

# Amino Acid Sequence of the Alpha Subunit of Human Leukocyte Adhesion Receptor Mol (Complement Receptor Type 3)

M. Amin Arnaout,\* Shalley K. Gupta,‡ Mark W. Pierce,§ and Daniel G. Tenen‡

\*Renal Unit, The Children's Hospital, §Diabetes Unit, Massachusetts General Hospital; †Hematology Division, Beth Israel Hospital, and Harvard Medical School, Boston, Massachusetts 02115

**Abstract.** Mol (complement receptor type 3, CR3; CD11b/CD18) is an adhesion-promoting human leukocyte surface membrane heterodimer (alpha subunit 155 kD [CD11b] noncovalently linked to a beta subunit of 95 kD [CD18]). The complete amino acid sequence deduced from cDNA of the human alpha subunit is reported. The protein consists of 1,136 amino acids with a long amino-terminal extracytoplasmic domain, a 26-amino acid hydrophobic transmembrane segment, and a 19-carboxyl-terminal cytoplasmic domain. The extracytoplasmic region has three putative Ca<sup>2+</sup>-binding domains with good homology and one with weak homology to the "lock washer" Ca<sup>2+</sup>-binding consensus

sequence. These metal-binding domains explain the divalent cation-dependent functions mediated by Mol. The alpha subunit is highly homologous to the alpha subunit of leukocyte p150,95 and to a lesser extent, to the alpha subunit of other "integrin" receptors such as fibronectin, vitronectin, and platelet IIb/IIIa receptors in humans and position-specific antigen-2 (PS2) in *Drosophila*. Mol<sub>α</sub>, like p150, contains a unique 187-amino acid stretch NH<sub>2</sub>-terminal to the metal-binding domains. This region could be involved in some of the specific functions mediated by these leukocyte glycoproteins.

THE ability of human granulocytes and monocytes to adhere in a divalent cation-dependent manner to particles coated with the activated complement C3 fragment, iC3b, is mediated by a surface glycoprotein heterodimer, Mol (complement receptor type 3 [CR3])<sup>1</sup> (Arnaout et al., 1983; Dana et al., 1984). Mol also mediates other divalent cation-dependent granulocyte adhesion functions such as binding to activated endothelium (Wallis et al., 1985), homotypic cell adhesion (Arnaout et al., 1985), phagocytosis (Arnaout et al., 1983), and chemotaxis (Beatty et al., 1983; Dana et al., 1986). Mol consists of two noncovalently associated subunits, each the product of a single gene (Kishimoto et al., 1987; Law et al., 1987; Arnaout et al., 1988). The alpha subunit has an apparent molecular mass of 155–165 kD (CD11b) and associates in a divalent cation-independent manner with a beta subunit of 95 kD (CD18) (Todd et al., 1982). The beta subunit is common to two other leukocyte surface glycoproteins, LFA-1 (alpha subunit 177 kD) and p150,95 (alpha subunit 130–150 kD) (Trowbridge and Omary, 1981; Sanchez-Madrid et al., 1983; Lanier et al., 1985). In addition to sharing the property of binding to

the same beta subunit, Mol, LFA-1, and p150,95 (leukocyte adhesion molecules [Leu-CAM]) require divalent cations to mediate their adhesion functions, have homologous NH<sub>2</sub> termini (Pierce et al., 1986; Miller et al., 1987), and presumably arose by gene duplication events (Arnaout et al., 1988).

One mechanism by which Mol and p150,95 mediate enhanced adhesiveness of activated phagocytes is through increased expression of these glycoproteins on the surface of activated cells. In granulocytes this occurs primarily by translocation of these antigens from intracellular storage pools present in secondary and tertiary granules (Arnaout et al., 1984) to the cell surface. The biologic importance of Leu-CAM is underscored by the finding that their inherited deficiency in humans impairs leukocyte adhesion-dependent inflammatory functions and predisposes to life-threatening bacterial infections (Dana et al., 1984). Leu-CAM are also members of a larger family of cell adhesion heterodimeric receptors "integrins", which include platelet IIb/IIIa, fibronectin and vitronectin receptors in humans, and position-specific adhesion receptors in *Drosophila* (Argaves et al., 1987; Bogaert et al., 1987; Corbi et al., 1987; Hynes, 1987; Kishimoto et al., 1987; Law et al., 1987; Fitzgerald et al., 1987; Poncz et al., 1987; Arnaout et al., 1988).

Recombinant DNA cloning studies revealed significant homologies among the beta subunits of Leu-CAM, fibronectin receptor, and platelet IIb/IIIa. Additional studies, includ-

M. A. Arnaout's present address is Renal Unit, Massachusetts General Hospital, Fruit Street, Boston, MA 02114, where reprint requests should be addressed.

1. *Abbreviations used in this paper:* CR3, complement receptor type 3; TM, transmembrane.

