Cloning and Expression of cDNAs for the α Subunit of the Murine Lymphocyte-Peyer's Patch Adhesion Molecule

Heinz Neuhaus,* Mickey C-T. Hu,*‡ Martin E. Hemler,§ Yoshikazu Takada,§ Bernhard Holzmann,* and Irving L. Weissman*‡

*Laboratory of Experimental Oncology, Department of Pathology; and ‡Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305; and §Dana-Farber Cancer Institute, Boston, Massachusetts 02115

Abstract. cDNA clones encoding the α chain of the murine lymphocyte-Peyer's patch adhesion molecule (LPAM), which is associated with lymphocyte homing, have been isolated by screening with the human VLA-4 (α_{4h}) probe. Several α_4 antigenic determinants were identified on COS-7 cells after transfection. From overlapping clones, ~5 kb of contiguous nucleotide sequence have been determined, encoding a protein sequence of 1039 amino acids for the LPAM α chain (α_{4m}). LPAM is a member of the integrin family of cell-surface heterodimers, and α_{4m} is the murine homologue of the human α_{4h} chain. The two proteins have a total sequence similarity of 84%, with an almost perfect conservation (31/32 amino acids) in the cytoplasmic domain. Like α_{4h}, α_{4m} is distinct from other integrin α chains because it has neither an I-domain nor a COOH-terminal cleavage site. The positions of the characteristic Cysteine residues are conserved, and a putative protease cleavage site is located near the middle of the protein sequence. The NH_2-terminal part of the protein contains seven homologous repeats, and three of them include putative divalent cation-binding sites. These sites are among the most conserved between the α_{4m} sequence and other α chains, and may therefore be involved in the binding of integrin α and β chains. An additional cDNA clone was isolated which shares a sequence of perfect homology with the α_{4m} encoding cDNAs, but has a unique 3' poly-A end. This observation correlates with the fact that three discrete murine RNA bands are observed in Northern blot experiments using α_{4m} as a probe, whereas only two human RNA species are described for α_{4h}, indicating a higher complexity for murine than for human sequences.

A critical step for the onset of an immune reaction resides in the ability of lymphocytes to cross blood vascular endothelium in order to enter various lymphoid organs, thus, recirculating between blood and lymph. The interaction between circulating lymphocytes and endothelium occurs at specialized sites in the postcapillary venules which display unusually high-walled endothelia and have therefore been termed high endothelial venules (HEVs; Stamper and Woodruff, 1976). This interaction is organ-specific. Lymphocytes from peripheral nodes will preferentially bind to peripheral node HEVs, and binding to HEVs in Peyer's patches occurs preferentially with lymphocytes from Peyer's patches (Kraal et al., 1983).

Several murine lymphomas specifically bind to Peyer's patch HEV only, and home to Peyer's patches in vivo (Butcher and Weissman, 1980). One of them, TK1, was used to raise a rat mAb, R1-2, which completely and selectively inhibits lymphocyte adhesion to Peyer's patch HEV (Holz-

1. *Abbreviations used in this paper*: HEV, high endothelial venule; LPAM, lymphocyte-Peyer's patch adhesion molecule.

mann et al., 1989). This antibody immunoprecipitates two molecular structures, lymphocyte-Peyer's patch adhesion molecules 1 and 2 (LPAM-1 and LPAM-2). Both are heterodimers composed of an α chain of ~200 kD, associated to either one of two β chains of ~150 kD (LPAM-1) and ~130 kD (LPAM-2), respectively (Holzmann and Weissman, 1989a). R1-2 inhibits lymphocyte adhesion to Peyer's patches by binding an epitope on the LPAM α chain, which is therefore likely to be directly involved in the binding of the cell surface homing receptors to their ligands on endothelial cells.

The LPAM α chain is antigenically related to the human VLA-4 α chain (Holzmann et al., 1989). VLA-4 is a member of the integrin family of cell-surface molecules, a family of $\alpha\beta$ heterodimers involved in cell–matrix and cell–cell interactions (Hynes, 1987, 1988). The α chain of the human VLA-4 (α_{4h}) is associated to the integrin β_1 chain (Hemler et al., 1987). The murine homologue to $\alpha_{4h}\beta_1$ is the LPAM-2 molecule ($\alpha_{4m}\beta_1$), whereas association of α_{4m} to β_p, a distinct and previously unknown β chain, forms LPAM-1 (Holzmann and Weissman, 1989a). Anti-α_{4h} antibodies successfully blocked adhesion of human lymphocytes to Peyer's patch HEV (Holzmann and Weissman, 1989b), indicating

that either there is a functioning human equivalent to the mouse LPAM-1 ($\alpha_4\beta_p$) molecule, or that the human VLA-4 ($\alpha_4\beta_1$) can play a role in lymphocyte homing to Peyer's patches. If the latter possibility is true, it is unlikely that VLA-4 would be a specific Peyer's patch homing receptor, since VLA-4 already has several adhesive functions (Hemler et al., 1990) which are not targeted towards a single lymphoid organ. For example, VLA-4 mediates adhesion to one or more domains within the alternatively spliced CS-1 region of fibronectin (Wayner et al., 1989; Guan and Hynes, 1990). VLA-4 recognizes the cellular ligand VCAM-1 on activated endothelial cells (Elices et al., 1990). VLA-4 triggers homotypic aggregation for most VLA-4-positive leukocyte cell lines (Bednarczyk and McIntyre, 1990; Campanero et al., 1990). VLA-4 may also participate in cytolytic T-cell interactions with target cells (Claybergerer et al., 1987; Takada et al., 1989).

Two molecular species have been described which have lymphoid organ homing characteristics in mouse—the lymph node homing receptor (LHR), a 90-kD glycoprotein detected by mAb MEL-14 (Gallatin et al., 1983), and $\alpha_{4m}\beta_p$, which is displayed on the cell surface of Peyer's patch-homing lymphocytes, either alone or in combination with $\alpha_{4m}\beta_1$. Other molecules may play an accessory role in homing, as for example, LFA1 (Hamann et al., 1988). In man, however, it has been reported that different epitopes of a single 90-kD glycoprotein (gp90[HERMES]), defined by mAbs, mediate adhesion of lymphocytes to HEV in peripheral lymph nodes, appendix or Peyer's patches, and inflamed synovium (Jalkanen et al., 1987). gp90[HERMES] is identical to CD44, but unrelated to the murine LHR molecule and its human homologue (Siegelman et al., 1989; Lasky et al., 1989; Stamenkovic et al., 1989; Goldstein et al., 1989; Siegelman and Weissman, 1989). PGP-1, the murine homologue of CD44, shows no correlation in its expression with the homing characteristics of lymphocytes to lymph nodes (unpublished data). There is no evidence so far that the CD44 glycoprotein is involved in homing of murine lymphocytes to Peyer's patches.

