Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 964–967, February 1990
Immunology

Primary structure of rat RT6.2, a nonglycosylated phosphatidylinositol-linked surface marker of postthymic T cells

(T-cell development/rat alloantigens/cDNA cloning/insulin-dependent diabetes mellitus)

FRIEDRICH KOCH, FRIEDRICH HAAG, ANDRIJKA KASHAN, AND HEINZ-GÜNTER THIELE*

Department of Immunology, University Hospital, Hamburg, Federal Republic of Germany

Communicated by Hans J. Müller-Eberhard, October 30, 1989 (received for review August 25, 1989)

ABSTRACT **RT6 is an unusual cell membrane protein that is expressed exclusively by postthymic T cells. The inherent defect in its expression has been correlated to lymphopenia and genetically determined susceptibility for insulin-dependent diabetes mellitus in the rat. We report here the primary structure of the RT6.2 alloantigen as deduced from the cDNA sequence. The predicted amino acid sequence of RT6.2 begins with a conventional leader of 20 amino acids and ends in a hydrophobic C-terminal extension peptide of 29 amino acids as is common for phosphatidylinositol-anchored proteins. Native RT6.2 is predicted to comprise 226 amino acids, with a calculated M_r of 26,036. Four cysteine residues account for two intrachain disulfide bonds. The sequence lacks potential N-glycosylation sites and contains an excess of positively charged residues. Homology searches in protein sequence data banks suggest that RT6.2 is not encoded by a member of the immunoglobulin supergene family. Moreover, these analyses did not reveal any close homologies of RT6.2 to known proteins: the highest homology found was 21.2% identity in a 52-amino acid overlap to the torpedo acetylcholinesterase precursor. Southern blot analyses indicate that RT6.2 is the product of a single-copy gene and provide evidence for closely related genes in the mouse and other species. The corresponding gene products remain to be identified.**

The rat T-cell alloantigenic system RT6 was originally discovered by several groups because of its potent immunogenicity in the allogeneic context (1–7). Two allelic gene products—designated RT6.1 and RT6.2—of M_r 25–30 × 10^3 could be discriminated by allotype-specific antibodies (8, 9). The corresponding alleles, *RT6*^{a} and *RT6*^{b}, were mapped to linkage group I between the loci for albinism and hemoglobin β chain complex (*Hbb*) (10), a region of high-linkage-synteny homologies to mouse chromosome 7 and human chromosome 11 (11, 12). Intriguingly, no comparable proteins have yet been described in mouse or man.

The RT6 antigenic system is of interest for studies of postthymic T-cell development (13) because it allows the distinction of mature peripheral T cells (RT6$^+$/Thy-1$^-$) from thymic lymphocytes and recent thymic migrants (RT6$^-$/Thy-1$^+$) (14, 15). Expression of RT6 is restricted to the final stages of postthymic T-cell development; these antigens are not expressed by any other hematopoietic cells or cells of other tissues (14, 15). Remarkably, it is quite difficult to maintain RT6 expression during *in vitro* cultivation of rat T cells (16).

RT6 is also of interest with respect to genetically determined insulin-dependent diabetes mellitus (IDDM) (15, 17–19). Thus, the diabetes-prone Bio-Breeding (dpBB) rat, an animal model for human autoimmune type I IDDM, has an inherent defect in the generation of RT6$^+$ cells (15, 18), and this defect has been associated with the pathogenesis of IDDM in these animals (18, 19).

Biochemically, the RT6 antigens represent unusual membrane proteins in that they occur in nonglycosylated (RT6.2) or in nonglycosylated and differently glycosylated (RT6.1) variants (8, 9). All forms of RT6 are anchored in the cell membrane by covalent linkage to a phosphatidylinositol moiety (9, 20). Treatment of T cells with RT6-specific antisera reportedly is mitogenic (21), similar to the case of other phosphatidylinositol-anchored T-cell surface proteins (22).

