The EMBO Journal vol.7 no.5 pp.1371 – 1378, 1988

Amino acid sequence of the murine Mac-1 α chain reveals homology with the integrin family and an additional domain related to von Willebrand factor

Robert Pytela

Basel Institute for Immunology, Grenzacherstrasse 487, Basel, Switzerland

Communicated by K.Simons

Clones encoding the Mac-1 α chain were selected from a mouse macrophage cDNA library by screening with oligonucleotide probes based on the sequence of a genomic clone encoding the N-terminus of the mature protein. The sequence of overlapping clones (4282 nt) was determined and translated into a protein of 1137 amino acids and a signal peptide of 15 amino acids. The Mac-1 sequence was found to be related to the α chain sequences of three other members of the integrin family of cell adhesion receptors, i.e. the fibroblast receptors for fibronectin and vitronectin and the platelet glycoprotein IIb/IIIa. All four sequences share a number of structural features, like the position of 13 cysteine residues, a transmembrane domain near the C-terminus and the location of three putative binding sites for divalent cations. Furthermore, a conserved sequence motif is repeated seven times in the N-terminal half of the molecule and three of these repeats include putative Ca/Mg-binding sites of the general structure DXDXDGXXD. On the other hand, Mac-1 contains a unique domain of 220 amino acids inserted into the N-terminal part of the integrin structure. This additional domain is homologous to a repeated domain found in von Willebrand factor, cartilage matrix protein and in the complement factors B and C2. In two of these proteins, the homologous domain is likely to be involved in binding to collagen fibrils. Therefore, Mac-1 may also bind to collagen, which could play a role in the interaction of leukocytes with the subendothelial matrix.

Key words: calcium-binding sites/cartilage matrix protein/complement factor B/collagen-binding sites/integrins

Introduction

Mac-1 is a cell surface glycoprotein of monocytes, macrophages and granulocytes which has been implicated in various adhesive interactions of these cells as well as in mediating the uptake of complement-coated particles (reviewed by Anderson and Springer, 1987). It is identical with CR-3, the receptor for the iC3b fragment of the third complement component (Beller *et al.*, 1982). The receptor protein is composed of two non-covalently associated polypeptide chains (α chain, M_r = 170 000 and β chain, M_r = 95 000). Two other heterodimeric leukocyte surface proteins, LFA-1 and gp 150/95, consist of the same 95 kd β chain in association with distinct α chains of 180 and 150 kd, respectively (Sanchez-Madrid *et al.*, 1983). Even though the three different α chains are not cross-reactive immunologically, they are clearly the products of homologous genes since their N-terminal sequences are closely related (Springer *et al.*, 1985; Miller *et al.*, 1987). Recently, it has been recognized (Suzuki *et al.*, 1986) that this family of leukocyte receptors is part of a superfamily of adhesive cell surface proteins that includes the receptors for fibronectin and vitronectin (Pytela *et al.*, 1985a,b), platelet gp IIb/IIIa (Pytela *et al.*, 1986), the VLA family (Takada *et al.*, 1987), and the position-specific antigens of *Drosophila* (Leptin *et al.*, 1987). The collective name 'integrins' has been proposed for the proteins encoded by this multigene family (Hynes, 1987). All the integrins that have been characterized are heterodimeric proteins with α chains ranging from 135 to 210 kd and β chains from 95 to 130 kd. In the same way that the leukocyte adhesion receptor subfamily is defined by a shared 95 kd β subunit, two other subfamilies are defined by shared β chains of 130 (FNR, VLA-1 to VLA-5) or 110 kd (VNR, gp IIb/IIIa) (Ginsberg *et al.*, 1987). Complete protein sequences deduced from cDNA have been published for all three β chains (Kishimoto *et al.*, 1987; Law *et al.*, 1987; Tamkun *et al.*, 1986; Argraves *et al.*, 1987a; Fitzgerald *et al.*, 1987), and the α chains of FNR, VNR and gp IIb/IIIa (Argraves *et al.*, 1986, 1987a; Suzuki *et al.*, 1986, 1987; Poncz *et al.*, 1987). It was found that all β chains share a high degree of identity (40–50%) with each other, and that the α chains form a separate family of homologous polypeptides that is apparently not related to the β chain family (reviewd by Ruoslahti and Pierschbacher, 1987). Both α and β chains are transmembrane polypeptides with a short (25–50 amino acids) C-terminal cytoplasmic domain. Evidence for interaction with the cytoskeleton (Horwitz *et al.*, 1986) and phosphorylation on tyrosine (Hirst *et al.*, 1986) has been obtained. Therefore, it has been proposed that integrins are generally involved in the formation of adhesive complexes between the cytoskeleton and extracellular components.

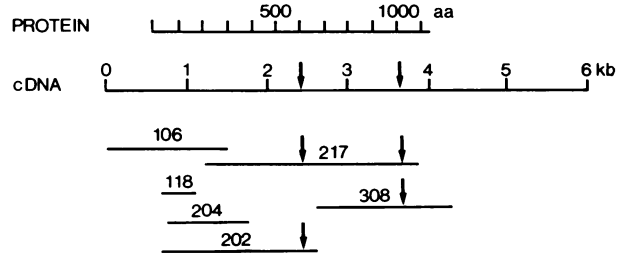

Fig. 1. Overlapping clones used for sequence analysis. Clones 106 and 118 were obtained in the initial screening, clones 217, 204 and 202 were selected by screening the same library with a 30mer oligonucleotide from the 3′ end of clone 106, and clone 308 by screening with a 17mer from the 3′ portion of clone 217. The size of the murine Mac-1 mRNA as determined by Northern blotting is 6 kb (Sastre *et al.*, 1986). Arrows denote *Eco*RI cleavage sites.

