

# Isolation, Tissue Distribution, and Chromosomal Localization of a Novel Testis-Specific Human Four-Transmembrane Gene Related to CD20 and Fc∈RI-B

Mark D. Hulett,\*,1 Eloisa Pagler,\* June R. Hornby,\* P. Mark Hogarth,† Helen J. Eyre,‡ Elizabeth Baker, # Joanna Crawford, # Grant R. Sutherland, # Stephen J. Ohms,§ and Christopher R. Parish\*

\*Division of Immunology and Cell Biology and \Division of Biochemistry and Molecular Biology, John Curtin School of Medical Research, ANU, P.O. Box 334, Canberra, ACT 2601, Australia; †Austin Research Institute, Austin and Repatriation Medical Centre, Melbourne, Victoria 2084, Australia; and ‡Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, Adelaide, South Australia 5006, Australia

Received November 27, 2000

CD20 and the  $\beta$  subunit of the high affinity receptor for IgE (FcεRIβ) are related four-transmembrane molecules that are expressed on the surface of hematopoietic cells and play crucial roles in signal transduction. Herein, we report the identification and characterization of a human gene, TETM4, that encodes a novel four-transmembrane protein related to CD20 and Fc $\epsilon$ RI $\beta$ . The predicted TETM4 protein is 200 amino acids and contains four putative transmembrane regions, N- and C-terminal cytoplasmic domains, and three inter-transmembrane loop regions. TETM4 shows 31.0 and 23.2% overall identity with CD20 and Fc $\epsilon$ RI $\beta$  respectively, with the highest identity in the transmembrane regions, whereas the N- and C-termini and inter-transmembrane loops are more divergent. Northern blot and RT-PCR analysis suggest that TETM4 mRNA has a highly restricted tissue distribution, being expressed selectively in the testis. Using fluorescence in situ hybridization and radiation hybrid analysis, the TETM4 gene has been localized to chromosome 11q12. The genes for CD20 and FceRIB have also been mapped to the same region of chromosome 11 (11q12-13.1), suggesting that these genes have evolved by duplication to form a family of fourtransmembrane genes. TETM4 is the first nonhematopoietic member of the CD20/FcεRIβ family, and like its hematopoietic-specific relatives, it may be involved in signal transduction as a component of a multimeric receptor complex. © 2001 Academic Press

Key Words: four-transmembrane; TM4SF; tetraspanin; CD20; FcεRIβ; testis; gene localization; chromosome 11; signal transduction.

CD20, the  $\beta$  subunit of the high affinity receptor for IgE, and HTm<sub>4</sub>, comprise a family of related proteins that contain four membrane spanning regions. All three proteins are expressed specifically in hematopoietic cells; CD20 on B cells (1), FcεRIβ on mast cells and basophils (2), and HTm<sub>4</sub> on cells of myeloid and lymphoid origin (3). Both CD20 and Fc $\epsilon$ RI $\beta$  have been well characterized as playing important roles in initiating signal transduction events as components of multimeric receptor complexes. CD20 has been shown to have the capacity to regulate B cell proliferation and differentiation as part of a large cell surface complex with MHC-I, MHC-II, CD40, and the tetraspanins CD53, 81 and 82 (1, 4). Fc $\epsilon$ RI $\beta$  is a key component of the tetrameric  $\alpha\beta\gamma_2$  Fc $\epsilon$ RI complex on mast cells and basophils, and plays a crucial role in enhancing cell surface expression of the complex and amplifying signal transduction events mediated upon the interaction of receptor-bound IgE with multivalent allergen (2, 5, 6). The functional role of HTm<sub>4</sub> is unknown, but as for CD20 and FcεRIβ, it is likely to contribute to the signalling of a multimeric receptor complex. In this study, we report the isolation, tissue distribution and chromosomal localization of a human gene that encodes a novel member of the CD20/FcεRIβ/HTm<sub>4</sub> family.

