PROTEIN SEQUENCE MOTIFS

An adhesive domain detected in functionally diverse receptors

Many extracellular proteins and receptors consist of a variety of domains or modules[1–3] brought together by exon shuffling[4]. This modular organization is an elegant way of fine-tuning a network of defined binding interactions between functionally diverse proteins. Thus, modules are often the only link between otherwise distinct proteins. Usually, they carry out a particular subfunction and have a common structural framework. Here, we define a new domain that we have identified in the functionally diverse receptors meprin, A5 protein and protein tyrosine phosphatase μ (rptpμ) (Fig. 1). We therefore propose the name MAM (Meprin, A5, μ) for this domain and suggest that it is likely to be widespread among various adhesive proteins.

MAM spans about 170 amino acids and contains four conserved Cys residues that probably form disulfide bridges, as in most other extracellular modules[1–3]. Hydrophobicity patterns (Fig. 1a), as well as secondary structure prediction methods, reveal the presence of a β-sheet (data not shown). Two conserved regions: tChtFahhxtt and ttGhhxhD-hxh (for nomenclature see Fig. 1a legend) appear to be unique in sequence databases and were therefore chosen as signature sequences for their addition to the PROSITE motif database[5].

All the proteins in which MAM has been detected have a modular, receptor-like architecture with a signal peptide, followed by a large amino-terminal extracellular part and a single transmembrane region (Fig. 1b). Although all these proteins appear to be expressed only in specialized tissues, they have little in common. Meprins are cell-surface glycoproteins that contain a domain related to zinc-metalloproteases, which can degrade a variety of polypeptides[6]. Two distinct subunits (α and β) have been identified and sequenced[6] [G. D. Johnson and L. D. Hersh, submitted (EMBL accession number, M88601)], which form homo- and heterotetramers. A5 is a developmentally regulated cell-surface molecule predominately expressed in the visual centers of the diencephalon and mesencephalon during nerve innervation[7]. Receptor protein tyrosine phosphatases (rptps) play an essential role in the regulation of various cell activities[8]. Cytosolic dephosphorylation is initiated by binding of respective ligands to defined extracellular domains. These domains are predominantly fibronectin type III modules and immunoglobulin-like domains, but recently a carbonic anhydrase-like domain has also been found in rptpζ (Ref. 9). Some extracellular regions of rptps (e.g. in CD45 or rptpζ) have not yet been classified by sequence analysis[9], but it is likely that they contain modules, such as MAM, which remain to be identified.

