with chloroform, and then dialyzed against 1 mM Tris·HCl, pH 7.5/0.1 mM EDTA. DNA concentration was calculated from the absorbancy at 260 nm, and the yields were \approx 10 μ g per ml of blood.

Restriction Endonuclease Cleavage Analysis. Genomic DNA (30 μ g) was digested with 60 units (unless otherwise specified) of restriction endonuclease at 37°C overnight (≈16 hr) under the conditions suggested by the manufacturers. Digested DNA was separated on an 0.8 or 2% agarose gel and transferred to nitrocellulose filters by the method of Southern (10). The filter was incubated with 14 ml of prehybridizing solution containing $3 \times \text{NaCl/Cit}$ (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate, pH 7), 10× Denhardt solution (1× is 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.2% bovine serum albumin), 50

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: bp, base pair(s); kb, kilobase(s).

mM sodium phosphate, alkali-denatured salmon testis DNA at 200 μ g/ml, poly(A) at 10 μ g/ml, poly(C) at 10 μ g/ml, and 0.1% NaDodSO₄ at 65°C overnight and then hybridized with 2×10^6 cpm of human insulin cDNA labeled with ³²P by nick-translation (11) in 5 ml of the above solution at 65°C for 3 days. After hybridization, the filter was washed with 0.1× NaCl/Cit at 65°C, dried, and exposed to x-ray film at -70° C for 7 days with an intensifying screen (DuPont Cronex Lightning Plus).

Molecular Cloning of the Insulin Genes. Genomic DNA (270 μ g) was digested with *Eco*RI at 37°C overnight and was separated on an 0.8% agarose gel. Three 5-mm slices covering the size ranges 16.5–23.5 (fragment A), 12.3–16.5 (fragment B), and 9.8–12.3 kilobases (kb) (fragment C) were cut from gel and isolated by electroelution (12); 1- μ g aliquots of each fragment were applied to an 0.8% agarose gel again, along with a marker lane of 30 μ g of *Eco*RI-digested genomic DNA. The autoradiogram of the Southern blot showed that fragment B contained the insulin gene, and this fragment was used for cloning.

 λ gtWES λB phage was grown, and DNA was isolated as described (13). The phage DNA was digested with EcoRI, and the phage arms were prepared by sucrose gradient centrifugation, dialyzed, and precipitated with ethanol (13). The purified EcoRI fragment containing the insulin gene was ligated to the phage arms in equimolar ratios; 1 μ g of arms and 0.65 μ g of DNA fragment B were hybridized at 42°C for 2 hr in 8 μ l of 1.25× ligase buffer (12). After cooling slowly, 1 μ l of 10 mM ATP, 1 μ l of 10 mM dithiothreitol, and 0.5 μ l of T4 ligase (2 units) were added, and the mixture was incubated at 9°C overnight. The ligated DNA was then packaged in vitro (14), yielding $\approx 5 \times 10^{5}$ recombinant phages per μg of arms. The packaged phage was infected into Escherichia coli LE392, and phage plaques were screened by an *in situ* plaque hybridization method (15) with ³²P-labeled and cloned human insulin cDNA as a probe. Positive plaques were picked and purified by screening at low density through two more cycles. DNA from selected clones were isolated from 400-ml lysates and analyzed by restriction endonuclease mapping. An Xho I fragment containing the insulin gene was then subcloned into the Sal I site of pBR322, and the nucleotide sequence was determined by the method of Maxam and Gilbert (16).

RESULTS

In the characterization of the gene encoding mutant [Leu^{B25}]insulin, the substitution of leucine for phenylalanine at residue 25 of the B chain resulted in the loss of a Mbo II site in the coding region, which was easily detected by the appearance of a band of ~900 base pairs (bp) on Southern blot hybridization (8). Similarly, when leukocyte DNA from patient J.F. was analyzed with Mbo II, a 920-bp band was detected in addition to the expected 1,600-, 580-, and 340-bp fragments. The data shown in Fig. 1 indicate that even prolonged digestion with excess Mbo II did not fully cleave normal DNA at this site, possibly due to modifications in the DNA. Nonetheless, under these conditions we were able to clearly distinguish the patient's cleavage pattern from that of the normal. This result suggested that the patient is heterozygous, with both a normal and abnormal allele in the insulin locus, and indicated that the mutation resides, at least in part, in the codons for residues B24 or B25.

