The EMBO Journal vol.6 no.13 pp.4023–4028, 1987

cDNA cloning and complete primary structure of the α subunit of a leukocyte adhesion glycoprotein, p150,95

Angel L.Corbi, Linda J.Miller, Karen O'Connor, Richard S.Larson and Timothy A.Springer

Laboratory of Membrane Immunochemistry, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

Communicated by A.F.Williams

The leukocyte adhesion receptors, p150,95, Mac-1 and LFA-1 are integral membrane glycoproteins which contain distinct α subunits of 180 000–150 000 M_r associated with identical β subunits of 95 000 M_r in αβ complexes. p150,95 α subunit tryptic peptides were used to specify oligonucleotide probes and a cDNA clone of 4.7 kb containing the entire coding sequence was isolated from a size-selected myeloid cell cDNA library. The 4.7-kb cDNA clone encodes a signal sequence, an extracellular domain of 1081 amino acids containing 10 potential glycosylation sites, a transmembrane domain of 26 amino acids, and a C-terminal cytoplasmic tail of 29 residues. The extracellular domain contains three tandem homologous repeats of ~60 amino acids with putative divalent cation-binding sites, and four weaker repeats which lack such binding sites. The cDNA clone hybridizes with a mRNA of 4.7 kb which is induced during *in vitro* differentiation of myeloid cell lines. The p150,95 α subunit is homologous to the α subunits of receptors which recognize the RGD sequence in extracellular matrix components, as has previously been shown for the β subunits, supporting the concept that receptors involved in both cell–cell and cell–matrix interactions belong to a single gene superfamily termed the integrins. Distinctive features of the p150,95 α subunit include an insertion of 126 residues N-terminal to the putative metal binding region and a deletion of the region in which the matrix receptors are proteolytically cleaved during processing.

Key words: p150,95/leukocyte adhesion/RGD-receptors

Introduction

The leukocyte adhesion receptors are a family of three cell surface glycoproteins which mediate cell–cell interactions in inflammation. These glycoproteins, LFA-1, Mac-1 and p150,95, each contain an α subunit (CD11a,b and c respectively) non-covalently associated with a β subunit (CD18). The α subunits of 180 000, 170 000 and 150 000 M_r have homologous N-terminal sequences, while the β subunits of 95 000 M_r are identical (Sanchez-Madrid *et al.*, 1983a, b: Springer *et al.*, 1986; Miller *et al.*, 1987). The importance of these glycoproteins was first defined by the ability of monoclonal antibodies either to the specific α subunits or to the common β subunit to inhibit adhesion-dependent leukocyte functions (Sanchez-Madrid *et al.*, 1982; Beller *et al.*, 1982). Subsequently, a genetic disease was defined, leukocyte adhesion deficiency (LAD) in which LFA-1, Mac-1 and p150,95 are deficient and a wide range of lymphocyte, neutrophil and monocyte adhesion-dependent functions are defective (Anderson and Springer, 1987). The primary defect in LAD is in the β subunit; association with the β subunit is required for correct processing and expression of the α subunits. LAD patients have recurrent bacterial infections and often die in the first years of childhood. The most important defect clinically is the inability of circulating neutrophils and monocytes to bind to endothelial cells at sites of inflammation and migrate into the inflamed tissue.

p150,95 is expressed on myeloid cells and on activated but not resting lymphocytes (Caligaris-Cappio *et al.*, 1985; Miller *et al.*, 1986; Keizer *et al.*, 1987). In blood neutrophils and monocytes, p150,95 is present on the cell surface and in intracellular vesicles which fuse with the plasma membrane after stimulation with inflammatory mediators (reviewed in Anderson and Springer, 1987), correlating with increased adhesion of these cells to endothelial cells. Inhibition by α subunit-specific monoclonal antibodies has shown that of the three leukocyte adhesion proteins, p150,95 is the most important in monocyte adhesion and chemotaxis (Te Velde *et al.*, 1987), whereas Mac-1 is most important in neutrophils (Anderson *et al.*, 1986). Monocyte extravasation and differentiation into tissue macrophages is accompanied by increased p150,95 expression (Schwarting *et al.*, 1985), as is differentiation of myelomonocytic precursor cells *in vitro* (Miller *et al.*, 1986). On lymphoid cells p150,95 appears to be a marker of cell activation and is particularly strongly expressed on hairy leukemia cells, transformed B lymphocytes which home to the spleen. Although most cytolytic T lymphocyte clones express higher amounts of LFA-1 than p150,95, a subset expresses equal quantities and functional studies show that both p150,95 and LFA-1 mediate conjugate formation with target cells (Keizer *et al.*, 1987).

Recently, the amino acid sequence of the leukocyte adhesion glycoprotein common β subunit has been determined (Kishimoto *et al.*, 1987; Law *et al.*, 1987). In the present paper we characterize the molecular biology and complete amino acid sequence of the p150,95 α subunit. Furthermore, we report extensive homology between the α subunit of the extracellular matrix receptors and the p150,95 leukocyte adhesion receptor.

Results and discussion

The p150,95 αβ complex was purified from hairy cell leukemia spleens by monoclonal antibody affinity chromatography (Miller *et al.*, 1987). The α subunit was isolated by preparative SDS–PAGE, digested with trypsin, and the peptides were separated by reverse-phase HPLC and subjected to amino acid sequencing. A total of 214 residues were determined from 13 peptides (Table I). Oligonucleotides based on two tryptic peptides (underlined, Table I) and the N-terminal sequence (Miller *et al.*, 1987) were used to screen a size-selected cDNA library constructed from phorbol myristyl acetate (PMA)-stimulated HL-60 cell mRNA. A full-length cDNA clone of 4.7 kb was isolated and sequenced by the Sanger dideoxynucleotide method (Figure 1). The nucleotide sequence of 4704 bp contains a 5′ untranslated region of 64 bp followed by a putative initiation codon and the entire coding sequence of 3492 nucleotides. The 1140-bp 3′ untranslated region includes an Alu sequence between nucleotides 3893 and 4054 and ends with a consensus

polyadenylation signal followed by a stretch of 50 adenosines (Figure 1B).