tggcttctgtggttctcagtggtgctgcaacccctgggtcactcctcc  
 aggttctggctcctccagccatggctctcagagctcctctgtaacagccttgacctatgcatggg  
 N A L R V L L L T A L T L C H G  
 ttcaactggacactgaaacgcaatgacctccaagagaacgcaagggtctcgggagagcgtggtc -1  
 F N L D T E N A M T F Q E N A R G F G Q S V V 23  
 cagcttcaggatccaggtgggtggttggagcccccagagagatggtggtgccaaccaaggggcagc  
 Q L Q G S R V V V G A P Q E I V A A N Q R G S 46  
 ctctacagtgagactacagacaggtcctcagcagccatccgctcagaggtcccgctggagccgtg  
 L Y Q C D Y S T G S C E P I R L Q V P V E A V 69  
 aqatgctccctgggctcctcctggcagccaccagccccctcagctggtggctggtgcccacc  
 N\* M S L G L S L A A T T S P P Q L L A C G P T 92  
 gtgcaccagactgagcagcagctatgtgaaaggctcgtcctcctggttggatccaactcgg  
 V H Q T C S E N T Y V K G L C F L F G S N L R 115  
 cagcagccccagaagtccacagagccctccgaggggtcctcctcaagaggatagtgacctgctctctg  
 Q Q P Q K F P E A L R G C P Q E D S D I A F L 138  
 attgatggtcctggtgagcatccacatgacttccggggatgaggaggtttgtctcaactgtgatg  
 I D G S G S I I P H D F R R M K E F V S T V M 161  
 gagcaatataaaagtcctcaaaccttctctcttggatgagctactcgaagaattccggattcacttt  
 E Q L K K S K T L F S L M Q Y S E E F R I H F 184  
 acctcaaaagtgccagaaacacctcaaacagatcactggtgagcaataacagcagctggtggg  
 T F K E F Q N N P N P R S L V K P I T Q L L G 207  
 cggacacacagccagggcctccgcaaggtgacagagctgcttaagcatccaacagcagccgca  
 R T H T A T C I R K V V R E L F N\* I T N G A R 230  
 aagaatgctttaagatcctagttgcatcagcaggtggagaaaagtggcgatccctgggatagag  
 K N A F K I L V V I T D G E K F G D P L G Y E 253  
 gatgtcctcctgagcagcagagagggagtcctcagctcagctcattgggggggagatgctcctccg  
 D V I P E A D R E G V I R Y V I G V G D A F R 276  
 agtgagaaaatcccgcaagagcttaataccatcagcagccgctcgtgacagctgctccaggtg  
 S E K S R Q E L N T I A S K P P R D H V F Q V 299  
 aataactttgagctcgaagaccattcagaaccagctcgggagaagatccttgagcagaggtact  
 N N F E A L K T I Q N Q L R E K I F A I E G T 322  
 cagacaggaagtagcagctcctttgagcagatgctcaggaaggctcagcagcctcagcctcct  
 Q T G S S S S F E H E M S Q E G F S A A I T S 345  
 aatggcccttctgagcagctggtgggagctatgactggcgtgggagctcttctataacacaaag  
 N G P L L S T V G S Y D W A G G V F L Y T S K 368  
 gagaagaacaccttcaacatgaccagagtgattcagacatgaatgatgcttacttgggtatgct  
 E K S T F I N\* M T R V D S D M N D A Y L G Y A 391  
 gccgcatcctctcaggaacccgggtcaaaagcctggttctgggggacctcagatcagcagatcggc  
 A A I I L R N R V Q S L V L G A P R Y Q H I G 414  
 ctggtagcagatgctcagcagaaactggatggtggaggtccaacgctaagtccaaggcaccagatc  
 L V A M F R Q N T G M W E S N A N V K G T Q I 437  
 ggcgctactcgggctccctcctcctcctggcagctggcagcagcagcagcagcagcagcagcagc  
 G A Y F G A S L C S V D V D S N\* G S T D L V L 460  
 atcgggccccctactactacagcagcagccagggggccaggtgctcggctgctcccttggccagggg  
 I G A P H Y E Q T R G G Q V S V C P L P R G 483  
 agggcctgggtggcagtgatgctgctcctcaggggagcagggccaacctggggcggcttggggca  
 R A R W Q C D A V L Y G E Q G Q P W G R F G A 506  
 gccctaacagtgctggggagcgaatggggacaagctgacggagctggccattggggccccagagag  
 A L T V L G D V N G D K L T D V A I G A P G E 529  
 gaggaacaacgggggtgctgttctcctgtttcaggaacccctcaggtatcggcagcagccctccatg  
 E D N R G A V Y L F H G T S G S G I S P S H S 552  
 cagcggatagcaggtcccaagctcctcagcagctccagattttggctcagctcagtgagtgggggcag  
 Q R I A G S K L S P R L Q Y F G Q S L S G G Q 575  
 gacctcaaatggatggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 D L T M D G L V D L T V G A Q G H V L L L R S 598  
 cagccagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 Q P V L R V K A I M E F N P R E V A R N V F E 621  
 tgaatgacagtggtggaaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 C N D Q V V K G K E A G E V R V C L H V Q K S 644  
 acacgggatcggcagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 T R D R L R E G Q I Q S V V T Y D L A L D S G 667  
 cgcccaatcccgccgctcctcagtgagacaagaacagcagcagcagcagcagcagcagcagcagcagc  
 R P H S R A V F N\* E T K N\* S T R R R Q T Q V L G 690  
 ctgaccagactgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 L T Q T C E T L K L Q L P N C I E D P V S P I 713  
 gtgctcggctgagctcctcctcctggtgggaacccattgctgcttccgggaalccrpgvtgctg  
 V L R L N\* F L V G T P L S A F F G N L R P V L 736  
 gggagagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 A E D A Q R L F T A L F P F E K N C G N D N I 759  
 tggcagatgacctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 C Q D D L S I T F S F M S L D C L V V G G P R 782  
 gactcagc  
 E F N\* V T V T V R N D G E D S Y R T Q V T F F 805  
 tccogcttggctcctaccggaggtgtccacactccagaaccagcagcagcagcagcagcagcagc  
 F P L D L S Y R K V S T L Q N Q R S Q R S W R 828  
 ctggcctgtagctcctcctccacagaggtctggggccttgaagagcaccagctgagcagcagcagc  
 L A C E S A S S T E V S G A L K S T S C S I N 851  
 caccocatctcccgaaaactcagaggtcacttgaatcagcttggatgagcctcagagctcc  
 H P I F P E N S E V T F N\* I T F D V D S K A S 874  
 cttggaaacaaactgctcctcaaggccaagtgtgaccagtgagaacaacatgcccagaaccaaacaac  
 L G N K L L L K A N\* V T S E N N M P R T N\* K T 897  
 gaattccaactggagctcgggtgaaatgctgctacatggtggtcaccagcagcagcagcagcagc  
 E F Q L E L P V K Y A V Y M V V T S H G V S T 920  
 aaatctcaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 K Y L N\* F T A S E N\* T S R V M Q H Q Y Q V S N 943  
 ctggggcagagggagccccccatcagcagctggtgcttctgggtccggctcagcagcagcagcagc  
 L G Q R S P P I S L V F L V P V R L N\* Q T V I 966  
 tgggagccccagctcctcctcaggaacctcagcagcagcagcagcagcagcagcagcagcagcagc  
 W D R P Q V T F S E N\* L S S T C H T K E R L P 989  
 tctcactccgacttctggctgagcctcgggaagcccccgggtgagctcctcagcagcagcagcagc  
 S H S D F L A E L R K A P V V N\* C S I A V C Q 1012  
 agaactcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 R I Q C D I P F F G I Q E E F N\* A T L K G N\* L 1035  
 tctgttggctgacatacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 S F D W Y I K T S H N H L L I V S T A E I L F 1058  
 aacgattccgctgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 N D S V F T L P F G Q G A F V R S Q T E T K V 1081  
 gagcagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 E P F E V P N P L P L I V G S S V G G L L L L 1104  
 gccctcagc  
 A L I T A A L T K L G F F K R Q Y K D M H S E 1127  
 gggggtccccggggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc  
 G G P P G A E P Q . 1136

