The EMBO Journal vol.6 no.6 pp.1599–1604, 1987

The amino acid sequence of protein II and its phosphorylation site for protein kinase C; the domain structure Ca^{2+}-modulated lipid binding proteins

Klaus Weber, Nils Johnsson, Uwe Plessmann, Phuc Nguyen Van[1], Hans-Dieter Söling[1], Cristophe Ampe[2] and Joel Vandekerckhove[2]

Max Planck Institute for Biophysical Chemistry, 3400 Goettingen, [1]University of Goettingen, Department of Clinical Biochemistry, 3400 Goettingen, FRG, and [2]Rijksuniversiteit Gent, Laboratory of Genetics, B 9000 Gent, Belgium

Communicated by K.Weber

Protein II isolated from porcine intestinal epithelium is a Ca^{2+}-modulated lipid-binding protein. The amino acid sequence of porcine protein II reported here sheds new light on the properties of a multigene protein family which includes the tyrosine kinase substrates of the sarc gene (p36) and of the EGF-receptor (p35). The sequence consolidates the structural principle in which an amino-terminal tailpiece of variable length is followed by a core built from four internally homologous segments for those proteins in the 35–40 kd range. Sequence data also show that the core can now be described as two domains each containing one low and one high homology segment. This view accounts for two Ca^{2+} sites, lipid aggregation and F-actin bundling — when present — and suggests that properties of the cores in which protein II differs from p36 and p35 arise primarily from segments 1 and 2. The protease-sensitive tailpiece of protein II is very short and lacks the phosphorylatable tyrosine present in the larger tail domains of p36 and p35. It harbors, however, like the p36 domain, the major site for *in vitro* phosphorylation by the Ca^{2+}- and lipid-activated protein kinase C. In protein II this site is most likely threonine 6. The sequence alignment also explains why protein II does not interact with a unique p11, a property probably specific for p36. Our results further suggest that liver endonexin may reflect two protein species both closely related to protein II.

Key words: calcium/EGF-receptor/kinase C/phospholipid/phosphorylation/sarc gene

Introduction

The highly ordered microfilament organization of the intestinal brushborder is usually isolated in the presence of EGTA to suppress the Ca^{2+}-activated F-actin severing activity associated with villin, one of the two bundling proteins present in the microvillus cores (for review see Weber and Glenney, 1982; Mooseker, 1985). Thus it was a surprise when we found that cytoskeletal preparations obtained in the presence of Ca^{2+} contained two novel proteins which were easily released by the addition of EGTA. Proteins I and II are Ca^{2+}-binding proteins (Gerke and Weber, 1984). While protein II is a monomer, protein I is a hetero-tetramer as it seems to contain two copies of each p36 and p11. p36 was immediately shown to be identical to the major cytoplasmic target substrate (p36) of retrovirally coded tyrosine-specific protein kinases present in transformed fibroblasts (see Erikson and Erikson, 1980; Radke *et al.*, 1980). Once available in milligram quantities p36 was shown to interact in a Ca^{2+}-dependent manner with F-actin as well as non-erythroid spectrin (Gerke and Weber, 1984, 1985a).

During the last years several laboratories have developed the concept of a family of related Ca^{2+}-binding proteins present in various cells and tissues (Shadle *et al.*, 1985; Geisow *et al.*, 1986; Geisow, 1986; Huang *et al.*, 1986; Saris *et al.*, 1986; Weber and Johnsson, 1986; see also Moore and Dedman, 1982; Creutz *et al.*, 1983; Südhof *et al.*, 1984). Some of these proteins have been shown to display Ca^{2+}-modulated phospholipid-binding activity (see for instance Südhof *et al.*, 1984; Glenney, 1985, 1986a,b; Geisow *et al.*, 1986; Johnsson *et al.*, 1986a; Shadle and Weber, 1987) and may therefore be called Ca^{2+}/lipid proteins. Recent evidence suggests that such proteins are characterized by several internal repeats of homologous segments (see above references) and this concept was proved when the full cDNA sequences of p36 and p35 (lipocortin) were compared (Huang *et al.*, 1986; Kristensen *et al.*, 1986; Saris *et al.*, 1986). p35 adds to the functional complexity of this group of proteins. It is the major target of the tyrosine-protein kinase associated with the EGF receptor (Fava and Cohen, 1984; Sawyer and Cohen, 1985; Barun *et al.*, 1986; Pepinsky and Sinclair, 1986) and was shown to inhibit phospholipase A_2 (Wallner *et al.*, 1986). This ties in with the concept of steroid-induced inhibitors of phospholipase A_2 which may be responsible for the anti-inflammatory effect of steroids. Although such effects seem also to hold for p36 (Huang *et al.*, 1986; Johnsson *et al.*, 1986b), such studies have so far only been performed with pancreatic rather than intracellular phospholipase. In spite of common properties there are also noticeable differences particularly since relatively few comparative studies have been done. For instance whereas p36 and p35 both show Ca^{2+}-induced F-actin binding (Gerke and Weber, 1984, 1985a; Glenney, 1986a) protein II does not under similar experimental conditions (Gerke and Weber, 1984; Shadle *et al.*, 1985). p36 but not p35 forms the hetero-tetramer protein I type complex (Erikson *et al.*, 1984; Gerke and Weber, 1984, 1985a; Glenney and Tack, 1985; Barun *et al.*, 1986; Glenney, 1986a) with a unique p11, which by sequence is homologous to the S100 proteins (Gerke and Weber, 1985b; Glenney and Tack, 1985; Hexham *et al.*, 1986) found primarily but not exclusively in glial cells. It seems that a short amino-terminal tail, which differs distinctly for the two proteins, provides a binding site important for the interaction between p36 and p11 (Glenney, 1986b; Glenney *et al.*, 1986; Johnsson *et al.*, 1986a,b). In the absence of lipids protein II shows tighter binding of Ca^{2+} than p36 (Gerke and Weber, 1984; Shadle *et al.*, 1985). In the presence of some phospholipids two tight Ca^{2+} sites were reported for p36 (Glenney, 1986b). Lipid binding is always Ca^{2+} dependent. While p36 is found by immunofluorescence microscopy in the submembraneous actin-spectrin mesh (see for instance Gerke and Weber, 1984; Gould *et al.*, 1984), molecules related to protein II are thought to bind to the endoplasmic reticulum (Geisow *et al.*, 1984, 1986) and p35 has been shown to occur on endosomes (Sawyer and Cohen, 1985) and