To begin an analysis of molecules involved directly or as accessory molecules in homing and other lymphocyte adhesion processes, the molecular species for all that are thought to play a role must be identified in a species, mouse, where both homing in vivo and adhesion in vitro can be tested. Here we describe the structure of the mouse α_4 molecule which was isolated by α_{4h} cDNA selection. It correlates in its expression to Peyer's patch homing cells, transfects several α_{4m} determinants to COS-7 cells, shows a high degree of homology to α_{4h}, and, together with α_{4h}, forms a unique subgroup of integrin α chains that are not disulfide-linked and have no I-domain.

Materials and Methods

cDNA Cloning

5 μg total poly-A$^+$ RNA from cell line TK1 were reverse transcribed using AMV reverse transcriptase (Molecular Genetic Resources, Tampa, FL), either with oligo-dT or random hexamer primers. For second strand synthesis, the RNaseH/PolymeraseI method was applied (Gubler and Hoffman, 1983), and NotI/EcoRI linker-adaptors (Invitrogen, San Diego, CA) were ligated to the double stranded cDNA. The DNA was size selected on Agarose gels for molecules over 2 kb in size (oligo-dT primed), or over 1 kb (random primed), respectively. After electroelution, cDNA was ligated to

EcoRI-cut and phosphatase treated λ ZAP vector DNA (Stratagene, La Jolla, CA), packaged in vitro, and phage particles were used to infect *E. coli* host cells. Complexity of the libraries was 0.75×10^6 (oligo-dT primed) and 4.5×10^6 (random primed). The libraries were screened in a first round using pGEM4.10, a cDNA clone encoding human α_4 (Takada et al., 1989), and in a second round using a fragment isolated from a mouse α_4 cDNA clone. Labeled probes for hybridizations were made using hexamer primers and Klenow enzyme (Feinberg and Vogelstein, 1983). Clones screened as positive were plaque purified twice, and the insert excised in vivo into the pBluescript vector by infection with the helper phage R408, according to the supplier's instructions (Stratagene).

Transient Expression of α_{4m} cDNA in COS-7 Cells

The α_{4m} cDNA clone of pZ5/6 plasmid was inserted into the mammalian expression vector pME18S (K. Maruyama, unpublished data), which primarily contains the promoter of the SRα vector (Takobe et al., 1988), in either sense or antisense orientation. COS-7 cells were grown to semiconfluence on 100-mm petri dishes in a CO_2 incubator at 37°C for 48 h, then transfected with either sense or antisense α_{4m}/pME18S plasmid DNA by the standard electroporation technique (Chu et al., 1987). 20 μg of plasmid DNA were used for each transfection. After electroporation, transfected COS-7 cells were incubated on ice for 10 min and then grown in DME containing 10% FBS for another 48 h at 37°C before harvesting. Cell surface expression of the α_{4m}-specific antigens was examined by specific antibody staining and fluorescence-activated cytometric analysis (FACS).
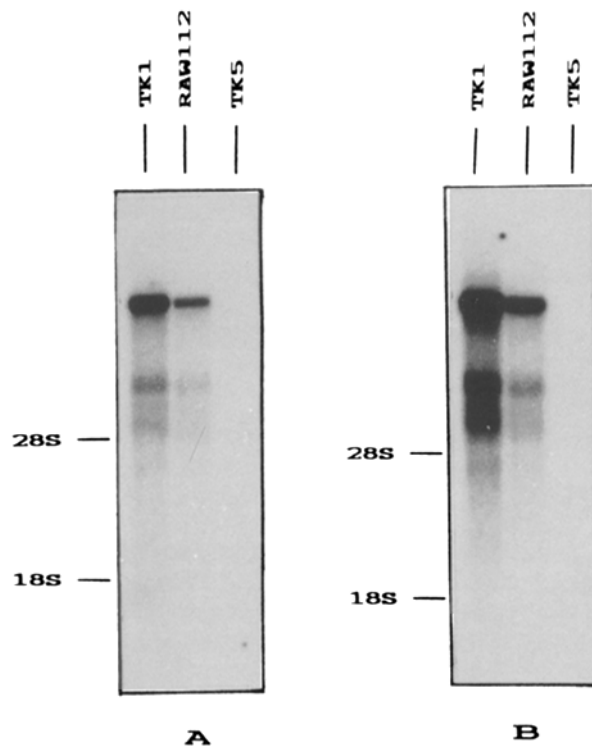

Figure 1. Comparison of RNA blot hybridization with human and murine cDNA probes. 2.5 μg of poly-A$^+$ RNA from cell lines TK1, RAW112, and TK5 were electrophoresed in formaldehyde-containing agarose gels and transferred to nylon membranes. These Northern blots were hybridized with a human α_4 cDNA clone (*A*) and a clone isolated from a TK1 cDNA library (*B*). Hybridization and washing in *A* was performed under reduced stringency (40% Formamide, 5× SSC, 5× Denhardt, 0.1% Na-Pyrophosphat, 0.1% SDS, 100 μg/ml salmon sperm DNA at 42°C; washes with 2× SSC, 0.1% SDS at room temperature and 42°C). Blot *B* was hybridized as *A* except that salt concentration was decreased to 4× SSC. Washes for blot *B* were at high stringency (2× SSC, 0.1% SDS at room temperature, followed by 0.1× SSC, 0.1% SDS at 65°C). Positions of 28S and 18S RNAs are shown on the left margins.
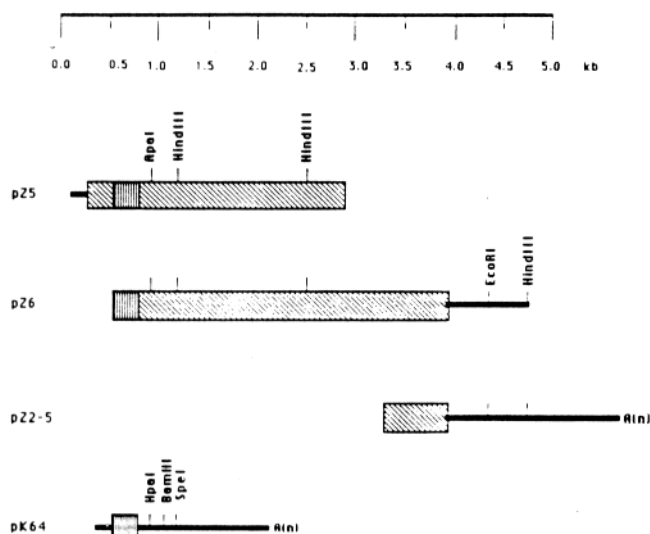
Figure 2. Schematic representation of four cDNA clones, analyzed in detail. The scale on top of the figure is in kilobase. Non-coding regions are represented by solid bars, coding regions by diagonally hatched boxes, and the shared exon in clones pZ5, pZ6, and pK64 by vertically hatched boxes. Recognition sites for restriction enzymes and locations of poly-A tails are as indicated.