Figure 1. cDNA nucleotide sequence and deduced amino acid sequence of the open reading frame of human Mol<sub>α</sub>.<sup>2</sup> Peptide sequences (including NH<sub>2</sub>-terminal sequence) determined from the purified antigen are underlined. Potential N-linked glycosylation sites (\*) are indicated. The overlined regions represent, respectively, the 5' hydrophobic leader sequence and the putative transmembrane hydrophobic domain. Three putative metal-binding domains and a fourth weak site are underlined by broken lines. The 378-bp Mol<sub>α</sub> clone used to isolate the rest of the cDNA, corresponds to nucleotides 2,823–3,201 in the full-length Mol<sub>α</sub> cDNA. The 126-amino acid peptide encoding this region differs by one amino acid (P<sup>949</sup>→L) from the deduced amino acid sequence derived from another cDNA library (Arnaout et al., 1988). This could represent a polymorphism in the alpha subunit. Cysteine residues are circled.

2. These sequence data have been submitted to the EMBL/GenBank Data Libraries under the accession numbers X07421 and M18044, respectively.

ing isolation of a partial human Mol $\alpha$  cDNA clone (Arnaout et al., 1988), characterization of the alpha subunit of p150,95 (Corbi et al., 1987) and of *Drosophila* PS-2 antigen (Bogaert et al., 1987), confirmed the structural relatedness and overall homology of the alpha subunits of integrins. The specific functions mediated by these adhesion receptors will be better understood once the complete amino acid sequence of each of these molecules is elucidated. In this report, we describe the complete primary structure of the alpha subunit of human Mol $\alpha$  and elucidate its homologous as well as its distinctive structural features in relation to the other alpha subunits in the integrin supergene family.

## Materials and Methods

### Protein Purification and Sequencing

Human Mol $\alpha$  was purified from granulocytes by affinity chromatography using lentil-lectin Sepharose and anti-Mol $\alpha$  monoclonal antibody-linked Sepharose as previously described (Pierce et al., 1986; Arnaout et al., 1987). Mol $\alpha$ -derived tryptic peptides were separated by reverse-phase HPLC on a C18 column and eluted with a gradient of acetonitrile in 0.1% TFA (Pierce et al., 1986). Individual peptides were subjected to microsequencing on gas phase sequencer (Applied Biosystems Inc., Foster City, CA).

### Screening a Human Phorbol-Ester-induced HL-60 cDNA Library

A human cDNA library constructed from poly(A)<sup>+</sup> messenger RNA derived from PMA-differentiated promyelocytic cell line HL-60 was a generous gift from Drs. Brian Seed and David Simmons, Massachusetts General Hospital, Boston, MA. This library was constructed in the Bst XI-cloning site of plasmid  $\pi$  H3M and plated after transfection into *Escherichia coli* MC1061/p3 host strain (Seed and Aruffo, 1987). 1-million colonies plated on 150-mm plates containing LB-agar, 7.5  $\mu$ g/ml ampicillin, and 12.5  $\mu$ g/ml tetracycline were screened according to Maniatis and co-workers (Maniatis et al., 1982). A human Mol $\alpha$  cDNA probe, 378 bp long (Arnaout et al., 1988) and several synthesized oligonucleotides (17-mers) were random hexanucleotide-<sup>32</sup>P labeled or 5'-end <sup>32</sup>P labeled, respectively, and used to probe nitrocellulose replicates using standard hybridization and washing conditions. Positive colonies were purified and plasmid DNA isolated as described (Maniatis et al., 1982). The dideoxy method for DNA sequencing (Sanger et al., 1977) was carried out using plasmid DNA or cDNA fragments cloned into M13 vectors and, respectively, plasmid vector- or M13-specific primers as well as primers derived from the generated sequences. Both strands of the isolated cDNA inserts were sequenced. Oligonucleotides were synthesized using a 380 DNA synthesizer (Applied Biosystems Inc.).

### Sequence Analysis and Homologies

Sequence data analysis was performed using BIONET System of Intelligenetics, Inc. (Palo Alto, CA). Sequence alignment was done using ALIGN and GENALIGN programs from BIONET. Protein homology was sought using the PIR data base and sequence similarities determined by using the FASTP program of Lipman and Pearson (1985).