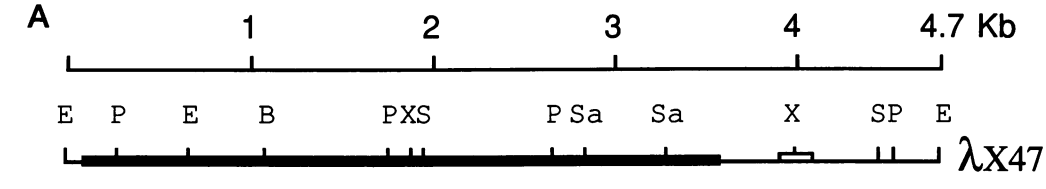
The p150,95 α subunit mRNA was examined in U937 and HL-60 myelomonocytic cell lines (Figure 2A), which can be induced with PMA to differentiate along the macrophage pathway. Concomitantly, synthesis of the p150,95 α subunit precursor and surface expression of the p150,95 αβ complex are induced (Miller *et al.*, 1986). PMA treatment for 3 days of both cell lines induced the appearance of the p150,95 α subunit mRNA of ~4.7 kb, in agreement with the size of the cDNA isolated. This demonstrates that the level of p150,95 α subunit mRNA regulates p150,95 protein expression. The β subunit appears to be synthesized in excess over the α subunit both before and after PMA stimulation (Miller *et al.*, 1986). Southern blots on *Eco*RI- and *Hin*dIII-digested DNA showed single strongly hybridizing bands, suggesting that the p150,95 α subunit is encoded by a single gene (Figure 2B). This was confirmed by isolation of a genomic clone, which contains both the strongly and weakly hybridizing fragments apparent in Figure 2B. (Data not shown.)

The translated amino acid sequence was confirmed by the determination of 237 N-terminal and tryptic peptide amino acid residues which represent 21% of the sequence. The amino acid sequence shows that the p150,95 α subunit is a transmembrane protein with a large extracellular domain (Figure 3, and Figure 5 below). The previously reported p150,95 α subunit N-terminal sequence (Miller *et al.*, 1987) begins at nucleotide 122. Only one potential translation initiation codon is found upstream at position 65, predicting the signal peptide MTRTRAALLLFTALATSLG. The mature protein contains 1144 amino acids and the predicted M_r is 125 908, in agreement with the previous estima-

Table I. Tryptic peptide sequences[a]

147–169	N F A T M M N F V(R) A V I S Q F Q R P S T Q F
190–210	X S N P L S L L A S V H Q L(Q) G F T Y(S) A
265–275	X A I G V G(L) A(F) Q N
297–310	V E D F X A L X(D) I Q N X(L)
383–394	X S Y L G Y S T E L A L
502–521	F G A A L T V L G D V N G D K L T(D) V V
532–551	G A V Y L F H G V L G P S I S P S H S Q
562–588	L Q Y F G Q A L S G G Q D L T Q D G L V D L A V G A(R)
596–615	T R P V L W V G V S M Q F I P A E I P(R)
741–750	Y F T A S L P F E K
893–902	T T F Q L E L P V K
945–969	D L P V S I N F(W) V P V E L N Q E A V X M X V E(V)
1081–1090	Y K V H N(P S P)X(I)

[a]The position in the translated cDNA sequence is indicated. Parentheses indicate uncertainty in the amino acid assignment. Positions where no residue could be assigned are denoted by X. Regions of peptides used for the design of oligonucleotide probes are underlined.



