Molecular Cloning of CD44R1 and CD44R2, Two Novel Isoforms of the Human CD44 Lymphocyte "Homing" Receptor Expressed by Hemopoietic Cells

By Graeme J. Dougherty,* Peter M. Lansdorp,*‡ David L. Cooper,§ and R. Keith Humphries*‡

*From the *Terry Fox Laboratory for Hematology/Oncology, British Columbia Cancer Research Centre, Vancouver, British Columbia, V5Z 1L3, Canada; the ‡Department of Medicine, University of British Columbia, Vancouver, British Columbia, V6T 1W5, Canada; and the §Department of Pathology, Duke University Medical Center, Durham, North Carolina 27710*

Summary

In addition to the 85–95 kD CD44 species found on most hemopoietic cell types, the human myelomonocytic cell line KG1a expresses proteins of approximately 115 kD and 130 kD that react with monoclonal antibodies belonging to CD44. The possibility that these higher molecular weight species may represent novel CD44 isoforms containing additional protein sequence was investigated. CD44 cDNA clones were isolated from a plasmid-based expression library prepared from KG1a mRNA. One of the three clones obtained (clone 2.3) was found to encode a CD44 molecule of approximately 130 kD in transfected COS cells. Sequence analysis indicated that the molecule encoded by this cDNA clone, designated CD44R1, was essentially identical to CD44 except for the presence of an additional 132 amino acids inserted into the extracellular domain. This inserted region is rich in serine and threonine residues that may serve as sites of O-linked glycosylation, and contains a potential site of N-linked glycosylation and a potential site of chondroitin sulphate attachment. PCR analysis using primers that flank the inserted region present within CD44R1 identified an additional CD44 isoform, designated CD44R2, that contains only the last 69 amino acids present within the unique region of CD44R1. Peripheral blood mononuclear cells and granulocytes from normal individuals and patients with chronic myelogenous leukemia, polycythemia vera, or acute myelomonocytic leukemia, express both CD44R1 and CD44R2. In contrast, CD44R1 and CD44R2 appear to be differentially expressed in various CD44-positive cell lines. Thus KG1a, and the Epstein-Barr Virus–transformed B cell lines WalkDR4 and Way-1 express both CD44 and the CD44 isoforms CD44R1 and CD44R2, while the myeloid cell lines HL60 and U937 express high levels of CD44, but only very low levels of CD44R1 and CD44R2. The CD44-negative cell lines DHL-4, DHL-10, Jurkat, and K562 are also negative for CD44R1 and CD44R2.

The entry of PBLs into lymphoid tissues and sites of inflammation depends upon initial adhesion to vascular endothelium. Various proteins expressed on the surface of leukocytes have been implicated in this interaction including CD44, LFA-1 (CD11a/CD18), Mac-1 (CD11b/CD18), LeuM8/LAM-1, and the fibronectin receptor VLA-4 ($\alpha_4\beta_1$, CD49d/CD29) (1–4). Particular attention has focused on CD44 since antibodies directed against distinct epitopes on this molecule selectively block the binding of lymphocytes to endothelial cells in some lymphoid tissues but not others (5, 6). The mechanism by which CD44 mediates such tissue-specific "homing" remains to be determined. There are, however, several possibilities. First, there may be a single CD44 protein with the specificity of binding determined by cell-specific differences in the degree and/or type of glycosylation. Alternatively, there may be multiple CD44 proteins each sharing extensive sequence homology but differing from one another in cellular distribution and binding specificity. Until now, available evidence has supported only the first of these possibilities. For example, the alternative forms of CD44 expressed by certain cell types have been shown, in several studies, to be produced as a result of the differential modification of a common polypeptide core (7, 8). Similarly, the sequencing of CD44 cDNAs isolated from different lymphoid cell lines has revealed only one major species (9, 10). Moreover, Northern blot analysis, has demonstrated that CD44 mRNA transcripts of identical size are present in a variety of hemopoietic cell types (9, 10). Such findings have lead to the suggestion that the tissue-specificity of lymphocyte-endothelial cell interactions is not mediated by CD44 alone, but is the

J. Exp. Med. © The Rockefeller University Press • 0022-1007/91/07/0001/05 $2.00
Volume 174 July 1991 1-5

result of an interaction between CD44 and combinations of other differentially expressed adhesion molecules (1, 3, 10, 11) (Dougherty, G.J., P. Lansdorp, C. Smith, and R.K. Humphries, manuscript submitted for publication).