These antigenic determinants were assayed with the following antibodies: R1-2, a mAb that detects an α_{4m} determinant on LPAM-1 and LPAM-2 (Holzmann and Weissman, 1989a; Holzmann et al., 1989); PS/2, an independent mAb detecting α_{4m}, generously provided by P. Kincade (Oklahoma Medical Research Foundation, Oklahoma City, OK) (Miyake et al., 1991); and a polyclonal rabbit antiserum raised in this laboratory against R1-2 immunoisolated LPAM-1 molecules from TK-1 cells. FACS analysis was carried out as previously described (Holzmann et al., 1989).
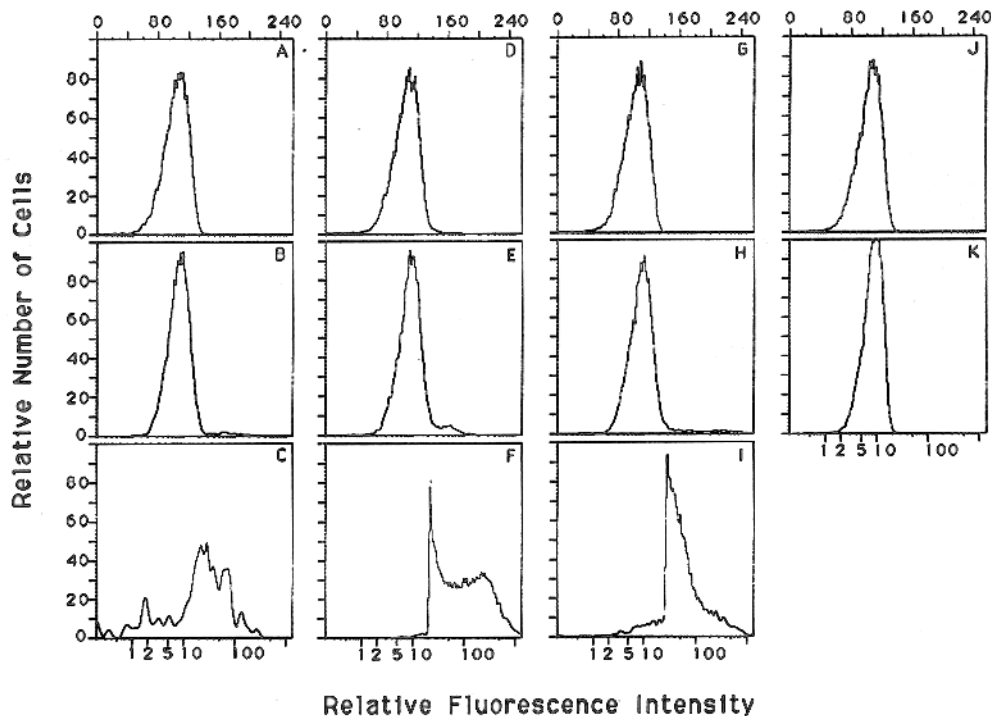
RNA and DNA Extractions, Northern and Southern Blots

Total RNA from frozen cells previously grown in vitro or from tissues was isolated by homogenizing in 7 M Guanidinium Isothiocyanate and centrifuging over a cushion of CsCl (5.7 M) at 35 krpm for 18 h in a Beckman SW41 rotor. For selection of poly-A$^+$ RNA, total RNA was passed twice over an oligo(dT)-cellulose column. Cytoplasmic RNA was isolated by lysing the cells in 0.5% NP-40 and removing the nuclei by centrifugation. For DNA isolation, the cells or carefully homogenized tissues were lysed in 0.5% SDS/1 mg/ml proteinase K at 65°C, extracted with phenol/chloroform, and treated with DNAse-free RNAse. RNA was electrophoresed in MOPS buffer on formaldehyde-agarose gels and transferred on nylon membranes using standard laboratory techniques (Davis et al., 1986). Probes were labeled as described for library screening.

DNA Sequencing

Ends of complete clones or of subcloned restriction fragments in pBluescript were sequenced using a double-stranded DNA sequencing method as described (Neuhaus et al., 1987), or after isolation of single-stranded template DNA by infection of a logarithmically growing culture with a helper phage according to the supplier's instructions (Stratagene). For the sequencing of the complete clones, overlapping 300-bp deletions were made using the ExoIII/Mung Bean Nuclease method (Henikoff, 1984), the resulting fragments cut back with EcoRI, size selected on Agarose gels, and ligated into EcoRI/SmaI cut M13 mp18 vector DNA. Single-stranded DNA was isolated from resulting clones and sequenced using a Sequenase kit according to the manufacturer's instructions (USB, Cleveland, OH). Each strand of the cDNA clones was sequenced at least once. Sequencing data were interpreted with the help of the sequence analysis software package of the University of Wisconsin (Devereux et al., 1984). Homology search to the EMBL/NBRF databases was also done using the PC/GENE sequence analysis software (IntelliGenetics/Genofit).

Results

cDNA Cloning

Unamplified cDNA libraries were screened with the human

Figure 3. FACS analyses of COS-7 cells transiently transfected with α_{4m} cDNA. As controls, the top row (*A, D, G,* and *J*) shows that antibodies used in this study did not stain untransfected COS-7 cells. Specifically, untransfected COS-7 cells were stained with anti-LPAM-1 antiserum (*A* and *J*), with anti-α_{4m} mAb R1-2 (*D*), and another anti-α_{4m} mAb, PS/2 (*G*). The α_{4m} cDNA-transfected COS-7 cells were stained with anti-LPAM-1 antiserum (*B*), R1-2 (*E*), and PS/2 (*H*). In the bottom row (*C, F,* and *I*), the antibody-stained cells were sorted by flow cytometer and the positive cells were reanalyzed, using anti-LPAM-1 antiserum (*C*), R1-2 (*F*), and PS/2 (*I*). As an additional control, the antisense α_{4m} cDNA-transfected COS-7 cells were stained with anti-LPAM-1 antiserum (*K*).

A

[Nucleotide sequence and deduced amino acid sequence data — dense sequence blocks with position numbers in the right margin from 100 through 4866.]

