Common domain structure of Ca^{2+} and lipid-binding proteins

Michael J. Geisow

Laboratory of Protein Structure, National Institute for Medical Research, Mill Hill, London NW7 1AA, England

Received 30 May 1986

The phospholipase A_2 inhibitor, lipocortin, shares common sequences with three abundant Ca^{2+}-regulated membrane binding proteins of unknown function which are present in many cell and tissue types. A two-domain model for the structure of lipocortin is described and it is suggested that the new Ca^{2+}-regulated proteins each contain at least one lipocortin domain. The structural and biochemical properties of each protein indicate that they all directly interact with phospholipids. Potential sites of interaction with the lipocortin domain are identified on the basis of homology with phospholipid transfer proteins and phospholipase A_2.

Lipocortin *Ca^{2+} binding* *Endonexin* *Calelectrin* *Phospholipase A_2* *Membrane binding*

1. INTRODUCTION

Phospholipids play a central role in signal transduction [1]. Inositol phospholipids are the source of two signals. First, inositol trisphosphate, which mobilizes intracellular Ca^{2+}. Second, diacylglycerol, which activates C-kinase. The same lipids, together with phosphatidylcholine, are the source of prostaglandins and leurotrienes via released *cis*-unsaturated fatty acids [2]. We have identified members of a new class of Ca^{2+}-regulated membrane binding proteins, whose membrane targets appear to be lipids [3–6]. Each protein binds to liposomes containing acidic phospholipids at moderate levels of Ca^{2+} (1–10 μM) and aggregates liposomes or membrane vesicles at higher Ca^{2+} levels (>200 μM) [4–6]. Creutz et al. [7] discovered an analogous protein synexin and, independently, the lipid-binding proteins described here. In view of their similar biochemical and biophysical properties, we have coined the generic term 'annexin' for such proteins [8]. Recently, we published amino acid sequence data for four 'annexins'; endonexin, calelectrin, protein II and p36, which established the presence of a common amino acid sequence – the 'endonexin fold' [6,9]. This feature was also found in the sequence of a phospholipase A_2 inhibitor, lipocortin, deduced from cDNA [10]. Lipocortin is present in many cell types and blocks the production of arachidonate and the consequent prostaglandin/leukotriene-mediated inflammatory response [10,11].

Lipocortin and p36 are strongly implicated in signal transduction. Lipocortin is phosphorylated by protein kinase C [12,13]. Lipocortin and p36 are substrates for EGF receptor and $pp60^{src}$ protein tyrosine kinases [13–15]. Each protein binds Ca^{2+} and shares with protein kinase C the property of Ca^{2+}- and lipid-dependent recruitment by membranes. The functions of the annexin group as a whole are not clear. It is therefore important to make a critical comparison of these molecules, to identify the extent to which they are really related in view of their potential role in transduction of the signals described. I report here a model for the domain structure of lipocortin. I hope this will provide a basis for comparison with the complete structures of the other proteins, when they become known.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

2. MATERIALS AND METHODS

Calelectrin, endonexin, p70 and p36 were purified as described [5,6] and sequences of both cyanogen bromide and trypsin fragments determined using an Applied Biosystems model 470A sequencer. Circular dichroism spectra were obtained using Jouan dichrographs with proteins dissolved either in Hepes or borate buffered solutions. Limited proteolysis of p70 was at 37°C for 30 min using a 1:500 (w/w) ratio of enzyme to substrate. Western blotting of limited proteolytic digests in SDS was performed as in [5].

Sequence homology searches were made using the algorithms of Wilbur and Lipman [16]. Secondary structure analysis was made using a computer implementation of the algorithms of Chou and Fasman [17,18].

3. RESULTS AND DISCUSSION

Specific antisera [4,5,19], peptide mapping and biophysical measurements [4,19] have established the independent identities of the proteins listed in table 1. There are, nevertheless, compelling similarities. The amino acid compositions, particularly of the structurally important hydrophobic residues, are similar [4,20]. Each protein gives a far-ultraviolet circular dichroism spectrum showing a high content of α-helix. Although there are specific antisera for each of the mammalian proteins, two groups have found that polyclonal antisera raised against endonexin will recognise p70, but not vice versa, suggesting that common, but not immunodominant epitopes exist in both proteins ([4], and Raeymaekers, L., University of Leuven, Belgium, personal communication).

