Tanford, C. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6527–6531.
Verjovski-Almeida, S., Kurzmack, M., & Inesi, G. (1978) *Biochemistry 17*, 5006–5013.
Weber, G. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3000–3003.
Yates, D. W., & Duance, V. C. (1976) *Biochem. J. 159*, 719–728.

Primary Structure of Bovine Calpactin I Heavy Chain (p36), a Major Cellular Substrate for Retroviral Protein-Tyrosine Kinases: Homology with the Human Phospholipase A_2 Inhibitor Lipocortin†

Torsten Kristensen,‡ Chris J. M. Saris,§ Tony Hunter,§ Lori J. Hicks,‡ Daniel J. Noonan,‡ John R. Glenney, Jr.,§ and Brian F. Tack*,‡

Department of Immunology, Scripps Clinic and Research Foundation, La Jolla, California 92037, and Molecular Biology and Virology Laboratory, The Salk Institute, San Diego, California 92138

Received April 25, 1986; Revised Manuscript Received June 9, 1986

ABSTRACT: An amplified Okayama–Berg plasmid cDNA library was constructed from total poly(A)+ RNA isolated from the Madin–Darby bovine kidney cell line MDBK. This library was screened with a partial murine calpactin I heavy chain (p36) cDNA clone, the identification of which was based on bovine p36 tryptic peptide sequences generated during the course of these studies. The largest p36 cDNA insert (p36/6 of 1.6 kilobase pairs) was fully sequenced by the dideoxy method. The DNA sequence of this insert had an open reading frame of 1014 base pairs and coded for a protein with a molecular weight of 38 481. The deduced protein sequence of 338 residues was concordant with 173 residue positions of p36 determined at the protein level. The 5′- and 3′-ends of p36/6 contained 54 and 307 base pairs of untranslated sequence, respectively. Examination of poly(A)+ RNA prepared from the Madin–Darby cell line indicated a p36 mRNA species of about 1.6 kilobases. Four regions of internal homology, each about 70 amino acid residues in length, were observed in the deduced protein sequence for p36. Thirty-three of the 70 residue positions were conserved in at least three of the four repeating units. A comparison of the derived amino acid sequence for bovine p36 with that previously determined for human lipocortin [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., & Pepinsky, R. B. (1986) *Nature (London) 320*, 77–81] revealed extensive homology (66% overall) and the presence of four repetitive regions in the lipocortin structure. Alignment of the four lipocortin repeats revealed that 32 of the 70 residue positions were conserved. Of these, 26 were observed in corresponding positions within the p36 consensus sequence. A highly conserved region within each of the repeating units of p36 and lipocortin further correlated with a 17 amino acid consensus sequence repeat [Geisow, M. J., Fritsche, U., Hexham, J. M., Dash, B., & Johnson, T. (1986) *Nature (London) 320*, 636–638] recently identified in several calcium-binding proteins, including p36, which interact with biomembranes. These data indicate that p36 and lipocortin are members of the same gene family.

Many of the oncogene products and growth factor receptors are known to be active protein-tyrosine kinases (Bishop & Varmus, 1982; Erikson et al., 1980; Hunter & Cooper, 1985; Hunter & Sefton, 1982; Sefton & Hunter, 1984). Among the major substrates for protein-tyrosine kinases in vivo are polypeptides with apparent molecular weights of 34 000–39 000 as estimated by SDS–PAGE[1] (Radke & Martin, 1979; Erikson & Martin, 1980; Radke et al., 1980; Cooper & Hunter, 1981; Erikson et al., 1981a,b; Fava & Cohen, 1984). These include at least two distinct but related proteins, which have recently been termed calpactins (for calcium-dependent phospholipid- and actin-binding proteins) (Glenney, 1986b). Calpactin I, a complex of p36 and p10, has been shown to be located in the cortical skeleton underlying the plasma membrane in fibroblasts (Greenberg & Edelman, 1983; Nigg et al., 1983; Lehto et al., 1983; Radke et al., 1983) and intestinal epithelial cells (Greenberg et al., 1984; Gould et al., 1986; Gerke & Weber, 1984). Calpactin I isolated from porcine or bovine intestine is a heterotetrameter comprised of two molecules each of p36 and p10 (Gerke & Weber, 1984; Glenney & Tack, 1985). Recently, three independent studies have demonstrated that the 10 000 molecular weight chain (p10) is closely related through its primary structure to the Ca^{2+}-binding S-100 proteins of the brain (Glenney & Tack, 1985; Gerke & Weber, 1985a; Hexham et al., 1986). Cal-

† This is Publication No. 4345 IMM from the Research Institute of Scripps Clinic. This work was supported by U.S. Public Health Service Grants AI 19222 and AI 22214 to B.F.T., CA 17096 and CA 28458 to T.H., and GM 32866 to J.R.G.

* Author to whom correspondence should be addressed.
‡ Scripps Clinic and Research Foundation.
§ The Salk Institute.

[1] Abbreviations: poly(A)+, poly(adenylic acid) containing; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; kDa, kilodalton(s); bp, base pair(s); EDTA, ethylenediaminetetraacetic acid; TLCK, N^{α}-*p*-tosyllysine chloromethyl ketone; TPCK, *N*-tosyl-L-phenylalanine chloromethyl ketone.

0006-2960/86/0425-4497$01.50/0

pactin I has been shown to bind Ca^{2+}, but the Ca^{2+}-binding activity has been assigned to the p36 moiety rather than p10 (Gerke & Weber, 1985b; Glenney, 1986a). Of particular interest with respect to Ca^{2+} regulation is that the affinity of calpactin I for Ca^{2+} can be greatly enhanced by anionic phospholipids in a manner similar to the effect of phospholipids and Ca^{2+} on protein kinase C activity (Glenney, 1985, 1986a).