In the present study, we describe the molecular cloning of two novel CD44 isoforms, designated CD44R1 and CD44R2. CD44R1, differs from CD44, in that it contains an additional 132 amino acids inserted into the extracellular domain close to the membrane spanning domain. CD44R2 contains a similarly positioned insert that corresponds to the last 69 amino acids present within the unique region of CD44R1. Sequence analysis indicates that trypsin-like proteases could perhaps cleave within the inserted regions present within CD44R1 and CD44R2, generating soluble forms of CD44 that could play a role in regulating CD44-mediated adhesive interactions. Alternatively, these inserted regions may instead contribute to CD44 heterogeneity and the resultant tissue specificity of CD44-mediated leukocyte "homing".

Materials and Methods

Cell Lines and Primary Cell Populations. The myelomonocytic cell line KG1 and its derivative KG1a, the promyelocytic cell line HL60, the histiocytic cell line U937, the T cell leukemia line Jurkat, and the chronic myelogenous leukemia cell line K562 were all obtained from the American Type Culture Collection, (Rockville, MD). The B cell lymphoma lines DHL-4 and DHL-10, and the EBV-transformed B cell lines WalkDR4 and Way-1 were obtained from colleagues in the Terry Fox Laboratory. PBMCs were isolated from normal individuals and those with chronic myelogenous leukemia, polycythemia vera, and acute myelomonocytic leukemia by centrifugation through Ficoll-Hypaque. PMN cells were purified from peripheral blood Ficoll-Hypaque pellets as previously described (12).

Monoclonal Antibodies (mAbs). The production and characterization of the CD44 mAbs 7f4, 8d8 and 1a5 has been described in detail elsewhere (Dougherty, G.J., P. Lansdorp, C. Smith, and R.K. Humphries, manuscript submitted for publication). mAbs F10-44-2 (CD44), G19-1 (CD43) and 8F5 (CD54) were obtained from the Myeloid panel of the Fourth International Conference and Workshop on Human Leukocyte Differentiation antigens (13). mAb 50B4 (CD44) (14) was a generous gift from Dr. Michelle Letarte (Department of Immunology, Children's Hospital, Toronto, Canada).

cDNA Clones. pCDw44, a full length human CD44 cDNA clone in expression vector CDM8 (9), was a generous gift from Dr. Brian Seed (Massachusetts General Hospital, Boston, MA).

Construction of the KG1a cDNA Library. A plasmid-based cDNA library of approximately 2×10^6 recombinants with a mean insert size of 2 kb was prepared from Poly A^+ mRNA isolated from KG1a cells using the expression vector CDM8 as previously described (15).

Isolation of cDNA Clones Encoding CD44R1. Bacterial colonies comprising the primary cDNA library were pooled and used to transfect approximately 10^6 COS7 cells by protoplast fusion. After a 72-h incubation period to allow replication of the transfected plasmids and expression of their cDNA inserts, cells expressing surface proteins reactive with mAbs 7f4, 8d8, and 1a5 were isolated by panning (16). Plasmids were recovered from these cells by Hirt extraction, transformed back into bacteria and subjected to a further round of protoplast fusion and selection. Plasmids recovered from cells isolated in the second round of selection were again transformed into bacteria and individual colonies picked. COS7 cells transfected with plasmid DNA isolated from 4 of 10 clones examined reacted with the CD44 mAbs 7f4, 8d8, 1a5, F10-44-2, and 50B4 but not with control mAbs belonging to CD43 (G19-1) and CD54 (8F5). None of these antibodies reacted with COS7 cells transfected with the CDM8 vector alone.

DNA Sequencing. Restriction fragments of cDNA inserts were subcloned into pUC 19 and sequenced by the dideoxy chain termination method, using denatured double stranded templates and T7 DNA polymerase (17), with reaction conditions as suggested by the manufacturer (Pharmacia Fine Chemicals).

Western Blot Analysis. Western blot analysis of CD44 expression in transfected COS cells and primary cell lines was carried out as previously described (18) except that peroxidase-conjugated rabbit anti-mouse IgG (Dako Corporation, Carpinteria, CA) was used as a second antibody.