Figure 4. (A) Combined nucleotide sequence of clones pZ5, pZ6, and pZ2-5, and the deduced amino acid sequence (single letter code) of α_{4m}. Cysteines in the protein sequence are marked by asterisks, putative N-glycosylation sites are typed in italics and underlined. The signal peptide and transmembrane domain are also underlined. Ca-binding domains are underlined by an interrupted line. End = translation stop codon. The underlined nucleotide sequence in the 3' non-coding region shows 60–80% homology to "R" repetitive elements (see text for further details). The nucleotide sequence from residue 478 to 714 matches a sequence found in clone pK64. This sequence is bordered by GT at the 5' end, and AG at the 3' end (indicated by ! above the sequence). 42 adenine residues form the poly-A tail in clone pZ2-5. The sequence data reported have been deposited in the EMBL, GenBank and Nucleotide/Protein Sequence Databases and will appear under the accession numbers X53176 (α_{4m}) and X53177 (clone pK64). (B) Nucleotide sequence of clone pK64. The underlined sequence (residues 82–318) is identical to a sequence in the coding region of the α_{4m} cDNAs. This sequence is bound by consensus splicing sites, including GT on the 5' end and AG on the 3' end (typed in bold). The 3' end of the pK64 cDNA sequence is formed by 37 adenine residues.

B

```
GAATCAGGGTCTAGCTTTAACTACTCTATGCCTTTAGAAAAATTTTGCAACAAAACTGATGAGATTTTCCCAAATACTGCAGGTAGCCCCAGTGGAGAGC  100
CTTGTGGGAAGACATGCCTGGAGGAGAGGGATAACCAGTGGCTGGGGGTCACCTTTCCAGACAGCCTGGAGAAAATGGCTCTATCGTGACTTGTGGGCAC  200
AGGTGGAAAAATATTTTTTACATGAAGAGCGATAACAAACTCCCCACTGGCATTTGCTACGTCATGCCTTCTGATTTGCGGACAGAACTGAGTAAAAGGA  300
TGGCCCCGTGTTACAAAGGTAAGTGGGACTTGACCGAACTAGAGGAAGAGAAGCAAAGGAAGGAGTCCACTCGTTTGCATTTGCAGTTTTGCCTGCTCAA  400
AGACGAAGAATATCCTTCATAAGGGCATTGTTAACTAACAGTATCCTTCCTTTGTTTACTTTGTGCAGTATCGATACTTTATACAAACATGCCATGGTTT  500
TGAAAAGCAATTCTTCATTTTCATTTGAATAATTGTGGAAGCTAGTTGGCAACTATCCAACTTTGTACTAAATTTATATTTTCAAATCAGTCTAAGCATG  600
TCTAATGTCAGGGAAACCATAGGATCCTCAAAGTCATGAAATTACATACAAAATTCTTATTAGAAAAAAAATGACCATAATTCTTGTTTTCCTAGAAACT  700
TGTAGGGGTCTTCAAAAGGATTCAAGGCACTAGTCAGAGCAAAGCCTTCCCACTTCCTGTTCCCCAAGATATTCCTTGAAGATGTAAGGAATTACTGTTT  800
ATAGCAACACTTGGTTCTCAGCCTGTGTATCACATCTCCTTTGGGGCTCAGATAACAATTTGCCTGGGGTCACCTAAGACCATCAGAAAGCACAGATATT  900
TACATTACAATGTATAACATTAGCAAAGTTTCACTTGTGAACCCTCAGCAAAACTAATTTTACAATGGGGGGGGGTCGTCACAACATGAGGACCTTTATG  1000
AAAAGGTTATAGCATTAGGAAGGTTGAAACCACTGGTCTAAATTTGGGTAACCTCTGTCTCAATGGAAACCTGTATTTTATGTACAGAATTAAGATATAT  1100
ATTTAGATTTTGGGACAAATGTCAGTAGCTGTCAAGATAATATTTATAACAATCAGTAACTGTCTTTAGAACACCAAAAATATGATGAACAC  1200
CAATTCATTTTAATCATAAAAAATACTATTATTATCTCCCTTATCCTCCAAAGATAACCTTGATTACCTGTTAGAACTTTGAACTCATACTTAAATTTAG  1300
GCACTGGTTCTGATCATTTTAGGATCTTTCATATCTGGATTCTGATGTCCTGGAGAAGTCCATTTAATAACTGTTTATATTTTACCTTAAATGCAAGAAG  1400
ATAAAATATGCATATTAAATTTTCAATTGCAATAAAAA(37)  1480
```

Figure 4.

VLA-4α cDNA clone pGEM4.10 (Takada et al., 1989). Selected plaque-purified clones were identified by hybridization with Northern blots of RNA isolated from murine T cell lines known to express either LPAM-1 (TK1), LPAM-2 (RAW112), or none (TK5), and compared to the hybridization pattern of the VLA-4α cDNA, as shown in Fig. 1. A second round of screening was performed using the insert from an isolated murine cDNA clone. Isolates were tested as described above, and the ends of some clones were sequenced. Based upon cross-hybridization and preliminary sequencing data, four clones were chosen for further characterization: pZ5 and pZ6 from the first round of screening and pZ2-5 from the second round, respectively, as well as pK64, a clone obtained in an independent screening of an oligo-dT primed cDNA library in λgt10 (see Fig. 2 for a schematic representation of the clones). All four clones were completely sequenced.

An α_{4m} cDNA Clone Transfects at Least Three LPAM-1 Determinants to COS-7 Cells

An α_{4m} cDNA clone was inserted into the vector pME18S to create the plasmid α_{4m}/pME18S (see Materials and Methods). Because the plasmid, derived from SRα, contains SV40 elements that allow it to respond to transcription factors in COS-7 cells, transient transfection experiments were performed, assaying for cell-surface expression of LPAM-1 determinants. Two mAbs specific for α_{4m} determinants, R1-2 (Holzmann et al., 1989) and PS/2 (Miyake et al., 1991), and a polyclonal rabbit antiserum raised against immunoisolated $\alpha_{4m}\beta_p$ and which detects both α_{4m} and β_p antigens, were used to screen transfectants using the FACS. In Fig. 3, transfectants stained with the antiserum (Fig. 3 B), R1-2 (Fig. 3 E), and PS/2 (Fig. 3 H) showed small numbers of cells staining well above background (Fig. 3, A, D, and G). The positive subsets from cells shown in Fig. 3 (B, E, and H) were sorted by the FACS, and reanalyzed, respectively, in C, F, and I. The antisense form of α_{4m}/pME18S did not transfect determinants detected by the anti-LPAM-1 antiserum (Fig. 3 K).