to be somehow released from the living cell (for references see Wallner *et al.*, 1986). A peptide study has further raised the possibility that liver endonexin (p32.5), which seems particularly closely related to protein II, could be built from five rather than four homologous segments (Geisow *et al.*, 1986). To clarify the relation between the various proteins, to rule out possible proteolytic derivation of some members and to begin to understand their molecular differences, more chemical information is necessary. Here we report the amino acid sequence of porcine protein II from intestinal epithelium which we showed previously to bind Ca^{2+} and to have Ca^{2+}-modulated interaction with lipids but not with F-actin or spectrin (Gerke and Weber, 1984; Shadle *et al.*, 1985; Shadle and Weber, 1987). Our work was originally aimed at deciding whether protein II could be derived from either p36 or p35 by proteolysis. Our results rule out this possibility and we demonstrate some important differences in the sequences of these three proteins.

Results

The sequence of porcine protein II

Most of the sequence data were derived from two enzymatic digests of the carboxymethylated protein. Soluble tryptic and chymotryptic peptides were separated by two-dimensional fingerprinting. Peptides recognized as pure were subjected to stepwise sequencing. For manual analysis the DABITC technique was used, while larger fragments were directly characterized by automated sequencing. Mixtures of peptides as well as the pH 6.5 insoluble material were separated by h.p.l.c. and characterized as before. Major points of orientation along the emerging sequence were obtained from CNBr fragments separated by h.p.l.c. and characterized by automated sequencing. As protein II has no free amino terminus and a single tryptophan (Table I) it was cleaved with BNPS-skatol and without a separation step directly sequenced. This provided the 31 residues following the single tryptophan (Figure 1). The combined results pointed to a unique sequence, which could be aligned along the homologous p36 and p35 sequences.

The amino-terminal CNBr fragment was identified by its lack of a free N terminus. It was cleaved with trypsin and thermolysin to provide overlapping peptides suitable for sequencing. The amino-terminal tryptic peptide contained one lysine and two alanines and lacked a free amino terminus. When available in larger amounts (see below) we found it to be electrophoretically neutral at pH 6.5. Thus the amino-terminal blocking group must be neutral. It is most likely the acetyl group as in many eukaryotic proteins. The carboxy-terminal 10 residues of protein II (Figure 1) were present in the only tryptic peptide lacking both lysine and arginine. With this assignment the carboxyl end was fixed and shown to be very similar to the corresponding region of p36 (see Huang *et al.*, 1986; Kristensen *et al.*, 1986; Saris *et al.*, 1986).

Two minor ambiguities remain in the sequence proposed in Figure 1. In segment 1 the tryptic peptide covering residues 57–61 is overlapped only with the subsequent and not with the preceding sequence block. In segment 3 we lack a chymotryptic overlap between residues 240 and 241. However, the analogy with p35 and p36 leaves little doubt about the sequence proposed.

The sequence proposed in Figure 1 is in good agreement with the amino acid composition obtained by acid hydrolysis (Table I; see also Gerke and Weber, 1984). The calculated chemical mol. wt is ~35 700 and thus nearly 4 kd higher than the apparent mol. wt deduced by gel electrophoresis (Gerke and Weber, 1984). Similar discrepancies are known for p35 and p36 since the cDNA sequences predict values of ~39 000 and 38 400 respectively (Saris *et al.*, 1986; Wallner *et al.*, 1986). Nevertheless we have used the original nomenclature as it is widespread in the literature. Other names are lipocortin I and II or calpactin II and I for p35 and p36 respectively (Huang *et al.*, 1986; Kristensen *et al.*, 1986; Saris *et al.*, 1986; Wallner *et al.*, 1986).