Recent work on the calpactin heavy chain (p36) has shown that it is comprised of two discrete domains, an amino-terminal 3-kDa "tail" that contains the major site of tyrosine phosphorylation (Tyr-23) (Glenney & Tack, 1985) as well as the site of phosphorylation by protein kinase C (Ser-25) (Gould et al., 1986) and a carboxy-terminal core that binds Ca^{2+} and phospholipid (Glenney, 1986a). We report here the complete coding sequence and deduced protein sequence for bovine p36. These studies have indicated that p36 is composed of multiple domains with four internal repeating units, each about 70 amino acid residues in length. Pronounced sequence similarities between p36 and the human phospholipase A_2 inhibitor lipocortin (Wallner et al., 1986) were further observed, thereby identifying a similar 4-fold internal repeat for this molecule. Interestingly, the sequence comparison of p36 with lipocortin revealed greatest variation for the amino-terminal ends, suggesting a potentially different function for this region of each molecule.

Experimental Procedures

Isolation of p36. Calpactin I was isolated from bovine intestine as previously described (Glenney & Tack, 1985). A proteolytic fragment of p36 with a molecular weight of 33 000 (33 kDaII), present in variable amounts and cross-reacting with anti-p36 antiserum (Glenney, 1986a), was identified as a separate peak on a Sephacryl S-200 (Pharmacia Fine Chemicals, Piscataway, NJ) column. Fractions containing this fragment were pooled, dialyzed, and further fractionated on a Whatman DE-52 column as described (Glenney, 1986a). A related 33-kDa core (33 kDaI) was also produced by digestion of calpactin I with limited amounts of TLCK-treated chymotrypsin (Sigma, St. Louis, MO) and purified as previously described (Glenney et al., 1986).

Purification and Sequence Analysis of Tryptic Peptides Derived from the p36 Core Fragment 33 kDaII. Twenty-five nanomoles of salt-free 33 kDaII was reduced in 0.1 M sodium pyrophosphate buffer (pH 8.0) containing 6.0 M guanidinium chloride, 1.0 mM dithiothreitol, and 1.0 mM EDTA for 2 h at 37 °C and radioalkylated with [2-^{3}H]iodoacetic acid (New England Nuclear, Boston, MA; 232.4 mCi/mmol). The protein was then desalted on a prepacked Sephadex G-25 M column (PD-10) (Pharmacia Fine Chemicals) equilibrated with 0.15 M ammonium bicarbonate (pH 8.25), digested with 1% (w/w) TPCK-treated trypsin (Worthington, Freehold, NJ) overnight at room temperature, and lyophilized. The tryptic digest was fractionated on a Pharmacia Mono Q column (Pharmacia Fine Chemicals) as described in the legend to Figure 1A. Selected peptide pools were further fractionated by reverse-phase high-pressure liquid chromatography (HPLC) on a Vydac C4 column (see legend to Figure 1B). Selected peptides were sequenced on a Beckman 890M microsequenator by using the 0.1 M Quadrol program. Phenylthiohydantoin-(PTH-) amino acids were analyzed on an analytical Du Pont Zorbax ODS column essentially as previously described (Zimmerman & Pisano, 1977). One-tenth volume of each PTH residue derived from a ^{3}H-labeled peptide was analyzed for radioactivity on a Beckman 6800 scintillation counter to identify PTH-[^{3}H](carboxymethyl)cysteine. Sequence analysis of the amino terminus of the 33-kDaI core fragment of p36 was performed in a manner analogous to that described for the above peptides.

Isolation and Sequence Analysis of Putative p36 cDNA Clones. A cDNA library was prepared according to the method of Okayama and Berg (1982, 1983) from poly(A)+ selected RNA isolated from the Madin–Darby bovine kidney cell line MDBK (Madin & Darby, 1958). The library was screened on blotted duplicate filters (Birnboim & Doly, 1979) as described (Grunstein & Hogness, 1975) with 1 μg of a ^{32}P-labeled (nick-translated) murine p36 cDNA clone (Saris et al., 1986) as the probe. Colonies that hybridized with this probe were purified by standard methods and subjected to restriction enzyme analysis. Preparative isolation of one of the largest putative p36 cDNA inserts (p36/6) of about 1600 bp was performed essentially as described (Maniatis et al., 1982) by digesting the plasmid with the restriction endonuclease *Bam*H1 (Boehringer-Mannheim, Mannheim, FRG). p36/6 was separated from vector DNA by preparative agarose gel electrophoresis, further purified by phenol/chloroform extraction and ethanol precipitation, and then sequenced by employing shotgun cloning into M13mp8 (Bankier & Barrell, 1983; Heidecker et al., 1980) and the dideoxynucleotide termination procedure of Sanger et al. (1977). All nucleotide sequence data were analyzed with the DBCOMP, DBUTIL, and Diagon programs (Staden, 1980, 1982a,b).

RNA Isolation and Analysis. Total cytoplasmic RNA was isolated from the Madin–Darby cell line, poly(A)+ selected, and analyzed by Northern blot analysis (Maniatis et al., 1982) with the 551-bp nick-translated murine p36 probe.

