Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 993–997, February 1990
Immunology

Molecular cloning of the murine BP-1/6C3 antigen: A member of the zinc-dependent metallopeptidase family

(cDNA cloning/lymphocyte differentiation/pre-B cell/aminopeptidase/Abelson murine leukemia virus)

QI WU*†‡, JILL M. LAHTI*§¶, GILLIAN M. AIR†, PETER D. BURROWS*†, AND MAX D. COOPER*†§¶‖

*Division of Developmental and Clinical Immunology and Departments of †Microbiology, ‖Pediatrics, and §Medicine, University of Alabama at Birmingham, and ¶Howard Hughes Medical Institute, Birmingham, AL 35294

Contributed by Max D. Cooper, October 30, 1989

ABSTRACT **The BP-1/6C3 antigen is a phosphorylated cell surface glycoprotein that can be identified by monoclonal antibodies on mouse pre-B cells, immature B cells, and certain stromal cell lines from bone marrow. Expression of this antigen is increased in stromal-dependent pre-B cell lines and retrovirally transformed pre-B cells. Expression of the BP-1/6C3 antigen thus correlates with proliferation and transformation of immature B-lineage cells. In this study, we report the isolation and characterization of cDNAs encoding the BP-1/6C3 antigen. Northern blot analysis revealed a major 4.1-kilobase mRNA in all BP-1/6C3$^+$ mouse pre-B lines and in a human pre-B cell line. BP-1/6C3 mRNA was either absent or truncated in BP-1/6C3$^-$ cell lines. The cDNA sequence predicts a type II integral membrane protein of 945 amino acids with an intracytoplasmic amino terminus of only 17 amino acids and a typical zinc-binding motif in its extracellular domain. BP-1/6C3 has significant homology to aminopeptidase N and is the second member of the zinc-dependent metallopeptidase gene family to be found on the surface of early B-lineage cells.**

A cell surface glycoprotein formed by two identical disulfide-linked polypeptide chains is identified in mice by the monoclonal alloantibody BP-1 and the rat monoclonal antibody 6C3 (1–3). The BP-1/6C3 molecule is expressed on early B-lineage cells in hemopoietic tissues but is not found on mature lymphocytes in peripheral lymphoid tissues. The subunits of the BP-1/6C3 antigen have a protein backbone of 110 kDa, but differentially glycosylated forms of this antigen may be expressed by different cell types (3–5). The molecule expressed on early B-lineage cells is a 140-kDa homodimer, whereas lymphogenic bone marrow-derived stromal cell lines express a 135-kDa homodimer, and a subpopulation of thymic cortical epithelial cells may express a 130-kDa homodimeric form of this molecule (1, 4, 5).

Several observations suggest that the BP-1/6C3 molecule may play an important role in regulating growth and differentiation of early B-lineage cells: (*i*) expression is limited to early stages of B-lineage differentiation (1, 3); (*ii*) interleukin 7, a bone marrow stromal cell cytokine, induces both proliferation and expression of the BP-1/6C3 antigen by marrow-derived precursors (6); (*iii*) expression of this antigen on bone marrow-derived stromal cell lines correlates with their ability to support pre-B cell growth (5); (*iv*) neoplastic pre-B cells typically express high levels of the antigen (3, 7, 8); and (*v*) the BP-1/6C3 antigen is a phosphoprotein, as are a number of other cell surface molecules that may influence cell activation (3, 9–11).

To explore further the nature and functional role of this molecule, we have purified it, obtained partial sequences of random peptides, and used them to identify cDNA clones encoding the BP-1/6C3 antigen. The predicted amino acid sequence** suggests that this protein is a member of the zinc-dependent metallopeptidase family (12).

MATERIALS AND METHODS

Protein Purification and Sequencing. BALB/c mice were inoculated subcutaneously with 1H6A, a pre-B cell line transformed by Abelson murine leukemia virus (gift of W. Michael Kuehl, National Cancer Institute, Bethesda, MD). Tumor cells (100 g) were homogenized in isotonic Tris/NaCl buffer [150 mM NaCl and 10 mM Tris·HCl (pH 7.5)] containing a mixture of protease inhibitors (20 mM ε-aminocaproic acid, antipain at 1 μg/ml, benzamidine at 10 μg/ml, 0.1 mM EDTA, 10 mM iodoacetamide, leupeptin at 1 μg/ml, aprotinin at 2.5 units/ml, chymostatin at 1 μg/ml, pepstatin A at 1 μg/ml, and 1 mM phenylmethylsulfonyl fluoride). A crude cell membrane fraction was solubilized in extraction buffer [150 mM NaCl, 10 mM Tris·HCl (pH 7.5), 1% Nonidet P-40, and the protease inhibitors], and insoluble material was removed by centrifugation (13). Solubilized membrane proteins were stirred overnight at 4°C with ricin agglutinin-conjugated Sepharose and, after extensive washing, bound glycoproteins were eluted sequentially with extraction buffer containing 0.2 M and 0.5 M β-galactoside. This eluate was stirred overnight at 4°C with 3 ml of BP-1 antibody-conjugated Sepharose (1.5–2 mg of antibody per ml) pretreated with both 0.5% diethylamine (pH 11.5) containing 0.1% Nonidet P-40 and extraction buffer. An antigen-loaded column was then prepared and washed with (*i*) Tris/NaCl buffer containing 0.1% Nonidet P-40, (*ii*) 0.5% diethylamine and 0.1% Nonidet P-40 (pH 9.0), and (*iii*) 0.5% diethylamine and 0.1% octyl glucoside (pH 10.5) before elution with 0.5% diethylamine and 0.1% octyl glucoside (pH 11.5). Fractions containing the BP-1 antigen were pooled, concentrated, and dialyzed against 20 mM phosphate buffer (pH 6.4) containing 4 M guanidine hydrochloride. The crude BP-1/6C3 antigen was reduced, carboxymethylated, and further purified by gel filtration HPLC with a Bio-Rad TSK400 column (14). The purified BP-1 antigen (1 nmol) was digested with staphylococcal V8 protease and fractionated by reverse-phase HPLC with a gradient of 0–70% acetonitrile in 0.1% trifluoroacetic acid. Individual peaks were analyzed by an Applied Biosystems 430A gas-phase protein sequencer.