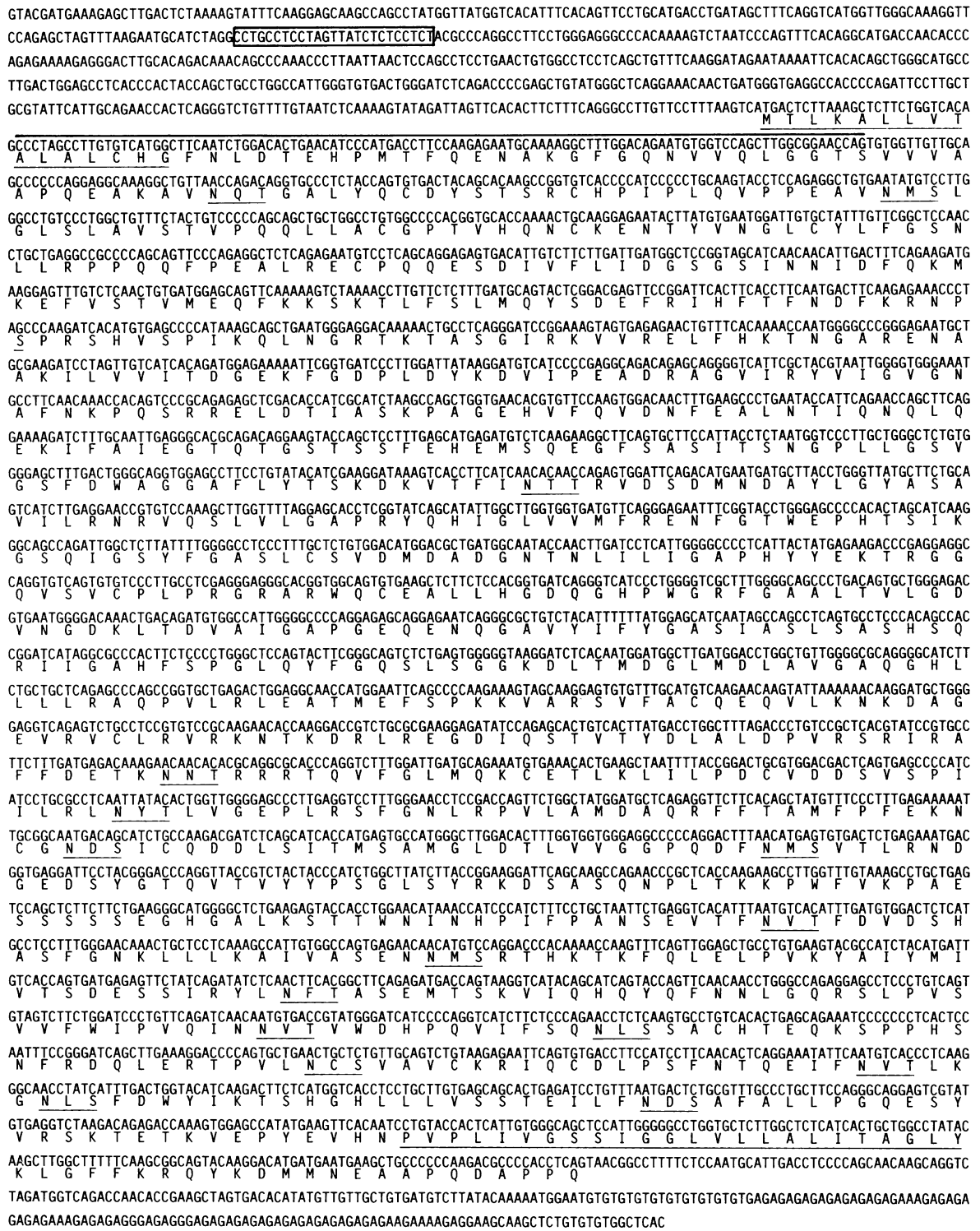

Fig. 2. Composite nucleotide sequence of Mac-1 cDNA clones and deduced amino acid sequence. The amino acid sequence is shown in the one letter code below the respective codons. The putative signal peptide and transmembrane domain are underlined. The overlined nucleotide sequence (residues 602–707) corresponds to the 106 nt exon sequence of a genomic Mac-1 clone (Sastre *et al.*, 1986). A putative interferon-regulatory element in the 5′ untranslated region is boxed. Underlined NXT/S sequences are potential acceptor sites for *N*-linked glycosylation.

All the characterized extracellular ligands for integrins contain the sequence Arg-Gly-Asp (RGD), and it appears that this sequence is generally used as an adhesive signal in integrin–ligand interactions (Ruoslahti and Pierschbacher, 1987). The only known ligand for Mac-1 is the iC3b fragment of C3, which contains an RGD sequence in its C-terminal domain, and binding studies using a synthetic peptide suggest that this site is involved in the interaction of Mac-1 with iC3b (Wright *et al.*, 1987). In addition, Mac-1 is known to be involved in the adhesion of macrophages and granulocytes to surfaces and endothelial cells, and these interactions, which are independent of iC3b, may be mediated by a distinct binding site on the Mac-1 molecule (Dana *et al.*, 1986).

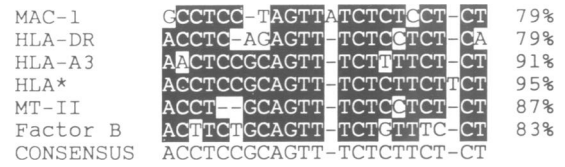

Fig. 3. Alignment of a 22 base sequence from the 5′ region of the Mac-1 cDNA (box in Figure 2) with homologous sequences of five interferon-regulated human genes (Friedman and Stark, 1985; Wu *et al.*, 1987). HLA-DR, class II MHC gene; HLA-A3 and HLA*, class I MHC genes; MT-II, metallothionin-II;B, complement factor B. Inverse printing indicates nucleotides that agree with the consensus sequence. Percentages indicate the degree of identity with the consensus sequence.