Sequence Analysis of the α_{4m} cDNAs

The full complementary DNA sequence (noncoding strand), from the 3' end (poly-A tail in the mRNA) to a region ~5 kb upstream which includes the ATG codon for the initiation of translation, and the deduced amino acid sequence for the murine α_4 protein (α_{4m}) are shown in Fig. 4 A. Comparison to clone pK64 (Fig. 4 B) shows that a fragment of 236 bp (underlined in Fig. 4 B) is identical to the α_{4m} cDNA se-

quence from residue 478 to 714. This fragment is in both cases bordered by the splice donor and acceptor sites, GT and AG, respectively. Introns of higher eukaryotic pre-mRNAs are bordered by the consensus sequence $(C/U)_{11}$ NCAG' at the 3' end and (C/A)AG'GU(G/A)AGU at the 5' end (Aebi et al., 1986, 1987; Green, 1986). The first motif is found immediately preceding the α_m homologous domain in pK64, and the second motif directly following this domain. A deficient or alternative splicing mechanism may therefore be at the origin of the structure observed in clone pK64.

The 3' poly-A tails as cloned consist of 42 adenine residues in the α_{4m} cDNA clone pZ2-5, and of 37 residues in clone pK64. In Northern blot experiments, the size of the most abundant RNA species is 9.5 kb (see below). Therefore, the 5-kb cDNA sequence as shown in Fig. 4 A either does not include a large 5' untranslated region of 4.5 kb, or the true α_{4m} mRNA is shorter and less abundant.

At the 3' end of the α_{4m} cDNA an untranslated sequence of ~1,600 bp is found. This untranslated region includes a repetitive element, underlined in Fig. 4 A, which shows extensive homology to the "R" repetitive element associated with mouse immunoglobulin genes, the mouse alpha interferon pseudogene, and mouse thymocyte extrachromosomal circular DNA (data not shown). Southern blot hybridization of mouse genomic DNA with a probe covering this region shows the high degree of repetition of this sequence in the mouse genome. Probes covering the other regions of all four cDNAs give a hybridization pattern on Southern blots consistent with a single or very low copy gene in the mouse genome. Furthermore, no rearrangement can be observed between DNAs from cell lines or tissues expressing or not expressing the LPAM α chain mRNA (data not shown).

Amino Acid Sequence of α_{4m} and Comparison to Other Integrins

The amino acid sequence as deduced from the open reading frame in Fig. 4 A encodes a protein of a total length of 1,039 amino acids. This includes a presumed signal peptide of 40 amino acids (underlined in Fig. 4 A). The mature protein, starting with the sequence YNLD, is therefore composed of 999 amino acids. The putative transmembrane domain is 24 amino acids long, followed by a short cytoplasmic region of 32 amino acids. The extracellular domain contains 13 potential N-glycosylation sites, 9 of them in the transmembrane-proximal half of the protein. Furthermore, the sequence contains 25 Cysteine residues, all but one of them in the ex-

tracellular portion. Seven homologous repeats (I to VII in Fig. 5) are found in the NH₂-terminal part of the protein. Repeats V-VII include metal binding loops, as is shown by the presence of the consensus sequence Dx(D/N)x(D/N)GxxD (underlined in Fig. 4 A). Since the association of α_{4m} with either β_1 or β_p requires Ca^{2+} (Holzmann et al., 1989), the interaction between the two chains of the receptor molecule probably involves these divalent cation-binding sites.

The amino acid sequence of α_{4m} was compared to other members of the integrin family (Fig. 5). As expected, the highest degree of homology is obtained with the α_{4h} sequence. Of a total of 999 amino acids in the mature protein, 843, corresponding to 84.4%, are identical. There is an almost perfect homology between the two proteins, especially in the transmembrane (20 out of 24 amino acids conserved) and cytoplasmic domains (all except 1 out of 32 conserved). At the nucleic acid sequence level, the homology is 84% for the coding regions of the two proteins, but drops to 52.3% for the part of the 3' noncoding sequence that could be compared (data not shown).

The integrin α subunits have originally been subdivided into two categories, those with I-domains, and those with proteolytically cleaved disulfide-linked COOH-terminal fragments (Takada and Hemler, 1989). It was shown, however, that α_{4h} does not belong to either of these (Takada et al., 1989). All the features in the protein sequence which distinguish it from other integrin α chains are conserved in α_{4m}. It has neither an I-domain, nor a cleaved, disulfide-linked COOH-terminal fragment. Furthermore, the cysteine residues at positions 165, 278, and 462 are conserved, as well as the potential protease cleavage site formed by the sequence KKEK (residues 573-576) immediately followed by an α_4-specific insert of five amino acids.

The overall homology between α_{4m} and the other integrin α chains (with the exception of α_{4h}) is <30%, but some regions are rather well conserved. For the repeated domains IV-VII, the homologies to α_{4m} are 30-60%. The pentapeptide GFFKR in the cytoplasmic domain (residues 970-974 in α_{4m}) is 100% identical in all the α chains. Moreover, the positions of Cysteine residues are well conserved. Of the 25 Cysteines found in the protein sequence, 13 are conserved in all 9 proteins (52%). If the 3 which are conserved in all but one of the α chains are included, this number increases to 16, or 64%.

No sequence homology was found between the lymph node homing receptor (Siegelman et al., 1989) and α_{4m} (data not shown). α_{4m} was also compared to all protein sequences available in the NBRF/EMBL sequence libraries. No significant homology to nonintegrin proteins was found.

Expression of the Different mRNAs in Tissues and Cell Lines

Expression of α_{4m} mRNA was analyzed in lymphoid tissues (total RNA) and cell line TK1 (poly-A⁺ selected cytoplasmic and total RNA), using probes from clones pZ6 and
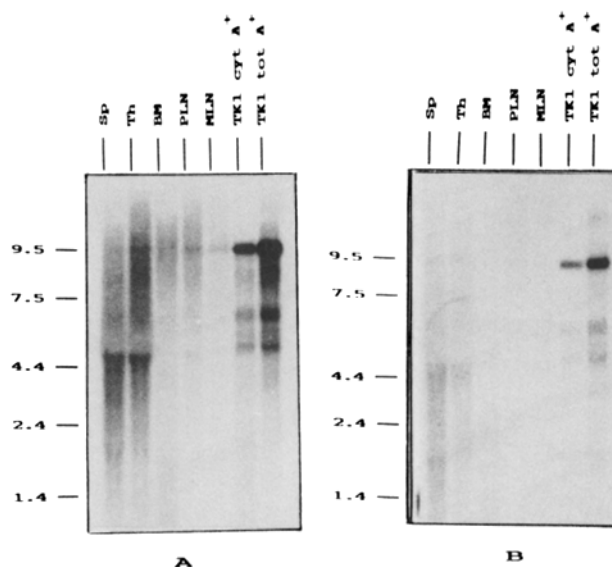