Results and Discussion

Amino Acid Sequence Analysis of 33-kDa Core Fragments of p36. The purity of p36 core fragments (33 kDaI and 33 kDaII) investigated in these studies was greater than 90% as assessed by SDS–PAGE (data not shown) and amino-terminal sequence analysis. In an earlier report (Glenney & Tack, 1985), we presented sequence data (underlined in Table I) for (a) positions 1–8 of the p36 core fragment 33 kDaI, produced on digestion of calpactin I with limited amounts of chymotrypsin, (b) positions 10–23 of the blocked N-terminal p36 chymotryptic peptide (NTp36CT) obtained from the above chymotryptic digest, and (c) positions 1–12 of the p36 core fragment 33 kDaII obtained as a byproduct during the isolation of calpactin I. Reanalysis of 33 kDaI confirmed and extended the former sequence by 38 amino acid residues (see Table I) and thereby increased the number of contiguous residue positions elucidated for bovine p36 at the protein level to 66 (i.e., residue positions 10–75 in Figure 3).

The separation of tryptic peptides of 33 kDaII by chromatography on a Pharmacia Mono Q column is shown in Figure 1A. The amino acid sequences of selected peptides obtained after a final step of purification by reverse-phase HPLC on a Vydac C4 column (see Figure 1B for a representative elution profile) are shown in Table I. A total of 11 tryptic peptides derived from peptide pools B, C, and F (see legend to Figure 1A) were sequenced and collectively spanned 145 residues of this 33 000 molecular weight core fragment of p36. The overall yield of each purified peptide varied between 4% and 48% (data not shown).

The framework-protein sequence for bovine p36 established in these studies provided absolute confirmation for the identity of the first partial p36-specific cDNA clone (Saris et al., 1986). This clone was obtained from a cDNA library constructed in λgt11 with size-selected poly(A)+ RNA prepared from the ANN-1 Abelson murine leukemia virus transformed NIH3T3 cell line by screening with a rabbit polyvalent antiserum against

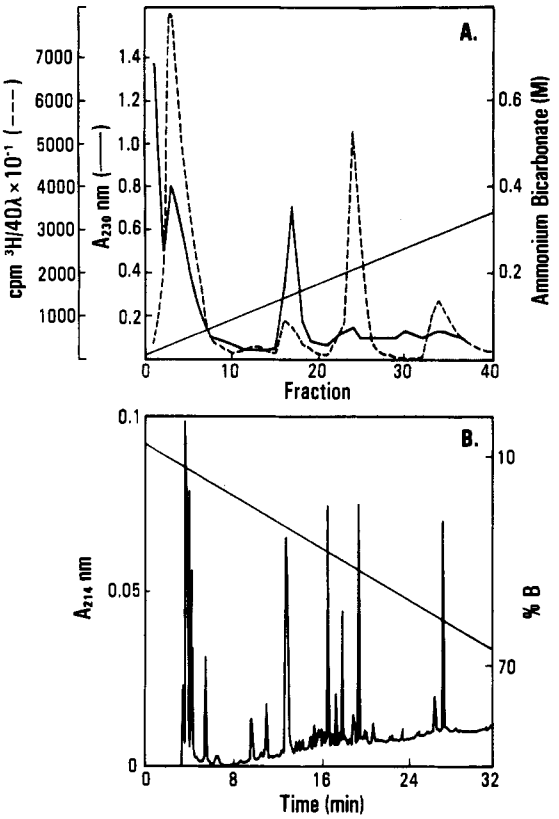

FIGURE 1: (A) Fractionation of tryptic peptides from a 25-nmol digest of the p36 core fragment 33 kDaII on a Pharmacia Mono Q column. The column was equilibrated in 0.01 M ammonium bicarbonate (pH 8.25). Following an initial wash with 20 mL of the equilibration buffer, peptides were eluted with a linear gradient of ammonium bicarbonate to a limit concentration of 0.5 M. The flow rate was maintained at 1.0 mL/min for 120 min (time of gradient elution), and 1.0-mL fractions were collected. The separation was monitored by measuring the absorbance at 230 nm (—) and by measuring the cpm of ^{3}H (---) in 40-μL aliquots of each fraction. Individual fractions were combined into pools as follows: fractions 3 and 4 = pool B, fractions 5 and 6 = pool C, and fractions 21–26 = pool F. (B) Separation of tryptic peptides present in pool C by reverse-phase HPLC on a 8 × 250 mm Vydac C4 column (5 μm; 300-Å pore size) equilibrated in 92% solvent A (0.1% TFA) and 8% solvent B (0.06% TFA, 85% acetonitrile). Peptides were eluted with a linear gradient from 8% to 70% solvent B applied over 90 min at a flow rate of 1.0 mL/min. Peptides were named (see Table I) Tp36, followed by the letter of the Pharmacia Mono Q column pool from which they were derived and the number of the peak collected from the reverse-phase C4 column.

purified chicken p36 (Cooper & Hunter, 1983). A total of six bovine p36 tryptic peptides (Tp36-B-12, -B-20A, -B-20B, -B-41, -C-39, and -F-15; see Table I) spanning 81 amino acid residues were located within the amino acid sequence deduced from the murine cDNA sequence. Only one substitution was noted, which corresponded to a glutamic acid/aspartic acid interchange at the amino terminus of the bovine p36 tryptic peptide Tp36-B-20B when compared to the homologous position in the derived mouse p36 sequence. In constrast, when the above bovine peptide sequences were aligned with homologous regions in the human lipocortin structure (Wallner et al., 1986), 45 substitutions were observed.