In order to analyse further the similarities and differences between Mac-1 and the other known integrins, it was important to obtain the sequence of the α chain, which is expected to play a decisive role in the receptor−ligand interaction. This report describes the cloning and sequencing of the Mac-1 α cDNA and the comparison of the deduced amino acid sequence with the sequences of other integrins.

Results

cDNA cloning and nucleotide sequence

Partial cDNA clones encoding the Mac-1 α chain were selected from a mouse macrophage λgt11 library by screening with oligonucleotides synthesized according to the sequence of a 106 nt exon from a genomic Mac-1 clone (Sastre *et al.*, 1986). The identified clones were sequenced, and after two additional rounds of screening, a set of overlapping clones was obtained as shown in Figure 1. The composite nucleotide sequence covers 4282 nt and includes the 106 nt exon from the genomic Mac-1 clone (Figure 2), with only two substitutions (positions 606 and 656), both of which occur in the third position of a codon and do not alter the deduced amino acid sequence. Comparison of the cDNA sequence with the genomic sequence also confirms that the intron/exon boundaries were correctly predicted by Sastre *et al.*, 1986. Translation of the cDNA sequence (Figure 2) reveals an open reading frame of 3459 nt encoding the 1153 amino acids of the putative Mac-1 α chain sequence. The N-terminal sequence of the mature protein (Springer *et al.*, 1985) is preceded by 15 mostly hydrophobic amino acids that fulfil the requirements for a signal peptide (Watson, 1984), and a methionine most likely encoded by the initiator ATG, since it is preceded by an in-frame stop codon just 3 nt upstream. The sequence of clone 106 extends for another 573 nt in the 5′ direction and contains a site (box in Figure 2) that is very similar to the interferon-regulatory sequences found in the 5′ regions of the human genes encoding three different HLA proteins, metallothionin and complement factor B (Figure 3) (Friedman and Stark, 1985; Wu *et al.*, 1987).

The 4.28 kd cDNA sequence shown in Figure 2 obviously does not include a large portion of the 3′ untranslated region, since the size of the Mac-1 α mRNA as determined by Northern blotting is 6 kb (Sastre *et al.*, 1986). Consequently, the 250 nt sequence following the TAA termination codon does not contain a poly(A) tail or polyadenylation signal.

Amino acid sequence of the mature Mac-1 α chain

The open reading frame shown in Figure 2 encodes a mature protein of 1137 amino acids, starting with the published N-terminal sequence, FNLDTEHPMTFQ (Springer *et al.*, 1985). The predicted mol. wt is 126 000, which is close to the experimental value of 130 000 for the unglycosylated α chain (Sastre *et al.*, 1986). The sequence contains 17 potential *N*-glycosylation sites (underlined in Figure 2), and assuming that all of them are occupied with oligosaccharides having an average mol. wt of 2500, the theoretical mol. wt of the mature glycoprotein would be 168 500, which is almost identical to the observed value of 170 000. Further examination of the amino acid sequence reveals a typical membrane spanning domain of 24 residues close to the C-terminus (underlined in Figure 2) followed by a short segment including three positively charged residues. Thus, the Mac-1 α chain can be divided into a 24 residue cytoplasmic domain and a 1089 residue extracellular portion containing all the potential glycosylation sites, most of them clustered near the transmembrane domain (see Figure 6).

Sequence comparison of Mac-1 and three other integrin α chains

Figure 4 shows the alignment of the Mac-1 sequence with the α chain sequences of vitronectin receptor (VNR), fibronectin receptor (FNR) and gp IIb/IIIa. The overall degree of similarity between Mac-1 (excluding the inserted 220 amino acids) and the other three integrins is much lower (20−22% identity) than between VNR and FNR (46%), and VNR and gp IIb (36%) or FNR and IIb (36%). There are, however, a number of structural features shared by all four sequences. The putative transmembrane domains are in very similar locations and share a high degree of similarity. Interestingly, the pentapeptide sequence, GFFKR, at the cytoplasmic side of the transmembrane domain is identical in all four polypeptides, which makes this region the most highly conserved part of the molecule. Furthermore, there are 13 conserved cysteine residues, one of which (residue 754 of Mac-1) is not conserved in the gp IIb sequence. Also, short stretches of sequence adjacent to these cysteines are generally well conserved. The most striking similarities, however, are evident upon further analysis of the most homologous regions in the N-terminal half of the polypeptides. Comparison of the sequences shown in the large boxes (Figure 4) reveals a very similar pattern of glycines and hydrophobic residues (Figure 5), and an almost invariant LLLGAP (L often replaced by V, I or A) at the C-terminal end of each domain. Four of these repeated domains (IV, V, VI and VII) contain in their central part a hydrophilic sequence with a characteristic pattern of aspartic acid residues (DXDXDGXXD) that is very similar to the EF-hand type of calcium-binding sites found in calmodulin and several other calcium-binding protein (Szebenyi *et al.*, 1981; see also Argraves *et al.*, 1987a; Poncz *et al.*, 1987). This hydrophilic sequence is not conserved in domain IV of Mac-1 and entirely lacking in domains I−III of all four proteins. Comparison of the integrin consensus sequence (Figure 5) with the EF-hand consensus shows that only the central part is identical, while the flanking regions, which are known to be α-helical in calmodulin, are completely different. According to secondary structure prediction (Garnier *et al.*, 1978), the peripheral parts of the integrin repeats would be expected to take up an extended conformation. Also, the integrin repeats are lacking the conserved glutamic acid residue of the EF-hand consensus, which is significant because the side-chain oxygen of this residue is known to coordinate with the metal in the calcium−calmodulin complex.