#### MATERIALS AND METHODS

Isolation of RNA and first strand cDNA synthesis. Total cellular RNA was prepared by homogenising 100 mg of tissue in 1 ml of Trizol reagent (Gibco-BRL, Grand Island, NY), upon which the aqueous fraction was recovered and RNA precipitated using isopropanol. First strand cDNA was produced from 5  $\mu g$  of total RNA by priming with an oligo dT primer (NotdT, 5'-AACTGGAAGAATT-CGCGGCCGCAGGAAT18-3') using a First Strand cDNA synthesis system (Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer's instructions.



<sup>&</sup>lt;sup>1</sup> To whom correspondence should be addressed. Fax: 61-2-6249-2595. E-mail: mark.hulett@anu.edu.au.

PCR and nucleotide sequence analysis. PCR was performed on 10 ng of first strand cDNA in the presence of 100 ng of each oligonucleotide primer, 1.25 mM dNTPs, 50 mM KCl, 10 mM Tris-HCl pH 8.3 and 1.5 mM MgCl $_2$ , and 1 unit of Taq DNA polymerase (Gibco-BRL, Gaithersburg, MD) for 35 amplification cycles. 3'-RACE was performed by PCR as described above with the oligonucleotide primer TET-1 (5'-GTCATCTCCTTTCAAATTATCAC-3', hybridizes to nucleotides 24–46 of the TETM4 cDNA) and the NotdT primer (see above). Nucleotide sequencing was performed by direct sequencing of amplified cDNA fragments using an Applied Biosystems 377 sequencer.

Northern blot analysis. Northern analysis of multiple human tissue blots (Clonetech, Palo Alto, CA) was performed by probing membranes with the full length TETM4 cDNA, labelled by random priming (Megaprime DNA labelling system; Amersham, Buckinghamshire, UK), using Expresshyb solution (Clonetech, Palo Alto, CA) as specified by the manufacturers. Membranes were washed in  $1\times$  SSC for 40 min at room temperature followed by  $0.1\times$  SSC for 40 min at  $60^{\circ}\text{C}$  and exposed to X-ray film.

Southern blot analysis. 10  $\mu g$  of genomic DNA was digested with a range of restriction enzymes, separated on a 1% agarose gel, and transferred to a Hybond-N nylon filter (Amersham, Buckinghamshire, UK). The blot was probed with the full length TETM4 cDNA labelled by random priming and hybridized in a 50% formamide, 6× SSC, 0.5% SDS, 5× Denhardt's solution and 100  $\mu g/ml$  salmon sperm DNA at 42°C. The membrane was washed in 1× SSC for 40 min at room temperature followed by 0.1× SSC for 40 min at 65°C and exposed to X-ray film.

Fluorescence in situ hybridization. A 1100-bp genomic fragment of the TETM4 gene, produced by PCR amplification with oligonucleotide primers TET-2 (5'-TTCAACTCAAAGCCCCTTGC-3', hybridizes to nucleotides 155–174 of the TETM4 cDNA) and TET-4 (5'-CCTTGGATATGGTTTTAACAAAG-3', nucleotides 290–268), was nick-translated with biotin-14-dATP and hybridized in situ at a final concentration of 15 ng/ml to metaphases from two normal males. The fluorescence in situ hybridization (FISH) method was as previously described (7), with the exception that chromosomes were stained before analysis with both propidium iodide (as counterstain) and diaminophenylindole (DAPI) (for chromosome identification). Images of metaphase preparations were captured by a cooled CCD camera using the chromoScan image collection and enhancement system (Applied Imaging Int. Ltd.).

Radiation hybrid analysis. The TETM4 gene was mapped using the medium resolution Stanford G3 panel of 83 clones. Screening of the panel was performed by PCR amplification of a 1100-bp TETM4-specific fragment using oligonucleotide primers TET-2 and TET-4 (see above). Amplifications were performed on 10 ng of each sample DNA under the above conditions.