Construction and Screening of cDNA Libraries and DNA Sequencing. Poly(A)$^+$ RNA from the 1H6A cell line was made into double-stranded cDNA (15) and cloned into λgt11 and λZAPII. A 56-mer oligonucleotide probe corresponding to one of the polypeptide sequences was used to screen the λgt11 library. Inserts isolated from positive clones II_{1-6} and

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: APN, aminopeptidase N; CALLA, common acute lymphoblastic leukemia antigen; NEP, neutral endopeptidase.

‡To whom reprint requests should be addressed.

**The sequence reported in this paper has been deposited in the GenBank data base (accession no. M29961).

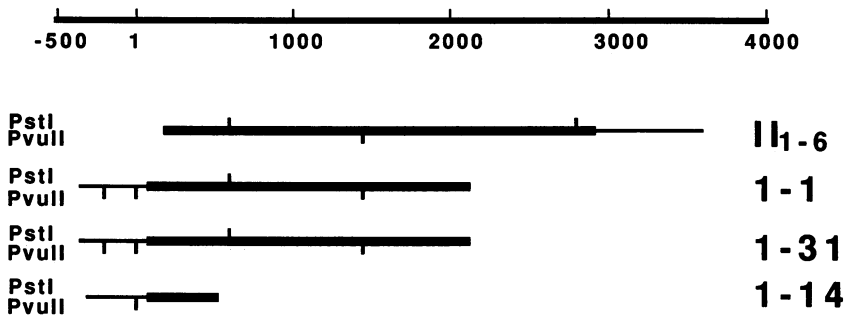

FIG. 1. Schematic representation of BP-1/6C3 cDNA clones. Independently isolated overlapping clones II_{1-6}, 1-1, 1-31, and 1-14 are shown. Clones 1-1, 1-31, and 1-14 begin ≈450 bp 5′ to the initiation methionine (ATG). The untranslated regions are indicated by lines, and the coding sequences are indicated by a solid bar. *Pst* I and *Pvu* II restriction sites are shown.

II_{2-9} were subcloned, and the DNA sequence was determined (16). The 5′ *Eco*RI/*Pst* I 340-base-pair (bp) fragment of II_{1-6} was used to screen the λZAPII library (Strategene) for clones containing the 5′ end of the coding region. DNA from three hybridizing clones was excised from λZAPII as pBluescript recombinant plasmids, subcloned, and sequenced as described above.

RNA and DNA Blot Analysis. Poly(A)$^+$ RNA (5 μg) was isolated, electrophoresed in 1% agarose/formaldehyde gels, and transferred to nitrocellulose membranes (17, 18). The filters were then hybridized with the 3363-bp II_{1-6} cDNA insert that was ^{32}P-labeled by random hexamer priming (Amersham) (19).

RESULTS

Purification and Amino Acid Sequencing of the BP-1/6C3 Protein. The BP-1/6C3 glycoprotein was isolated from the Abelson murine leukemia virus-transformed pre-B cell line, 1H6A, which expresses high levels of this antigen (mean immunofluorescence intensity is 1000 times the background). BP-1/6C3 antigen was isolated by sequential elution from ricin agglutinin and BP-1 antibody affinity columns and gel filtration HPLC to yield a highly enriched BP-1/6C3 preparation (>95%) when analyzed by $NaDodSO_4$/PAGE. All five reverse-phase HPLC peaks selected for amino acid sequencing from V8 protease digestion contained a single major sequence (See Fig. 3).