Polymerase Chain Reaction (PCR) Analysis. PCR analysis was used to identify CD44R2 and to determine the expression of both CD44R1 and CD44R2 in various cell lines and primary cell populations. Briefly, total cellular RNA was isolated from the cell types indicated using the guanidinium isothiocyanate procedure (19). cDNA was synthesized from 5 μg of total cellular RNA by reverse transcription using Moloney leukemia virus-reverse transcriptase and oligo-dT primers, and PCR reactions carried out exactly as described in the Clonetech Laboratories Inc. Cytokine Mapping process (Clonetech Laboratories Ltd., Palo Alto, CA). The oligonucleotide primers used were 23mers 5'-CCGACAGCACAGACAGAATCCCT-3' and 5'-ATCTGATTCAGATCCATGAGTGG-3'. 5 μl aliquots of the PCR reaction were separated on a 2% agarose gel, transferred to nylon membranes and probed as described (20–22).

Results and Discussion

The predominant CD44 species expressed by lymphocytes and other hemopoietic cell types has a mol wt of approximately 85–95 kD (1–7, 9, 10). In the myelomonocytic cell line KG1a however, additional species of approximately 115 kD and 130 kD are also evident (Fig. 1). We considered the possiblity that these additional species may represent novel isoforms of CD44 containing additional peptide sequence. To investigate this possibility, CD44 cDNAs were cloned from a plasmid-based expression library prepared from KG1a mRNA. The CD44 cDNAs obtained were transfected into COS7 cells and the proteins produced examined by Western blot.

As shown in Fig. 1, two of the three cDNAs isolated (clones 2.2 and 2.7) encoded CD44 molecules of approximately 85 kD. Although somewhat smaller and more heterogeneous than the form of CD44 expressed by KG1a cells, these molecules are nevertheless similar in size to the CD44 protein produced in COS7 cells transfected with the human CD44 cDNA clone pCDw44 (9). In contrast, the third clone isolated, designated 2.3, encoded a molecule of approximately 130 kD similar in size to the larger of the two unusual CD44 species expressed in KG1a.

Clone 2.3 was sequenced and shown to encode a molecule that differed from previously described CD44 cDNAs (9, 10) by the presence of an additional 132 amino acids inserted into the extracellular domain proximal to the transmembrane re-

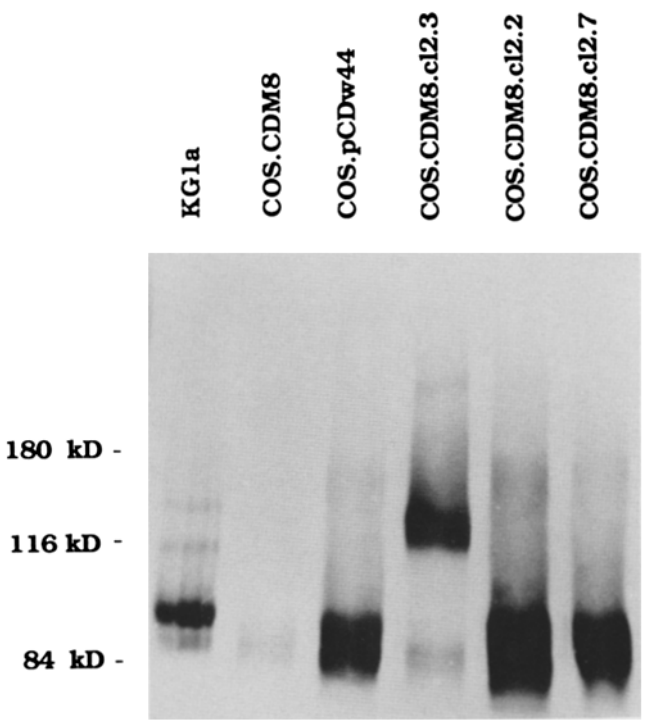

Figure 1. Western blot analysis of CD44 expression in KG1a cells, and COS cells transfected with human CD44 cDNAs. COS cells were transfected with plasmid DNA by electroporation. 72 h later, cells were harvested with PBS containing 2 mM EDTA. Whole cell detergent lysates were separated on a 5% SDS-PAGE gel, transferred to nitrocellulose and CD44 proteins identified by staining with mAbs 8d8 or 50B4, followed by a peroxidase-conjugated rabbit anti-mouse IgG second antibody. The reaction was developed by incubating the blots in PBS containing 0.06% (w/v) diaminobenzidine and 0.012% (v/v) H_2O_2.

