Differential targeting of closely related ECM glycoproteins: the pherophorin family from *Volvox*

Klaus Godl, Armin Hallmann¹, Stephan Wenzl and Manfred Sumper

Lehrstuhl Biochemie I, Universität Regensburg, D-93053 Regensburg, Germany

¹Corresponding author

K.Godl and A.Hallmann made equal contributions to this work

The alga Volvox carteri represents one of the simplest multicellular organisms. Its extracellular matrix (ECM) is modified under developmental control, e.g. under the influence of the sex-inducing pheromone that triggers development of males and females at a concentration below 10⁻¹⁶ M. A novel ECM glycoprotein (pherophorin-S) synthesized in response to this pheromone was identified and characterized. Although being a typical member of the pherophorins, which are identified by a C-terminal domain with sequence homology to the sex-inducing pheromone, pherophorin-S exhibits a completely novel set of properties. In contrast to the other members of the family, which are found as part of the insoluble ECM structures of the cellular zone, pherophorin-S is targeted to the cellfree interior of the spherical organism and remains in a soluble state. A main structural difference is the presence of a polyhydroxyproline spacer in pherophorin-S that is linked to a saccharide containing a phosphodiester bridge between two arabinose residues. Sequence comparisons indicate that the self-assembling proteins that create the main parts of the complex Volvox ECM have evolved from a common ancestral gene.

Keywords: ECM glycoproteins/green algae/ hydroxyproline-rich glycoproteins/pherophorins/*Volvox*

Introduction

Some of the simplest multicellular organisms are found among the green algae of the genus *Volvox. Volvox carteri* is composed of only two cell types: 2000–4000 biflagellate *Chlamydomonas*-like somatic cells are arranged in a monolayer at the surface of a hollow sphere and 16 much larger reproductive cells ('gonidia') lie just below the somatic cell sheet (Starr, 1969). Due to its simplicity, *Volvox* is an ideal model for a biochemical analysis of developmental processes.

Many developmental responses of cells are mediated by the extracellular matrix (ECM) with which those cells are in contact. In *Volvox*, this is very clearly demonstrated in the fascinating process of sexual differentiation triggered by the sex-inducing pheromone. *Volvox* cells are surrounded and held together by a glycoprotein-rich ECM (for a review see Kirk *et al.*, 1986). The *Volvox* ECM shows a distinct structural architecture. The outermost area, called the boundary zone (BZ; for nomenclature see Figure 1A) (Kirk *et al.*, 1986), contains those components of the ECM that appear to be continuous over the surface of the organism. The area lying internal to the boundary zone, called the cellular zone (CZ), exhibits specializations around individual cells. The deep zone (DZ) of the ECM consists of components that fill the cell-free interior of the spherical organism. This zone contains highly viscous polysaccharide-rich amorphous components (Kirk *et al.*, 1986).

The chemical composition of the Volvox ECM is strongly modified under the influence of the sex-inducing pheromone. The pheromone, a glycoprotein (Starr and Jaenicke, 1974; Tschochner et al., 1987; Mages et al., 1988), converts asexually growing males and females to the sexual pathway. This pheromone is among the most potent biological effector molecules known: it exhibits full effectiveness at 6×10^{-17} M (Starr, 1970; Gilles *et al.*, 1984). Many lines of evidence indicate that the ECM plays a key role in this sexual induction process. The earliest biochemical responses to the pheromone detected so far are structural modifications within the ECM (Wenzl and Sumper, 1982, 1986b; Gilles et al., 1983). The CZ of the Volvox ECM contains members of the newly described pherophorin family. Pherophorins are glycoproteins that contain a C-terminal domain with homology to the sexual pheromone. Pherophorin I is constitutively expressed and represents a main component of the cellular zone of the ECM. Under the influence of the pheromone, synthesis of pherophorin II is initiated and its C-terminal domain becomes proteolytically liberated from the parent glycoprotein (Sumper et al., 1993). It has been proposed that this modification of the ECM is part of a signal amplification process required to obtain the exquisite sensitivity of this sexual induction system. So far, the DZ has not been investigated in detail for any biochemical changes in response to the sexual pheromone, although some early observations indicate a modification of this ECM region (Gilles et al., 1983).

In this paper, we describe a novel pherophorin that is synthesized in response to the sex-inducing pheromone. Although highly homologous to other members of the family that are located exclusively within the somatic cell sheets, this novel pherophorin (pherophorin-S) is specifically targeted to the DZ of the ECM. Pherophorin-S exhibits a unique glycosylation pattern among the pherophorins: it contains a phosphodiester bridge between two arabinose residues.

Results

Identification of pherophorin-S

Pheromone-induced changes in the composition of the ECM have been characterized in detail only within the



Fig. 1. (A) Highly stylized drawing emphasizing the main compartments of the *Volvox* ECM. BZ, boundary zone; CZ, cellular zone; DZ, deep zone, according to the nomenclature of Kirk *et al.* (1986); S, somatic cells; G, gonidium (reproductive cell). (**B**) Identification of the soluble pherophorin-S (Phero-S) and comparison with insoluble pherophorin II (Phero II). Fluorogram of an SDS–polyacrylamide gel loaded with different extracts from asexual (veg) or sexually induced *Volvox* spheroids pulse and pulse-chase labelled with [35 S]sulfate. Sexual induction was initiated by the application of pheromone (~0.1 pM) for 40 min. After a pulse labelling period of 30 min, aliquots of *Volvox* spheroids were removed and analysed (ind). After a chase of 18 h, aliquots of the populations were again analysed (chase). First, a DZ extract of the ECM was prepared by carefully disrupting the spheroids. After centrifugation, the somatic cell sheets were further treated with detergent followed by EDTA/detergent (Sumper *et al.*, 1993) to extract pherophorin II from the CZ of the ECM. Extracts from 50 *Volvox* spheroids were applied to a 4–15% SDS–polyacrylamide gel. In contrast to pherophorin II, which is proteolytically processed (42 and 30 kDa; Sumper *et al.*, 1993; Godl *et al.*, 1995), pherophorin-S remains stable during the chase period.