B

```
GAATTCCTGCCACTCTTCCTGCAACGGCCCAGGAGCTCAGAGCTCCACATCTGACCTTCTAGTCATGACCAGGACCAGGGCAGCACTCCTCCTGTTCACAGCCTTAGCAACTTCTCTAGGTTTCAACTTGGACACAGAGGAGCTGACAGC   150
CTTCCGTGTGGACAGCGCTGGGTTTGGAGACAGCGTGGTCCAGTATGCCAACTCCTGGGTGGTGGTTGGAGCCCCCCAAAAGATAACAGCTGCCAACCAAACGGGTGGCCTCTACCAGTGTGGCTACAGCACTGGTGCCTGTGAGCCCAT   300
CGGCCTGCAGGTGCCCCCGGAGGCCGTGAACATGTCCCTGGGCCTGTCCCTGGCGTCTACCACCAGCCCTTCCCAGCTGCTGGCCTGCGGCCCCACCGTGCACCACGAGTGCGGGAGGAACATGTACCTCACCGGACTCTGCTTCCTCCT   450
GGGCCCCCACCCAGCTCACCCAGAGGCTCCCGGTGTCCAGGCAGGAGTGCCCAAGACAGGAGCAGGACATTGTGTTCCTGATCGATGGCTCAGGCAGCATCTCCTCCCGCAACTTTGCCACGATGATGAACTTCGTGAGAGCTGTGATAAG   600
CCAGTTCCAGAGACCCAGCACCCAGTTTTCCCTGATGCAGTTCTCCAACAAATTCCAAACACACTTCACTTTCGAGGAATTCAGGCGCACGTCAAACCCCCTCAGCCTGTTGGCTTCTGTTCACCAGCTGCAAGGGTTTACATACACGGC   750
CACCGCCATCCAAAATGTCGTGCACCGATTGTTCCATGCCTCATATGGGGCCCGTAGGGATGCCCACCAAAATTCTCATTGTCATCACTGATGGGAAGAAAGAAGGCGACAGCCTGGATTATAAGGATGTCATCCCCATGGCTGATGCAGC   900
AGGCATCATCCGCTATGCAATTGGGGTTGGATTAGCTTTTCAAAACAGAAATTCTTGGAAAGAATTAAATGACATTGCATCGAAGCCCTCCCAGGAACACATATTTAAAGTGGAGGACTTTGATGCTCTGAAAGATATTCAAAACCAACT  1050
GAAGGAGAAGATCTTTGCCATTGAGGGTACGGAGACCACAAGCAGTAGCTCCTTCGAATTGGAGATGGCACAGGAGGGCTTCAGCGCTGTGTTCACACCTGATGGCCCCGTTCTGGGGGCTGTGGGGAGCTTCACCTGGTCTGGAGGTGC  1200
CTTCCTGTACCCCCCAAATATGAGCCCTACCTTCATCAACATGTCTCAGGAGAATGTGGACATGAGGGACTCTTACCTGGGTTACTCCACCGAGCTGGCCCTCTGGAAAGGGGTGCAGAGCCTGGTCCTGGGGGCCCCCCGCTACCAGCA  1350
CACCGGGAAGGCTGTCATCTTCACCCAGGTGTCCAGGCAATGGAGGATGAAGGCCGAAGTCACGGGGACTCAGATCGGCTCCTACTTCGGGGCCTCCCTCTGCTCCGTGGACGTAGACACCGACGGCAGCACCGACCTGGTCCTCATCGG  1500
GGCCCCCCATTACTACGAGCAGACCCGAGGGGGCCAGGTGTCTGTGTGTCCCTTGCCCAGGGGGTGGAGAAGGTGGTGGTGTGATGCTGTTCTCTACGGGGAGCAGGGCCACCCCTGGGGTCGCTTTGGGGCGGCTCTGACAGTGCTGG  1650
GGATGTGAATGGGGACAAGCTGACAGACGTGGTCATCGGGGCCCCAGGAGAGGAGGAGAACCCGGGGTGCTGTCTACCTGTTTCACGGAGTCTTGGGACCCAGCATCAGCCCCTCCCACAGCCAGCGGATCGCGGGCTCCCAGCTCTCCTC  1800
CAGGCTGCAGTATTTTGGGCAGGCACTGAGCGGGGGTCAAGACCTCACCCAGGATGGACTGGTGGACCTGGCTGTGGGGGCCCGGGGCCAGGTGCTCCTGCTCAGGACCAGACCTGTGCTCTGGGTGGGGGTGAGCATGCAGTTCATACC  1950
TGCCGAGATCCCCAGGTCTGCGTTTGAGTGTCGGGAGCAGGTGGTCTCTGAGCAGACCCTGGTACAGTCCAACATCTGCCTTTACATTGACAAACGTTCTAAGAACCTGCTTGGGAGCCGTGACCTCCAAAGCTCTGTGACCTTGGACCT  2100
GGCCCTCGACCCTGGCCGCCTGAGTCCCCGTGCCACCTTCCAGGAAACAAAGAACCGGAGTCTGAGCCGAGTCCGAGTCCTCGGGCTGAAGGCACACTGTGAAAACTTCAACCTGCTGCTCCCGAGCTGCGTGGAGGACTCTGTGACCCC  2250
CATTACCTTGCGTCTGAACTTCACGCTGGTGGGCAAGCCCCTCCTTGCCTTCAGAAACCTGCGGCCTATGCTGGCCGCACTGGCTCAGAGATACTTCACGGCCTCCCTACCCTTTGAGAAGAACTGTGGAGCCGACCATATCTGCCAGGA  2400
CAATCTCGGCATCTCCTTCAGCTTCCCAGGCTTGAAGTCCCTGCTGGTGGGGAGTAACCTGGAGCTGAACGCAGAAGTGATGGTGTGGAATGACGGGGAAGACTCCTACGGAACCACCATCACCTTCTCCCACCCCGCAGGACTGTCCTA  2550
CCGCTACGTGGCAGAGGGCCAGAAACAAGGGCAGCTGCGTTCCCTGCACCTGACATGTGACAGCGCCCCCAGTTGGGAGCCAGGGCACCTGGAGCACCAGCTGCAGAATCAACCACCTCATCTTCCGTGGCGGCGCCCAGATCACCTTCTT  2700
GGCTACCTTTGACGTCTCCCCCAAGGCTGTCCTGGGAGACCGGCTGCTTCTGACAGCCAATGTGAGCAGTGAGAACAACACTCCCAGGACCAGCAAGACCACCTTCCAGCTGGAGCTCCCGGTGAAGTATGCTGTCTACACTGTGGTTAG  2850
CAGCCACGAACAATTCACCAAATACCTCAACTTCTCAGAGTCTGAGGAGAAGGAAAGCCATGTGGCCATGCACAGATACCAGGTCAATAACCTGGGACAGAGGGACCTGCCTGTCAGCATCAACTTCTGGGTGCCTGTGGAGCTGAACCA  3000
GGAGGCTGTGTGGATGGATGTGGAGGTCTCCCACCCCCAGAACCCATCCCTTCGGTGCTCCTCAGAGAAAATCGCACCCCCAGCATCTGACTTCCTGGCGCACATTCAGAAGAATCCCGTGCTGGACTGCTCCATTGCTGGCTGCCTGCG  3150
GTTCCGCTGTGACGTCCCCTCCTTCAGCGTCCAGGAGGAGCTGGATTTCACCCTGAAGGGCAACCTCAGCTTTGGCTGGGTCCGCCAGATATTGCAGAAGAAGGTGTCGGTCGTGAGTGTGGCTGAAATTACGTTCGACACATCCGTGTA  3300
CTCCCAGCTTCCAGGACAGGAGGCATTTATGAGAGCTCAGACGACAACGGTGCTGGAGAAGTACAAGGTCCACAACCCCACCCCCCCTCATCGTAGGCAGCTCCATTGGGGGTCTGTTGCTGCTGGCACTCATCACAGCGGTACTGTACAA  3450
AGTTGGCTTCTTCAAGCGTCAGTACAAGGAAATGATGGAGGAGGCAAATGGACAAATTGCCCCAGAAAACGGGACACAGACCCCCGCCCAGTGAGAAATGATCCCTCTTTGCCTTGGACTTCTTCTCCCGCGATTTTCCCCACTT  3600
ACTTACCCTCACCTGTCAGGCTGACGGGGAGGAACCACTGCACCACCGAGAGAGGCTGGGATGGGCCTGCTTCCTGTCTTTGGGAGAAAACGTCTTGCTTGGGAAGGGGCCTTTGTCTTGTCAAGGTTCCAACTGGAAACCCTTAGGACA  3750
GGGTCCCTGCTGTGTTCCCCAAAAGGACTTGACTTGCAATTTCTACCTAGAAATACATGGACAATACCCCCAGGCCTCAGTCTCCCTTCTCCCATGAGGCACGAATGATCTTTCTTTCCTTTCCTTTTTTTTTTTTTTCTTTTCTTTTTT  3900
TTTTTTTTTGAGACGGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCAATGGCGTGATCTCGGCTCGCTGCAACCTCCGCCTCCCGGGTTCAAGTAATTCTGCTGTCTCAGCCTCCTGCGTAGCTGGGACTACAGGCACACGCCACCTGGC  4050
CCGGCCCGATCTTTCTAAAATACAGTTCTGAATATGCTGCTCATCCCCACCTGTCTTCAACAGCTCCCCATTACCCTCAGGACAATGTCTGAACTCTCCAGCTTCGCGTGAGAAGTCCCCTTCCATCCCAGAGGGTGGGCTTCAGGGCGC  4200
ACAGCATGAGAGCCTCTGTGCCCCCATCACCCTCGTTTCCAGTGAATTAGTGTCATGTCAGCATCAGCTCAGGGCTTCATCGTGGGGCTCTCAGTTCCGATTCCCCAGGCTGAATTGGGAGTGAGATGCCTGCATGCTGGGTTCTGCACA  4350
GCTGGCCTCCCGCGGTTGGGTCAACATTGCTGGCCTGGAAGGGAGGAGCGCCCTCTAGGGAGGGACATGGCCCCGGTGCGGCTGCAGCTCACCAGCCCCAGGGGCAGAAGAGACCCAACCACTTCCTATTTTTTGGAGGCTATGAATATAG  4500
TACCTGAAAAAATGCCAAGCACTAGATTATTTTTTTAAAAAGCGTACTTTAAATGTTTGTGTTAATACACATTAAAACATCGCACAAAAACGATGCATCTACCGCTCCTTGGGAAATAATCTGAAAGGTCTAAAATAATAAAAAGCCTTCT  4650
GTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                                                                                                  4704
```