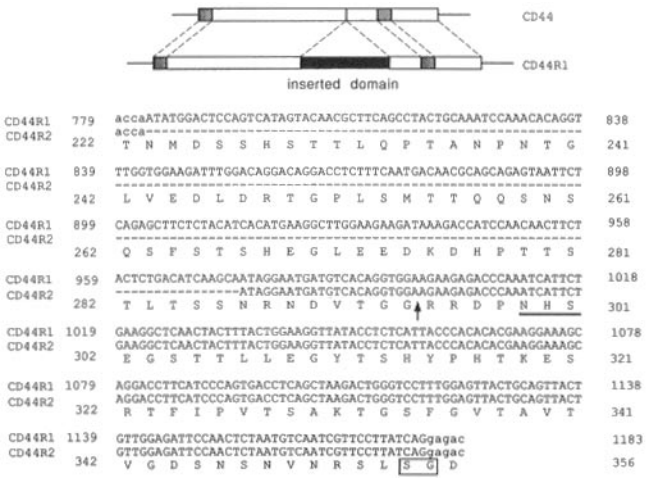

Figure 2. Nucleotide and deduced amino acid sequence of the inserted regions present in CD44R1 and CD44R2. CD44 and CD44R1 are drawn to scale. Shaded areas indicate the putative hydrophobic leader and membrane-spanning domains. A potential N-linked glycosylation site (Asn-X-Ser/Thr) is underlined. A potential chondroitin sulphate attachment site (Ser-Gly) is boxed. The arginine dipeptide that may serve as a site of cleavage by trypsin-like proteases is arrowed.

gion (Fig. 2). This inserted region is rich in serine and threonine residues which could potentially serve as sites of O-linked glycosylation. It also contains an additional N-linked glycosylation site and an additional serine-glycine dipeptide that may serve as a site of attachment for chondroitin sulphate (23). Interestingly, this region also contains an arginine dipeptide, a motif that has been shown in other cell surface proteins to serve as a site of posttranslational cleavage by trypsin-like proteases (24–26). Cleavage at this site would produce a soluble form of CD44. A soluble protein related to CD44 has been identified in the serum (27). Although the function of this molecule is not known, it has been suggested to play a role in regulating CD44-mediated cellular interactions (2). We have designated this novel form of CD44, CD44R1.

RNA isolated from various primary cell populations and cell lines was analyzed for the expression of CD44 transcripts by PCR using primers that flank the unique region present within CD44R1. Ethidium bromide staining of the PCR products produced, gave a major band of approximately 90 bp in the CD44-positive cell lines KG1a, HL60, U937, WalkDR4, Way-1, and in PBMC and granulocyte preparations (data not shown). This species was not present in the CD44-negative cell lines DHL-4, DHL-10, Jurkat, and K562, and corresponds to the expected size of material produced as a result of the amplification of CD44 (9, 10). Several less highly expressed products were also evident in material amplified from some of the CD44-positive cell populations. As shown in Fig. 3, two of these species (487 bp and 395 bp) were found to hybridize with a probe spanning the CD44R1 insert. These PCR products were subcloned into pUC-based vectors and sequenced. The sequence of the larger of the two PCR products was found to correspond exactly to CD44R1. The smaller PCR product in contrast, contained an identically positioned insert that encoded only the most COOH-terminal 69 amino acids present within the inserted region of CD44R1 (Fig. 2). We have designated this latter CD44 isoform CD44R2. The inserted region of CD44R1 appears to be encoded by three exons present within a single CD44 gene located on the short arm of chromosome 11 (data not shown). Only the third of these three exons is present in CD44R2. Presumably, CD44, CD44R1 and CD44R2 are generated by alternative slicing of the CD44 gene.

