Journal of Cell Science 111, 1909-1919 (1998)
Printed in Great Britain © The Company of Biologists Limited 1998
JCS4569

Characterisation of two articulins, the major epiplasmic proteins comprising the membrane skeleton of the ciliate *Pseudomicrothorax*

Irm Huttenlauch[1,*], Robert K. Peck[2], Uwe Plessmann[1], Klaus Weber[1] and Reimer Stick[1]

[1]Max Planck Institute for Biophysical Chemistry, Department of Biochemistry, Am Faßberg 11, D-37077 Göttingen, Federal Republic of Germany
[2]University of Geneva, Protistologie, 154, Route de Malagnou, CH-1224 Chêne-Bougeries, Switzerland
*Author for correspondence

Accepted 1 May; published on WWW 30 June 1998

SUMMARY

Most protists possess a unique membrane skeleton, the epiplasm, which is involved in pattern forming processes of the cell cortex and functions in maintaining cell shape. Articulins, a novel class of cytoskeletal proteins, are major constituents of the epiplasm. We have isolated cDNAs encoding the two major articulins of the ciliate *Pseudomicrothorax dubius*. Peptide sequence data confirm the identity of the cloned cDNAs encoding articulins 1 and 4. With the data presented here sequence information for all major articulins of ciliates as well as the distantly related euglenoids is available. Sequence comparison of the two newly characterised ciliate articulins with the previously determined sequences of p60, a minor articulin of the same species, and the two euglenoid articulins reveals general sequence principles and uncovers new features of this protein family. The hallmark of articulins is a central core domain of repetitive motifs of alternating valine and proline residues, the VPV-motif. These VPV-motif repeats are either 12-residues, or in some places, six residues long. Positively and negatively charged residues segregate in register with valine and proline positions. The VPV-motif is unique to articulins. The terminal domains flanking the core are generally hydrophobic and contain a series of hexa- or heptapeptide repeats rich in glycine and hydrophobic residues. The sequences of these short repeats are very similar in articulins of the same species but are not conserved between euglenoids and ciliates.

Key words: Cytoskeleton, Membrane skeleton, Epiplasm, Articulin, *Pseudomicrothorax*

INTRODUCTION

Many protists possess a characteristic cell shape and a highly elaborate surface architecture. Ciliates are among those that display the most sophisticated spatial arrangement of their cortical organelles (for review see Aufderheide et al., 1980; Frankel, 1989). In ciliates usually hundreds or even thousands of cortical units are aligned in longitudinal rows. Each unit consists of one or two cilia, each implanted on a basal body, and a set of membranous and cytoskeletal elements. The units are polar structures, and their precise and asymmetrical arrangement defines cell polarity i.e. the anterioposterior polarity (and thereby the swimming direction) and the left-right or circumferential asymmetry of these unicellular organisms. The cortical patterns are species specific and are faithfully reproduced at each cell division (Sperling et al., 1991). There is experimental evidence that the membrane skeleton is involved in pattern forming processes of the cortical elements and that it functions in maintaining cell shape (Peck, 1977, 1986; Dubreuil and Bouck, 1988). The membrane skeleton is part of the cytoskeleton. Unlike metazoans and a few protists such as *Dictyostelium* and *Physarum*, where the membrane skeleton is dominated by microfilamentous structures containing actin and a variety of actin binding proteins (for review see Schleicher and Noegel, 1992; Stockem and Brix, 1994), ciliates, dinoflagellates and euglenoids possess a unique cortical cytoplasm called the epiplasm.

The epiplasm is a rigid and often prominent proteinaceous layer organised as an almost continuous sheet or separated into individual plates (Peck, 1977; Dubreuil and Bouck, 1985; Grain, 1986; Bouck and Ngô, 1996). It is always in close contact with a membrane. In euglenoids it is the plasma membrane, in ciliates it is the inner membrane of the alveolus, a membrane system that is closely apposed to the plasma membrane (Peck, 1977; Bricheux and Brugerolle, 1986, 1987) and that is thought to be involved in regulation of the intracellular calcium level (Stelly et al., 1991). The epiplasm possesses a highly elaborate surface structure with different morphologies depending upon the position in the cell. Examination of purified epiplasm fragments by scanning electron microscopy reveals precisely positioned openings of defined sizes at sites, where in intact cells, the cortical organelles, such as the basal bodies, the parasomal sacs and the trichocysts, are positioned (Peck et al., 1991). The function of the epiplasm as a scaffold for the positioning of cortical organelles is highlighted by observations on trichocyst

formation. At points of emergence of trichocysts the epiplasm is interrupted forming trichocyst docking sites of characteristic configuration and position. Significantly, formation of docking sites is independent of whether mature trichocysts are actually inserted into the cortex (Peck, 1986).

The epiplasm of the ciliate *Pseudomicrothorax dubius* consists of two major groups of polypeptides, 78-80 kDa and 11-18 kDa, as well as a series of minor components of 18-62 kDa (Peck et al., 1991). The group of 78-80 kDa consists of at least three different polypeptides and numerous isoelectric variants. The quantitatively most prominent components of the epiplasmic polypeptides are polypeptides 1 and 4, while others (polypeptides 2 and 3) are isoelectric variants of polypeptide 1 (Peck et al., 1991).

Starting from existing peptide sequence information we determined the complete cDNA sequences for the major epiplasmic polypeptides 1 and 4 of the ciliate *P. dubius*. Both polypeptides are members of the articulin protein family, a novel family of cytoskeletal proteins first described in the euglenoid *Euglena gracilis* (Marrs and Bouck, 1992). With the two articulins characterised here, all members of this protein family are known, which comprise the major protein constituents of the epiplasm in both euglenoids (Marrs and Bouck, 1992) and ciliates (this study). Moreover, sequence comparison of the two major articulins of *P. dubius* with a previously characterised minor articulin of the same species (Huttenlauch et al., 1995) and the two major articulins of the euglenoid *E. gracilis* identifies several general features of this novel type of cytoskeletal protein. All articulins have a tripartite organisation. The hallmark of these proteins is a central domain of repetitive motifs rich in valine and proline. The consensus of these repeats is an alternation of valine and proline, the VPV-motif. Charged residues, when present, segregate in a characteristic manner. The P positions show nearly only negative charges, while the V positions harbour the positive charges. The terminal domains of articulins are generally hydrophobic. Moreover, the terminal domains of the four major articulins contain short (6-7 residues long) hydrophobic repeats. Although the sequences of these short repeats are not conserved between ciliate and euglenoid articulins, their occurrence suggests a common function.

MATERIALS AND METHODS

Cultures and epiplasm preparation of *P. dubius*

P. dubius strain N5b was cultivated, harvested, and the epiplasm was prepared as described (Peck et al., 1991).

Electrophoretic procedures and immunoblotting

SDS-PAGE using 8-10% (w/v) polyacrylamide slab gels was performed as described (Huttenlauch and Peck, 1991). The molecular mass values of the polypeptides were determined from mobility on SDS-PAGE compared with the following protein standards: phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), and carbonic anhydrase (30 kDa). Two-dimensional gel electrophoresis and electrophoretic transfer of proteins to nitrocellulose sheets was as described (Huttenlauch and Peck, 1991) except that transfer was carried out at 4 mA/cm^2 for 1 hour.

Immunolabeling of gel blots was carried out as described (Stick, 1988). Blots were incubated overnight at 4°C with culture supernatant of the mouse monoclonal antibody 4B5F3 (Curtenaz and Peck, 1992) diluted 1:15. Secondary antibody was peroxidase-conjugated goat anti-mouse IgG (Dianova, FRG) diluted 1:2,500. The immunocomplexes were detected by chemiluminescence using the SuperSignal CL-HRP substrate system (Pierce, USA) according to the instructions of the manufacturer and ECL Hyperfilm (Amersham, UK). For detection of ^{35}S-labelled proteins dried gels or blots were exposed to ECL Hyperfilm (Amersham).

Isolation of epiplasmic proteins, preparation of fragments and peptide sequences

Isolation of epiplasmic proteins, preparation of fragments and peptide sequencing was essentially done as described (Huttenlauch et al., 1995).