Fig. 1. (A) Restriction map of the p150,95 α subunit cDNA clone λX47. Restriction sites are *Eco*RI (E), *Bgl*II (B), *Pst*I (P), *Sac*I (Sa), *Sph*I (S) and *Xma*I (X). The open reading frame is indicated as a thick line. The empty box represents an Alu sequence. **(B)** Nucleotide sequence of the λX47 α subunit cDNA clone. The translation initiation codon is indicated by an asterisk. The termination codon and the Alu sequence in the 3′ untranslated region are underlined. The putative polyadenylation signal is boxed.

tion for the deglycosylated p150,95 α chain of 132 000 M_r (Miller and Springer, 1987). The putative extracellular domain contains 1081 amino acids with 10 potential N-glycosylation sites. Previous studies on the α subunit of p150,95 suggest that five or six sites are glycosylated, confirming that this is the extracellular domain and accounting for the rest of the molecular mass (Miller and Springer, 1987). A hydrophobic sequence with features of a transmembrane region separates the extracellular domain from a 30-amino-acid hydrophilic cytoplasmic tail.

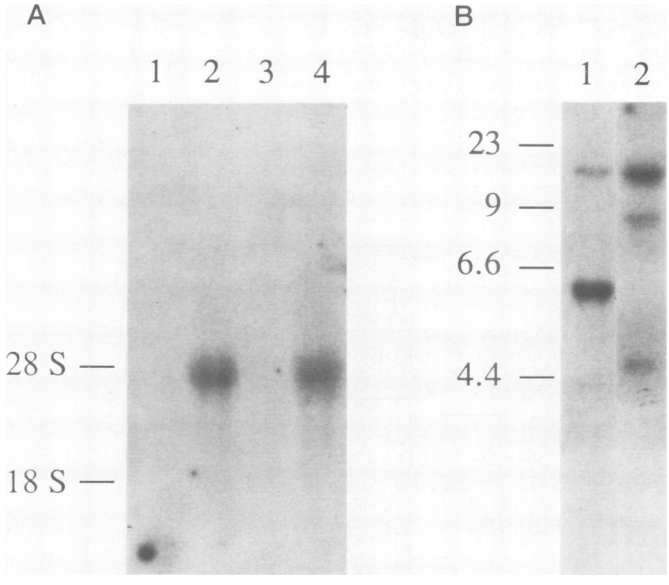


Fig. 2. Northern and Southern blots. (**A**) Northern blot analysis of 20 μg of total RNA from U937 (**lanes 1** and **2**) and HL-60 (**lanes 3** and **4**) cells before (**lanes 1,3**) and after (**lanes 2,4**) culture for 3 days with 2 or 10 μg/ml of PMA respectively. The positions of 28S and 18S rRNA are indicated. (**B**) Southern blot of *Eco*RI- (**lane 1**) and *Hin*dIII-digested (**lane 2**) hairy cell leukemia DNA. *Hin*dIII-cut λ DNA markers are indicated in kb.

A search for internal homologies in the amino acid sequence revealed the existence of three tandem homologous repeats of ~65 residues located between residues 424 to 619 in the extracellular domain (Figure 4) as well as weaker repeats (see below). The homology between the three stronger repeats is 25–33%, significant at $P < 10^{-6}$. The similarity is highest for the central 27 amino acids of the repeats, which are 40–51% identical. At least two of the repeats (II and III) are separated by an intervening sequence (data not shown), suggesting the possibility that the repeats are encoded by separate exons.

A requirement for divalent cations has been demonstrated for leukocyte adhesion receptor-mediated cell–cell interaction (Anderson and Springer, 1987), and for binding of purified Mac-1 and p150,95 to the complement component iC3b (Wright *et al.*, 1983; Micklem and Sim, 1985). Examination of the p150,95 α subunit sequence reveals three regions, located in the center of the tandem repeats, which resemble the 'EF hand' loop structure of the Ca^{2+}- and Mg^{2+}-binding proteins parvalbumin, calmodulin and troponin C (Szebenyi *et al.*, 1981) in the position of oxygen-containing side chains which coordinate with metal and in the frequent occurrence of intervening glycine residues (Figure 4).

Together with previous studies on the β subunit, this report defines the first complete primary structure of a leukocyte adhesion receptor, which is summarized in Figure 3. There is no homology between the α and β subunits. Overall, the structure is very similar to that of the extracellular matrix receptors. Previous studies demonstrated a striking homology between the β subunit of the leukocyte adhesion glycoproteins (integrin β_2) and the β subunit of the human and chicken fibronectin receptor (integrin β_1) (45–46% identity) and the β subunit common to

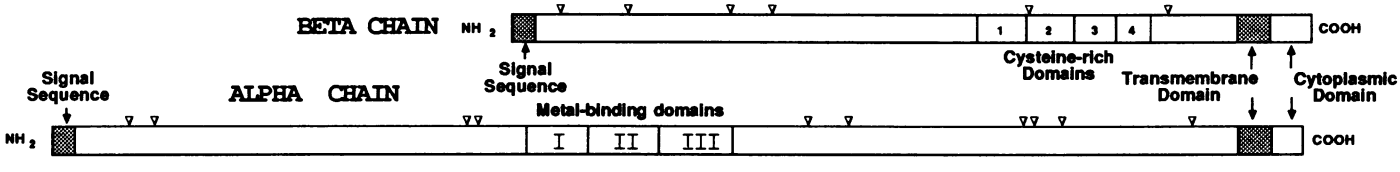


Fig. 3. Schematic representation of the complete p150,95 molecule, based on results reported here and by Kishimoto *et al.* (1987) and Law *et al.* (1987) for the β subunit. Potential N-glycosylation sites are indicated by triangles.