Isolation and Sequencing of the BP-1/6C3 cDNA. To isolate cDNA clones, a 56-mer oligonucleotide was synthesized corresponding to the amino acid sequence of one of the V8 proteolytic peptides: Gly-Glu-Trp-Lys-Asn-Phe-Arg-Leu-Pro-Asp-Pro-Ile-Asn-Pro-Val-His-Tyr-Asp-Leu. Inosine was utilized to compensate for third-base degeneracy. When the 1H6A cDNA library was screened with the 56-mer oligonucleotide, two clones with inserts of ≈3.5 kilobases (kb) were identified. Restriction mapping indicated that the inserts of these two clones were identical. DNA sequence analysis of clone II_{1-6} revealed that this cDNA contained a 3′ poly(A) tail (22 bp) and an open reading frame of 2679 bp (Fig. 1). The predicted amino acid sequence deduced from this cDNA contained the sequences of all five V8 peptides. At the 3′ end of the insert, a typical polyadenylylation signal, AATAAA, was identified (Fig. 2). Since the II_{1-6} cDNA

```
   1  tccaattgaa aagggaagtc agctgacaag gaaagttagt taaatttaac atcctttcat gtgtaacatc tgacttcggg ggagaaaATG AACTTTGCAG
 101  AGGAAGAGCC CTCCAAGAAA TACTGCATTA AAGGCAAACA CGTGGCCATC ATCTGTGGAG TAGTGGTGGC AGTTGGATTA ATAGTGGGAC TTTCTGTGGG
 201  TTTGACCAGG TCGTGTGAGC AGGACACAAC CCCAGCTCCT TCCCAACCTC CTCCGGAAGC CAGCACTGCC CTCCCTCCTC AGGACCAGAA TGTCTGCCCT
 301  GACAGTGAAG ATGAAAGCGG AGAATGGAAA AACTTCAGGC TGCCTGACTT CATCAATCCA GTTCACTACG ACCTGGAGGT GAAGGCCCTG ATGGAGGAAG
 401  ACAGGTATAC GGGAATAGTG ACCATCTCTG TCAACTTGAG CAAACCCACT CGTGACCTAT GGCTCCACAT CAGGGAGACC AAGATCACCA AGCTGCCGGA
 501  GCTAAGGAGG CCCTCTGGGG AGCAGGTACC AATTCGACGG TGCTTCGAGT ATAAGAAGCA GGAGTACGTG GTGATCCAGG CTGCAGAAGA CCTTGCGGCC
 601  ACCAGTGGGG ACAGTGTCTA TCGGCTGACC ATGGAGTTTA AAGGCTGGCT GAACGGTTCC CTTGTGGGTT TTTACAAAAC CACCTACATG GAGGACGGGC
 701  AAATCAGGAG CATAGCTGCC ACTGACCATG AACCAACAGA TGCCAGGAAG TCCTTCCCTT GTTTCGACGA ACCCAACAAG AAGTCAACTT ACAGTATATC
 801  CATCATCCAC CCAAAAGAAT ACTCAGCACT TTCTAATATG CCAGAAGAGA AATCAGAGAT GGTGGATGAC AACTGGAAGA AAACCACTTT TGTGAAGTCT
 901  GTCCCAATGA GCACTTACCT GGTGTGCTTT GCTGTGCATC GGTTCACTGC TATAGAGAGA AAATCCAGGA GCGGCAAACC ACTCAAGGTC TATGTCCAGC
1001  CCAATCAGAA GGAGACAGCA GAGTATGCGG CAAACATAAC CCAAGCTGTA TTTGATTACT TCGAAGACTA CTTCGCTATG GAGTATGCGC TTCCTAAACT
1101  GGATAAAATT GCTATTCCAG ATTTTGGCAC CGGCGCCATG GAAAATTGGG GACTTGTCAC TTACCGAGAA ACAAACCTGC TTTACGACCC CCTACTATCG
1201  GCCTCATCTA ACCAGCAGAG AGTGGCCAGC GTGGTTGCCC ACGAACTTGT ACACCAGTGG TTTGGAAATA CTGTGACCAT GGACTGGTGG GACGACTTGT
1301  GGCTAAATGA AGGATTTGCT TCGTTCTTCG AGTTCTTGGG AGTAAACCAC GCAGAAAAAG ACTGGCAGAT GCTCAGTCAG GTGCTGCTTG AAGATGTGTT
1401  CCCCGTGCAA GAGGACGACT CCCTGATGTC TTCACATCCA GTGGTCGTCA CCGTGTCCAC GCCAGCTGAA ATAACATCTG TGTTTGATGG GATATCATAC
1501  AGCAAGGGAG CTTCTATTCT GAGAATGCTC CAAGACTGGA TAACACCAGA GAAATTCCAA AAAGGCTGTC AGATTTACTT GAAAAAATTC CAGTTCGCGA
1601  ATGCAAAAAC TTCCGACTTT TGGGATTCAC TGCAAGAGGC AAGCAATCTG CCAGTGAAAG AAGTGATGGA CACCTGGACT AGCCAGATGG GTTATCCTGT
1701  GGTCACTGTG AGTGGAAGGC AGAACATCAC CCAGAAACGC TTTCTGTTGG ACTCCAAAGC TGATCCTTCG CAGCCACCGT CAGAGCTCGG TTACACATGG
1801  AATATTCCAG TCAGATGGGC TGATAATGAC AACTCAAGGA TCACCGTGTA CAATAGGTTA GACAAAGGAG GAATCACTCT GAATGCTAAT CTTAGCGGAG
1901  ATGCTTTTCT CAAAATCAAC CCAGATCACA TTGGGTTTTA TCGTGTAAAT TATGAAGGAG GAACGTGGGA TTGGATAGCC GAGGCTCTCT CCTCAAACCA
2001  CACGAGATTC TCCGCTGCTG ACCGGTCAAG TTTTATTGAT GATGCTTTTG CTTTGGCAAG AGCTCAACTT CTGAATTATA AAATAGCTCT GAACTTGACC
2101  ATGTATCTCA AATCAGAAGA GGATTTCCTA CCATGGGAGA GAGTCATTTC ATCTGTAAGC TACATCATTA GCATGTTTGA AGATGACAGA GAGCTGTACC
2201  CCATGATAGA GACGTACTTT CAAGGCCAAG TGAAGCCCGT TGCAGATTTG CTGGGATGGC AGGATACCGG AAGCCACATC ACAAAGTTAC TCCGGGCCTC
2301  TATCTTAGGA TTCGCATGCA AGATGGGGGA CAGAGAAGCC TTGGGCAATG CTTCCCAGTT ATTTGACTCC TGGCTGAAAG GGAGCGCAAG TATTCCGGTA
2401  AACCTCAGGC TGCTTGTGTA CCGCTATGGA ATGCAGAACT CTGGCAATGA GGCAGCGTGG AACTATACCC TAGAGCAGTA TCAGAAAACA TCGCTTGCCC
2501  AAGAGAAAGA AAAACTGCTC TATGGGTTAG CTTCAGTGAA GGATGTTAAA CTCTTGGCAA GGTATCTGGA AATGCTCAAA GACCCCAATA TTATTAAAAC
2601  TCAGGATGTA TTTACCGTCA TCCGCTACAT CTCCTACAAC AGTTATGGGA AGACAATGGC CTGGAATTGG ATACAACTCA ACTGGGACTA TCTGGTCAGC
2701  AGATTTACAA TCAATGACAG ATACCTTGGC CGGATCGTCA CCATAGCTGA GCCCTTCAAC ACTGAACTGC AGCTCTGGCA GATGCAGAGC TTTTTTGCAA
2801  AATATCCAAA TGCTGGCGCT GGAGCAAAAC CTAGAGAGCA AGTGCTGGAG ACGGTGAAGA ACAACATCGA GTGGCTAAAC GTGAACAGAC AGTCCATCAG
2901  AGAGTGGTTC GCTAGcctgc cgtaggacgg acagtgaaaa cacgtgcaag ggtgaaggca cgtcatctgt gcctctgctg caccgtgggc ggagcctgag
3001  atggcctgtt ttacagaccg cagaggggag cctttgaaac cagttctgct tttgctaagt actgtattta tgtcacgcaa agcttttaaa ttattcctct
3101  ttgtttttgc aggaagatac taatagagta tttaatatac tcagggattt ctgctaagtg ttcttcatgg gaaagtcttt acaaactgaa gggtttaata
3201  gtcacattaa tgtctttaat atcattcctt gtatcttaga tctgtgaaaa tggagcaatt tggtcctagc tttacatttt tagtaccaaa gaaacaccac
3301  ttaatcttct ctcttgattg atttttaaag tttaaatacc agtacttgca gggtgtattg ccatcttaat ctttatttta ttattcagaa ttacatagac
3401  ctaatatcat tttattcaca aatgtctcac tgctttaaat cctaccaaag taaccttggc ttaagttttg atcaaggatt gaaggaatga gccaaagaaa
3501  tgactgactg ttgcctttgg tcttacaaaa tcagacaatg ctgtattctt ttttttttt aaataaattt aaattattac cttgaaaaaa aaaaaaaaaa
3601  aaaaaa  3606
```

