THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 263, No. 25, Issue of September 5, pp. 12403-12411, 1988
Printed in U.S.A.

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*

(Received for publication, April 15, 1988)

Angel L. Corbi, Takashi K. Kishimoto, Linda J. Miller, and Timothy A. Springer‡

From the Laboratory of Membrane Immunochemistry, Dana-Farber Cancer Institute, Boston, Massachusetts 02115

Mac-1 (CD 11b/CD18) is a leukocyte adhesion heterodimeric glycoprotein which functions both as a receptor for iC3b (CR3) and in several cell-cell and cell-substrate adhesive interactions. We describe full-length cDNA clones for the α subunit of Mac-1. Mac-1 α subunit message was detected in blood monocytes and phorbol-12-myristate acetate-induced myeloid cell lines, but not in cells of the T or B lineages, correlating with Mac-1 protein surface expression. The α subunit of Mac-1 is a transmembrane protein of 1137 residues with a long extracellular domain (1092 residues) and a 19-amino acid cytoplasmic tail. The extracellular domain contains three putative divalent cation-binding sequences and 19 potential *N*-glycosylation sites. The amino acid sequence of Mac-1 α shows that it is a member of the integrin superfamily; Mac-1 α shows 63% identity to the α subunit of the leukocyte adhesion glycoprotein p150,95 and 25% to the α subunits of the extracellular matrix receptors platelet glycoprotein IIb/IIIa, the fibronectin receptor, and the vitronectin receptor. The Mac-1 α subunit putative divalent cation-binding sites and the flanking regions exhibit a high degree of identity both to the p150,95 α subunit (87% identity at the amino acid level) and to the rest of the integrin α subunits (38%). The α subunit of Mac-1, like the p150,95 α subunit, contains a domain of 187 amino acids in the extracellular region which is absent in other integrins. This leukocyte or "L" domain is homologous to the A domains of von Willebrand factor, which in turn are homologous to regions of the C3-binding proteins factor B and C2. These findings draw attention to this region of Mac-1 as a potential binding site for iC3b.

The leukocyte adhesion receptors Mac-1, LFA-1, and p150,95 are cell surface glycoproteins which play a fundamental role in the adhesive interactions of myeloid and lymphoid cells (1). The three molecules are α/β heterodimers composed of a common β subunit (M_r 95,000) noncovalently associated to unique α subunits (LFA-1α, M_r 180,000; Mac-1α, M_r 170,000; p150,95α, M_r 150,000) (2). Mac-1, LFA-1, and p150,95 are structurally related (3) but differ in their function and cellular distribution (4). LFA-1 is present on all leukocytes, whereas Mac-1 and p150,95 are almost exclusively expressed on cells of the myeloid lineage (5, 6). The sequences of the common β subunit (7, 8) and the p150,95 α subunit (9) are related to those of the extracellular matrix receptors (ECM)[1] (10–14), defining a family of adhesion receptors termed integrins (15).

The functional role of Mac-1 was first illustrated by the ability of anti-Mac-1 α subunit monoclonal antibodies (mAb) to block the rosetting of iC3b-coated erythrocytes to macrophages and polymorphonuclear leukocytes (16), demonstrating that Mac-1 is indistinguishable from the complement receptor type three (CR3). Subsequently, the involvement of Mac-1 in inflammatory processes was evidenced by the inhibition of neutrophil aggregation and adhesion to endothelial cells by anti-Mac-1 α subunit and anti-β subunit-specific mAb (4, 17, 18). Recent epitope mapping studies have suggested that the sites involved in iC3b binding are distinct from those involved in neutrophil aggregation and adherence to protein-coated plastic (4, 17, 19). Therefore, Mac-1 appears to be a multivalent receptor with at least two independent adhesion-related functions. The functional importance of Mac-1 has been further emphasized by the clinical symptoms exhibited by leukocyte adhesion deficiency (LAD) patients (1). LAD leukocytes are deficient in Mac-1, LFA-1, and p150,95 due to mutations in the common β subunit (20). As a consequence, neutrophil and monocyte recruitment into the inflamed tissue is absent in LAD patients, leading to recurrent bacterial infections, impaired wound healing, and childhood morbidity (1).

The expression and functional activity of Mac-1 is regulated during leukocyte differentiation and activation. Differentiation and maturation of myelomonocytic cell lines results in increased Mac-1 expression (5), while blood monocyte differentiation into tissue macrophages is accompanied by a considerable decrease in the amount of Mac-1 on the cell surface (6). The expression of Mac-1 on the surface of circulating neutrophils and monocytes is upregulated by inflammatory stimuli; Mac-1 is stored in an intracellular vesicular compartment which is rapidly mobilized to the cell surface by chemoattractants (21, 22). Although the augmented expression of Mac-1 can lead to increased adhesiveness, qualitative changes

* This work was supported by National Institutes of Health Grants CA 31799 and AI 20316, an American Cancer Society Faculty Award (to T. A. S.) and a Fundación Juan March and a Fogarty International Fellowship (to A. L. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J03925.

‡ To whom correspondence should be addressed: Center for Blood Research, 800 Huntington Ave., Boston, MA 02115.

[1] The abbreviations used are: ECM, extracellular matrix; CR3, complement receptor type 3; LAD, leukocyte adhesion deficiency; mAb, monoclonal antibody; PMA, phorbol-12-myristate acetate; vWF, von Willebrand factor; SDS, sodium dodecyl sulfate; kb, kilobase pair(s); HPLC, high performance liquid chromatography.

after cell activation may also be important in regulating ligand binding (23, 24). Both the qualitative and quantitative changes may be important in regulation of leukocyte binding to postcapillary endothelium at inflammatory sites.

We have previously described the N-terminal sequence of the murine and human Mac-1 α subunits (3, 25) and a murine genomic clone encoding a short N-terminal exon (26). In the present report, we describe the cDNA cloning and complete primary structure of the α subunit of human Mac-1. Most of the Mac-1 α subunit is similar to the α subunits of the extracellular matrix receptor integrins, with an additional domain which is related to the A repeats of von Willebrand factor and to two C3-binding proteins, Factor B and C2.

EXPERIMENTAL PROCEDURES

Protein Purification and Sequencing—Human Mac-1 was affinity-purified using anti-Mac-1 α subunit LM2/1 mAb-Sepharose, as described in detail (3). Additionally, human Mac-1 was purified from lysates of Sendai virus-induced leukocytes (predominantly neutrophils) by affinity chromatography on an IB4.5 (27) anti-β mAb Protein A-Sepharose column (28). The preparation of the lysate has been previously described (3). The lysate was passed through a 3-ml protein A-Sepharose precolumn connected in series with the 3.5-ml IB4.5 column after pre-washing of both columns with 50 ml of 0.5% Triton X-100, 0.5% sodium deoxycholate, 20 mM Tris-HCl, pH 8. After loading the lysate, the column was washed with 40 ml of the same solution and the elution of the bound material was carried out by successive washings with buffers of increasing pH or ionic strength: 1) 10 ml of 0.1% Triton X-100, 0.1 M glycine, pH 9; 2) 20 ml of 0.1% Triton X-100, 0.1 M glycine, pH 10; 3) 50 ml of 0.1% Triton X-100, 0.1 M triethylamine, pH 11.5; and 4) 40 ml of 0.14 M NaCl, 0.5% Triton X-100, 0.01 M Tris-HCl, pH 8. SDS-polyacrylamide gel electrophoresis of the fractions showed that most of the bound material was eluted at pH 11.5. Immunoprecipitation with subunit-specific mAb showed that Mac-1 was the prevalent molecule eluted from the anti-β mAb-Sepharose column, as expected because neutrophils are the primary cell in the leukocyte lysate and express much more Mac-1 than LFA-1 or p150,95.