CZ (Figure 1A) of the ECM (Wenzl and Sumper, 1982, 1986b; Ertl et al., 1989; Sumper et al., 1993; Godl et al., 1995). In order to extend this analysis to the DZ (Figure 1A) of the ECM, which may constitute >90% of the total volume of the organism, the composition of this ECM compartment from asexual and sexually induced organisms was compared after pulse labelling with radioactive sulfate. Mild mechanical stress as may be exerted by forcing Volvox spheroids through a hypodermic needle fragments the spheroids, producing hemispheres or smaller fragments of cellular sheets. The material of the DZ is thereby selectively released. This mild fragmentation of the organism does not affect viability of the cells. After low speed centrifugation, the cell-free supernatant, containing the material from the DZ was subjected to SDS-PAGE. About 30 min after application of the sexual pheromone, synthesis of a previously unobserved ³⁵S-labelled component, with an apparent molecular mass of 110 kDa, becomes detectable in the fluorogram of the SDS-polyacrylamide gel (Figure 1B, DZ, ind). This component is synthesized only transiently. Maximum expression is found 120 min after application of the pheromone. For reasons explained below, this component of the DZ was named pherophorin-S. Pherophorin-S is only detectable in the DZ of the ECM and is not detectable in asexually growing organisms. As pherophorin-S is quantitively extracted without any additives (detergents, salt or EDTA), it is a soluble component of the ECM.

In contrast to other members of the pherophorin family which are proteolytically processed (Sumper *et al.*, 1993; Godl *et al.*, 1995), pherophorin-S is a stable protein, as demonstrated by a pulse–chase labelling experiment (Figure 1B).

Pulse labelling experiments using radioactive phosphate revealed that, in contrast to all other known members of the pherophorin family that have been studied previously, pherophorin-S also incorporates phosphate. The chemical nature of the incorporated phosphate will be described below.

Purification of pherophorin-S

Mild mechanical disruption of *Volvox* spheroids liberates the material of the DZ, including pherophorin-S. This property was used to purify pherophorin-S; the corresponding extract will be denoted in the following as 'deep zone extract'. After filtration and centrifugation, the colourless extract containing pherophorin-S was fractionated by anion exchange chromatography (Q-Sepharose followed by Mono Q). Final purification was achieved by preparative SDS–PAGE.

Purified pherophorin-S exhibits different apparent molecular masses on SDS–PAGE, depending on the percentage of the acrylamide used. The observed values range from 90 to 130 kDa (~100 kDa on an 8% gel). This is a property of some glycosylated proteins. Treatment of pherophorin-S with anhydrous hydrogen fluoride at 0°C, a procedure that selectively deglycosylates glycoproteins but does not cleave polypeptides (Mort and Lamport, 1977), reduces the apparent molecular mass by ~10 kDa (8% SDS–PAGE). Therefore, pherophorin-S is a glycoprotein.

To obtain amino acid sequence data, purified pherophorin-S glycoprotein was digested with trypsin and the resulting peptide mixture was separated by reversed phase C_2/C_{18} HPLC. The material of well-separated peaks was directly subjected to amino acid sequence analysis on an



40



Fig. 2. (A) Strategy applied to collect the complete nucleotide sequence of pherophorin-S cDNA. Amino acid sequence information of tryptic peptide IYPSVGSSSIVTPSWTAIGG was used to design sense and antisense primers for amplification of a cDNA fragment of 56 bp in length (probe). Completion of the cDNA was achieved by 5'- and 3'-RACE-PCR and by sequence analysis of a genomic clone. The positions of known introns are indicated by arrowheads. The EMBL data library accession No. is Y07752. (B) Deduced amino acid sequence of pherophorin-S. The characteristic polyproline stretch is shown in white letters on a black background. Amino acid sequences confirmed from isolated peptides are underlined. An arrowhead marks the potential signal peptidase cleavage site. Potential N-glycosylation sites are boxed. The broken line marks the C-terminal part of the polypeptide found to be homologous to the sex-inducing pheromone.

automated gas phase sequencer. The amino acid sequence data obtained are underlined in Figure 2B.

Cloning of the pherophorin-S gene

The amino acid sequence of the tryptic peptide IYPSVGSSSIVTPSWTAIGG was used to synthesize an antisense oligonucleotide primer to reverse transcribe mRNA isolated from sexually induced Volvox algae. A sense primer derived from the same peptide allowed amplification by PCR of a cDNA probe of 56 bp in length (Figure 2A, probe), which was cloned into the SmaI site of pUC18 by blunt ligation. Sequencing of this insert revealed a nucleotide sequence coding for the amino acid sequence of the peptide mentioned above. The RACE-PCR technique (Frohman et al., 1988) was used to obtain additional sequence information. A 3'-RACE-PCR produced the 3'-end of the mRNA, whereas successive 5'-RACE-PCRs did not arrive at the 5'-end of the mRNA (Figure 2A). Similar problems observed with hydroxyproline-rich glycoproteins from Volvox (Ertl et al., 1989, 1992) suggested that the pherophorin-S mRNA might also contain a C-rich stretch that causes premature termination of reverse transcription. Thus a genomic library of V.carteri constructed in the replacement vector λ EMBL3 was screened to obtain the missing sequence data from a genomic clone. Nine positive clones were identified out of 60 000 phages screened. The ~17 kb insert of one of these clones was subcloned and sequenced. A GC-rich section coding for a proline-rich domain was indeed identified. The sequence upstream of this GC-rich section was then established by two additional 5'-RACE-PCRs (Figure 2A), each resulting in the same 5'-end. Comparison of genomic and cDNA sequences revealed the presence of two introns. The strategy applied to collect the complete cDNA sequence is schematically summarized in Figure 2A. The deduced amino acid sequence for pherophorin-S is given in Figure 2B. A molecular mass of 63.4 kDa was calculated for the polypeptide chain of pherophorin-S.