FIG. 2. Nucleotide sequence of the BP-1/6C3 cDNA. The sequence shown is a composite of clones 1-1, 1-31, 1-14, and II_{1-6}. Position 1 is 87 bp 5′ of the ATG initiation codon and represents the position at which clone 1-14 becomes identical to clones 1-1 and 1-31. The coding region is shown in uppercase letters, while the 5′ and 3′ untranslated region sequences are in lowercase letters. The initiation codon (ATG) and the stop codon (TAG) are in boldface and the polyadenylylation signal (AATAAA) is underlined.

clone was too small to encode the complete BP-1/6C3 molecule and no methionine residue was found at the 5′ end of the open reading frame, we screened another random-primed 1H6A cDNA library and identified three additional clones, 1-1, 1-31, and 1-14, which annealed to the 340-bp 5′ *Pst* I fragment of clone II_{1-6} (Fig. 1). Nucleotide sequencing revealed that these clones contained the remainder of the coding region (Fig. 2). All three contained an open reading frame starting from the same methionine and overlapping the initial cDNA (clone II_{1-6}). Two of the clones, 1-1 and 1-31, had an identical cDNA sequence, which included a 459-bp 5′ untranslated sequence, while one clone, 1-14, had a unique 322-bp sequence in the 5′ untranslated region (data not shown).

The BP-1/6C3 Antigen Is a Type II Integral Membrane Protein with Significant Homology to Zinc-Dependent Metallopeptidases. The complete nucleotide sequence deduced from the overlapping BP-1/6C3 cDNA clones revealed a single open reading frame of 945 amino acids (Fig. 3). Fifty-five of the 56 amino acid residues identified by amino acid sequencing of the V8 proteolytic fragments were included in this reading frame, indicating that this cDNA represents the authentic BP-1/6C3 (Fig. 3). The existence of three in-frame stop codons 5′ to the first methionine confirmed the initiation codon assignment (Fig. 2). The BP-1/6C3 cDNA sequence predicts a protein containing a polypeptide core of 107.8 kDa with nine potential N-linked glycosylation sites (Figs. 3 and 4). This prediction is consistent with the previous demonstration that the BP-1/6C3 protein has a molecular mass of 140 kDa and that removal of N-linked oligosaccharides yields a 110-kDa core protein (1, 3, 5). In addition, Tyr-328 fulfills all the criteria for predicting the occurrence of tyrosine sulfation (23).

The predicted translation product of the BP-1/6C3 cDNA lacks an amino-terminal hydrophobic segment, which would function as a signal peptide. However, the hydrophobicity plot identifies a 22-amino acid hydrophobic segment at amino acid residues 18–40. The domain amino-terminal to residue 18 contains a stretch of basic residues typical of the sequences found on the cytoplasmic side of transmembrane regions, and Ser-9 and Tyr-12 are candidates for cytoplasmic phosphorylation sites. The only other region with significant hydrophobicity occurs between residues 271 and 288, but the hydrophobicity of this region is less than that of characteristic membrane-spanning regions. All nine potential N-linked glycosylation sites are located carboxyl-terminal to the hydrophobic segment at amino acid residues 18–40. Thus topographic features predicted by the nucleotide sequence are characteristic of type II integral membrane proteins [i.e., inverted membrane orientation (extracellular carboxyl terminus) relative to classical type I integral membrane proteins] (ref. 24; Fig. 4).