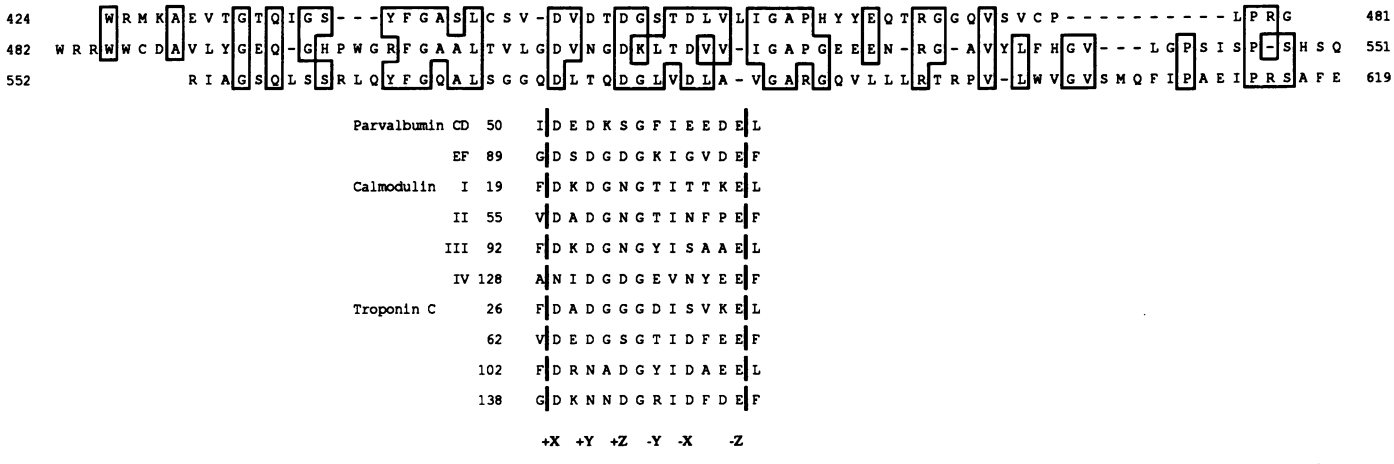


Fig. 4. Homologous repeats and putative divalent cation-binding sites. Common residues in the three homologous repeats are boxed. The sequences of known Ca^{2+}- and/or Mg^{2+}-binding sites (Szebenyi *et al.*, 1981) are shown below the p150,95 α subunit homologous repeats and are aligned with the putative divalent cation-binding sites. The coordination axes of residues ligating the metal are shown. The vertical dashed line demarcates the border between the EF-hand loop and the surrounding α helical segments (Szebenyi *et al.*, 1981).