Oligonucleotides

Articulin 1: the previously determined amino acid sequence of a peptide of articulin 1 (peptide *a* of spot 1; Huttenlauch et al., 1995) was used to synthesize a pair of degenerated oligonucleotides for RT-PCR: sense primer P1 5′CARGARCCIGTIGCNGTNCC3′ (QEPVAVP) and antisense primer P2 5′TGIACIGCRTANGGYTG-NGG3′ (PQPYAVQ). Oligonucleotides deduced from the nucleotide sequence of the PCR product in the first round of RT-PCR for 5′-RACE I: P3 5′GACCTCGAATGGAGTGTC3′, 3′-RACE: P5 5′GCCGGAGTCGCTGGCGTC3′. Oligonucleotide for the second 5′-RACE: P4 5′GGAGGCACTGGATGCCTGGACGACGC3′, Oligonucleotides designed to amplify the complete open reading frame of articulin 1 flanked by restriction sites *Stu*I (5′) and *Xba*I (3′) for insertion into the pCS2+ vector: sense primer P5 5′CGGGCAGGCCTTTCTAATCAAATGGCTCAG3′, antisense primer P6 5′ATGGTCTAGAGCTTAGAGGACTCTGCTGTTGCC3′. 5′-oligonucleotide including an *Eco*RI restriction site designed to amplify the open reading frame of articulin 1 for insertion into the pCS2+Flag-tag vector: sense primer P7 5′CGGAATTCA-ATGGCTCAGGTTGTAGGAACTCC3′.

Articulin 4: two previously determined amino acid sequences of peptides of articulin 1 (spot 1; Huttenlauch et al., 1995) were used to synthesize a pair of degenerated oligonucleotides for RT-PCR: peptide *f* of spot 1 (DVPVERP) sense primer P8 5′GAYGTHC-CHGTHGARAGACC3′, peptide *l* spot 1 (DRVVPVP) antisense primer P9 5′GGDACDGGDACDACTCTRTC3′. Oligonucleotides deduced from the partial cDNA 8-11 for the amplification of the 5′ end of the cDNA: antisense primer P10 5′CCTGGA-GACCGACTCTGGAGCCGACG3′; antisense nested primer P11 5′GGACCTGCTGTCCGATGCGTGGGCTG3′. Oligonucleotides designed to amplify the open reading frame of articulin 4 flanked by restriction sites *Bam*HI and *Hin*dIII for insertion into the pINDU expression vector: sense primer with *Bam*HI site P12 5′CGGG-ATCCATGGCTTTTGTTAACAACAACC3′; antisense primer with *Hin*dIII site P13 5′CCCAAGCTTGGATATTATCAGAG-CTAGAGGGC3′. Oligonucleotides designed to amplify the open reading frame of articulin 4 flanked by restriction sites *Eco*RI and *Xho*I to amplify the open reading frame of articulin 4 for insertion into the pCS2+Flag-tag vector: sense primer P14 5′CGGAATTCAATGGCTTTTGTTAACAACAACC3′; antisense primer P15 5′CCGCTCGAGGGATATTATCAGAGCTAGAGGG3′.

Polymerase chain reaction (PCR)

Total RNA and poly(A)$^+$ RNA from *P. dubius* was isolated as described (Huttenlauch et al., 1995). First strand cDNA was synthesized from total RNA using the Gene Amp RT-PCR kit from Perkin-Elmer Cetus (USA). 1 μg total RNA was primed per reaction with a mixture of random hexanucleotides following the manufacturer's protocol. PCR was carried out with combinations of degenerate primers (0.2-0.4 μM each) deduced from the peptide sequences as specified above. Cycling parameters for the primer pair P1/P2 were: initial denaturation 2 minutes, 94°C followed by 35

cycles (1 minute, 94°C; 1 minute 56°C, 30 seconds, 72°C) and a final polymerisation step (10 minutes, 72°C). Reaction products were separated on an agarose gel (2.3%), the product of the expected size was electroeluted, reamplified under identical conditions and directly cloned using the pGEM-T vector cloning kit (Promega, USA) according to the manufacturer's instructions. Cycling parameters for the primer pair P8/P9 and subcloning into the *Eco*RV restriction site of the Bluescript KS vector (Stratagene Corp., USA) were as described (Huttenlauch et al., 1995). Amplification of the 5′ end of the partial cDNA of articulin 4 was done using the 5′-Amplifinder RACE Kit (Clontech, USA), and isolation of the complete cDNA sequence of articulin 1 was done using a Marathon cDNA amplification kit and an Advantage cDNA PCR kit (Clontech, USA) following the protocol of the manufacturer.

cDNA library construction, screening, and sequencing

Construction and screening of a *P. dubius* cDNA library was as described (Huttenlauch et al., 1995). Double stranded sequencing was done with internal primers using a Prism Dye Deoxy Terminator Cycle Sequencing kit (Applied Biosystems, USA) and an ABI 392 DNA Sequencer.

Plasmid constructions and protein expression

To obtain a clone containing the entire coding sequence of articulin 4, the *Xba*I/*Apa*I restriction fragment of plasmid pBst SK 8-11 containing the bases +34 to +1,727 of the cDNA (numbering with respect to the cDNA sequence) including polylinker sequences of the pBst SK vector was cloned into pGEM-T 5′ articulin 4 that had been linearized with *Xba*I/*Apa*I.

To generate expression constructs of either articulin 1 or 4 the entire open reading frames were amplified from an Amplifinder-ligated cDNA (articulin 1) or from the pGEM-T plasmid containing the full-length cDNA (articulin 4) by PCR using the primer pairs listed above. Amplicons were cloned into the *Eco*RI/*Xba*I sites (articulin 1) and the *Eco*RI/*Xho*I sites (articulin 4) of the pCS2+Flag-tag vector (Rupp et al., 1994). The resulting proteins contain at their NH_2-termini the flag-epitope followed by two vector derived amino acids (MDYKDDDDKNS) (Flag-tag underlined). Details of the PCR amplification, cloning, and in vitro translation of proteins using the TNT-coupled transcription translation system (Promega Corp., USA), and bacterial expression in *E. coli* BL 21 have been described (Huttenlauch et al., 1995).

RESULTS

The protein pattern of the isolated epiplasm of *Pseudomicrothorax dubius* consists of two major groups of polypeptides with 78-80 kDa and 11-18 kDa, respectively, as well as several minor components of 18-62 kDa. The group of 78-80 kDa resolves into at least six spots by two dimensional gel electrophoresis (spots 1-6 in Fig. 1; Peck et al., 1991). Peptide map analysis indicated that spots 1, 2, and 3 represent isoelectric variants of the same polypeptide, while spot 4 represents a different polypeptide (Peck et al., 1991). Thus, polypeptide 1 including its isoelectric variants together with polypeptide 4 comprise the two major constituents of the high molecular mass group of epiplasmic proteins of *P. dubius*. Based on peptide sequences obtained earlier and the cDNA sequence information to be reported here we named the two polypeptides articulin 1 and articulin 4, respectively.

Molecular cloning of articulin 1 cDNA

Previously we obtained peptide sequences from both articulins 1 and 4 as well as a complete cDNA sequence encoding a minor epiplasmic polypeptide, named p60 (Huttenlauch et al., 1995). The cDNA encoding part of the articulin p60 was isolated starting from a PCR product generated with primers deduced from peptide sequences of articulin 4. From the analysis of this cDNA sequence we inferred that the cross hybridisation of the PCR primers was due to the high degree of amino acid similarity between these polypeptides. Moreover, cDNA library screening using an articulin 1-encoding PCR fragment resulted in the isolation of a cDNA encoding articulin 4 (see below). Since this cross hybridisation occurred even under highly stringent hybridisation conditions, in the present study we used a different PCR based method to clone the cDNA encoding articulin 1.