Table I. Amino acid composition of protein II by hydrolysis (A) and sequence (B; see Figure 1)

	A	B
Asp	38.4	28
Asn		9
Thr	18.8	18
Ser	23.8	24
glu	33.3	21
Gln		13
Pro	2.4	2
Gly	25.8	25
Ala	24.6	25
Val	14.8	15
Met	9.2	10
Ile	17.4	19
Leu	33.3	33
Tyr	13.0	13
Phe	10.6	10
Lys	21.5	21
His	2.2	2
Arg	23.9	25
Trp	1.0	1
Cys	3.8	4

The sequence arrangement used in Figure 1 follows that used by Kristensen *et al.* (1986) in their comparison of p36 and p35. Differences are due primarily to the attempt to optimize the internal homology between the four core segments of the three proteins for which full sequences are available. In addition the separation of the variable amino-terminal tail from the core becomes more noticeable with the third sequence (see below). In the core domain of ~310 amino acids the number of identical residues is 170, 155 and 144 in comparison with p36/p35, protein II/p36 and protein II/p35 respectively. Thus p36 and p35 are the most closely related protein pair (55% sequence identity). Table II and Figure 1 also provide a direct measure of the homology for each of the four core segments in all three proteins. The homology values differ significantly. Segments 1 and 3 are low homology segments, while segments 2 and 4 are now identified as high homology segments (see Discussion). In all three proteins segments 2 are acidic, while segments 4 are basic.

Protein II has a very short and protease sensitive tail

During the earlier phase of this study several attempts were made to digest protein II under native conditions with different protease including trypsin, chymotrypsin and V8 protease. SDS gel electrophoresis always indicated no obvious change in the apparent mol. wt. When the sequence work was more advanced, it became obvious (Figure 1) that a protease sensitive tail if present on protein II could only be very small. Consequently a native digest of protein II (Figure 2A) with trypsin was analyzed by h.p.l.c. We identified a resistant core domain starting at residue 9. The core was sequenced for 37 residues. This result positioned segment 1 unambiguously and identified the tail as residues 1–8