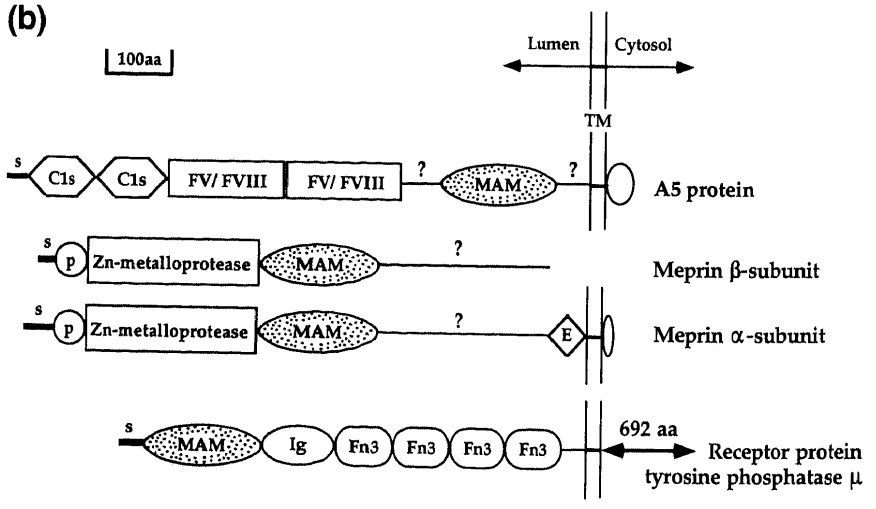
(a)

```
Rtpm/Mouse   22 TFSGGCLFD..EPYSTCGYSQADEDDFNWEQVNTL......TKPTSDPWMPSGSFMLVNTSGKPEGQRAHLEKGLLYPKFGFQC
Rtpm/Human   22 TFSGGCLFD..EPYSTCGYSQSEGDDFNWEQVNTL......TKPTSDPWMPSGSLMLVNASGRPEGQRAHLLLPQLKEN.DTHC
  A5/Xenla  646 HSDLDCKFGWGSQKTVCNWQHDISSDLKWAVLNSK......TGPVQD.HTGDGNFIYSEADERHEGRAARLMSPVVSSSRSAHC
Mpra/Mouse  276 TLLDHCDFE...KTNVCGMIQGTRDDADWAHGDSS.QPEQVDHTLVEQCKGAGYFMFFNTSLGARGEAALLESRILYPKRKQQC
  Mprb/Rat  261 SFMDSCDFE...LENICGMIQSSQDSADWQRLSQVLSGPENDHSNMGQCKDSGFFMHFNTSTGNGGITAMLESRVLYPKRGFQC
consensus:      t   tC Ft     thCth tt ttthtW  htt       tt   t h ttG hhh ttt tttG  AhL  t h tt t tC

IKFYYFVSSKSNAAPGLLNVYVKVN.NGPLGNPIWNIS....GDPTRTWHRAELAISTFWPNFYQVIFEVVTS..G HQGYLAIDEVKVLGH....PCTR
IDFHYFVSSKSNSPPGLLNVYVKVN.NGPLGNPIWNIS..:..GDPTRTWNRAELAISTFWPNFYQVIFEVITS..G HQGYLAIDEVKVLGH....PCTR
LTFWYHMDG...SHVGTLSIKLKYEMEEDFDQTLWTVS....GNQGDQWKEARVVLHKTM KQYQVIVEGTVG.KG.SAGGIAVDDIIIANHISPSQCRA
LQFFYKMTG...SPADRFEVWVRRDDNAGKVRQLAKIQT.FQGDSDHNWKIAHVTLNEE. KKFRYVFLGTKGDPGNSSGGIYLDDITLTET....PCPA
VEFYLYNSG...SGNGQLNVYTREY TAGHQDGVLTLQREIRDIPTGSWQLYYVTLQVT. EKFRVVFEGVGG.PGASSGGLSIDDINLSET....RCPH
htFahh tt   tt t hthhh+   ttth tthhtht    t ttttWt h hhht    t athhh hhtt  G ttGhh hD-h h tt    tCtt
```

Figure 1

(a) Alignment of the MAM domains. The beginning of the domain in the respective sequences is given. The two signature sequences are underlined. Positions in which four out of five sequences have identical amino acids are shown in bold. The consensus highlights conserved features: capitals, strictly conserved residues; h, hydrophobic; a, aromatic; t, turn-like or polar, probably located at the surface; –/+, charged. The pair-wise overall identities range between 21% [A5/rptpμ], 28% [meprin/rptpμ] and 31% [A5/meprin]. The significance of the homology was also assessed by other methods: FASTA searches[10] with A5 protein result in high optimized scores for meprin (161) and rptpμ (173). In a PROFILESEARCH[11] with the presented alignment, all MAM domains have z-scores above 17.7, whereas other sequences score lower than 5.9. Finally, a PROPAT database screening[12] with the tChtFahhtt motif alone discriminates non-related sequences by at least six mismatches.



(b) Modular architecture of A5, meprin and rptpμ. Ig, immunoglobulin-like domain; Fn3, fibronectin type III domain; p, propeptide; s, signal peptide; TM, transmembrane region; FV/FVIII, a domain found in several proteins including coagulation factors V and VIII; C1s, a domain first identified in complement subcomponents C1s/C1r, but it also occurs in several developmentally regulated proteins. In most of them the C1s domain (Fig. 1b) is located next to a zinc-metalloprotease which is present in meprin but not in A5.