This is much less than the apparent molecular mass seen on SDS-polyacrylamide gels (Figure 1B, DZ, ind). The amino acid sequence of pherophorin-S exhibits five N-glycosylation sites, but glycosylation only accounts for an increase in molecular mass of ~10 kDa (8% SDS-PAGE). Most likely, the presence of a domain with a very high proline/hydroxyproline content explains the difference between the observed and calculated molecular masses of deglycosylated pherophorin-S, because such a domain has a reduced ability to bind SDS (Andres et al., 1993).

Pherophorin-S is a member of the pherophorin family

A striking feature of the deduced amino acid sequence is a central domain, 88 amino acid residues in length, that is composed almost exclusively (89%) of proline residues. Most probably, the secondary structure of this domain is a polyproline II helix that separates the N- and C-terminal domains. A BLASTP search (Altschul et al., 1990) of the SwissProt Protein Sequence Database revealed significant identities of the deduced amino acid sequence to the pherophorin family from Volvox (Figure 3). The region of identity covers nearly the total length of the polypeptide chain. For instance, the N-terminal part of the polypeptide (amino acids 42-207) shows 55.3% identity to the N-terminal part of pherophorin II (amino acids 1-168) (Sumper et al., 1993), the C-terminal half (amino acids 308-596) exhibits 40.4% identity to the C-terminal half of pherophorin II (amino acids 186-484). The name pherophorin-S was chosen to indicate the fact that this polypeptide is the first member of the pherophorin family remaining highly soluble within the ECM. Pherophorin-S contains 29 cysteines, most of them conserved among the different pherophorins, indicating conservation of the three-dimensional structure.

Although pherophorin-S is without any doubt a member of the pherophorin family, its chemical properties are



Fig. 3. Alignment of the amino acid sequences surrounding the central polyproline spacers of pherophorin-S and pherophorins I–III. The polyproline stretches are given in bold type. Gaps (–) are introduced to allow maximal alignment.

strikingly different from all the other members known so far. First, pherophorin-S is targeted to the DZ compartment of the ECM, whereas pherophorins I–III are located within the CZ of the ECM. Second, all other pherophorins are highly insoluble components of the ECM and, third, only pherophorin-S incorporates phosphate.

It is the proline-rich domain that is completely different in pherophorin-S: this domain is ~88 amino acids in length, whereas corresponding elements in pherophorins I–III are only ~10 amino acids in length (Figure 3). As demonstrated below, it is this polyproline-domain of pherophorin-S that incorporates phosphate.

The promoter of the pherophorin-S gene mediates pheromone-dependent transcription

To examine regulation of the pherophorin-S gene, the pherophorin-S 5'-untranslated region (~1 kb) was placed in front of a reporter gene, the arylsulfatase gene from *Volvox* (Hallmann and Sumper, 1994b) (Figure 4A). In wild-type *Volvox*, arylsulfatase is only expressed under sulfur starvation; no activity is detectable in organisms grown in sulfate-containing medium (Hallmann and Sumper, 1994a).

After transformation of *Volvox* with the chimeric pherophorin-S–arylsulfatase gene, the reverse transcription-PCR technique was used to verify the existence of hybrid mRNA in transformants (see Materials and methods). *Volvox* transformants containing the arylsulphatase gene under the control of the pherophorin-S promoter were incubated with or without the pheromone in the presence of the chromogenic enzyme substrate 4-nitrocatechol sulfate. Arylsulfatase activity was determined photometrically by measuring the absorbance of the liberated 4-nitrocatechol. Only transformants treated with the sex-inducing pheromone exhibited enzyme activity (Figure 4B). Since arylsulfatase is extremly stable, activity could even be assayed in SDS–polyacrylamide gels using the chromogenic substrate 5-bromo-4-chloro-3-indolyl sulfate (Figure 4C).

Thus, the ~ 1 kb DNA fragment isolated from the upstream region of the phorophorin-S gene mediates transcription of the arylsulphatase reporter gene in response to the sex-inducing pheromone.

Transgenic Volvox expressing pherophorin-S

To obtain sufficient amounts of pherophorin-S for structural studies, transgenic *Volvox* were generated that express the pherophorin-S gene under the control of the strong *Volvox* β -tubulin promoter (Figure 5A). Stable transformants were produced as previously described (Hallmann and Sumper, 1994b, 1996; Schiedlmeier *et al.*, 1994). The expression rate was 20- to 30-fold higher in vegetatively grown *Volvox* transformants than in sexually induced wild-type algae (Figure 5B). The transgenic *Volvox* strain did not have any visible phenotype.

The overexpression of pherophorin-S in clone PheroS-T1 allowed a much simpler purification protocol. Asexually grown PheroS-T1 algae were disrupted. After centrifugation the supernatant was brought to 10% acetonitrile and passed over a C_{18} cartridge. The flow-through was concentrated and applied to preparative SDS–PAGE. Pure pherophorin-S could be eluted from the gel. Recombinant pherophorin-S was used for all further chemical characterizations.

The characteristics of pherophorin-S from asexually growing transformant PheroS-T1 were compared with that obtained from sexually induced wild-type algae. Neither targeting into the DZ of the ECM nor post-translational modification causing incorporation of phosphate is affected in the transformant, indicating that these properties are not under the control of the sex-inducing pheromone. The apparent molecular masses of pherophorin-S from wildtype algae and from transformant PheroS-T1 are identical (Figure 5C).

Identification of a phosphodiester between arabinose residues in pherophorin-S

The carbohydrate composition of pherophorin-S was determinded by radio gas chromatography. Pherophorin-S purified from the transgenic *Volvox* strain grown in the presence of $[^{14}C]$ bicarbonate was hydrolysed and the resulting monosaccharides were converted to the corresponding alditol acetates. Pherophorin-S contains the neutral sugars arabinose and galactose in a 1:1 ratio (Figure 6A).

