The EMBO Journal vol.12 no.2 pp.667-675, 1993

Clathrin assembly protein AP180: primary structure, domain organization and identification of a clathrin binding site

S.A.Morris[1,3], S.Schröder[1,4], U.Plessmann[2], K.Weber[2] and E.Ungewickell[1,5,6]

[1]Max-Planck Institute for Biochemistry, D-8033 Martinsried and [2]Max-Planck Institute for Biophysical Chemistry, Department of Biochemistry, D-3400 Göttingen, Germany
[3]Present address: Department of Pharmacy, School of Pharmacy, University of California, San Francisco, CA 94143-0446, USA
[4]Present address: Biocenter of the University of Basel, Klingenbergstrasse 10, CH-4056 Basel, Switzerland
[5]Present address: Department of Pathology, Washington University School of Medicine, 660 South Euclid Ave, St Louis, MO 63110, USA
[6]Corresponding author

Communicated by K.Weber

Binding of AP180 to clathrin triskelia induces their assembly into 60-70 nm coats. The largest rat brain cDNA clone isolated predicts a molecular weight of 91 430 for AP180. Two cDNA clones have an additional small 57 bp insert. The deduced molecular weight agrees with gel filtration results provided the more chaotropic denaturant 6 M guanidinium thiocyanate is substituted for the weaker guanidinium chloride. The sequence and the proteolytic cleavage pattern suggest a three domain structure. The N-terminal 300 residues (pI 8.7) harbour a clathrin binding site. An acidic middle domain (pI 3.6, 450 residues), interrupted by an uncharged alanine rich segment of 59 residues, appears to be responsible for the anomalous physical properties of AP180. The C-terminal domain (166 residues) has a pI of 10.4. AP180 mRNA is restricted to neuronal sources. AP180 shows no significant homology to known clathrin binding proteins, but is nearly identical to a mouse phosphoprotein (F1-20). This protein, localized to synaptic termini, has so far been of unknown function.

Key words: AP180/assembly proteins/clathrin-coated vesicles/mouse phosphoprotein F1-20

Introduction

Clathrin-coated vesicles (CCV) are dynamic organelles that are involved in intracellular membrane trafficking (Goldstein *et al.*, 1985; Brodsky, 1988; Morris *et al.*, 1989). To understand the complex function(s) of CCV on a molecular level, one major approach has been the identification and characterization of the protein components that make up the cytoplasmic coat of CCV. Progress is most advanced for CCV purified from bovine brain. Biochemically well characterized coat components of this tissue include in addition to clathrin, the adaptor complexes (Golgi and plasma membrane adaptor), AP180 and auxilin (Morris *et al.*, 1989, Keen, 1990). While the two adaptor complexes are widely distributed proteins, AP180 and auxilin have so far been described only in neuronal cells or tissues (Ahle and Ungewickell, 1986, 1989; Kohtz and Puszkin, 1988; Ungewickell and Oestergaard, 1989). The adaptors are believed to link receptors to the clathrin network (Pearse, 1988) and to promote assembly of clathrin. Apart from their effects on the polymerization of clathrin, the function(s) of auxilin and AP180 are not known. AP180 was independently discovered in three laboratories and named AP180 (Ahle and Ungewickell, 1986), AP3 (Keen and Black, 1986; Keen, 1987) and NP185 (Kohtz and Puszkin, 1988). It has been only very recently formally established that the three proteins are probably identical or at least very closely related (Murphy *et al.*, 1991). The molar concentration of AP180 in brain-coated vesicle preparations is about half that of the plasma membrane adaptor and three times that of the Golgi adaptor and that of auxilin (Lindner and Ungewickell, 1992). In bovine brain only ~30% of AP180 is associated with membranes. The bulk forms a cytosolic pool (Ungewickell and Oestergaard, 1989). NP185 can be induced in PC12 cells by nerve growth factor and appears to be concentrated in nerve endings (Kohtz and Puszkin, 1988; Su *et al.*, 1991).

AP180 has some unusual molecular features. In most SDS-PAGE systems it migrates anomalously with an M_r of 155 000-185 000, although its molecular weight appeared to be between 112 and 123 kDa, as determined by other techniques (Ahle and Ungewickell, 1986; Prasad and Lippoldt, 1988; Murphy *et al.*, 1991). AP180 has a very low extinction coefficient, it is only poorly stained by Coomassie blue and is very sensitive to proteolytic attack (Prasad and Lippoldt, 1988; Murphy *et al.*, 1991). AP180 is a phosphoprotein that is phosphorylated at serine residues *in vivo* (Keen and Black, 1986) and *in vitro* (Morris *et al.*, 1990). The native protein was shown to be a monomer that interacts with clathrin triskelia with a stoichiometry of one and thereby induces clathrin assembly into a homogeneous population of 60-70 nm coats (Ahle and Ungewickell, 1986; Prasad and Lippoldt, 1988). Recent work in our laboratory has demonstrated that AP180 is about four times more active in inducing clathrin assembly than the adaptor complexes or auxilin (Lindner and Ungewickell, 1992). AP180 is readily cleaved by trypsin and elastase into fragments of apparent M_r of ~30 000 and ~100 000, respectively (Murphy *et al.*, 1991). Gel filtration of the unfolded polypeptide in 6 M guanidinium hydrochloride confirmed the molecular weight of the smaller fragment. In contrast the 100 000 apparent M_r fragment, which displays the same poor affinity for Coomassie blue as does the parental chain, eluted as a 50 000 Da species. Only the ~30 kDa moiety retained the ability to interact with clathrin cages, but it does not promote assembly of clathrin triskelia.

Knowledge of the primary structure of AP180 should aid our understanding of the function of this protein. Here we report the complete primary structure of AP180 deduced from sequencing rat brain cDNA clones one of which has an open reading frame that codes for a 91.4 kDa polypeptide. Partial protein sequencing of the M_r ~30 000 clathrin

binding fragment showed it to be located in the N-terminal third of AP180.