## Results and Discussion

A 378-bp human Mol $\alpha$  cDNA clone (Arnaout et al., 1988), was used to isolate three additional cDNA clones that together contained the 3,408-base nucleotide coding sequence of Mol $\alpha$  gene (Fig. 1). The coding sequence is preceded by a single translation initiation methionine. The translation product of the single open reading frame began with a 16-amino acid hydrophobic peptide representing a leader sequence, followed by the NH<sub>2</sub>-terminal phenylalanine residue (Pierce et al., 1986). The translation product also contained all eight

tryptic peptides isolated from the purified antigen, the amino-terminal peptide, a 26-amino acid hydrophobic domain representing a potential transmembrane (TM) region, and a short 19-amino acid carboxyl-terminal cytoplasmic domain (Fig. 1). The coding region of Mol $\alpha$  (1,136 amino acids) is eight amino acids shorter than the alpha subunit of p150,95 (1,144 amino acids) (Corbi et al., 1987). This suggests that the larger apparent molecular mass of Mol $\alpha$  (155–165 kD by PAGE versus 130–150 kD for the alpha subunit of p150,95) is largely due to the greater number of N-glycosylation sites (18 potential sites in Mol $\alpha$  versus 10 in p150). The cytoplasmic region of Mol $\alpha$  contains one serine residue that could serve as a potential phosphorylation site. The cytoplasmic region is also relatively rich in acidic residues and in proline (Fig. 1). It is interesting to note that the cytoplasmic domain of endocytic receptors (Goldstein et al., 1985) and some integral lysosomal membrane receptors (e.g., cation-dependent and cation-independent mannose-6-phosphate receptors [Dahms et al., 1987; Lobel et al., 1987]) are also rich in acidic residues and in proline. Since Mol is involved in the process of phagocytosis and is also targeted to intracellular storage pools (Arnaout et al., 1984), these residues may be important in mediating these functions. The TM region is highly homologous to the TM region in other integrins (Fig. 2) and could therefore serve additional common functions such as in the formation or stabilization of the  $\alpha/\beta$  complex. The long extracytoplasmic amino-terminal region has several interesting features. First, it contains four putative metal-binding domains (outlined by broken lines in Fig. 1) that are similar to Ca<sup>2+</sup>-binding sites (consensus sequence: DXDXDGXXDXXE) found in calmodulin (Williams, 1986), intestinal vitamin D-binding protein (Fullmer and Wasserman, 1981), parvalbumin (Coffee and Bradshaw, 1973), myosin light-chain (Reinach et al., 1986), troponin (Kretsinger, 1976), thrombospondin (Lawler and Hynes, 1986) with two exceptions: (a) the predicted secondary structure (Chou and Fasman, 1978) shows that the metal-binding sites are surrounded by beta structure rather than by alpha helical segments (not shown); (b) the invariant glutamyl residue at position 12 of the metal-binding domains of EF-hand-loop proteins is not present in Mol $\alpha$  domains or in metal-binding sites of other integrins. This finding suggests that each Mol $\alpha$ -metal-binding site may be composed of two noncontiguous peptide segments (lock-washer configuration) (Vyas et al., 1987). The two peptide segments could lie within the receptor protein itself, as occurs in galactose-binding protein (Vyas et al., 1987) or between the receptor