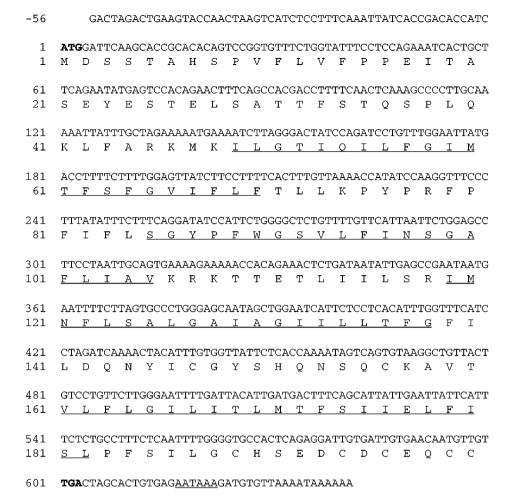
#### RESULTS AND DISCUSSION

### Identification and Isolation of the TETM4 cDNA

In order to investigate the possible existence of additional novel members of the CD20/FcεRIβ/HTm<sub>4</sub> family, the human dbEST (public expressed sequence tags, GenBank database) was searched with a consensus peptide sequence corresponding to a conserved region of the second transmembrane (TM) region of the three known human family members (GYPFWGAIFF-SISG) (3, 8, 9). A number of ESTs were identified, all from testis libraries (GenBank Accession Nos. AI149899, AA416972, AA411806, AA707529, AA470059, AA436088, AA781801, AI002083, AA435988), which

contained a region homologous to the conserved search peptide. Sequence analysis of the ESTs suggested that they were fragments of a single gene which was related to, vet distinct from, CD20/FceRIβ/HTm<sub>4</sub>. The EST sequences were assembled into a single contig of 695 bp, and examination of the compiled sequence, suggested an open reading frame that encoded for a putative protein of 200 amino acids. An oligonucleotide primer was designed to the predicted 5' untranslated region of the cDNA (TET-1, 5'-GTCATCTCCT-TTCAAATTATCAC-3', hybridizes to nucleotides 24-46 of the TETM4 cDNA) and used in 3' rapid amplification of cDNA ends (RACE)-PCR with the oligo-dT primer NotdT (see Materials and Methods), on first-strand cDNA generated from human testis total RNA. A product of 707 bp was amplified, that upon direct sequencing, was determined to encode the complete coding region predicted from the EST contig, confirming that the cDNA sequence was derived from a single mRNA. The cDNA was also cloned into the vector pCR2.1 (Invitrogen, Carlsbad, CA), and multiple clones were analysed, which revealed an identical sequence to that determined from direct sequencing of the PCR product. The nucleotide and deduced amino acid sequence encoded by the full length cDNA, designated TETM4 (for testis expressed transmembrane-4, see below), is shown in Fig. 1.

The complete TETM4 cDNA is 695 bp long and contains a canonical polyadenylation signal sequence at nucleotides 669-673 (Fig. 1). The cDNA encodes for a deduced protein of 200 amino acids with a predicted molecular weight of 22.2 kDa. Hydropathy analysis indicates the presence of four hydrophobic regions that represent four putative transmembrane domains. Using the TMpred program (http://www.ch.embnet.org/ software/TMPRED form.html), which predicts membrane spanning regions and their orientation, TETM4 is predicted to have four strong transmembrane helices which are likely to adopt a membrane topology with both the N- and C-termini intracellular. On the basis of this prediction, the TETM4 protein can be divided into the following domains; four transmembrane domains (TM-1, TM-2, TM-3 and TM-4) of 22, 21, 20 and 22 amino acids respectively, N- and C-terminal cytoplasmic domains of 48 and 18 amino acids, respectively, two extracellular loops of 14 and 22 amino acids and a short intracellular loop of 13 amino acids (Fig. 1). Significantly, both CD20 and Fc $\epsilon$ RI $\beta$  have been shown experimentally to have a topology on the cell surface as predicted here for TETM4 (1, 2). However, it should be noted that it remains possible that TETM4 may adopt an alternate topology with both the N- and C-termini extracellular. Furthermore, it also needs to be considered that TETM4 may not be expressed on the cell surface but instead is localized on a subcellular membrane(s).