Searches of the GenBank data base (release 60.0), EMBL data base (release 19.0), and National Biochemical Research Foundation protein data base (release 21.0) using the WORDSEARCH program of the University of Wisconsin Genetics Computer Group package (25) revealed no significant homology of BP-1/6C3 with either the immunoglobulin gene superfamily or with the family of receptors with inverted membrane orientation (e.g., the Lyb-2 B-cell differentiation antigen, the asialoglycoprotein receptor, and the B-cell-specific Fc receptor for IgE) (26, 27). However, it did show significant amino acid homology to human aminopeptidase N (APN or CD13; EC 3.4.11.2; 36%) (20, 21), rat APN (34%) (22), APN in *Escherichia coli* (21%) (28), and to human leukotriene A4 hydrolysase (21%) (29), all of which are members of the zinc-dependent metallopeptidase gene family (12). The homology is especially striking in the region containing a zinc-binding motif (Figs. 3 and 4).

```
          1                                                  50
BP-1/6C3 MNFAEEEPSK KYCIKGKHVA IICGV VVA VGLIVGLSVG LTRSCEQDTT
HuAPN          MAK GFY. SKSLG .LGILLG.A. .CT.IA...V Y   SQEKNK
RtAPN          MAK GFY. SKTLG .LGILLG.A. .CT.IA...V Y   AQEKNR
         51                                                 100
BP-1/6C3 PAPSQPPPEA STALPPQDQN VCPDSEDESG EWKNFRLPDF INPVHYDLEV
HuAPN    N.N.S.VAST TP SASATT. PAS  ATTLD QS.AWNRYRL P.TLKP.SYQ
RtAPN    N.ENSAIAPT LPGSTSATTS TTN   PAID ES.PWNQYRL PKTLIP.SYQ
         101                                                150
BP-1/6C3 KALMEEDR.. ...YTGIVTI SVNLSKPTRD LWLHIRETKI TKLPE
HuAPN    VT.RPYLTPN DRGLYVFKGS .TVRFTCKEA TDVI.IHS.K LNYTLSQGHR
RtAPN    VT.RPYLTPN EQGLYIFKGS .TVRFTCNET TNVI.IHS.K LNYTNKGNHR
         151                                                200
BP-1/6C3   LRRPSGEQ VPIRRCFEY. KKQEYVVIQA AEDLAATSGD SVYRLTMEFK
HuAPN    VV..GVG.SQ P.DIDKT.LV EPT..L.VHL KGS.VK    . .Q.EMDS..E
RtAPN    VA..ALGDTP A.NIDTT.LV ERT..L.VHL QGS.VKGHQ     .EMDS..Q
         201                                                250
BP-1/6C3 GWLNGSLVGF YKTTYMEDGQ IRSIAATDHE PTDARKSFPC FDEPNKKSTY
HuAPN    .E.ADD.A.. .RSE...GNV RKVV.T.QMQ AA..... ....AM.AEF
RtAPN    .E.ADD.A.. .RSE...GGN KKVV.T.QMQ AA..... ....AM.ASF
         251                                                300
BP-1/6C3 SISIIHPKEY SALSNM PEE KSEMV  DDN WKKTTFVKSV PMSTYLVCFA
HuAPN    N.TL....DL T.....L.KG P.TPLPE.PN .NVTE.HTTP K.....LA.I
RtAPN    N.TL...NNL T.....L.KD SRT LQA.PS .NVTE.HPTP K.....LAYI
         301                                                350
BP-1/6C3 VHRFTAIERK SRSGKPLKVY VQPN  QKET AEYAANITQA VFDYFEDYFA
HuAPN    .SE.DYV.KQ ASN.VLIRIW AR.SAIAAGH GD..L.V.GP ILNF.AGHYD
RtAPN    .SE.KYV.AV .PNRVQIRIW AR.SAIDEGH GD..LQV.GP ILNF.AQHYN
         351                                                400
BP-1/6C3 MEYALPKLDK IAIPDFGTGA MENWGLVTYR ETNLLYDPLL .ASSNQQRVA
HuAPN    TP.P...S.Q .GL...NA.. .......... .NS.LF...S SS...KE..V
RtAPN    TA.P.E.S.Q ..L...NA.. .......... .SA.VF..QS .SI..KE..V
         401                                                450
BP-1/6C3 SVVAAELVEQ WFGNTVTMDW WDDLWLNEGF ASFFEFLGVN HAEKDWQMLS
HuAPN    T.I..... ....L..IE. .N........ ..YV.Y..AD Y..PT.NLKD
RtAPN    T.I..... ....L..V.. .N........ ..YV.F..AD Y..PT.NLKD
         451                                                500
BP-16C3  QVLLEDVFPV QEDDSLMSSH PV   VVTVS TPAEITSVFD GISYSKGASI
HuAPN    LMV.N..Y.. MAV.ALA... .LSTPASEIN ...Q.SEL.. A.......V
RtAPN    LIV.N..Y.. MAV.ALA... .LSSPANE.N ...Q.SEL.. S.T.....V
         501                                                550
BP-1/6C3 LRMLQDWITP EKFQKGCQIY LKKFQFANAK TSDFWDSLQE ASN
HuAPN    ....SSFLSE DV.KQ.LAS. .HT.AYQ.TI YLNL..H... .V.NRSIQLP
RtAPN    ....TSFL.E DL.K..LSS. .HT..YS.TI YL.L.EH..Q .VDSQTAIKL
         551                                                600
BP-1/6C3  LPVKEVMDT WTSQMGYPVV TV SGRQNIT QKRFLLDSKA DPSQPPSELG
HuAPN    T T.RDI.NR ..L...F..I ..DTSTGTLS .EH....PDS NVTR ...FN
RtAPN    PAS.STI..R .IL...F..I ..NTSTGE.Y .EH....PTS K.TR ..DFN
         601                                                650
BP-1/6C3 YTWNIPVRWA DNDNSRITVY NRLDKGGITL NANLSGDAFL KINPDHIGFY
HuAPN    .V.IV.ITSI RDGRQQQD.W LIDVRAQND. FST ..NEWV L..LNVT.Y.
RtAPN    .L.IV.IPYL KNGKEDHY W LETE.NQSAE FQT .SNEWL L..INVT.Y.
         651                                                700
BP-1/6C3 RVNYEGGTWD WIAEALSSNH TRFSAADRSS FIDDAFALAR AQLLNYKIAL
HuAPN    ....DEEN.. KTQTQ.QRD. SAIPVIN.AQ I.N...N..S .HKVPVTL..
RtAPN    Q...DENN.. K.QNQ.QTDL SVIPVIN.AQ I.H.S.N..S .GK.SITLP.
         701                                                750
BP-1/6C3 NLTMYLKSEE DFLPWERVIS SVSYIISMFE DDRELYPMIE TYFOGQVKPV
HuAPN    .N.LF.IE.R QYM...AA.. .L..FKL..  .RS.V.GPMK N.LKK..T.L
RtAPN    SN.LF.A..T EYM...AA.. .LN.FKL..  .RS.V.GPMK R.LKK..T.L
         751                                                800
BP-1/6C3       ADLL GWQDTGSHIT KLLRASILGF ACKMGDREAL GNASQLFDSW
HuAPN    FIHFRNNTNN WREIPENLMD QYSEVNAIST ..SN.V..CE EMV.G..KQ.
RtAPN    FAYFKIKTNN WLDRPPTLME QYNEINAIST ..SS.LE.CR DLVVG.YSQ.
         801                                                850
BP-1/6C3 LKGSASIPV   NLRLLVYRY GMQNSGNEAA WNYTLEQYQK TSLAQEKEKL
HuAPN    MENPNNN.IH P...ST..CN AIAQG.E. E .DFAW..FRN AT.VN.AD..
RtAPN    MNN.DNN.IH P...ST..CN AIASCEE  . ..FAWATVPE RT.VN.AD..
         851                                                900
BP-1/6C3 LYGLASVKDV KLLARYLEML KDPNIIKTQD VFTVIRYISY NSYGKTMAWN
HuAPN    RAA.ACSKEL WI.N...SYT LN.DF.RK.. ATST.IS.TN .VI.QGLV.D
RtAPN    RSAVGRSNEV WI.N...SYT LN.DY.RK.. ATST.VS.AN .VV.QTLV.D
         901                                                950
BP-1/6C3 WIQLNWDYLV SRFTINDRYL GRIV TIAEP FNTELQLWQM QSFFAKYPNA
HuAPN    FV.S..KK.F NDYGGGSFSF SNLIQAVTRR .ST.YE.Q.L EQ.KKDNEET
RtAPN    FVRS..KK.F EDYGGGSFSF ANLIQGVTRR .SS.FE.Q.L EQ.KEDNSAT
         951                                          989
BP-1/6C3 ..GAGAKPRE QVLETVKNNI EWLNVNROSI REWFASLP.
HuAPN    GF.S.TRAL. .A..KT.A.. K.VKE.KEVV LQ..TENSK
RtAPN    GF.S.TRAL. .A..KT.A.. K.VKE.KDVV LK.. TENS*
```