The α subunit of Mac-1 was isolated by preparative SDS-polyacrylamide gel electrophoresis of the affinity-purified antigen. After electroelution, the isolated α subunit was precipitated with ethanol and reduced and alkylated. 50 μg of purified α subunit was dissolved in 0.1 M ammonium bicarbonate, 0.1 mM calcium chloride, 0.3% zwittergent 3-14, and digested with 2% (w/w) trypsin at 37 °C for 6 h, with further additions of 1% trypsin every 2 h. The resulting peptides were separated by reverse-phase HPLC on a C4 column (Vydac) and eluted with a gradient of acetonitrile (0–60%) in 0.1% trifluoroacetic acid for 2 h. Collected fractions were concentrated and subjected to microsequencing on an Applied Biosystems gas-liquid phase sequencer.

Isolation and Sequencing of cDNA Clones—Peptide sequences were used to design four single sequence oligonucleotide probes. The oligonucleotides were end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP (29) and used to screen a size-selected cDNA library from PMA-induced HL-60 cells (9). 5 × 10^6 primary recombinants were plated, transferred to duplicate nitrocellulose filters, and prehybridized overnight as described (9). Hybridization with the oligonucleotides was done in 6 × SSC, 0.1% SDS, 0.05% sodium pyrophosphate, and 100 μg/ml of tRNA at 37 °C overnight. Filters were washed in the same solution without the tRNA at room temperature for 30 min and at 45 °C for 15 min. Wet filters were exposed overnight to preflashed x-ray film with intensifying screen. Phage plaques giving duplicate positive signals were obtained after screening with a 42-mer oligonucleotide (5′-ACCCAGGTGACCTTCTTCTTCCCCCT-AGACCTGTCCTACCGG-3′) and were subjected to three additional rounds of subcloning and screening with the same probe. Isolation of full-length cDNA clones was carried out by re-screening the filters with an end-labeled oligonucleotide (5′-GGATGGACTGGTAGA-CCTGACTGTAGGAGC-3′) and nick-translated probes from the 5′ end of the partial cDNA clone λM14.

Restriction Mapping and Sequencing—DNA from the positive phages was purified, cloned into pUC13, 18, or 19, and restriction mapped using standard procedures (29). Restriction fragments were subcloned into M13 mp18 and mp19 and sequenced by the dideoxy termination method (30). Oligonucleotide-primed DNA sequencing was used in cases where no convenient restriction sites were available. The whole coding region, the 5′-untranslated region, and more than 60% of the 3′-untranslated region were sequenced in both strands. The 3′-untranslated regions of the cDNA clones λM23 and λM90 were subjected to dideoxy sequencing of plasmid DNA and using the Erase-a-base system (Promega) to make deletions according to the manufacturer's instructions (31).

Northern Blot—Adherent mononuclear cells from peripheral blood were isolated by Ficoll-Hypaque centrifugation and incubation of the mononuclear cells in tissue culture plates with RPMI 1640 and 10% fetal calf serum for 30 min at 37 °C. Nonadherent cells were removed from the plates by extensive washing with RPMI 1640. Adherent cells were detached from the plates by incubation with 10 ml of phosphate-buffered saline, 5 mM EDTA, for 15 min at 37 °C. More than 95% of the adherent cells were positive for the presence of Mac-1 as detected by indirect immunofluorescence. PMA treatment of the cell lines HL-60 and U937 was performed as described (5). Total RNA was extracted from peripheral blood adherent cells, as well as from the cell lines SKW3, JY, HL-60, and U937, using guanidine isothiocyanate (29). 10 or 20 μg from each sample was run on a 1% formaldehyde-agarose gel and transferred onto a nylon membrane. Hybridizations were performed using the 2.3-kb *Eco*RI fragment from λM23 as a probe.

Sequence Homologies—The sequence of Mac-1 α was compared with the protein sequence database in the National Biomedical Research Foundation (NBRF) (Washington, D. C.). The ALIGN program was used for alignment of sequences (32). The statistical significance of the alignments was assessed obtaining the alignment scores for 100 random permutations of the aligned sequences and calculating the number of standard deviations between the mean of the scores for the randomized comparisons and the score of the actual alignment. The scoring used the 250 PAM mutation data matrix, with a gap penalty = 6 and a bias = 6.

RESULTS AND DISCUSSION

Isolation of cDNA Clones—The Mac-1 α/β complex was purified from leukocyte Triton X-100 lysates by mAb-affinity chromatography. After preparative SDS-polyacrylamide gel electrophoresis and electroelution, the Mac-1 α subunit was homogeneous as described previously (3). The purified α subunit was digested with trypsin and the resulting peptides separated by reverse-phase HPLC and subjected to protein microsequencing (Table I). Comparison of the Mac-1 α subunit peptide sequences with the sequence of the p150,95 α subunit (9) showed a high degree of similarity. The peptide 88 was selected to design a 42-mer oligonucleotide specific for

TABLE I

Mac-1 α subunit tryptic peptide sequences

The peptides and their positions in the cDNA-derived sequence are shown. Positions without amino acid assignment are denoted by an *X*. Uncertain assignments are indicated between parentheses. Peptides were obtained from Mac-1 purified on IB4 mAb protein A-Sepharose, except for peptide 86 that was obtained from Mac-1 purified on LM2/1 mAb-Sepharose and peptides 88 and 90 that were obtained from both sources. The peptide sequence used for the design of the oligonucleotide probe is underlined.

		Residues
14	TIQNQLR	307–313
32	VQSLVLGAPR	400–409
43	YVIGVGDAFR	267–277
44	YQHIGLVAMFR	410–420
53A	WQC(D)AVLYGEQGQPXGR	488–504
53B	EFV(S)XX(MEQL)	155–164
54A	F(G)DPLGYEDVIPEADR	246–261
54B	S(L)VK(P)ITQLLG(R)	197–208
71	LFTALFPFEK	744–753
79A	VDSDMNDAYLG(Y)	379–390
79B	XQ(C)XIPFFGIQE	1015–1026
86	GCPQEDSDIAFLIDGSGSIIPHDFR	127–151
88	TQVTFFFPLDLSYR	801–814
90	LXFSLVGTPLSAFGNLRPVLAEDAQR	718–743

the α subunit of Mac-1 because of its low level of redundancy and its homology to a region of the p150,95 α subunit close to the C terminus (*i.e.* towards the 3′ end of a cDNA).