Figure 6. Comparison of Northern blot hybridization of clones pZ6 (A) and pK64 (B). 10 μg of total RNA from spleen (Sp), thymus (Th), bone marrow (BM), peripheral lymph nodes (PLN), and mesenteric lymph nodes (MLN), and 2.5 μg of cytoplasmic (cyt) or total (tot) TK1 poly-A⁺ RNA, were electrophoresed and blotted as in Fig. 1. Positions of RNA molecular weight standards (RNA ladder, BRL) are shown. Both blots were hybridized and washed at high stringency, as described in Fig. 1.

pK64, respectively (Fig. 6, A and B). The most prominent RNA species is always ∼9.5 kb long, but two other discrete bands, corresponding to RNA lengths of ∼5.5 and 6.5 kb, are also observed. This contrasts to the situation found in human cells, where α_{4h} hybridizes only to two RNAs of ∼5 and 6 kb (Takada et al., 1989). Assuming that these are the mRNAs encoding α_4 both in human and mouse tissues, the observed 9.5-kb-long RNA may be pK64 specific and its hybridization to α_{4m} cDNAs due to cross-hybridization of homologous sequences. To test this hypothesis, total tissue RNA and poly-A⁺-selected TK1 RNA was hybridized to nonoverlapping probes covering various regions of clones pZ5, pZ6, pZ2-5, and pK64, respectively (Fig. 7).

A 9.5-kb-long RNA form is the most prominent in each of the hybridizations with the 5' ApaI fragment from pZ5 (Fig. 7 A), an internal fragment of clone pZ6 covering part of the coding region (Fig. 7 B), and the 3' HindIII fragment from the noncoding region in pZ2-5 (Fig. 7 C). Shorter RNA forms are present with unequal distribution. The 5.5-kb-long RNA is not visible in the blot probed with the pZ2-5 fragment. The pZ5 and pZ6 fragments (5' and intermediate, respectively), which correspond to the coding region, give the canonical hybridization pattern (5.5, 6.5, and 9.5 kb), which was also observed with the human clone pGEM4.10 (see Fig. 1). The two bands visible at ∼5 kb in the lanes loaded with tissue RNAs are probably due to cross-

---

Figure 5. Comparison of amino acid sequences of integrin α chains. α_{4m} (LPAM) is compared to the α chains of the following proteins: VLA-4 (Takada et al., 1989), VLA-2 (Takada and Hemler, 1989), fibronectin receptor FNR (Argraves et al., 1987), Glycoprotein IIb (Poucz et al., 1987), vitronectin receptor VNR (Suzuki et al., 1987), murine (m), and human (h) MAC1 (Pytela, 1988; Corbi et al., 1988), and protein p150 (Corbi et al., 1987). All sequences are of human origin, except for LPAM and mMAC1. Amino acid homologies with LPAM are typed in bold and underlined. The seven homologous repeats (I-VII) are indicated.
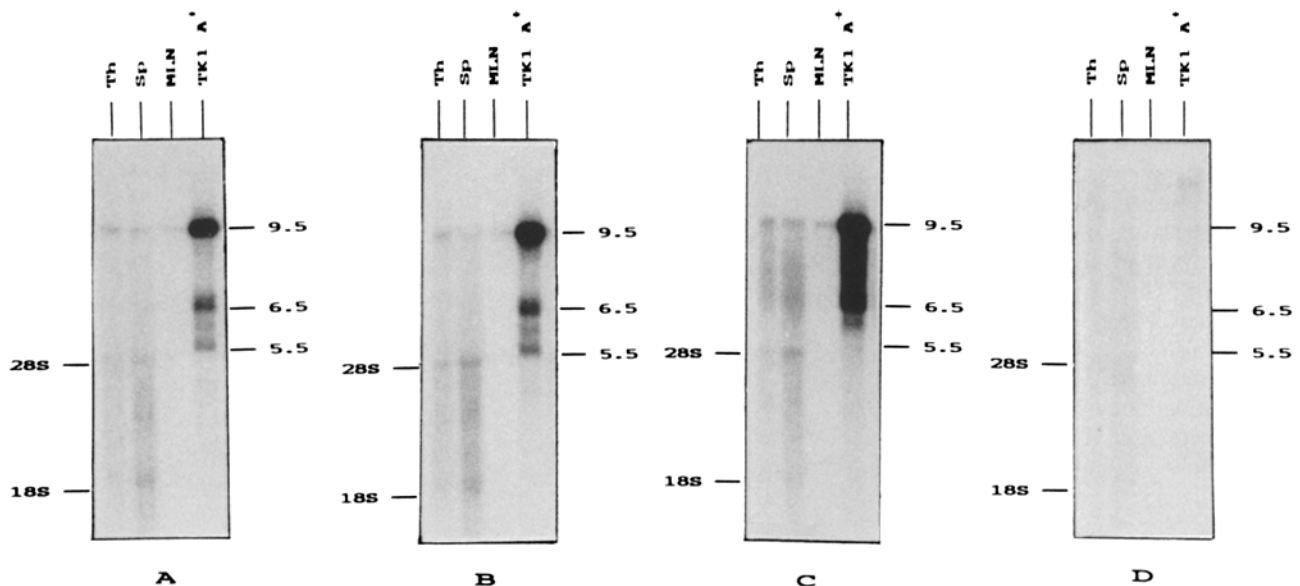
Figure 7. Comparison of mRNA expression by hybridization with various cDNA probes. 10 μg total RNA from thymus (*Th*), spleen (*Sp*), and mesenteric lymph nodes (*MLN*), and 2.5 μg poly-A$^+$ selected total TK1 RNA were electrophoresed and transferred as described. Northern blots were hybridized with the 5' ApaI fragment from clone pZ5 (*A*), an internal restriction fragment from the coding region of clone pZ6 (*B*), the 3' HindIII fragment from clone pZ2-5 (*C*), and the 3' BamHI fragment from clone pK64 (*D*). All hybridizations and washes were done at high stringency. Positions of 28S and 18S RNAs are given on the left margins, and calculated RNA molecular masses of 9.5, 6.5, and 5.5 kb on the right margins.

hybridization of a contaminant to 28S rRNA, since the position of these bands corresponds to 28S rRNA on the ethidium bromide–stained gel before transfer (data not shown).