One of the peptide sequences obtained for articulin 1 (peptide *a* of spot 1 in Huttenlauch et al., 1995) was of sufficient length to allow the design of a primer pair for amplification of the respective cDNA fragment by RT-PCR. Size prediction allowed the isolation and subsequent cloning of the relevant cDNA fragment out of a complex mixture of (false) PCR products (data not shown). The identity of this fragment was verified by sequencing. The fragment yielded sufficient nucleotide sequence information (35 nucleotides) to synthesize a sense and an antisense primer for subsequent PCR amplifications using an anchor ligated cDNA library in both the 5′ and 3′ directions (see Materials and Methods for details). The 5′ cDNA fragment, 1,492 nucleotides in length, encoded

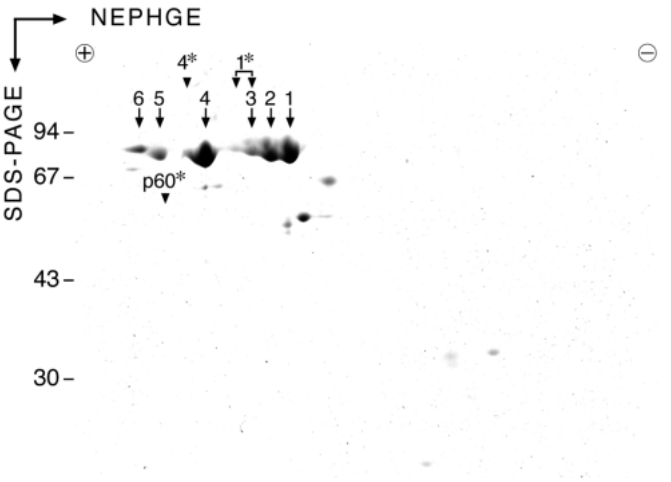

Fig. 1. Two-dimensional gel electrophoretic analysis of epiplasm from the ciliate *P. dubius* and in vitro translated acticulins. Epiplasmic polypeptides were mixed with radiolabeled in vitro translated articulins 1, 4 (this report), and p60 (Huttenlauch et al., 1995) prior to electrophoresis. They were separated by NEPHGE in the first dimension, and by SDS-PAGE (10%) in the second dimension. A Coomassie Blue stained gel is shown. Major spots of the 78-80 kDa group of polypeptides are indicated by arrows and numbered 1-6. Positions of the in vitro translated articulins in the NEPHGE dimension are indicated by arrowheads and marked by an asterisk. Note that in vitro translated articulin 1 is separated into two spots. In vitro translated articulins 1 and 4 migrate identically in the SDS dimension to the corresponding polypeptides of the epiplasm. The arrowheads for these two polypeptides (1* and 4*) have been placed at a different level for clarity. The relative mobilities of marker proteins: phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), and carbonic anhydrase (30 kDa) are indicated.

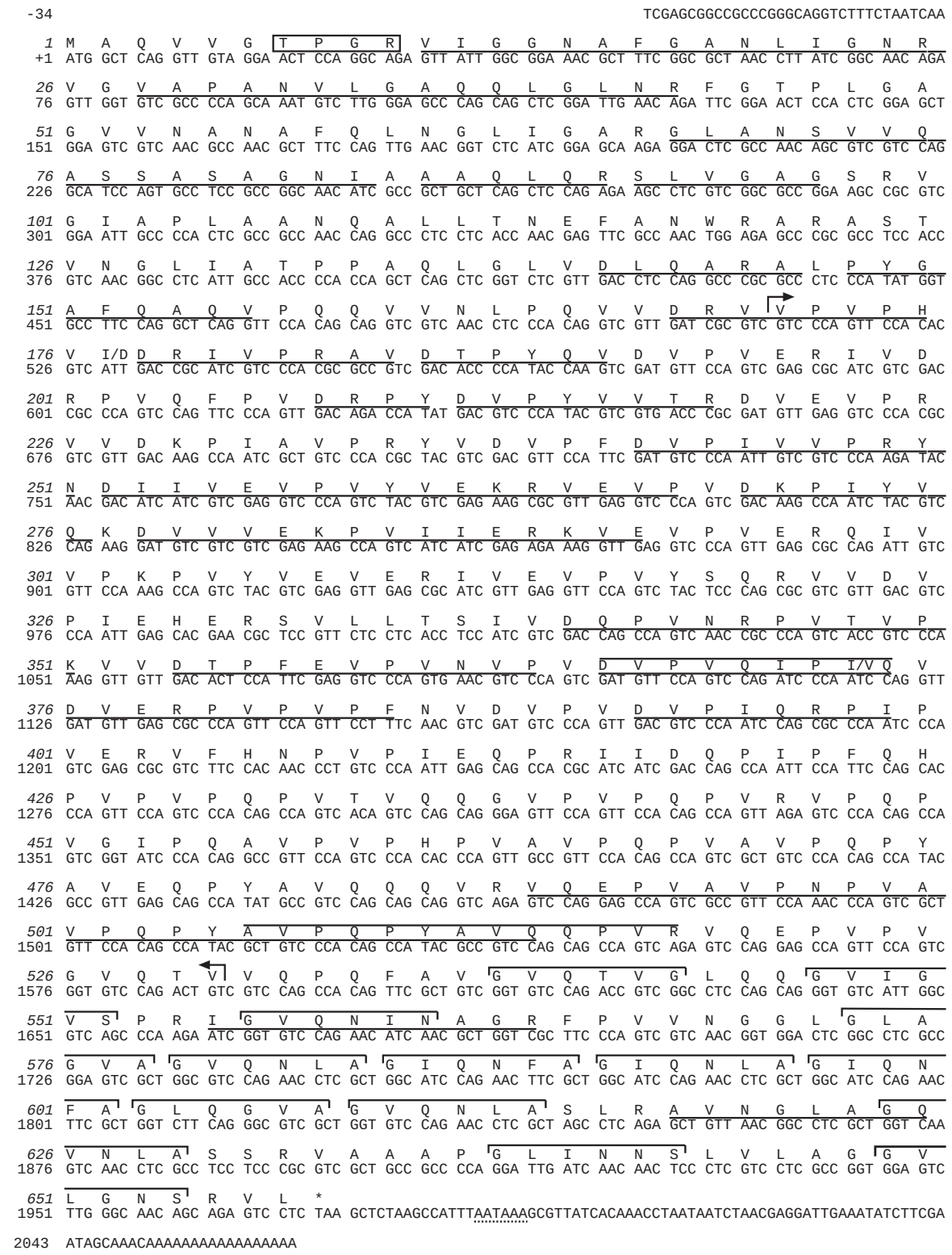

Fig. 2. cDNA sequence and deduced amino acid sequence of *P. dubius* articulin 1. Amino acid residues are represented by single letters above the first base of each codon. Numbering of amino acids is in italics. Numbering of nucleotides starts with +1 at the ATG initiation codon. The stop codon TAA is designated by an asterisk. The borders of the central VPV-motif domain (residues 171-530) are marked by arrows. Sequences determined by peptide sequence analysis are underlined. Overlapping sequences that were independently determined from two different peptides are overlined. Two sequence deviations including one from the predicted cDNA sequence are given after a slash. Hexapeptide repeats in the tail domain are indicated by brackets above the sequence. A putative $p34^{cdc2}$ protein kinase consensus site (*TPGR)* (Moreno and Nurse, 1990) in the head domain is boxed. A polyadenylation signal (*AATAAA*) is designated by a dotted line. The sequence data are available from GenBank under accession number AF039086.