FIG. 3. Alignment of the predicted BP-1/6C3 protein with the human and rat APN amino acid sequences (20–22). The translated BP-1/6C3 cDNA sequence predicts a 945-amino acid protein with nine potential N-linked glycosylation sites (3 amino acids are underlined to indicate each site). The five BP-1/6C3 V8 proteolytic polypeptides identified by protein microsequencing are also underlined. The single 22-amino acid hydrophobic segment is double underlined. The potential zinc-binding motif is boldfaced. The sequences were aligned with the sequence of human APN and rat APN using GAP. The gaps were left as spaces, and amino acids identical to those of BP-1/6C3 are presented as dots in human APN and rat APN sequences. (Note that the amino acid sequence numbering in the text is based on the linear sequence, not the gapped version shown here.)

RNA Blot Analysis. Poly(A)$^+$ RNA from a panel of BP-1/6C3$^+$ and BP-1/6C3$^-$ cell lines was analyzed by probing with the cDNA insert of clone II_{1-6} (Fig. 5*A*). A major 4.1-kb mRNA was detected in all cell lines that expressed the BP-1/6C3 antigen. These included a bone marrow-derived

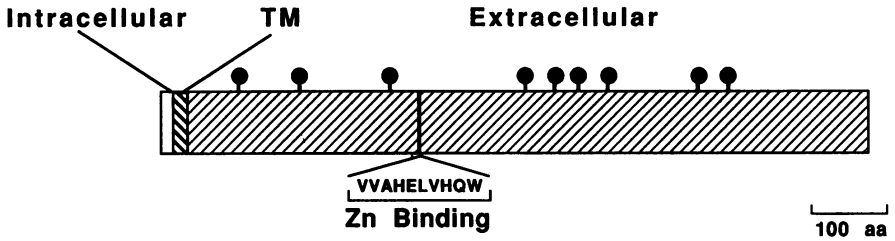

FIG. 4. Schematic representation of features of the predicted BP-1/6C3 protein. The BP-1/6C3 protein is composed of three major domains: the amino-terminal intracellular domain (open box), the transmembrane (TM) domain (▨), and the carboxyl-terminal extracellular domain (▨) within which the zinc-binding site is located (solid bar). The nine potential N-linked glycosylation sites are indicated as knobs. aa, Amino acids.

lymphopoietic stromal cell line and several pre-B cell lines (3, 6, 30). In some of the pre-B cell lines, a minor mRNA of 3.8 kb was also detected (Fig. 5*A*). The 40E1-17 pre-B cell line (Fig. 5, lane 6) is noteworthy because this BP-1/6C3$^-$ subclone was derived from a BP-1/6C3$^+$ cell line and contains a truncated 2.6-kb BP-1/6C3 message. All of the other BP-1/6C3$^-$ cell lines examined, including certain pre-B cell lines, mature B-cell lines, a pre-B/plasmacytoma hybrid, and T-cell lines, contained no detectable BP-1/6C3 mRNA.

RNA from two cell lines derived from children with acute lymphoblastic leukemia (31) was also analyzed using the murine II_{1-6} insert as a probe. A 4.1-kb transcript was detected in 697, a cell line retaining its pre-B cell characteristics (i.e., cytoplasmic μ-chain expression; ref. 31), suggesting that the BP-1/6C3 gene is conserved in humans (Fig. 5*B*, lane 2).

DISCUSSION

In this study, cDNA clones encoding the BP-1/6C3 antigen have been isolated. The overlapping clones of BP-1/6C3 cDNA contained the majority of the 5′ untranslated segment, the complete coding region, and the complete 3′ untranslated segment. A difference in the 5′ untranslated region sequences was noted between cDNA isolates, which may represent alternative splicing. A similar variation has been found in cDNA clones of the common acute lymphoblastic leukemia antigen (CALLA), CD10, another tumor-associated, early lymphocytic differentiation marker (32).

The predicted BP-1/6C3 protein contains three domains, a 17-amino acid amino-terminal domain, a 22-amino acid hydrophobic domain, and a 906-amino acid carboxyl-terminal domain. The amino-terminal cytoplasmic tail is unlikely to be a catalytic domain, such as a tyrosine kinase, since it is only 17-amino acids long (33). As with the other type II integral membrane proteins, the hydrophobic domain near the amino terminus would function as both an internal noncleavable signal peptide and a membrane-spanning domain (24).

An interesting feature of the BP-1/6C3 protein is the presence of the 10-amino acid motif for zinc binding (12) and its significant homology to human and rat APN. The CALLA protein (CD10), recently shown to be identical with neutral endopeptidase (NEP or enkephalinase; EC 3.4.24.110) also shares homology with BP-1/6C3 within the zinc-binding motif (33–35). These homologies infer a possible evolutionary relationship and underscore the importance of the zinc-binding region in enzymatic activity.

In addition to the sequence homology, the APN and the BP-1/6C3 molecules share several biochemical characteristics, including molecular mass and number of potential N-linked glycosylation sites. APN is found primarily on the brush border of the small intestine, renal proximal tubules, and placenta, but it is also found on the synaptic membranes of the central nervous system, the surface of cells of granulocyte–monocyte progenitors and their mature cell progeny (36–38). Leukemic blasts from patients with acute myeloid leukemia or with acute lymphocytic leukemia may also express APN (37, 39). Although expression of the BP-1/6C3 protein has been characterized primarily in hemopoietic and lymphopoietic tissues, it can also be found in nonlymphoid cells, such as the brush border of the small intestine, glomeruli, and proximal renal tubules (ref. 4; Q.W. and M.D.C., unpublished observations). Thus, both BP-1/6C3 and APN may exhibit lineage- and stage-specific expression by cells of both hemopoietic and nonhemopoietic tissue origin.