Northern Analysis of Bovine RNA with a Murine p36-Specific Probe. When an *Eco*RI restriction fragment (551 bp) of the above murine p36 clone was used as a specific hybridization probe, a bovine p36 mRNA species of approximately 1.6 kb was detected in RNA from the Madin–Darby cell line. The size of bovine p36 mRNA compared favorably with the corresponding mRNA present in human A431 cells, the ANN-1 cell line, and various murine tissues including

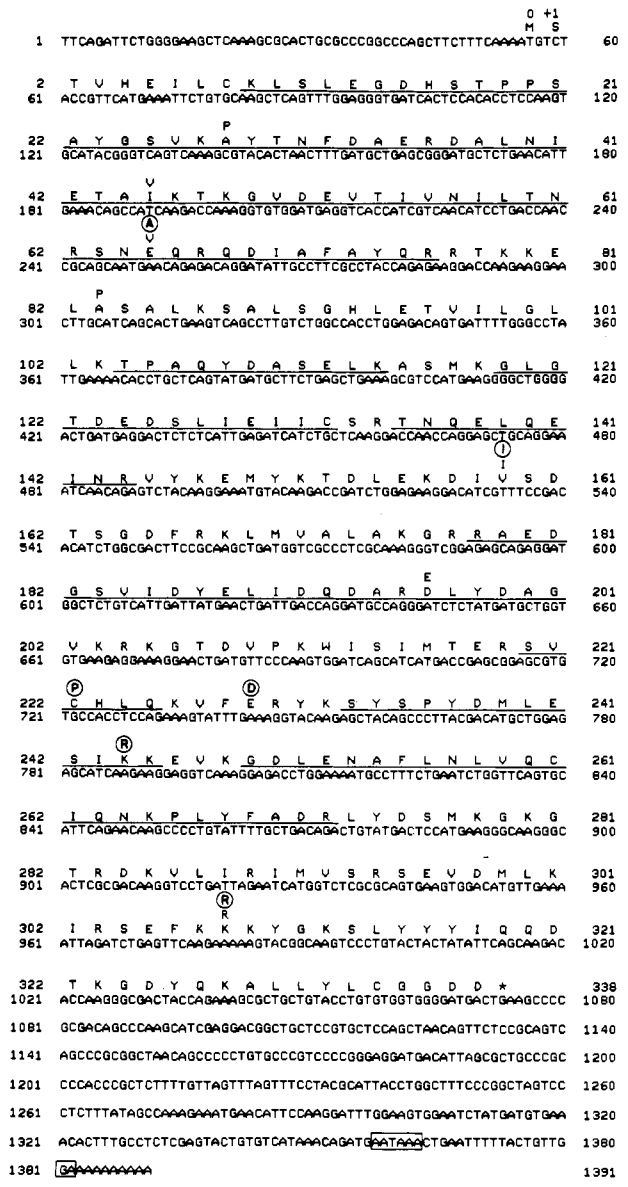

FIGURE 2: Complete nucleotide coding sequence of bovine p36 and deduced amino acid sequence in single letter code. Underlined regions of the deduced sequence of p36 correspond to sequences determined at the protein level (see Table I). The raised letters correspond to amino acid differences observed on comparing the primary structure of murine p36 (Saris et al., 1986) with that of bovine origin. The raised letters that are circled represent differences between human (Huang et al., 1986) and bovine p36 sequences. The polyadenylation recognition signal AATAAA is boxed.

kidney, small intestinal epithelium, lung, and thymus (Saris et al., 1986).

Isolation and Sequence Analysis of a Full-Length Bovine p36-Specific cDNA Clone. The Madin–Darby bovine kidney cell library (30 000 recombinants) was plated out at a density of 8000 colonies per nitrocellulose filter and screened with the *Eco*RI cut murine p36 cDNA clone. One of the largest putative p36 cDNA inserts (p36/6) was oligomerized and sonicated to generate random DNA fragments, which were blunt-end ligated into the *Sma*I site of M13mp8. Following transfection into JM101, templates were prepared from 45 transformants and used for sequence analysis. This approach yielded a contiguous sequence of 1377 bp exclusive of the poly(A) tail. On average, each base position was sequenced 2 times in both the (+) and (−) strands. Translation of the nucleotide sequence of p36/6 indicated an open-reading frame of 1014 bp (Figure 2). The deduced protein sequence (338

Table I: Sequence Data for the N-Terminal Chymotryptic Peptide (NTp36CT) of p36, the N-Termini of p36 Core Fragments 33 kDaI and kDaII, and Tryptic (T) Peptides Derived from 33 kDaII[a]

fragment peptide			amino acid residues	residue positions in p36 (see Figure 3)
NTp36CT		Leu Ser Leu Glu Gly Asp His Ser Thr Pro Pro Ser Ala Tyr Gly Ser Val Lys Ala Tyr Thr Asn Phe Asp Ala Glu Arg Asp Ala Leu		
33 kDaI		Asn Ile Glu Thr Ala Ile Lys Thr Lys Gly Val Asp Glu Val Thr Ile Val Asn Ile Leu	66	10–75
33 kDaII		Thr Asn Arg Ser Asn Glu Gln Arg Gln Asp Ile Ala Phe Ala Tyr Gln		
Tp36-B-12		Ser Val Cys His Leu Gln Lys	7	220–226
Tp36-B-20	A.	Thr Asn Gln Glu Leu Gln Glu Ile Asn Arg	10	135–144
	B.	Asp Leu Tyr Asp Ala Gly Val Lys Arg	9	196–204
Tp36-B-21		Thr Pro Ala Gln Tyr Asp Ala Ser Glu Leu Lys	11	104–114
Tp36-B-28		Gln Asp Ile Ala Phe Ala Tyr Gln Arg	9	68–76
Tp36-B-41		Ser Tyr Ser Pro Tyr Asp Met Leu Glu Ser Ile Lys Lys	13	233–245
Tp36-C-17		Ala Tyr Thr Asn Phe Asp Ala Glu Arg	9	28–36
Tp36-C-29		Ala Tyr Thr Asn Phe Asp Ala Glu Arg Asp Ala Leu Asn Ile Glu Thr Ala	17	28–44
Tp36-C-39		Gly Asp Leu Glu Asn Ala Phe Leu Asn Leu Val Gln Cys Ile Gln Asn Lys Pro Leu Tyr Phe Ala Asp Arg	24	249–272
Tp36-F-15		Arg Ala Glu Asp Gly Ser Val Ile Asp Tyr Glu Leu Ile Asp Gln Asp Ala Arg	18	178–195
Tp36-F-19		Gly Leu Gly Thr Asp Glu Asp Ser Leu Ile Glu Ile Ile Cys	14	119–132
			total 173	