Results

Partial amino acid sequences of AP180 and molecular cloning of its cDNA

For direct sequencing AP180 was prepared either from bovine brain cytosol by affinity chromatography using the monoclonal antibody mAb AP180-I or from clathrin-coated vesicles using standard chromatographic techniques (Ahle and Ungewickell, 1986). It was initially planned to use this information for the synthesis of specific oligonucleotide probes and for confirming the identity of candidate AP180 cDNA clones. Attempts to sequence intact AP180 were unsuccessful, presumably because of a blocked N-terminus. To obtain internal sequences from AP180 the protein was, after a final purification step by SDS–PAGE, transferred on to a polyvinylfluoride membrane and digested with trypsin and/or Asp-N-protease. Peptides released were purified by HPLC and subjected to microsequencing (Table I). Controlled digestion of native AP180 (apparent M_r 180 000) with trypsin rapidly converts the protein into fragments of apparent 148, 126, 107 and 33/32 kDa (Figure 1). mAb AP180-I stained zones of 180 000, 148 000, 126 000 and 107 000 apparent M_r, but not the 33/32 kDa species. The 107 kDa fragment was only barely visible on Coomassie-stained gels and appeared upon silver staining as a negatively stained zone (not shown). This digestion pattern is very similar to that described by Murphy *et al.* (1991) who reported cleavage of native AP3 (AP180), which behaves in their gel system like a 155 kDa polypeptide, into fragments of 135, 100, ~30 and 14 kDa. We only saw the 14 kDa fragment clearly when the gel was stained with silver. The monoclonal antibody CLAP3 of Murphy *et al.* (1991) reacted like our monoclonal only with fragments of apparent M_r of 100 000 and higher. The ~30 000 Da fragment was shown by Murphy *et al.* (1991) to retain the ability of intact AP180 to bind to immobilized clathrin. In contrast the 100 kDa fragment was shown to interact with clathrin only if still associated with the ~30 kDa fragment. We essentially confirmed these results in a clathrin cage binding assay. For this, purified clathrin was assembled into cages and added in excess to an unfractionated tryptic digest of AP180 composed of ~100 and ~30 kDa fragments. The cages were pelleted by ultracentrifugation, and supernatant and pellet fractions were analysed by SDS–PAGE and immunoblotting for the presence of AP180 fragments (Figure 2). In one experiment, in which intact AP180 was added to the digest prior to addition of clathrin cages, we noted that the intact protein associated quantitatively with the clathrin (Figure 2B), while 30–40% of the ~30 kDa fragment remained in the supernatant (Figure 2A). Most of the 107–148 kDa fragments partitioned equally between supernatant and pellet fractions (Figure 2B).

Because of its role in clathrin binding we decided to obtain as much peptide sequence information as possible from the

Table I. Peptide sequences obtained from bovine AP180[a]

Origin	Cleavage	Peptide	Sequence
AP180	Trypsin	T180-18	SGSHGYDXSTFIR
		T180-19	QMAFDFAR
		T180-24	ATNSSXV
		T180-26	FIQYLASR
		T180-32	NTLFNLSNFL
		T180-36	ALGSDLDSSLASLVENL-EISE
		T180-T9[b]	KLTGGAN
		180-T34/1[b]	KPPAKDPL
		T180-T34/3[b]	HLDYLIQATNETNVNIP-QMADTL
AP180	Asp-N protein	ASP180-7	DYLXQATN
		ASP180-10	DTFVAPSPATTASPA
		ASP180-11	DLTQAPSSLM
		ASP180-16	DAFAASPGEAPAASEGA
		ASP180-29/1[b]	DMSTFIRRYSRYLNEK-AFSYRQMAFD
E38	none	N-terminal	SVTGSAVARAVCKATT-HEVMGPKKK
T107	none	N-terminal	SSPATXVT
T33	none	N-terminal	IAAAQYSV
	Trypsin	T33-T5	RYSR
		T33-T8	TMAPEK
		T33-T10	ATTHEVMGPK
		T33-T12	YLNEK
		T33-T14	LLK
		T33-T19	KPGNNEGSGAPSPLSK
		T33-T21	FLTR
		T33-T22	AFSYR
		T33-T24	IAAAQYSVTGSAVA
		T33-T30	QMAFDFAR
		T33-T32	FIQYLASR
		T33-T35	SGSHGYDMSTFIR
		T33-T39	ATNSSXVVVF
		T33-T44	NTLFNLSNFLDK
		T33-T18/2[b]	SGQTLTD
		T33-T25/2[b]	ALVTTHHLMVHGNER
		T33-T26/1[b]	VSEFLK
		T33-T26/2[b]	DALEIYKR
		T33-T42/1[b]	HLDYLIQATNETNVN
		T33-T42/2[b]	VAEQVGIXKGDIPDL
	Asp-N protein	T33-ASP8	SMPILQGQI
		T33-ASP16	DALLEF

[a]The 107 and 33 kDa apparent molecular weight tryptic fragments and the 38 kDa elastase fragment of AP180 were subjected to N-terminal sequencing. Internal sequences of AP180 and its proteolytic derivatives were obtained by cleavage with trypsin or Asp-N protease as indicated.
[b]Indicates sequences obtained from peptide mixtures.

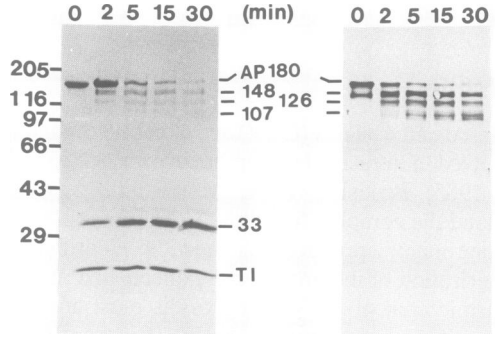

Fig. 1. Limited digestion of AP180 with trypsin. Purified AP180 was treated with trypsin (w/w 1:375) on ice for the times indicated. The digest was fractionated on 7.5–19% polyacrylamide gradient gels (see Materials and methods for details). Left panel shows a Coomassie blue stained gel and the right panel shows the corresponding immunoblot that was reacted with mAb AP180-I. TI, soybean trypsin inhibitor.

33 kDa fragment. From its N-terminus we obtained the sequence IAAAQYSV. Internal sequences in this fragment were obtained by successive cleavages with trypsin and endoprotease Asp-N (Table I). Elastase digestion of native AP180 produced a M_r 38 000 fragment in high yields (not shown), which started with the sequence SVTGSAVARA VCKATTHEVM GPKKKHL. Its ability to interact with clathrin was not tested. The apparent 107 kDa trypsin fragment started with the sequence SSPATXVT, where 'X' stands for an unidentified residue.

Initial attempts to screen cDNA libraries with oligonucleotide probes were abandoned in favour of screening a λ-ZapII expression library with the mAb AP180-I. Three positive clones were identified and one of them, AP180-1A, was sequenced. It contained an open reading frame predicting a region highly homologous to peptide T180-36 (Table I). A full length clone (AP180-36) was subsequently identified by screening the same library with a *Pst*I restriction fragment consisting of 952 nucleotides from the 5′ end of clone AP180-1A. Clone AP180-36 has an open reading frame for a protein of a predicted molecular mass of 91 430 (Figure 3), which is ~20 000 Da smaller than expected for AP180 on the basis of gel filtration and sedimentation studies (Ahle and Ungewickell, 1986; Prasad and Lippoldt, 1988; Murphy *et al.*, 1991). The cDNA clones AP180-15 and −19 differed from AP180-36 in respect to an insertion of 57 bp in position 1894, which increases the molecular weight of the predicted protein to 93 518. This heterogeneity could result from either alternative splicing or the transcription of a related gene. We also noted some sequence heterogeneity at nucleotide position 1704, corresponding to a region unusually rich in alanine and threonine. Here clones AP180-12 and −19 contained two extra CCA nucleotide triplets giving an extra two threonines in the amino acid sequence. A summary of the identified clones is shown in Figure 4.