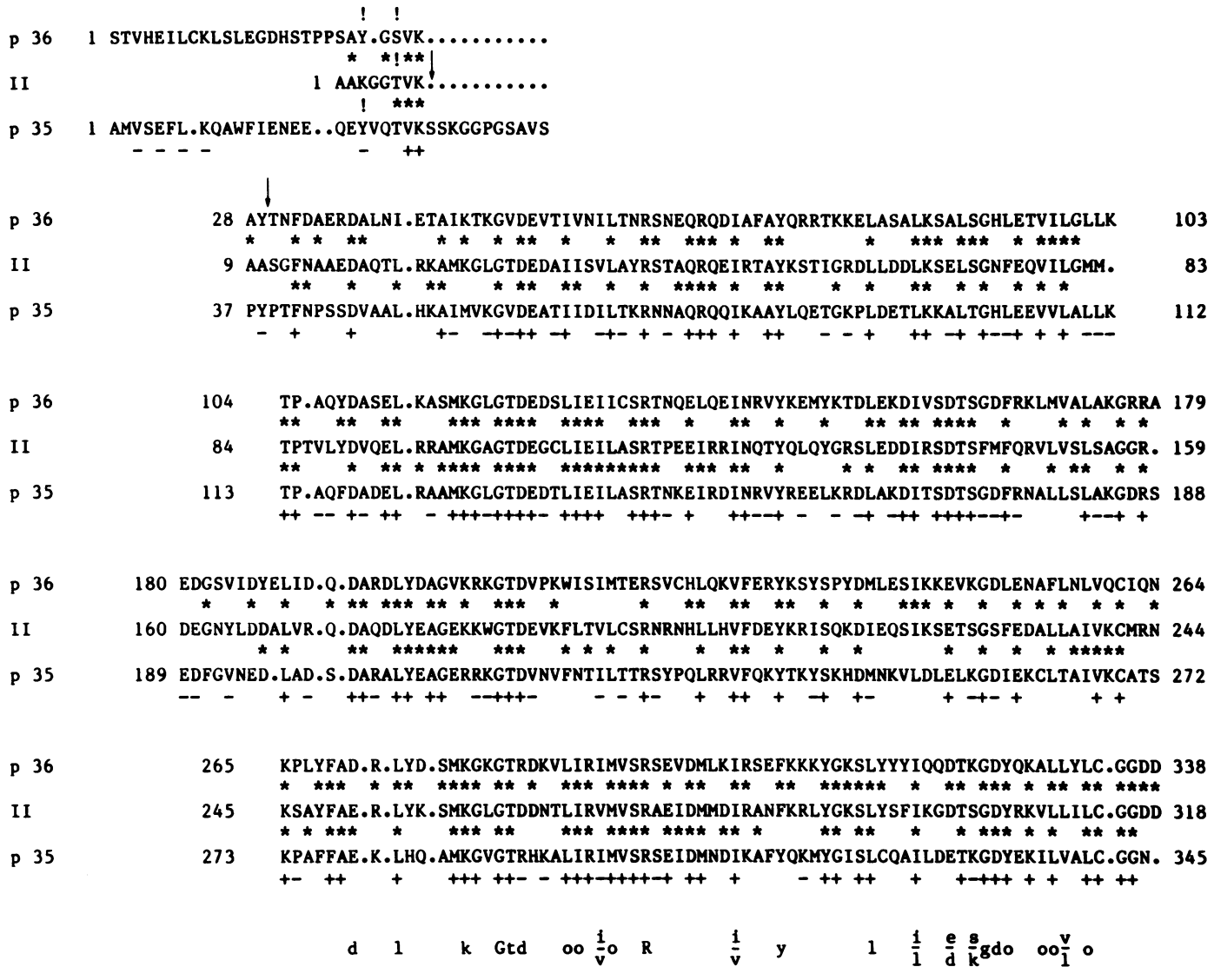

Fig. 1. Sequence alignments of porcine protein II (this study) with bovine p36 (Kristensen *et al.*, 1986) and human p35 (Huang *et al.*, 1986). Stars above and below the protein II sequence mark residues shared with p36 and p35 respectively. Symbols below the p35 sequence give identical residues in all three proteins (+) or residues where only p36 and p35 are identical (−). The arrangement emphasizes the structural principle of Ca^{2+}/lipid-binding proteins: a variable amino-terminal tail (**block 1**) followed by a protease-resistant core, which reflects four tandemly repeated segments with internal homology (**blocks 2−5**; for references see text). Dots allow for a better alignment. The bottom line indicates consensus residues present in all four core segments of the three proteins. Capital letters give invariant residues. Small letters indicate high but not absolute preference and circles point to residues being hydrophobic but not identical. Exclamation marks identify the known tyrosine phosphorylation sites of p36 and p35 (Glenney and Tack, 1985; Barun *et al.*, 1986) and the major serine phosphorylation site used by protein kinase C on p36 (Gould *et al.*, 1986; Johnsson *et al.*, 1986b) and protein II (see Results). Arrows mark the separation between the variable tail regions and the cores indicated by *in vitro* proteolysis with chymotrypsin on p36 (Glenney and Tack, 1985; Johnsson *et al.*, 1986a) and with trypsin on protein II (this study).

(Figure 1). The two small tryptic peptides (residues 1−3 and 4−8) and their overlap were also obtained from the h.p.l.c. column. Thus it seems that lysine 8 is the first site of trypsin action and the octapeptide released is subsequently cleaved at lysine 3. As shown in Figure 1 the start of the tryptic core domain of protein II aligns very well with the protease-resistant chymotryptic core of p36 (Glenney and Tack, 1985; Johnsson *et al.*, 1986a). This result lends further support to the suggested alignment of the three quite distinct tailpieces.

Kinase C phosphorylation site

During earlier work on p36 (Johnsson *et al.*, 1986b) we noted that protein II is a good *in vitro* substrate for the lipid activated protein kinase C (Figure 2B) but not for the catalytic subunit of cAMP-dependent kinase or for the calmodulin-dependent kinase (data not shown). Phosphorylated protein II was treated for various times with trypsin and the time course was monitored by gel electrophoresis and autoradiography (Figure 2C). There is a rapid loss of phosphate although protein II does not change its electrophoretic mobility and is not further cleaved. This behaviour corresponds to unphosphorylated protein II where trypsin removes only the first two peptides (see above, Figure 2A). Within these eight residues there is no serine and thus the sole threonine at residue 6 is the potential target for a serine/threonine specific kinase. In agreement, phosphoamino acid analysis (Figure 2B) done on the gel-excised protein yielded primarily phosphothreonine in addition to some phosphoserine, which may possibly arise from a contaminating fragment of the auto-phos-

phorylated kinase preparation. The alignment made in Figure 1 emphasizes that the rather different tail regions of the three proteins are targets for various protein kinases. Thus the major site for kinase C-dependent phosphorylation of p35 *in vitro* and *in vivo* is serine 25 (Gould *et al.*, 1986; Johnsson *et al.*, 1986b) and tyrosine phosphorylation is known to occur on tyrosines 23 and 20 of p36 and p35 respectively (Glenney and Tack, 1985; Barun *et al.*, 1986). Although protein II or very similar proteins are present in many cell types known to express p36 and/or p35 (our unpublished results; see also Geisow *et al.*, 1986), we are not aware of any report demonstrating a tyrosine phosphorylation of a p32 polypeptide. This may have a very simple explanation. No tyrosine is present in the very short tail of protein II (Figure 1), while the longer domains of p35 and p36 do contain tyrosine and can act as substrates of EGF or sarc kinase respectively.

Discussion

The chemical characterization of protein II consolidates the sequence principles of a group of recently discovered Ca^{2+}/lipid-binding proteins and provides more insight into their structural organization. A variable amino-terminal tailpiece precedes the core domain which consists of four internally homologous segments, each ~78 residues long. Molecular weight variability seems essentially restricted to the tailpiece, which is 8, 27 or 36 residues in protein II, p36 and p35 respectively. Protein II has the shortest tailpiece. This lacks a tyrosine residue and so cannot be phosphorylated by the EGF receptor kinase which acts on p35 or by retrovirally coded kinases which use p36. Two of the three Ca^{2+}/lipid-binding proteins display in the tail region a strong phosphorylation site for the lipid- and Ca^{2+}-dependent protein kinase C. The alignment suggested in Figure 1 brings threonine 6 of protein II (see Results) into a similar sequence environment as the major phosphorylation site of p36, which is serine 25 (Gould *et al.*, 1986; Johnsson *et al.*, 1986b). Although p35 seems a good substrate for kinase C the phosphorylation site has not been mapped (Barun *et al.*, 1986). The alignment in Figure 1 makes the tailpiece of p35 an obvious candidate for future experiments. The variability in sequence and in length of the three tailpieces conforms with the finding that the p11 polypeptide interaction site seems specific for p36 and is absent both in p35 and protein II (Gerke and Weber, 1984; Shadle *et al.*, 1985; Glenney *et al.*, 1986; Glenney, 1986a,b; Johnsson *et al.*, 1986a). p36 is isolated as a heterotetramer with p11 from intestinal epithelial cells, while protein II is a monomer (Gerke and Weber, 1984) and also cannot bind exogenously added p11 (Shadle *et al.*, 1985).