[a] One to three nanomoles of each sample was sequenced in a Beckman 890M microsequenator. The average initial yield was 64%. A repetitive yield of 97% was generally observed. Underlined residues are those determined in an earlier study, which correspond to the amino-terminal region of p36 (i.e., residues 10–37). Peptide Tp36-B-20 gave an equimolar double sequence that was interpretable from inspection of the deduced protein sequence. Peptides Tp36-C-17 and Tp36-C-19 overlap with each other and correspond to residues 28–44 of p36.

residues) was fully concordant with all p36 protein sequences (see Table I) determined in these and other studies (Glenney & Tack, 1985).

The predicted sequence of bovine p36 is identical in length and shows 98% homology with the predicted murine p36 sequence (Saris et al., 1986). There are seven single amino acid changes (see Figure 2) of which four were highly conserved. Within the coding region the nucleotide sequences of murine and bovine cDNA clones show 91% homology. The 3′ untranslated regions of the bovine and murine clones are 307 and 280 bp, respectively. Three small apparent deletions in the murine 3′ untranslated region account for the size difference of 27 bp. Not counting the deletions, the homology between the murine and bovine 3′ untranslated regions is about 70%. Comparison of the bovine p36 sequence with that of the corresponding human protein (Huang et al., 1986) indicated six amino acid changes (see Figure 2), four of which were highly conserved.

In an earlier report (Glenney & Tack, 1985), we suggested that the N-terminus of p36 was blocked. We were able, however, to isolate the putative N-terminal tryptic peptide, which was estimated from compositional data to be a lysyl peptide, approximately nine residues in length. The published composition of this peptide (Thr, Ala, Cys, Val, Ile, Leu, His, Glu or Gln, Lys) is in close agreement (eight out of nine matches) with the first nine amino acid residues specified by the deduced sequence of p36 if the translational initiation codon (AUG) at bp positions 55–57 (see Figure 2) is used. As the predicted N-terminal residue of p36 is a serine, we suggest that this blocked residue was missed earlier, probably due to high hydrolytic loss, and that alanine was overestimated due to high background.

Homologous Regions within the p36 Sequence. Analysis of the complete p36 sequence by Diagon, an interactive graphics program for comparing and aligning DNA or protein sequences, indicated the presence of four homologous regions (see Figure 3A). Each region (labeled I–IV in Figure 3A) was approximately 70 amino acids in length, the first of which commenced with the glutamic acid residue at position 35 (see Figure 2). By utilizing a grid overlay to simplify the interpretation of the p36 diagon, it was possible to see that the rank order of relatedness between the repetitive units of p36 was II/IV > I/II > III/IV, II/III, I/III > I/IV. A more detailed residue by residue comparison of each repeating unit in p36 is shown in Figure 4. While significant sequence variability was apparent between the individual 70 amino acid repeats, 33 residue positions were conserved in at least 3 of the 4 repeats (see p36 consensus sequence in Figure 4). Within the p36 consensus sequence, the most highly clustered of the conserved residue positions (underlined in Figure 4) corresponded to residues 45–63, 117–135, 202–220, and 277–295 in repeats I–IV, respectively. Very recently, a 17 amino acid consensus sequence repeat was identified in several Ca^{2+}-binding proteins, which have an unusual mode of interaction with biomembranes (Geisow et al., 1986). The consensus sequence Lys-Gly-(fob)-Gly-Thr-Asp-Glu-(var)-(var)-Leu-Ile-(fil)-Ile-Leu-Ala-(fil)-Arg was deduced from analysis of two calelectrin (a 34-kDa protein isolated from the electroplax tissue of the ray *Torpedo marmorata*) peptides, one peptide each from two preparations of p36 (obtained from pig mesenteric lymph nodes

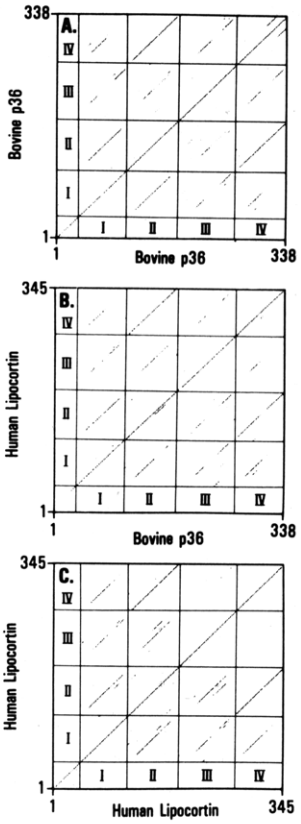

FIGURE 3: Diagon analysis of the deduced amino acid sequences of bovine p36 and human lipocortin (Wallner et al., 1986). In panel A, p36 is compared with itself; in panel B, p36 is compared with lipocortin; in panel C, lipocortin is compared against itself. Regions of internal homology in both structures are indicated by Roman numerals, and their boundaries are as shown in Figure 4. A percent score of 132 and a span length of 11 were used for generating these diagrams.

and intestinal epithelia) and five peptides of p32.5 purified from bovine liver (also referred to as endonexin). This sequence shows a high degree of similarity with the most conserved regions (underlined in Figure 4) in each of the four repeating units of bovine p36.

