

BACTERIA MAY INDUCE THE SECRETION OF MUCIN-LIKE PROTEINS BY THE DIATOM *PHAEODACTYLUM TRICORNUTUM*

Matthias T. Buhmann,² Birgit Schulze, Alexander Förderer,³ David Schleheck, and Peter G. Kroth⁴

Fachbereich Biologie, Universität Konstanz, 78457 Konstanz, Germany

Benthic diatoms live in photoautotrophic/heterotrophic biofilm communities embedded in a matrix of secreted extracellular polymeric substances. Closely associated bacteria influence their growth, aggregation, and secretion of exopolymers. We have studied a diatom/bacteria model community, in which a marine *Roseobacter* strain is able to grow with secreted diatom exopolymers as a sole source of carbon. The strain influences the aggregation of *Phaeodactylum tricornutum* by inducing a morphotypic transition from planktonic, fusiform cells to benthic, oval cells. Analysis of the extracellular soluble proteome of *P. tricornutum* in the presence and absence of bacteria revealed constitutively expressed newly identified proteins with mucin-like domains that appear to be typical for extracellular diatom proteins. In contrast to mucins, the proline-, serine-, threonine-rich (PST) domains in these proteins were also found in combination with protease-, glucosidase- and leucine-rich repeat-domains. Bioinformatic functional predictions indicate that several of these newly identified diatom-specific proteins may be involved in algal defense, intercellular signaling, and aggregation.

Key index words: biofilm; biofouling; diatom; EPS; extracellular enzymes; interactions; mucins; proteomics; signaling

Abbreviations: EPS, extracellular polymeric substances; LRR, leucine-rich repeats; PST, proline-, serine-, threonine-

Some of Earth's most diverse and productive aquatic ecosystems are found in coastal and shallow-water regions where primary production is frequently dominated by diatoms. While open water habitats are dominated by planktonic, centric radially symmetric diatoms, pennate bilaterally symmetric diatoms are mostly benthic, moving on or attaching to surfaces via secretion of extracellular polymeric substances (EPS) through a slit-like raphe in their silicified cell

walls (Edgar 1983). These EPS consist mainly of polysaccharides and of proteins, and their secretion is modulated in response to various stimuli, such as tidal fluctuation, diurnal rhythm, or light levels (Hoagland et al. 1993, Smith and Underwood 1998). The carbohydrate and amino acid composition of the EPS secreted by different diatoms has been investigated intensively, but very little is known on the amino acid sequences of the secreted protein fraction and of their corresponding genes (Abdullahi et al. 2006, Poulsen et al. 2014).

Algae live closely associated with heterotrophic bacteria that may utilize secreted algal-derived organic substances, particularly the EPS, as carbon source (Nalewajko et al. 1980, Schäfer et al. 2002). These bacteria can have influence on the EPS secretion and aggregation, as well as on the growth of diatoms (Bruckner et al. 2008, 2011, Gärdes et al. 2010). Independent from the habitat, particularly alphaproteobacteria are ubiquitously associated with algae and diatoms, and bacteria of the *Roseobacter* clade were identified in many algae/bacteria associations (Schäfer et al. 2002, Buchan et al. 2005, Wagner-Döbler and Biebl 2006). Recently, a *Roseobacter* strain has been isolated and characterized from a *P. tricornutum* culture that had been isolated in Xiamen, China (Chen et al. 2015).

To investigate direct interactions between bacteria and diatoms, we have established a distinct diatom/bacteria laboratory model community, which is constituted of a marine alphaproteobacterium of the *Roseobacter* clade, *Roseovarius* sp. strain 217 (referred to as "R217") and an axenic, marine pennate diatom, *Phaeodactylum tricornutum* (Bohlin). The typical habitats of *P. tricornutum* are brackish water environments like rock pools or estuaries, where it lives both planktonically or adherent in phototrophic biofilms (De Martino et al. 2007). Dependent on strain and growth conditions, *P. tricornutum* may change its appearance between a triradiate, fusiform, and oval cell morphology, with only oval cells being able to adhere strongly to surfaces and producing completely silicified cell walls (Borowitzka and Volcani 1978, Stanley and Callow 2007). Isolates of *P. tricornutum* from several locations worldwide show a different behavior concerning the transformation from one to another morphotype, thus the morphotypic variability may represent an adaptation to instable environmental conditions (De Martino et al. 2007, 2011).