As shown in Fig. 3, PBMC and granulocytes (PMN) from normal individuals and patients with chronic myelogenous leukemia, polycythemia vera, or acute myelomonocytic leukemia express both CD44R1 and CD44R2. In contrast, CD44R1 and CD44R2 appear to be differentially expressed in various CD44-positive cell lines. Thus KG1a, and the EBV-transformed B cell lines WalkDR4 and Way-1 express both CD44 (Dougherty, G.J., P. Lansdorp, C. Smith, and R.K. Humphries, manuscript submitted for publication) and the CD44 isoforms CD44R1 and CD44R2, while the myeloid cell lines HL60 and U937 express high levels of CD44 (Dougherty, G.J., P. Lansdorp, C. Smith, and R.K. Humphries, manuscript submitted for publication), but only very low levels of CD44R1 and CD44R2. The CD44-negative cell lines DHL-4, DHL-10, Jurkat and K562 are also negative for CD44R1 and CD44R2.

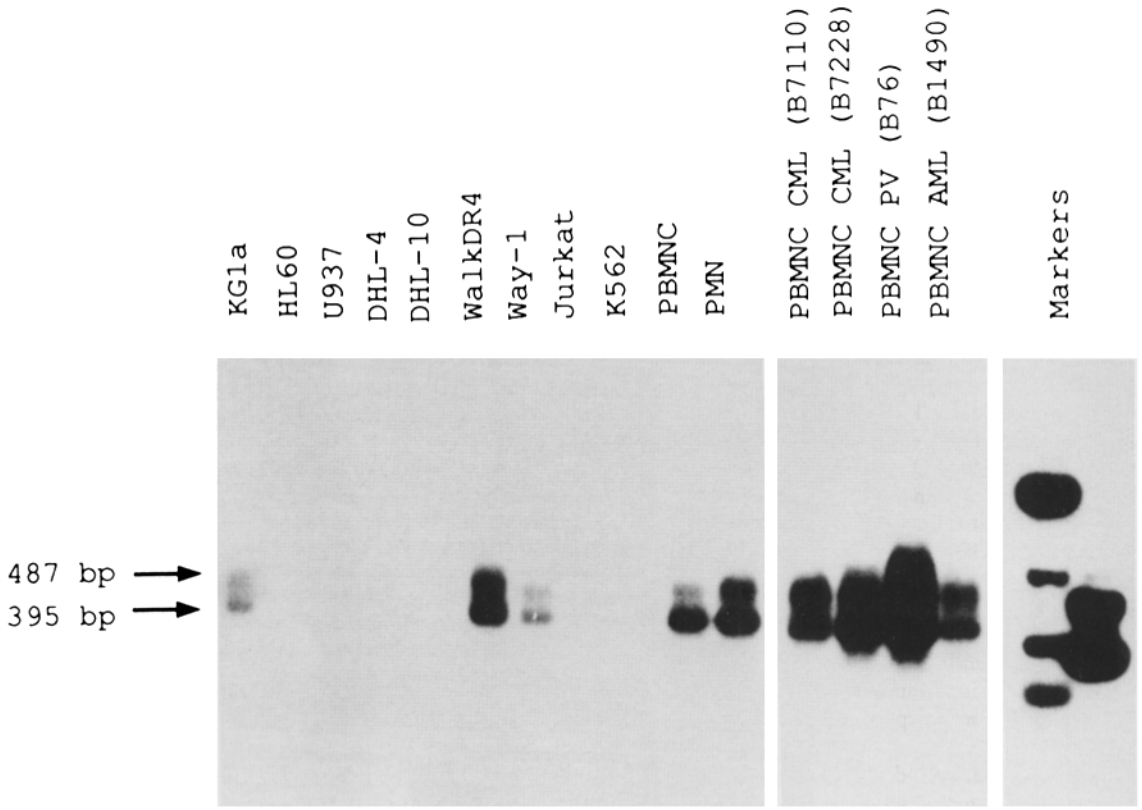

Figure 3. PCR analysis of CD44R1 and CD44R2 expression in various cell lines and primary cell populations. PCR was performed on cDNA prepared from the indicated cell lines using primers that flank the inserted region present within CD44R1. Amplified material was separated on a 2% agarose gel, transferred to nylon and hybridized with a probe that encompassed the unique region present in CD44R1 (Hinf1 788 - Hinf1 1145). PBMNC, peripheral blood mononuclear cells; PMN, polymorphonuclear cells; CML, chronic myelogenous leukemia; PV, polycythemia vera; AML, acute myelomonocytic leukemia.