We report here the molecular cloning of RT6.2-specific cDNAs and the entire coding sequence for RT6.2.† The present data provide tools for elucidating the molecular basis of the marked polymorphism in this alloantigenic system in the rat (9) and the defect of its expression in dpBB rats (15, 19). Moreover, they will facilitate the identification of the presumptive RT6-homologues in the human and other species.

MATERIALS AND METHODS

Isolation, Characterization, and Sequencing of RT6.2-Specific cDNA Clones. Standard protocols for the isolation of nucleic acids and DNA cloning (23, 24) were followed throughout. Poly(A)$^+$ RNA from the rat T-T hybridoma EpD3 (16) was used to construct a phage λgt11 cDNA expression library for immunoscreening with RT6-specific antisera as described (25). The 330-base-pair (bp) cDNA insert of an immunoreactive phage was subcloned into pBluescript plasmid (Stratagene). The insert was purified from the restriction digest of a plasmid "miniprep" by agarose gel electrophoresis and was radiolabeled with [^{32}P]dCTP (Amersham) to a high specific activity (5 × 10^8 cpm/μg) by the random prime-labeling procedure. This probe was used to rescreen the library by the colony hybridization protocol, and eight additional cross-hybridizing clones were isolated. The cDNA inserts of these clones were amplified with *Thermus aquaticus (Taq)* polymerase (Perkin–Elmer) by the polymerase chain reaction (26) and characterized with respect to insert size and restriction sites by agarose gel electrophoresis. Appropriate restriction fragments were subcloned into plasmid (pBluescript, Stratagene) and phage (M13, Boehringer Mannheim) sequencing vectors. DNA sequencing was performed on single-stranded and double-stranded DNA templates by the Sanger dideoxy chain termination method with Sequenase (United States Biochemical) according to the manufacturer's protocol. The sequence of the entire coding region for RT6.2 was determined in both directions, and each nucleotide was determined an average of four times. Hydropathy plots and molecular weight and isoelectric point calculations were performed on a Macintosh personal computer with the MACMOLLY program (Softgene,

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: IDDM, insulin-dependent diabetes mellitus; dpBB rats, diabetes-prone Bio-Breeding rats.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M30311).

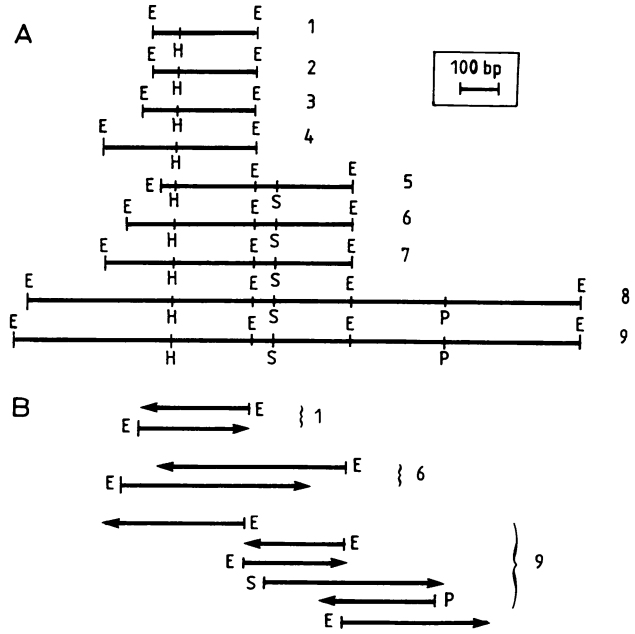

FIG. 1. Restriction maps (*A*) and sequencing strategy (*B*) for RT6.2-specific cDNAs. (*A*) Clones 1 and 2 were isolated from a rat T-cell hybridoma cDNA expression library in phage λgt11 by immunoscreening with RT6.2-specific antisera. Clones 3–9 were isolated from the same library by using the cDNA insert of clone 1 as a probe. The sites of the restriction enzymes *Eco*RI, *Hin*cII, *Sst* I, and *Pst* I are indicated by E, H, S, and P, respectively. (*B*) For sequence analyses, the inserts of clones 1, 6, and 9 were subcloned intact or after restriction enzyme digestion. The restriction sites used as sequencing start points are marked as in *A*. The extent and direction of sequencing are indicated by the arrows.