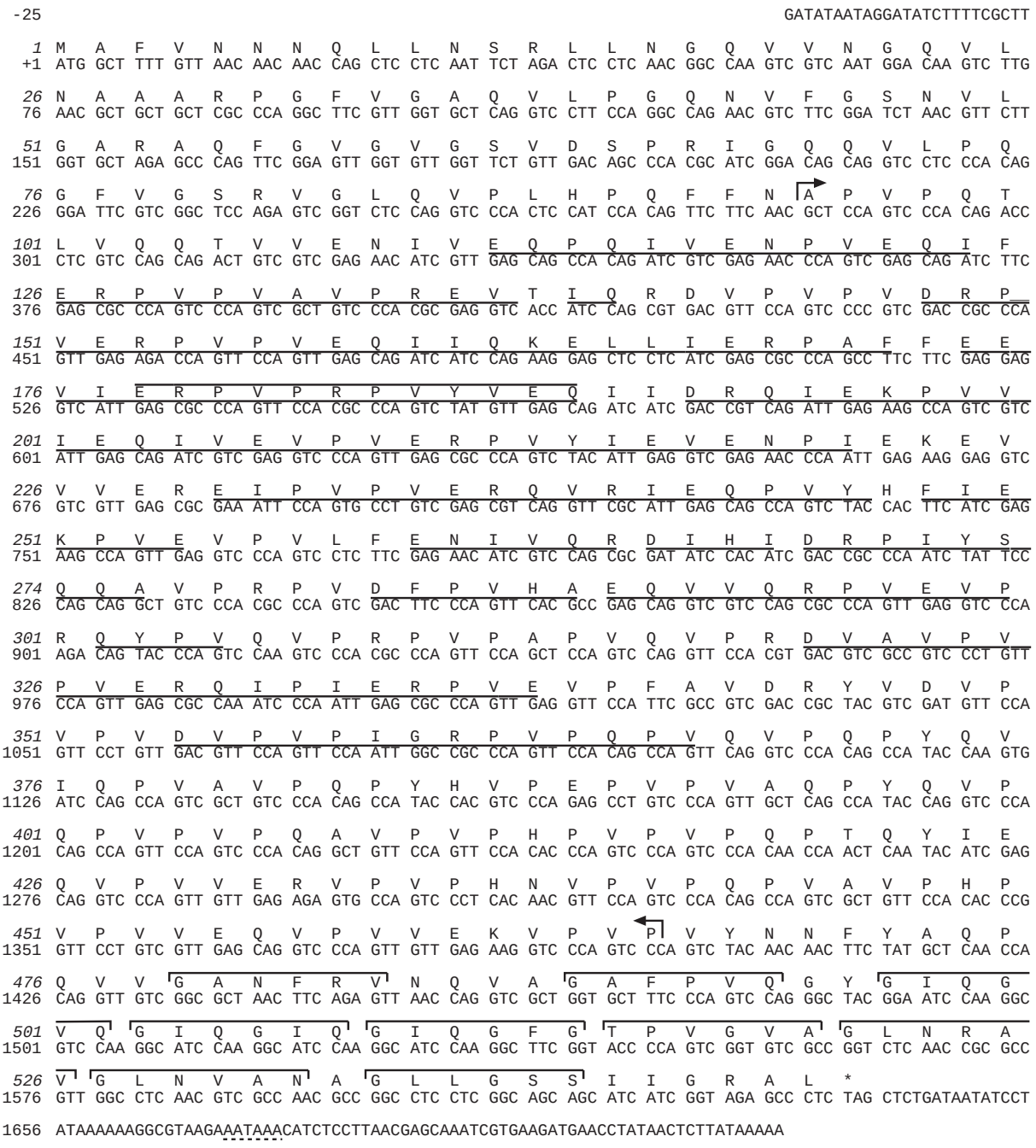

Fig. 3. cDNA sequence and deduced amino acid sequence of *P. dubius* articulin 4. Amino acid residues are represented by single letters above the first base of each codon. Numbering of amino acids is in italics. Numbering of nucleotides starts with +1 at the ATG initiation codon. The stop codon TAG is designated by an asterisk. The borders of the central VPV-motif domain (residues 95-466) are marked by arrows. Sequences determined by peptide sequence analysis are underlined. An overlapping sequence that was independently determined from two different peptides is overlined. Hexapeptide repeats in the tail domain are indicated by brackets above the sequence. A polyadenylation signal (*AATAAA*) is designated by a dotted line. The sequence data are available from GenBank under accession number AF039087.

a VPV-repetitive protein fragment that showed perfect matches with several of the previously determined peptide sequences of articulin 1. However, it did not contain any methionine residue, indicating that it lacked the NH_2 terminus of the polypeptide. To obtain the 5′ end of the coding region, we performed a second round of 5′ PCR with the anchor ligated cDNA library, this time with a gene specific oligonucleotide priming approximately 250 nucleotides downstream of the 5′ end of the cDNA fragment. This amplification added another 38 nucleotides to the existing sequence information, including the start methionine codon, and 34 nucleotides of 5′ untranslated region (for the criteria of the completeness of the open reading frame, see below). Sequencing of the 3′ cDNA fragments revealed heterogeneity of the amplified products. While two of

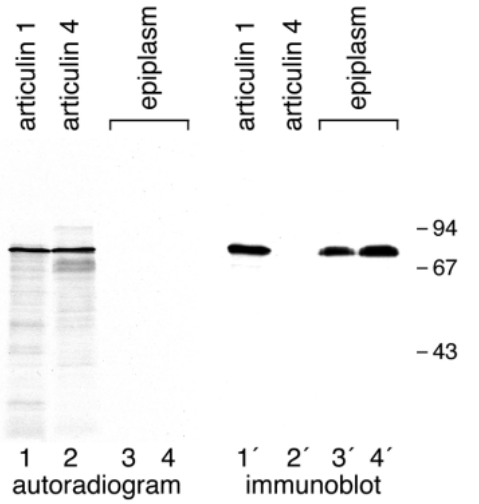

Fig. 4. Gel electrophoretic comparison of in vitro translated *P. dubius* articulins 1 and 4 and epiplasmic articulins. Articulin 1 (lane 1) and articulin 4 (lane 2) translated in a coupled in vitro transcription-translation system in the presence of [^{35}S]methionine, and two different amounts of polypeptides of an epiplasmic isolate of *P. dubius* (lanes 3, 4) were separated on SDS-PAGE (8%). Polypeptides were electrophoretically transferred to nitrocellulose membrane and first detected by autoradiography (autoradiogram, lanes 1-4) and then probed and detected with monoclonal antibody 4B5F3 as described in Materials and Methods (immunoblot, lanes 1´-4´). Note that in vitro translated articulins 1 and 4 as well as epiplasmic articulin 1 as detected by antibody 4B5F3 all migrate with nearly the same mobility. Articulin 4 is not detected by antibody 4B5F3. The relative mobility of marker proteins phosphorylase b (94 kDa), bovine serum albumin (67 kDa), and ovalbumin (43 kDa) are indicated.

the three clones analysed encoded the desired cDNA fragment of articulin 1, another clone encoded the corresponding fragment of articulin 4, for which we had isolated a cDNA by other techniques (see below). The latter observation points again to the high degree of sequence similarity among genes encoding members of the articulin protein family. The 3′ cDNA fragment encoded the complete COOH terminus of articulin 1, a 79 nucleotide long 3′ untranslated region and a short poly(A) tail. The anchor ligated cDNA library was finally used to generate a cDNA encoding the entire coding region of articulin 1 (see Materials and Methods). This cDNA codes for a polypeptide of 657 amino acid residues.

Amino acid sequence information obtained by peptide sequence analysis (Huttenlauch et al., 1995) covers about 40% of the polypeptide and is from regions scattered throughout the entire polypeptide (underlined in Fig. 2) The peptide sequences are in perfect agreement with the amino acid sequence deduced from the cDNA with two exceptions. At position 177 an aspartic acid residue is found in the peptide sequence, while the cDNA predicts an isoleucine residue. Neither of the residues violates the sequence rules of articulin VPV-motifs (see below). The biological significance of this discrepancy, if any, remains unknown. The other exception demonstrates the existence of two different alleles or two closely related genes for articulin 1. Two peptide sequences were obtained, one encodes an isoleucine residue at position 373 that matches the sequence deduced from the cDNA (Fig. 2), the other peptide shows a valine residue at this position. The otherwise perfect match between the peptide sequence determination and the cDNA sequence demonstrates that we have isolated a cDNA encoding articulin 1 of *P. dubius*.

Molecular cloning of articulin 4 cDNA

A partial cDNA clone (pBst SK 8-11) encoding articulin 4 was isolated by screening a cDNA library of *P. dubius* under highly stringent hybridisation conditions. The DNA probe used in this screening was obtained by RT-PCR amplification with degenerate primers deduced from peptide sequences of articulin 1. As mentioned above, the close sequence similarity of articulin proteins, especially in the VPV-repeat motifs, made cross hybridisation common in cDNA screening and PCR amplification reactions (see also Huttenlauch et al., 1995). cDNA 8-11, 1707 nucleotides in length, encoded the COOH terminus of a VPV-repetitive protein including a 87 nucleotide long 3′ untranslated region with a poly(A) addition motif AAAUAA, 47 nucleotides upstream of a short poly(A) tail. However, the protein did not contain any methionine residue, indicating that it lacked the NH_2 terminus. We obtained the complete coding region by carrying out a primer extension reaction in conjunction with a PCR amplification with poly(A^+) RNA, ligation of an anchor oligonucleotide to the 3′ end of the single stranded cDNA, followed by PCR with a cDNA specific nested primer and a primer complementary to the anchor sequence (see Materials and Methods). This amplification added 25 nucleotides to the 5′ end of the cDNA including the start methionine codon and a 40 nucleotide long 5′ untranslated region with two in frame stop codons. The complete cDNA encodes a polypeptide of 545 amino acid residues (Fig. 3). Peptide sequences obtained for articulin 4 (Huttenlauch et al., 1995) cover about 30% of the polypeptide and are clustered within the central VPV-motif domain (underlined in Fig. 3). They are in perfect agreement with the amino acid sequence deduced from the cDNA. Thus, the cloned cDNA encodes *P. dubius* articulin 4. Moreover, the presence of two in frame stop codons upstream of the start methionine codon, as well as the presence of all three stop codons (TAG, TGA, TAA) immediately downstream of the open reading frame, proves that the cDNA covers the entire coding region.