Screening of 5×10^6 primary recombinants from a PMA-induced HL-60 cDNA library with the 42-mer yielded 16 positive clones and the longest one (λM14, 2.9 kb) was selected for sequencing (Fig. 1). The λM14 cDNA-derived amino acid sequence encodes four of the tryptic peptides derived from the purified Mac-1 α subunit. To isolate a full-length Mac-1 α cDNA clone, the library was re-screened with a 1.0-kb *Eco*RI fragment and a 30-mer derived from the 5′ end of λM14, and 24 new cDNA clones extending towards the N terminus of the protein were selected. Isolation of the inserts of these 24 cDNA clones showed that three of them (λM23, λM42, and λM90) extend 2 kb 5′ of λM14 (Fig. 1 and not shown). λM23 is a full-length Mac-1 α cDNA clone. It encodes the protein N terminus (3) and the tryptic peptides not detected in λM14. λM14 and λM23 exhibit identical restriction maps in their overlapping regions.

Nucleotide Sequence—The composite cDNA sequence of λM14 and λM23 contains 4740 base pairs and has a long open reading frame of 3534 nucleotides, leaving a 5′-untranslated region of 72 base pairs and a 3′-untranslated region of 1.2 kb (Fig. 2). The 3′-untranslated region of λM14 contains an inverted stretch of poly(CA) (nucleotides 3667–3862) (33), a partial *Kpn*I interspersed repeat (nucleotides 4566–4631) (33), and ends with a stretch of more than 40 adenosines. There are two consensus polyadenylation signals at nucleotides 4191 and 4678. Analysis of the 3′-untranslated regions from λM14, λM23, and λM90 indicates that the first polyadenylation signal is used in λM23 and the second one in λM14 and λM90, indicating that both polyadenylation signals are functional. Restriction mapping of 15 additional cDNA clones suggests that both polyadenylation sites are used with equal frequency. The 3′-untranslated region of λM90 (and λM42, data not shown) lack 440 base pairs found between nucleotides 3629 and 4070 in λM14 and λM23 (Fig. 1). The sequences GAA/GTATCC and AAG/A at the boundaries of this deletion (*arrows*, Fig. 2) conform to the GT/AG rule for splicing sites (34), and thus the two different classes of cDNAs appear to correspond to alternatively spliced mRNAs.

Mac-1 α Subunit mRNA Expression—The cell surface expression of the Mac-1 α/β complex is almost exclusively restricted to cells of the myeloid lineage (5). Northern blots showed that the Mac-1 α subunit mRNA is 4.7 kb and is present in monocytes and myeloid cell lines, but not in T or B cell lines (Fig. 3, *A* and *B*).

Mac-1 expression is regulated during leukocyte differentiation and can be induced in myelomonocytic cell lines by culture with phorbol esters for 1–3 days (5). Northern blot analysis reveals that the steady-state level of Mac-1 α RNA in the HL-60 and U937 myelomonocytic cell lines is extremely low or nil (Fig. 3*B*, *lanes 2* and *4*) (5). PMA treatment of both cell lines induced expression of the Mac-1 α subunit mRNA (Fig. 3*B*, *lanes 1* and *3*) and increased the expression of the β subunit mRNA (Fig. 3*C*). These findings are concordant with previous studies on the biosynthesis of the Mac-1 α and β subunits (35), on the surface expression of the Mac-1 α/β complex on these cell lines (5), and suggest that the cell surface expression of Mac-1 is regulated by mRNA level. Similarly, we have previously shown that treatment of the murine premyelocytic cell line M1 with γ-interferon induces expression of the murine Mac-1 α subunit mRNA (26). Two Mac-1 α subunit messages of approximately 4.7 and 4.4 kb were resolved in the myelomonocytic cell lines HL-60 and U937 upon prolonged electrophoresis of the RNA (Fig. 3*B*). The presence of multiple Mac-1 α mRNA species may be the result of the alternative use of the two polyadenylation signals or alternative splicing. The size of the mRNA species is consistent with these possibilities. Southern blot analysis under stringent conditions on human DNA demonstrates that Mac-1 α is encoded by a single copy gene (36).

Amino Acid Sequence—The open reading frame from λM14/λM23 translates into a protein of 1137 residues, with a signal peptide of 16 amino acids defined by the previously reported Mac-1 α subunit N-terminal sequences (3, 37) (Fig. 2). In addition to the agreement with the protein N-terminal sequence, the 186 residues determined by tryptic peptide sequencing (Table I) agreed perfectly with the translated sequence. This confirmed the isolation of authentic Mac-1 α subunit cDNA clones. The amino acid sequence of the Mac-1 α subunit has the characteristics of a classical transmembrane protein, with an N-terminal 1092-residue domain, a 26-residue hydrophobic putative transmembrane domain, and a 19-residue C-terminal hydrophilic domain (Fig. 2). The presence in the N-terminal 1092-residue domain of 19 potential *N*-glycosylation sites (Asn-Xaa-Ser/Thr), one of which was sequenced in peptide 90 (residues 718–743), confirms that this is the extracellular domain. The predicted molecular mass of the protein is 125,611 daltons, consistent with previous estimations after *N*-glycanase treatment of the α subunit of Mac-1 (M_r 137,000) (35). Assuming M_r 2,500 per high mannose carbohydrate, M_r = 173,111 is predicted for the Mac-1 α subunit precursor, compared to the observed M_r = 160,000 (35). After carbohydrate processing, the Mac-1 α subunit is M_r 170,000.

The primary structure of the α subunit of Mac-1 suggests the presence of seven internal repeats (Fig. 4). Repeats V, VI, and VII show the highest degree of similarity to one another which is statistically significant ($p < 10^{-2}$ to $p < 10^{-4}$) and contain sequences similar to the divalent cation-binding EF-hand loop motif of proteins like calmodulin and parvalbumin (38) (Fig. 4), correlating with the divalent cation requirements of the Mac-1-mediated adhesion (27). Repeats I–IV lack the EF-hand loop-like sequences but contain the highly conserved sequences YFGAS/AL and LVTVGAP flanking the center of

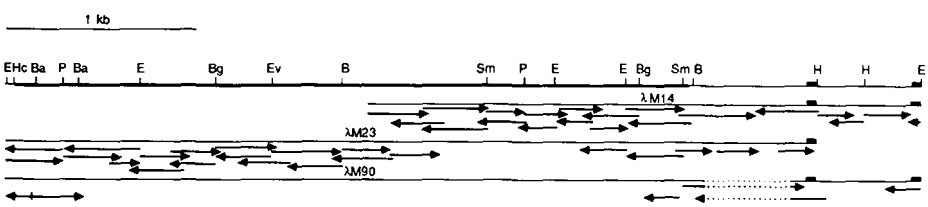

FIG. 1. **Restriction map of Mac-1 α cDNA clones and sequencing strategy.** The *thick line* represents the coding region of the cDNA. The polyadenylation signals are shown as *black boxes*. The region deleted in λM90 is indicated as a *dotted line*. The indicated restriction sites are *Bal*I (*B*), *Bam*HI (*Ba*), *Bgl*II (*Bg*), *Eco*RI (*E*), *Evo*RV (*Ev*), *Hinc*II (*Hc*), *Hin*dIII (*H*), *Sma*I (*Sm*), and *Pst*I (*P*).