West Berlin). A homology search of the predicted amino acid sequence of RT6.2 in the National Biomedical Research Foundation data bank was performed with the Fast P program from the HUSAR program package (27).

Protein Purification and Amino Acid Sequence Analyses. Purification of RT6.2 and amino acid sequence analyses of proteolytic peptide fragments were performed as described (28).

Southern Blot Analysis. Southern blots (on Hybond-N membranes from Amersham) were prepared as described (25) or purchased from Clontech. Radiolabeled probe was prepared as described above, and hybridization was performed for 18 hr in 6× SCC (1× SCC = 0.15 M NaCl/0.015 M sodium citrate, pH 7) containing 5× Denhardt's solution (1× = 0.02% Ficoll/0.02% bovine serum albumin/0.02% polyvinylpyrrolidone), 10% dextran sulfate, and 0.1% sodium dodecyl sulfate (SDS) at 60°C. Blots were washed several times at low (2× SSC, 50°C), intermediate (2× SSC, 60°C), and high (0.2× SSC, 65°C) stringencies and subjected to autoradiography at −80°C for 8–96 hr.

RESULTS AND DISCUSSION

Cloning and Characterization of RT6.2-Specific cDNAs. A RT6.2-specific cDNA was isolated from a rat T-T hybridoma cDNA expression library by immunoscreening as described (25). This 330-bp cDNA insert hybridized to a single 1.5- to 1.6-kilobase (kb) mRNA species, indicating that it represents only a fragment of the full-length cDNA (25). By using this fragment as a probe to rescreen the library, seven additional cross-hybridizing clones were isolated. Restriction mapping of these cDNA inserts (Fig. 1*A*) revealed the presence of two internal *Eco*RI sites that evidently had been protected only partially by the treatment with *Eco*RI methylase during cDNA construction. Fig. 1*B* illustrates the strategy used for obtaining the RT6.2 cDNA sequence.

Structural Features of RT6.2 as Deduced from the cDNA. The composite nucleotide sequence and the predicted amino acid sequence of the single open reading frame are shown in Fig. 2. The amino acid sequence corresponds precisely to all amino

A

```
TTTGACGATG CCATCAAATA TTTGCAAGTT CTTCCTAACT TGGTGGTTAA TTCAGCAGGT
GACTGGGCTG ACAGGGCCTT TGATGCTAGA CACGGCTCCC AATGCATTTG ATGACCAGTA
TGAGGGCTGT GTCAACAAAA TGGAGGAAAA GGCACCCCTG CTTTTACAGG AAGACTTTAA
TATGAATGCG AAATTAAAAG TTGCGTGGGA AGAGGCAAAG AAACGATGGA ACAACATAAA
ACCTAGTAGG AGTTATCCCA AAGGTTTCAA TGATTTCCAT GGAACGGCTT TAGTTGCCTA
CACTGGGAGT ATCGCTGTAG ATTTTAACAG AGCTGTTAGG GAATTCAAGG AAAATCCTGG
TCAATTCCAC TACAAGGCCT TCCATTACTA CTTAACAAGA GCTCTTCAGC TTTTGAGTAA
CGGGGATTGT CATTCAGTCT ACCGAGGCAC TAAGACCAGG TTTCACTATA CTGGAGCTGG
CTCCGTGCGA TTCGGGCAGT TCACGTCTTC ATCTTTATCT AAGAAAGTAG CTCAATCTCA
AGAGTTTTTC AGTGATCATG GGACGCTGTT CATCATCAAA ACCTGCTTGG GGGTTTATAT
CAAAGAATTC TCTTTCCGTC CTGACCAAGA GGAGGTGTTA ATTCCAGGCT ATGAGGTATA
TCAGAAAGTC AGGACACAAG GCTACAACGA AATTTTCCTG GACTCCCCGA AGAGGAAGAA
GAGCAACTAC AATTGCTTAT ACAGCAGCGC AGGAGCCAGA GAAAGCTGTG TATCCCTGTT
CCTTGTGGTT CTCCCCAGTC TTCTGGTCCA GCTGCTTTGT CTTGCTGAGC CGTAGCCCTC
TCCTGTCTGC AGTTTGTGGG TCTGAGTGTT AAAGGGAGGC CAAGGAAGCA AGCACTCAAT
TGGTTGTCTA AGATCTGTGG TGGAGGGGGA GGGCGGATCC TATCAGGGAA AGTCTGCTCT
AGCATGCACG ATTCTGTAAG CATCCAAGGA TCCGACTGGT CGACCATTAA ATTACACTAA
CTTCAACTAT ACTACTTCAC GAGTTGC
```

