Somatic Mutation in Genes for the Variable Portion of the Immunoglobulin Heavy Chain

Abstract. *The size of the gene pool potentially encoding antibodies to* p-*azophenyl arsonate has been examined. A heavy chain–specific full-length complementary DNA clone has been constructed with the use of messenger RNA from a hybridoma that produces antibodies to the arsonate hapten and bears nearly a full complement of the determinants comprising the cross-reactive idiotype (CRI). The sequences of both the complementary DNA clone and the corresponding immunoglobulin heavy chain have been independently determined. A probe for the variable region gene was prepared from the original heavy chain complementary DNA clone and used to analyze, by Southern filter hybridization, genomic DNA from both A/J (CRI positive) and BALB/c (CRI negative) mice. Approximately 20 to 25 restriction fragments containing "germline" variable region gene segments were detected in both strains, and many are shared by both. Since 35 CRI-positive heavy chains have been partially sequenced thus far and 31 are different, the results of the hybridization analysis suggest that somatic mutation events involving the variable region gene segments of the heavy chain play a role in the origin of the amino acid sequence diversity seen in this system.*

Antibody molecules recognize antigen by means of a binding site made from the variable (V) regions of the immunoglobulin heavy (H) and light (L) chains. The antibodies recognizing any one antigen may be classified into families whose members share closely related V-region amino acid sequences. One such family is that participating in the response of A/J mice to the hapten *p*-azophenyl arsonate (arsonate) first described by Kuettner *et al.* (*1*). Analysis of the arsonate-specific serum antibodies (*2*), B cell hybridomas (*3*), and T cells (*4*) and of the regulation (*5*) of this system has increased our understanding of the induction and control of humoral immune responses.

One of the primary issues concerning the arsonate system is the nature of the genetic information stored in the germline and its relation to the observed diversity of the immune response to this simple hapten. Initial structural and serological studies on the serum antibodies suggested that a limited number of germline genes, coupled with a small amount of somatic mutation in the V regions, might account for the diversity of the response to arsonate (*2*). However, serologic studies of B cell hybridoma products showed that the potential repertoire was far more diverse than had been expected from analysis of serum antibodies (*6–8*). In addition, amino acid sequence studies on cross-reactive idiotype (CRI) positive hybridoma products provided evidence for a minimum of 31 different heavy chain V-segment (V_H) sequences (*3, 6, 9, 10*).

The most direct analysis of the origin of amino acid sequence diversity will most likely derive from studies of the genes themselves. We have chosen a strongly idiotype positive hybridoma (93G7), which produces a monoclonal antibody to arsonate (anti-arsonate), and from it constructed a full-length heavy chain complementary DNA (cDNA) clone. The sequence of the cDNA clone and the independently derived protein sequence of the 93G7 heavy chain variable region are shown in Fig. 1. The presumptive leader or signal sequence was deduced from the cDNA sequence. We have assumed that the methionine codon at position −19 represents the protein synthesis initiation codon by comparison with other immunoglobulin signal sequences.