Pherophorin-S incorporates [³³P]phosphate in pulselabelling experiments. Incorporated radioactivity is quantitatively removed from the polypeptide chain upon treatment with anhydrous HF, indicating that phosphate is not linked to a hydroxyamino acid. In addition, hydrolysis of pherophorin-S in 0.5 M trifluoroacetic acid at 100°C for 2 h quantitatively liberates bound phosphate as a low molecular mass derivative. Analysis of this hydrolysate on polyethyleneimine thin-layer plates resulted in the detection of two radioactive spots. One radioactive product migrated like a phosphomonoester, the main degradation product migrated like a phosphodiester (Figure 7A). The latter substance stained with orcinol reagent and was completely hydrolysed after 2 h in 6 M HCl at 100°C. Sugar analysis by gas chromatography identified arabinose as the only sugar present in this derivative (Figure 6C). In earlier studies, the phosphodiester arabinose-5-phospho-5'-arabinose was identified as a structural element in the ECM glycoprotein SSG 185 from Volvox (Holst et al., 1989). Proof for the existence of the same structural component in pherophorin-S was obtained by mass spectrometry as follows. Pherophorin-S (supplemented



Fig. 4. A 1 kb promoter region of the pherophorin-S gene mediates transcription control by the sex-inducing pheromone. (**A**) Structure of the chimeric gene containing the pherophorin-S promoter region and the arylsulfatase reporter gene (genomic clone). (**B**) Photometric arylsulfatase activity assay (Hallmann and Sumper, 1994a) from transformants. Algae were disrupted by ultrasonic treatment and the lysates assayed using 4-nitrocatechol sulfate as a chromogenic substrate (OD produced in 30 min by 50 algae/ml). (**C**) Activity stain of arylsulfatase (ARS) from transformants in an SDS–polyacrylamide gel (8%) using 5-bromo-4-chloro-3-indolyl sulfate as the chromogenic substrate. DZ extracts from asexually grown (–) or sexually induced (+) *Volvox* algae are shown. Transformants (T) are compared to the *nitA*⁻ strain (N⁻) used as DNA recipient and to wild-type HK10 algae (WT).

with trace amounts of labelled material) was hydrolysed and reduced. The resulting products were separated by high performance anion exchange chromatography (HPAEC) (Figure 7B). Radioactive fractions were subjected to thinlayer chromatography (data not shown) and to mass spectrometry. After reduction of the anomeric carbon atoms with NaBH₄, the phosphodiester produced a mass signal at 365.0 (Figure 7C). This exactly corresponds to the calculated mass for the reduced phosphodiester.

The polyproline domain carries the phosphodiester

Pherophorin-S is not completely digested if treated with proteases like pronase, proteinase K or subtilisin. Rather, a resistant core with an apparent molecular mass of ~50 kDa (8% SDS–PAGE) remains (Figure 8A). Proteolytic degradation of ³³P-labelled pherophorin-S results in a core material that still contains all of the originally incorporated radioactivity. Consequently, the phosphodiester is located within this protease-resistant core material. In order to define this core material, purified pherophorin-S was digested with subtilisin. The 50 kDa core material was eluted from a preparative SDS–PAGE, deglycosylated with anhydrous HF (because the glycosylated material

could not be analysed by Edman degradation) and purified by reversed phase HPLC (C_2/C_{18}). A single peptide eluted at 20% acetonitrile. Edman degradation of this material resulted in the sequence shown in Figure 8B. As expected, the resistant core material represents the proline-rich domain of pherophorin-S. The amino acid sequence analysis also confirms that the prolines at the very beginning of this domain (residues 211 and 213) remain unmodified, whereas prolines 215, 218, 219, 221 and 223 (and probably all the following) become hydroxylated and can serve as saccharide attachment sites.

The carbohydrate composition of the protease-resistant 50 kDa core material was determined by radio gas chromatography. The core material contains the neutral sugars arabinose and galactose in a 1:1 ratio (Figure 6B), exactly as found for intact pherophorin-S.

Discussion

Sequence homology proves that pherophorin-S is a member of the pherophorin family of *Volvox* ECM proteins. Like pherophorin II, it is synthesized in response to the sex-inducing pheromone. However, pherophorin-S



Fig. 5. Transgenic *Volvox* constitutively expressing pherophorin-S. (A) Structure of the chimeric gene containing the *Volvox* β -tubulin promoter region and the pherophorin-S coding region (genomic clone). (B) Constitutive expression of pherophorin-S (Phero-S) in transgenic *Volvox*. Silver stain of a SDS–PAGE gel (8%) loaded with DZ extracts from vegetatively grown wild-type *Volvox* (lane 1) and transgenic *Volvox* (clone PheroS-T1) (lane 2). (C) [³³P]Phosphate incorporation into pherophorin-S (Phero-S). Purified ³³P-labelled pherophorin-S from sexually induced wild-type *Volvox* (lane 1) and DZ extracts from asexual wild-type *Volvox* (lane 2) and the transformant (lane 3), prepared after pulse labelling (30 min) with [³³P]phosphate. Autoradiogram of a SDS–PAGE gel (8%). X-Ray film after 0.5 days' exposure.

exhibits unique properties: it is accumulated in the DZ of the Volvox ECM in a completely soluble state, in contrast to all the other members known so far, which are insoluble and restricted to the CZ. Biogenesis of ECMs occurs by self-assembly, which means that each component contains within its structure the information necessary for this fascinating process. In pherophorins I-III the N- and C-terminal domains are separated by a short polyhydroxyproline spacer. It is this spacer element that is strikingly different in pherophorin-S: a stretch of ~90 amino acid residues that are almost exclusively hydroxyproline residues separates the terminal domains. This spacer is glycosylated and, of particular interest, contains a phosphodiester bridge between two arabinose residues. This type of modification was originally discovered in another ECM glycoprotein, namely SSG 185 from Volvox (Ertl et al., 1989). SSG 185 is the monomeric precursor of a polymeric substructure within the CZ of the ECM that surrounds individual cells, creating honeycomb-like chambers. Remarkably, SSG 185 contains exactly the same type of polyhydroxyproline spacer. Since SSG 185 and pherophorin-S are found in completely different regions of the ECM, it is unlikely that this particular spacer provides the signal for targeting pherophorin-S to the DZ. Using the newly established system of Volvox transformation (Schiedlmeier et al., 1994), chimeras of pherophorin domains should allow identification of structural elements that are responsible for specific targeting within the ECM.