B

```
-20         -10         1           11          21          31
MPSNICKFFL TWWLIQQVTG LTGPLMLDTA PNAFDDQYEG CVNKMEEKAP LLLQEDFNMN

41          51          61          71          81          91
AKLKVAWEEA KKRWNNIKPS RSYPKGFNDF HGTALVAYTG SIAVDFNRAV REFKENPGQF

101         111         121         131         141         151
HYKAFHYYLT RALQLLSNGD CHSVYRGTKT RFHYTGAGSV RFGQFTSSSL SKKVAQSQEF

161         171         181         191         201         211
FSDHGTLFII KTCLGVYIKE FSFRPDQEEV LIPGYEVYQK VRTCGYNEIF LDSPKRKKSN

221         231         241         251
YNCLYSSAGA RESCVSLFLV VLPSLLVQLL CLAEP
                                    PI
```

FIG. 2. Primary structure of RT6.2 cDNA (*A*) and the deduced amino acid sequence in single-letter code (*B*). (*A*) The initiation and stop codons that delineate the single long open reading frame are boxed. The two internal *Eco*RI restriction sites are marked by solid underlines, and the single *Hin*cII, *Sst* I, and *Pst* I sites are marked by broken underlines. (*B*) Amino acid assignments that were confirmed by direct sequence analyses of RT6.2 peptides (28) are underlined. The N-terminal leucine is marked as residue 1. The proposed cleavage point of the C-terminal tail and phosphatidylinositol-anchor attachment site is marked by the wavy line behind serine-226. The four internal cysteine residues are marked by arrows. The two glutamic acid-phenylalanine dimers that are encoded by the internal *Eco*RI sites are boxed.

acid assignments from amino acid sequence analyses (28) (underlined residues in Fig. 2*B*). The RT6.2 cDNA encodes a protein of 275 amino acids, including leader and C-terminal extension peptides (Fig. 2). Fig. 3 shows a hydropathy plot for the predicted amino acid sequence and schematically illustrates its main structural features. The presumptive AUG initiator codon is located 20 codons upstream of the codon for the N-terminal leucine of RT6.2; the intervening hydrophobic (phob)-amino-acid-rich stretch is characteristic for a leader sequence. The localization of the two known immunogenic epitopes of RT6.2 (9) to the N-terminal 93 amino acids (Fig. 3*B*) can be deduced from the immunoreactivity of the recombinant fusion protein of clone 1 (25), which encodes only this portion of RT6.2 (Figs. 1 and 2).