At the DNA level, the genes for V_H regions are constructed from separate segments encoding the V, the D (diversity), and the J (joining) segments (*11–13*). The V segment extends from the Met (*14*) codon −19 to codon 98, the D segment from 99 to 106 (shown in bold face

```
                                                                  Pst I
AC ACA CTG ACT CAA ACC ATG GGA TGG AGC TTC ATC TTT CTC TTC CTC CTG TCA GTA ACT GCA GGT GTC CAC TCT
                                   Met Gly Trp Ser Phe Ile Phe Leu Phe Leu Leu Ser Val Thr Ala Gly Val His Ser
                                                                  -10

GAG GTT CAG CTT CAG CAG TCT GGA GCT GAG CTG GTG AGG GCT GGG TCC TCA GTG AAG ATG TCC TGC AAG GCT TCT
GLU VAL GLN LEU GLN GLN SER GLY ALA GLU LEU VAL ARG ALA GLY SER SER VAL LYS MET SER CYS LYS ALA SER
                                                                                                  25

GGA TAT ACA TTC ACA AGC TAC GGT ATA AAC TGG GTG AAA CAG AGG CCT GGA CAG GGC CTG GAA TGG ATT GGA TAT
GLY TYR THR PHE THR SER TYR GLY ILE ASN TRP VAL LYS GLN ARG PRO GLY GLN GLY LEU GLU TRP ILE GLY TYR
                                                                                                  50

                                                                                  Acc I
ATT AAT CCT GGA AAT GGT TAT ATT AAC TAC AAT GAG AAG TTC AAG GGC AAG ACC ACA CTG ACT GTA GAC AAA TCC
ILE ASN PRO GLY ASN GLY TYR ILE ASN TYR ASN GLU LYS PHE LYS GLY LYS THR THR LEU THR VAL ASP LYS SER
                                                                                                  75

                                                                  Pst I
TCC AGC ACA GCC TAC ATG CAG CTC AGA AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TTC TGT GCA AGA TCC CAT
SER SER THR ALA TYR MET GLN LEU ARG SER LEU THR SER GLU ASP SER ALA VAL TYR PHE CYS ALA ARG SER HIS
                                                                                                 100

TAC TAT GGT GGT AGC TAC GAC TTT GAC TAC TGG GGC CAA GGC ACC CCT CTC ACA GTC TCC TCA
TYR TYR GLY GLY SER TYR ASP PHE ASP TYR TRP GLY GLN GLY THR PRO LEU THR VAL SER SER
```

Fig. 1. The cDNA sequence and the protein sequence of the 93G7 heavy chain variable region. The assignment of the leader or signal sequence (lower case) is based on homology with other sequences. The Pst I and Acc I restriction sites are shown. The bold type indicates the amino acids that constitute the D segment. The idiotypic characterization of the 93G7 molecule has been described (*6, 8*). Automated sequencing was performed with a Beckman 890C amino acid sequencer, modified by the addition of a cold trap (*18, 25*). Polysomal polyadenylated [poly(A)+] RNA was isolated from hybridoma cells (*26*). Double-stranded cDNA was synthesized from this RNA according to the method of Wickens *et al.* (*27*) and, after concentration with 2-butanol (*28*), was passed over a Sephadex G-100 column (*29*). The cDNA-containing fractions were digested at 15°C with a fivefold excess of *S1* nuclease and then passed over a column of Bio-Gel A150 M (*30*). Fractions containing cDNA predominantly larger than 1000 base pairs (bp) were pooled, and an average of 50 residues of adenylate [poly(A) tailing] were added with the use of terminal transferase (*31, 32*). The poly(A)-tailed cDNA was annealed to Bam HI–cleaved, polythymidylate-tailed plasmid pBR322 and transformed into *Escherichia coli* X1776 by a high-efficiency transformation procedure (*33*). Ampicillin-resistant, tetracycline-sensitive transformants were first screened by colony hybridization (*34*) with a mouse γ-1 probe (*35*) (93G7, like most CRI+ anti-arsonate hybridomas, is an immunoglobulin G1), and the plasmids of positive clones were analyzed by agarose-gel electrophoresis to determine the size of the cDNA inserts. One plasmid was chosen for further characterization. The entire V region was sequenced (*36*). To provide specific V_H probes, subfragments were generated by cleavage with Pst I (amino acid residues −4 to 92) or with Pst I and Acc I (residues −4 to 72) and subcloned into pBR322 and M13.

0036-8075/82/0416-0309$01.00/0

type), and the J segment from 107 to 121. Like all anti-arsonate heavy chains completely sequenced to date (*10*), the J_H sequence of 93G7 is most closely homologous with BALB/c J_{H2} (*13*) although different by two amino acids. One, an Asp for a Tyr in the first position of the J segment may be attributed to junctional diversification during the joining, at the DNA level, of D- and J-gene segments. The other, a Pro for Thr in the tenth position, is peculiar to 93G7 and is not shared by any of the other sequences of A/J antibodies to arsonate (*10*). When compared at the DNA level, the 93G7 cDNA J_H sequence is again more homologous to J_{H2} than to any of the other BALB/c J_H segments (*13*).