Within the core domain length variability is minimal (Figure 1). Although protein II shows a single residue deletion at the carboxyl end of segment 1 when compared with p36 and p35, this is counterbalanced three residues later by the addition of a single residue. At the carboxyl end of segment 2 protein II has a single deletion in comparison with the other two proteins and p35 has a unique deletion early in segment 3. Finally, at the carboxyl end p35 differs by a single deletion from protein II and p36, which are particularly closely related in this region. Thus the core domain length in all three proteins is kept constant at 309−311 residues (Figure 1).

With the third core sequence available, the number of residues identical in all four segments is restricted to an invariant glycine separated by 12 residues from an invariant arginine. A second glycine separated by 29 residues from the arginine is not, as previously thought, invariant (Geisow, 1986; Saris *et al.*, 1986). The second segment of protein II displays a phenylalanine in this position (see Figure 1). Because of this deviation particular emphasis was put on overlapping peptides to ascertain this feature of protein II. Along the four homologous segments the three proteins reveal several positions with a very strong preference

Table II. Relative homology in the four core segments of the three related proteins aligned in Figure 1

Number of identical residues	Segment number (residue length in parentheses)			
	1 (76)	2 (77)	3 (85)	4 (74)
p36 versus p35	39	52	37	42
II ersus p36	32	38	36	49
II versus p35	29	43	32	40
All three residues	23	34	22	35

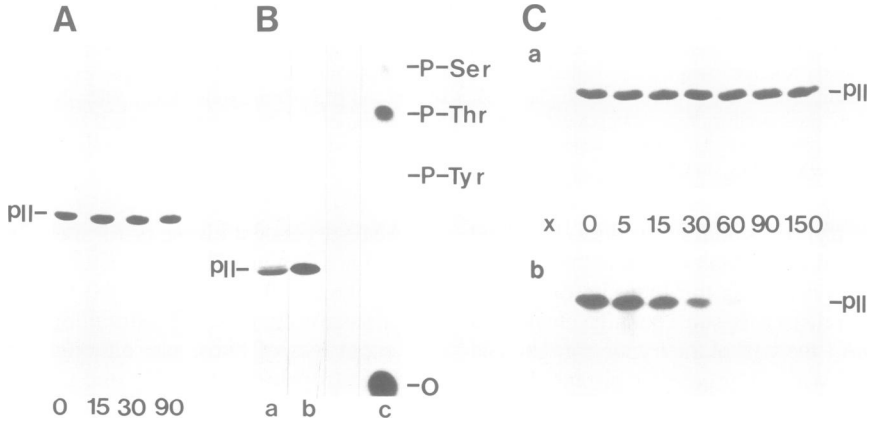

Fig. 2. Limited proteolysis of protein II and its phosphorylation by protein kinase C. Protein II was treated with trypsin (see Materials and methods) for 15, 30 and 90 min and then subjected to gel electrophoresis **(A)**. Although the protein seems resistant direct protein chemical data (see Results) show that trypsin removes the amino-terminal 8 residues. Phosphorylation of protein II by kinase C is shown in **B**. Slots **a** and **b** are the dye-stained gel and the autoradiograph respectively. The result of phospho-aminoacid analysis in strip c indicate >80% threonine phosphate in addition to some serine phosphate. O marks the origin of the thin-layer electrophoresis. *In vitro* proteolysis of the phosphorylated protein II by trypsin is shown in **C**. Protein II was phosphorylated with kinase C for 45 min at room temperature. After addition of EGTA to 1 mM trypsin was added and aliquots were removed after different times (see slots in C). The dye-stained gels are in **a** while corresponding autoradiographs are in **b**. Note the increasing loss of phosphate label in **b** and the resistant core in **a**. **Lane X** shows the equivalent part of the gel for the autophosphorylated kinase preparation.

for a certain amino acid type and these positions are summarized at the bottom of Figure 1; many of them involve hydrophobic residues. In a previous analysis of p35, particular emphasis was put on six hydrophobic residues close to the carboxyl end of each segment (Geisow, 1986). These are also present in protein II. It now becomes obvious that the highest hydrophobicity in this region belongs in all three residues to segment 4, where the subsequent cysteine−glycine−glycine sequence is also preserved. Future experiments have to decide whether this region could contribute to Ca^{2+}-induced phospholipid binding.