We previously reported that antiserum to calelectrin combines with p70, p36 and endonexin on Western blots [4,5] and that this may reflect the presence of a common epitope in a conserved sequence [6]. The partial sequences of endonexin, calelectrin and p36 show that there are multiple copies of this consensus sequence in each protein [6]. To look for the distribution of immunoreactive peptides, each protein was subjected to limited proteolysis by V-8 protease in 0.1% SDS. After electrophoresis, each digest was stained by anticalelectrin on immunoblots. Whereas essentially all calelectrin peptides were immunostained, only a small number of fragments of endonexin, p36 and p70 were recognised. The relative molecular masses of the cross-reacting fragments were distributed around one half and one quarter the initial molecular masses (fig.1a). Under non-denaturing conditions, p70 was rapidly cleaved by both α-chymotrypsin and V-8 protease to 33 kDa fragments and further to 15 and 10 kDa fragments by the former enzyme (fig.1b).

Recently, the complete sequence of human lipocortin was deduced from cloned cDNA [10]. Similarity matrix analysis of the sequence indicated four regions of internal homology. Alignment of these regions, with only two insertions, shows lipocortin to be composed of four similar 70 amino acid folding units (fig.2). Together, these account for most of the 345 amino acid sequence, the remainder being in the unique N-terminus and

Table 1

Biochemical characteristics of annexins

Protein	M_r	K_D' (Ca^{2+}) (M)	Composition (mol%)				α-Helix (%)	Ref.
			Met	Tyr	Phe	Cys		
p70	67000	1.2×10^{-6}	2.6	2.6	3.0	1.1	35	33
Lipocortin	40000	–	2.6	3.2	3.2	0.6	(35)	10
p36	36000	1.0×10^{-4}	2.3	4.3	3.5	–	42	21
Endonexin	32500	2.5×10^{-5}	2.1	3.2	3.6	1.4	36	19
Calelectrin	34000	1.0×10^{-4}	1.5	2.9	2.9	1.5	54	–

Ca^{2+} affinities are reproduced from the references cited. Helix content for lipocortin was derived from structure prediction using the Chou-Fasman method. Amino acid compositions were calculated from sequence (lipocortin) or cited from [4,20]