Hybridization of the Northern blot with a fragment from the 3' end of the clone pK64 shows a strikingly different pattern (Fig. 7 D). Indeed, the most prominent RNA is longer than 9.5 kb. The 9.5-kb RNA as well as some other RNAs between 7 and 9 kb give a very faint signal only, whereas the two shortest forms observed with all other probes are missing. This RNA, whose size is estimated to be ∼12 kb, is absent in hybridizations with probes that do not contain sequences cross-hybridizing with pK64 (see Figs. 7, *B* and *C* and 1, *A* and *B*). It seems to be an RNA that shares at least one exon with the α_{4m} encoding mRNA (outlined in Fig. 4, *A* and *B*), but with a specific 3' end, as is shown by sequencing and hybridization data. The exact role of this RNA in the cell is as of yet undetermined.

Discussion

The murine T cell line TK1 expresses LPAM-1, a Peyer's Patch-homing–associated cell surface molecule, at high levels, as defined by binding of the blocking antibody R1-2. This mAb detects an epitope on the α chain of both LPAM-1 ($\alpha_{4m}\beta_p$) and LPAM-2 ($\alpha_{4m}\beta_1$). α_{4m} is antigenically related to the human VLA-4 α chain (α_{4h}), since the respective antibodies and/or antisera are crossreactive (Holzmann et al., 1989). A cDNA clone encoding α_{4h} was used to isolate cDNA clones from libraries made by reverse transcribing poly-A$^+$ selected RNA from the murine T-cell lymphoma TK1. An α_{4m} cDNA clone inserted into the expression vector pME/18S transfects the coding information for several α_{4m} antigenic determinants to COS-7 cells. This confirms that the α_{4h} homologue in mice, α_{4m}, encodes all known LPAM α molecules.

Comparison of the newly isolated murine sequences with the human sequence show that the two are 84% homologous in the coding region. The same degree of homology is observed at the amino acid sequence level. In the murine system, LPAM molecules expressed at the cell surface of lymphocytes mediate the organ-specific binding to endothelial cells in Peyer's patches or other mucosal sites. Based on the high degree of homology both on DNA and protein levels between the α chains of LPAM-1/2 and LVA-4, it is reasonable to assume that at least one of the murine molecules (probably LPAM-1) would have a human homologue that serves the same function. Supporting this, anti-α_{4h} antibodies were able to block Peyer's patch HEV adhesion by human lymphocytes (Holzmann and Weissman, 1989b). It is also reasonable to assume that one of the mouse heterodimers (probably LPAM-2) will have several different cell–cell and cell–matrix adhesive functions, analogous to those seen for VLA-4 in the human. In fact, R1-2 and PS/2 antibody blocking studies reveal that α_{4m}-bearing early hematolymphoid cells utilize this molecule to mediate binding to the hematopoiesis-regulating bone marrow stromal cells with which they interact (Miyake et al., 1991).

The sequence of the LPAM α chain was compared to other members of the integrin family for which protein sequences were available. Interestingly, four of the seven homologous repeats in the NH$_2$-terminal half of the protein are rather well conserved, suggesting a common function for this domain in all the molecules. A possible hypothesis is that the interaction between the α and β chains of the integrin heterodimers occurs in this region. This assumption is supported by the fact that these domains include divalent cation-binding sites, and the binding of the β by the α chain is known to be dependent on divalent cations (Holzmann et al., 1989). Accordingly, the other less homologous regions of the proteins would be responsible for the functional specificity. An α_{4m}

exon-homologous domain in an independent clone, pK64, is bordered on both ends by intron splice donor and acceptor sequences. Such a structure may have arisen by an alternative splicing mechanism. This hypothesis is currently being investigated by using genomic clones from a mouse DNA library.

The described cDNA clones contain the sequence encoding the α chain for the two different LPAM heterodimers. The interaction of the rat mAb R1-2 with an epitope on this protein completely inhibited binding of lymphocytes to Peyer's patch HEV; the R1-2 antibody also blocks cell–cell adhesion between hematopoietic and bone marrow stromal cells. The cloned cDNAs, along with the recently cloned murine β_p gene (M. Hu, unpublished data), now make it possible to study sequence/function relationships for both of the LPAM heterodimers, $\alpha_{4m}\beta_p$ and $\alpha_{4m}\beta_1$. It should be possible to investigate if expression of this α chain alone, or in combination with one of the two described β chains, is sufficient to transform lymphocytes from a Peyer's patch non-binding state into binding cells, thus inducing a true "homing" activity to mucosal sites. Also, it should now be possible to determine the precise biochemical relationship between Peyer's patch homing and the several other functions of VLA-4/LPAM-2, including adhesion to activated endothelium (Elices et al., 1990), binding to fibronectin (Wayner et al., 1989; Guan and Hynes, 1990), triggering of homotypic aggregation (Bednarczyk and McIntyre, 1990; Campanelo et al., 1990), and participation in T cell–mediated cytolysis of target cells (Clayberger et al., 1987; Takada et al., 1989).

References

Aebi, M., H. Hornig, R. A. Padgett, J. Reiser, and C. Weissmann. 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell.* 47:555–565.

Aebi, M., H. Hornig, and C. Weissmann. 1987. 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. *Cell.* 50:237–246.

Argraves, W. S., S. Suzuki, H. Arai, K. Thompson, M. D. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the human fibronectin receptor. *J. Cell Biol.* 105:1183–1190.

Bednarczyk, J. L., and B. W. McIntyre. 1990. Induction of lymphocyte aggregation by mAb binding to a member of the integrin superfamily. *J. Immunol.* 144:777–784.

Butcher, E. C., and I. L. Weissman. 1980. Cellular, genetic, and evolutionary aspects of lymphocyte interactions with high endothelial venules. *Ciba Foundation Symposium.* 71:265–286.

Campanero, M. R., R. Pulido, M. A. Ursa, M. Rodriquez-Moya, M. O. De Landazuri, and F. Sanchez-Madrid. 1990. An alternative leukocyte homotypic adhesion mechanism, LFA-1/ICAM-1-independent, triggered through the human VLA-4 integrin. *J. Cell Biol.* 110:2157–2165.

Chu, G., H. Hayakawa, and P. Berg. 1987. Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Res.* 15:1311.

Clayberger, C., A. M. Krensky, B. W. McIntyre, T. D. Koller, P. Parham, F. Brodsky, D. J. Linn, and E. L. Evans. 1987. Identification and character-

ization of two novel lymphocyte function-associated antigens, L24 and L25. *J. Immunol.* 138:1510–1514.