The sequence of RT6.2 ends in a stretch of 20 hydrophobic amino acids (Fig. 3*A*), as is common for phosphatidylinositol-anchored surface proteins (29). This accords well with the phosphatidylinositol-anchor predicted for RT6.2 on the basis of biochemical analyses (20). The site of posttranslational cleavage and phosphatidylinositol-anchor attachment has been determined directly for some of these proteins and in each case has been shown to lie approximately 10 amino acid residues upstream of the hydrophobic stretch (29). Amino acid sequence analysis of the C-terminal peptide of RT6.2 (28) suggests that the serine residue at position 226 may be the cleavage point for RT6.2. Thus, the sequence of this peptide (Ser-Asn-Tyr-Asn-Cys-Leu-Tyr-Ser; Fig. 2*B*) ends after the eighth sequencing cycle, whereas the copurified N-terminal peptide yielded clearly distinguishable amino acid derivatives for more than 10 additional sequencing cycles (28).

The predicted amino acid composition and sequence also confirm other biochemical features of RT6.2 (8, 9). Thus, the sequence contains no Asn-Xaa-(Ser or Thr) potential glycosylation site, accounting for the lack of any classical membrane protein oligosaccharide side chains in native RT6.2 (8). The predicted M_r of 26,036 for native RT6.2 accords with the molecular weight determined by SDS/PAGE (8, 9). The excess of positively charged residues (calculated isoelectric point of 8.2) may explain the basic behavior of RT6.2 in nonequilibrium pH-gradient gel electrophoresis (NEPHGE) (9). Finally, the four cysteine residues account for the two intrachain disulfide bonds predicted by the shift in molecular weight in SDS/PAGE after treatment of the native protein with reducing agents or NaOH (8).

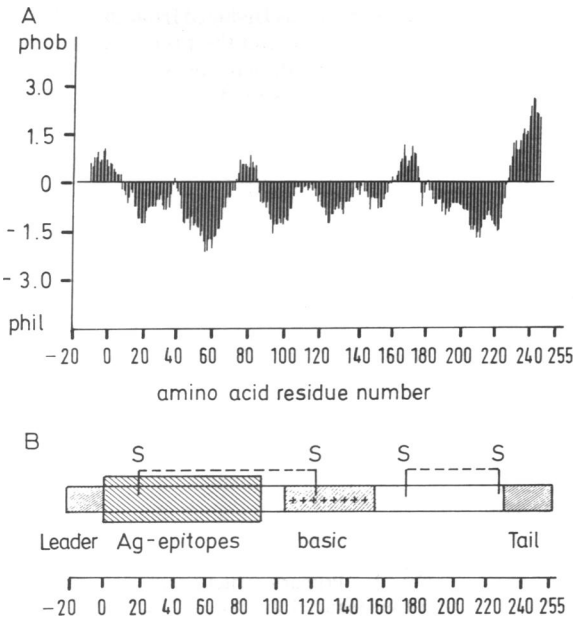

FIG. 3. Structural features of the predicted amino acid sequence for RT6.2. (*A*) The hydropathy plot of the predicted amino acid sequence was obtained with the MACMOLLY Program (Softgene) by using the Kyte–Doolittle algorithm and a window setting of 19 amino acid residues. phob, Hydrophobic; phil, hydrophilic. (*B*) Schematic outline of the main structural features of the predicted amino acid sequence. The relative position and lengths of the cleaved N-terminal leader and C-terminal tail peptides are indicated by shaded boxes. The 93-amino acid N-terminal domain containing the two known antigenic epitopes (a RT6.2 allotype-specific epitope as well as an epitope shared with RT6.1) is boxed. The relative localization of the internal cysteine residues is marked by S, and one of three possible pairs of intrachain disulfide bonds is indicated by S-----S. An internal segment rich in basic amino acids is also outlined.

Comparison of the Deduced Amino Acid Sequence of RT6.2 with Other Known Protein Sequences. A homology search (27) with the deduced amino acid sequence of RT6.2 in the National Biomedical Research Foundation protein sequence data base did not reveal any significant sequence homologies to known T-cell surface proteins or CD markers. The highest homology score obtained was for the torpedo acetylcholinesterase precursor: initial score 69/optimized score 83 (21.2% identity in a 55-amino acid overlap). A detailed comparison of the amino acid sequences of RT6.2 and known members encoded by the immunoglobulin supergene family (30) suggests that RT6 does not belong to this family (analysis kindly performed by A. F. Williams, Medical Research Council Immunology Unit, Oxford).