In order to determine the sequence of the mutant allele, we cloned both alleles of the insulin gene in bacteriophage λ gtWES λ B. For this purpose genomic DNA was digested with *EcoRI*, and a 14-kb size fraction containing the insulin genes was partially purified by agarose gel electrophoresis, ligated to *EcoRI* arms of λ gtWES λ B, and packaged *in vitro*. After 5 \times 10⁵ plaques were screened, two recombinants containing the abnormal insulin allele and two containing the normal insulin



FIG. 1. Demonstration of the loss of a restriction endonuclease *Mbo* II cleavage site in genomic DNA from patient J.F.; 30 μ g of genomic DNA from a normal subject and from the patient were digested with various units (U) of *Mbo* II at 37°C as indicated. Digestion was carried out for 16 hr with half of the stated amount of enzyme and was continued for an additional 8 hr after addition of the remaining amount. Digested DNA was analyzed as described; ϕ X174 DNA digested with *Hae* III was used as a molecular weight standard.

allele as identified by the *Mbo* II hybridization pattern were obtained (Fig. 2). Subsequently, phage DNA from one recombinant of each allele was isolated, and ≈ 5.0 -kb *Xho* I fragments containing the insulin gene were subcloned into pBR322 for sequence analysis (see the restriction map in Fig. 3).

Within the coding regions, the DNA sequences of both the normal and abnormal alleles were in agreement with published results (18, 19) except in the codon for amino acid B24 of the abnormal allele. Here a single nucleotide transition was found (TTC \rightarrow TCC) that would result in the production of an abnormal insulin having a serine-for-phenylalanine substitution at this position (Fig. 4). Outside the coding region, sequence data established that both genes contained three out of the four nucleotide changes associated with the α -type allele as described by Ullrich *et al.* (18) (the fourth nucleotide change of the α -type



FIG. 2. Mbo II cleavage analysis of the cloned insulin gene alleles. Phage DNA (1 μ g) from each clone was digested with 10 units of Mbo II at 37°C for 4 hr. Digested DNA was electrophoresed, blotted, and hybridized to human insulin cDNA as described.



FIG. 3. Restriction map and sequence analysis strategy for the normal and mutant human insulin genes. DNA from phage recombinants with the 14-kb EcoRI fragment containing the normal or mutant insulin gene (Fig. 2) were digested with Xho I, and the 4.8- to 4.9-kb fragment was subcloned into the Sal I site of pBR322 for restriction mapping and sequence analysis. The length of the polymorphic insert was estimated based on the size of restriction fragments generated with Pst I, Nco I, Bgl I, and Bgl II and by comparison with a representative sequence published by Bell et al. (17). Horizontal arrows indicate segments with sequences determined (5'-to-3' direction) by the Maxam-Gilbert procedure (16). Two allelic forms of the insulin gene, which differ at four positions, have been identified (18), and sequence analysis of both the normal and mutant alleles showed identity with the α type in three sites (circled). The sequence of the fourth site within the second intervening sequence was not determined. Remaining segments sequenced in both the normal and mutant gene were identical and in agreement with published results (18, 19) with two exceptions: (i) in the mutant gene, there is a nucleotide transition (T \rightarrow C) that changes the codon of residue B24 from phenylalanine to serine and (ii) there is a nucleotide substitution/deletion 45 nucleotides upstream from the "TATA" box in the 5' flanking region.

gene occurs in the second intron, the sequence of which was not determined in this work) (Fig. 3). Restriction mapping showed that both alleles contain small inserted sequences in the 5' polymorphic region (20) of \approx 530 bp for the normal and 400 bp for the abnormal alleles (see Fig. 3). Of the remaining regions, the only other variation in sequence from the published



FIG. 4. Demonstration of nucleotide sequence change in the abnormal insulin gene. The autoradiographs of portions of the sequence analysis gels for both the normal and abnormal genes are shown. The sense strand DNA was end-labeled with $[^{32}P]ATP$ at the *Nco* I site (see Fig. 3), and the sequence was determined by the method of Maxam and Gilbert (16).

gene sequence was a single nucleotide deletion/addition that was observed 45 nucleotides upstream from the Goldberg-Hogness ("TATAA") box. These findings are summarized in Fig. 3.