²Present address: Empa. Materials Science and Technology, Lerchenfeldstrasse 5, St. Gallen 9014, Switzerland.

³Present address: Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, Köln 50829, Germany.

⁴Author for correspondence: e-mail Peter.Kroth@uni-konstanz.de.

We determined the effects of the presence and absence of the *Roseovarius* sp. strain 217 on biofilm formation and cellular aggregation of *P. tricornutum*. In a proteomic approach, we investigated whether extracellular diatom proteins are induced by the presence or absence of bacteria via peptide fingerprinting-mass spectrometry and differential metabolic labeling (^{15}N). We found that most of the identified proteins were constitutively expressed, and we were able to identify several novel diatom proteins with proline-, serine-, and threonine-rich domains (PST-domains) similar to those of mucins. However, these PST-domains were often combined with biochemically functional (catalytic) domains (e.g. protease or glucosidase domains), or leucine-rich repeat (LRR) domains. Detailed sequence analyses allowed a first functional classification of these newly identified extracellular diatom proteins, unveiling a novel class of secreted diatom proteins.

MATERIALS AND METHODS

Chemicals. Standard chemicals were purchased in p.a. quality from Sigma Aldrich (Munich, Germany) or Carl Roth (Karlsruhe, Germany). Artificial sea salts were purchased from Tropic Marin (Wartenberg, Germany).

Biological material. *P. tricornutum* strains Pt1 (CCAP 1055/1, CCMP2561) and PT4 (UTEX646, CCAP 1052/6) were grown in half strength seawater f/2 medium (Guillard and Ryther 1962) on a rotary shaker at 70 rpm, illuminated at 20°C, with a 12/12 h light/darkness cycle (30 W fluorescent tubes; Philips, Munich, Germany) and a light intensity of 40 $\mu\text{mol photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$. For differential metabolic labeling, either $\text{Na}^{14}\text{NO}_3$ or $\text{Na}^{15}\text{NO}_3$ (98 mol % ^{15}N ; Sigma Aldrich, Munich, Germany) was provided as the sole source of available nitrogen. Axenic cultures of *P. tricornutum* were obtained by a 3 day antibiotics treatment (1 mg \cdot mL $^{-1}$ penicillin G, 0.5 mg \cdot mL $^{-1}$ streptomycin and 0.1 mg \cdot mL $^{-1}$ chloramphenicol) and subsequent spray plating of the washed and diluted cell suspensions. For this, in a sterile air cabinet, 20 μL of serial dilutions of washed diatom cell suspensions were taken up by a 20 μL pipette, and the tip was placed into the sterile air stream produced by an air pump with a 0.2 μm sterile filter and focused by a 0.7 mm syringe (BD Biosciences, Heidelberg, Germany). Half strength f/2 medium agar dishes (1.3% agar) were placed in a distance of 30 cm, collecting droplets of the sprayed cell suspension. Single colonies were picked and cultivated in liquid medium. Axenic cultures were maintained in fluid half strength f/2 medium at 8°C and low light conditions ($<5 \mu\text{mol photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$). Absence of bacteria was confirmed by absence of bacterial colonies after plating on conventional, half and quarter strength LB medium and on Marine Broth medium (Difco, BD, Heidelberg, Germany) and by bright field and epifluorescence microscopy of SYBR Green and Dioc6 stained samples (Life Technologies GmbH, Darmstadt, Germany). Selection for oval Pt1 cells was performed by incubation in flat bottom polystyrene tissue culture flasks with vented lids (Sarstedt, Nürnberg, Germany) and removal of fusiform and nonattached cells via medium exchange every 4 d. Fusiform cells grew in stationary cultures on a rotary shaker at 70 rpm.