To determine the size and expression pattern of the AP180 message, total RNA from rat brain, kidney, heart and liver were incubated with the labelled *Pst*I fragment of clone AP180-1A. Brain tissue gave a strong signal at ~5 kb (Figure 5). The broadness of the zone may reflect the heterogeneity of the original mRNA rather than degradation. While mRNA preparations from kidney gave a very weak, but sharp signal, neither heart nor liver produced a signal. It is possible that the signal from kidney comes from contaminating adrenal gland tissue, since we detected by immunoblotting small amounts of AP180 in the medulla (not shown), which is a tissue derived from neural crest.

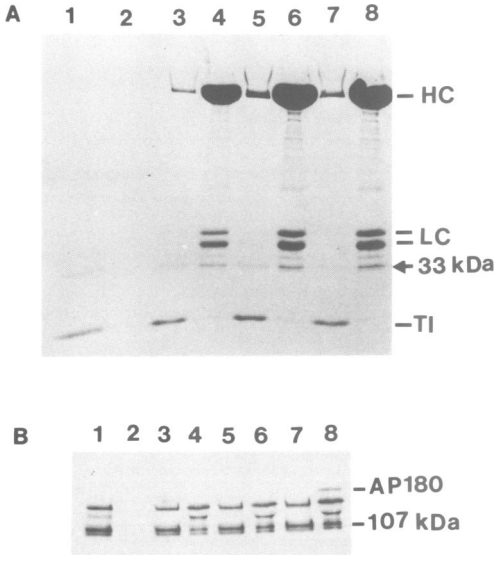

Fig. 2. Binding of the M_r 33 kDa fragment to clathrin cages. 4 μg of mildly digested AP180 were incubated for 30 min on ice in the absence of clathrin (lanes 1 and 2) or with 25 μg (lanes 3 and 4) or 50 μg (lanes 5 and 6) of pre-assembled clathrin cages. In one experiment 1 μg of intact AP180 was also added (lanes 7 and 8). Cages were pelleted by ultracentrifugation and analysed for the presence of bound AP180 fragments by SDS−PAGE. Odd numbers correspond to the supernatants; even numbers to pellets. Pellets in (**A**) were loaded at double concentrations. (**B**) Immunoblot of the binding experiment using mAb AP180-I. HC and LC mark clathrin heavy and light chains. TI is trypsin inhibitor. Arrow points at the 33 kDa apparent M_r clathrin binding fragment. Note that this fragment binds less efficiently to clathrin than intact AP180 (lanes 7 and 8 in panels A and B).

Molecular weight of AP180 and its domain organization

Previously we reported that AP180 elutes in 6 M guanidinium chloride like a 123 kDa polypeptide (Ahle and Ungewickell, 1986). Keen and colleagues, who used the same method, determined 112 kDa for AP3/AP180 (Murphy *et al.*, 1991). Both values are substantially higher than the molecular weights of 91 430 or 93 518 (with insertion) calculated from the derived sequence for AP180. This could suggest (i) that we have not cloned AP180 but a highly related smaller protein, (ii) that AP180 undergoes extensive post-translational modification or (iii) that reduced AP180 does not behave as a random coil in 6 M guanidinium chloride. To address these issues we repeated the molecular weight measurement of AP180 by gel filtration on an analytical Superose 6 column. However, in contrast to previous measurements, this time the protein was analysed in 6 M guanidinium thiocyanate rather than 6 M guanidinium chloride, since the thiocyanate anion is a more potent protein denaturant than chloride (Gratzer and Beaven, 1968). A set of marker proteins was used to demonstrate linearity between the logarithm of molecular weight and elution volume in this particular solvent (Figure 6). Total coated vesicle protein was applied to the Superose 6 gel filtration column. The eluent was fractionated and analysed by SDS−PAGE and immunoblotting. Under these conditions AP180 eluted at a position approximately equivalent to that of the 91.4 kDa γ subunit of the Golgi adaptor (Robinson, 1990,) but slightly behind the 104.7 kDa β subunit of the plasma membrane adaptor (Kirchhausen *et al.*, 1989) (Figure 7). This result shows that the molecular weight of AP180 in 6 M guanidinium thiocyanate is close to that calculated from the sequence of the protein encoded by cDNA clone AP180-36. This and the fact that the deduced sequence of AP180 contains all 39 sequenced peptides listed in Table I strongly suggests that we have cloned and sequenced AP180 and not a related protein.

All peptides from the bovine brain 33 kDa clathrin binding fragment of AP180 were located within the first 304 residues. Determination of the N-terminal sequence of the M_r 33 000 fragment gave the sequence IAAAQYSV, which is located only 10 residues from the proposed N-terminus of AP180. The N-terminus of the apparent 107 kDa fragment is defined by the sequence SSPATXVT (peptide T107), which is