Southern filter hybridization (*15*) of liver DNA from A/J and BALB/c mice using a specific V-region fragment as probe (Fig. 2) revealed 20 to 25 distinct bands. While it is possible that some bands may contain more than one hybridizing unit and while the bands vary considerably in intensity, the number of bands does not appear to differ significantly between digests whether Eco RI or Bam HI is used. Most of the bands have the same sizes in both strains, but there are also strain-specific bands. Similar results were also obtained with either the full-length (V + C) cDNA clone or a second V-region subclone comprising codons −4 to 72 (data not shown) as probes. These results suggest that, while many of the germline V segments may be shared between the two strains, unique V segments exist and may be responsible for the differential expression of the CRI. The hybridization results (Fig. 2) were obtained under conditions that allow the cross-hybridization of V genes which differ by as much as 20 percent in nucleotide sequence (*16, 17*). It is conceivable, therefore, that not all of the hybridizing bands represent arsonate-specific V-gene segments and thus 20 to 25 represents the upper limit for the number of idiotype-positive anti-arsonate genes.

Analyses have indicated that every monoclonal antibody to the arsonate hapten contains distinct antigenic determinants (idiotypes) (*8*). In addition, 35 individual anti-arsonate CRI-positive H chains have been at least partially sequenced, and 31 of the sequences differ by at least one amino acid (*3, 6, 9, 10, 18*). No two completely sequenced anti-arsonate heavy or light chains are identical in their V segments. Statistical analysis would suggest that several hundred germline genes coding for anti-arsonate would have to be present in order to accommodate the available data by a strict germline mechanism. The data from Southern filter hybridization are inconsistent with that possibility and suggest that somatic mutation must occur in the V_H-gene segment or segments encoding this family of highly related antibodies. In fact, hybridization analysis of various hybridoma DNA's suggests that a single V_H-gene segment may be used for the expression of both CRI-positive and some CRI-negative phenotypes within this highly conserved family of antibodies (*18a*).

Several reports have suggested that somatic mutation may play an important role in the generation of antibody diversity. Two other idiotypic systems have been analyzed by experimental procedures analogous to those we describe. The NP^b system, an induced idiotypic system in C57BL/6 mice, is thought to be composed of a large repertoire of expressed sequences (*19, 20*). However, studies by Bothwell *et al.* (*21*) indicate that these expressed protein sequences are encoded by relatively few V-gene segments. Similarly, in the phosphorylcholine (PC) idiotypic system (*22*), amino acid sequence analysis of myelomas and hybridomas indicates an extensive array of expressed antibodies to PC despite a limited number of germline V_H segments (*23*). In humans, the number of germline V genes has been examined for both heavy chains (*12*) and light chains (*16*); in these cases, a comparison with known protein sequences again suggests the occurrence of somatic mutation within the V segments. Furthermore, direct evidence for somatic mutants of mouse V_L genes has been obtained from comparison of germline and expressed genes (*24*). Taken together, the results show that somatic variation of V-gene sequences plays a significant role in the expansion of antibody diversity.