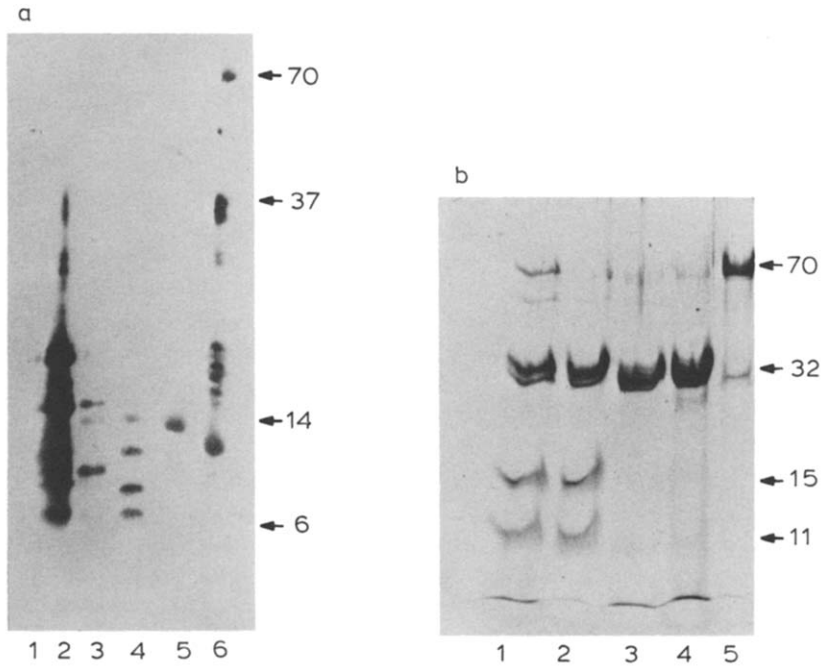

Fig.1. Limited proteolysis of annexins. (a) Immunoblots of annexins digested with *Staphylococcus aureus* V-8 protease in the presence of 0.1% SDS. Lanes: 1, non-annexin control (38 kDa protein present in adrenal membrane EGTA extracts [5]); 2, *Torpedo* calelectrin; 3, p35 (not identified); 4, p32.5 (endonexin); 5, p36 (*src*-gene kinase substrate); 6, p70. The blot was immunostained with anti-calelectrin as described [5]. (b) Digestion of p70 under non-denaturing conditions. Lanes: 1,2, α-chymotrypsin; 3;4, V-8 protease; 5, no protease. In addition, 1 and 3 contained 1 mM Ca^{2+}, 2 and 4, 1 mM EGTA.

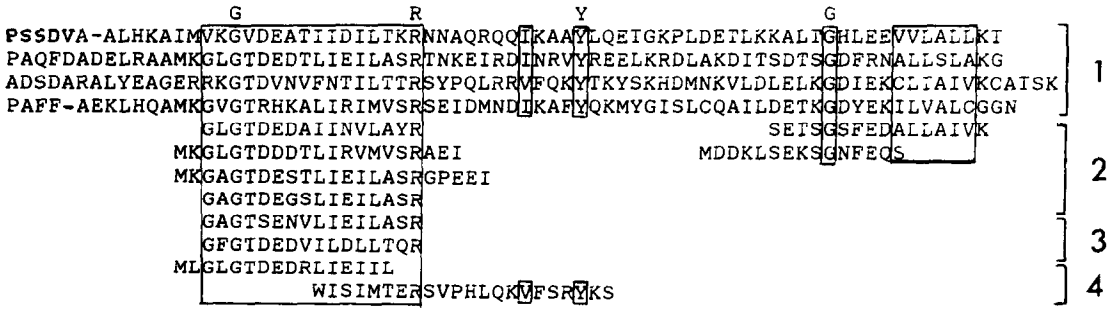

Fig.2. Repeated folding units present in annexins. (1) Lipocortin; the amino-terminus (residues 1–43) and the connection between units 2 and 3 (residues 184–199) are not shown. (2) Endonexin peptides. (3) Calelectrin peptides. (4) p36 peptides. The last p36 sequence is reproduced from [23]. Conserved sequences are shown in boxes and invariant residues on the top line.

a loop between the C-terminus of unit 2 and the N-terminus of unit 3. At this stage, all partial sequences determined for calelectrin, endonexin and p36, together with two independent partial sequences published for the latter protein [22,23] were compared with lipocortin. Only those sequences with significant homology (extending over 7 residues or more) were selected. The alignment highlights two regions of close homology within each 70 amino acid unit (boxed in fig.2). The first is that previously identified – the endonexin fold [6,9]. This extends for 17 residues and contains a

strongly conserved subsequence GhGTDE, where h is usually a hydrophobic residue. The second glycine is invariant and an invariant arginine forms the C-terminal boundary of the consensus sequence.

The second homologous region also occurs around an invariant glycine. The 11 amino acids around this glycine represent the closest approach to an E-F hand [9] in the molecules (with the reservation that only the lipocortin structure is complete). None of the sequences would be predicted to bind Ca^{2+}, since at the '-z' position of a putative octahedral co-ordination complex, where an oxygen-containing side chain must occur for Ca^{2+} binding [24], there is a hydrophobic side chain. This is the first amino acid of a conserved, 6-residue hydrophobic sequence. Sequences interconnecting the two closely homologous regions described have lower, but significant homology. In particular, an invariant β-branched hydrophobic residue and tyrosine occur in the centre of each sequence.

The N-terminus of lipocortin (residues 1–32) and the connecting polypeptide chain between units 2 and 3 are predicted to be relatively unstructured (not shown), and would be expected to be particularly susceptible to proteolysis. This appears to be the case, since lipocortin is easily cleaved into fragments with relative molecular masses of 30 and 15 kDa [10,25]. A 15 kDa fragment (macrocortin) has the full biological activity of the parent protein [26]. p36 is also cleaved by proteases to give a blocked amino terminal fragment and a 33 kDa protein which retains most of the properties of the intact protein [22,23]. The present biochemical and structural evidence are all consistent with the idea that p70, p36, endonexin and lipocortin contain closely similar folding units, two of which, corresponding to a total molecular mass of 15 kDa, make up the membrane binding domains of each protein.