and a natural ligand, accounting for some of the metal-dependent adhesion-promoting functions of Leu-CAM and the cation-dependent formation of  $\alpha/\beta$  complex in other integrins. Secondly, the extracytoplasmic region contains a unique 187-amino acid sequence (beginning at R<sup>151</sup>) which is not present in the homologous alpha subunits of fibronectin, vitronectin, or platelet IIb/IIIa receptors (Fig. 2) or in PS-2 antigen in *Drosophila* (Bogaert et al., 1987). Interestingly, this sequence is present in the highly homologous alpha subunit of leukocyte p150,95 with 57% of the amino acids identical and 34% representing conserved substitutions (Fig. 2). It is known that both Mol and p150,95 have a binding site for complement fragment iC3b (CR3 activity) (Arnaout et al., 1983; Micklem et al., 1985). It is tempting to speculate that this unique region is involved in iC3b binding. Even

```

Mola FNLDTENAMITFOENARG--FGGSVVQLQGR---VVVGAFOQIV--AANQGLVYCCDYST--GSCB
p150 FNLDTTELTAEVRDASG--FGGSVVQVQANW---VVVGAFOQIV--AANQGLVYCCDYST--GACE
FNRa FNLDTAAP-AVLGGPPSPSFFGSEVEFYRPGTDGVS-VVVGAKANTSQPGVTLGGAVLCPWASPTQIT
VNRa FNLNDVDSI-AEYSGPPSPSFGFVAVDFVHASSRMFLVVGAKANTQPGIVEGVOVLKCTWSS--TRRD
IIB FNLDPVQLT-FYAGPPSPSFGSGLDFHKDGHGRVA-IVVGARPTLG--PSQETGQVFLCPWRAEGGQD
*
Mola PTLQVHVEAVNM--SLGLS-----IDNTSPPO-LLACGPTVHTG-----SENVVVG
p150 PTLQVHVEAVNM--SLGLS-----IASITSPQ-LLACGPTVHHEG-----GRMMLLIG
FNRa PI-EPDSKGS-RLLGSLSSSEEGEPEVYKSLQWFGATVRAHGSSTLACAPLYSWRTEKEPL--SDPVG
VNRa PI-EFDATGN-RDYAR-----DDPLEFKSHQWFGASVRSKQOKLLACAPLYSWRTEKQY--REPVG
IIB FSLIDFDLRDSTRNVCS-----QLTQTFKARQGLGASVVSWSVIVVACAEVQELHWVLEKTEAEKTPVQ
*
Mola LCFILFCSNLRQPKPFPEARLCPQESLDTAP-LIDGSGSLPHDFRHKLKEFVSTVMEQLKSKKMLFSLM
p150 LCFILFPTQLTRLRPSRQE--CFRQEQDIVE-LIDGSGSLSSRNFAIMNFVRAVISQFORPSTQESLM
FNRa TCMIS-TDNFTRILEYAP--CRSDHS--WAAGQCY--CQG-----
VNRa TCFIQ--DGTKTVEYAP--CRS-QDI--DADQCF--CQG-----
IIB SCFLAQP-ESGRRAEYSP--CRGNTLSRIYVENLHPSWDKRYCEA-----
*
Mola QISEFTRHITFKKFNQNNHRSILVKPITDILGHTATHTKRVRELNITNGARKNAKIKIIVITDGE
p150 QISNKEQHTIFTEERRTNSISLILASVHLDGFTMTATNIVVHRLPHASYGARRATKILIVITDGE
FNRa -----
VNRa -----
IIB -----
*
Mola KCFDHLKEDVITPADRECMIRYVIGVGLARFSEKSRQELMIFASKPRIMVIVMNFALKNTQNLKE
p150 KFGDHLKEDVITPADRECMIRYVIGVGLAFQNRNSWELNDIASKFSRITRIVVFDLALKNTQNLKE
FNRa -----
VNRa -----
IIB -----
*
Mola KIFATEGTLGSSSSFEHMSQEGFSAALISNG-PLLSIVGSLVWGGVFLMI-----SGC
p150 KIFATEGTLGSSSSFEHMSQEGFSAVFPDQ-FLVLAUGSEFWGCGALVME-----PNM
FNRa -----GFSAEFTKCRVVLGGPSPFVWCGDILSADIQEQAIESYYPEYLINLV
VNRa -----GFSIDFKARVVLGGPSPFVWCGDILSADVAEIVSKYDPNVYSIVY
IIB -----GFSVVVTOAEFELVLAGAPSEVYFVGLAQAPVADIFSSYRPGIILHWV
*
Mola KSTFINMTRVLSDMNDAVLYGVAADILRN--VQSLVLCAPRYOHL-GVAMAFQNT-GMSESNAVNG
p150 SHFINMGOENVDMSYLYGVELALWKG--VQSLVLCAPRYOHL-GKAVLITQVS-RQRMKNAVNG
FNRa OGO-LQTRQASLYVDSYLVAVGFEFSGD-DTEDFVAGVFNGLTYGVVTLNGSDIRSLYNF--SG
VNRa NNQ-LATRTAQEIFDYSYLVAVGDFNGD-GIDDFVSCPAARTLGMVYLDGKMSLSLYNF--TG
IIB SSSLSFDSSNPEYFVDSYLVAVGFEFDGDLNTEYVCAFTWSWTIGAVELDSYQRLHRLR--G
*
Mola TQIAYFGASLCSVDVDSGSDLVLLIGAPHYE----QTRGGQVSVCP-PRGTRHICDAVLYGEG
p150 TQIAYFGASLCSVDVDSGSDLVLLIGAPHYE----QTRGGQVSVCP-PRGTRHICDAVLYGEG
FNRa EQMASYFCYVAATDMDGGLDL-IMGALIMDRTPDGRGQEVVVMVLOHPLGLEPPTLTLTGHE
VNRa EQMASYFCYVAATDMDGGLDL-IMGALIMDRTPDGRGQEVVVMVLOHPLGLEPPTLTLTGHE
IIB EQMASYFCYVAATDMDGGRDL-IMGALIMDRTPDGRGQEVVVMVLOHPLGLEPPTLTLTGHE
*
Mola QPWGRFGAALTVLGVDNGDKLTDVAIGAP-GEENRGAAYLPHG-TSGCISPSHSQRIAGSKLSRIG
p150 QPWGRFGAALTVLGVDNGDKLTDVAIGAP-GEENRGAAYLPHG-VIQFISPSHSQRIAGSKLSRIG