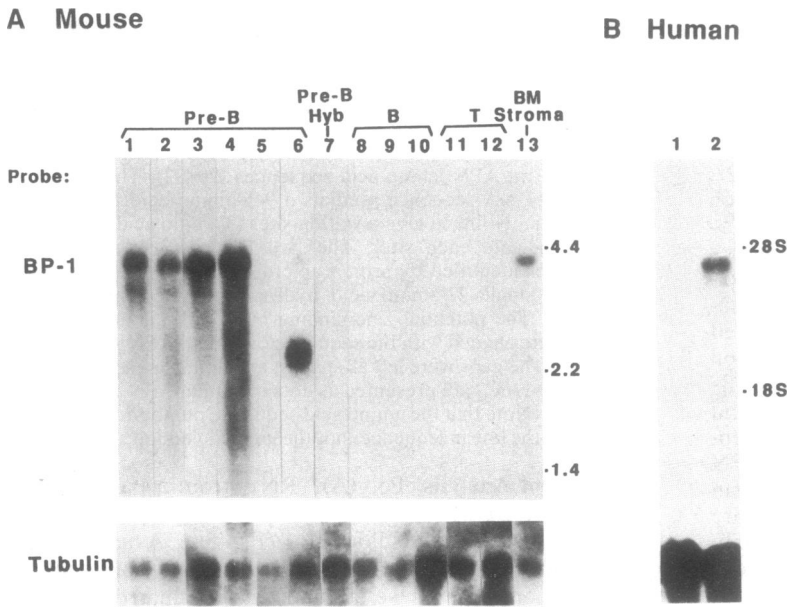

FIG. 5. Northern blot analysis of BP-1/6C3 expression. (*A*) Five micrograms of poly(A)$^+$ RNA was isolated from murine pre-B cell lines (lanes: 1, 18.81; 2, L1-2; 3, Erb#10; 4, NSF5.3; 5, 38B9; and 6, 40E1-17), a pre-B hybridoma (lane 7, 56.3), B-cell lines (lanes: 8, 4.5; 9, CH31; and 10, BCL_1), T-cell lines (lanes: 11, BW5147; and 12, YAC-1), and a bone marrow-derived stromal cell line (lane 13, BHM557). The RNA blot was hybridized with a ^{32}P-labeled 3.3-kb cDNA insert of clone II_{1-6} (*Upper*) and then stripped and rehybridized with a β-tubulin probe (*Lower*). Molecular size markers (in kDa) are indicated at the right. (*B*) Five micrograms of poly(A)$^+$ RNA for this blot was isolated from human acute-lymphoblastic leukemia cell lines 207 (lane 1) and 697 (lane 2). The same probes were used as described in *A*.

Although the BP-1/6C3 antigen closely resembles APN, the two molecules are distinct entities: (*i*) the overall sequence homology between rat APN and mouse BP-1/6C3 is only 34% in comparison with 77% homology between rat and human APN (30), (*ii*) the well-defined Ser/Thr-rich junctional domain in APN (human: Ser-43–Ser-68; rat: Ser-44–Ser-68) is totally absent in the BP-1/6C3 molecule (20), (*iii*) only two of nine cysteine residues are conserved in both molecules (Fig. 3), and (*iv*) the BP-1/6C3 molecule is not expressed on mouse myeloid cells (1–3). Thus BP-1/6C3 does not appear to be the murine APN but rather a cell surface molecule that is genetically, biochemically, and perhaps functionally related.

The function of the homologous molecules may provide insight into the biological role of BP-1/6C3. APN catalyzes the removal of amino-terminal amino acids with a preference for neutral amino acids (36). Oligonucleotides are a more efficient substrate for APN than either proteins or dipeptides (36). CALLA/NEP, an important cell surface marker on leukemic cells of pre-B phenotype, catalyzes the cleavage of peptides at the amino side of hydrophobic residues and is expressed on a broad range of normal tissues such as kidney, brain, intestine, lung, and male genital tract (40).

Both CALLA/NEP and APN have been shown to play an important role in the final degradation of small peptides on the surface of epithelial cells of the small intestine (41, 42). In other tissues, they may inactivate biologically active regulatory hormones; e.g., both NEP and APN have been observed to inactivate enkephalins and tuftsin in brain and the chemotactic peptide *N*-formylmethionylleucylphenylalanine (fMLF) in kidney and neutrophils (41–43).

The membrane-associated peptidases on lymphoid and myeloid cells could affect microenvironmental factors involved in the regulation of cell lineage progression. The cell membrane-bound BP-1/6C3 and CALLA/NEP molecules may be specifically involved in the inactivation or activation of cytokine signals that influence precursor B cells. Interleukin 7, a cytokine product of bone marrow stromal cells (44), induces both expression of cell surface BP-1/6C3 and growth by pre-B cells (6). Subsequent B-cell differentiation is accompanied by the loss of BP-1/6C3 expression and a return to the resting state as the newly formed B cell leaves the hemopoietic microenvionment (1, 45). In contrast, pre-B cells raised on feeder layers of interleukin 7-producing stromal cells continue to express high levels of BP-1/6C3 as do many neoplastic pre-B cells (1, 3, 6, 46). It may therefore be more than coincidence that neoplastic and cultured pre-B cells that express unusually high levels of the BP-1/6C3 molecule characteristically fail to differentiate further. Finally, the broad tissue distribution of BP-1/6C3 is an indication of functional versatility.

We thank Dr. Peter Neame for his help in protein purification and sequencing, Ann Brookshire for help in preparing this manuscript, and Drs. Stephen C. Peiper and Vincent J. Kidd for their critical review. This work was supported by Grants CA16693, CA13148, and P30 AI27767 awarded by the National Institutes of Health. M.D.C. is a Howard Hughes Medical Institute investigator; P.D.B. is a Scholar of the Leukemia Society of America.