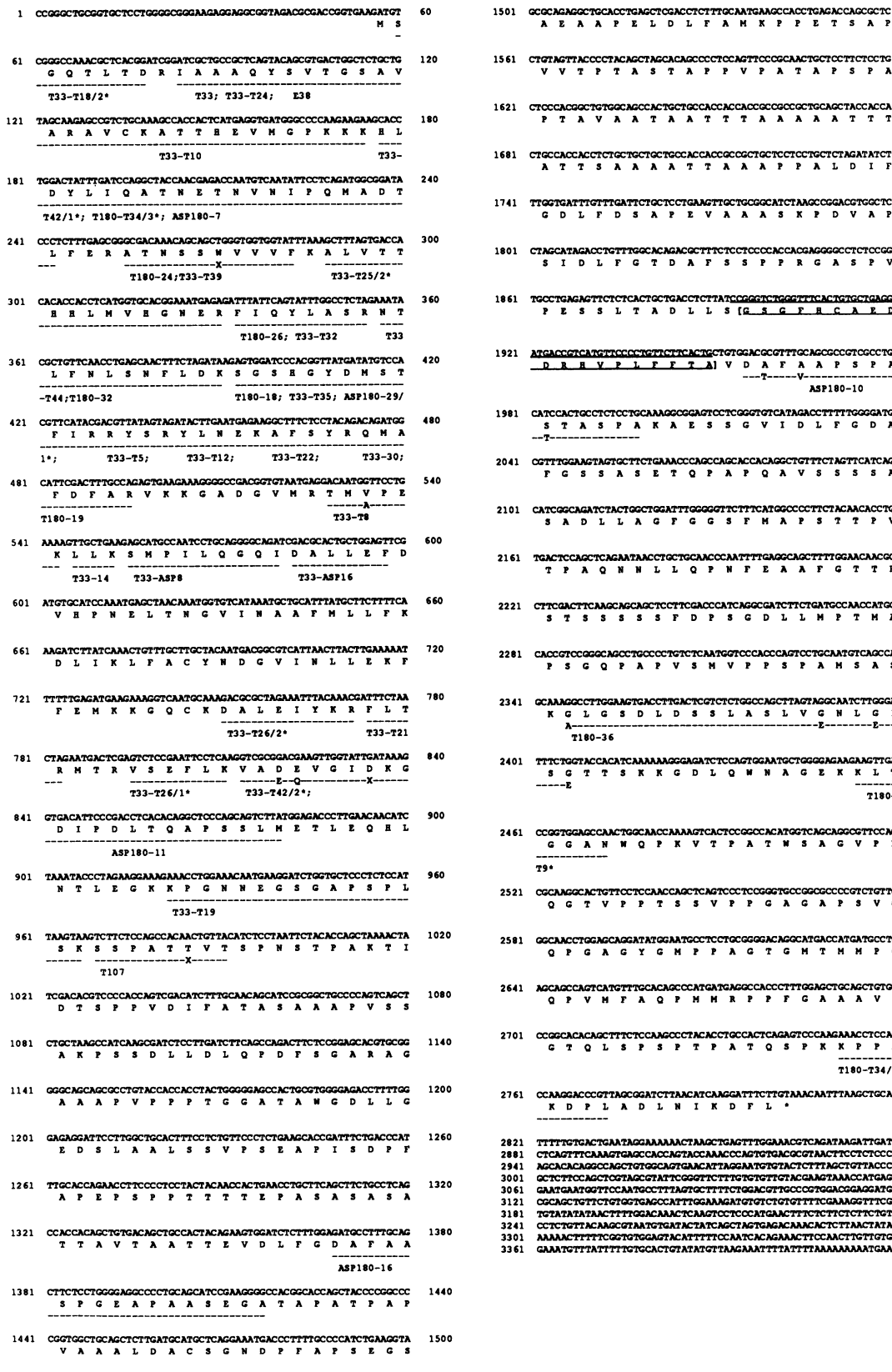

Fig. 3. Complete nucleotide and deduced amino acid sequence of rat AP180. Numbers refer to nucleotide positions. Beneath the nucleotide sequence is the deduced amino acid sequence coded for by the open reading frame between nucleotides 55 and 2799. The broken lines indicate the positions of peptides obtained by direct sequencing of bovine AP180 (from Table I). Differences between the rat and bovine sequence are indicated (X = undetermined amino acid). Underlined is the 57 nucleotide insert sequence between nucleotides 1894 and 1950, which is present in two of our eight clones. The corresponding deduced 19 amino acid sequence is shown in square brackets and is also underlined. The above sequence has been deposited in the EMBL Sequence Data Library under the accession numbers X68877 (rat mRNA for AP180) and X68878 (rat + insert for AP180).

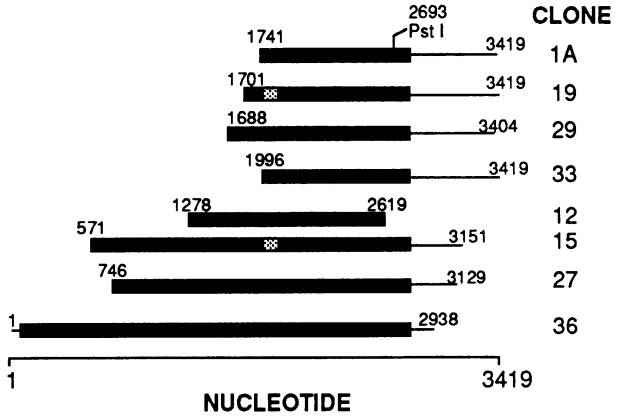

Fig. 4. Alignment of AP180 clones. Thick lines designate coding and thin lines non-coding sequence. Clone 1A was obtained from a λ-Zap II expression library by screening with mAb 180-I. This was cut with restriction endonuclease *PstI* at position 2693 to produce a 952 nucleotide 5′ fragment which was labelled by random hexanucleotide priming and used to rescreen the same library. This rescreen provided the other clones shown. The 57 nucleotide insert present in clones 15 and 19 is indicated by the shaded area. Clone 36 contains an open reading frame of 2688 nucleotides, coding for a protein of 896 amino acids.

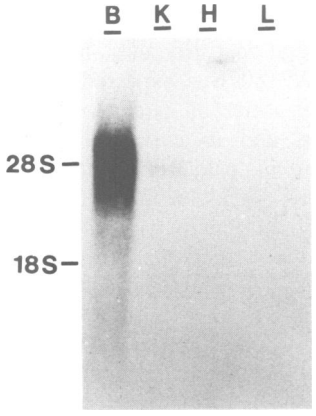

Fig. 5. Northern blot analysis. A blot of total RNA (~15 μg) from rat brain (B), kidney (K), heart (H) and liver (L) was incubated with the labelled 952 nucleotide 5′-end *PstI* restriction fragment of clone AP180-1A. The positions of 28S and 18S ribosomal RNA are shown as markers. In brain a strong diffuse signal at 5 kb is seen. Kidney gives a much weaker, but sharper signal. No mRNA for AP180 was detected in heart and liver.

adjacent to the peptide T33-T19 obtained from the 33 kDa fragment. Thus, trypsin cuts at Lys304, which results in a molecular weight of 33 173 Da for the clathrin binding fragment. This is in agreement with its electrophoretic mobility in SDS-containing gels. In contrast to peptides from other regions of AP180, the peptides from the M_r 33 000 fragment are with only one exception (T33-T8) 100% identical to the corresponding sequences of AP180 from rat. This conservation across species might be expected for a region involved in binding to a highly conserved protein such as clathrin. The distribution of charged amino acids and the distribution of alanine and proline suggests a three domain structure for AP180 (Figure 8). The first 304 residues that bind to clathrin are predominantly basic (pI 8.7). Then follows a domain of 450 residues that is overall very acidic (pI 3.6). This domain is interrupted by an uncharged segment

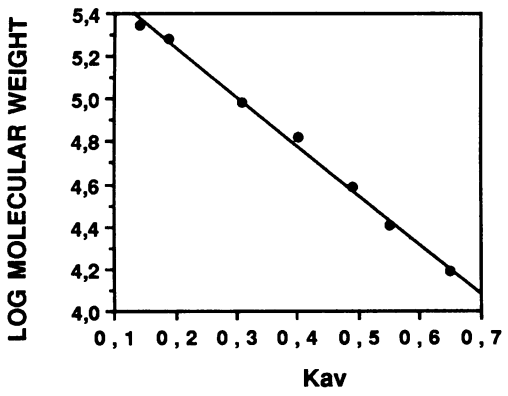

Fig. 6. Calibration of a Superose 6 gel filtration column in 6 M guanidinium thiocyanate. Protein standards were denatured by dialysis against 6 M guanidinium thiocyanate, 50 mM Tris−HCl, pH 8.0, 2 mM EDTA and 5 mM DTT. The column was calibrated with myosin (222 kDa), clathrin heavy chain (192 kDa), phosphorylase b (97 kDa), BSA (66 kDa), aldolase (39 kDa), clathrin light chains a and b (25.6 kDa average) and haemoglobin (15.5 kDa). The void volume was determined with clathrin cages that were lightly fixed with glutaraldehyde. Kav, partition coefficient of the protein between the mobile and stationary phases.