Is there evidence to indicate which regions of the folding units described are responsible for Ca^{2+} and membrane binding? Our experiments with endonexin suggest a minimal requirement for the lipid phosphoryl group for membrane binding [6]. The head group specificity appears to vary, and phosphatidic acid, phosphatidylinositol lipids and phosphatidylserine appear to be bound by different annexins [6,23,27].

It was considered that a binding site for lipid phosphorus might resemble that for (pyro)-phosphate present in dinucleotide-binding proteins. In these proteins, the phosphate makes a close contact with invariant glycine residues in the sequence XhXhGXGXX(G); X = any residue; h = a hydrophobic residue and the final G is optional [28,29]. In fact, precisely this pattern was discovered in the protein data base in the sequences of phospholipase A_2 (PLA_2) and a viper PLA_2 inhibitor [30]. In the X-ray structure of pig PLA_2 this sequence is present in a surface loop of the polypeptide chain. The carbonyl oxygens of the invariant glycine residues act as two ligands of the bound Ca^{2+} upon which the activity of the enzyme depends [31]. The amino acids preceding and following the invariant glycines are in α-helical and extended chain conformation, respectively. The sequence around the two conserved glycines of the endonexin fold (fig.2) differs slightly from that present in dinucleotide-binding proteins. It conforms to the pattern XhXX(G)hGTXXX, where the second hydrophobic group in the dinucleotide protein pattern is replaced by a charged residue in all sequences of fig.2 except for the first unit of lipocortin and the p36 fragment. The second glycine residue is closest to the phosphate ester group in dinucleotide-binding proteins. In the sequences in fig.2, the glycine at this position is invariant. If this glycine were close to the phosphoryl group of bound lipid, the location of the conserved threonine could be significant. Lipocortin is known to be phosphorylated by kinase C and, on the basis of substrate specificity, this threonine residue was predicted to be the phosphate acceptor [10].

PLA_2, a viper PLA_2 inhibitor, and lipid transfer proteins also contain hydrophobic 6-residue sequences. In PLA_2, this sequence contributes side chains to the hydrophobic cavity surrounding the catalytic residues in the active site [31]. In the phosphatidylcholine transfer protein (PC-TP) the corresponding six-residue sequence makes a close contact with the 2-acyl fatty acid chain of bound lipid [32].

We previously suggested that the bound Ca^{2+} in endonexin might be near what appears to be a fragment of the consensus sequence -WGTDEK (not shown in fig.2), on the evidence of efficient energy transfer from Tb^{3+} in the Ca^{2+} site to protein tryp-

tophan [6]. Calcium binding to p36 markedly affects the environment of the single tryptophan molecule and this also appears to be in the conserved endonexin fold (fig.2). In view of the absence of any E-F hands in lipocortin, a tentative suggestion can be made that a complex Ca^{2+}-binding site exists in this protein, perhaps using side-chain carbonyl oxygens as in PLA_2 [31]. In view of the low Ca^{2+} affinity of some of these proteins in the absence of membrane lipids [19,20], it is possible that lipid binding directly affects the coordination of Ca^{2+}, so that affinity for both ion and lipid is increased in the ternary complex.

In summary, lipocortin probably folds into two related 15 kDa domains. At least one of these domains, consisting of two similar 'folding units', is likely to be present in the proteins p70, p36, calelectrin and endonexin. Each of these domains contains sequences which resemble parts of the lipid-binding sites of PLA_2 and lipid transfer proteins.

ACKNOWLEDGEMENTS

I thank Professor R.J. Flower (University of Bath) and Dr N.M. Green (National Institute for Medical Research) for helpful discussions. I am endebted to Herr Ulrich Fritsche and Dr V. Witzemann for the measurement of calelectrin helix content (Max Planck Institute for Biophysical Chemistry, Abteilung Neurochemie, Göttingen, FRG).