**FIG. 1.** Nucleotide and deduced amino acid sequence of human TETM4. The nucleotide sequence is numbered with the first nucleotide of the translational initiation codon as  $\pm 1$ . The amino acid sequence is presented below the nucleotide sequence in single letter code, with the four putative transmembrane regions underlined. The predicted initiation codon and stop codon are in bold type and the polyadenylation signal sequence is underlined. GenBank Accession No. AF321127.

#### Analysis of the TETM4 Amino Acid Sequence

The alignment of the predicted amino acid sequence of TETM4 to that of human CD20 (8), FcεRIβ (9), and HTm<sub>4</sub> (3), indicates an overall identity of 31.0% (55.4% similarity), 23.2% (47.0%) and 26.4% (51.9%), respectively (Fig. 2). The identity of TETM4 to CD20/FcεRIβ/ HTm<sub>4</sub> is highest in the transmembrane regions, with the N- and C-termini and intra-transmembrane loop regions showing little homology. TETM4 contains a number of charged/polar residues in its TM regions, including a glutamine residue (Q54) in the first TM domain, an asparagine in each of the second (N97) and third (N121) TM domains, and a glutamic acid (E177) in the forth TM domain. All of these residues, with the exception of N97, are also conserved in CD20/FcεRIβ/ HTm<sub>4</sub> (Fig. 2). Interestingly, charged/polar residues are also common in the transmembrane regions of other multi-membrane spanning proteins such as the tetraspanins, which like CD20 and Fc $\epsilon$ RI $\beta$ , associate with other membrane molecules (10). Other interesting structural features of the TETM4 protein include two cysteine residues in its second extracellular loop region (C147 and C156), which are also found in CD20/  $Fc \in RI\beta/HTm_4$ , and may be involved in forming an intra- or inter-chain disulphide bond(s). The C-terminal cytoplasmic tail of TETM4 contains a cluster of five cysteine residues in a 12 residue stretch (C189, 194, 196, 199 and 200), which may also be involved in disulphide bond formation.

#### Identification of a TETM4 Splice Variant

The PCR amplification of the TETM4 cDNA with oligonucleotide primers TET-1 and NotdT (see Materials and Methods) from human testis also led to the isolation of a putative splice variant. Nucleotide sequence analysis indicated that this cDNA is 542 bp long and is identical to the TETM4 cDNA, however, it is missing the coding region for the third trans-

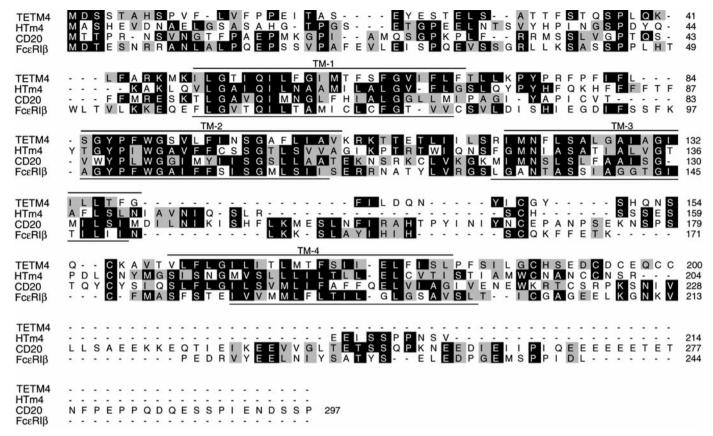
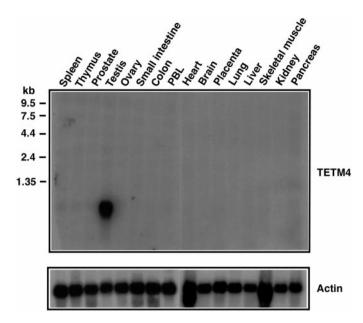