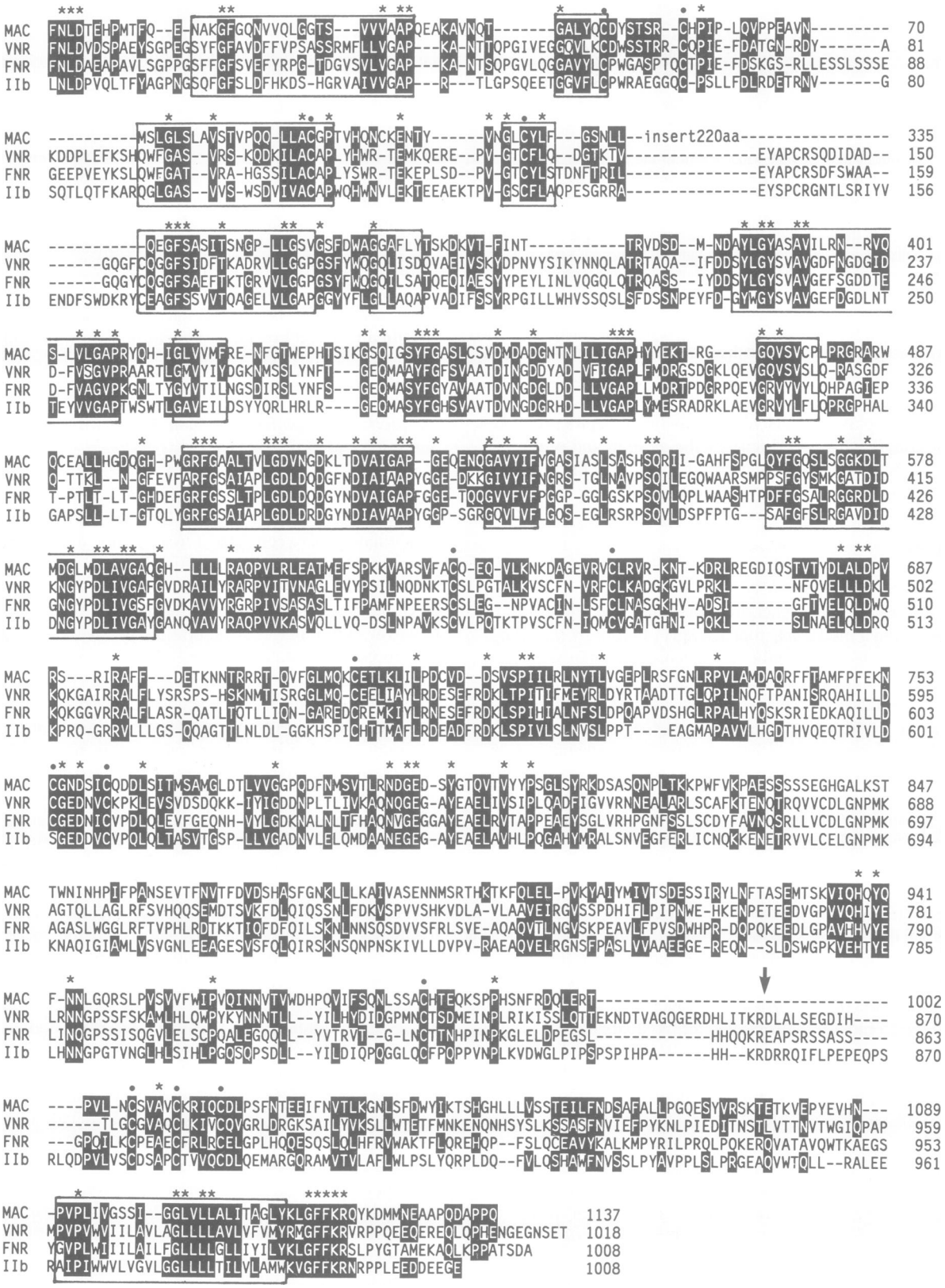

Fig. 4. Alignment of the α chain protein sequences of Mac-1, VNR (Suzuki *et al.*, 1987), FNR (Argraves *et al.*, 1987a) and gp IIb (Poncz *et al.*, 1987). Putative signal peptides are not included. Inverse printing indicates homology with the Mac-1 sequence, including the conservative substitutions L/I/V/A, Y/F, S/T, K/R and E/D/Q/N. Identities of the VNR/FNR/gpIIb sequences that are not shared with Mac-1 are not shown. Boxes represent the seven homologous repeats, as well as the putative transmembrane domains near the C-terminus. Asterisks indicate identical amino acids in all four sequences, and the 13 conserved cysteines are marked by dots. The arrow denotes the putative site of post-translational cleavage preceded by a double basic (KR) sequence, which is not conserved in Mac-1.