REFERENCES

[1] Majerus, P.W., Neufeld, E.J. and Wilson, D.B. (1984) Cell 37, 701–703.
[2] Nishizuka, Y. (1984) Nature 308, 693–698.
[3] Geisow, M.J. and Burgoyne, R.D. (1982) J. Neurochem. 38, 1735–1741.
[4] Südhof, T.C., Ebbecke, M., Walker, J.H., Fritsche, U. and Boustead, C. (1984) Biochemistry 23, 1103–1109.
[5] Geisow, M.J., Childs, J., Dash, B., Harris, A., Panayotou, G., Südhof, T. and Walker, J.H. (1984) EMBO J. 3, 2969–2974.
[6] Geisow, M.J., Fritsche, U., Hexham, J.M., Dash, B. and Johnson, T. (1986) Nature 320, 636–638.
[7] Creutz, C.E., Dowling, L.G., Sando, J.J., Villar-Palasi, C., Whipple, J.H. and Zaks, W.J. (1983) J. Biol. Chem. 258, 14664–14674.
[8] Geisow, M.J. and Walker, J.H. (1986) Trends Biochem. Soc., in press.
[9] Kretsinger, R.H. and Creutz, C.E. (1986) Nature 302, 573–574.
[10] Wallner, B.P., Mattaliano, R.J., Hession, C., Cate, R.L., Tizard, R., Sinclair, L.K., Foeller, C., Pingchang Chow, E., Browning, J.L., Ramachandran, K.L. and Pepinsky, R.B. (1986) Nature 320, 77–81.
[11] Flower, R.J. (1985) Agents Actions 17, 255–262.
[12] Hirata, F. (1981) J. Biol. Chem. 256, 7730–7733.
[13] Hirata, F., Matsuda, K., Notsu, Y., Hattori, T. and Del Carmine, R. (1984) Proc. Natl. Acad. Sci. USA 81, 4717–4721.
[14] Erikson, E., Shealy, D.J. and Erikson, R.L. (1981) J. Biol. Chem. 256, 11381–11384.
[15] Erikson, E. and Erikson, R.L. (1980) Cell 21, 829–836.
[16] Wilbur, W.J. and Lipman, D.J. (1983) Proc. Natl. Acad. Sci. USA 80, 726–730.
[17] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251–276.
[18] Geisow, M.J. and Roberts, R.D.B. (1980) Int. J. Biol. Macromol. 2, 387–389.
[19] Shadle, P.J., Gerke, V. and Weber, K. (1985) J. Biol. Chem. 260, 16354–16360.
[20] Gerke, V. and Weber, K. (1984) EMBO J. 3, 227–233.
[21] Gerke, V. and Weber, K. (1985) J. Biol. Chem. 260, 1688–1695.
[22] Glenney, J.R. and Tack, B.F. (1985) Proc. Natl. Acad. Sci. USA 82, 7884–7888.
[23] Johnsson, N., Vandekerckhove, J., Van Damme, J. and Weber, K. (1986) FEBS Lett. 198, 361–364.
[24] Kretsinger, R.H. (1980) CRC Crit. Rev. Biochem. 8, 119–174.
[25] Hirata, F., Notsu, Y., Iwata, M., Parente, L., Di Rosa, M. and Flower, R.J. (1982) Biochem. Biophys. Res. Commun. 109, 223–230.
[26] Blackwell, G.J., Carnuccio, R., Di Rosa, M., Flower, R.J., Parente, L. and Persico, P. (1980) Nature 287, 147–149.
[27] Glenney, J.R. (1986) FEBS Lett. 192, 79–82.
[28] Moller, W. and Arions, R. (1985) FEBS Lett. 186, 1–7.
[29] Wierenga, R.K., Terpsha, P. and Hol, W.G.J. (1986) J. Mol. Biol. 87, 101–107.
[30] Mancheva, I., Kleinschmidt, T., Aleksiev, B. and Braunitzer, G. (1984) Hoppe-Seyler's Z. Physiol. Chem. 365, 885–894.
[31] Dijkstra, B.W., Kalk, K.H., Hol, W.G. and Drenth, J. (1981) J. Mol. Biol. 147, 97–123.
[32] Wirtz, K.W.A., Moonen, P., Van Deenen, L.L.M., Radhakrishnan, R. and Khorana, H.G. (1980) Ann. NY Acad. Sci. USA 348, 244–255.
[33] Owens, R.J. and Crumpton, M.J. (1984) Biochem. J. 219, 309–316.