Articulins 1 and 4 migrate aberrantly in SDS-PAGE

Comparison of the molecular masses predicted from the cDNAs with the relative molecular masses determined by SDS-PAGE revealed an obvious discrepancy for both articulins. While the calculated molecular masses of articulin 1 and 4 are 69,728 and 59,858, respectively, both proteins migrate, depending on the percentage of the polyacrylamide gel, with nearly the same mobility at 78-80 kDa in SDS-PAGE (Fig. 1; Peck et al., 1991). We therefore translated both polypeptides in vitro and compared their migration with that of articulins isolated from the epiplasm. Since our initial attempts to express the two polypeptides initiating at the endogenous start methionine codons failed, we cloned the cDNAs in frame into the pCS2+Flag vector downstream of the Flag-epitope, that carries its own start methionine. The flag-epitope is eight amino acids long and together with three vector derived amino acids adds only 1.3 kDa to the mass of the resulting fusion protein (for details of the sequence see Materials and

Methods). Both constructs were efficiently translated in vitro in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine (Fig. 4, lanes 1 and 2). More significantly, both polypeptides migrated with nearly the same mobility of about 75-80 kDa. For a direct comparison of the migration of the in vitro translated proteins with articulins isolated from the epiplasm we ran aliquots of epiplasmic preparations and radiolabeled in vitro translated polypeptides on the same gel. The separated proteins were transferred to nitrocellulose and radiolabeled polypeptides were detected by autoradiography (Fig. 4, lanes 1 and 2). The filter was then processed for immunodetection. Monoclonal antibody 4B5F3, which specifically recognises articulins (Curtenaz and Peck, 1992; Curtenaz et al., 1994; I. Huttenlauch and R. Stick, unpublished), shows strong reactivity with articulin 1 (Fig. 4, lane 1′) but does not react with articulin 4 (Fig. 4, lane 2′). The immunoblot reveals that in vitro translated articulin 1 (Fig. 4, lane 1′) has nearly the same mobility as articulin 1 of the epiplasm (Fig. 4, lanes 3′ and 4′). Thus we conclude, that the cDNA used to program the in vitro translation system represents the entire coding region of articulin 1. Moreover, the striking difference between the calculated and the apparent masses in SDS-PAGE (about 25% in the case of articulin 4) seem to be intrinsic properties of the polypeptides and are probably not due to extensive posttranslational modifications in vivo. This is also supported by experiments with bacterially expressed articulin 4. For this purpose, the complete coding region of articulin 4 was cloned into a bacterial expression vector (pINDU; Bujard et al., 1987). The resulting fusion protein contained 545 amino acids of articulin 4, as well as four vector derived amino acids (MRGS) at the NH_2 terminus. This polypeptide also migrated with an aberrant mobility of 78-80 kDa in SDS-PAGE (data not shown). In addition, the latter result demonstrates, that the aberrant mobility is not due to the presence of the flag-epitope.

Two dimensional gel electrophoresis resolves epiplasmic articulin 1 into at least three isoelectric variants (Fig. 1, spots 1-3; Peck et al., 1991). Co-electrophoresis in two dimensional separations shows that epiplasmic articulins migrate at more basic positions in the pH-gradient than the in vitro translated polypeptides (Fig. 1, compare 1-3 with 1*, and 4 with 4*). This is especially evident for articulin 1 and cannot be explained by the presence of the flag-epitope, which shifts the pI by only 0.1 pH units. In vitro translated articulin 1 is resolved into two

***P. dubius* articulin 1** **30 repeats**

1	V	P	V	P	H	V	I	D	R	I	V	P
2	R	A	V	D	T	P	Y	Q	V	D	V	P
3	V	E	R	I	V	D	R	P	V	Q	F	P
4	V	D	R	P	Y	D	V	P	Y	V	V	T
5	R	D	V	E	V	P	R	V	V	D	K	P
6	I	A	V	P	R	Y	V	D	V	P	F	D
7	V	P	I	V	V	P	R	Y	N	D	I	I
8	V	E	V	P	V	Y	V	E	K	R	V	E
9	V	P	V	D	K	P	I	Y	V	Q	K	D
10	V	V	V	E	K	P	V	I	I	E	R	K
11	V	E	V	P	V	E	R	Q	I	V	V	P
12	K	P	V	Y	V	E	V	E	R	I	V	E
13	V	P	V	Y	S	Q	R	V	V	D	V	P
14	I	E	H	E	R	S	V	L	L	T	S	I
15	V	D	Q	P	V	N	R	P	V	T	V	P
16	K	V	V	D	T	P	F	E	V	P	V	N
17	V	P	V	D	V	P	V	Q	I	P	I	Q
18	V	D	V	E	R	P	V	P	V	P	F	N
19	V	D	V	P	V	D	V	P	I	Q	R	P
20	I	P	V	E	R	V	F	H	N	P	V	P
21	I	E	Q	P	R	I	I	D	Q	P	I	P
22	F	Q	H	P	V	P	V	P	Q	P	V	T
23	V	Q	Q	G	V	P	V	P	Q	P	V	R
24	V	P	Q	P	V	G	I	P	Q	A	V	P
25	V	P	H	P	V	A	V	P	Q	P	V	A
26	V	P	Q	P	Y	A	V	E	Q	P	Y	A
27	V	Q	Q	Q	V	R	V	Q	E	P	V	A
28	V	P	N	P	V	A	V	P	Q	P	Y	A
29	V	P	Q	P	Y	A	V	Q	Q	P	V	R
30	V	Q	E	P	V	P	V	G	V	Q	T	V
	1		3		5		7		9		11	
consensus	**V**	**p**	**V**	**P**	**V**	**p**	**V**	**p**	**v**	**p**	**V**	**p**
V	21	2	15	1	16	2	17	2	10	2	16	1
P	0	12	0	15	0	11	0	10	0	13	0	11
-	0	10	1	9	0	5	0	7	1	5	0	4
+	4	0	2	0	7	1	6	0	3	1	4	3
charge consensus	**+**	**-**		**-**	**+**	**-**	**+**	**-**	**+**	**-**	**+**	

***P. dubius* articulin 4** **31 repeats**

	1		3		5		7		9		11	
consensus	**v**	**p**	**V**	**p**	**V**	**P**	**-**	**p**	**V**	**p**	**V**	**p**
V	12	0	16	0	17	2	10	1	16	2	16	4
P	0	15	0	12	0	16	0	12	0	14	0	11
-	2	6	0	10	0	5	3	6	1	9	0	7
+	3	2	6	0	2	0	4	1	7	0	5	0
charge consensus		**-**	**+**	**-**	**+**	**-**		**-**	**+**	**-**	**+**	**-**

***P. dubius* articulin p60** **29 repeats**

	1		3		5		7		9		11	
consensus	**V**	**p**	**v**	**p**	**V**	**P**	**V**	**-**	**v**	**-**	**V**	**p**
V	17	1	12	0	16	1	16	1	12	3	15	2
P	0	11	0	11	0	10	0	8	0	9	0	12
-	0	11	0	10	0	6	0	7	0	6	0	6
+	5	0	3	1	7	0	4	0	2	0	6	1
charge consensus	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**

articulin repeat motif

	1		3		5		7		9		11	
sequence	**V**	**p**	**V**	**p**	**V**	**p**	**v**	**p**	**v**	**p**	**V**	**p**
charge	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**	**+**	**-**

Fig. 5. Sequence alignment of VPV-motifs of the central domain of *P. dubius* articulin 1 and quantitation of the 12-residue VPV-motifs of all three known *P. dubius* articulins. The sequence of the core domain of articulin 1 has been arranged as consecutive 12-residue repeats without linker sequences. Note, however, that in the COOH-terminal third of the core domain the repeat is six rather than 12 residues. Amino acid sequences for articulin 1 (residues 171-530) and articulin 4 (residues 95-466) were deduced from the nucleotide sequences of the corresponding cDNAs (Figs 2, 3). Data for articulin p60 (residues 142-489) were taken from Huttenlauch et al. (1995). A consensus sequence is given above each quantitation. Uppercase letters, residues in ≥50% of the repeats; lowercase letters, residues in ≥33% of the repeats. In the consensus '-' indicates that none of the residues reaches ≥33% occupancy. The number of valine (V) and proline (P), as well as positively (+) and negatively (–) charged residues for each VPV-motif position is recorded. A charge consensus is given below each quantitation. +, majority of charged residues positively charged; –, majority of charged residues negatively charged. A consensus 'articulin repeat motif' for all known articulins (*P. dubius* articulins 1, 4, and p60, and *E. gracilis* 80 kDa and 86 kDa) is given at the bottom. Data from *E. gracilis* were taken from Marrs and Bouck (1992). 12-residue VPV-motifs of all articulins were aligned consecutively without any linker sequences, resulting in 29 motifs for articulin p60, 30 motifs for articulin 1, 31 motifs for articulin 4, 35 motifs for *E. gracilis* articulin p80 kDa (residues 77-496), and 35 motifs for p86 kDa (residues 113-532). Note that all uneven positions with the exception of position 9 show ≥50% occupancy by valine, that all even positions show ≥33% occupancy by proline, and that the prevalence of charged residues is strictly alternating.