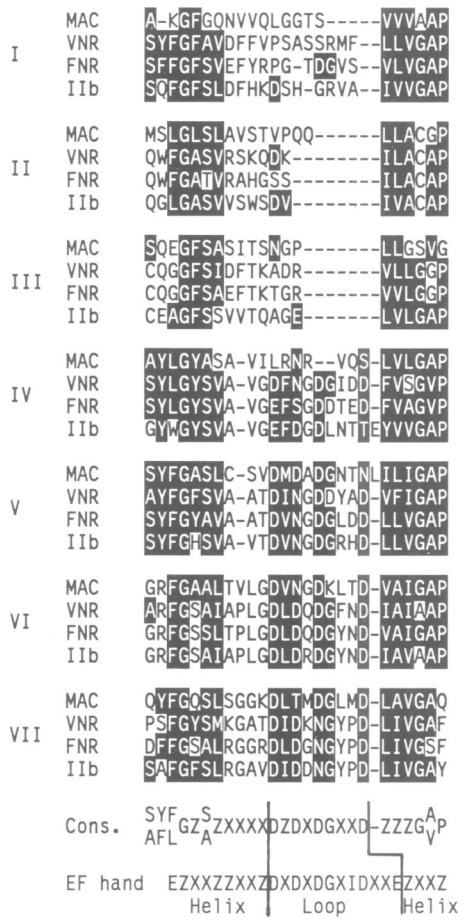

Fig. 5. Alignment of the seven repeating units shown as large boxes in Figure 4. Amino acids in agreement with the consensus sequence are shown in inverse print. The EF-hand consensus sequence is taken from Szebenyi *et al.* (1981). In the consensus sequences, Z stands for any hydrophobic amino acid, X for any amino acid, and D may be replaced by E, N, S or T.

The first six of the seven homologous repeats are followed by a short conserved segment consisting of a glycine followed by 5 hydrophobic residues (small boxes in Figure 4). These short segments are probably part of the seven-repeat structure. However, the distance between the large and the small boxes is quite variable, and the sequences between them bear no similarity.

In contrast to these shared properties, the Mac-1 α chain has a number of unique features. It contains six cysteines that are not matched by cysteines in the other integrins (see Figure 6). Moreover, four cysteines that are conserved among the other integrins (positions 141, 155, 668 and 681 of VNR) have no correspondence in Mac-1. A highly conserved domain of the other integrins (VNR 138−155) has been deleted from the Mac-1 sequence, and another one (VNR 674−688) has been replaced by a completely unrelated sequence. Generally, the C-terminal part of the extracellular portions has diverged significantly, and also the cytoplasmic tails are of different size and structure.

Furthermore, the domain that contains the Lys-Arg sequence that precedes the site of post-translational cleavage (Argraves *et al.*, 1986, 1987a; Suzuki *et al.*, 1986, 1987; Ponz *et al.*, 1987) has been deleted from the Mac-1 sequence (arrow in Figure 4), which agrees with the fact that this

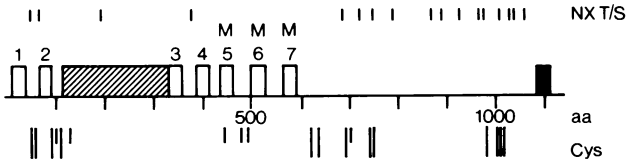

Fig. 6. Domain structure of the Mac-1 α chain. Open boxes indicate repeating units, the cross-hatched box represents the inserted domain homologous to vWF and the filled box stands for the transmembrane domain. M, putative binding sites for divalent cations; NXT/S, potential acceptor sites for *N*-linked glycosylation; Cys, cysteine residues (long lines: cysteines conserved between all four integrin α chains; short lines: cysteines unique for the Mac-1 sequence).

cleavage does not occur in Mac-1. The most obvious difference, however is found in the N-terminal portion, between the second and the third of the homologous repeats (Figure 4), where 220 amino acids are inserted in the Mac-1 sequence. This additional domain is reponsible for the larger size of the unglycosylated Mac-1 polypeptide compared to the other integrins (127 versus 110 kd).

Comparison of the additional domain of Mac-1 with von Willebrand factor (vWF), cartilage matrix protein (CMP) and factor B/C2

A computer search of the NBRF protein sequence database using a short stretch (positions 134−149) from the Mac-1 additional domain revealed a significant homology with the published sequence of vWF (Titani *et al.*, 1986). Further analysis showed that the entire additional domain is homologous to the type A homologous repeats of vWF, shown as cross-hatched boxes in Figure 7 (Sadler *et al.*, 1986; Shelton-Inloes *et al.*, 1986). As noted by these authors, the same domain is homologous to a region of the complement factors B and C2. Furthermore, the sequence (Argraves *et al.*, 1987b) of CMP contains two units of the same homologous repeat. Alignment of these sequences with the additional domain of Mac-1 (Figure 8) shows that they all share a significant degree of homology, which is most pronounced in one large (Mac-1 residues 133−165) and two small subdomains (230−243 and 264−272). The overall degree of identity ranges between 14−22%, with the exception of factor B and C2, which are known to be closely related to each other (34%). However, in the most conserved region (Mac-1 133−165), a much higher degree of identity of 28−41% can be observed, and this number rises to 60−80% if conservative replacements are included.

Discussion

This report presents the complete amino acid sequence, deduced from cDNA, of the murine macrophage cell surface protein, Mac-1. It is clear that this sequence is authentic because it includes the previously described 28 N-terminal residues of the mature Mac-1 protein (Sastre *et al.*, 1986). Moreover, the predicted protein sequence is undoubtedly the sequence of a member of the integrin α chain family, which was to be expected since the Mac-1 β chain is also very similar to the other integrin β chains (Kishimoto *et al.*, 1987; Law *et al.*, 1987). Finally, the size of the predicted protein closely agrees with the experimental value.

An unusual feature of the Mac-1 cDNA is that it contains a very large 5′ untranslated region (at least 573 nt). Interest-

ingly, a consensus sequence of interferon-regulated genes is present in this region (Figure 3). Since it is known that Mac-1 mRNA expression in myelomonocytic cell lines is stimulated by interferons (Sastre *et al.*, 1986), it is possible that this consensus element plays a role in the positive regulation of transcription. However, it should be noted that in all other cases this regulatory site is found upstream of the transcription start site and thus is not included in the cDNA sequence.