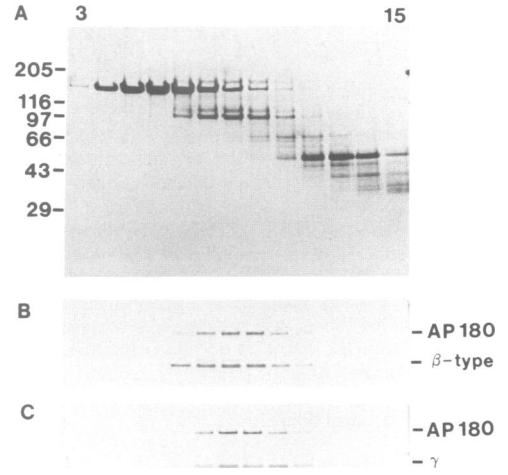

Fig. 7. Molecular weight estimation of AP180 by gel filtration in 6 M guanidinium thiocyanate. 0.2 ml total coat protein unfolded by dialysis against 6 M guanidinium thiocyanate, 50 mM Tris−HCl pH 8.0, 2 mM EDTA and 5 mM DTT were applied to a calibrated Superose 6 gel filtration column equilibrated in the same buffer. 0.5 ml fractions were collected and their composition analysed by SDS−PAGE and immunoblotting (see experimental section for details). (**A**) Coomassie blue stained gel of fractions 3−15; (**B**) corresponding immunoblot which was reacted simultaneously with mAbs AP180-I and 100/1 (anti-β-type adaptor subunits); (**C**) immunoblot that was stained simultaneously with mAbs AP180-I and 100/3 (anti-γ adaptor subunit). Myosin, clathrin heavy chain, phosphorylase b, BSA, aldolase, clathrin light chains and haemoglobin eluted in fractions 5, 6, 10, 12, 15, 17 and 18, respectively.

of 59 residues that is unusually rich in alanine (29 residues). A small C-terminal domain (150 residues) is again overall basic (pI 10.4). Numerous potential phosphorylation sites for casein kinase II (12), protein kinase c (6), cdc2 kinase (4) and cyclic AMP-dependent protein kinase sites (2) are present. From previous studies it is already known that at least one of the casein kinase II sites is phosphorylated by a coat-associated kinase *in vitro* (Morris *et al.*, 1990).

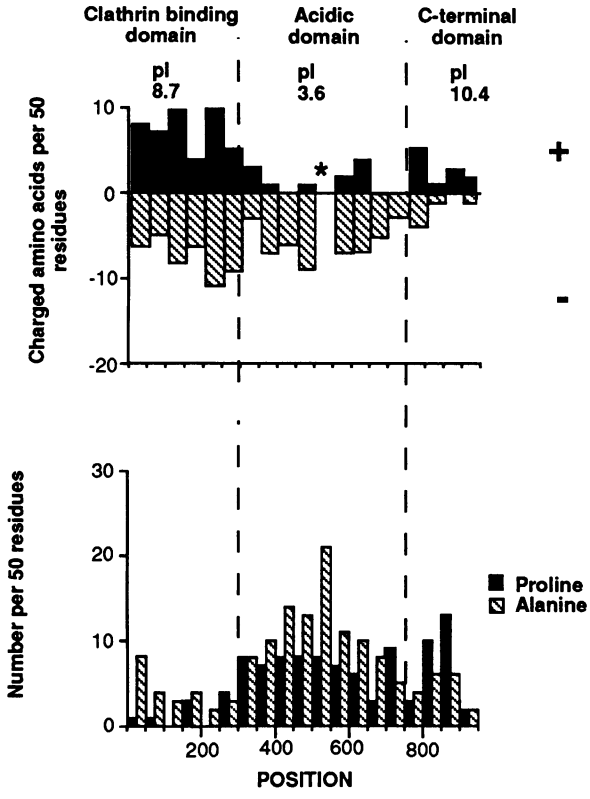

Fig. 8. Domain structure of AP180. Shown is an analysis of the amino acid sequence of AP180 including the 19 amino acid insert. Absence of the insert has no significant influence on the overall domain structure. The isoelectric points of the domains were calculated using the UWGCG program PEPTIDESORT. The upper panel shows a plot of charged amino acid residues. The asterisk indicates the stretch of 59 uncharged residues. The lower panel shows a plot of the proline and alanine residues. Both forms of analysis suggest a three-domain structure for AP180 as indicated by the two vertical broken lines. The basic N-terminal ~30 kDa domain is followed by an acidic ~50 kDa middle domain, which is primarily responsible for the anomalous Coomassie and SDS binding of AP180. A basic ~15 kDa region forms the C-terminal domain.

Relationship of AP180 to other proteins

A search conducted using FASTA showed that the coding sequence of clone AP180-2 is >97% identical to an unpublished sequence coding for a neuron-specific mouse phosphoprotein F1-20 (Genbank databank entry MUSF 120a) that is localized to synaptic terminals (Lafer *et al.*, 1989). The degree of identity between the two proteins increases on the amino acid level to 98%, since most differences at the nucleotide level affect the third codon position only. The major difference is a short 15 nucleotide insertion in the mouse phosphoprotein F1-20 at position 2254 resulting in an extra five amino acids at position 714 (Figure 9), which does not occur in any of our cDNA clones. Also, the mouse phosphoprotein lacks the 57 bp insertion that is present in clones AP180-15 and 19.

Apart from the apparent identity of AP180 to the mouse phosphoprotein, no significant homologies between AP180 and any other proteins listed in the data banks were noted. A weak similarity (22% identity) between AP180 and the γ subunit of the Golgi adaptor exists in a 116 residue overlap, which in both proteins is rich in proline. AP180 also does not share any significant sequence homology with the clathrin binding protein auxilin (data not shown).

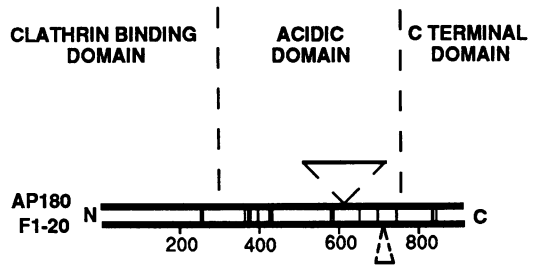

Fig. 9. Comparison between rat AP180 and mouse phosphoprotein F1-20. Both deduced protein sequences are shown diagrammatically. Based on the BESTFIT (UWGCG) program they are 98% identical. Connecting lines mark different amino acids. Note that very few of these occur in the clathrin binding domain. Also indicated is the position of the 19 amino acid insert in rat AP180 (upper triangle) that is absent from the reported forms of mouse phosphoprotein F1-20 and the five amino acid insert in mouse phosphoprotein F1-20 (lower triangle), which is absent from all of our AP180 clones.