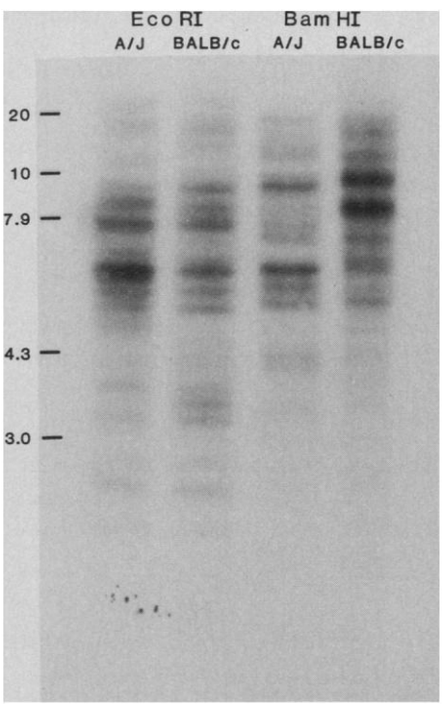

Fig. 2. Southern filter hybridization (*15*) of mouse liver DNA with the V_H-specific Pst I subfragment of the cDNA probe. Both Eco RI (left) and Bam HI (right) digests of A/J liver DNA and BALB/c liver DNA are shown. Total A/J and BALB/c liver DNA was prepared essentially as described (*37*) and further purified by CsCl density equilibrium gradients. The DNA's were completely digested with either Eco RI or Bam HI (overnight at 37°C) and fractionated on a 1 percent agarose gel in 4× Hellings buffer (final concentration 160 m*M* tris, 180 m*M* sodium acetate, 70 m*M* NaCl, 8 m*M* EDTA, *p*H 8.0). The DNA was denatured in the gel for 1 hour in 1*M* NaCl, 0.5*M* NaOH. The gel was neutralized for 1 hour in 3*M* NaCl, 0.5*M* tris, *p*H 7.4 prior to the transfer of DNA to nitrocellulose in 10× SSC (1.5*M* NaCl, 0.15*M* sodium citrate, *p*H 7.2). The filters were baked under vacuum at 80°C, and one of two procedures was used for hybridization. (i) The filters were first washed at 42°C for 24 hours in hybridization solution containing 5× SSC, 0.02 percent polyvinylpyrrolidone (PVP), 0.02 percent Ficoll, 0.02 percent bovine serum albumin (BSA) (*38*), herring sperm DNA (100 μg/ml), and 50 percent deionized formamide. Hybridization was carried out under the same conditions with the addition of poly(rA) (30 μg/ml) and the nick-translated (*39*) probe (1 × 10^6 cpm/ml). After a 48-hour hybridization, the filter was washed sequentially with 3× SSC, 1× SSC, and 0.1× SSC at 42°C for 3 hours (1 hour for each solution). (ii) Washing and hybridization were carried out in 6× SSC plus 0.1 percent sodium dodecyl sulfate, and 0.2 percent PVP, Ficoll, and BSA at 65°C; and subsequently washed in 1× SSC (the low stringency procedure previously described) (*16*). Identical results were obtained with both procedures.

JOHN SIMS
TERENCE H. RABBITTS
MRC Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge, CB2 2QH England
PILA ESTESS
CLIVE SLAUGHTER
PHILIP W. TUCKER
J. DONALD CAPRA*
Department of Microbiology, Southwestern Medical School, University of Texas Health Science Center at Dallas, Dallas 75235

References and Notes

1. M. G. Kuettner, A-L. Wang, A. Nisonoff, *J. Exp. Med.* **135**, 579 (1972).
2. J. D. Capra and A. Nisonoff, *J. Immunol.* **123**, 279 (1979).
3. P. Estess, A. Nisonoff, J. D. Capra, *Mol. Immunol.* **16**, 1111 (1979).
4. M-S. Sy, A. R. Brown, B. Benacerraf, M. I. Green, *J. Exp. Med.* **151**, 896 (1980).
5. A. Nisonoff, S. T. Ju, S. F. Owen, *Immunol. Rev.* **34**, 89 (1977).
6. P. Estess, E. Lamoyi, A. Nisonoff, J. D. Capra, *J. Exp. Med.* **151**, 863 (1980).