DISCUSSION

In this report, we describe the identification of a point mutation in the insulin gene of a patient with hyperinsulinemic diabetes. The presence of a mutation in the B-chain coding region was first suggested by restriction endonuclease Mbo II cleavage analysis of genomic DNA isolated from leukocytes of the patient. The enzyme Mbo II recognizes the nucleotide sequence T-C-T-T-C, which is present in the codons for the B-24 and -25 phenylalanine residues. Examination of the insulin gene regions coding for residues that are known to be invariant and essential for insulin's structural integrity or biological activity, or both, revealed a number of potential restriction enzyme cleavage sites that might be useful for screening. However, the majority of enzymes cleaving these sites would result in the production of large numbers of small fragments that would be difficult to analyze with presently available methods. Only three sites, as shown in Fig. 5, emerged as potentially useful in practice for this purpose; two of these are present in the A-chain and one is in the B-chain coding regions. Of these three, only the MboII site in the B chain is in a conserved region that is known to strongly influence receptor binding affinity (22). It is of interest that two of three patients so far identified as having mutant insulins have had a loss of this site. Thus, cleavage analysis by Mbo II of genomic DNA may be a useful first screening maneuver for detection of mutant insulin genes. A similar screening approach has already been applied to the detection of the sickle cell mutation in the β -globin gene by using the enzyme Dde II (23). The Mbo II cleavage defect also made it easier to identify the abnormal allele of the cloned insulin gene in this patient. Indeed, an equal number of clones of the normal and abnormal alleles were obtained as shown in Fig. 2.

Nucleotide sequence analysis revealed a point mutation within the codon for phenylalanine-B24 (TTC \rightarrow TCC) in the mutant gene. No other mutation was detected within the coding region of either allele of the insulin gene when these were compared with the previously reported nucleotide sequences for the normal gene (18, 19). This single nucleotide transition results in the substitution of serine for phenylalanine at position B24, which accounts for the less hydrophobic character of the mutant in-

Δ

sulin observed on reverse-phase HPLC analysis of the patient's serum insulin (3).

Because the region B23-26 has been proposed to lie near or within the receptor binding site of the insulin molecule (22), it is likely that amino acid substitutions in this region will result in a loss of biological activity. Studies using semisynthetic insulin analogues have confirmed this possibility (24-26). Thus, even the conservative substitution of leucine for phenylalanine at either B24 or B25 markedly reduces the biological activity of the resultant insulin. It is evident that the substitution of leucine for phenylalanine at B25 in the first patient and of serine for phenylalanine at B24 in this patient could be an important predisposing factor in producing diabetes. Both patients are mildly diabetic and exhibit fasting hyperinsulinemia, which may be a reflection of the impaired receptor-binding potency of the mutant insulin. Thus, although both normal and abnormal insulins were found in almost equal proportions in the pancreas of the first patient (3, 7), indicating the codominant expression of both alleles of the insulin gene, more than 90% of the circulating insulin was found to be the mutant form (3). These findings suggest that the metabolism of the mutant insulin is slowed and would account for the hyperinsulinemia in both patients. However, at the same time, both patients may suffer a relative insulin deficiency in terms of biologically active material and, thus, can be categorized as having mildly insulin-deficient diabetes. The markedly reduced biological activity of the insulin isolated from both patients' sera is consistent with this conclusion and indicates that the (receptor-mediated) degradation (27, 28) of the mutant insulin is selectively impaired in vivo.