Discussion

AP180 was first discovered as a specific coat component of clathrin-coated vesicles from neuronal tissue (Ahle and Ungewickell, 1986; Keen and Black, 1986). It binds to the clathrin heavy chain (Ahle and Ungewickell, 1986) with a stoichiometry of ~1 per triskelion (Prasad and Lippoldt, 1988), and thereby promotes very efficient assembly of clathrin into regular coat structures. Molecular cloning and sequencing of AP180 was expected to shed light on its structural organization, its relationship to other clathrin binding proteins and to explain some of its unusual properties, such as anomalous behaviour in SDS–PAGE and low affinity for Coomassie blue. The 91.4 kDa protein encoded in the full-length cDNA clone AP180-36 is ~20–30 kDa smaller than expected for the protein on the basis of previous molecular weight determinations, which included methods such as gel filtration in 6 M guanidinium chloride, generally believed to provide reliable estimates for the molecular weights of unfolded and reduced proteins. This unexpected result was of primary concern to us since we wanted to rule out the possibility of having cloned a related smaller protein, not AP180. Thus we reinvestigated the molecular weight of AP180 and explored the possibility that 6 M guanidinium chloride does not suffice to completely unfold the protein. One of the first examples of a protein that resists complete unfolding by 6 M guanidinium chloride is thyroxine-binding pre-albumin (Branch *et al.*, 1972). These authors concluded that 6 M guanidinium chloride cannot be considered as a universal denaturant. We therefore used to 6 M guanidinium thiocyanate. By virtue of its more chaotropic anion, which ranks highest in the Hofmeister series, guanidinium thiocyanate is a far more potent denaturant than other guanidinium salts (Gratzer and Beaven, 1968). The demonstrated almost perfect linear relationship between the logarithm of the molecular weight of marker proteins and their elution volume from Superose 6 (Figure 6) makes guanidinium thiocyanate a very useful solvent for molecular weight analysis of protein subunits. The elution position of AP180, as determined relative to those of the adaptor subunits γ and β, was clearly between 91 and 105 kDa and thus in agreement with the molecular mass of the cloned protein. We do not know yet why intact AP180 fails to become completely unfolded by 6 M guanidinium hydrochloride. However, the regions of the molecule that

are responsible for this unusual behaviour can now be identified by expressing distinct parts of the protein and analysing their elution characteristics in guanidinium chloride. The fact that all 39 sequenced bovine peptides are present within the proposed rat sequence strongly also suggests that clone AP180-36 encodes the complete AP180.

Overall, AP180 is an acidic protein with a calculated pI of 4.6 and a low content of aromatic residues (Trp 0.6%, Tyr 1%) accounting for its low extinction coefficient (E1%, 1 cm, 280 nm = 4.5).

The distribution of proline, alanine and charged residues in AP180 suggest a three-domain organization for the protein: a central acidic domain is flanked by basic domains (Figure 8). This readily explains the ~50, ~30 and ~14 kDa fragments, which are obtained by limited digestion with trypsin (Murphy *et al.*, 1991). The first cut, which generates the ~14 kDa fragment (Murphy *et al.*, 1991), probably occurs at Lys744 resulting in the liberation of the basic C-terminal domain. The next one which we have shown here to occur at Lys304, severs the covalent linkage between the 33.5 kDa clathrin binding domain and the acidic central domain.

Clone 1A was obtained by screening an expression library with mAb180-1. Hence the epitope for this antibody must be within the C-terminal 353 amino acids. Given that mAb180-1 does not stain the 14 kDa tryptic fragment (data not shown), the epitope has to reside in between residue 562 and residue 744.

Murphy *et al.* (1991) noted that under non-denaturing conditions the 33.5 kDa clathrin binding domain and the acidic central domain remain partially associated. This was confirmed by our observation that in addition to the 33 kDa fragment some of the 50 kDa fragment also (seen on the immunoblots as an apparent ~107 kDa species) co-sedimented with clathrin in the cage binding assay. When compared with intact AP180, the 33 kDa fragment appears to have a somewhat reduced affinity for clathrin. Murphy *et al.* (1991) indicated that the fragment had also lost clathrin assembly promoting activity. AP180 which binds with a stoichiometry of one to clathrin triskelia (trimers of heavy chains), may function as an assembly protein by cross-linking adjacent triskelia. The reduction in affinity and loss of assembly promoting activity could be readily explained by the loss of a second binding site, which conceivably might be contained within the 14 kDa C-terminal domain.

The central 50 kDa domain is characterized by a highly acidic isoelectric point (pI 3.6) and by an unusual stretch of 59 uncharged residues, composed of 42% alanine, 27% threonine and 17% proline. It is probable that the acidic nature of the 50 kDa domain interferes with SDS binding, resulting in anomalous migration in SDS–PAGE, and is also responsible for the reduced binding of Coomassie blue. In two of our clones the uncharged region is followed by an insertion of 19 residues. So far we do not know if AP180 from bovine brain-coated vesicles carries this insertion or not. However, since the insert adds a fifth cysteine to the protein, it should be possible to answer this question by determining the number of cysteines per AP180 molecule experimentally.

AP180 shares no significant sequence homologies to other known clathrin binding proteins such as the subunits of the adaptor complexes or auxilin. Considering the differences between AP180 and adaptors in respect to subunit and domain organization, and in light of the reported differences in the clathrin binding stoichiometries between auxilin and AP180, this is not necessarily surprising (Prasad and Lippoldt, 1988; Ahle and Ungewickell, 1990). From competition binding experiments conducted with adaptors, auxilin and AP180, we also concluded that AP180 binds to a site distinct from that of the adaptors and auxilin (R.Lindner and E.Ungewickell, in preparation). Moreover, Murphy and Keen (1992) have recently reported that the removal of the terminal domain from the clathrin heavy chain affects binding of AP180 and adaptors differently.

A database search identified an as yet unpublished sequence for a protein, referred to as mouse phosphoprotein F1-20. Within the coding region the nucleotide sequences are so similar that one must consider AP180 and F1-20 to be highly related. F1-20 is defined by a monoclonal antibody that was used to demonstrate its specific location in synapses (Lafer *et al.*, 1989). The homology extends even into the non-coding regions with the small differences readily attributable to species variation. The limited information available on F1–20 is in agreement with the reported properties of AP180. We and others have shown that AP180 is *in vitro* and *in vivo* a phosphoprotein. Puszkin and colleagues have reported that AP180 is specific for neuronal tissue and is enriched in nerve endings (Su *et al.*, 1991). Our Northern blot analysis also showed that the AP180 message is abundant only in neuronal tissue.