Complete amino acid sequences of three other integrin α chains have been published, and it was found that they are closely related to each other, with an overall degree of identity of 36–46%. Mac-1 is much more distantly related to the other three α chains (Figure 4), sharing only 20–22% identical residues. Also, a much larger number of gaps has to be introduced to achieve optimal alignment. This contrasts with the high degree of identity between the corresponding β chains (~45%). A possible explanation for this discrepancy could be that during evolution α chains diverged at a time when the same β chain was still used for all the integrins, or alternatively, that the differences in the α chain sequences reflect unique functions of Mac-1 that are only dependent on the α subunit.

Despite the low overall degree of similarity, there can be no doubt that the Mac-1 α chain is a member of the integrin α chain family, because of the very similar location of the transmembane domain, conserved positions of 13 cysteine residues and a number of short domains that are very similar in all four polypeptides. Some of the most conserved domains (40–65% identity) in the N-terminal half of the molecule can be arranged in seven repeating units (Figures 4 and 5). Therefore, it is likely that this part of the protein was originally made up of seven tandem repeats, each ~65 amino acids long, that arose by a series of duplication events but subsequently diverged from each other leaving only the most essential structural features unchanged. As shown in Figure 5, putative binding sites for divalent cations are present in some of these repeats. This finding is not unexpected

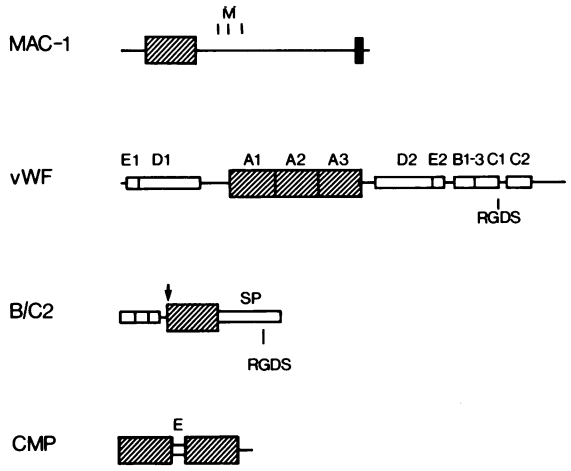

Fig. 7. Domain structure of Mac-1, vWF, factors B/C2 and CMP. Homologous domains are shown as cross-hatched boxes. The filled box indicates the transmembrane domain, open boxes represent type B, C, D and E repeats in vWF and the 3-fold N-terminal repeat in B/C2. SP, serine protease domain; M, putative metal binding sites; RGDS, Arg-Gly-Asp-Ser sequences representing potential interaction sites with adhesive cell surface receptors, E, EGF-like domain. The arrow denotes the site where the zymogens B and C2 are cleaved into the Ba/C2b fragments and the active serine proteases Bb/C2a.

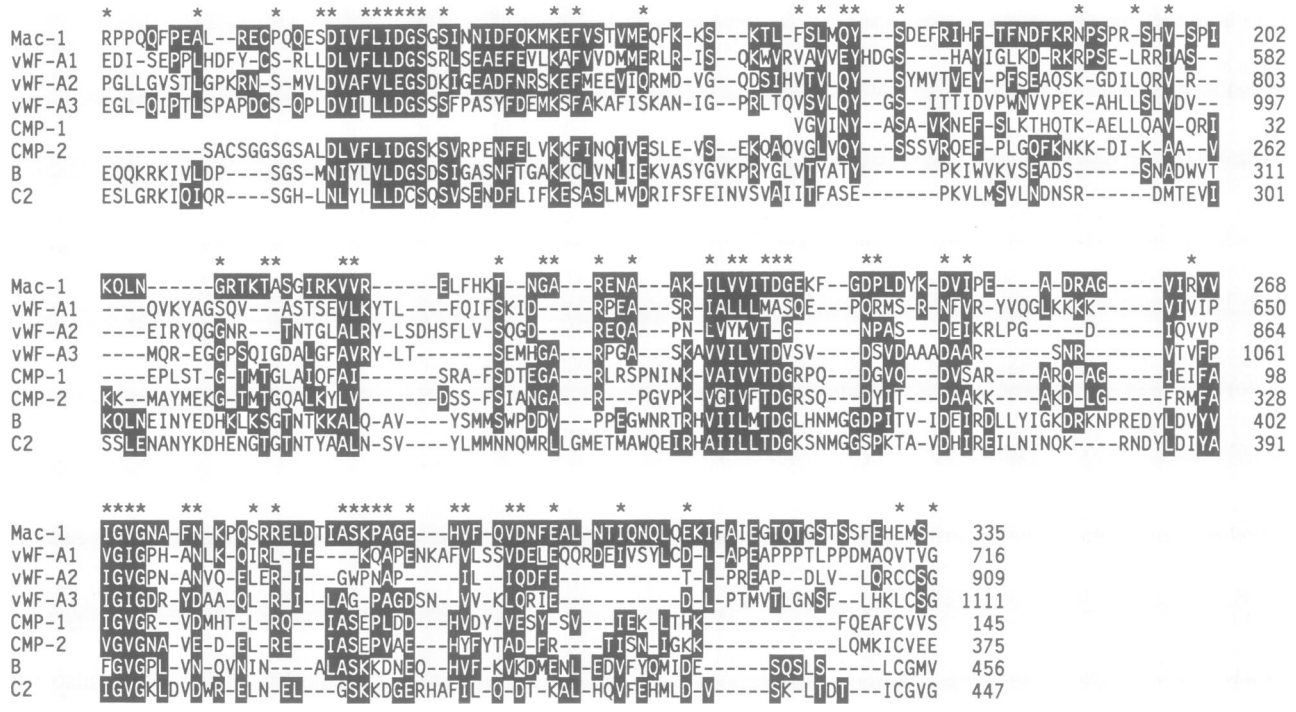

Fig. 8. Alignment of the inserted domain of Mac-1 (residues 116–335) with the repeated domains A1 to A3 (residues 497–1111) from the sequence of vWF (Sadler *et al.*, 1985; Titani *et al.*, 1986), residues 1–145 (CMP-1) and 184–375 (CMP-2) of CMP (Argraves *et al.*, 1987b) and residues 230–456 of the human factor B sequence (Mole *et al.*, 1984) as well as residues 220–447 of the C2 sequence (Bentley, 1986). Rules for inverse printing of residues conserved between Mac-1 and the other seven sequences are as in Figure 4. Positions of identical amino acids shared by at least four of the eight sequences are indicated by asterisks.