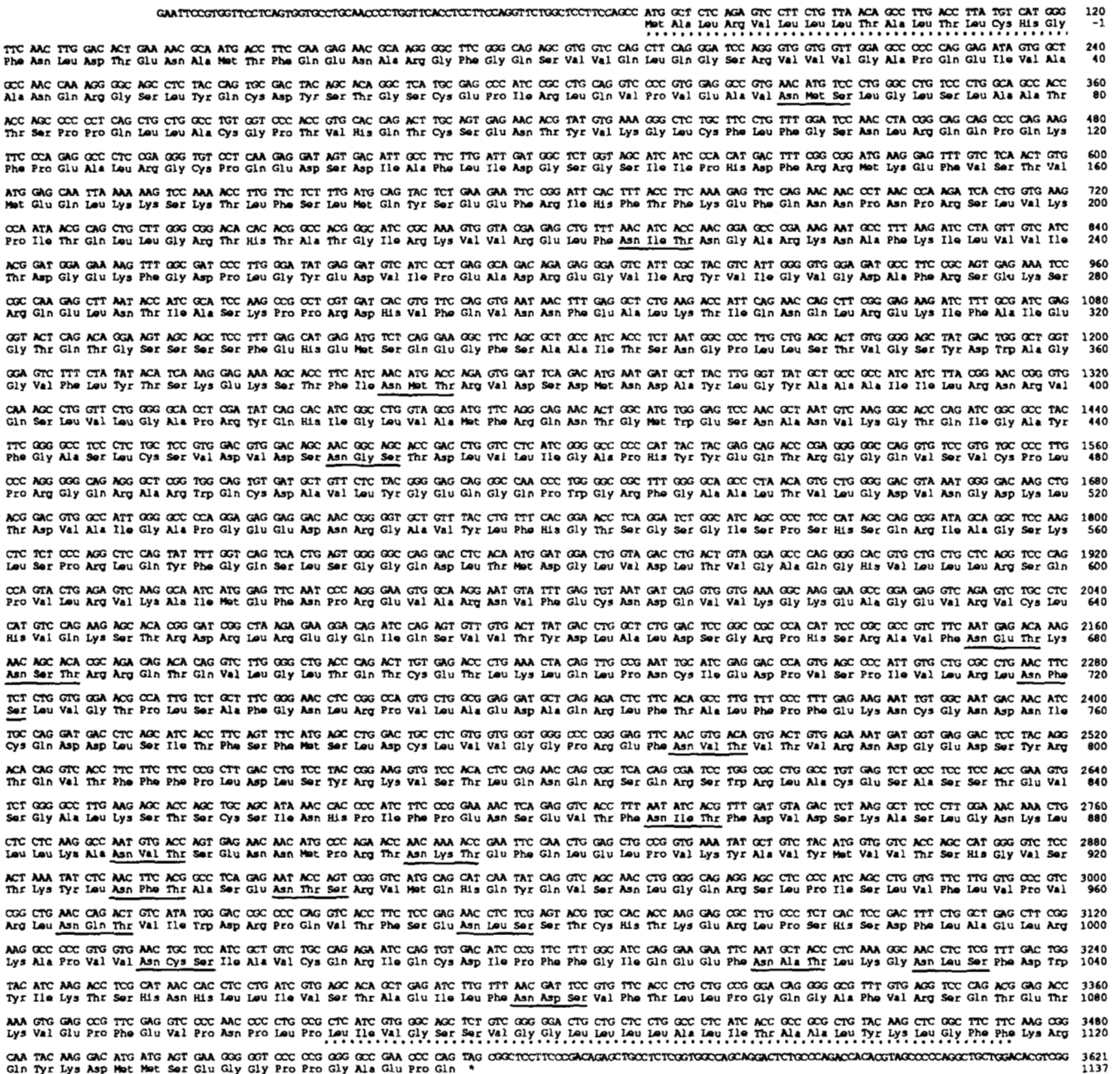

FIG. 2. **Nucleotide and cDNA-derived amino acid sequence.** The putative *N*-glycosylation sites are *underlined*. The signal sequence and the putative transmembrane region are indicated by a *dotted line*. The termination codon is labeled with an *asterisk*. *Boxes* in the 3′-untranslated region show the polyadenylation signal sequences. *Arrows* mark the boundaries of a potential intron spliced out of some of the cDNAs.

the repeats (Fig. 4). These consensus flanking sequences are also conserved in other integrins (Fig. 4). The presence of the seven repeats suggest that much of the N-terminal portion of the Mac-1 α subunit may have arisen by duplication events.

Mac-1 as a Member of the Integrin Gene Superfamily—The determination of the primary structure of the Mac-1/LFA-1/p150,95 common β subunit (7, 8) and the α subunit of p150,95 (9) has shown that the leukocyte adhesion receptors are evolutionary related to the extracellular matrix (ECM) receptors and led to the concept of a gene superfamily of cell-cell and cell-matrix receptors termed integrins (15). Three subfamilies of integrin molecules, each with a distinct β subunit, have been defined, namely the fibronectin receptor subfamily (sharing integrin β_1), the leukocyte adhesion recep-

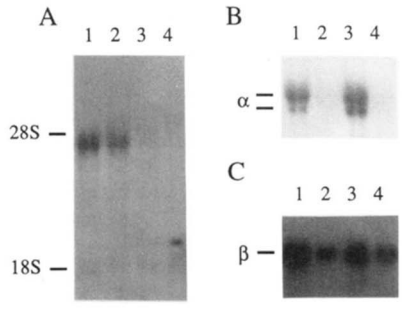

FIG. 3. **Expression of Mac-1 α subunit mRNA.** *A*, Northern blot analysis of total RNA from peripheral blood adherent cells (20 μg, *lane 1*; 10 μg, *lane 2*) and 20 μg of total RNA from SKW3 cells (T lineage) (*lane 3*) and JY cells (B lymphoblastoid) (*lane 4*). The position of the 28 S and 18 S ribosomal RNA markers are indicated. *B*, Northern blot performed on 20 μg of total RNA from HL-60 (*lanes 1* and *2*) and U937 (*lanes 3* and *4*), before (*lanes 2* and *4*) and after (*lanes 1* and *3*) treatment with PMA for 3 days. The gel was overrun to separate both mRNA species. *C*, same as in *B* after probing with the cDNA for the common β subunit. Similar amounts of RNA were present in each lane as shown by ethidium bromide staining.

tor subfamily (sharing integrin β_2), and the vitronectin receptor-IIb/IIIa subfamily (sharing integrin β_3).