articulin 1 — 13 repeats

	1	2	3	4	5	6									
1	G	V	Q	T	V	G	L	Q	Q						
2	G	V	I	G	V	S	P	R	I						
3	G	V	Q	N	I	N	A	G	R	F	P	V	V	N	G	G	L
4	G	L	A	G	V	A									
5	G	V	Q	N	L	A									
6	G	I	Q	N	F	A									
7	G	I	Q	N	L	A									
8	G	I	Q	N	F	A									
9	G	L	Q	G	V	A									
10	G	V	Q	N	L	A	S	L	R	A	V	N	G	L	A
11	G	Q	V	N	L	A	S	S	R	V	A	A	A	P	
12	G	L	I	N	N	S	L	V	L	A	G				
13	G	V	L	G	N	S	R	V	L	*					
consensus	**G**	**v/l/i**	**Q**	**N/g**	**l/v**	**A**									
frequency	13	12	8	8	8	8									

articulin 4 — 9 repeats

	1	2	3	4	5	6									
1	G	A	N	F	R	V	N	Q	V	A					
2	G	A	F	P	V	Q	G	Y							
3	G	I	Q	G	V	Q									
4	G	I	Q	G	I	Q									
5	G	I	Q	G	F	G									
6	T	P	V	G	V	A									
7	G	L	N	R	A	V									
8	G	L	N	V	A	N	A								
9	G	L	L	G	S	S	I	I	G	R	A	L	*		
consensus	**G**	**l/i**	**n/q**	**G**	**v**	**q**									
frequency	8	6	6	5	3	3									

Fig. 6. Sequence alignment of the hexapeptide repeats in the tail domains of *P. dubius* articulins 1 and 4. Amino acid sequences for articulin 1 (residues 538-657) and articulin 4 (residues 479-545) were deduced from the nucleotide sequences of the corresponding cDNAs (Figs 2, 3). A sequence consensus is given below each alignment together with the frequency of the consensus residues. Uppercase letters, residues in ≥50% of the repeats; lowercase letters, residues in ≥33% of the repeats. Note that in the repeats of both articulins, position 1 is occupied almost exclusively by glycine, while positions 2 and 5 are generally hydrophobic.

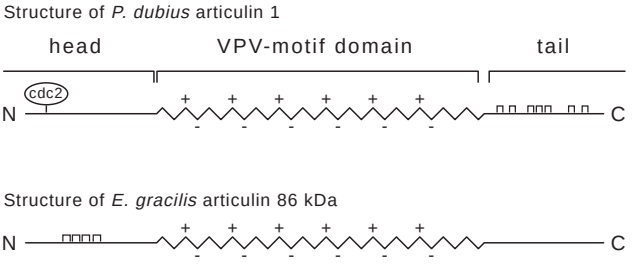

Fig. 7. Diagrammatic model of two representatives of articulins as deduced from the primary sequence data. The central VPV-motif domain is drawn as a zig-zag line to indicate the predicted β-sheet. Short peptide repeats within the terminal domains are indicated as brackets. For clarity fewer than the actual number of repeats is drawn in the tail domain of *P. dubius* articulin 1. Note that the sequences of the repeats within the terminal domains differ between *Pseudomicrothorax* and *Euglena* articulins. cdc2: position of putative $p34^{cdc2}$ protein kinase consensus site in the head domain of *P. dubius* articulin 1. Another cdc2 consensus site is present in *P. dubius* articulin p60. + and −: charge segregation within the central VPV-repeat motif domains in a hypothetical β-sheet structure.

spots. The more basic spot comigrates exactly with isoelectric variant 3 (Fig. 1). The differences in isoelectric points could be due to posttranslational modifications that do not occur, or only partially occur in the rabbit reticulocyte lysate. The nature of these modifications is currently unknown. However, previous data show that *P. dubius* articulins are neither glycosylated (Huttenlauch et al., 1991) nor phosphorylated. The latter conclusion is based on two observations: firstly, articulins are not labelled when cells are grown in [^{32}P]orthophosphate (Peck et al., 1991), and secondly, the pattern of isoelectric variants of articulin 1 remains unchanged when epiplasmic proteins are treated with alkaline phosphatase prior to two-dimensional gel electrophoresis (data not shown).

Sequence features of articulins 1 and 4 of *P. dubius*

Based on the presence of VPV-motifs, revealed by partial peptide sequences, both polypeptides were classified as articulins (Huttenlauch et al., 1995). To date, two articulins have been described from the membrane skeleton of the euglenoid *Euglena* (Marrs and Bouck, 1992), and one articulin from the epiplasmic membrane skeleton of the ciliate *Pseudomicrothorax* (Huttenlauch et al., 1995). The latter articulin, designated p60, is a quantitatively minor component of the epiplasm. Sequence determination of the complete coding regions of the two major articulins of the ciliate *P. dubius*, reported here, allows a more detailed analysis of the features of the two polypeptides and identifies the general sequence principles of articulins.

All known articulins show a three domain organisation. The central domain is characterised by the tandem arrangement of 12 residue long VPV-motifs. Articulins 1 and 4 of *P. dubius* show 30 and 31 of these motifs. They are arranged without any linker sequences. The motifs show an alternating pattern of valine and proline residues. Fig. 5 shows a sequence alignment of the VPV-motifs of articulin 1 and gives the abundance of these two residues at individual V- and P-positions together with a consensus for all three known *P. dubius* articulins. Uppercase letters represent occupancy in >50% of the repeats, while lowercase letters indicate >33% occupancy (see Fig. 5 for details). While the majority of the V-positions shows an occupancy of >50% valine only a few of the P-positions reach 50% occupancy by proline. More significantly, however, V-positions never show a proline while exclusion of valine from P-positions is less strict. The second remarkable feature of the VPV-motifs is the segregation of charged residues. The P-positions harbour the majority of negatively charged residues, while positively charged residues are found preferentially in V-positions. On average, valine plus positively charged residues or proline plus negatively charged residues make up nearly two thirds of the residues of the respective positions. Of all other amino acid residues only two, i.e. glutamine and isoleucine, are relatively abundant in the central core domain. It should be noticed, however, that in some parts the repeat length might be six residues rather than 12 residues as for example in the COOH-terminal third of the

central domain of articulin 1. The consensus of this repeat is VPQPVA (Fig. 5).

The central domain is flanked by hydrophobic NH_2-terminal and COOH-terminal domains in both polypeptides. These terminal domains are rich in glycine, alanine, leucine, and asparagine residues. They contain no (articulin 4) or very few (articulin 1) negatively charged residues. Positively charged residues are exclusively arginine, and they are more or less evenly distributed throughout the terminal domains of both articulins.

In the two *Euglena* articulins a heptapeptide around the sequence APVTYGA is found four times in either the head (*E. gracilis* p86) or in the tail domain (*E. gracilis* p80). We therefore screened the terminal domains of both *P. dubius* articulins for the presence of repeated sequence motifs. As mentioned above, the terminal domains of the two polypeptides are rich in glycine residues. Most glycine residues are followed by a hydrophobic residue. Although a certain similarity between short sequence motifs grouped around these glycine residues is noted, no clear consensus of a repeated motif is found in the head domain of either articulin 1 or 4. In contrast, the tail domains of both *P. dubius* articulins contain a series of hexapeptide repeats of similar sequence. Alignments of these repeats are shown in Fig. 6 together with the consensus sequence. The tail domain of articulin 1 harbours 13 short repeats while 9 short repeats are present in the tail domain of articulin 4. In both polypeptides a region in the middle of the hexapeptide repeats is found, where seven (articulin 1) and six (articulin 4) hexapeptides, respectively, are arranged consecutively without any linker sequences (Fig. 6). Position one of the repeats is almost exclusively occupied by glycine residues while positions two and five are generally hydrophobic. There is a preference for glutamine residues in position three which is especially clear within the consecutive hexapeptide repeats (Fig. 6). In position four asparagine is most frequently found in articulin 1 and glycine is preferred in articulin 4. It should be noticed, that three hexapeptide repeats (hexapeptides 3-5 in Fig. 6) of articulin 4 could also be broken up into a series of five tripeptides around the sequence GIQ, a motif that is also present nine times in the tail domain repeat region of articulin 1, although here these motifs are not arranged in a consecutive pattern.