FIG. 2. Amino acid alignment of TETM4 with CD20, FcεRIβ and  $HTm_4$ . The amino acid sequences are presented in single letter code. The alignment was performed using CLUSTALW (GCG package) and adjusted manually. Gaps (–) have been introduced to maximise alignment of the sequences. Identical or similar residues between at least two sequences are shaded in black or grey, respectively. Similar residues are defined as: D, E (acidic); A, G, I, L, V (aliphatic); N, Q (side-chain containing amide group; F, W, Y (aromatic); R, H, K (basic); S, T (side-chain containing hydroxyl group). The positions of the four putative transmembrane regions for TETM4 and FcεRIβ are overlined and underlined, respectively. The predicted TM regions for  $HTm_4$  and CD20 are very similar to that shown for FcεRIβ, with the exception that CD20 contains a continuous hydrophobic stretch between TM-1 and TM-2. The GenBank Accession Nos. are: TETM4, AF321127; CD20, AAA35581; FcεRIβ, AAA60269; HTm4, AAA62319.

membrane and second extracellular domains (nucleotides 394 to 547 encoding amino acids L113 to F163). This splice variant is also represented in the EST database by two clones (GenBank Accession Nos. AA411806 and AA781801). The deduced polypeptide encoded by this cDNA would contain only three transmembrane regions, and would therefore be predicted to have a membrane topology with the C-terminal domain extracellular, as opposed to intracellular for the four-transmembrane form. A putative TETM4 protein with this different topology would be likely to have an altered function. The lack of the fourth transmembrane region may influence possible association(s) with other membrane molecules, and the shifting of the C-terminal domain from intracellular to extracellular may change any potential signalling capacity mediated through interactions with intracellular signalling proteins. Clearly, it would be of interest to determine if this variant encodes for a functional protein.

#### Tissue Distribution of TETM4 mRNA

The tissue distribution of TETM4 mRNA was investigated by Northern blot analysis of a range of human tissues. A strong band centered around 0.7 kb was detected only in testis and not in spleen, thymus, prostate, ovary, small intestine, colon, peripheral blood leukocyte, heart, brain, placenta, liver, lung, skeletal muscle, kidney or pancreas (Fig. 3). Prolonged exposure of the Northern blot failed to reveal any significant signal in any tissue other than testis. Reversetranscriptase (RT)-PCR analysis of TETM4 mRNA was also performed on first strand cDNA made from mRNA isolated from the above range of human tissues. Amplification with oligonucleotides TET-2 (see Materials and Methods) and TET-3 (5'-CAGTAACAGCCTTA-GACTGAC-3', hybridizes to nucleotides 537-516 of the TETM4 cDNA) produced the expected product of 383 bp only in testis, but not any other tissue (data not shown). These data suggest that TETM4 shows an



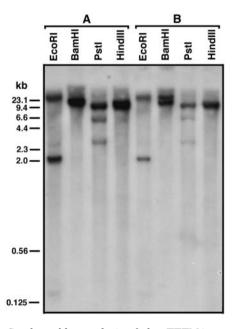
**FIG. 3.** Northern blot analysis of TETM4 mRNA expression. Multiple tissue Northern blot filters (Clonetech, Palo Alto, CA) were hybridized with  $^{32}\text{P-labelled}$  full-length TETM4 cDNA in ExpressHyb solution (Clonetech, Palo Alto, CA) as specified by the manufacturer. The filters were rehybridized with a control  $^{32}\text{P}$  β-actin cDNA and show approximately equal amounts of mRNA loaded per lane; heart and skeletal muscle have two β-actin transcripts. The positions of molecular weight makers (in kilobases) are indicated. Exposure times were 12 h for both TETM4 and β-actin.

extremely specific tissue distribution, being found only in the testis. The identification of TETM4 ESTs derived only from human testis libraries, also supports the proposed testis specific expression. Thus, in contrast to CD20/Fc $\epsilon$ RI $\beta$ /HTm $_4$ , which are all hematopoietic specific, TETM4 is the first member of this family that is expressed in a non-hematopoietic tissue.