7. A. Marshak-Rothstein, J. D. Benedetto, R. L. Kirsch, M. Gefter, *J. Immunol.* **125**, 1987 (1980).
8. E. Lamoyi, P. Estess, J. D. Capra, A. Nisonoff, *J. Exp. Med.* **124a**, 2834 (1980).
9. A. Marshak-Rothstein, M. Siekevitz, M. N. Margolies, M. Mudgett-Hunter, M. L. Gefter, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1120 (1980); S. S. Alkan, R. Knecht, D. G. Braun, *Hoppe-Seylers Z. Physiol. Chem.* **361**, 191 (1980); M. Margolies, A. Marshak-Rothstein, M. L. Gefter, *Mol. Immunol.* **18**, 1065 (1981); M. Siegelman, C. A. Slaughter, L. J. McCumber, P. Estess, J. D. Capra, in *Immunoglobulin Idiotypes*, C. Janeway, E. Sercarz, H. Wigzell, Eds. (Academic Press, New York, 1981), pp. 135–158; C. A. Slaughter and J. D. Capra, unpublished observations.
10. C. A. Slaughter, M. Siegelman, P. Estess, I. Barasoain, A. Nisonoff, J. D. Capra, in *Developmental Immunology: Clinical Problems and Aging*, E. L. Cooper and M. A. B. Brazier, Eds. (Academic Press, New York, in press).
11. P. Early, H. Huang, M. Davis, K. Calame, L. Hood, *Cell* **19**, 981 (1980).
12. G. Matthyssens and T. H. Rabbitts, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6561 (1980).
13. H. Sakano, R. Maki, Y. Kurosawa, W. Roeder, S. Tonegawa, *Nature (London)* **286**, 676 (1980).
14. Abbreviations for the amino acid residues are: Ala, alanine; Arg, arginine; Asp, aspartic acid; Asn, asparagine; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine.
15. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
16. D. L. Bentley and T. H. Rabbitts, *Cell* **24**, 613 (1981).
17. T. H. Rabbitts, G. Matthyssens, P. H. Hamlyn, *Nature (London)* **284**, 238 (1980).
18. C. A. Slaughter and J. D. Capra, in preparation.
18a. P. Estess, personal communication.
19. R. S. Jack, T. Imanishi-Kari, K. Rajwesky, *Eur. J. Immunol.* **8**, 559 (1977).
20. M. Reth, T. Imanishi-Kari, K. Rajewsky, *ibid.* **9**, 1004 (1979).
21. A. L. M. Bothwell, M. Paskind, M. Reth, T. Imanishi-Kari, K. Rajewsky, D. Baltimore, *Cell* **24**, 625 (1981).
22. R. Lieberman, M. Potter, E. B. Mushinski, W. Humphrey, S. Rudikoff, *J. Exp. Med.* **13**, 983 (1974); J. L. Claflin, *Eur. J. Immunol.* **6**, 669 (1976); P. J. Gearhart, N. D. Johnson, R. Douglas, L. Hood, *Nature (London)* **291**, 29 (1981).
23. S. Crews, J. Griffin, H. Huang, K. Calame, L. Hood, *Cell* **25**, 59 (1981).
24. E. Selsing and U. Storb, *ibid.*, p. 47; M. Pech, J. Hochtl, H. Schnell, H. G. Zachan, *Nature (London)* **291**, 668 (1981).
25. L. J. McCumber, M. Qadeer, J. D. Capra, *Methods in Peptide and Protein Sequence Analysis*, C. Birr, Ed. (Elsevier/North-Holland, Amsterdam, 1980), pp. 165–172.
26. T. H. Rabbitts, J. M. Jarvis, C. Milstein, *Cell* **6**, 5 (1975).
27. M. P. Wickens, G. N. Buell, R. T. Schimke, *J. Biol. Chem.* **253**, 2483 (1978).
28. D. W. Strafford and D. Bieber, *Biochim. Biophys. Acta* **378**, 18 (1975).
29. T. E. Shenk, C. Rhodes, P. W. J. Rigby, P. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 989 (1975).
30. N. M. Gough, E. A. Webb, S. Cory, J. M. Adams, *Biochemistry* **19**, 2702 (1980).
31. M. W. McDonnell, M. N. Simon, F. W. Studier, *J. Mol. Biol.* **110**, 119 (1977).
32. R. Roychouhury, E. Jay, R. Wu, *Nucleic Acids Res.* **3**, 105 (1976).
33. D. Hanahan, personal communication.
34. M. Grunstein and D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961 (1975).
35. W. Dunnick, T. H. Rabbitts, C. Milstein, *Nucleic Acids Res.* **8**, 1475 (1980).
36. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
37. N. Blin and D. W. Stafford, *Nucleic Acids Res.* **3**, 2303 (1976).
38. D. T. Denhardt, *Biochem. Biophys. Res. Commun.* **23**, 641 (1966).
39. P. J. W. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).

40. Supported in part by NIH grants AI12127 and AI18016 and NSF grant PCM 76-22411. J.S. was supported by Damon Runyon–Walter Winchell Cancer Fund grant DRG-298F and NIH grant AI-06020. We thank Ana Garcia-Spain, Sandy Graham, Martha Danhof, and Alan Forster for technical assistance and Kathy Able for assistance in typing the manuscript.

* Address requests for reprints to J.D.C.