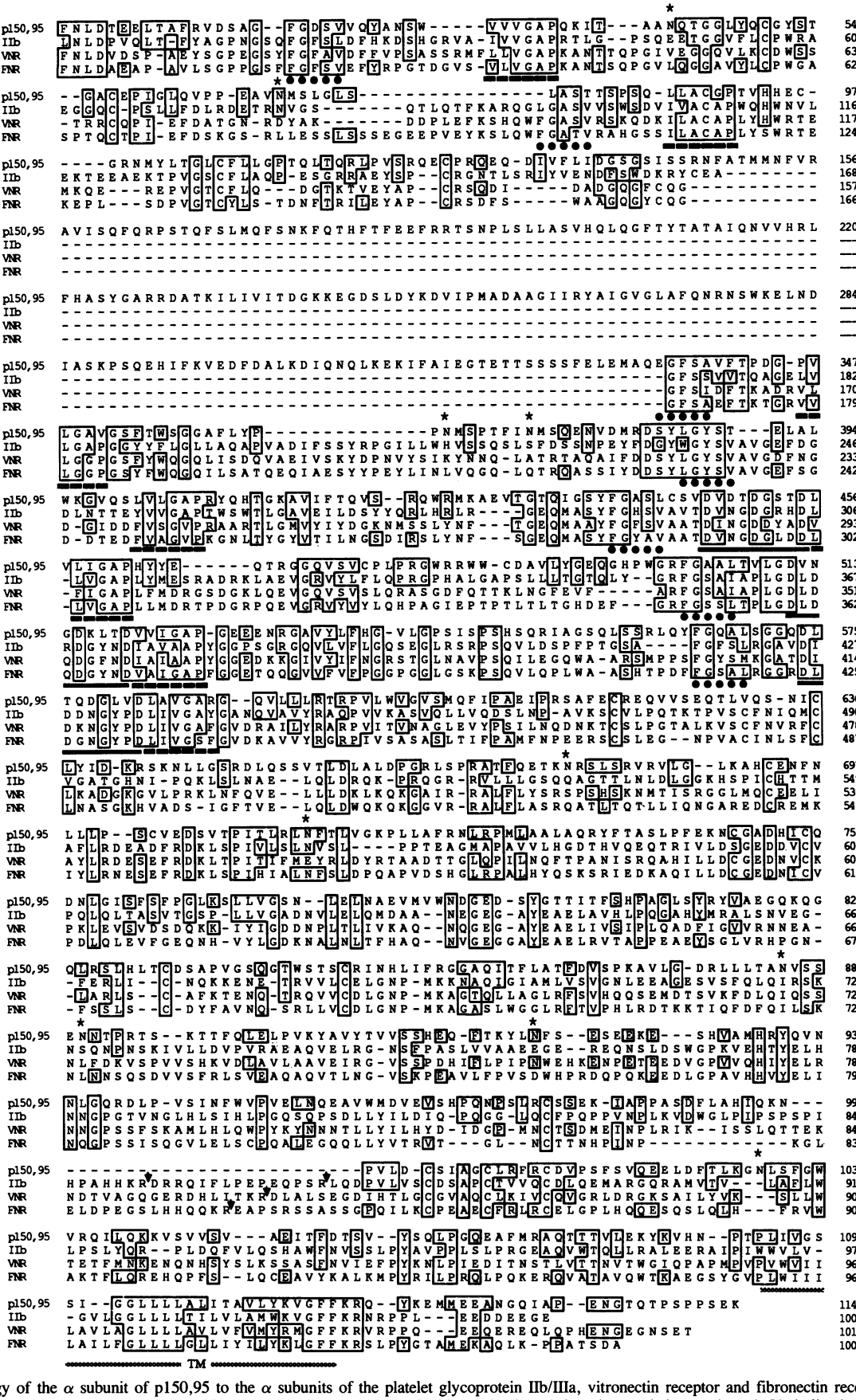


Fig. 5. Homology of the α subunit of p150,95 to the α subunits of the platelet glycoprotein IIb/IIIa, vitronectin receptor and fibronectin receptor (Argraves *et al.*, 1987; Poncz *et al.*,1987; Suzuki *et al.*, 1987). Identities between the p150,95 α subunit and the other α chains are boxed. Underlined areas correspond to the putative divalent cation-binding sites in the p150,95 α subunit. The conserved flanking sequences are indicated by dashed and dotted lines. The cleavage sites for the RGD-receptor α subunits are indicated by arrowheads (Charo *et al.*, 1986; Argraves *et al.*, 1987; Poncz *et al.*, 1987; Suzuki *et al.*, 1987). Potential N-glycosylation sites in the p150,95 α subunit are indicated by asterisks. The transmembrane region is underlined by a shaded line.

the human platelet gpIIb/IIIa and vitronectin receptor (integrin β_3) (37% identity) (Tamkun *et al.*, 1986; Fitzgerald *et al.*, 1987; Kishimoto *et al.*, 1987; Law *et al.*, 1987). These latter β subunits are present in $\alpha\beta$ complexes which mediate cell−substrate adhesion by binding to the sequence Arg-Gly-Asp (RGD) within extracellular matrix components such as fibronectin, vitronectin and fibrinogen. Thus three subfamilies of adhesion receptors have been defined, which have been designated the integrins (Hynes, 1987). We examined the structural relationship between the p150,95 leukocyte adhesion receptor α subunit and the other integrin α subunits which previously had been suspected based only on short N-terminal sequence comparisons (Charo *et al.*, 1986; Miller *et al.*, 1987; Takada *et al.*, 1987).

The p150,95 α subunit sequence was compared with the recently obtained platelet glycoprotein IIb/IIIa (Poncz *et al.*, 1987), vitronectin and fibronectin receptor α subunit sequences (Argraves *et al.*, 1987; Suzuki *et al.*, 1987) (Figure 5). The overall homology between the α subunit of p150,95 and each of the three extracellular matrix receptor α subunits is 25% and is highly significant ($P < 10^{-36}$). N-terminal sequencing (Miller *et al.*, 1987) and preliminary cDNA sequencing data (R.S.Larson, A.L.Corbi and T.A.Springer, unpublished) show that the α subunits of the leukocyte adhesion receptors, p150,95, Mac-1 and LFA-1 are 35−64% identical. In comparison, the vitronectin receptor and gpIIb/IIIa, which share the integrin β_3 subunit, are 37% homologous to one another, whereas the β_3-associated α subunits and the β_1-associated fibronectin receptor α subunit are 36−44% homologous to one another.

There are two main distinctive features of the p150,95 α subunit. A region of 187 amino acids (residues 149−335) present in the p150,95 α chain is absent in the other three receptors, accounting for most of the difference in length between the α subunit of p150,95 (1144 residues) and the matrix receptors (1018−1007 residues). In contrast to p150,95, the α subunits of the fibronectin and vitronectin receptors and gpIIb/IIIa consist of two disulfide-bonded chains (heavy and light) which are formed by proteolytic processing of a precursor (arrowheads in Figure 5) (Charo *et al.*, 1986; Argraves *et al.*, 1987; Poncz *et al.*, 1987; Suzuki *et al.*, 1987). The cleavage sites in the three RGD receptor α subunits are located at different positions within a non-homologous stretch of 26−35 amino acids. This non-homologous region is deleted in the p150,95 α subunit (gap after residue 998), correlating with the lack of proteolytic processing of the p150,95 α subunit (Miller and Springer, 1987). These distinctive features of p150,95 are shared with Mac-1 (A.L.Corbi and T.A.Springer, unpublished), suggesting that they are characteristic of leukocyte adhesion receptor α subunits.

A short region of high conservation (62−67%) is found in the inner 2/3 of the transmembrane region (residues 1096−1116). The transmembrane regions of the integrin β subunits are also highly conserved (Kishimoto *et al.*, 1987; Law *et al.*, 1987) and are unrelated to the α subunit transmembrane sequences. The high conservation of both α and β subunit transmembrane sequences suggests that some important interaction may take place in this region, either between the α and β subunits or with other cellular components, possibly involving transmembrane signalling.

The longest region of extensive homology between p150,95 and the other three α subunits (residues 432−599, 32−38% identify) corresponds closely to the three tandem homologous repeats which contain the putative divalent cation-binding sites in the p150,95 α subunit. The vitronectin receptor, fibronectin receptor and gpIIb/IIIa contain very similar putative divalent cation binding sites (Argraves *et al.*, 1987; Poncz *et al.*, 1987; Suzuki *et al.*, 1987). Although thus far only the internal homologies in the putative divalent cation binding sites have been pointed out for these receptors, our analysis shows that the flanking sequences also have internal homologies, suggesting the presence of tandem repeats of ~65 amino acids similar to those described here for the p150,95 α subunit. For example, consensus sequences containing $FG\begin{smallmatrix}F\\Y\\A\end{smallmatrix}S\begin{smallmatrix}V\\L\end{smallmatrix}$ and $\begin{smallmatrix}LV\\VL\end{smallmatrix}\begin{smallmatrix}V\\I\\L\end{smallmatrix}$ GAP are found before and after the putative divalent cation-binding sites respectively, and at four additional positions in all four receptors (residues 18−35, 76−91, 337−351 and 386−406, dots and dashes in Figure 5) in which the intervening cation-binding sites are lacking except for one instance. A putative metal ion-binding site distinctive in its content of glutamic acid is found for the matrix receptors but not p150,95 α at residues 386−406. The presence of the three and four weak homologous repeats suggests that much of the N-terminal half of the α subunit may have evolved by tandem duplication; the 187 residues present in p150,95 α without counterpart in the three matrix receptor α subunits does not contain similar elements.