An unexpected result provided by the third sequence is a better understanding of the relative relation of the four internally homologous segments. As shown in Figure 1 the relative homology among the three proteins is distinct in the four segments. Segments 2 and 4 are high homology regions with 34 and 35 identical residues respectively. In contrast, segments 1 and 3 are low homology regions revealing 23 and 22 identical residues respectively (Table II). The theme of two repeating units each built from a low and a high homology segment has important implications as it could explain several previous observations. The core of p36 has only moderate Ca^{2+} binding but displays in the presence of certain phospholipids two tight Ca^{2+} sites (Gerke and Weber, 1984, 1985a; Glenney, 1985, 1986a,b; Johnsson *et al.*, 1986a). Protein II shows in the absence of lipids two Ca^{2+} sites which differ in affinity (Shadle *et al.*, 1985). In a dual domain structure each of the two units proposed could provide one Ca^{2+} site. The Ca^{2+}-modulated F-actin bundling activity displayed *in vitro* by p36 and p35 again requires two sites (Gerke and Weber, 1984, 1985a; Glenney, 1986a,b). Also, a possible function of Ca^{2+}-modulated lipid binding proteins in exocytosis (Geisow *et al.*, 1984, 1986) or liposome aggregation would argue for two interaction sites. Finally, as noted by Geisow (1986), phospholipase A_2 inhibition activity of certain preparations of lipocortin (Flower *et al.*, 1984; Hirata, 1984; Hirata *et al.*, 1984) thought to reflect p35 (Pepinsky and Sinclair, 1986) is retained in proteolytic fragments, which from their mol. wt of ~15 000 might correspond to two neighbouring segments, i.e. to one of the low−high homology units proposed above. The three Ca^{2+}/lipid-binding proteins lack the EF-hand Ca^{2+} site readily recognizable along linear sequences (Kretzinger and Creutz, 1986). Also the Ca^{2+}-modulated F-actin binding sites specific for p36 and p35 cannot yet be deduced from the linear sequences. The significantly higher relative homology displayed by p36/p35 in comparison with protein II locates primarily to segments 1 and 2 (Figure 1; Table II). This could indicate that unit 1 may account for those properties of the cores in which protein II differs from the other two members. In spite of current difficulties in locating the Ca^{2+} sites, our previous spectroscopic studies show a direct influence of Ca^{2+} binding on the environment of the single tryptophan present in the core of p36 and protein II (Gerke and Weber, 1985a; Shadle *et al.*, 1985; Johnsson *et al.*, 1986a). This is located to segment 3 where it occurs at a slightly different relative position in the loop region which also harbors the invariant glycine and arginine residue (Figure 1). Further experiments are needed to decide whether these tryptophans participate in the weak Ca^{2+}-binding site or are only influenced by it. The lack of F-actin binding for protein II (Gerke and Weber, 1984) is not understood. It could, however, be related to a relatively less basic nature of protein II in comparison with p36 and p35.

While it is obvious that the Ca^{2+}/lipid-binding proteins discussed here have evolved by gene duplication events from a sequence segment only 70−80 amino acids long, the number of such segments has been a matter of debate. Five related peptide sequences of the monomeric bovine liver endonexin lead to the prediction of a 5-fold segmented structure for this protein (Geisow *et al.*, 1986). These bovine sequences are now unambiguously aligned with porcine protein II. Allowing for 1−3 species-specific amino acid substitutions per peptide four of them correspond to defined stretches of porcine intestinal protein II. They cover, as earlier proposed, the same region within the segmented structure of the core. While lines 4 and 5 in Figure 3 of Geisow *et al.* (1986) clearly correspond to segments 1 and 4 respectively, lines 6 and 7 do not cover different segments as previously thought but both arise from the same region of segment 2. Evidence for segment 3 is, however, unambiguously seen by the short sequence given in line 3 of the original report. Although this heptapeptide was subsequently dropped in a review (Geisow, 1986) favouring a four-segment structure of endonexin it is identical to the corresponding region in segment 3 of protein II. The presence of two slightly different sequences both covering the same region of segment 2 of bovine liver endonexin seems to indicate that this protein preparation is not homogenous. It probably reflects two closely related protein species, each rather similar to protein II as also suggested by immunological results (Geisow *et al.*, 1984, 1986; Shadle *et al.*, 1985). Whether one of them truly reflects protein II requires corresponding sequences to be established either on porcine liver endonexin or on bovine protein II isolated from intestinal epithelium. Although these results indicate that the complexity of Ca^{2+}/lipid proteins of a particular tissue can be underestimated, they should not be taken as independent support for the suggestion that some cell types such as chromaffin granules may contain 11 different members of this protein family (Creutz *et al.*, 1983; Saris *et al.*, 1986). Proteolysis can clearly indicate a too high complexity. When encountered to a moderate extent it always involves the amino-terminal tail region and various proteases can yield differently processed derivatives as known for p36, p35 (Glenney and Tack, 1985; Johnsson *et al.*, 1986a; Barun *et al.*, 1986) and now also for protein II. Rigorous proof of molecular identity demands some chemical characterization. In our hands cleavage at the unique tryptophan followed by automated sequencing has been very helpful (Johnsson *et al.*, 1986a and Results) as at least the well-characterized proteins p36, p35 and II carry an amino-terminal blocking group. Such characterization done on other related proteins should lead to a consensus as to the number of different proteins and thus allow an appreciation of the cell type-specific expression patterns of the multigene family. It also would give a base from which one could explore whether some of the proteins bind to biomembranes distinct from the plasma membranes (Geisow *et al.*, 1986).