22 January 1982

Decreased Insulin Receptors but Normal Glucose Metabolism in Duchenne Muscular Dystrophy

Abstract. *Compared to matched controls, 17 patients with Duchenne muscular dystrophy showed decreased insulin binding to monocytes due to decreased receptor concentration. These patients showed no signs of altered glucose metabolism and retrospective analysis of the clinical records of a further 56 such patients revealed no modification in carbohydrate metabolism. These data suggest that reduced insulin receptor number does not produce overt modifications of glucose metabolism in Duchenne muscular dystrophy.*

The concentration or the affinity of the insulin receptor changes in response to physiological, pharmacological, or pathological variations in glucose metabolism. However, it is not known whether modification of the insulin receptor per se provokes variation in glucose metabolism. In this report we describe a clinical situation in which modification of the insulin receptor status does not appear to produce overt disturbances of glucose metabolism.

The monocytes of an apparently normal 5-year-old boy showed fewer insulin receptors compared to the monocytes of other matched boys. No reason could be found to explain this finding. About 2 years later, however, the boy was found to have muscular dystrophy of the Duchenne type (DMD). DMD is a systemic lesion that is inherited as an autosomal dominant trait (*1*) and is characterized by marked membrane defects (*2*); a modification of the insulin receptor is thus to be expected. However, DMD is also characterized by normal carbohydrate metabolism (*1*), which would not lead one to expect an alteration of the insulin receptor; to our knowledge, in fact, defects in the insulin receptor in the absence of overt changes of carbohydrate metabolism have not been reported. This led us to study the insulin receptor in males with DMD.

The studies were carried out on 17 males with DMD (aged 8 to 23 years) from the Istituto S. Stefano in Porto Potenza Picena (Italy). All patients had a history of pseudohypertrophy plus progressive proximal weakness, marked increase in creatine phosphokinase (CPK), and electromyography typical of DMD. Diagnosis had been made 2 to 15 years previously. The control group consisted of 17 males (aged 8 to 25 years) whose body weights were 91 to 112 percent of the ideal (*3*) and whose family histories indicated no diabetes or obesity. Insulin binding to monocytes in DMD patients was lower than in controls ($P < .001$) (Fig. 1); conversely, no differences were found in plasma glucose or serum insulin values (data not shown). Six handicapped subjects with neurogenic atrophy (non-DMD) from the same hospital (receiving the same diet and leading the same physical life as the DMD patients) showed normal insulin binding to monocytes, thus confirming that the results were not due to the life-style of the DMD patients (data not shown). The wide range of data in the DMD patients (Fig. 1) was not due to abnormal insulin bind-

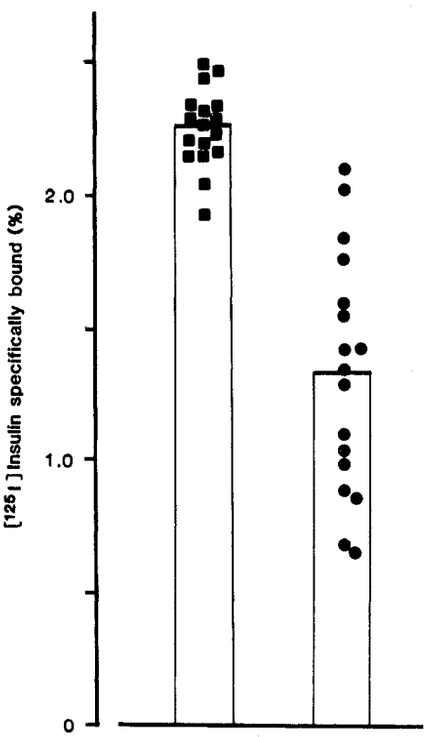

Fig. 1. Insulin binding to monocytes (4×10^6 per milliliter) from normal subjects (■) and DMD patients (●). All subjects were on an unrestricted diet and were not taking any drug known to affect carbohydrate or insulin metabolism. Blood (70 ml) was collected at 0800 hours after overnight fasting and samples were centrifuged on a Ficoll-Angiografin gradient. Mononuclear cells were isolated, resuspended in tris-HCl buffer (*p*H 7.8) to a final concentration of 3×10^7 per milliliter, and subsequently incubated for 100 minutes at 15°C with [^{125}I]insulin (0.6 ng/ml; CEA-IRE Sorin, Italy; 180 to 200 μCi/μg) in the absence or presence of increasing amounts of porcine insulin (Organon, Holland) (*9*). Student's *t*-test for unpaired data was used in the statistical analysis.

0036-8075/82/0416-0311$01.00/0