Despite the functional diversity of the proteins shown in Fig. 1, MAM is likely to have an adhesive function, as do most other well-characterized extracellular modules. Although the exact function of MAM remains to be established, the identification of this new domain and the characterization of conserved regions might stimulate NMR studies, mutation experiments and genomic screening for the presence of homologous domains in other proteins.

References

1 Bork, P. (1991) *FEBS Lett.* 286, 47–54
2 Patthy, L. (1991) *Curr. Opin. Struct. Biol.* 1, 351–361
3 Bork, P. (1992) *Curr. Opin. Struct. Biol.* 2, 413–421
4 Rogers, J. H. (1990) *FEBS Lett.* 268, 339–343
5 Bairoch, A. (1992) *Nucleic Acids Res.* 20, 2013–2018
6 Jiang, W. *et al.* (1992) *J. Biol. Chem.* 267, 9185
7 Takagi, S. *et al.* (1991) *Neuron* 7, 295–307
8 Fischer, E. H., Charbonneau, H. and Tonks, N. K. (1991) *Science* 253, 401–406
9 Krueger, N. X. and Saito, H. (1992) *Proc. Natl Acad. Sci. USA* 89, 7417–7421
10 Pearson, W. R. and Lipmann, D. J. (1988) *Proc. Natl Acad. Sci. USA* 85, 2444–2448
11 Gribskov, M., McLachlan, A. D. and Eisenberg, D. (1987) *Proc. Natl Acad. Sci. USA* 84, 4355–4358
12 Bork, P. and Grunwald, C. (1990) *Eur. J. Biochem.* 191, 347–358

GEORG BECKMANN AND PEER BORK

Max-Delbrück-Centre of Molecular Medicine, 1115 Berlin-Buch, Germany and EMBL, W-6900 Heidelberg, Germany.

A putative fatty acid-binding domain of the NMDA receptor

Since the demonstration that arachidonic acid and other fatty acids directly regulate K^+ channels[1] and other ion channel types[2] it has not been possible to determine whether these amphiphilic molecules regulate channel activity by an alteration of the 'lipid environment' in the vicinity of the ion channel or, as we have previously suggested[1,2], by an interaction with the ion channel protein itself. For one ion channel, the NMDA receptor, whose activity is modulated by arachidonic acid[3], the deduced amino acid sequence is now available[4]. Using this sequence we asked whether any region of the NMDA receptor resembled the known fatty acid-binding proteins, some of which have not only been cloned but also studied with high-resolution X-ray crystallography[5,6]. Such a region, we supposed, might confer sensitivity to fatty acids on the NMDA receptor itself.

We compared the amino acid sequence of the NMDA receptor with the amino acid sequences of fatty acid-binding proteins and other members of the intracellular lipid-binding protein family. This comparison revealed a significant degree of similarity between a 131-residue domain of the NMDA receptor (residues 263–393 inclusive) and representative members of the intracellular lipid-binding protein family, suggesting that the 131-residue domain is indeed homologous to the proteins in this family (Fig. 1).

Members of the intracellular lipid-binding protein family share a common structural motif even though they display as little as 20% amino acid identity[7]. X-ray crystallographic studies have demonstrated