Southern Blot Analyses of Genomic DNA from the Rat and Other Species. Southern blot analyses of restriction enzyme-fragmented genomic rat DNA with RT6.2 cDNA probes indicate a single-copy gene and reveal a distinct restriction fragment length polymorphism in the *RT6*^{a} and *RT6*^{b} alleles with several different restriction enzymes (25). These results suggest extensive genetic events rather than a single point mutation as the basis for the allelic divergence of *RT6*^{a} and *RT6*^{b}. Corresponding analyses with genomic DNA of dpBB rats reveal a pattern indistinguishable from that of diabetes-resistant Bio-Breeding (BB) and other *RT6*^{a} rats, indicating that this gene is not grossly altered in dpBB rats (15). Sequencing of the dpBB rat genes and normal *RT6*^{a} genes and their regulatory sequences may help to clarify why dpBB rats are unable to generate RT6$^+$ cells. Such analyses will be greatly facilitated by the *RT6*-specific probes described here. In Southern "zoo-blot" analyses, bands that cross-react with RT6.2 cDNA probes can be detected in many different species under low-stringency conditions (Fig. 4*A*). It is

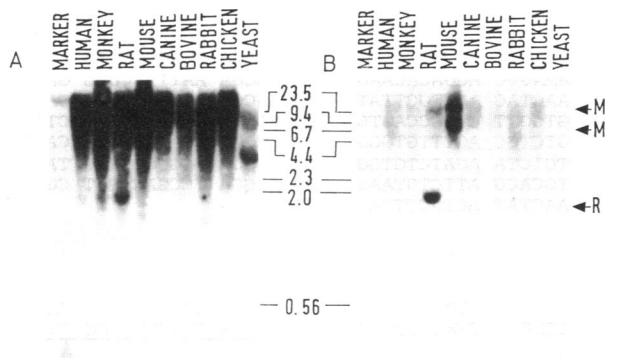

FIG. 4. Southern "zoo-blot" analysis with an RT6.2 probe. The cDNA insert of clone 1 (coding for the 93 N-terminal amino acids of RT6.2) was radiolabeled and used to probe a "zoo-blot" of *Eco*RI-digested genomic DNAs of the indicated species. (*A*) Autoradiograph of this blot after low-stringency washing (2× SSC at 50°C). Autoradiography was performed with Kodak X-Omat AR film and two intensifying screens for 12 hr at −80°C. (*B*) Autoradiograph of the same blot after washing at higher stringency (2× SSC at 60°C). Autoradiography was performed as in *A* but for 48 hr. Sizes are shown in kb.

noteworthy that the mouse bands resist relatively stringent washing conditions (Fig. 4*B*). These results provide evidence for the existence of a closely related gene in the mouse and for genes of lesser homology in more distantly related species. With the RT6.2-specific cDNAs described in this report, characterization of the homologous genes and gene products in other species should now be possible.

Note Added in Proof. Recent sequence analyses of cDNAs for RT6.1 (F.H., unpublished data) and the RT6-homologue of the BALB/c mouse (F.K., unpublished data) indicate that a cluster of point mutations in the *RT6*^{a} vs. the *RT6*^{b} allele creates a single potential N-glycosylation site in RT6.1 (amino acid residues 38–40), which is absent in RT6.2 and in the RT6 homologue of the BALB/c mouse.

We thank S. Kreiker for technical assistance; F. Buck (Hamburg) for help with amino acid sequence analyses; N. Suhai (Heidelberg) for help with data bank searches; G. Butcher (Babraham, U.K.), K. Wonigeit (Hannover, F.R.G.), and W. Wurst (Göttingen, F.R.G.), for providing antibodies and genomic DNA; and H. H. Arnold (Hamburg) and J. Howard (Babraham) for laboratory space and helpful advice.