SSG 185 and pherophorin I represent the main components of the cellular zone of the ECM in asexually growing *Volvox* (Godl *et al.*, 1995). Under the influence of the sex-inducing pheromone, pherophorin II is deposited within the CZ of the ECM and newly synthesized pherophorin-S modifies the composition of the DZ. As the primary structures of all these ECM glycoproteins are known, it is possible to compare their modular composition



Fig. 6. Sugar analysis by radio gas chromatography. The scan at the top of each chromatogram records the mass signal of internal standards. (**A**) ¹⁴C-labelled alditol acetates derived from pherophorin-S after acid hydrolysis. (**B**) ¹⁴C-labelled alditol acetates derived from the subtilisin-resistant core peptide of pherophorin-S. (**C**) ¹⁴C-labelled alditol acetates derived from the subtilisin derived from the purified phosphodiester that is a structural element of pherophorin-S.

and to search for structural homologies among the modules. Figure 9A presents the domain structure of these glycoproteins in diagrammatic fashion. The C-terminal domains of all pherophorins (B-type domain) are characterized by sequence homology with the sex-inducing pheromone. The N-terminal domains of all pherophorins (A-type domain) are related to both the N- and C-terminal domains of SSG 185 (Figure 9B) (Godl et al., 1995). Moreover, a sequence comparison of the A- and B-type domains of pherophorin II reveals that even these two regions exhibit 24% sequence identity over a stretch of 179 amino acid residues. Thus, the main parts of the complex ECM of *Volvox*, and even the species-specific signalling molecule (sex-inducing pheromone), appear to have been derived from the same ancestral gene (Figure 9B). Differently modulated domains have been linked together via spacer elements that are composed of polyhydroxyproline

D



Fig. 7. Identification of a ³³P-labelled phosphomono- and phosphodiester of arabinose in acid hydrolysates of pherophorin-S. (A) Autoradiogram of a polyethyleneimine thin-layer chromatogram. Lane 1, phosphodiester and phosphomonoester obtained from the ECM-protein SSG 185 (Holst et al., 1989) plus inorganic phosphate; lane 2, hydrolysed pherophorin-S; lane 3, inorganic phosphate. (B) The compounds of pherophorin-S released by acid hydrolysis were purified by HPAEC under strong alkaline conditions. The eluent was monitored by pulsed amperometric detection (PAD). For details see Materials and methods. D, phosphodiester; M, phosphomonoester; P, inorganic phosphate. (C) Characterization of the phosphodiester derived from pherophorin-S by electrospray mass spectrometry.

М

8 10 12 14 16 18

Fig. 8. The polyproline domain of pherophorin-S carries the phosphodiester of arabinose. (A) Purified ³³P-labelled pherophorin-S (Phero-S) subjected to 8% SDS-PAGE. Silver stain (lane 1) and fluorogram of the same sample before (lane 2) and after digestion with subtilisin (lane 3). (B) Sequence analysis of the subtilisin-resistant fragment (50 kDa) from pherophorin-S. Non-hydroxylated prolines are indicated by arrowheads. The section containing no prolines, but only hydroxyprolines, is marked by a horizontal arrow.

sequences and it is only the introduction of these spacers that qualifies these Volvox ECM proteins for membership in the class of hydroxyproline-rich glycoproteins (HRPGs) typical of plant cell walls. These molecular data offer

strong support for the idea of a gene superfamily of hydroxyproline-rich glycoproteins (for reviews see Kieliszewski and Lamport, 1994; Woessner and Goodenough, 1994) from which new ECM proteins could

220

A-type domains: identities / similarities (% / %)





B-type domains: identities / similarities (% / %)

Fig. 9. Domain conservation in *Volvox* ECM proteins and the species-specific sex-inducing pheromone. (A) Each protein is represented in diagrammatic fashion with the N- and C-terminal domains separated by rod-like spacers. Dark boxes indicate the polyhydroxyproline spacers. (B) The degree of sequence identity/similarity observed between each A-type domain of pherophorins and SSG 185 and between each B-type domain of pherophorins and the sexual pheromone. For SSG 185 only the N-terminal A-type domain is given.

constantly evolve. Volvocine radiation, i.e. the transition from unicellular *Chlamydomonas* to multicellular *Volvox*, is a recent event, probably having occurred within the past 50–75 million years (Rausch *et al.*, 1989). This recent transition to multicellularity converted a simple cell wall into a complex ECM. Thus, this independent development of an ECM should offer an attractive model for studying the mechanisms operating during this process. The development of the *Volvox* ECM also appears to provide another example of the combinatorial advantage of shuffling modules, as is so evident in the evolution of the metazoan ECM (for a review see Doolittle, 1995).

Materials and methods

Volvox strains

Wild-type *V.carteri* strain HK10 (female) was obtained from the Culture Collection of Algae at the University of Texas (R.C.Starr). Mutant strain 153-48 of *V.carteri* (Adams *et al.*, 1990), obtained from D.L.Kirk (Washington University, St Louis, MO) was used as the DNA recipient in transformation experiments. This strain with wild-type morphology carries a stable loss-of-function mutation in *ntA*, the structural gene encoding nitrate reductase (Gruber *et al.*, 1996).

Culture conditions

Synchronous cultures were grown in *Volvox* medium (Provasoli and Pintner, 1959) at 28°C in a 8 h dark/16 h light (10 000 lux) cycle (Starr and Jaenicke, 1974). The non-selective medium used in transformation experiments was *Volvox* medium supplemented with 1 mM NH₄Cl; selective medium was *Volvox* medium lacking NH₄Cl and containing only nitrate as a nitrogen source.

Radioactive labelling of pherophorin-S with [³⁵S]sulfate or [³³P]phosphate

Pulse labelling with [35 S]sulfate was performed as described by Wenzl and Sumper (1981). In pulse and pulse-chase experiments with [33 P]phosphate, *Volvox* spheroids were washed thoroughly with and then suspended in 1 ml glycerophosphate-free *Volvox* medium. After the addition of 50 µCi [33 P]phosphate, incubation under standard conditions was continued for 0.5 or 1 h.