DISCUSSION

Most protists possess a unique membrane skeleton called the epiplasm. Articulins are the major protein constituents of the epiplasm of ciliates and euglenoids. We have characterised cDNAs encoding the complete coding regions of the two major articulins of the ciliate *Pseudomicrothorax dubius*. Comparison with previously determined articulin sequences allows us to describe some general sequence features of this novel class of cytoskeletal proteins.

As observed earlier, the presence of repeated amino acid sequence motifs led to some puzzling cross hybridisation between cDNAs of individual articulins in the course of molecular cloning although the overall similarity at the nucleotide level is lower than in many other multigene families (Huttenlauch et al., 1995). The unambiguous assignment of the cloned cDNAs coding for articulin 1 and 4, respectively, was possible by extensive peptide sequence results.

Articulin 1 contains a single tryptophan, located in the head domain, while articulin 4 lacks tryptophan. This observation explains and confirms earlier peptide map analysis in which treatment of articulins with N-chlorosuccinimide, which specifically cleaves at tryptophanyl residues, results in a fragment of articulin reduced in mass by ~15 kDa, while articulin 4 is not cleaved (Peck et al., 1991).

Both articulins show an aberrant migration on SDS-PAGE compared to the molecular masses predicted from the open reading frames. This is particularly striking for articulin 4 where the predicted molecular mass is ~25% less than expected from SDS-PAGE. A similar phenomenon has been observed for the two articulins of the euglenoid *E. gracilis*. Also these two polypeptides show a much slower migration in SDS-PAGE than expected from their calculated molecular masses (Marrs and Bouck, 1992). Direct comparison of the migration of articulins from isolated epiplasm and the in vitro translated polypeptides demonstrate that the complete coding regions of articulins 1 and 4 have been obtained and that the aberrant migration in SDS-PAGE is an intrinsic property of these polypeptides. However, this property cannot be assigned to any particular sequence feature of articulins since articulin p60, which shows all the characteristic features of the protein family, has a perfect match between the calculated molecular mass and the molecular weight determined by SDS-PAGE (Huttenlauch et al., 1995).

General sequence features of articulins

With the characterisation of *P. dubius* articulins 1 and 4 reported here, all major articulins of ciliates and euglenoids are known. In total, the sequences of five members of the articulin protein family are known (Marrs and Bouck, 1992; Huttenlauch et al., 1995; this report). All five show the same tripartite domain organisation (Fig. 7). The hallmark of these proteins is the central core domain of repetitive motifs of alternating valine and proline residues, the VPV-motif. The following rules can be deduced from the available sequence data. The majority of V-positions shows an occupancy of >50% valine, while proline in P-positions shows >30% occupancy but rarely reaches 50% occupancy. There is a strict exclusion of proline residues from V-positions. This exclusion is so invariant that it can be taken to define the borders of the central core domain of VPV-motifs. Within the VPV-motif there is a clear segregation of charged residues with respect to V- and P-positions. Although not as strict as the segregation of valine and proline residues, negatively charged residues are preferentially found in P-positions, while the majority of positive charges is found in V-positions. There is some variation of the actual VPV-motif within individual polypeptides and often closely related motifs are clustered.

Subdivision of articulin core domains into 12-residue long VPV-motifs separated by linkers of variable length, as earlier proposed (Marrs and Bouck, 1993; Huttenlauch et al., 1995) is not as conclusive as it might have previously appeared. The core domains of both articulins 1 and 4 can be arranged as consecutive 12-residue motifs without any linker sequences, and the same holds for articulin p60 and the two articulins of *E. gracilis*. For *P. dubius* p60 a consecutive arrangement of the VPV-motifs gives rise to 29 instead of 24 VPV-motifs in the previous alignment containing linker sequences (Huttenlauch et al., 1995). A quantitation of valine and proline residues as

well as of charged residues for the new alignment of p60 is given in Fig. 5. Similarly, a consecutive alignment for *E. gracilis* p80 and p86 gives rise to 35 instead of 31 12-residue VPV-motifs in both proteins. The regularity of the distribution of valine and proline residues as well as the charge segregation is not significantly poorer in such a continuous arrangement compared with a pattern with occasional insertions. Moreover, in some regions of the core domain the repeat length is six residues rather than 12 residues. Therefore, the general sequence principle of the articulin core domain might be the simple alternation of valine and proline residues together with an in register segregation of charge.

Secondary structure prediction roles of the sequence in the central domain of articulins indicate a high content of β-sheet. In a β-sheet structure, the charge distribution would result in the separation of negatively and positively charged residues into two layers above and below the polypeptide backbone (Fig. 7). Previous FTIR spectroscopical measurements of bacterially expressed articulin p60 of *P. dubius* showed about 30% β-sheet (Huttenlauch et al., 1995) and, as discussed by Marrs and Bouck (1992) for *Euglena* articulins, an extended conformation would be consistent with the filaments that reconstituted from fractions highly enriched in articulins in the presence of plasma membranes (Dubreuil and Bouck, 1988). Alternatively, based on the analysis of hydrophobic cluster distribution within articulin sequences, Coffe et al. (1996) proposed a coiled coil configuration for articulins.

Electron microscopic analysis of purified recombinant articulin p60 revealed rod-shaped, short filamentous structures with diameters of 15-20 nm as well as 25-30 nm thick filaments (Huttenlauch et al., 1995). Articulins 1 and 4 that were expressed in mammalian cells by transient transfection show a strong tendency to aggregate to form larger assemblies but electron microscopic analysis does not reveal any particular substructure of these assemblies (our unpublished results). More experimentation will, however, be necessary before detailed models of how articulins assemble can be proposed.

Peptide repeats in the terminal domains of articulins

The central core domain of articulins is flanked by hydrophobic head and tail domains. The abundance of most amino acids shows a characteristic segregation between the central core domain and the two terminal domains. Even more striking, in four of the five known articulins, repeated sequence motifs are found in either the head or the tail domains (Fig. 7). These repeats, either hexapeptides or heptapeptides, are generally hydrophobic and often rich in glycine residues. The actual sequences, although very similar in articulins of one species, are not conserved between euglenoids and ciliates. In *Pseudomicrothorax* articulins 1 and 4, regions with consecutive hexapeptide repeats can be distinguished from regions where hexapeptide repeats are separated by insertions of variable length. Another difference between the two *Euglena* articulins and the two major articulins of *Pseudomicrothorax* is the location of the repeats with respect to head and tail domains (Fig. 7). While in *Euglena* 86 kDa and 80 kDa articulins these repeats are present in the head and the tail domain, respectively, they are located within the tail domain of both the *Pseudomicrothorax* polypeptides.

It has been speculated that in *Euglena* the repeats in the terminal domains might be involved in head to tail interaction between the two articulins (Marrs and Bouck, 1992). The presence of similar repeats in the tail domains of both major *Pseudomicrothorax* articulins suggests common function of these domains in *Euglena* and *Pseudomicrothorax*.

The epiplasm tightly interacts with membranes. In *Euglena* a major integral membrane protein, IP39, has been characterised that binds and anchors one (80 kDa) of the two articulins to the membrane (Rosiere et al., 1990). While a peptide fragment of IP39 has been identified as the 80 kDa articulin binding domain, it is not known to which domain of the articulin IP39 binds. The isolation of cDNAs for the major articulins will allow an experimental analysis using protein engineering to address the functional significance of the three domain structure deduced from the sequence analysis. In addition, we have successfully expressed articulins 1 and 4 in mammalian cells at high levels. These type of experiments together with in vitro studies on the interaction of individual articulin protein domains should allow further insight into the assembly properties of this novel class of cytoskeletal proteins.

The head domain of articulin 1 contains the motif TPGR at position 7-10 (Fig. 2). This sequence conforms to the proposed consensus site for phosphorylation by $p34^{cdc2}$ protein kinase (Moreno and Nurse, 1990). An identical $p34^{cdc2}$ consensus site was previously found in the tail domain of *P. dubius* articulin p60 (Huttenlauch et al., 1995). It is tempting to speculate, that phosphorylation at specific sites of the terminal domains might influence the assembly properties of articulins. *P. dubius* articulins are not labelled when cells are grown in [^{32}P]orthophosphate (Peck et al., 1991). Phosphorylation at these sites might, however, be restricted to those regions of mitotic cells, that are in the process of reproducing their membrane skeleton. Consequently, phosphorylated articulins might comprise only a minor fraction and might be difficult to detect. That epiplasmic proteins might indeed be phosphorylated during mitosis has been demonstrated with a phosphoprotein specific monoclonal antibody in *Paramecium* (Keryer et al., 1987).