Recently it has been demonstrated that CD44 mediates the attachment of lymphocytes to endothelium via recognition of hyaluronate (28, 29). The amino terminal region of CD44 is composed of two domains with extensive sequence homology to cartilage link proteins (9, 10) which function to stabilize large aggregates of cartilage proteoglycans via recognition of hyaluronate (30). Interestingly, mAb Hermes 3, which inhibits the binding of lymphocytes to mucosal endothelium but not synovial or peripheral lymph node endothelium (5, 6), recognizes an epitope on CD44 that is distinct from the putative hyaluronate-binding domain (10). This finding raises the possibility that other regions of CD44 may contribute to the specificity of hyaluronate binding perhaps through the recognition of other cell surface molecules (10). Although the functional importance of the inserted region present in CD44R remains to be determined, it is tempting to speculate that this domain may somehow contribute to the tissue specificity of CD44 mediated "homing".

We thank Patricia Rosten and Cam Smith for expert technical assistance.

This work was supported by grants from Medical Research Council (MRC) of Canada and the National Cancer Institute (NCI) of Canada with core support from the British Columbia Cancer Agency and the British Columbia Cancer Foundation.

Address correspondence to Graeme J. Dougherty, Terry Fox Laboratory for Hematology/Oncology, British Columbia Cancer Research Centre, 601 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada.

Received for publication 1 November 1990 and in revised form 19 March 1991.