Purification of pherophorin-S

Sexually induced wild-type Volvox spheroids from six 201 cultures were harvested by filtration on a 100 µm mesh nylon screen. The spheroids

The disrupted spheroids were centrifuged at 20 000 g for 30 min. In order to remove any remaining insoluble components, the supernatant was brought to 25 mM Tris-HCl, pH 7.5, 0.9 M NaCl, 1 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and applied to a QAE-Sephadex A-25 column (Pharmacia) equilibrated with the same buffer. Pherophorin-S does not bind and was therefore detected in the flowthrough. After this filtration step, the material containing pherophorin-S was diluted to 0.3 M NaCl and applied to a Q-Sepharose FF anion exchange column (Pharmacia). Elution was performed with a linear gradient of 0.3-0.9 M NaCl in 25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM PMSF. Pherophorin-S elutes at ~0.7 M NaCl. The fraction containing pherophorin-S was diluted to 0.35 M NaCl and applied to a MonoQ HR 5/5 FPLC anion exchange column (Pharmacia). Again, elution was performed with a linear gradient of 0.35-1.5 M NaCl in 25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM PMSF. Pherophorin-S was recovered at ~0.80-0.85 M NaCl. Fractions containing pherophorin-S were concentrated by precipitation with deoxycholate and trichloroacetic acid (Mahuran et al., 1983). Final purification was achieved by preparative SDS-PAGE (7%). Proteolytic digestion and separation of peptides

were broken up by forcing them through a 0.5 mm hypodermic needle.

Aliquots of 20–30 µg pherophorin-S were applied to a 7% SDS–PAGE gel and stained with Coomassie brilliant blue. The gel slice containing pherophorin-S was cut into small pieces. Further treatment and digestion with trypsin was performed as decribed by Selmer *et al.* (1996). The resulting peptides were eluted from the gel by diffusion in 0.2 M (NH₄)HCO₃/50% acetonitrile. The eluate was passed through a 0.22 µm filter (Millipore), brought to 0.1% trifluoroacetic acid and dried by lyophilization. The peptides were dissolved in 6 M guanidine–HCl/ 0.1% trifluoroacetic acid and fractionated by reversed phase HPLC (SMART system; Pharmacia) on a 3 µm µRPC C₂/C₁₈ column (Pharmacia). Peptides were eluted with a 30 min linear gradient of 5–40% acetonitrile in 0.1% trifluoroacetic acid with a flow rate of 200 µl/min. Peptides were sequenced by Edman degradation using an automated gas phase peptide sequencer (Applied Biosystems, Foster City, CA).

Generation of a cDNA probe by PCR

Generation of a cDNA probe for pherophorin-S was performed using the degenerate antisense oligonucleotide primer CCKATNGCNGTCCA and the degenerate sense oligonucleotide primer ATHTAYCCNAGYGT. The resulting 56 bp cDNA fragment was ligated into the *Sma*I site of vector pUC18 by blunt ligation and sequenced.

Cloning of the pherophorin-S gene

The RACE-PCRs were performed as described by Frohman *et al.* (1988). The *V.carteri* genomic library in λ EMBL3 (Frischauf *et al.*, 1983)

described by Ertl *et al.* (1989) was used to clone the pherophorin-S gene. The screening procedure followed standard techniques (Sambrook *et al.*, 1989). DNA sequencing was performed by the chain termination method (Sanger *et al.*, 1977) using T7 DNA polymerase (Pharmacia).

Construction of the chimeric pherophorin-S-arylsulfatase gene

The *Volvox* arylsulfatase reporter gene (Hallmann and Sumper, 1994b) was placed under the control of the pherophorin-S gene 5'-region. Additional restriction sites were introduced by PCR to facilitate ligation of the parent DNAs. An *Eco*RV site was generated directly in front of the start codons of both the pherophorin-S gene and the *Volvox* arylsulfatase gene. A *Kpn*I site was introduced into the pherophorin-S 5'-region ~1 kb upstream of the start codon. Then, a *Kpn*I–*Eco*RV fragment covering ~1 kb upstream sequence of the pherophorin-S gene was ligated to an ~10 kb *Eco*RV–*Sal*I fragment containing the *Volvox* arylsulfatase gene (Hallmann and Sumper, 1994a) with its 15 introns. The complete construct was confirmed by sequencing.

Construction of the chimeric β -tubulin-pherophorin-S gene

To achieve high pherophorin-S production, the pherophorin-S gene was placed under the control of the *Volvox* β -tubulin promoter. For construction of the chimeric gene, genomic clones of *Volvox* β -tubulin (Harper and Mages, 1988) and pherophorin-S were used. An additional *Eco*RV site was generated by PCR directly in front of the start codons of both the pherophorin-S gene and the *Volvox* β -tubulin gene to facilitate ligation of the parental DNAs. A *KpnI* site was introduced into the β -tubulin promoter region ~0.5 kb upstream of the start codon. Then, a *KpnI–Eco*RV fragment bearing the ~0.5 kb *Volvox* β -tubulin promoter region was ligated to an ~9 kb *Eco*RV–*SmaI* fragment containing the pherophorin-S gene. During cloning a T nucleotide from the *Eco*RV site. The complete construct was confirmed by sequencing.

Stable nuclear transformation of Volvox carteri

Volvox carteri strain 153-48 was transformed by using a particle gun to bombard cells with DNA-coated gold particles as described previously (Schiedlmeier *et al.*, 1994). Plasmids carrying the artificial gene constructs were introduced into *V.carteri nitA*⁻ strain 153-48 by co-transformation with plasmid pVcNR1 (Gruber *et al.*, 1992; Schiedlmeier *et al.*, 1994), containing the coding region of the *V.carteri nitA* gene plus downstream and upstream DNA. Bombarded cultures were cultivated in selective *Volvox* medium containing only nitrate as a nitrogen source.