For epiplasmins, a major group of epiplasmic proteins of *Paramecium*, a glutamine rich coil structure has been discussed on the basis of short partial sequences (Coffe et al., 1996). Although these short peptide sequences are rich in valine and show alternating proline and valine residues they do not display the characteristics of typical VPV-motifs. One has to await the completion of epiplasmin sequences before firm conclusions on the relatedness of these two types of proteins can be drawn.

A global phylogenetic tree of eukaryotes based on rRNA sequence data can be subdivided into an ‘early’, ‘middle’ and ‘terminal crown’ zone. According to such a phylogenetic tree euglenoids as a subgroup of the Euglenozoa are members of the ‘middle’ zone while ciliates belong to the ‘terminal crown’ zone indicating a quite distant relationship between euglenoids and ciliates (Sogin et al., 1986). Data from protein coding genes, on the contrary, suggest a much closer relationship between euglenoids and ciliates (Keeling and Doolittle, 1996; for a critical discussion see Hasegawa and Hashimoto, 1993; Adoutte et al., 1996, and references therein). How useful the data on articulins that are now available will be for evolutionary considerations will critically depend on whether articulins are restricted to these two groups of protists or are more widespread. An answer to the latter question might be facilitated using the new molecular tools. In addition to cDNA

library screening and PCR based approaches articulin specific antibodies might be generated either by affinity selection against recombinant VPV-motif domain polypeptides or by epitope mapping to select those monoclonal antibodies that recognize articulin specific sequence motifs. These antibodies might then be used in protein blotting and immunoscreening.

A data bank search reveals that articulins do not show significant sequence similarity to known proteins of bacterial, archebacterial or metazoan origin. The 12-residue long VPV-motif is unique to articulins. It will be interesting to see how the sequence principles that have been uncovered by comparison of the primary sequences of articulins will guide the three-dimensional structure and the supramolecular assembly of these novel cytoskeletal proteins.

We thank Klaus Schwamborn, Institut für Biochemie und molekulare Zellbiologie, University of Göttingen, for help with the transfection experiments, Heinz-Jürgen Dehne, this institute, for help with the electron microscopic experiments and Mary Osborn, this institute, for critically reading the manuscript. I.H. acknowledges the support of the Deutsche Forschungsgemeinschaft (Hu 549/1-3 and 549/1-4).

REFERENCES

Adoutte, A., Germont, A., Le Guyader, H. and Philippe, H. (1996). Que savons-nous de l'histoire évolutive des eukaryotes? 2. De la diversification des protistes à la radiation des multicellulaires. *Médicine/sciences:* **12**, I-XVII.

Aufderheide, K., Frankel, J. and Williams, N. E. (1980). Formation and positioning of surface-related structures in Protozoa. *Microbiol. Rev.* **44**, 252-302.

Bouck, G. B. and Ngô, H. (1996). Cortical structure and function in euglenoids with reference to trypanosomes, ciliates, and dinoflagellates. *Int. Rev. Cytol.* **169**, 267-318.

Bricheux, G. and Brugerolle, G. (1986). The membrane cytoskeleton complex of euglenoids. I. Biochemical and immunological characterization of the epiplasmic proteins of *Euglena acus. Eur. J. Cell Biol.* **40**, 150-159.

Bricheux, G. and Brugerolle, G. (1987). The pellicular complex of euglenoids. II. A biochemical and immunological comparative study of major epiplasmic proteins. *Protoplasma* **140**, 43-54.

Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.-T. and Dobberstein, B. (1987). A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. *Meth. Enzymol.* **155**, 416-433.

Coffe, G., Le Caer, J.-P., Lima, O., Adoutte, A. (1996). Purification, in vitro reassembly, and preliminary sequence analysis of epiplasmins, the major constituent of the membrane skeleton of *Paramecium*. *Cell Motil. Cytoskel.* **34**, 137-151.

Curtenaz, S. and Peck, R. K. (1992). A monoclonal antibody study of protein distribution in the membrane skeleton of the ciliate *Pseudomicrothorax*. *J. Cell Sci.* **103**, 1117-1125.

Curtenaz, S., Nahon, P., Iftode, F. and Fleury, A. (1994). Interspecific immunological cross-reactions among cortical proteins of four ciliates. *Eur. J. Protist.* **30**, 440-450.

Dubreuil, R. R. and Bouck, G. B. (1985). The membrane skeleton of a unicellular organism consists of bridged articulating stripes. *J. Cell Biol.* **101**, 1884-1896.

Dubreuil, R. R. and Bouck, G. B. (1988). Interrelationships among the plasma membrane, the membrane skeleton and surface form in a unicellular flagellate. *Protoplasma* **143**, 50-164.

Frankel, J. (1989). *Pattern Formation. Ciliate Studies and Models*. New York, Oxford. Oxford University Press.

Grain, J. (1986). The cytoskeleton in protists: nature, structure and functions. *Int. Rev. Cytol.* **104**, 153-249.

Hasegawa, M. and Hashimoto, T. (1993). Ribosomal RNA trees misleading? *Nature* **361**, 23.

Huttenlauch, I. and Peck, R. K. (1991). The membrane skeleton of *Pseudomicrothorax.* II. Biochemical and immunological characterization of the glycosylated cytoskeletal proteins. *J. Cell Sci.* **100**, 707-715.

Huttenlauch, I., Geisler, N., Plessmann, U., Peck, R. K., Weber, K. and Stick, R. (1995). Major epiplasmic proteins of ciliates are articulins: cloning, recombinant expression, and structural characterization. *J. Cell Biol.* **130**, 1401-1412.

Keeling, P. J. and Doolittle, W. F. (1996). Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. *Mol. Biol. Evol.* **13**, 1297-1305.

Keryer, G., Davis, F. M., Rao, P. N. and Beisson, J. (1987). Protein phosphorylation and dynamics of cytoskeletal structures associated with basal bodies in *Paramecium*. *Cell Motil. Cytoskel.* **8**, 44-54.

Marrs, J. A. and Bouck, G. B. (1992). The two major membrane skeletal proteins (articulins) of *Euglena gracilis* define a novel class of cytoskeletal proteins. *J. Cell Biol.* **118**, 1465-1475.

Moreno, S. and Nurse, P. (1990). Substrates for $p34^{cdc2}$: in vivo veritas? *Cell* **61**, 549-551.

Peck, R. K. (1977). The cortical ultrastructure of the somatic cortex of *Pseudomicrothorax dubius*: structure and function of the epiplasm in ciliated protozoa. *J. Cell Sci.* **25**, 367-385.

Peck, R. K. (1986). The trace element selenium induces trichocyst formation in the ciliated protozoan *Pseudomicrothorax dubius*. *Eur. J. Cell Biol.* **41**, 174-181.

Peck, R. K. Duborgel, F., Huttenlauch, I. and de Haller, G. (1991). The membrane skeleton of *Pseudomicrothorax,* I. Isolation, structure, and composition. *J. Cell Sci.* **100**, 693-706.

Rosiere, T. K., Marrs, J. A. and Bouck, G. B. (1990). A 39-kD plasma membrane protein (IP39) is an anchor for the unusual membrane skeleton of *Euglena gracilis*. *J. Cell Biol.* **110**, 1077-1088.

Rupp, R. A. W., Snider, L. and Weintraub, H. (1994). *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev.* **8**, 1311-1323.

Schleicher, M. and Noegel, A. A. (1992). Dynamics of the *Dictyostelium* cytoskeleton during chemotaxis. *New Biol.* **4**, 461-472.

Sogin, M. T., Elwood, H. J. and Gunderson, J. H. (1986). Evolutionary diversity of eukaryotic small-subunit rRNA genes. *Proc. Nat. Acad. Sci. USA* **83**, 1383-1387.

Sperling, L., Keryer, G., Ruiz, F. and Beisson, J. (1991). Cortical morphogenesis in *Paramecium*: a transcellular wave of protein phosphorylation involved in ciliary rootlet disassembly. *Dev. Biol.* **148**, 205-218.

Stelly, N., Mauger, J.-P., Claret, M. and Adoutte, A. (1991). Cortical alveoli of *Paramecium:* a vast submembranous calcium storage compartment. *J. Cell Biol.* **113**, 103-112.

Stick, R. (1988). cDNA cloning of the developmentally regulated lamin L_{III} of *Xenopus laevis*. *EMBO J.* **7**, 3189-3197.

Stockem, W. and Brix, K. (1994). Analysis of microfilament organization and contractile activities in *Physarum*. *Int. Rev. Cytol.* **149**, 145-215.