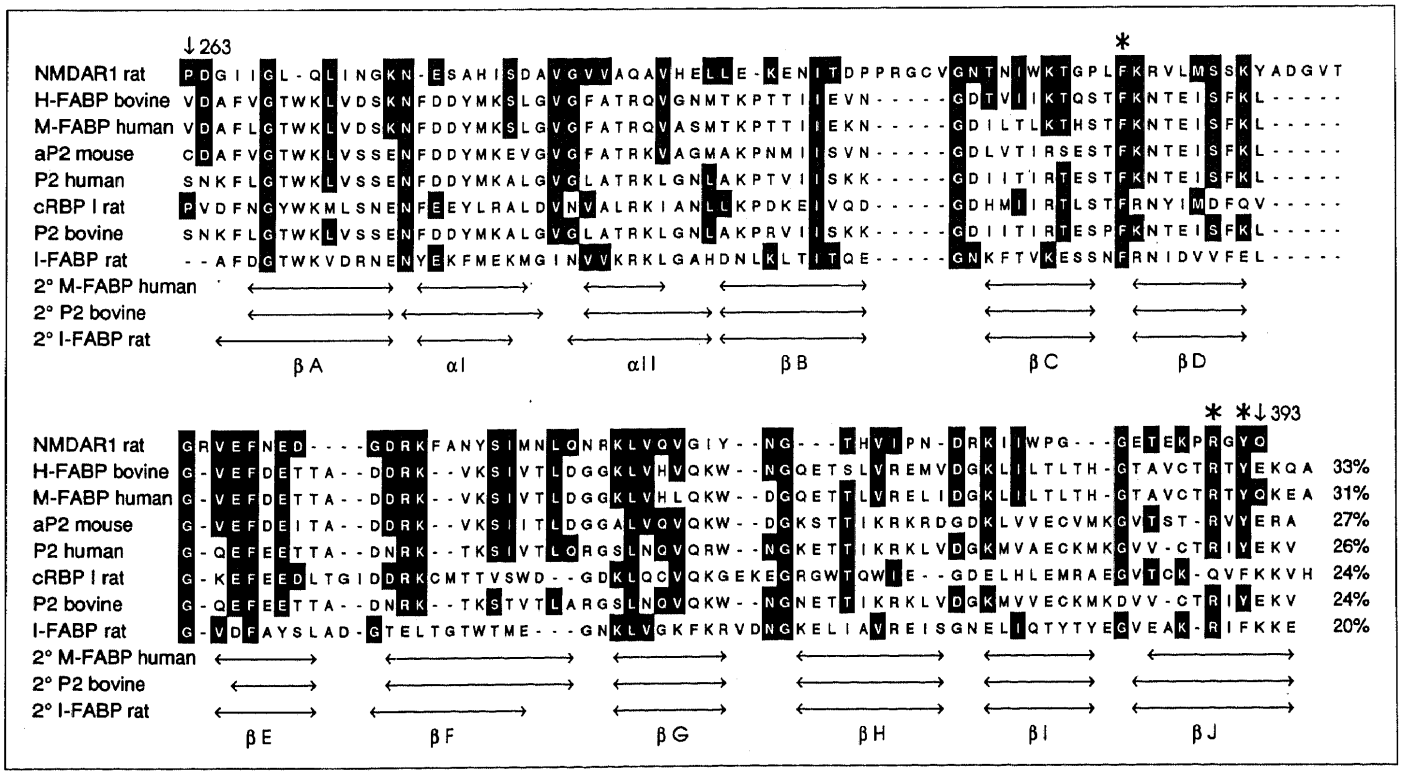


Figure 1

Alignment of the amino acids of the putative fatty acid-binding domain of the NMDA receptor (residues 263–393 inclusive) with representative members of the intracellular lipid-binding protein family. Residues that are identical to those in the NMDA receptor sequence are highlighted. Percentage identity to the 131-residue domain of the NMDA receptor is indicated for bovine heart and human muscle fatty acid-binding proteins (H-FABP and M-FABP), mouse adipocyte lipid-binding protein (aP2), human P2 myelin protein, rat cellular retinol-binding protein type I (cRBP I), bovine P2 myelin protein and rat intestinal fatty acid-binding protein (I-FABP). The positions of Phe55 of rat intestinal fatty acid-binding protein, Phe57 and Arg126 of human muscle fatty acid-binding protein and Tyr128 of bovine P2 myelin protein, which are thought to be important for fatty acid-binding, are marked by an asterisk. The secondary structures of human muscle fatty acid-binding protein, bovine P2 myelin protein and rat intestinal fatty acid-binding protein are also shown. Initial searching for sequence similarity was performed using the PROSCAN program in the DNASTAR package. Multiple sequence alignment was then performed using the MACAW program[11].