1. Butcher, G. W. & Howard, J. C. (1977) *Rat News Lett.* **1,** 12.
2. DeWitt, C. W. & McCullough, M. (1975) *Transplantation* **19,** 310–317.
3. Howard, J. C. & Scott, D. W. (1974) *Immunology* **27,** 903.
4. Miller, C. & DeWitt, C. W. (1972) *J. Immunol.* **109,** 919.
5. Thiele, H. G., Arndt, R., Stark, R. & Wonigeit, K. (1979) *Transplant. Proc.* **11,** 1636–1638.
6. Wonigeit, K. (1979) *Transplant. Proc.* **11,** 1631–1635.
7. Lubaroff, D. M., Greiner, D. L. & Reynolds, C. W. (1979) *Transplant. Proc.* **11,** 1092–1096.
8. Thiele, H. G., Koch, F., Hamann, A. & Arndt, R. (1986) *Immunology* **59,** 195–201.
9. Koch, F., Kashan, A. & Thiele, H. G. (1988) *Immunology* **65,** 259–265.
10. Butcher, G. W., Clarke, S. & Tucker, E. M. (1979) *Transplant. Proc.* **11,** 1629–1630.
11. Matsumoto, K., McCafferty, E., Neilson, E. G. & Gasser, D. (1984) *Immunogenetics* **20,** 117–123.
12. Nadeau, J. H. & Reiner, A. H. (1988) *Curr. Top. Microbiol. Immunol.* **137,** 39–40.
13. Stutman, O. (1986) *Immunol. Rev.* **91,** 159–194.
14. Thiele, H. G., Koch, F. & Kashan, A. (1987) *Transplant. Proc.* **19,** 3157–3160.
15. Thiele, H. G., Koch, F., Haag, F. & Wurst, W. (1989) *Thymus* **13,** 137–154.
16. Koch, F., Kashan, A. & Thiele, H. G. (1988) *Hybridoma* **7,** 341–353.
17. Jackson, R. A., Buse, J. B., Ritai, R., Pelletier, D., Milford, E. L., Carpenter, C. B., Eisenbarth, G. S. & Williams, R. M. (1984) *J. Exp. Med.* **159,** 1629–1635.
18. Greiner, D. L., Handler, E. S., Nakano, K., Mordes, J. P. & Rossini, A. (1986) *J. Immunol.* **136,** 148–151.
19. Greiner, D. L., Mordes, J. P., Handler, E. S., Angelillo, M., Nakamura, N. & Rossini, A. (1987) *J. Exp. Med.* **166,** 461–475.
20. Koch, F., Thiele, H. G. & Low, M. (1986) *J. Exp. Med.* **164,** 1338–1343.
21. Wonigeit, K. & Schwinzer, R. (1987) *Transplant. Proc.* **19,** 29–33.
22. Low, M. G. & Saltiel, A. R. (1988) *Science* **239,** 268–275.
23. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
24. Ausubel F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G. & Struhl, K. (1988) *Current Protocols in Molecular Biology* (Wiley, New York).
25. Koch, F., Haag, F., Kashan, A. & Thiele, H. G. (1989) *Immunology* **67,** 344–350.
26. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
27. Lipman, D. J. & Pearson, W. R. (1985) *Science* **227,** 1435–1441.
28. Kashan, A., Buck, F., Haag, F., Koch F. & Thiele, H. G. (1989) *Immunol. Lett.* **23,** 133–138.
29. Ferguson, M. A. J. & Williams, A. F. (1988) *Annu. Rev. Biochem.* **57,** 285–320.
30. Williams, A. F. & Barclay, A. N. (1988) *Annu. Rev. Immunol.* **6,** 381–405.