Polymorphism of the human insulin gene and its possible relationship to diabetes have been reported (20, 29-31). The length polymorphism is found in the 5' flanking region of the insulin gene (Fig. 3) and is composed of a tandemly repeated 14-nucleotide G+C-rich family of sequences (17). Two main modal size distributions have been observed in the human population that are inherited in a Mendelian fashion (32). These consist of shorter inserts of 0-600 bp or longer inserts, predominantly of 1.6 kb (31). Several recent reports indicate that longer inserts are found more frequently in the genes of noninsulin-dependent diabetic patients than in normal subjects, but this subject remains controversial (29-31). In this patient, both alleles have shorter inserts, as indicated in Fig. 3. It is not known whether this polymorphism affects the expression of the insulin gene. The only change found in the 5' region in this study



FIG. 5. Potentially useful restriction enzyme cleavage sites within highly conserved regions of the human insulin molecule. Residues in solid circles are those that are invarient in the known vertebrate insulins (21).

was a single nucleotide deletion in the normal gene just 45 bp upstream from the TATA box (see Fig. 3).

Both genetic and environmental factors can influence the development of diabetes and the genetic factor(s) contributing to the disease are considered to be heterogeneous (1, 2). From the clinical course and the family history (5), this patient could be classified as having maturity-onset diabetes of the young (MODY), in which an autosomal dominant pattern of inheritance has been proposed (33, 34). The hyperinsulinemia of this patient clearly sets it apart from the majority of nonobese maturity-onset diabetics, who exhibit moderate reduction in circulating immunoreactive insulin. However, the present study along with those on two other individuals with mild diabetes associated with other mutations in one allele of the insulin gene, giving rise to a biologically defective insulin (3, 9), strongly suggest that altered insulin gene dosage in man (only one functional allele rather than two) can significantly impair normal insulin production and lead to diabetes in susceptible individuals. Thus, other kinds of mutations within the preproinsulin coding sequence that prevent gene expression, block precursor synthesis, or impair its transport and maturation (35, 36) may produce a similar phenotypic picture. The incidence of such abnormalities in the human (diabetic or normal) population is unknown.

The authors are indebted to Dr. William N. Wishner and Dr. Adina Zeidler for their assistance in identifying the patient and obtaining samples for this study and to Ms. Lisa Fuller for her expert assistance in preparing this manuscript. This work was supported by Grants AM 13914 and AM 20595 from the United States Public Health Service. S.C.M.K. is the recipient of United States Public Health Service Postdoctoral Fellowship AM 06577.

- Rotter, J. I. & Rimoin, D. L. (1978) Diabetes 27, 599-605. 1.
- Pyke, D. A. (1979) Diabetologia 17, 333-340. 2.
- 3. Shoelson, S., Haneda, M., Blix, P., Nanjo, A., Sanke, T., Inouye, K., Steiner, D., Rubenstein, A. & Tager, H. (1983) Nature (London) 302, 540-543.
- Given, B. D., Mako, M. E., Tager, H. S., Baldwin, D., Markese, 4. J., Rubenstein, A. H., Olefsky, J., Kobayashi, M., Kolterman, O. & Paucher, R. (1980) N. Engl. J. Med. 302, 129-135.
- 5. Haneda, M., Bergenstal, R., Friedenberg, G., Wishner, W., Blix, P., Provow, S., Polonsky, L. & Rubenstein, A. H. (1982) Diabetes 31, 4A (abstr.).
- Olefsky, J. M., Saekow, M., Tager, H. & Rubenstein, A. H. (1980) J. Biol. Chem. 255, 6098-6105. 6.
- 7. Tager, H., Given, B., Baldwin, D., Mako, M., Markese, J., Rubenstein, A., Olefsky, J., Kobayashi, M., Kolterman, O. & Paucher, R. (1979) Nature (London) 281, 122-125.
- 8. Kwok, S. C. M., Chan, S. J., Rubenstein, A. H., Paucher, R. & Steiner, D. F. (1981) Biochem. Biophys. Res. Commun. 98, 844-849