Homology between Bovine p36 and Human Lipocortin Sequences: Identification of a Similar 4-Fold Repeat Structure in Lipocortin. A comparison of the deduced protein sequence of bovine p36 (338 amino acid residues) with the previously determined human lipocortin sequence [345 residues; see Wallner et al., (1986)] by Diagon analysis (Figure 3B) indicated a high degree of overall sequence similarity and the presence of four repeating units in the lipocortin structure (see also Figure 3C, where lipocortin is compared against itself). Interestingly, a much reduced level of homology was indicated for the amino-terminal ends of each molecule (see Figure 3B), i.e., positions 1–34 for p36 and positions 1–43 for lipocortin. A considerably more detailed comparison for the full structures of p36 and lipocortin is shown in Figure 4, where the two sequences have been aligned invoking minimal gapping. In order to evaluate the relatedness of repeating units I–IV in p36 and lipocortin, a residue by residue comparison was made. Corresponding positions were considered homologous only if they were occupied by the same amino acid or chemically similar residues. By these criteria, the overall homology between repeating units I, II, III, and IV in p36 compared to lipocortin was 62%, 83%, 60% and 72%, respectively. Thirty-two residue positions were conserved in at least three of the four lipocortin repeats (see lipocortin consensus sequence in Figure 4), 26 of which were observed in corresponding positions within the p36 consensus sequence. The consensus sequence reported by Geisow et al. (1986) was also present in each of the four repeating units of lipocortin.

An attempt was made to align the amino-terminal ends of p36 and lipocortin (see Figure 4), for which the overall homology (35%) is much lower than for the repeating units. This was accomplished by aligning the N-terminal serine of p36 with the corresponding position (alanine) in lipocortin and introducing a ten-residue deletion following residue position 11 in the p36 sequence and a one-residue deletion following position 7 in lipocortin. Two interesting features emerged from this method of alignment. First, the major sites of pp60$^{v\text{-}src}$

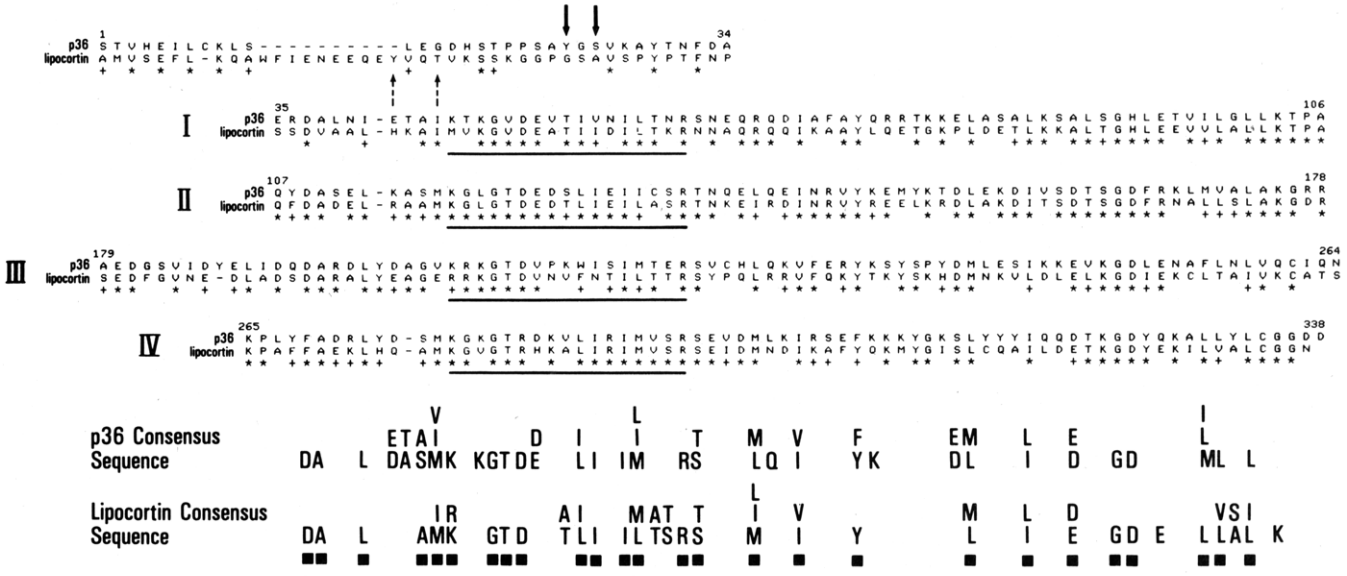

FIGURE 4: Alignment of bovine p36 and human lipocortin sequences. The symbols * and + denote identically placed and chemically similar residues, respectively, between both molecules, where E = D, A = S = T, Q = N, F = Y = H, K = R, and V = I = L = M. Consensus sequences for the repeating units of p36 and lipocortin are shown in bold single-letter code. The symbol ■, shown beneath the lipocortin consensus sequence, denotes those residue positions that are invariant between the lipocortin and p36 consensus sequences. Underlined regions in repeats I–IV of both molecules correspond to the 17 amino acid consensus sequence identified by Geisow et al. (1986). The solid arrows denote established sites of phosphorylation (Tyr-23 and Ser-25) in bovine p36, and the dashed arrows indicate possible sites of phosphorylation (Tyr-20 and Thr-23) in human lipocortin.

and protein kinase C phosphorylation in p36, Tyr-23 (Glenney & Tack, 1985) and Ser-25 (Gould et al., 1986), respectively, are not present in the corresponding positions of the lipocortin structure. While the amino-terminal ends of p36 and lipocortin may be aligned in such a manner as to place Tyr-20 and Thr-23 of lipocortin in positions homologous with the corresponding sites of phosphorylation in p36, more extensive gapping is required to do so. Second, the sequence residing between positions 10 and 18 in lipocortin is related to the nine-residue segment preceding it (positions 1–9) and may therefore have arisen by duplication.