FNRa F--GRFCSLTHGLLDLQGVNVAIGAHRGPTQCVFVFPQGGKSPQVQLPWA-ASHTPD
VNRa F--RRFCSLTHGLLDLQGVNVAIGAHRGPTQCVFVFPQGGKSPQVQLPWA-ARMMP
IIB LY-GRFCSLTHGLLDLQGVNVAIGAHRGPTQCVFVFPQGGKSPQVQLPWA-ARMMP
*
Mola YFGQSLGGQDLTMDGLVDLIVGAGG-HVLLL-RSOPVIVVADLKEFNRVAVRVEFCNIVVVLKAG
p150 YFGQSLGGQDLTMDGLVDLIVGAGG-HVLLL-RIRPVIVVADLKEFNRVAVRVEFCNIVVVLKAG
FNRa YFGQSLGGQDLTMDGLVDLIVGAGG-HVLLL-RIRPVIVVADLKEFNRVAVRVEFCNIVVVLKAG
VNRa YFGQSLGGQDLTMDGLVDLIVGAGG-HVLLL-RIRPVIVVADLKEFNRVAVRVEFCNIVVVLKAG
IIB YFGQSLGGQDLTMDGLVDLIVGAGG-HVLLL-RIRPVIVVADLKEFNRVAVRVEFCNIVVVLKAG
*
Mola EV-NYCIHVQ-RSTRDHTREGQISVVITMDLALDSCRPHRAVEMTKNSRRTOTVIG-LPQICETLK
p150 QS-NYCIYID-RRSKNILGSRDLSVITMDLALDSCRPHRAVEMTKNSRRTOTVIG-LKAKCENFN
FNRa INLSFCINASGVAES-IGFTVE--LQIDRQRKQCVR-RALFLASRQMTLQDILLONGAREDEMK
VNRa FNVKCCIKADQGVLELQVLEKQVVE--LLLDLKLQKCAIR-RALFLYSRSPSHSKNMTISRGGLMOCHE
IIB FNIQVIGATGHNL-PKLSNLAEE--LQIDRQRK-PROGR-BVLLGSGQAQTLNLNLDGKHSPIQHTM
*
Mola LILP-NQIEDVSPVILRNFSLVGLPISAFNLRPVLAEADAORLFTALMFFERKCNQNDICODDLSL
p150 LILP-NQIEDVSPVILRNFSLVGLPISAFNLRPVLAEADAORLFTALMFFERKCNQNDICODDLSL
FNRa LILP-NQIEDVSPVILRNFSLVGLPISAFNLRPVLAEADAORLFTALMFFERKCNQNDICODDLSL
VNRa LILP-NQIEDVSPVILRNFSLVGLPISAFNLRPVLAEADAORLFTALMFFERKCNQNDICODDLSL
IIB LILP-NQIEDVSPVILRNFSLVGLPISAFNLRPVLAEADAORLFTALMFFERKCNQNDICODDLSL
*
Mola FSPMSLGLVGGP-HEVNYVIMHNDGSDS-RRTQVTFHFLDLSYRMSFTIIONGSRWRNDCESAS
p150 FSPMSLGLVGGP-HEVNYVIMHNDGSDS-HGTTITRSHFAGLSYRMAEGQKQQLRSLHLHDCSA-
FNRa VFGQNHVYLDGKNAELNLFPAHQ-NVEEGAEAEALRVTAPEAEVGLVRRHGNFESIS--CDYFA
VNRa VSDQKTYIGDDNPLTLKKAQ-NVEEG-AVEAELIYSHPQADFGVNRNEALARIS--CAFKT
IIB AVTGSPLVYADNVLEQMDAA-NVEEG-AVEAELAVHLEGGAHVTRALSNVGFEKLI--GNQKK
*
Mola STEVSAALKSTSCSINHPIFENSEVTRNITFDVSKASLANKLLKANVISENNMPTNKTFPOLELPV
p150 PVGSGTW-SLTSQVHFLFEGGACITPLATFDVSHKAVLQDRLELIVANVSENNIPRTKTFPOLELPV
FNRa ---NQ-SRLLVLDLGNL-MKAGASLWGLLRPLPHLRDTRKTKTQDFQIILSNLNSGSDVVSFRLSVE
VNRa ---ENQ-TRQVLDLGNL-MKAGTQLLAGLRSVHQOSEMDTSVKFDLQSSNLFDKVSPVSHKVDLA
IIB ---ENE-TRVVLDELGNL-MKNAQIGIAMLVSNLEAEGESVSPQLQIRSNLQSNPNSEIVLLDVPVR
*
Mola KYA-VYVNVVSH-VSTKYLNFTAS-ENT-----SHVMQRQYQVSNLQGRSHF-ISIMLVVPLNIV
p150 KYA-VYVNVVSHQFKYLNFTASEEKE-----SHVAMHRYQVNNLQGRDLP-VSINFWVPELVNCEA
FNRa AQAVMLNGVSKPEAVLFPVSDWHPDQPKQEDLGPVHVMVLELNGPSSLSQCVLDSQDALEGQQ
VNRa VLAQVIRGVSSPDHIFLPIPNWHEKENPETEEDVGFVHCHIVELRNGPSSFSKAMLHLOWPKYVNTT
IIB AEPQVLRGNSFPASVVAEEGE--REQNSLDSWGPVCHIVELRNGPCTVNGLHLSIHLGOSQPSD
*
Mola IADRPQVTSNLSSTHTKERL-FHSDFLAELRKA-----PVVN
p150 VMDVESHQVDFSLRSEKIA-FPASDFLAHIQKN-----PVLD
FNRa LLYVTRVLI-GL-NCTNHPIME-----KGLELDPEGLHHQOKREAPSSASSGQTLK
VNRa LLYIHLVD-IDGF-MNCTSDMEINFLRIK--ISSLQTEKNDTVAGQGERDHLTKRDLALSEGDIHTLG
IIB LLYIILDIQ-POGG-LQFQPPVNLKVDMLGLPLTPSPPIHPAHKRRRQIFLPEQPSRQLQDPVILS
*
Mola CSTAVLQRIQCDI-FMFCIQEENATLKGNSLPLVYIKTSHNHLIVSIA--ETLFDNMTL-LPGQ
p150 CSTAVLQRIQCDI-FESVQVQELDFTLKGNSLPLVYVRLQKQVSVVSM--EINFDNMTL-LPGQ
FNRa CPBACFPLRLELPHLQ-DEFQSILH--FRVAKTFLQREHOPFS--LQCEAVYKALMYPYRILPRQ
VNRa CVAQLKIVDQVGRDRKSAIIVYK--SLIMLETFMKNENQNHVYSLKSSASFNWIDPEKMLPIE
IIB CSAPVYVQCDLQEMARQAMVTV--LAEFLMLSPVLYQR--PLQDFVLSHAWNNSLPLVAVHLS
*
Mola GAFVRSQTHVPPFVFNPLPLIVGSSV--GCLLLALITALIKLGFQKQ--YKMMSE-GE-H
p150 EAFVRSQTHVPPFVFNPLPLIVGSSV--GCLLLALITALIKLGFQKQ--YKMMSE-GE-H
FNRa LPQKERVAVVQVQVKAESYGVPLIILAILLCLLLLILIKLGFQKQSLPYGTAM-KKAQLK
VNRa DITNSTLVITMVTNGICGAFVFNVVIILAVLACILLAVLVFVYVYKGFQKVRPPO--EEDEREQL
IIB LPRGEAVVQDILLRALEBRALVWVVLV--GVLCGLLLILVILVAVYKGFQKVRPPL--EEDDEGE
*
Mola FCAEPQ
p150 F--ENGTQTPSPPEK
FNRa FATSDA
VNRa QPHENEGNSSET
IIB F