Comparisons with the Mac-1 α subunit were undertaken to define the relationships among the α subunits of the different integrin subfamilies (Fig. 5). The α subunits of Mac-1 and p150,95 are 63% identical at the amino acid level and 68% identical at the nucleotide level. The high degree of structural similarity between the α subunits of Mac-1 and p150,95 is reflected at the functional level: Mac-1 and p150,95 exhibit iC3b-binding ability (16, 39), and both proteins are known to play a role in neutrophil aggregation and neutrophil and monocyte adhesion to endothelial cells (4, 18, 40). The α subunit of LFA-1 is 35% identical to the α subunits of Mac-1 and p150,95.[2] The α subunits of the fibronectin receptor, vitronectin receptor, and the glycoprotein IIb are 40% identical to one another. Since the α subunits of Mac-1 and p150,95 are 25% identical to the α subunits of the three ECM receptors, the α subunits of Mac-1, p150,95, and LFA-1[2] are more closely related to each other than to the rest of the integrin α subunits. The leukocyte α subunits also resemble one another in containing a segment of 187 residues not found in the ECM α subunits (amino acids 150–338 in the α subunit of Mac-1), and in lacking a region of 28 amino acids (gap at residue 1002 in Mac-1) where the ECM receptor α subunits are proteolytically cleaved during processing to generate two disulfide-linked chains (41).

Recently, we have located the Mac-1, LFA-1, and p150,95 α and β subunit genes by Southern blot on mouse x human somatic cell hybrids and by chromosomal *in situ* hybridization using cDNA probes (36). The genes encoding the α subunits of LFA-1, Mac-1, and p150,95 map to chromosome 16, between bands p11–p13.1, defining a gene cluster involved in leukocyte adhesion (36). The common structural characteristics and the close proximity of the three α subunit genes strongly suggest that the genes for the α subunits of Mac-1, p150,95, and LFA-1 evolved by gene duplication events and that these gene duplications took place after the divergence of the different integrin α subunit subfamilies.

The area of highest extended identity between Mac-1 α and the rest of the integrin α subunits lies between residues 434 and 592, precisely the boundaries of the three internal repeats containing putative divalent cation-binding sequences. Over this region, Mac-1 α shows 88% identity to p150,95 α at the amino acid level and 90% at the nucleotide level, and the percentage of identity to the ECM receptor integrin α subunits is 38%. The ECM receptor α subunits have four putative divalent cation-binding sites within their primary structure (12–14) which align with the repeats IV–VII in the α subunits of Mac-1 and p150,95; repeat IV in Mac-1 and p150,95 does not contain the putative divalent cation-binding sequence (Fig. 4). Ligand binding by the ECM receptor integrins is calcium-dependent (41) and in the case of the α subunit of the glycoprotein IIb/IIIa the binding of radioactive calcium has been verified (41). The high degree of conservation of these regions suggests a role in maintaining receptor conformation or a direct involvement in ligand binding.

An additional region of high conservation among the integrin α subunits are the membrane spanning regions. The transmembrane domain of Mac-1 α exhibits 88% identity with the one in p150,95 α and 40–50% identity with those of the IIb, VNR, and FNR α subunits (10, 13, 14) (Fig. 5). A similar conservation has been noted in the transmembrane and cytoplasmic domains of the different β subunits (7, 8, 10, 11, 13). Based on this high degree of conservation, it is conceivable that these regions may play a role in the regulation of the ligand binding (18, 24) through interactions with other membrane components or with the cytoskeleton. The finding that the fibronectin receptor binds talin (42), that LFA-1-mediated interactions are dependent on the integrity of the cytoskeleton (43), and that LFA-1 cocaps with talin after phorbol ester stimulation (44) further supports the involvement of these domains in cytoskeletal interactions and in signal transduction.

[2] R. S. Larson, A. L. Corbi, L. Berman, and T. A. Springer, manuscript submitted.

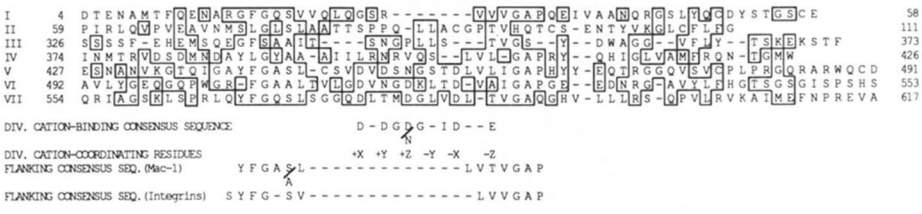

FIG. 4. **Mac-1 α homologous repeats.** Common residues are *boxed*. The three repeats containing putative divalent cation-binding sites (*V–VII*) are aligned with additional N-terminal related sequences lacking the putative Ca^{2+}- or Mg^{2+}-binding sites (*I–IV*). Based on the frequency of each residue in the seven repeats of the α subunits of Mac-1 and of other integrins (p150,95, the fibronectin receptor, the vitronectin receptor, and the platelet glycoprotein IIb) (9, 12–14), consensus sequences were derived for the regions flanking the putative divalent cation-binding sites. Consensus residues were defined as those appearing in at least 30% of the analyzed sequences. The consensus divalent cation-binding sequence and the coordination axes of the residues ligating the divalent cation are based on the sequence and the 3-dimensional structure of the Ca^{2+}- and Mg^{2+}-binding proteins parvalbumin, troponin C, and calmodulin (38) and are shown below the alignments of the seven repeats.