References

1. Duijvestijn, A., and A. Hamann. 1989. Mechanisms and regulation of lymphocyte migration. *Immunol. Today.* 10:23.
2. Haynes, B.F., M.J. Telen, L.P. Hale, and S.M. Denning. 1989. CD44-A molecule involved in leukocyte adherence and T-cell activation. *Immunol. Today.* 10:423.
3. Stoolman, L.M. 1989. Adhesion molecules controlling lymphocyte migration. *Cell.* 56:907.
4. Osborn, L. 1990. Leukocyte adhesion to endothelium in adhesion. *Cell.* 62:3.
5. Jalkanen, S.T., A.C. Steere, R.I. Fox, and E.C. Butcher. 1986. A distinct endothelial cell recognition system that controls lymphocyte traffic into inflamed synovium. *Science (Wash. DC).* 233:556.
6. Jalkanen, S.T., R.F. Bargatze, J. de los Toyos, and E.C. Butcher. 1987. Lymphocyte recognition of high endothelium: antibodies to distinct epitopes of an 85-95 kD glycoprotein antigen differentially inhibit lymphocyte binding to lymph node, mucosal or synovial endothelial cells. *J. Cell. Biol.* 105:983.
7. Jalkanen, S.T., M. Jalkanen, R. Bargatze, M. Tammi, and E.C. Butcher. 1988. Biochemical properties of glycoproteins involved in lymphocyte recognition of high endothelial venules in man. *J. Immunol.* 141:1,615.
8. St. John, T., W.M. Gallatin, and R.L. Idzerda. 1989. CD44 expression results in the synthesis of unusual RNAs and multiple protein species containing core polypeptide sequences. *Reg. Immunol.* 2:300.
9. Stamenkovic, I., M. Amiot, J.M. Pesando, and B. Seed. 1989. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell.* 46:1,057.
10. Goldstein, L.A., D.F.H. Zhou, L.J. Picker, C.N. Minty, R.F. Bargatze, J.F. Ding, and E.C. Butcher. 1989. A human lymphocyte homing receptor, the Hermes antigen, is related to cartilage proteoglycan core and link proteins. *Cell.* 56:1,063.
11. Holzmann, B., B.W. McIntyre, and I.L. Weissman. 1989. Identification of a murine Peyer's patch-specific lymphocyte homing receptor as an integrin molecule with an α chain homologous to human VLA-4α. *Cell.* 56:37.
12. Dooley, D.C., J.F. Simpon, and H.T. Merymann. 1982. Isolation of large numbers of fully viable human neutrophils: a preparative technique using Percoll. *Exp. Hematol.* 10:591.
13. Knapp, W., B. Dorken, W.R. Gilks, E.P. Rieber, R.E. Schmidt. 1989. Leukocyte Typing IV. *In* White Cell Differentiation Antigens. H. Stein and A.E. G. Kr. von dem Borne, editors. Oxford University Press, Oxford.
14. Letarte, M., S. Iturbe, and E.J. Quackensbush. 1985. A glycoprotein of molecular weight 85,000 on human cells of B-lineage: Detection with a family of monoclonal antibodies. *Mol. Immunol.* 22:113.
15. Dougherty, G.J., R.J. Kay, and R.K. Humphries. 1989. Molecular cloning of 114/A10, A cell surface antigen containing highly conserved repeated elements, which is expressed by murine hemopoietic progenitor cells and interleukin-3 dependent cell lines. *J. Biol. Chem.* 264:6,509.
16. Kay, R., F. Takei, and R.K. Humphries. 1990. Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens. *J. Immunol.* 145:1,952.
17. Hattori, M., and Y. Sakaki. 1985. Dideoxy sequencing methods using denatured plasmid templates. *Anal. Biochem.* 152:232.
18. Dougherty, G.J., S.T. Dougherty, R.J. Kay, P. Lansdorp, and R.K. Humphries. 1989. Identification and characterization of 114/A10, an antigen highly expressed on the surface of murine myeloid and erythroid progenitor cells and IL-3–dependent cell lines. *Exp. Hematol.* 17:877.
19. Davis, L.G., M.D. Dibner, and J.F. Battey. 1986. Basic Methods in Molecular Biology. Elsevier, New York.
20. Hodgson, C.P., and R.Z. Fisk. 1987. Hybridization probe size control: optimized 'oligolabelling'. *Nucleic Acids Res.* 15:6,295.
21. Kay, R.J., R.J. Boissey, R.H. Russnak, and E.P.M. Candido. 1986. Efficient transcription of a *Caenorhabditis elegans* heat shock gene pair in mouse fibroblasts is dependent on multiple promoter elements which can function bidirectionally. *Mol. Cell. Biol.* 6:3,134.
22. Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, New York.
23. Mann, D.H., Y. Yamaguchi, M.A. Bourdon, and E. Ruoslahti. 1990. Analysis of glycosaminoglycan substitution in decorin by site-directed mutagenesis. *J. Biol. Chem.* 265:5,317.
24. Tempest, P.R., M.R. Stratton, and C.S. Cooper. 1988. Structure of the met protein and variation of met protein kinase activity among human tumour cell lines. *Br. J. Cancer.* 58:3.
25. Hemler, M.E., Y. Takada, M. Elices, and C. Crouse. 1989. Structure and functions for the adhesion receptors VLA-2 and VLA-4: comparisons to other members of the integrin superfamily. *In* Leukocyte Adhesion Molecules. Structure, Function, and Regulation. T.A. Springer, D.C. Anderson, A.S. Rosenthal, and R. Rothlein, editors. Springer-Verlag, New York. pp. 44–57.
26. Nophar, Y., O. Kemper, C. Brakebusch, H. Engelmann, R. Zwang, D. Aderka, H. Holtmann, and D. Wallach. 1990. Soluble forms of tumor necrosis factor receptors (TNR-Rs). The cDNA for the type 1 TNF-R, cloned using amino acid data of its soluble form, encodes both the cell surface and a soluble form of the receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3,269.
27. Lucas, M.G., A.M. Green, and M.J. Telen. 1989. Characterization of the serum In(Lu)-related antigen: Identification of a serum protein related to erythrocyte p80. *Blood.* 73:596.
28. Aruffo, A., I. Stamenkovic, M. Melnick, C.B. Underhill, and B. Seed. 1990. CD44 is the principal cell surface receptor for hyaluronate. *Cell.* 61:1,303.
29. Miyake, K., C.B. Underhill, J. Lesley, and P.W. Kincade. 1990. Hyaluronate can function as a cell adhesion molecule and CD44 participates in hyaluronate recognition. *J. Exp. Med.* 172:69.
30. Goetinck, P.F., N.S. Stripe, P.A. Tsonis, and D. Carlone. 1987. The tandemly repeated sequences of cartilage link protein contain the sites of interaction with hyaluronic acid. *J. Cell. Biol.* 105:2,403.