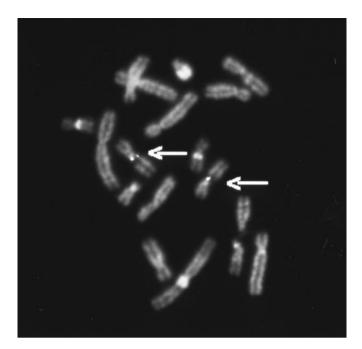
## Southern Blot Analysis and Chromosomal Localization of the TETM4 Gene

Southern blot analysis was performed on restricted human genomic DNA using the full length TETM4 cDNA as a probe. A simple banding pattern was produced that is consistent with the TETM4 gene being a single copy gene (Fig. 4). To determine the chromosome localization of the TETM4 gene, fluorescence in situ hybridization (FISH) was performed on metaphase chromosomes of two normal males using a 1100-bp TETM4-specific genomic fragment as a probe. Twenty metaphases from the first male were examined for a fluorescent signal, which was present in all 20 metaphases in the region 11q12-11q13, with 57% of the signal located in the central portion of band 11q12 (Fig. 5). Similar results were obtained from hybridizations of the probe to metaphases from the second normal male. Radiation hybrid mapping of the TETM4 gene was also performed using the medium resolution Stanford G3 panel of 83 clones. Screening of the panel by PCR amplification of a 1100 bp TETM4-specific fragment using oligonucleotide primers TET-2 and TET-4 (see Materials and Methods), indicated that *TETM4* is most closely associated with the Stanford Human Genome Centre marker SHGC-20674, with a LOD score of 13.23 (data not shown). SHGC-20674 is not ordered on a Stanford map; however, Stanford has linked it to marker SHGC-35409 which is ordered on the Stanford Radiation Hybrid Map. The markers most closely associated with TETM4 are flanked by markers D11S335 and D11S4363. Searches of the Cytogenetic Yac Bank (http://sgiweb.ncbi.nlm.nih.gov/Zjing/yac.html) placed the flanking markers D11S1335 and D11S4363 on Yac WC11.5 which spans the 11q12 cytogenetic band. The genes for CD20, FceRIB and HTm4 have also been mapped to the same region of chromosome 11 (11q12-13) (3, 11, 12). These data suggest that the TETM4, CD20, FcεRIβ, and HTm<sub>4</sub> genes have evolved by duplication and divergence of the same ancestral gene to form a family of four-transmembrane genes.

The tetraspanins comprise a distinct family of four-transmembrane molecules that are expressed on both hematopoietic and non-hematopoietic cells. In contrast to the CD20/FcεRIβ/HTm<sub>4</sub>/TETM4 family, the tetraspanins appear to form a far more extensive family and are found in species ranging from schistosomes to humans. At least 20 members have been described in



**FIG. 4.** Southern blot analysis of the *TETM4* gene. 10  $\mu g$  of human genomic DNA was isolated from two male individuals (A) and (B), restricted with a range of enzymes, and Southern analysis performed by hybridizing with random primed <sup>32</sup>P labelled TETM4 cDNA in 50% formamide, 6× SSC, 0.5% SDS and 5× Denhardt's solution. The blot was washed under high stringency conditions and exposed to X-ray film. The positions of molecular weight makers (in kilobases) are indicated.



**FIG. 5.** Chromosomal localization of the *TETM4* gene by FISH. Partial metaphases are displayed showing FISH with a TETM4 probe. Normal male chromosomes have been counterstained with DAPI. Hydridization sites on chromosome 11 are indicated by arrows. FISH signals and the DAPI banding pattern have been merged.

the human, including CD9, CD37, CD53, CD81, and CD82 (10). It is possible that like the tetraspanins, the CD20/Fc $\epsilon$ RI $\beta$ /HTm $_4$ /TETM4 family may be much larger. Indeed, we have recently identified a number of additional family members that we are currently characterising (M. Hulett, manuscript in preparation).