Taking together the observations made in several laboratories, AP180 can be considered as a multifunctional protein (Ahle and Ungewickell, 1986; Keen and Black, 1986; Kohtz and Puszkin, 1988, 1989; Murphy *et al.*, 1991; Su *et al.*, 1991). Independently confirmed properties include its presence in the coat of clathrin-coated vesicles from neuronal sources, its function as a clathrin binding protein with assembly promoting properties and its ability to interact with membranes. As yet unconfirmed are reports on its ability to interact with tubulin, clathrin light chains and the plasma membrane adaptor (Kohtz and Puszkin, 1988, 1989; Su *et al.*, 1991). We expect that by expressing AP180 or individual domains in bacteria, large enough quantities of this protein will become available to investigate its interactions with cytoskeletal components and membranes using biochemical techniques. Its function may be further explored in transformed cells that normally do not express this protein and in transgenic animals no longer expressing functional AP180.

Materials and methods

Materials

Fresh bovine brains were obtained from the local abattoir and processed within 2 h of slaughter. Cyanogen bromide-activated Sepharose 4B, Ficoll 400 and Superose 6 gel filtration column (HR10/30) were from Pharmacia; 2-(*N*-morpholino)ethane sulfonic acid, trypsin and phenylmethylsulfonyl fluoride were from Sigma; guanidinium thiocyanate (reagent grade) was from Merck; ethyleneglycol *bis*(2-aminoethylether)*N,N,N',N'*-tetraacetic acid (EGTA) and sucrose were from Serva (Heidelberg, FRG); T4 DNA ligase, Klenow fragment, calf intestine alkaline phosphatase and endoproteinase Asp-N were from Boehringer; T7 sequencing kit and Deaza G/A Sequencing mixes were from Pharmacia; an oligo (dT) λ-ZAP II rat brain cDNA library was from Stratagene; Geneclean II for plasmid DNA purification was from Bio 101 (La Jolla, USA); DNA grade agarose was from Bio-Rad; reagents for SDS–PAGE and urea were from LKB Instruments GmbH (Gräfelfing, FRG); horseradish peroxidase-conjugated IgGs to mouse immunoglobulins were from Dakopatts GmbH (Hamburg, FRG); nitrocellulose transfer membranes (BA83, 0.2 μm) were from Schleicher and Schuell, Dassel

(FRG); radiochemicals and Hybond-N nylon membranes were from Amersham. All other chemicals and reagents were of analytical grade.

Purification of AP180 and protein sequence analysis

AP180 was either purified from bovine brain clathrin-coated vesicles as described previously by Ahle and Ungewickell (1986) or affinity purified from 1 l bovine brain cytosol [the supernant obtained from the first ultracentrifugation step in standard coated vesicle preparations, Campbell *et al.* (1984)] using 3 mg mAb AP180-I (Ahle and Ungewickell, 1986) attached to 1.5 ml Sepharose 4B. Intact AP180 was eluted from the column either with 0.1% SDS at 60°C for 5 min or with ice-cold 0.2 M glycine pH 2.5 (Su *et al.*, 1991) followed by rapid neutralization of the eluate with 1/10 vol of 1.5 M Tris−HCl pH 8.8. A M_r 38 000 fragment of AP180 was prepared by incubating the AP180-I resin with 2.5 μg elastase for 20 min at room temperature. The fragment which did not adhere to the antibody was released from the column during a wash step. A ~30 kDa fragment of AP180 that retains the ability to bind to clathrin (Murphy *et al.*, 1991), was prepared by treatment of the purified protein (0.1−0.2 mg/ml) in 50 mM Tris−HCl, pH 8.0 with trypsin at a weight ratio of 1:375 for 20 min on ice. The reaction was stopped by addition of a 10-fold molar excess of soybean trypsin inhibitor. For sequencing, intact AP180 or its fragments were purified by SDS−PAGE on mini gels (Hoefer Mighty Small II unit). The gels were briefly stained and destained. Protein zones of interest were excised from the gel. Identical species were pooled, concentrated into one slot of a second gel (12×14×0.2 cm) and electrophoretically transferred on to a polyvinyldifluoride membrane (Bauw *et al.*, 1989). The resulting blots were treated with trypsin and in some cases also with Asp-N protease for 4 h at 37°C. The eluted peptides were subjected to HPLC (Applied Biosystems, Aquapore C8 column, elution with 7−42% acetonitrile in 0.1% trifluoroacetic acid or a Vydac 218 TP52 column, elution with 7−65% acetonitrile in trifluoroacetic acid). All peptides obtained were subjected to automated gas phase sequencing using either an Applied Biosystems Sequenator (Model 470A) or a Knauer Sequenator (Model 810). Both instruments were equipped with an on-line phenylthiohydantion amino acid analyser.

Cage binding experiments

Cages were assembled from clathrin triskelia by dialysing the protein at concentrations ranging from 2−5 mg/ml against 0.1 M 2(*N*-morpholino)ethanesulphonic acid, 1 mM EGTA and 0.5 mM $MgCl_2$ pH 6.5 in the presence of 3 mM $CaCl_2$. Binding experiments were performed by incubating 25 or 50 μg of clathrin cages with 4 μg of mildly trypsinized AP180 in 100 μl isolation buffer for 30 min on ice. In one experiment 1 μg of intact AP180 was also added. Binding was analysed by ultracentrifugation in a fixed angle rotor followed by SDS−PAGE of supernatant and pellet fractions as described in detail elsewhere (Lindner and Ungewickell, 1991).

Isolation and sequence analysis of cDNA

A rat brain λ-ZAP II cDNA expression library (40 000 p.f.u.) was screened according to the manufacturer's instructions with mAb AP180-I. Three clones which proved positive after two rounds of subcloning were further characterized. The Bluescript SK+ plasmid was automatically excised from the vector according to the manufacturer's instruction and subjected to sequencing by the dideoxy-chain termination method of Sanger *et al.* (1977). One clone, AP180-1A, contained the deduced sequence GLGSDLDSSLASLVGNLGISG, which differed in only four positions from a peptide sequence that was determined by microsequencing of bovine brain AP180. 32 more clones, including a full-length one (AP180-36), were identified in the same library by screening with a 0.9 kb nucleotide probe that corresponded to the 5′-end of clone 1A. This was labelled by complementary strand synthesis primed by random hexanucleotide mixtures in the presence of 50 μCi [α-^{32}P]ATP and Klenow enzyme. Again, 40 000 p.f.u. were plated and screened with the probe according to standard procedures (Sambrook *et al.*, 1989). Synthetic 18mer oligonucleotides were used as internal primers to obtain complete sequences. Regions showing compressions on the sequencing gels were sequenced with 7-deaza dATP and 7-deaza dGTP mixtures. For Northern blotting experiments total RNA was extracted from rat brain, kidney, heart and liver with 8 M guanidinium hydrochloride (Sambrook *et al.*, 1989). ~15 μg RNA were loaded into the slots of a 1% agarose gel containing formaldehyde. Electrophoresis, transfer of RNA to Hybond-N nylon membrane and incubation with a labelled restriction fragment as a probe was according to the manufacturer's instructions. A final concentration of 50% formamide was used in the hybridization step. The incubation temperature was 42°C. An equivalent blot was stained with methylene blue (Sambrook *et al.*, 1989) to make sure that the ribosomal RNA zones from the different samples were of approximately equal intensities. Sequence data were analysed on a VAX/VMS computer using programs of the University of Wisconsin Genetics Computer Group (Devereux *et al.*, 1984). The following protein and DNA databases were used in search of related proteins or motifs: Genbank, release 70; EMBL, release 30; MIPSX Protein, release 32; PIR 1/2/3 Protein, release 32; SWISS protein, release 21 and ProSite, release 8.1.