because the function of all integrins that have been characterized is dependent on Ca^{2+} or Mg^{2+}. In VNR, FNR and gp IIb, repeats IV, V, VI and VII contain metal binding loops, whereas only three of them (V–VII) are conserved in Mac-1. Repeats I–III are lacking part of the metal binding domain. It appears likely that in the primordial integrin, all the repeats contained a metal-binding domain and that in the course of evolution some of these were lost by deletion events.

In addition to the sequences homologous to the other integrin α chains, Mac-1 contains a 220 amino acid domain that apparently has been inserted between the second and the third of the homologous repeats. The sequence of these 220 amino acids is related to domains that occur in four other extracellular proteins, i.e. von Willebrand factor (vWF), cartilage matrix protein, complement factor B and C2 (Figures 7 and 8). In vWF, it occurs in the form of three tandem repeats between residues 496 and 1111 of the mature protein. The same region has been described to contain at least two independent collagen-binding sites (Pareti *et al.*, 1987). Therefore, it is likely that each one of the three repeats has collagen-binding activity. This interaction is believed to be important for the function of vWF as a mediator of platelet adhesion to the subendothelial matrix, which is a critical step in hemostasis. A very similar function has been ascribed to Mac-1, which is known to be involved in the adhesion of phagocytes to endothelial cells *in vitro*, as well as in the recruitment of macrophges to inflammatory stimuli *in vivo* (Rosen and Gordon, 1987). The matrix component that is recognized by Mac-1 has not been identified, but based on its homology with vWF one could speculate that Mac-1 also binds to collagen fibrils. This would imply that the adhesive function of Mac-1 is mediated by the extra domain, while the iC3b binding site would probably reside in the integrin structure. The latter prediction is also suggested by the finding that iC3b binding is mimicked by an RGD-containing peptide and therefore may be similar to the activity of other integrins (Wright *et al.*, 1987).

The hypothesis that the extra domain of Mac-1 may mediate collagen-binding is further supported by the fact that two similar domains are present in CMP, which is known to interact with collagen (Argraves *et al.*, 1987b). Since the two homologous repeats account for ~80% of the CMP sequence (Figure 7), it is very likely that they include the region(s) that interact with collagen.

Factors B and C2 are closely related to each other and play important roles in the alternative (B) and classical (C2) pathways of complement activation. Both proteins are zymogens containing a C-terminal serine protease domain which is activated by a cleavage that removes the N-terminal 234 amino acids (Figure 7). Between the protease domain and the cleavage site, there is a stretch of 220 amino acids of previously unknown function. This sequence is homologous to the A domains of vWF and to the extra domain of Mac-1 (Figure 8), which provides a new clue as to the possible function of this domain in factor B and C2. Based on the homology with vWF, one could speculate that this domain confers collagen-binding activity, which might play a role in complement localization to inflammatory sites. Alternatively, it could be important for some previously unrecognized function of these complement components, which are synthesized by many cell types including macrophages and fibroblasts. It is interesting to note that factor B, like vWF, contains an RGDS sequence close to its C-terminal end (Figure 7). Thus, factor B could potentially mediate cell adhesion to extracellular matrix by interacting with RGD-dependent receptors on the cell surface and simultaneously with collagen fibers via its vWF-related domain. This type of interaction could possibly play a role in the adhesion of phagocytes. Indeed, it has been reported that factor B is able to bind to the cell surface and to induce spreading of adherent macrophages (Götze *et al.*, 1979).

The occurrence of a homologous domain in the otherwise completely unrelated proteins Mac-1, vWF, CMP and factors B/C2 probably represents another example of information transfer ('exon shuffling') between different genes. It will be important to investigate whether these sequence homologies are reflected in functional similarities between the respective proteins.

Materials and methods

DNA cloning

An oligo (dT)-primed mouse macrophage cDNA library in the λgt11 vector was obtained from Clontech Laboratories, Palo Alto, CA. The initial screening was done with two 33mer oligonucleotides, TTCAATCTGGACACTGAACATCCCATGACCTTC and GCTTTGGACAGAATGTGGTCCAGCTTGGCGGAA, following standard techniques for labeling and hybridization (Maniatis *et al.*, 1982). Two further rounds of screening were performed with oligonucleotides corresponding to the 3′ end of previously identified clones. Inserts were excised using *Eco*RI and purified by electrophoresis on low melting point agarose.

DNA sequencing

cDNA inserts were subcloned into the phage vector M13mp18 or into the bluescript plasmid (Stratagene, San Diego, CA). The insert sequence was determined according to Sanger *et al.* (1987), using modified T7 DNA polymerase (Sequenase) obtained from USB, Cleveland, Ohio, or the Klenow fragment of *E.coli* DNA polymerase I (Pharmacia). Both ends of each subcloned DNA fragment were sequenced using universal primers that hybridize to the vector DNA (Pharmacia), and these sequences were extended using 17mer oligonucleotide primers. The sequence reported here is the result of sequencing both strands of the cDNA inserts, and most of it was determined from at least two independent cDNA clones.