```

Figure 2. Alignment of the alpha subunits of human integrins Mol<sub>a</sub>, p150,95 (Corbi et al., 1987), fibronectin receptor (Argaves et al., 1987), vitronectin receptor (Suzuki et al., 1987), and IIB (Poncz et al., 1987). Identical amino acids are boxed. The 187-amino acid sequence unique to leukocyte p150 and Mol<sub>a</sub> is apparent. The proteolytic cleavage sites occurring in the alpha subunits of fibronectin receptor, vitronectin receptor, and IIB are indicated (arrowheads).

more intriguing is the possibility that this region may be involved in some heterotypic or homotypic adhesive interactions mediated by Leu-CAM. This is suggested by the fact that this region of Mol $\alpha$  (residues 150–337) has significant homology (17.1% identity and 52.9% conserved substitutions) to the collagen/heparin/platelet GpI binding region (A domain) (residues 530–713) of the mature von Willebrand factor (Verweij et al., 1986; Girma et al., 1987). Thirdly, the region just NH<sub>2</sub>-terminal to the TM site contains the consensus (CXXXXCXXXXC) that is conserved in all alpha chains of integrins (Fig. 2 and Bogaert et al., 1987). In other integrins, it is proposed that these cysteines are involved in the covalent binding of the heavy and light chains of the alpha subunit that are generated by proteolytic cleavage at usually dibasic sites just NH<sub>2</sub>-terminal to this sequence. It is interesting to note that no such cleavage occurs in the alpha subunit of Mol despite occurrence of a dibasic cleavage site at amino acid positions 999 and 1,000. The large deletion introduced in this region to maximize homology to other integrins (Fig. 2) suggests that differences in secondary structure may explain lack of susceptibility of Mol (and other Leu-CAM) to proteolytic cleavage. This region (Q<sup>900</sup>-E<sup>1025</sup>) also contains an epitope—recognized by a monoclonal antibody (M26) directed against guinea pig Mol $\alpha$ , which binds to precursor or mature monomeric Mol $\alpha$  but does not bind to the alpha subunit after formation of the  $\alpha/\beta$  complex (Remold-O'Donnell and Savage, 1988; Arnaout et al., 1988). This suggests that the M26 epitope becomes inaccessible to the antibody in the  $\alpha/\beta$  complex or is involved in its formation. Fourthly, internal homology blots (not shown) (Dayhoff et al., 1983) revealed four internal tandem repeats (amino acid residues: 358–412; 426–483; 487–553; 554–614), each containing one of the four putative metal-binding sites and their flanking regions. This suggests that these regions of the alpha subunit might have arisen by tandem duplication events.

Future studies using site-directed mutagenesis should further elucidate the structural features that underlie the specific functions mediated by Mol and the other evolutionary related integrins. Given the recent findings that anti-Mol $\alpha$  monoclonal antibodies block leukoaggregation (Arnaout et al., 1985) and reduce the size of myocardial infarct in dogs by 50% (Simpson et al., 1988), a detailed structure-function analysis of Mol may provide useful reagents for controlling disease processes characterized by neutrophil-mediated tissue injury.

We thank Drs. Brian Seed and David Simmons for providing the HL60 library and Linda Clayton for helpful discussions.

This research was supported by National Institute of Health Grants AI 21963, CA 41456, and a March of Dimes grant No. 1010. D. G. Tenen is a Fellow of the Leukemia Society of America. Part of this work was done during Dr. Arnaout's tenureship as an Established Investigator of the American Heart Association.

Received for publication 15 March 1988, and in revised form 28 March 1988.

## References

- 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.
- Arnaout, M. A., R. M. Hakim, R. F. Todd, N. Dana, and H. R. Colten. 1985. Increased expression of an adhesion-promoting surface glycoprotein in the granulocytopenia of hemodialysis. *N. Engl. J. Med.* 312:457–462.
- Arnaout, M. A., M. W. Pierce, N. Dana, and L. C. Clayton. 1987. Complement receptor type 3 (CR3): structure and function. *Methods Enzymol.* 150:602–615.
- Arnaout, M. A., E. Remold O'Donnell, M. W. Pierce, P. Harris, and D. G. Tenen. 1988. Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein. Mol: chromosomal localization and homology to the alpha subunits of Integrins. *Proc. Natl. Acad. Sci. USA.* 85: 2776–2780.
- Arnaout, M. A., H. Spitz, C. Terhorst, J. Pitt, and R. F. Todd. 1984. Deficiency of a surface glycoprotein (LFA-1) in two patients with Mol deficiency. *J. Clin. Invest.* 74:1291–1300.
- Arnaout, M. A., R. F. Todd, N. Dana, J. Melamed, S. F. Schlossman, and H. R. Colten. 1983. Inhibition of phagocytosis of complement C3- or immunoglobulin G-coated particles and of C3bi binding by monoclonal antibodies to a monocyte-granulocyte membrane glycoprotein (Mol). *J. Clin. Invest.* 72:171–179.
- Beatty, P. G., J. A. Ledbetter, P. J. Martin, T. H. Price, and J. A. Hansen. 1983. Definition of a common leukocyte cell-surface antigen (Lp95,150) associated with diverse cell-mediated immune functions. *J. Immunol.* 131: 2913–2919.
- Bogaert, T., N. Brown, and M. Wilcox. 1987. The Drosophila PS2 antigen is an invertebrate integrin that, like the fibronectin receptor, becomes localized to muscle attachments. *Cell.* 51:929–940.
- Chou, P. Y., and G. D. Fasman. 1978. Predictions of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* 47:25–148.
- Coffee, C. J., and R. A. Bradshaw. 1973. Carp muscle calcium-binding protein. *J. Biol. Chem.* 248:3305–3312.
- 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 subunit of a leukocyte adhesion glycoprotein, P150, 95. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:4023–4028.
- Dahms, N. M., P. Lobel, J. Breitmeyer, J. M. Chirgwin, and S. Kornfeld. 1987. 45 kd mannose 6-phosphate receptor: cloning, expression and homology to the 215 kd mannose 6-phosphate receptor. *Cell.* 50:181–192.
- Dana, N., B. Styrst, G. D. Griffin, R. F. Todd, III, M. S. Klempner, and M. A. Arnaout. 1986. Two functional domains in the phagocyte membrane glycoprotein Mol identified with monoclonal antibodies. *J. Immunol.* 137:3259–3263.
- Dana, N., R. F. Todd, T. Pitt, T. A. Springer, and M. A. Arnaout. 1984. Deficiency of a surface membrane glycoprotein (Mol) in man. *J. Clin. Invest.* 73:153–159.
- Dayhoff, M. O., W. C. Barker, and L. T. Hunt. 1983. Establishing homologies in protein sequences. *Methods Enzymol.* 91:524–545.
- Fitzgerald, L. A., B. Steiner, S. C. Rall, S. S. Lo, and D. R. Philips. 1987. Protein sequence of endothelial glycoprotein IIIa derived from a cDNA clone. *J. Biol. Chem.* 262:3936–3939.
- Fullmer, C. S., and R. H. Wasserman. 1981. The amino acid sequence of bovine intestinal calcium-binding protein. *J. Biol. Chem.* 256:5669–5674.
- Girma, J.-P., D. Meyer, C. L. Verweij, H. Pannekoek, and J. J. Sixma. 1987. Structure-function relationship of human von Willebrand factor. *Blood.* 70:605–611.
- Goldstein, J. L., M. S. Brown, R. G. W. Anderson, D. W. Russell, and W. J. Schneider. 1985. Receptor mediated endocytosis: concepts emerging from the LDL receptor system. *Annu. Rev. Cell Biol.* 1:1–39.
- Hynes, R. O. 1987. Integrins: a family of cell surface receptors. *Cell.* 48: 549–554.
- Kishimoto, T. K., K. O'Conner, A. Lee, T. M. Roberts, and T. A. Springer. 1987. Cloning of the beta subunit of the leukocyte adhesion proteins: homology to extracellular matrix receptor defines a novel supergene family. *Cell.* 48:681–690.
- Kretsinger, R. H. 1976. Structure and evolution of calcium modulated proteins. *Annu. Rev. Biochem.* 45:239–266.
- Lanier, L. L., M. A. Arnaout, R. Schwarting, N. L. Warner, and G. D. Ross. 1985. P150/95, third member of the LFA-1/CR3 polypeptide family identified by anti-Leu-M5 monoclonal antibody. *Eur. J. Immunol.* 15:713–718.
- Law, S. K. A., J. Gagnon, J. E. K. Hildreth, C. E. Wells, A. C. Willis, and A. J. Wong. 1987. The primary structure of the beta subunit of the cell surface adhesion glycoproteins LFA-1, CR3 and P150,95 and its relationship to the fibronectin receptor. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:915–919.
- Lawler, J., and R. O. Hynes. 1986. The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins. *J. Cell Biol.* 103:1635–1648.
- Lipman, D. J., and W. R. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science (Wash. DC).* 227:1435–1439.
- Lobel, P., N. M. Dahms, J. Breitmeyer, J. M. Chirgwin, and S. Kornfeld. 1987. Cloning of the bovine 215-dKa cation-independent mannose 6-phosphate receptor. *Proc. Natl. Acad. Sci. USA.* 84:2233–2237.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 545 pp.
- Mickle, K. J., and R. B. Sim. 1985. Ligand binding by the p150,95 antigen of U937 monocytic cells: properties in common with complement receptor type (CR3). *Biochem. J.* 231:233–236.
- Miller, L. J., M. Wiebe, and T. A. Springer. 1987. Alpha subunit N-terminal sequences of human Mac-1 and P150,95 leukocyte adhesion proteins. *J. Im-*