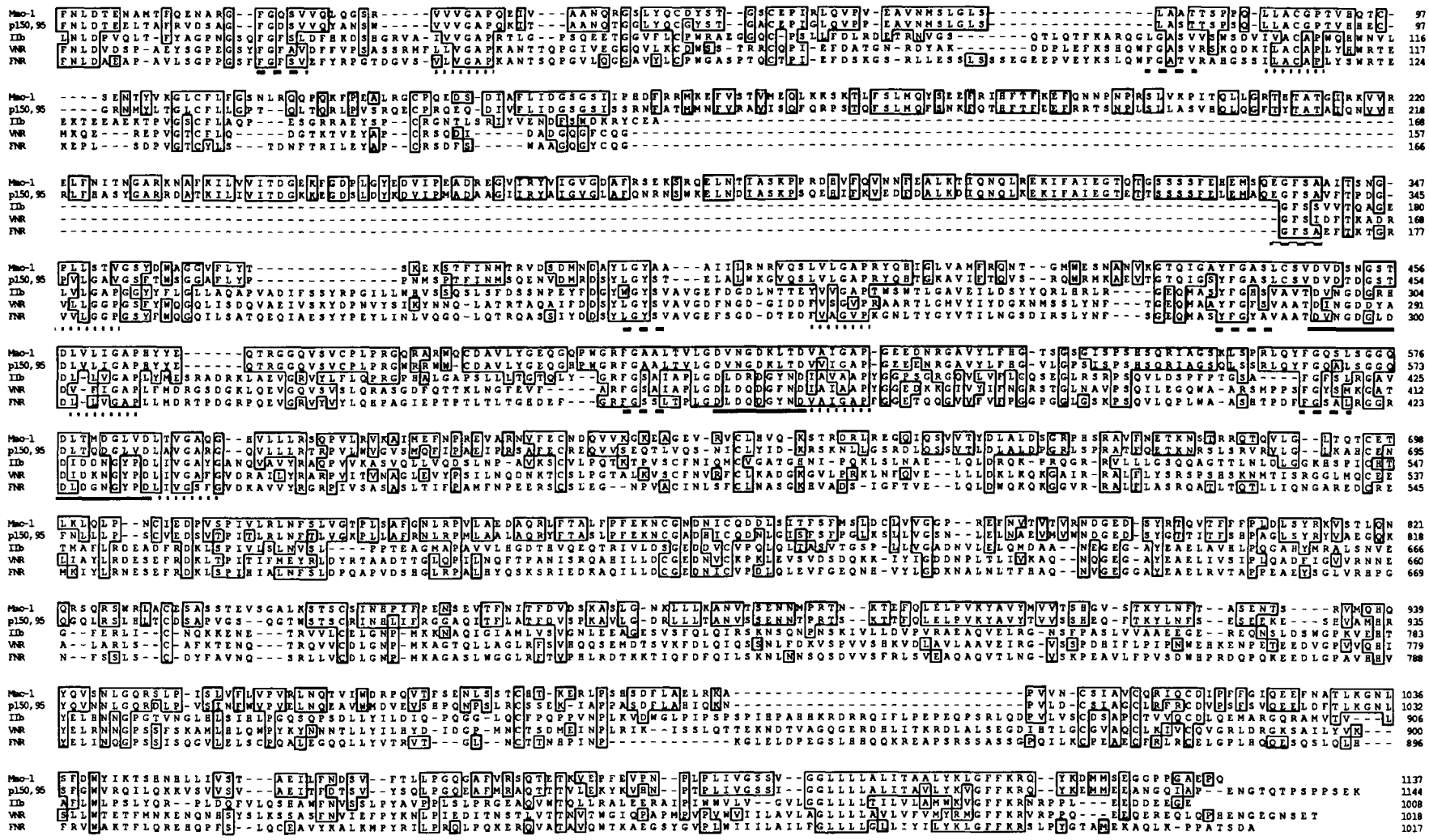

FIG. 5. **Comparison of the primary structure of the α subunits of Mac-1, p150,95, vitronectin receptor, fibronectin receptor, and glycoprotein IIb.** Identities between Mac-1 and the rest of the integrin α subunits are *boxed*. The putative divalent cation-binding sites and the conserved flanking sequences are *underlined* by *continuous* and *dotted lines*, respectively. The putative transmembrane regions are indicated by *TM*.

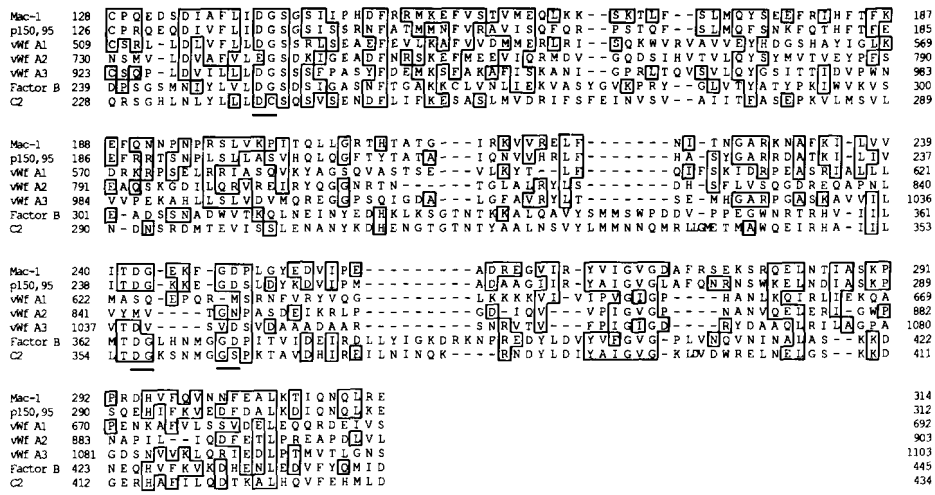

FIG. 6. **Comparison of the sequences of the leukocyte-specific domain of Mac-1 α and p150,95 α with the A repeats of von Willebrand factor, complement component C2, and factor B.** Common residues between the sequences of Mac-1 α and/or p150,95 and the rest of the proteins are *boxed*. The alignments were performed as described under "Experimental Procedures."

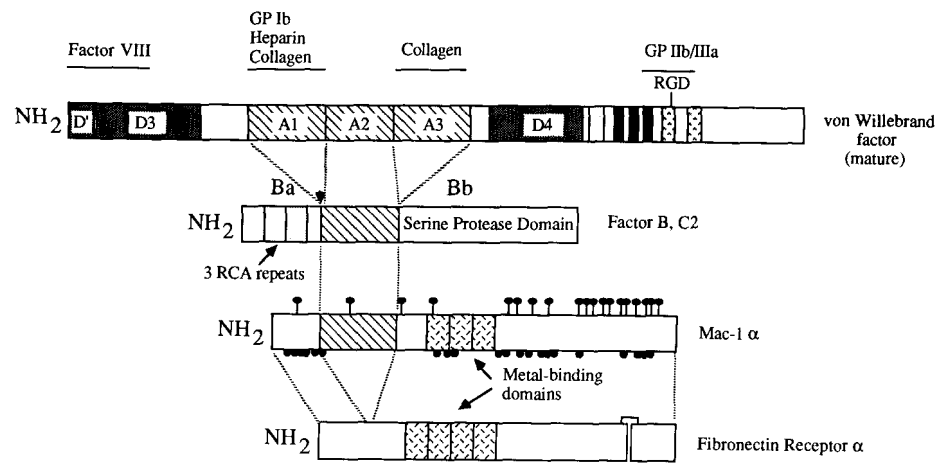

FIG. 7. **Evolutionary relationships of the Mac-1 α subunit, the A domains of von Willebrand factor, factor B, and C2.** The binding domains and the ligands for von Willebrand factor are indicated on top (45, 50). The structures of factor B and C2 are previously described (46, 47). Regulator of complement activation repeats (RCA) in factor B and C2 are 60 amino acids homologous repeats common to many complement components and receptors. The *N*-glycosylation sites and the cysteine residues in the α subunit of Mac-1 are shown on *top* and *bottom* of its schematic representation, respectively.

TABLE II

Significance of the alignment scores of the Mac-1 α subunit leukocyte-specific domain with the von Willebrand factor A repeats and factor B, and C2.