The complete cDNA sequences for the Ca^{2+}- and lipid-dependent protein kinases C (Parker *et al.*, 1986) show that these enzymes do not belong to the family of Ca^{2+}-phospholipid binding proteins discussed here. However the family of related proteins extends by various criteria to protein(s) of apparent mol. wt(s) ~68 000, which were first detected in lymphocytes (Owens *et al.*, 1984) and then in various tissues including liver (Owens *et al.*, 1984; Geisow *et al.*, 1986; Shadle *et al.*, 1985; Smith and Dedman, 1986). They could have more than four homologous segments or more likely use their extra mass to provide new autonomous domains of as yet unknown function. A recent report raises the possibility that an as yet poorly characterized 53-kd protein is immunologically related to the 32-kd protein(s) (Smith and Dedman, 1986). Our sequence result makes a direct precursor−product relation difficult to see, at least for protein II. It has a carboxy-terminal region nearly identical to that of p36, which was deduced by cDNA cloning. At the amino-terminal

end protein II seems to carry the acetyl blocking group also suggested for p36.

Earlier work shows that at least p35 is found in peritoneal exudates from various animals and that it seems to exist in intracellular as well as extracellular forms derived from various cell types (Flower *et al.*, 1984; Hirata *et al.*, 1984; see also Huang *et al.*, 1986 for other references). Although p36 (residues 61–63) and p35 (residues 42–44) each have one potential glycosylation site and protein II has even two (residues 124–126 and 244–246), there is no conservation of these putative sites along the otherwise well-aligned sequences. Thus it seems that p35 may be lost from certain cells by processes involving vesicularization and membrane loss rather than normal secretion.

Materials and methods

Protein II was purified from the mucosa of porcine small intestines as described by Gerke and Weber (1984). It is separated from protein I by chromatography on DE-52. While protein I passes through the column in the equilibration buffer previously specified, the protein II is retained and elutes at 20–35 mM NaCl upon gradient elution. Gel filtration on Sephadex G-100 provides a preparation which is electrophoretically homogeneous when analyzed by one-dimensional gel electrophoresis in the presence of SDS.

Protein II was carboxymethylated with iodoacetamide in 8 M urea or 6 M guanidine-HCl. Derivatized protein was digested with trypsin or chymotrypsin in 0.05 M NH_4HCO_3. Digests were lyophilized and processed for two-dimensional fingerprinting on paper as described. Peptides recognized to be pure by their amino acid composition and their Dansyl end group were subjected to sequence analysis. In manual analysis the DABITC technique (Chang *et al.*, 1978) was used; alternatively analysis was done by a gas-phase sequencer using standard procedures. Mixtures of peptides were subjected to reverse-phase h.p.l.c. separation on a C18 column and then characterized by automated sequencing. Derivatized protein II treated with CNBr was subjected to reverse-phase h.p.l.c. separation using a C4 column. Cleavage at the single tryptophan was done on derivatized protein with BNPS-skatole as described. It was freed from reagents by gel filtration in 50% acetic acid on Sephadex G25 and directly subjected to automated sequencing for 31 steps. Alternatively gel electrophoresis was used to separate the two fragments of apparent mol. wts 13 and 21 kd respectively. Fragments were recovered by electroelution and freed of detergent and dye. All enzymatic fragmentations were done by standard procedures.

Native protein II in 0.1 M NH_4HCO_3, pH 7.8, was treated at room temperature with trypsin at an enzyme to substrate ratio of 1:200. In the preparative experiment digestion was stopped after 30 min by lyophilization. The dried material was dissolved in 4 M guanidine-HCl and subjected to h.p.l.c. For analytical experiments digestion was for different time intervals and the resulting preparations were analyzed by gel electrophoresis (Figure 2).

The preparations of various protein kinases including protein kinase C from rat brain were obtained as before (Johnsson *et al.*, 1986b). For *in vitro* phosphorylation similar conditions were used as previously given for p36.

Acknowledgements

We thank W.Koch for expert technical assistance and purified protein II. N.J. was supported by a Max Planck predoctoral fellowship. P.N.V. was supported by a predoctoral fellowship from the Deutsche Forschungsgemeinschaft (project A3, SFB 236).