- munol.* 138:2381-2383.
- Pierce, M. W., E. Remold-O'Donnell, R. F. Todd, and M. A. Arnaout. 1986. N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species and homology to platelet IIb/IIIa. *Biochem. Biophys. Acta.* 874:368-371.
- Poncz, M., R. Eisman, R. Heidenreich, S. M. Silver, G. Vilaire, S. Surrey, E. Schwartz, and J. S. Bennet. 1987. Structure of the platelet membrane glycoprotein IIb. *J. Biol. Chem.* 262:8476-8482.
- Reinach, F. C., K. Nagai, and J. Kendrick-Jones. 1986. Site-directed mutagenesis of the regulatory light-chain  $Ca^{2+}/Mg^{2+}$  binding site and its role in hybrid myosins. *Nature (Lond.)*. 322:80-83.
- Remold-O'Donnell, E., and B. Savage. 1988. Characterization of macrophage adhesion molecule (MAM). *Biochem. J.* 27:37-41.
- Sanchez-Madrid, F., J. A. Nagy, E. Robbins, P. A. Simon, and T. A. Springer. 1983. A human leukocyte differentiation antigen family with distinct alpha subunits and a common beta subunit. *J. Exp. Med.* 158:1785-1799.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463-5467.
- Seed, B., and A. Aruffo. 1987. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. USA.* 84:3365-3369.
- Simpson, P. J., R. F. Todd, III, J. C. Fantone, J. K. Mickelson, J. D. Griffin, and B. R. Lucchesi. 1988. Reduction of experimental myocardial perfusion injury by a monoclonal antibody (anti-Mol, anti-CD11b) that inhibits leukocyte adhesion. *J. Clin. Invest.* 81:624-630.
- Suzuki, S. W. S. Argraves, H. Arai, L. R. Languino, M. D. Pierschbacher, and E. Ruoslahti. 1987. Amino acid sequence of the bitronectin receptor a subunit and comparative expression of adhesion receptor mRNAs. *J. Biol. Chem.* 262:14080-14085.
- Todd, R. R., A. van Aghoven, S. F. Schlossman, and C. Terhorst. 1982. Structural analysis of leukocyte differentiation antigens Mol and Mo2 on human monocytes. *Hybridoma.* 1:329-337.
- Trowbridge, I. S., and M. B. Omary. 1981. Molecular complexity of leukocyte surface glycoproteins related to the macrophage differentiation antigen Mac-1. *J. Exp. Med.* 154:1517-1524.
- Verweij, C. L., P. J. Diergaarde, M. Hart, and H. Pannekoek. 1986. Full length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein considerably larger than the mature vWF subunit. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:1839-1847.
- Vyas, N. K., M. N. Vyas, and F. A. Quioco. 1987. A novel calcium binding site in the galactose-binding protein of bacterial transport and chemotaxis. *Nature (Lond.)*. 327:635-638.
- Wallis, W. J., P. G. Beatty, H. D. Ochs, and J. M. Harlan. 1985. Human monocyte adherence to cultured endothelium: monoclonal antibody-defined mechanisms. *J. Immunol.* 135:2323-2330.
- Williams, R. J. P. 1986. Calcium and the cell. *Ciba Found. Symp.* 122:145-161.