All the integrins require Ca^{2+} or Mg^{2+} for ligand binding and the gpIIb/IIIa α subunit has been directly demonstrated to bind Ca^{2+} (Fujimura and Phillips, 1983). Thus it is likely that the EF-hand loop-like sequences in p150,95 bind divalent cations and reflect the divalent cation requirement for ligand binding. Secondary structure predictions (Chou and Fasman, 1974) suggest that the putative divalent cation-binding sites are embedded in β pleated sheet rather than between α helices as in the EF hand; furthermore, the proline in GAP is not consistent with an α helical segment following the loop (Figure 4). A further difference with the classic EF-hand loop structure is the absence of the E (glutamic acid) residue which the X-ray structures demonstrate is coordinated with the metal in the −Z position (Figure 4). This may leave the metal free to coordinate in the −Z position with the ligand. It is tempting to speculate that the metal bound to the receptor may coordinate with the D residue of RGD in the ligand; this sequence appears suited for metal binding since the sequences GD and DG occur frequently in the EF-hand loop and in the putative metal-binding domains of all four integrin α subunits. In addition to the requirement for Ca^{2+} or Mg^{2+} for ligand binding, the idea that the α subunit tandem repeats are important in ligand binding is attractive because the α subunits control ligand specificity, and there is evidence for distinct binding sites for different ligands (Anderson *et al.*, 1986; Dana *et al.*, 1986). The large insertion unique to p150,95 which is N-terminal to the putative divalent cation-binding sites may also be important in ligand binding. Studies are currently underway to test these ideas.

Our study demonstrates that the leukocyte adhesion receptor α subunit belongs to the same integrin protein family as the matrix receptor α subunits. It is interesting that similar $\alpha\beta$ receptor structures have evolved to mediate both cell−cell and cell−matrix interactions. These receptors may be important not only in facilitating cell contact and localization, but in modulating in a position-specific fashion responses to other stimuli. Further studies are required to determine whether the leukocyte adhesion receptors recognize RGD-like sequences within their ligands (Wright *et al.*, 1987), interact with the cytoskeleton, and have a regulatable affinity for ligand, as has been demonstrated for some of the matrix receptor integrins (Hynes *et al.*, 1987).

Materials and methods

Protein purification and sequencing

The p150,95 molecule was purified from detergent lysates of hairy cell leukemia spleens using monoclonal antibody affinity chromatography (Miller *et al.*, 1987). After preparative SDS–PAGE, the α chain was electroeluted (Hunkapiller *et* al., 1983), and reduced and alkylated. After ethanol precipitation at −20°C for 16 h, the protein was dissolved in 0.1 M ammonium bicarbonate (pH 8–8.5) containing 0.1 mM calcium chloride and 0.2% zwittergent 3–14 (Calbiochem). Digestion was carried out with 2% (w/w) trypsin at 37°C for 6 h, with addition of 1% (w/w) trypsin every 2 h. The tryptic fragments were separated by reverse-phase HPLC on a C4 column (Vydac) and eluted with a gradient of acetonitrile (0–60%) in 0.1% TFA. Individual peptides were concentrated on a Speed Vac concentrator (Savac) and subjected to microsequencing on an Applied Biosystems gas–liquid phase sequencer.

Construction and screening of the cDNA library

Total RNA was prepared from PMA-differentiated HL-60 cells as previously described (Chirgwin *et al.*, 1979) and poly(A)$^+$ RNA isolated by oligo(dT) chromatography (Aviv and Leder, 1972). Poly(A)$^+$ RNA (10 μg) was used to construct a cDNA library following the method of Gubler and Hoffman (1983). Synthesis of the cDNA first strand was carried out using oligo(dT) (Pharmacia) and reverse transcriptase (Life Sciences). Double-stranded cDNA synthesis was completed with DNA polymerase I (New England Biolabs), ribonuclease H (Boehringer Mannheim) and *Escherichia coli* DNA ligase (New England Biolabs). Following the blunting and methylation reactions, phosphorylated *Eco*RI linkers were ligated to the double-stranded cDNA. After *Eco*RI digestion, linkers were removed in a 5-ml Sepharose CL 4B column (Pharmacia) and the first two-thirds of the double-stranded cDNA peak recovered. To select the library for long inserts, the double-stranded cDNA was size-fractionated on a 0.8% low melting point agarose gel and the double-stranded cDNA between 2 and 7 kb was recovered by electroelution onto NA-45 paper (Schleicher and Schuell) and ligated into *Eco*RI-digested λgt10. Primary recombinants (3.2×10^6) were obtained after packaging and plating on *E.coli* c600 hfl.

Primary recombinants (5×10^5) were plated at 25 000 plaques/150-mm plate. Plaques were amplified *in situ* on duplicate nitrocellulose filters (Woo, 1979), processed (Benton and Davis, 1977) and prehybridized overnight at 37°C in 6 × SSC, 1 × Denhardt's, 0.5% SDS, 0.05% Pi/PPi and 100 μg/ml of tRNA. Hybridization was carried out at 37°C with a single sequence 46-mer oligonucleotide (5′-CCTCCTGGTTCAGCTCCACAGGCACCCAGAAGTTG-ATGGAGACAGG-3′). Filters were washed in 6 × SSC, 0.1% SDS, 0.05% Pi/PPi at room temperature for 30 min and 50°C for 15 min. After overnight autoradiographic exposure, plaques that gave positive signals on duplicate filters were purified by successive platings and rescreenings. Their DNA was purified (Benson and Taylor, 1984) and their inserts sized and probed with an independent 49-mer (5′CATCCTGGGTCAGGTCCTGGCCGCCAGACAGGGCCTGG-CCAAAGTACTG-3′) and an N terminus oligonucleotide probe (5′-AGCAG-AGTCCACCCGGAAGGCTGTCAGCTC-3′).

Restriction mapping and sequencing

Positive cDNA clones were restriction-mapped by the end-labeling partial digestion procedure (Smith and Bernstiel, 1976). Restriction fragments were subcloned into M13 mp18 and mp19 (Messing, 1983). Both strands were sequenced by the dideoxy termination method (Sanger *et al.*, 1977) using [^{35}S]dATP. All restriction sites were crossed in sequencing in both orientations. Oligonucleotide-primed dideoxy sequencing was used to confirm the sequence of some regions.

Southern and Northern blots

For Southern blots, 10 μg of hairy cell leukemia DNA were digested with *Eco*RI or *Hin*dIII. After electrophoresis on a 0.8% agarose gel, the DNA was transferred onto a nylon membrane (Zetaprobe, Biorad), prehybridized and hybridized following standard procedures (Maniatis *et al.*, 1982) using the 1.2-kb *Eco*RI–*Pst*I cDNA fragment as probe. Northern blots used 20 μg of total cellular RNA from either PMA-stimulated or untreated HL-60 and U937 cells. The transferred RNA was probed with the 0.8-kb *Pst*I fragment from the λX47 cDNA clone.