1. Cooper, M. D., Mulvaney, D., Coutinho, A. & Cazenave, P. (1986) *Nature (London)* **321**, 616–618.
2. Pillemer, E., Whitlock, C. & Weissman, I. L. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4434–4438.
3. Wu, Q., Tidmarsh, G. F., Welch, P. A., Pierce, J. H., Weissman, I. L. & Cooper, M. D. (1989) *J. Immunol.* **143**, 3303–3308.
4. Adkins, B., Tidmarsh, G. F. & Weissman, I. L. (1987) *Immunogenetics* **27**, 180–186.
5. Whitlock, C. A., Tidmarsh, G. F., Muller-Sieburg, C. & Weissman, I. L. (1987) *Cell* **48**, 1009–1021.
6. Welch, P. A., Gillis, S. & Cooper, M. D. (1989) *FASEB J.* **3**, A1271 (abstr.).
7. Morse, H. C. III, Tidmarsh, G. F., Holmes, K. L., Frederickson, T. F., Hartley, J. N., Pierce, J. H., Langdon, W. Y., Dailey, M. O. & Weissman, I. L. (1987) *J. Exp. Med.* **165**, 920–925.
8. Tidmarsh, G. F., Dailey, M. O., Whitlock, C. A., Pillemer, E. & Weissman, I. L. (1985) *J. Exp. Med.* **162**, 1421–1434.
9. Wu, Q., McNagny, K., Welch, P. A., Borzillo, G. & Cooper, M. D. (1987) *Mechanisms of B Cell Neoplasia* (Editiones Roche, Basel), pp. 148–151.
10. May, M. S., Jacobs, S. & Cuatrecasas, P. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2016–2020.
11. Gallis, B., Lewis, A., Wignall, J., Alpert, A., Mochizuki, D. Y., Cosman, D., Hopp, T. & Urdal, D. (1986) *J. Biol. Chem.* **261**, 5075–5080.
12. Jongeneel, C. V., Bouvier, J. & Bairoch, A. (1989) *FEBS Lett.* **242**, 211–214.
13. Turner, M. J., Cresswell, P., Parham, P. & Strominger, J. L. (1975) *J. Biol. Chem.* **250**, 4512–4519.
14. Allen, G. (1981) in *Sequencing of Proteins and Peptides: Laboratory Techniques in Biochemistry and Molecular Biology*, eds. Work, T. S. & Burdon, R. H. (Elsevier/North-Holland, New York), pp. 128–129.
15. Gubler, U. & Hoffman, B. J. (1982) *Gene* **25**, 263–269.
16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
17. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
18. Seed, B. & Aruffo, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3365–3369.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 202–203.
20. Olsen, J., Cowell, G. M., Kønigshofer, E., Danielsen, M., Møller, J., Laustsen, L., Hansen, O. C., Welinder, K. G., Engberg, J., Hunziker, W., Spiess, M., Sjöström, H. & Norén, O. (1988) *FEBS Lett.* **238**, 307–324.
21. Look, A. T., Ashmun, R. A., Shapiro, L. H. & Peiper, S. C. (1989) *J. Clin. Invest.* **83**, 1299–1307.
22. Malfroy, B., Kado-Fong, H., Gros, C., Giros, B., Schwartz, J. & Hellmiss, R. (1989) *Biochem. Biophys. Res. Commun.* **161**, 236–242.
23. Hortin, G., Folz, R., Gordon, J. I. & Strauss, A. W. (1986) *Biochem. Biophys. Res. Commun.* **141**, 326–333.
24. Singer, S. J., Maher, P. A. & Yaffe, M. P. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1960–1964.
25. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
26. Williams, A. F. & Barclay, N. (1988) *Annu. Rev. Immunol.* **6**, 381–405.
27. Nakayama, E., Höegen, I. V. & Parnes, J. R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1352–1356.
28. Bally, M., Foglino, M., Bruschi, M., Murgier, M. & Lazdunski A. (1986) *Eur. J. Biochem.* **155**, 565–569.
29. Funk, D. C., Rådmark, O., Fu, J. Y., Matsumoto, T., Jörnvall, H., Shimizu, T. & Samuelsson, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6677–6681.
30. Davidson, W. F., Fredrickson, T. N., Rudikoff, E. K., Coffman, R. L., Hartley, J. W. & Morse, H. C., III (1984) *J. Immunol.* **133**, 744–753.
31. Findley, H. W., Cooper, M. D., Kim, J. H., Alvarodo, C. & Ragab, A. H. (1982) *Blood* **50**, 1305–1309.
32. Shipp, M. A., Richardson, N. E., Sayre, P. H., Brown, N. R., Masteller, E. L., Clayton, L. K., Ritz, J. & Reinherz, E. L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4819–4823.
33. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241**, 42–52.
34. Malfroy, B., Kuang, W., Seeberg, P. H., Mason, A. J. & Schofield, P. R. (1988) *FEBS Lett.* **229**, 206–210.
35. Letarte, M., Vera, S., Tran, R., Addis, J. B. L., Onizuka, R. J., Quackenbush, E. J., Jongeneel, C. V. & McInnes, R. R. (1988) *J. Exp. Med.* **168**, 1247–1253.
36. McDonald, J. K. & Barrett, A. J. (1986) *Mammalian Proteases: A Glossary and Bibliography: Exopeptidase* (Academic, New York), Vol. 2, pp. 59–71.
37. Griffin, J. D., Ritz, J., Nadler, L. M. & Schlossman, S. F. (1981) *J. Clin. Invest.* **68**, 932–941.
38. Griffin, J. D., Ritz, J., Beveridge, R. P., Lipton, J. M., Daley, J. F. & Schlossman, S. F. (1983) *Int. J. Cell Cloning* **1**, 33–48.
39. Sobol, R. E., Mick, R., Royston, I., Davey, F. R., Ellison, R. R., Newman, R., Cuttner, J., Griffin, J. D., Colins, H., Nelson, D. A. & Bloomfield, C. D. (1987) *N. Engl. J. Med.* **316**, 1111–1117.
40. Erdös, E. G. & Skidgel, R. A. (1989) *FASEB J.* **3**, 145–151.
41. Semenza, G. (1986) *Annu. Rev. Cell Biol.* **2**, 255–313.
42. Turner, A. J., Matsas, R. & Kenny, A. J. (1985) *Biochem. Pharmacol.* **34**, 1347–1356.
43. Nagaokea, I. & Yamashita, T. (1981) *Biochim. Biophys. Acta* **675**, 85–93.
44. Namen, A. E., Schmierer, A. E., March, C. J., Overell, R. W., Park, L. S., Urdal, D. L. & Mochizuki, D. Y. (1988) *J. Exp. Med.* **167**, 988–1002.
45. Osmond, D. S. (1986) *Immunol Rev.* **93**, 103–124.
46. Witte, P. L., Burrows, P. D., Kincade, P. W. & Cooper, M. D. (1987) *J. Immunol.* **138**, 2698–2705.