- Kwok, S. C. M., Steiner, D. F., Rubenstein, A. H. & Tager, H. 9. S. (1983) Diabetes 32, 872-875.
- 10. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) J. 11. Mol. Biol. 113, 237–251. Kwok, S. C. M., Chan, S. J. & Steiner, D. F. (1983) J. Biol. Chem.
- 12. 258, 2357-2363.
- Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., 13. Quon, D., Sim, G. K. & Efstratiadis, A. (1978) Cell 15, 687-701. Hohn, B. (1979) Methods Enzymol. 68, 299-309. 14.
- Blattner, F. R., Blechl, A. E., Denniston-Thompson, K., Faber, 15 H. E., Richards, J. E., Slightom, J. L., Tucker, P. W. & Smithies, O. (1978) Science 202, 1279-1284.
- 16. Maxam, A. M. & Gilbert, W. (1979) Methods Enzymol. 65, 499-560
- 17. Bell, G. I., Selby, M. J. & Rutter, W. J. (1982) Nature (London) 295, 31-35.
- Ullrich, A., Dull, T. J., Gray, A., Brosius, J. & Sures, I. (1980) 18. Science 209, 612–615
- 19. Bell, G. I., Pictet, R. L., Rutter, W. J., Cordell, B., Tischer, E. & Goodman, H. M. (1980) Nature (London) 284, 26-32
- 20. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) Proc. Natl. Acad. Sci. USA 78, 5759-5763.
- Fasman, G., ed. (1976) Handbook of Biochemistry and Molecu-21. lar Biology, Vol. 3, Proteins (CRC, Cleveland), pp. 378-382.
- Pullen, R. A., Lindsay, D. G., Wood, S. P., Tickle, I. J., Blun-dell, T. L., Wollmer, A., Krail, G., Brandenburg, D., Zahn, H., 22. Gliemann, J. & Gammeltoft, S. (1976) Nature (London) 259, 369-373.
- 23. Geever, R. F., Wilson, L. B., Nallaseth, F. S., Milner, P. F., Bittner, M. & Wilson, J. T. (1981) Proc. Natl. Acad. Sci. USA 78, 5081-5085.
- Assoian, R. K., Thomas, N. E., Kaiser, E. T. & Tager, H. S. (1982) 24. Proc. Natl. Acad. Sci. USA 79, 5147-5151.
- 25. Kobayashi, M., Ohgaku, S., Iwasaki, M., Maegawa, H., Shigeta, Y. & Inouye, K. (1982) Biochem. J. 206, 597-603.
- 26. Keefer, L. M., Piron, M.-A., DeMeyts, P., Gattner, H.-G., Diaconescu, C., Saunders, D. & Brandenburg, D. (1981) Biochem. Biophys. Res. Commun. 100, 1229-1236.
- Terris, S. & Steiner, D. F. (1975) J. Biol. Chem. 250, 8389–8398. Terris, S. & Steiner, D. F. (1976) J. Clin. Invest. 57, 885–896. 27.
- 28
- 29. Rotwein, P., Chyn, R., Chirgwin, J., Cordell, B., Goodman, H. M. & Permutt, M. A. (1981) Science 213, 1117-1120.
- Owerbach, D. & Nerup, J. (1982) Diabetes 31, 275-277 30.
- 31. Rotwein, P. S., Chirgwin, J., Province, M., Knowler, W. C., Pettit, D. J., Cordell, B., Goodman, H. M. & Permutt, M. A. (1983) N. Engl. J. Med. 308, 65-71.
- 32. Owerbach, D., Poulsen, S., Billesbolle, P. & Nerup, J. (1982) Lancet i, 870-883.
- 33. Fajans, S. S., Cloutier, M. C. & Crowther, R. L. (1978) Diabetes 27, 1112-1125
- 34. Tattershall, R. B. & Fajans, S. S. (1975) Diabetes 24, 44-53.
- Gabbay, K. H., Bergenstal, R. M., Wolff, J., Mako, M. E. & 35. Rubenstein, A. H. (1979) Proc. Natl. Acad. Sci. USA 76, 2881-2885.
- 36. Robbins, D. C., Blix, P. M., Rubenstein, A. H., Kanazawa, Y., Kosaka, K. & Tager, H. S. (1981) Nature (London) 291, 679-681.