Acknowledgements

I am grateful to J.Hatton, R.Schulze and H.Kiefer for oligonucleotide synthesis, J.Schwager for help with cloning techniques, D.Grossberger for helpful discussion and advice, C.Servis, J.Schwager and J.F.Kaufman for critical reading of the manuscript, P.Riegert for technical assistance, C.Harley and J.Millar for typing the manuscript and H.Stahlberger for artwork. The Basel Institute for Immunology was founded and is supported by F.Hoffman-La Roche & Co. Ltd, Basel, Switzerland.

References

Anderson,D.C. and Springer,T.A. (1987) *Annu. Rev. Med.*, **38**, 75–194.
Argraves,W.S., Pytela,R., Suzuki,S.,Millan,J.L., Pierschbacher,M.D. and Ruoslahti,E. (1986) *J. Biol. Chem.*, **261**,12922–12924.
Argraves,W.S., Suzuki,S., Arai,H., Thompson,K., Pierschbacher,M.D. and Ruoslahti,E. (1987a) *J. Cell. Biol.*, **105**, 1183–1190.
Argraves,W.S., Deak,F., Sparks,K.J., Kiss,I. and Goetinck,P. (1987b) *Proc. Natl. Acad. Sci. USA*, **84**, 464–468.
Beller,D.I., Springer,T.A. and Schreiber,R.D. (1982) *J. Exp. Med.*, **156**, 1000–1009.
Bentley,D.R. (1986) *Biochem. J.*, **239**, 339–345.
Dana,N., Styrt,B., Griffin,J.D., Todd,R.F.,III, Klempner,M.S. and Arnaout,A.A. (1986) *J. Immunol.*, **137**, 3259–3263.
Fitzgerald,L.A., Steiner,B., Rall,S.C.,Jr, Lo,S.S. and Philips,D.R. (1987) *J. Biol. Chem.*, **262**, 3936–3939.
Friedman,R.L. and Stark,G.R. (1985) *Nature*, **314**, 637–639.
Garnier,J., Osguthorpe,D.J. and Robson,B. (1978) *J. Mol. Biol.*, **120**, 97–120.

Ginsberg,M.H., Loftus,J., Ryckwaert,J.J., Pierschbacher,M.D., Pytela,R., Ruoslahti,E. and Plow,E.F. (1987) *J. Biol. Chem.*, **262**, 5437–5440.
Götze,O., Bianco,C. and Cohn,Z.A. (1979) *J. Exp. Med.*, **149**, 372–386.
Hirst,R., Horwitz,A., Buck,C. and Rohrschneider,L. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6470–6474.
Horwitz,A., Duggan,K., Buck,C., Beckerle,M.C. and Burridge,K. (1986) *Nature*, **320**, 531–533.
Hynes,R.O. (1987) *Cell*, **48**, 549–554.
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.
Leptin,M., Aebersold,R. and Wilcox,M. (1987) *EMBO J.*, **6**, 1037–1043.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Miller,L.J., Wiebe,M. and Springer,T.A. (1987) *J. Immunol.*, **138**, 2381–2383.
Mole,J.E., Anderson,J.K., Davidson,E.A. and Woods,D.E. (1984) *J. Biol. Chem.*, **259**, 3407–3412.
Pareti,F.I., Niiya,K., McPherson,J.M .and Ruggeri,Z. (1987) *J. Biol. Chem.*, **262**, 13835–13841.
Poncz,M., Eisman,R., Heidenreich,R., Silver,S., Vilaire,G., Surrey,S., Schwartz,E. and Bennett,J.S. (1987) *J. Biol. Chem.*, **262**, 8476–8482.
Pytela,R., Pierschbacher,M.D. and Ruoslahti,E. (1985a) *Cell*, **40**, 191–198.
Pytela,R., Pierschbacher,M.D. and Ruoslahti,E. (1985b) *Proc. Natl. Acad. Sci. USA*, **82**, 5766–5770.
Pytela,R., Pierschbacher,M.D., Ginsberg,M.H., Plow,E.F. and Ruoslahti,E. (1986) *Science*, **231**, 1559–1562.
Rosen,H. and Gordon,S. (1987) *J. Exp. Med.*, **166**, 1685–1701.
Ruoslahti,E. and Pierschbacher,M.D. (1987) *Science*, **238**, 491–497.
Sadler,J.E., Shelton-Inloes,B.B., Sorace,J.M., Harlan,J.M., Titani,K. and Davie,E.W. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6394–6398.
Sanchez-Madrid,F., Nagy,J.A., Robbins,E., Simon,P. and Springer,T.A. (1983) *J. Exp. Med.*, **158**, 1785–1803.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
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. USA*, **83**, 5644–5648.
Shelton-Inloes,B.B., Titani,K. and Sadler,J.E. (1986) *Biochemistry*, **25**, 3164–3171.
Springer,T.A., Teplow,D.B. and Dreyer,W.J. (1985) *Nature*, **314**, 540–542.
Suzuki,S., Pytela,R., Arai,H., Argraves,W.S., Krusius,T., Pierschbacher,M.D. and Ruoslahti,E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8614–8618.
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., Huang,C. and Hemler,M.E. (1987) *Nature*, **326**, 607–609.
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.
Titani,K., Kumar,S., Takio,K., Ericsson,L.H., Wade,R.D., Ashida,K., Walsh,K.A., Chopek,M.W., Sadler,J.E. and Fujikawa,K. (1986) *Biochemistry*, **25**, 3171–3184.
Watson,M.E.E. (1984) *Nucleic Acids Res.*, **12**, 5145–5164.
Wright,S.D., Reddy,P.A., Jong,M.T.C. and Erickson,B.W. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1965–1968.
Wu,L.C., Morley,B.J. and Campbell,R.D. (1987) *Cell*, **48**, 331–342.

Received on December 23, 1987; revised on February 15, 1988