The elucidation of the primary structure for bovine p36 should allow some assessment of those structural features required for phospholipid and Ca^{2+} binding (Glenney, 1986a) as well as interaction with actin and spectrin (Gerke & Weber, 1984). Previous studies have demonstrated that, in the presence of phosphatidylserine, calpactin I binds 2 mol of Ca^{2+} per mole of the p36 subunit and that all of this activity resides within the 33-kDa core fragment of p36 (Glenney, 1986a). The phospholipid requirement distinguishes calpactin I from the classical intracellular Ca^{2+}-binding proteins such as calmodulin. Consistent with this difference in phospholipid dependency is the fact that we find very little homology between p36 sequence and those regions of calmodulin that confer Ca^{2+}-binding activity. The presence of four structural repeating units in p36 is consistent with multiple Ca^{2+}-binding sites, where each domain may contribute half of a binding site. The consensus sequence identified by Geisow et al. (1986) for several members of an emerging new family of Ca^{2+}-binding proteins that interact with biomembranes may constitute part of a Ca^{2+}-binding site (despite the lack of homology with the E-F hand) or a phospholipid-binding site. Kretsinger and Creutz (1986) have raised the very intriguing possibility that one or more members of the p36–calelectrin–p32.5 (endonexin)–lipocortin (p35?)–synexin family may be involved in the coupling events in secretory cells.

Previous work has shown that at least two distinct members of the calpactin family, I (p36–p10) and II (p35), can be detected in three cell lines (Glenney, 1986b). These proteins have been characterized as major substrates of protein-tyrosine kinases, with Ca^{2+} regulating both their ability to be phosphorylated in vitro (Fava & Cohen, 1984; Glenney, 1985) and their interaction with the membrane. Although most antibody preparations react with only one of the two calpactins detected in human A431 cells, antibodies to the bovine intestinal calpactin I heavy chain (p36) also react weakly with calpactin II (p35) (Glenney, 1986b). Consistent with this, peptide mapping of ^{125}I-labeled p36 and p35 revealed that two peptides (of 16 total) apparently comigrate, suggesting regions of sequence identity (Glenney, 1986b). Clearly lipocortin is a member of this family of proteins. Since the amino acid sequences of bovine and murine p36 are so highly conserved (98%), it is unlikely that lipocortin is the human form of the calpactin I heavy chain. It is possible that lipocortin may represent calpactin II (p35). What would remain unexplained, however, is that lipocortin has been characterized as an extracellular protein (Hirata, 1981; Hattori et al., 1983; Ruthhut et al., 1983) whereas both calpactin I and calpactin II are intracellular proteins (Radke & Martin, 1979; Erikson & Martin, 1980; Radke et al., 1980, 1983; Cooper & Hunter, 1981, 1982; Erikson et al., 1981a,b; Fava & Cohen, 1984; Chen & Chen, 1981; Greenberg & Edelman, 1983; Nigg et al., 1983; Lehto et al., 1983; Greenberg et al., 1984; Gould et al., 1986). This situation is reminiscent of another Ca^{2+}-dependent cytoskeletal protein, gelsolin. Whereas the biological activity of gelsolin (M_r 90000) is entirely consistent with its regulation of actin filaments inside the cell, a variant gelsolin (Yin & Stossel, 1979; Wang & Bryan, 1981; Rouayrenc et al., 1984) is also an abundant protein of blood plasma (Harris & Weeds, 1983; Yin et al., 1984).

The amino acid sequence homology between bovine p36 and lipocortin is not of a uniform nature. For example, within the four repeating units of bovine p36, the homology with lipocortin varies between 60% and 83%. A more striking difference, however, is in the region of the 3-kDa amino-terminal "tail" domain where there is significantly less similarity (35%) between the two proteins and lipocortin is nine amino acids longer than p36. Thus, it seems likely that whatever function is expressed in this tail domain is altered between these two proteins. This may be related to the difference in association state of calpactin I and calpactin II [calpactin I is associated with the 10000 molecular weight p10 component whereas calpactin II is not; see Glenney (1986b)]. Consistent with this is the observation that the tail region of p36 comprises the site through which the association with the p10 component occurs (Glenney et al., 1986). Studies are in progress to determine whether lipocortin is calpactin II and whether calpactin I has any ability to inhibit phospholipase activity.

ACKNOWLEDGMENTS

We are indebted to Dr. Lois Tack (The Salk Institute, La Jolla, CA), who first observed the presence of repeating units in the deduced protein sequence of p36, for alignment of the 4-fold repeats in the bovine p36 and human lipocortin structures. This was accomplished without the aid of a computer, requiring only the inducement and resolve afforded by a properly aged bottle of Callaway (Temecula, CA) Chenin Blanc. We thank Joan Celeste for typing the manuscript.

REFERENCES

Bankier, A. T., & Barrell, B. G. (1983) in *Techniques in Nucleic Acid Biochemistry* (Flavell, R. A., Ed.) Vol. 85-08, pp 1–34, Elsevier/North-Holland, Limerick, Ireland.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.

Bishop, J. M., & Varmus, H. E. (1982) in *RNA Tumor Virus, Molecular Biology of Tumor Viruses* (Weiss, R., Teich, N., Varmus, H. E., & Coffin, J., Eds.) 2nd ed., p 999, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Cheng, Y.-S. E., & Chen, L. B. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2388–2392.

Cooper, J. A., & Hunter, T. (1981) *Mol. Cell. Biol. 1*, 394–407.

Cooper, J. A., & Hunter, T. (1982) *J. Cell Biol. 94*, 287–296.