Gel filtration in 6 M guanidinium thiocyanate

To determine the subunit molecular weight of AP180, total coat protein was denatured by dialysis against 6 M guanidinium thiocyanate, 2 mM EDTA, 5 mM DTT and 50 mM Tris−HCl pH 8.0 and then applied to a Superose 6 gel filtration column, equilibrated in the same solvent. The column was operated at room temperature with a flow rate of 0.3 ml/min and calibrated with myosin, clathrin heavy and light chains, phosphorylase b, BSA, aldolase and hemoglobin. To determine the elution volume of AP180, β-type adaptor subunits and the γ subunit of the Golgi adaptor, fractions of 0.5 ml were collected and dialysed at room temperature against 8 M urea, 2 mM EDTA and 20 mM Tris−Cl pH 8.0, to remove the salt. Fractions were analysed by SDS−PAGE and immunoblotting.

Gel electrophoresis and immunoblotting techniques

Protein was regularly analysed on linear 7.5−19% gradient acrylamide gels with an increasing *bis* content of C = 1−2%. Previously we showed that gels with 1% *bis* resolve AP180 from the heavy chain of clathrin (Lindner and Ungewickell, 1991). The composition of buffers and the stacking gel corresponded to the method described by Laemmli (1970). Gels were either stained with Coomassie brilliant blue or electroblotted on to nitrocellulose paper for staining with antibodies. Antibody−antigen complexes were reacted with peroxidase-conjugated secondary rabbit antibodies against mouse IgG according to the manufacturer's recommendation (usually at dilutions of 1:500) and developed with 4-chloro-1-naphthol.

Acknowledgements

We are indebted to Dr V.Ohanian (University of Manchester, UK) for supplying us with high quality RNA preparations from various rat tissue. This work was supported in part by a project grant from the Deutsche Forschungsgemeinschaft (Schwerpunktprogramm 'Funktionelle Domänen zellulärer Strukturproteine') to E.U. (Un 43/3-5).

References

Ahle,S. and Ungewickell,E. (1986) *EMBO J.*, **5**, 3143−3149.
Ahle,S. and Ungewickell,E. (1990) *J. Cell Biol.*, **111**, 19−29.
Bauw,G., Van Damme,J., Puype,M., Vandekerckhove,J., Gesser,B., Ratz,G.P., Jauridsen,J.B. and Celis,J.E. (1989) *Proc. Natl. Acad. Sci USA*, **86**, 7701−7705.
Branch,W.T., Robbins,J. and Edelhoch,H. (1972) *Arch. Biochem. Biophys.*, **152**, 144−151.
Brodsky,F.M. (1988) *Science*, **242**, 1396−1402.
Campbell,C., Squicciarini,J., Shia,M., Pilch,P.F. and Fine,R.E. (1984) *Biochemistry*, **23**, 4420−4426.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387−395.
Goldstein,J.L., Brown,M.S., Anderson,R.G., Russell,D.W. and Schneider,W.J. (1985) *Annu. Rev. Cell Biol.*, **1**, 1−39.
Gratzer,W.B. and Beaven,G.H. (1968) *J. Phys. Chem.*, **73**, 2270−2275.
Keen,J.H. (1987) *J. Cell Biol.*, **105**, 1989−1998.
Keen,J.H. (1990) *Annu. Rev. Biochem.*, **59**, 415−438.
Keen,J.H. and Black,M.M. (1986) *J. Cell Biol.*, **102**, 1325−33.
Kirchhausen,T., Nathanson,K.L., Matsui,W., Vaisberg,A., Chow,E.P., Burne,C., Keen,J.H. and Davis,A.E. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2612−2616.
Kohtz,D.S. and Puszkin,S. (1988) *J. Biol. Chem.*, **263**, 7418−7425.
Kohtz,D.S. and Puszkin,S. (1989) *J. Neurochem.*, **52**, 285−295.
Laemmli,U.K. (1970) *Nature*, **227**, 680−685.
Lafer,E.M., Hrinya Tannery,N., Sousa,R.J., Zhou,S. and Groupp,E.R. (1989) In Rotundo,R.L. *et al.* (eds), *Advances in Gene Technology: Molecular Neurobiology and Neuropharmacology*. IRL Press, Oxford. pp. 73−79.
Lindner,R. and Ungewickell,E. (1991) *Biochemistry*, **30**, 9097−9101.
Lindner,R. and Ungewickell,E. (1992) *J. Biol. Chem.*, **267**, 16567−16573.
Morris,S.A., Ahle,S. and Ungewickell,E. (1989) *Curr. Opin. Cell Biol.*, **1**, 684−690.
Morris,S.A., Mann,A. and Ungewickell,E. (1990) *J. Biol. Chem.*, **265**, 3354−3357.

Murphy,J.-E., Pleasure,I.T., Puszkin,S., Prasad,K. and Keen,J.H. (1991) *J. Biol. Chem.*, **266**, 4401–4408.
Murphy,J-E. and Keen,J.H. (1992) *J. Biol. Chem.*, **267**, 10850–10855.
Pearse,B. (1988) *EMBO J.*, **7**, 3331–3336.
Prasad,K. and Lippoldt,R.E. (1988) *Biochemistry*, **27**, 6098–6104.
Robinson,M.S. (1990) *J. Cell Biol.*, **111**, 2319–2326.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sambrook,J., Fritsch,E.F., Maniatis,T. (1989) *Molecular Cloning, A Laboratory Manual*, 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Su,B., Hanson,V., Perry,D. and Puszkin,S. (1991) *J. Neurosci. Res.*, **29**, 461–473.
Ungewickell,E. and Oestergaard,L. (1989) *Anal. Biochem.*, **179**, 352–356.

Received on 14 September, 1992; revised on 26 October, 1992

Note added in proof

A paper describing the cloning and sequencing of F1–20 cDNA was recently published by Zhou *et al.* [(1992) *J. Neurosci*, **12**, 2144–2155].