References

Barun,K.D., Misono,K.S., Lukas,T.J., Mroczkowski,B. and Cohen,S. (1986) *J. Biol. Chem.*, **261**, 13784–13792.
Chang,J.Y., Brauer,D. and Wittmann-Liebold,B. (1978) *FEBS Lett.*, **93**, 205–214.
Cohen,S. and Fava,R.A. (1985) *J. Biol. Chem.*, **260**, 12351–12358.
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.
Davies,A.A. and Crumpton,M.J. (1985) *Biochem. Biophys. Res. Commun.*, **128**, 571–577.
Erikson,E. and Erikson,R.L. (1980) *Cell*, **21**, 829–836.
Erikson,E., Tomasiewicz,H.G. and Erikson,R.L. (1984) *Mol. Cell. Biol.*, **4**, 77–85.
Fava,R.A. and Cohen,S. (1984) *J. Biol. Chem.*, **259**, 2636–2645.
Flower,R.J., Wood,J.N. and Parente,L. (1984) *Adv. Inflam. Res.*, **7**, 61–69.
Geisow,M.J. (1986) *FEBS Lett.*, **203**, 99–103.
Geisow,M.J., Childs,J., Dash,B., Harris,A., Panayotou,G., Südhof,T. and Walker,J.H. (1984) *EMBO J.*, **3**, 2969–2974.
Geisow,M.J., Fritsche,U., Hexham,J.M., Dash,B. and Johnson,T. (1986) *Nature*, **320**, 636–638.
Gerke,V. and Weber,K. (1984) *EMBO J.*, **3**, 227–233.
Gerke,V. and Weber,K. (1985a) *J. Biol. Chem.*, **260**, 1688–1695.
Gerke,V. and Weber,K. (1985b) *EMBO J.*, **4**, 2917–2920.
Glenney,J.R., Jr (1985) *FEBS Lett.*, **192**, 79–82.
Glenney,J.R., Jr (1986a) *Proc. Natl. Acad. Sci. USA*, **83**, 4258–4262.
Glenney,J.R., Jr (1986b) *J. Biol. Chem.*, **261**, 7247–7252.
Glenney,J.R., Jr and Tack,B.F. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7884–7888.
Glenney,J.R., Jr, Boudreau,M., Galyean,R., Hunter,T. and Tack,B. (1986) *J. Biol. Chem.*, **261**, 10485–10488.
Gould,K.L., Cooper,J.A. and Hunter,T. (1984) *J. Cell Biol.*, **98**, 487–497.
Gould,K.L., Woodgett,J.R., Isacke,C.M. and Hunter,T. (1986) *Mol. Cell. Biol.*, **6**, 2738–2744.
Hexham,J.M., Totty,N.F., Waterfield,M.D. and Crumpton,M.J. (1986) *Biochem. Biophys. Res. Commun.*, **134**, 248–254.
Hirata,F. (1984) *Adv. Inflam. Res.*, **7**, 71–78.
Hirata,F., Mastsuda,K., Notsu,Y., Hattori,T. and del Carmine,R. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4717–4721.
Huang,K.-S., Wallner,B.P., Mattaliano,R.J., Tizard,R., Burne,C., Frey,A., Hession,C., McGray,P., Sinclair,L.K., Chow,E.P., Browning,J.L., Ramachandran,K.L., Tang,J., Smart,J.E. and Pepinsky,R.B. (1986) *Cell*, **46**, 191–199.
Johnsson,N., Vandekerckhove,J., Van Damme,J. and Weber,K. (1986a) *FEBS Lett.*, **198**, 361–364.
Johnsson,N., Van,P.N., Söling,H.D. and Weber,K. (1986b) *EMBO J.*, **5**, 3455–3460.
Kristensen,T., Saris,C.J.M., Hunter,T., Hicks,L.J., Noonan,D.J., Glenney,J.R., Jr and Tack,B.F. (1986) *Biochemistry*, **25**, 4497–4503.
Kretsinger,R.H. and Creutz,C.E. (1986) *Nature*, **320**, 573.
Moore,P.B. and Dedman,J.R. (1982) *J. Biol. Chem.*, **257**, 9663–9667.
Owens,R.J., Gallagher,C.J. and Crumpton,M.J. (1984) *EMBO J.*, **3**, 945–952.
Mooseker,M.S. (1985) *Annu. Rev. Cell Biol.*, **1**, 209–241.
Parker,P.J., Coussens,L., Totty,N., Rhee,L., Young,S., Chen,E., Stabel,S., Waterfield,M.D. and Ullrich,A. (1986) *Science*, **233**, 853–859.
Pepinsky,R.B. and Sinclair,L.K. (1986) *Nature*, **321**, 81–84.
Radke,K., Gilmore,T. and Martin,G.S. (1980) *Cell*, **21**, 821–828.
Saris,C.J.M., Tack,B.F., Kristensen,T., Glenney,J.R., Jr and Hunter,T. (1986) *Cell*, **46**, 201–212.
Sawyer,S.T. and Cohen,S. (1985) *J. Biol. Chem.*, **260**, 8223–8236.
Shadle,P.J., Gerke,V. and Weber,K. (1985) *J. Biol. Chem.*, **260**, 16354–16360.
Shadle,P.J. and Weber,K. (1987) *Biochem. Biophys. Acta*, **897**, 502–506.
Smith,V.L. and Dedman,J.R. (1986) *J. Biol. Chem.*, **261**, 15815–15818.
Südhof,T.C., Ebbecke,M., Walker,J.H., Fritsche,U. and Boustead,C. (1984) *Biochemistry*, **23**, 1103–1109.
Wallner,B.P., Mattaliano,R.J., Hession,C., Cate,R.L., Tizard,R., Sinclair,L.K., Foeller,C., Chow,E.P., Browning,J.L., Ramachandran,K.L. and Pepinsky, R.B. (1986) *Nature*, **320**, 77–81.
Weber,K. and Glenney,J.R., Jr (1982) *Phil. Trans. R. Soc. B*, **299**, 207–214.
Weber,K. and Johnsson,N. (1986) *FEBS Lett.*, **203**, 95–98.

Received on March 5, 1987; revised on March 30, 1987