Reverse transcription–PCR amplification and sequencing of chimeric transcripts

For reverse transcription–PCR the antisense primer 5'-TTTGAGGCGC-AATTCCG (pherophorin-S) was used for transformants containing the β -tubulin promoter–pherophorin-S chimeric gene. The sense primer was 5'-ATAACAAGCGACCACTAC (β -tubulin). Products of PCR amplification were ligated into the *Sma*I site of pUC18 and sequenced.

Preparation of recombinant pherophorin-S

Volvox transformants (clone PheroS-T1), constitutively expressing the pherophorin-S gene under the control of the β -tubulin promoter, were grown in 20 l glass flasks under standard conditions. Spheroids were disrupted and centrifuged as described before. After centrifugation ³³P-labelled pherophorin-S was added to facilitate identification. The supernatant was brought to 10% acetonitrile and passed over a C₁₈ (octadecylsilane) cartridge (Millipore). The flow-through was concentrated by lyophilization and applied to a preparative 8% SDS–PAGE gel. After autoradiography pherophorin-S was eluted with water by diffusion, dialysed and lyophilized.

Preparation of the 50 kDa subtilisin fragment

Purified and ³³P-labelled pherophorin-S was digested with 0.6 μ g/µl subtilisin (Carlsberg, type VIII; Sigma) in 50 mM Tris–HCl, pH 8.0, 0.5% SDS for 1 h at 30°C and applied to 10% SDS–PAGE. After autoradiography the 50 kDa subtilisin fragment was eluted with water, dialysed and lyophilized. Deglycosylation of the 50 kDa subtilisin fragment was performed with anhydrous hydrogen fluoride at 0°C as described by Mort and Lamport (1977). The deglycosylated peptide was purified by reversed phase HPLC (SMART system; Pharmacia) on a 3 μ m µRPC C₂/C₁₈ column (Pharmacia) by applying a 30 min linear gradient of 5–40% acetonitrile in 0.1% CF₃CO₂H at a flow rate of 200 µl/min.

Preparation of the phosphodiester of arabinose

Aliquots of 100 µg purified pherophorin-S was mixed with ³³P-labelled pherophorin-S and hydrolysed in 0.5 M trifluoracetic acid at 100°C for 2 h. The hydrolysate was dried *in vacuo*, redissolved in H₂O and extracted twice with 1 vol. *n*-butanol. The products were reduced with 0.5 M NaBH₄ at 30°C for 45 min. After removal of H₃BO₃ (Laine *et al.*, 1972) the sample was subjected to HPAEC using a CarboPac PA1 column (Dionex). A gradient of 0–0.5 M sodium acetate in 0.15 M NaOH was applied over 30 min at a flow rate of 1 ml/min. Monitoring of the eluate was performed by a pulsed amperometric detector (ED40; Dionex). Fractions of interest were passed through a column of AG 50W-X8 H⁺ ion exchange resin, lyophilized and redissolved in 0.1 mM ammonium acetate containing 50% acetonitrile and subjected to electrospray mass spectrometry.

Mass spectrometry

Molecular masses of fractions of interest were determined by electrospray mass spectrometry using the negative mode of a SSQ 7000 mass spectrometer (Finnigan).

Carbohydrate analysis

Volvox spheroids suspended in 1 ml *Volvox* medium lacking glycerophosphate were pulse-labelled with 0.4 mCi [¹⁴C]bicarbonate for 90 min. Pherophorin-S was isolated as described above. The neutral sugar composition of pherophorin-S and of the 50 kDa subtilisin fragment of pherophorin-S was determined by radio gas chromatography of the alditol acetates as described by Wenzl and Sumper (1986a). For carbohydrate analysis of the phosphodiester, ¹⁴C-labelled phosphodiester was isolated in the same way as described above. Complete hydrolysis of the phosphodiester was achieved in 6 M HCl at 100°C for 2 h. The liberated monosaccharides were identified by radio gas chromatography of alditol acetates. Unlabelled arabinose, xylose, mannose, galactose and glucose were added as internal standards.

Thin-layer chromatography

Thin-layer chromatography was performed on polyethyleneimine–cellulose plates (Schleicher & Schüll) according to Holst *et al.* (1989).

Acknowledgements

We wish to thank Dr R.Deutzmann and E.Hochmuth for mass spectrometry and for sequencing peptides. We also thank J.Nink for help with phosphodiester preparation and C.Friederich for expert technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 521).

References

- Adams, C.R., Stamer, K.A., Miller, J.K., McNally, J.G., Kirk, M.M. and Kirk, D.L. (1990) Patterns of organellar and nuclear inheritance among progeny of two geographically isolated strains of *Volvox carteri*. *Curr. Genet.*, 18, 141–153.
- Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. J. Mol. Biol., 215, 403–410.
- Andres, D.A., Goldstein, J.L., Ho, Y.K. and Brown, M.S. (1993) Mutational analysis of alpha-subunit of protein farnesyltransferase. Evidence for a catalytic role. J. Biol. Chem., 268, 1383–1390.
- Doolittle, R.F. (1995) The multiplicity of domains in proteins. *Annu. Rev. Biochem.*, **64**, 287–314.
- Ertl,H., Mengele,R., Wenzl,S., Engel,J. and Sumper,M. (1989) The extracellular matrix of *Volvox carteri*: molecular structure of the cellular compartment. *J. Cell Biol.*, **109**, 3493–3501.
- Ertl,H., Hallmann,A., Wenzl,S. and Sumper,M. (1992) A novel extensin that may organize extracellular matrix biogenesis in *Volvox carteri*. *EMBO J.*, **11**, 2055–2062.
- Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N. (1983) Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol., 170, 827–842.
- Frohman,M.A., Dush,M.K. and Martin,G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl Acad. Sci. USA*, 85, 8998–9002.