Sequence homologies

Alignment of the sequences of the α subunits of p150,95, vitronectin and fibronectin receptors and gpIIb was initially performed using the Microgenie DNA program (Beckman) and the ALIGN program (Bionet). The alignment was then maximized using the GENALIGN program (Bionet). Gaps were included in calculations of amino acid identity by scoring their beginning and end as mismatches.

Acknowledgements

The authors gratefully acknowledge Bill Lane for sequencing of the tryptic peptides and Drs J.S.Bennett and E.Ruoslahti for communicating prepublication manuscripts. We thank Dr Thomas Roberts for helpful discussion, advice and encouragement. This work was supported by NIH grant CA 31799, an American Cancer Society Award to T.S.A. and a Fundacion Juan March fellowship to A.L.C.

References

Anderson,D.C. and Springer,T.A. (1987) *Annu. Rev. Med.*, **38**, 175–194.
Anderson,D.C., Miller,L.J., Schmalstieg,F.C., Rothlein,R. and Springer,T.A. (1986) *J. Immunol*, **137**, 15–27.
Argraves,W.S., Suzuki,S., Arai,H., Thompson,K., Pierschbacher,M.D. and Ruoslahti,E. (1987) *J. Cell. Biol.*, **105**, 1183–1190.
Aviv,H. and Leder,P. (1982) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
Beller,D.I., Springer,T.A. and Schreiber,R.D. (1982) *J. Exp. Med.*, **156**, 1000–1009.
Benson,S.A. and Taylor,R.K. (1984) *Bio Tech.*, **4**, 126–127.
Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
Caligaris-Cappio,F., Pizzolo,G., Chilosi,M., Bergui,L., Semenzato,G., Tesio,L., Morittu,L., Malavasi,F., Gobbi,M., Schwarting,R., Campana,D. and Janossy,G. (1985) *Blood*, **66**, 1035–1042.
Charo,I.F., Fitzgerald,L.A., Steiner,B., Rall,S.C.,Jr, Bekeart,L.S. and Phillips,D.R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8351–8355.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Chou,P.Y. and Fasman,G.D. (1974) *Biochemistry*, **13**, 211–245.
Dana,N., Styrt,B., Griffin,J.D., Todd,R.F.,III, Klempner,M.S. and Arnaout,M.A. (1986) *J. Immunol.*, **137**, 3259–3263.
Fitzgerald,L.A., Steiner,B., Rall,S.C.,Jr, Lo,S. and Phillips,D.R. (1987) *J. Biol. Chem.*, **262**, 3936–3939.
Fujimura,K. and Phillips,D.R. (1983) *Thromb. Haemostasis*, **50**, 251a.
Gubler,U. and Hoffman,B.J. (1983) *Gene*, **25**, 263–269.
Hunkapiller,M.W., Lujan,E., Ostrander,F. and Hood,L.E. (1983) *Methods Enzymol.*, **91**, 227–236.
Hynes,R.O. (1987) *Cell*, **48**, 549–554.
Keizer,G.D., Borst,J., Visser,W., Schwarting,R., de Vries,J.E. and Figdor,C.G. (1987) *J. Immunol.*, **138**, 3130–3136.
Kishimoto,T.K., O'Connor,K., Lee,A., Roberts,T.M. and Springer,T.A. (1987) *Cell*, **48**, 681–690.
Law,S.K.A., Gagnon,J., Hildreth,J.E.K., Wells,C.E., Willis,A.C. and Wong,A.J. (1987) *EMBO J.*, **6**, 915–919.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Micklem,K.J. and Sim,R.B. (1985) *Biochem. J.*, **231**, 233–236.
Miller,L.J. and Springer,T.A. (1987) *J. Immunol.*, **139**, 842–847.
Miller,L.J., Schwarting,R. and Springer,T.A. (1986) *J. Immunol.*, **137**, 2891–2900.
Miller,L.J., Wiebe,M. and Springer,T.A. (1987) *J. Immunol.*, **138**, 2381–2383.
Poncz,M., Eisman,R., Heidenreich,R., Silver,S.M., Vilaire,G., Surrey,S., Schwartz,E. and Bennett,J.S. (1987) *J. Biol., Chem.*, **262**, 8476–8482.
Sanchez-Madrid,F., Krensky,A.M., Ware,C.F., Robbins,E., Strominger,J.L., Burakoff,S.J. and Springer,T.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7489–7493.
Sanchez-Madrid,F., Nagy,J., Robbins,E., Simon,P. and Springer,T.A. (1983a) *J. Exp. Med.*, **158**, 1785–1803.
Sanchez-Madrid,F., Simon,P., Thompson,S. and Springer,T.A. (1983b) *J. Exp. Med.*, **158**, 586–602.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schwarting,R., Stein,H. and Wang,C.Y. (1985) *Blood*, **65**, 974–983.
Smith,H.O. and Bernstiel,M.L. (1976) *Nucleic Acids Res.*, **3**, 2387–2398.
Springer,T.A., Miller,L.J. and Anderson,D.C. (1986) *J. Immunol.*, **136**, 240–245.
Suzuki,S., Argraves,W.S., Arai,H., Languino,L.R., Pierschbacher,M.D. and Ruoslahti,E. (1987) *J. Biol. Chem.*, **262**, 14080–14085.
Szebenyi,D.M.E. Obendorf,S.K. and Moffat,K. (1981) *Nature*, **294**, 327–332.
Takada,Y., Strominger,J.L. and Hemler,M.E. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 3239–3243.
Tamkun,J.W., DeSimone,D.W., Fonda,D., Patel,R.S., Buck,C., Horwitz,A.F. and Hynes,R.O. (1986) *Cell*, **46**, 271–282.
Te Velde,A.A., Keizer,G.D. amd Figdor,C.G. (1987) *Immunology*, **61**, 261–267.
Wood,S.L.C. (1979) *Methods Enzymol.*, **68**, 389–395.
Wright,S.D., Rao,P.E., Van Voorhis,W.C., Craigmyle,L.S., Iida,K., Talle,M.A., Westberg,E.F., Glodstein,G. and Silverstein,S.C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5699–5703.
Wright,S.D., Reddy,P.A., Jong,M.T.C. and Erickson,B.W. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1965–1968.

Received on September 25, 1987

Note added in proof

The sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession number Y00093.