Corbi, A. L., L. J. Miller, K. O'Connor, R. S. Larson, and T. A. Springer. 1987. cDNA cloning and complete primary structure of the alpha subunit of a leukocyte adhesion glycoprotein, p150, 95. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:4023–4028.

Corbi, A. L., T. K. Kishimoto, L. J. Miller, and T. A. Springer. 1988. The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) α subunit: cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B. *J. Biol. Chem.* 263:12403–12411.

Davis, L. G., M. D. Dibner, and J. F. Battey. 1986. Basic Methods in Molecular Biology. Elsevier, New York. 388 pp.

Devereux, J., P. Haeberli, and O. Smithers. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387–395.

Elices, M. J., L. Osborn, Y. Takada, C. Crouse, S. Luhowskyj, M. E. Hemler, and R. R. Lobb. 1990. VCAM-1 on activated endothelium interacts with the leukocyte integrin VLA-4 at a site distinct from the VLA-4/fibronectin binding site. *Cell.* 60:577–584.

Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13.

Gallatin, W. M., I. L. Weissman, and E. C. Butcher. 1983. A cell-surface molecule involved in organ-specific homing of lymphocytes. *Nature (Lond.).* 304:30–34.

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.* 56:1063–1072.

Green, M. R. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.* 20:671–708.

Guan, J. L., and R. O. Hynes. 1990. Lymphoid cells recognize an alternatively spliced segment of fibronectin via the integrin receptor $\alpha_4\beta_1$. *Cell.* 60:53–61.

Gubler, U., and B. Hoffman. 1983. A simple and very efficient method for generating cDNA libraries. *Gene.* 25:263–268.

Hamann, A., D. Jablonski-Westrich, A. Duijvenstijn, E. C. Butcher, H. Baisch, R. Harder, and H. G. Thiele. 1988. Evidence for an accessory role of LFA-1 in lymphocyte-high endothelium interaction during homing. *J. Immunol.* 140:693–699.

Hemler, M. E. 1988. Adhesive protein receptors on hematopoietic cells. *Immunol. Today.* 9:109–113.

Hemler, M. E., C. Huang, Y. Takada, L. Schwarz, J. L. Strominger, and M. L. Clabby. 1987. Characterization of the cell surface heterodimer VLA-4 and related peptides. *J. Biol. Chem.* 262:11478–11485.

Hemler, M. E., M. J. Elices, C. Parker, and Y. Takada. 1990. Structure of the integrin VLA-4 and its cell-cell and cell-matrix adhesion functions. *Immunol. Rev.* 114:45–65.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene.* 28:351–359.

Holzmann, B., and I. L. Weissman. 1989a. Peyer's patch-specific lymphocyte homing receptors consist of a VLA-4 like α chain associated with either of two integrin β chains, one of which is novel. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1735–1741.

Holzmann, B., and I. L. Weissman. 1989b. Integrin molecules involved in lymphocyte homing to Peyer's patches. *Immunol. Rev.* 108:45–61.

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.* 56:37–46.

Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48:549–554.

Jalkanen, S., N. Wu, R. F. Bargatze, and E. C. Butcher. 1987. Human lymphocyte and lymphoma homing receptors. *Annu. Rev. Med.* 38:467–476.

Kraal, G., I. L. Weissman, and E. C. Butcher. 1983. Differences in in vivo distribution and homing of T cell subsets to mucosal vs non-mucosal lymphoid organs. *J. Immunol.* 130:1097–1102.

Lasky, L. A., M. S. Singer, T. A. Yednock, D. Dowbenko, C. Fennie, H. Rodriguez, T. Nguyen, S. Stachel, and S. D. Rosen. 1989. Cloning of a lymphocyte homing receptor reveals a lectin domain. *Cell.* 56:1045–1055.

Miyake, K., I. L. Weissman, J. S. Greenberger, and P. W. Kincade. 1991. Evidence for a role of two adhesion molecules in lympho-hemopoiesis: the integrin VLA-4 on hematolymphoid cells and a novel stromal glycoprotein. *J. Exp. Med.* 173:599–607.

Neuhaus, H., A. Etter, F. Müller, and H. Tobler. 1987. Type-1 like intervening sequences are found in the rDNA of the nematode Ascaris lumbricoides. *Nucleic Acids Res.* 15:7689–7707.

Poncz, M., R. Eisman, R. Heidenreich, S. M. Silver, G. Vilaire, S. Surrey, E. Schwartz, and J. S. Benett. 1987. Structure of the platelet membrane glycoprotein IIb: homology to the alpha subunits of the vitronectin and fibronectin membrane receptors. *J. Biol. Chem.* 262:8476–8482.

Pytela, R. 1988. Amino acid sequence of the murine Mac-1 α chain reveals homology with the integrin family and an additional domain related to von Willebrand factor. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1371–1378.

Siegelman, M. H., and I. L. Weissman. 1989. Human homologue of the mouse lymph node homing receptor: evolutionary conservation of at tandem cell interaction domains. *Proc. Natl. Acad. Sci. USA.* 86:5562–5566.

Siegelman, M. H., M. Van de Rijn, and I. L. Weissman. 1989. Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing tandem

interaction domains. *Science (Wash. DC)*. 243:1165–1172.

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.* 56:1057–1062.

Stamper, H. B., and J. J. Woodruff. 1976. Lymphocyte homing into lymph nodes: in vitro demonstration of the selective affinity of recirculating lymphocytes for high-endothelial venules. *J. Exp. Med.* 144:828–833.

Suzuki, S., W. S. Argraves, H. Arai, L. R. Languino, M. D. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the vitronectin receptor alpha subunit and comparative expression of adhesion receptor mRNAs. *J. Biol. Chem.* 262:14080–14085.

Takada, Y., and M. E. Hemler. 1989. The primary structure of the VLA-2/collagen receptor α_2 subunit (platelet GP Ia): homology to other integrins and the presence of a possible collagen-binding domain. *J. Cell Biol.* 109:397–407.

Takada, Y., M. J. Elices, C. Crouse, and M. E. Hemler. 1989. The primary structure of the α_4 subunit of VLA-4: homology to other integrins and a possible cell-cell adhesion function. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:1361–1369.

Takobe, Y., M. Seiki, J. Fujisawa, P. Hoy, K. Yokota, K. Arai, M. Yoshida, and N. Arai. 1988. SRα promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol. Cell. Biol.* 8:466–472.

Wayner, E. A., A. Garcia-Pardo, M. J. Humphries, J. A. McDonald, and W. G. Carter. 1989. Identification and characterization of the lymphocyte adhesion receptor for an alternative cell attachment domain in plasma fibronectin. *J. Cell Biol.* 109:1321–1330.