The sequences of the Mac-1 α subunit leukocyte-specific domain (128–314), the A repeats of von Willebrand factor (A1:509–692; A2:730–903; A3:923–1103), factor B (240–445), and C2 (228–434) were compared using the ALIGN program. The alignment scores are presented in standard deviation units. The probability that the alignments (shown in Fig. 6) have occurred by chance is shown between parentheses. The analysis was performed as described under "Experimental Procedures."

	vWf A1	vWf A2	vWf A3	Factor B	C2
Mac-1	7.4 ($<10^{-11}$)	7.2 ($<10^{-11}$)	8.1 ($<10^{-15}$)	6.9 ($<10^{-10}$)	1.8 ($<10^{-1}$)
vWf A1		10.3 ($<10^{-23}$)	9.1 ($<10^{-18}$)	7.0 ($<10^{-11}$)	4.6 ($<10^{-5}$)
vWf A2			13.4 ($<10^{-23}$)	4.0 ($<10^{-4}$)	4.1 ($<10^{-4}$)
vWf A3				4.8 ($<10^{-5}$)	3.2 ($<10^{-2}$)
Factor B					23.8 ($<10^{-23}$)

Homology with von Willebrand Factor and Factor B—As previously described for p150,95 (9), there is a region of 187 amino acids in the Mac-1 α subunit (residues 150–338) without counterpart in the three sequenced ECM receptor integrin α subunits. We will refer to this region as the L domain, because of its presence in leukocyte integrins. A search of the National Biomedical Research Foundation sequence data base with the FASTP program revealed homology of residues 128–314 of the Mac-1 α subunit with von Willebrand factor (vWF) (Fig. 6). The homologous regions are of particular interest because they correspond in Mac-1 almost exactly to the L domain and in vWF to the type A domain, a domain previously defined by its presence in three homologous tandem repeats of 200 residues denoted A1 (497–716), A2 (717–909), and A3 (910–1111) (45). The p150,95 α subunit (Fig. 5) and the murine Mac-1 α subunit[3] also show homology to the vWF A domains. Since the vWF A domains have previously been shown to be homologous to factor B (45, 46), a component of the alternative complement pathway which interacts with C3, we also examined homologies with factor B and with its homologue in the classical pathway, C2 (47) (Fig. 6). The homologous domain in factor B is clearly demarcated on the N-terminal side by the site at which factor B is cleaved to give the active Bb factor and on the C-terminal side by the serine protease domain (46, 47) (Fig. 7). Evaluation with the ALIGN program (32) showed that the homology of Mac-1 with vWF domains A1, A2, and A3 and with factor B is highly statistically significant (Table II) and thus that the homolo-

[3] R. Pytela, personal communication.

gous amino acid segments in each of these proteins must have evolved from a single primordial domain. Although the homology of Mac-1 with C2 is not statistically significant, factor B shows significant homology to C2 and serves as an evolutionary link, showing that the segment in C2 evolved from the same primordial domain.

A FASTP search of the SWISS-PROT database on Bionet revealed a further homology of the Mac-1 α subunit L domain with the chicken cartilage matrix protein (55). Chicken cartilage matrix protein is an extracellular matrix protein known to interact with cartilage proteoglycan and collagen and is composed of two homologous repeats (55). The homology between Mac-1 and chicken cartilage matrix protein is highly statistically significant (+10.6 S.D.), and the area of homology corresponds precisely to the L domains in Mac-1 and LFA-1[2] and to both homologous repeats in chicken cartilage matrix protein.

Given the homology between factor B and the L domain and the ability of factor B to bind C3b, it is tempting to speculate that the L domain is the iC3b ligand-binding site of Mac-1 and p150,95. It is also of interest that binding of Bb to C3b requires Mg^{+2} (48). Similarly, binding of isolated Mac-1 and p150,95 to iC3b-sensitized cells or to iC3b-Sepharose requires divalent cations (27, 39). A putative divalent cation binding site in the L domain may be represented by one sequence motif containing DG and another containing DG and GD which are conserved in Mac-1, p150,95, and factor B (*underlined* in Fig. 6). If these sites were contiguous in the 3-dimensional structure, they could form a divalent cation binding site similar in sequence to those present in internal repeats V–VII of the α subunit of Mac-1. The L domain and the following region containing repeats IV–VII are relatively free of *N*-linked glycosylation sites and cysteines (Fig. 7), allowing them to be accessible and conformationally flexible. The idea that the L domain could be involved in recognition of iC3b raises the possibility of recognition of a sequence in iC3b distinct from RGD. Although RGD is present in iC3b, evidence that it is important in recognition by Mac-1 (49) remains to be confirmed by peptide inhibition data.

Both the A1 and A3 domains of vWF bind collagen (50) and the A1 domain binds the platelet glycoprotein Ib and heparin (50). The alternative complement pathway in which factor B plays an important role is the most primitive mechanism of the immune system for distinguishing self from nonself (48). Thus, the 200 amino acid domain appears to be a primitive recognition unit which has been duplicated and embedded in a number of proteins which have evolved to play diverse recognition functions in hemostasis (vWF), extracellular matrix structure (chicken cartilage matrix protein), complement activation (factor B and C2), and complement receptor and cell-cell interactions (leukocyte integrins).

The ECM receptor integrins are phylogenetically ancient as shown at the level of sequence homology in *Drosophila* (41) and by immunological cross-reaction with proteins of similar size in nematodes and fungi (51). Our findings demonstrate that the Mac-1 α subunit evolved by the introduction of a primordial recognition domain into the ECM receptor-type of α subunit (Fig. 7). The introduction of the extra domain may have increased the potential for recognition of diverse ligands by the leukocyte integrins and may explain their somewhat different ligand specificity, since recognition of the ICAM-1 ligand by LFA-1 does not involve RGD (52–54).

The availability of cDNA clones for both the α and β subunits of Mac-1 will allow important questions concerning the structure and function of this receptor to be addressed. These include the definition of the distinct ligand binding sites involved in iC3b binding and in cell-cell adhesion and testing of the hypothesis that cell stimulation results in a conformational change in the ligand binding site, transmitted from the cytoplasmic or membrane domains, that alters affinity for ligand.

Note Added in Proof—The murine Mac-1α subunit (Pytela, R. (1988) *EMBO J.* **7,** 1371–1378) is 74% identical to the human Mac-1α subunit and shows similar homologies to other proteins.

Acknowledgments—We gratefully acknowledge Bill Lane for sequencing of the tryptic peptides and Lisa Bermann, Karen O'Connor, and Scott Thompson for their invaluable help.