Cooper, J. A., & Hunter, T. (1983) *J. Biol. Chem. 258*, 1108–1113.

Erikson, E., Cook, R., Miller, G. J., & Erikson, R. L. (1981a) *Mol. Cell. Biol. 1*, 43–50.

Erikson, E., Shealy, D. J., & Erikson, R. L. (1981b) *J. Biol. Chem. 256*, 11381–11384.

Erikson, K., & Martin, G. S. (1980) *Cell (Cambridge, Mass.) 21*, 829–838.

Erikson, R. L., Purchio, A. F., Erikson, E., Collett, M. S., & Brugge, J. S. (1980) *J. Cell Biol. 87*, 319–325.

Fava, R. A., & Cohen, S. (1984) *J. Biol. Chem. 259*, 2636–2645.

Geisow, M. J., Fritsche, U., Hexham, J. M., Dash, B., & Johnson, T. (1986) *Nature (London) 320*, 636–638.

Gerke, V., & Weber, K. (1984) *EMBO J. 3*, 227–233.

Gerke, V., & Weber, K. (1985a) *EMBO J. 4*, 2917–2920.

Gerke, V., & Weber, K. (1985b) *J. Biol. Chem. 260*, 1688–1695.

Glenney, J. R., Jr. (1985) *FEBS Lett. 192*, 79–82.
Glenney, J. R., Jr. (1986a) *J. Biol. Chem. 261*, 7247–7252.
Glenney, J. R., Jr. (1986b) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4258–4262.
Glenney, J. R., Jr., & Tack, B. F. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7884–7888.
Glenney, J. R., Jr., Boudreaux, M., Galyean, R., Hunter, T., & Tack, B. F. (1986) *J. Biol. Chem.* (in press).
Gould, K. L., Woodgett, J., Isacke, C. M., & Hunter, T. (1986) *Mol. Cell. Biol. 6*, 2738–2744.
Greenberg, M. E., & Edelman, G. (1983) *Cell (Cambridge, Mass.) 33*, 767–774.
Greenberg, M. E., Brackenbury, R., & Edelman, G. M. (1984) *J. Cell Biol. 98*, 473–486.
Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3961–3965.
Harris, H. E., & Weeds, A. G. (1983) *Biochemistry 22*, 2728–2741.
Hattori, T., Hoffman, T., & Hirata, F. (1983) *Biochem. Biophys. Res. Commun. 111*, 551–559.
Heidecker, G., Messing, J., & Groneborn, B. (1980) *Gene 10*, 69–73.
Hexham, J. M., Totty, N. F., Waterfield, N. D., & Crumpton, M. B. (1986) *Biochem. Biophys. Res. Commun. 134*, 248–254.
Hirata, F. (1981) *J. Biol. Chem. 256*, 7730–7733.
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., & Pepinsky, R. B. (1986) *Cell (Cambridge, Mass.)* (in press)
Hunter, T., & Soften, B. M. (1982) in *The Molecular Actions of Toxins and Viruses* (Cohen, P., & Van Heyningen, S., Eds.) p 333, Elsevier, Amsterdam.
Kretsinger, R. H., & Creutz, C. E. (1986) *Nature (London) 320*, 573.
Lehto, V. P., Virtanen, I., Passivuo, R., Ralston, R., & Alitalo, K. (1983) *EMBO J. 2*, 1701–1705.
Madin, S. H., & Darby, N. B. (1958) *Proc. Soc. Exp. Biol. Med. 98*, 574.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Nigg, E. A., Cooper, J. A., & Hunter, T. (1983) *J. Cell Biol. 96*, 1601–1609.
Okayama, H., & Berg, P. (1982) *Mol. Cell. Biol. 2*, 161–170.
Okayama H., & Berg, P. (1983) *Mol. Cell. Biol. 3*, 280–289.
Radke, K., & Martin, G. S. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5212–5216.
Radke, K., Gilmore, T., & Martin, G. S. (1980) *Cell (Cambridge, Mass.) 21*, 821–828.
Radke, K., Carter, V. C., Moss, P., Dehazya, P., Schliwa, M., & Martin, G. S. (1983) *J. Cell Biol. 97*, 1601–1611.
Rothhut, B., Russo-Marie, F., Wood, J., DiRosa, M., & Flower, A. J. (1983) *Biochem. Biophys. Res. Commun. 117*, 878–884.
Rouayrenc, J. F., Fattourn, A., Gabrion, J., Audemard, E., & Kassab, R. (1984) *FEBS Lett. 167*, 52–58.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Saris, C. J. M., Tack, B. F., Kristensen, T., Glenney, J. R., Jr., & Hunter, T. (1986) *Cell (Cambridge, Mass.)* (in press).
Sefton, B. M., & Hunter, T. (1984) *Adv. Cyclic Nucleotide Protein Phosphorylation Res. 18*, 195–226.
Staden, R. (1980) *Nucleic Acids Res. 8*, 3673–3694.
Staden, R. (1982a) *Nucleic Acids Res. 10*, 2951–2961.
Staden, R. (1982b) *Nucleic Acids Res. 10*, 4731–4751.
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., & Pepsinsky, R. B. (1986) *Nature (London) 320*, 77–81.
Wang, L. L., & Bryan, J. (1981) *Cell (Cambridge, Mass.) 25*, 637–649.
Yin, H. L., & Stossel, T. P. (1979) *Nature (London) 281*, 583–586.
Yin, H. L., Kwiatkowski, D. J., Mole, J. E., & Cole, F. S. (1984) *J. Biol. Chem. 259*, 5271–5276.
Zimmerman, C. L., & Pisano, J. J. (1977) *Methods Enzymol. 47*, 45–51.