The testis-specific expression of TETM4 raises some intriguing questions as to its function. As described above, CD20 and  $Fc \in RI\beta$  are expressed specifically on hematopoietic cells where they form components of multimeric cell surface receptor complexes, and play important roles in signal transduction (1, 5, 6). It is therefore tempting to speculate that TETM4 may also associate with receptor complexes on the surface of specific cells in the testis and participate in signalling events. Clearly, to address these possibilities and to delineate the function of TETM4, further fundamental issues need to be addressed such as determining the cellular and subcellular distribution of TETM4 in the testis. These studies are currently in progress.

#### **ACKNOWLEDGMENTS**

This work was supported by the National Health and Medical Research Council of Australia and the Wellcome Trust.

#### REFERENCES

- 1. Tedder, T. F., and Engel, P. (1994) CD20: A regulator of cell-cycle progression of B lymphocytes. *Immunol. Today* **15**, 450–454.
- 2. Kinet, J.-P. (1999) The high affinity IgE receptor (Fc∈RI): From physiology to pathology. *Annu. Rev. Immunol.* **17**, 973–976.
- 3. Adra, C. N., Lelias, J. M., Kobayashi, H., Kaghad, M., Morrison, P., Rowley, J. D., and Lim, B. (1994) Cloning of the cDNA for a hematopoietic cell-specific protein related to CD20 and the  $\beta$  subunit of the high affinity IgE receptor: Evidence for a family of proteins with four membrane-spanning regions. *Proc. Natl. Acad. Sci. USA* **91,** 10178–10812.
- Szollosi, J., Horejsi, V., Bene, L., Angelisove, P., and Damjanovich, S. (1996) Supramolecular complexes of MHC class I, MHC class II, CD20, and tetraspan molecules (CD53, CD81, and CD82) at the surface of a B cell line JY. *J. Immunol.* 157, 2939–2946.
- 5. Lin, S., Cicala, C., Scharenberg, A. M., and Kinet, J.-P. (1996) The Fc $\epsilon$ RI  $\beta$  subunit functions as an amplifier of Fc $\epsilon$ RI  $\gamma$ -mediated cell activation signals. *Cell* **85**, 985–995.
- 6. Donnadieu, E., Jouvin, M.-H., and Kinet, J.-P. (2000) A second amplifier function for the allergy-associated Fc $\epsilon$ RI - $\beta$  subunit. *Immunity* **12**, 515–523.
- 7. Callen, D. F., Baker, E., Eyre, H. J., Chernos, J. E., Bell, J. A., Sutherland, G. R. (1990) Reassessment of two apparent deletions of chromosome 16p to an ins(11;16) and a t(1:16) by chromosome painting. *Ann. Genet.* **33**, 219–221.
- 8. Tedder, T. F., Streuli, M., Schlossman, S. F., and Saito, H. (1988) Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen of human B lymphocytes *Proc. Natl. Acad. Sci. USA* **85.** 208–212.
- 9. Kuster, H., Zhang, L., Brini, A. T., MacGlashan, W. J., and Kinet, J.-P. (1992) The gene and cDNA for the human high affinity Immunoglobulin E receptor  $\beta$  chain and expression of the complete human receptor. *J. Biol. Chem.* **267**, 12782–12787.
- 10. Maeker, H. T., Todd, S. C., and Levy, S. (1997) The tetraspanin superfamily: Molecular facilitators. *FASEB J.* **11**, 428–442.
- 11. Sandford, A. J., Shirakawa, T., Moffatt, M. F., Daniels, S. E., Ra, C., Faux, J. A., Young, R. P., Nakamura, Y., Lathrop, G. M., Cookson, W. O. C. M., and Hopkin, J. M. (1993) Localization of atopy and  $\beta$  subunit of high-affinity IgE receptor (FceRI) on chromosome 11q. *Lancet* **341**, 332–334.
- Tedder, T. F., Disteche, C. M., Louie, E., Adler, D. A., Croce, C. M., Schlossman, S. F., and Saito, H. (1989) The gene that encodes the human CD20 (B1) differentiation antigen is located on chromosome 11 near the t(11;14) (q13;q32) translocation site. *J. Immunol.* 142, 2555–2559.