- Gilles, R., Gilles, C. and Jaenicke, L. (1983) Sexual differentiation of the green alga *Volvox carteri*: involvement of extracellular phosphorylated proteins. *Naturwissenschaften*, **70**, 571–572.
- Gilles, R., Gilles, C. and Jaenicke, L. (1984) Pheromone-binding and matrix-mediated events in sexual induction of Volvox carteri. Z. Naturf., 39c, 584–592.
- Godl,K., Hallmann,A., Rappel,A. and Sumper,M. (1995) Pherophorins: a family of extracellular matrix glycoproteins from *Volvox* structurally related to the sex-inducing pheromone. *Planta*, **196**, 781–787.
- Gruber,H., Goetinck,S., Kirk,D.L. and Schmitt,R. (1992) The nitrate reductase-encoding gene of *Volvox carteri*: map location, sequence and induction kinetics. *Gene*, **120**, 75–83.
- Gruber,H., Kirzinger,S.H. and Schmitt,R. (1996) Expression of the Volvox gene encoding nitrate reductase: mutation-dependent activation of cryptic splice sites and intron-enhanced gene expression from a cDNA. *Plant Mol. Biol.*, **31**, 1–12.
- Hallmann,A. and Sumper,M. (1994a) An inducible arylsulfatase of Volvox carteri with properties suitable for a reporter-gene system. Purification, characterization and molecular cloning. *Eur. J. Biochem.*, 221, 143–150.
- Hallmann, A. and Sumper, M. (1994b) Reporter genes and highly regulated promoters as tools for transformation experiments in *Volvox carteri*. *Proc. Natl Acad. Sci. USA*, **91**, 11562–11566.
- Hallmann,A. and Sumper,M. (1996) The *Chlorella* hexose/H⁺ symporter is a useful selectable marker and biochemical reagent when expressed in *Volvox. Proc. Natl Acad. Sci. USA*, **93**, 669–673.
- Harper, J.F. and Mages, W. (1988) Organization and structure of *Volvox* β-tubulin genes. *Mol. Gen. Genet.*, **213**, 315–324.
- Holst,O., Christoffel,V., Fründ,R., Moll,H. and Sumper,M. (1989) A phosphodiester bridge between two arabinose residues as a structural element of an extracellular glycoprotein of *Volvox carteri. Eur. J. Biochem.*, 181, 345–350.
- Kieliszewski, M.J. and Lamport, D.T.A. (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J.*, 5, 157–172.
- Kirk,D.L., Birchem,R. and King,N. (1986) The extracellular matrix of Volvox: a comparative study and proposed system of nomenclature. J. Cell Sci. 80, 207–231.
- Laine, R.A., Esselman, W.J. and Sweeley, C.C. (1972) Gas-liquid chromatography of carbohydrates. *Methods Enzymol.*, 28, 159–167.
- Mages,H.-W., Tschochner,H. and Sumper,M. (1988) The sexual inducer of Volvox carteri. Primary structure deduced from cDNA sequence. *FEBS Lett.*, 234, 407–410.
- Mahuran, P., Clements, P., Carella, M. and Strasberg, P.M. (1983) A high recovery method for concentrating microgram quantities of protein from large volumes of solution. *Anal. Biochem.*, **129**, 513–516.
- Mort,A.J. and Lamport,D.T.A. (1977) Anhydrous hydrogen fluoride deglycosylates glycoproteins. Anal. Biochem., 82, 289–309.
- Provasoli, L. and Pintner, I.J. (1959) Artificial media for freshwater algae: problems and suggestions. In Tyron, C.A. and Hartman, R.T. (eds), *The Ecology of Algae*. Pymatuning Laboratory of Field Biology Special Publication no. 2, University of Pittsburgh, Pittsburgh, PA, pp. 84–96.
- Rausch,H., Larsen,N. and Schmitt,R. (1989) Phylogenetic relationships of the green alga *Volvox carteri* deduced from small-subunit ribosomal RNA comparisons. J. Mol. Evol., 29, 255–265.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
- Schiedlmeier, B., Schmitt, R., Müller, W., Kirk, M.M., Gruber, H., Mages, W. and Kirk, D.L. (1994) Nuclear transformation of Volvox carteri. Proc. Natl Acad. Sci. USA, 91, 5080–5084.
- Selmer,T., Hallmann,A., Schmidt,B., Sumper,M. and von Figura,K. (1996) The evolutionary conservation of a novel protein modification, the conversion of cysteine to serinesemialdehyde in arylsulfatase from *Volvox carteri. Eur. J. Biochem.*, 238, 341–345.
- Starr,R.C. (1969) Structure, reproduction, and differentiation in Volvox carteri f. nagariensis Iyengar, strains HK9 & 10. Arch. Protistenk., 111, 204–222.
- Starr,R.C. (1970) Control of differentiation in Volvox. Dev. Biol., suppl. 4, 59–100.
- Starr,R.C. and Jaenicke,L. (1974) Purification and characterization of the hormone initiating sexual morphogenesis in *Volvox carteri* f. *nagariensis* Iyengar. *Proc. Natl Acad. Sci. USA*, **71**, 1050–1054.

Sumper, M., Berg, E., Wenzl, S. and Godl, K. (1993) How a sex pheromone might act at a concentration below 10⁻¹⁶ M. *EMBO J.*, **12**, 831–836.

- Tschochner,H., Lottspeich,F. and Sumper,M. (1987) The sexual inducer of Volvox carteri: purification, chemical characterization and identification of its gene. EMBO J., 6, 2203–2207.
- Wenzl,S. and Sumper,M. (1981) Sulfation of a cell surface glycoprotein correlates with the developmental program during embryogenesis of *Volvox carteri. Proc. Natl Acad. Sci. USA*, **78**, 3716–3720.
- Wenzl,S. and Sumper,M. (1982) The occurrence of different sulphated cell surface glycoproteins correlates with defined developmental events in *Volvox. FEBS Lett.*, **143**, 311–315.
- Wenzl,S. and Sumper,M. (1986a) A novel glycosphingolipid that may participate in embryo inversion in *Volvox carreri. Cell*, **46**, 633–639.
- Wenzl,S. and Sumper,M. (1986b) Early event of sexual induction in Volvox: chemical modification of the extracellular matrix. Dev. Biol., 115, 119–128.
- Woessner, J.P. and Goodenough, U.W. (1994) Volvocine cell walls and their constituent glycoproteins: an evolutionary perspective. *Protoplasma*, **181**, 245–258.

Received on August 9, 1996; revised on September 16, 1996