REFERENCES

1. Anderson, D. C., and Springer, T. A. (1987) *Annu. Rev. Med.* **38,** 175–194
2. Sanchez-Madrid, F., Nagy, J., Robbins, E., Simon, P., and Springer, T. A. (1983) *J. Exp. Med.* **158,** 1785–1803
3. Miller, L. J., Wiebe, M., and Springer, T. A. (1987) *J. Immunol.* **138,** 2381–2382
4. Anderson, D. C., Miller, L. J., Schmalstieg, F. C., Rothlein, R., and Springer, T. A. (1986) *J. Immunol.* **137,** 15–27
5. Miller, L. J., Schwarting, R., and Springer, T. A. (1986) *J. Immunol.* **137,** 2891–2900
6. Hogg, N., Takacs, L., Palmer, D. G., Selvendran, Y., and Allen, C. (1986) *Eur. J. Immunol.* **16,** 240–248
7. Kishimoto, T. K., O'Connor, K., Lee, A., Roberts, T. M., and Springer, T. A. (1987) *Cell* **48,** 681–690
8. Law, S. K. A., Gagnon, J., Hildreth, J. E. K., Wells, C. E., Willis, A. C., and Wong, A. U. (1987) *EMBO J.* **6,** 915–919
9. Corbi, A. L., Miller, L. J., O'Connor, K., Larson, R. S., and Springer, T. A. (1987) *EMBO J.* **6,** 4023–4028
10. Tamkun, J. W., DeSimone, D. W., Fonda, D., Patel, R. S., Buck, C., Horwitz, A. R., and Hynes, R. O. (1986) *Cell* **46,** 271–282
11. Fitzgerald, L. A., Steiner, B., Rall, S. C., Jr., Lo, S., and Phillips, D. R. (1987) *J. Biol. Chem.* **262,** 3936–3939
12. Suzuki, S., Argraves, W. S., Arai, H., Languino, L. R., Pierschbacher, M. D., and Ruoslahti, E. (1987) *J. Biol. Chem.* **262,** 14080–14085
13. Argraves, W. S., Suzuki, S., Arai, H., Thompson, K., Pierschbacher, M. D., and Ruoslahti, E. (1987) *J. Cell Biol.* **105,** 1183–1190
14. 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
15. Hynes, R. O. (1987) *Cell* **48,** 549–554
16. Beller, D. I., Springer, T. A., and Schreiber, R. D. (1982) *J. Exp. Med.* **156,** 1000–1009
17. Dana, N., Styrt, B., Griffin, J. D., Todd, R. F., III, Klempner, M. S., and Arnaout, M. A. (1986) *J. Immunol.* **137,** 3259–3263
18. Vedder, N. B., and Harlan, J. M. (1988) *J. Clin. Invest.* **81,** 676–682
19. Rosen, H., and Gordon, S. (1987) *J. Exp. Med.* **166,** 1685–1701
20. Kishimoto, T. K., Hollander, N., Roberts, T. M., Anderson, D. C., and Springer, T. A. (1987) *Cell* **50,** 193–202
21. Todd, R. F., III, Arnaout, M. A., Rosin, R. E., Crowley, C. A., Peters, W. A., and Babior, B. M. (1984) *J. Clin. Invest.* **74,** 1280–1290
22. Miller, L. J., Bainton, D. F., Borregaard, N., and Springer, T. A. (1987) *J. Clin. Invest.* **80,** 535–544
23. Detmers, P. A., Wright, S. D., Olsen, E., Kimball, B., and Cohn, Z. A. (1987) *J. Cell Biol.* **105,** 1137–1145
24. Buyon, J. P., Abramson, S. B., Phillips, M., Slade, S. G., Ross, G. D., Weissmann, G., and Winchester, R. J. (1988) *J. Immunol.* **140,** 3156–3160
25. Springer, T. A., Teplow, D. B., and Dreyer, W. J. (1985) *Nature* **314,** 540–542
26. Sastre, L., Roman, J. M., Teplow, D. B., Dreyer, W. J., Gee, C. E., Larson, R. S., Roberts, T. M., and Springer, T. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83,** 5644–5648
27. Wright, S. D., Rao, P. F., Van Voorhis, W. C., Craigmyle, L. S., Iida, K., Talle, M. A., Westberg, E. F., Goldstein, G., and Silverstein, S. C. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 5699–5703
28. Schneider, C., Newman, R. A., Sutherland, D. R., Asser, U., and

Greaves, M. F. (1982) *J. Biol. Chem.* **257,** 10766–10769
29. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
30. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
31. Henikoff, S. (1984) *Gene (Amst.)* **28,** 351–359
32. Dayhoff, M. O., Barker, W. C., and Hunt, L. T. (1983) *Methods Enzymol.* **91,** 524–545
33. Sun, L., Paulson, K. E., Schmid, C. W., Kadyk, L., and Leinwand, L. (1984) *Nucleic Acids Res.* **12,** 2669–2690
34. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472
35. Miller, L. J., and Springer, T. A. (1987) *J. Immunol.* **139,** 842–847
36. Corbí, A. L., Larson, R. S., Kishimoto, T. K., Springer, T. S., and Morton, C. C. (1988) *J. Exp. Med.* **167,** 1597–1607
37. Pierce, M. W., Remold-O'Donnell, E., Todd, R. F., III, and Arnaout, M. A. (1986) *Biochim. Biophys. Acta* **87,** 368–371
38. Szebenyi, D. M. E., Obendorf, S. K., and Moffat, K. (1981) *Nature* **294,** 327–332
39. Micklem, K. J., and Sim, R. B. (1985) *Biochem. J.* **231,** 233–236
40. Te Velde, A. A., Keizer, G. D., and Figdor, C. G. (1987) *Immunology* **61,** 261–267
41. Ruoslahti, E., and Pierschbacher, M. D. (1987) *Science* **238,** 491–497
42. Horwitz, A., Duggan, K., Buck, C., Beckerie, M. C., and Burridge, K. (1986) *Nature* **320,** 531–533
43. Springer, T. A., Dustin, M. L., Kishimoto, T. K., and Marlin, D. (1987) *Annu. Rev. Immunol.* **5,** 223–252
44. Burn, P., Kupfer, A., and Singer, J. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 497–501
45. Shelton-Inloes, B. B., Titani, K., and Sadler, J. E. (1986) *Biochemistry* **25,** 3164–3171
46. Mole, J. E., Anderson, J. K., Davison, E. A., and Woods, D. E. (1984) *J. Biol. Chem.* **259,** 3407–3412
47. Bentley, D. R. (1986) *Biochem. J.* **239,** 339–345
48. Muller-Eberhard, H. J., and Schreiber, R. D. (1980) *Adv. Immunol.* **29,** 1–53
49. Wright, S. D., Reddy, P. A., Jong, M. T. C., and Erickson, B. W. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 1965–1968
50. Girma, J.-P., Meyer, D., Verweij, C. L., Pannekoek, H., and Sixma, J. J. (1987) *Blood* **70,** 605–611
51. Marcantonio, E. E., and Hynes, R. O. (1988) *J. Cell Biol.* **106,** 1765–1772
52. Marlin, S. D., and Springer, T. A. (1987) *Cell* **51,** 813–819
53. Staunton, D. E., Marlin, S. D., Stratowa, C., Dustin, M. L., and Springer, T. A. (1988) *Cell* **52,** 925–933
54. Simmons, D., Makgoba, M. W., and Seed, B. (1988) *Nature* **331,** 624–627
55. Argraves, W. S., Deak, F., Sparks, K. J., Kiss, I., and Goetinck, P. F. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 464–468