Roseovarius sp. strain 217 was obtained from Karin Denger, University of Konstanz. This strain originally had been isolated from seawater for its ability to utilize methyl halides

(Schäfer et al. 2005) and was cultivated in half strength f/2 culture medium supplemented with 30 mM sodium acetate, 0.5 ng \cdot L $^{-1}$ paraaminobenzoic acid and 0.2% (w/v) tryptone, as tested in growth experiments with half strength f/2 medium and individual components of the vitamin solution published by Widdel and Pfennig (1981).

Growth experiments. Biofilm growth experiments were performed in LED illuminated sterile biofilm growth chambers with continuous medium exchange and automated measurement of biofilm turbidity, as described previously (Buhmann et al. 2011). Growth of suspended *P. tricornutum* was assessed via cell counter measurements (Multisizer 3; Beckman Coulter GmbH, Krefeld, Germany), and bacterial growth was assessed via colony forming unit counts. The percentage of oval cells was determined via microscopic counts in a hemato cytometer (Thoma chamber).

Extracellular protein extraction and nanoLC MS/MS analysis. A detailed description is found in the Supplementary Methods section (see Appendix S1 in the Supporting Information). In brief, suspended stationary 500 mL cultures of axenic Pt1, of strain R217/Pt1 co cultures, and of Pt1 in medium supplemented with filter sterilized, cell free medium of strain R217 were cultivated in triplicates with either $\text{Na}^{14}\text{NO}_3$ (^{14}N) or $\text{Na}^{15}\text{NO}_3$ (^{15}N) as sole source of nitrogen. In addition, with out metabolic labeling, Pt4 was cultivated axenically in the presence and absence of cell free medium of R217. Exponentially growing ^{14}N and ^{15}N labeled cells were pooled based on chlorophyll *a* concentrations, and culture supernatants were collected by centrifugation and sterile filtration. Diatom chlorophyll *a* concentrations have been calibrated against cell counts performed by an automated cell counter (Multisizer; Beckmann Coulter, Krefeld, Germany). The contribution of bacteriochlorophyll to the overall measured chlorophyll *a* concentrations was negligible as the technical standard variation was higher than the slightly increased measured chlorophyll *a* concentrations of R217/Pt1 co cultures, compared to axenic cultures. Proteins were extracted via ammonium sulfate precipitation and ultrafiltration, followed by SDS PAGE separation, tryptic digestion and subsequent identification via reversed phase liquid chromatography nanospray tandem mass spectrometry (LC MS/MS) at the Proteomics Facility of the University of Konstanz.

Protein identification and bioinformatic analyses. Gene models were checked for completeness and redundancy. Identified proteins above a threshold of a MASCOT score of 25 were subjected to NCBI BLASTp and delta BLAST searches, and subsequent PSI BLAST iteration. Subsequently, sequences were examined for conserved domains with the SMART server (<http://smart.embl.heidelberg.de/>) and PROSITE domain prediction tools (<http://prosite.expasy.org/>). In case of existing characterized protein orthologues from other organisms, amino acid sequence alignments with the ClustalW 2.1 server (<http://www.clustal.org/clustal2/>) were performed. Furthermore, we examined the amino acid sequences for high abundance of cysteine amino acids, proline, serine, threonine or acidic amino acids. Proteins with high sequence similarity to intracellular proteins, e.g. nuclear and photosystem related proteins, were not subjected to further functional prediction analyses.

In addition, we screened for signal peptide sequences via SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>), WoLF PSORT (<http://wolffpsort.seq.cbrc.jp/>), and in some cases via TargetP 1.1 (<http://www.cbs.dtu.dk/services/TargetP/>). Further, sequences were analyzed for non classical secretion via SecretomeP 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>). Transmembrane domain predictions were performed with the TMHMM server (<http://www.cbs.dtu.dk/services/TMHMM/>). GPI (glycosylphosphatidylinositol)

membrane anchors were predicted with PredGPI (<http://gpcr.biocomp.unibo.it/predgpi/>). Putative proteases were analyzed with help of the MEROPS database (<http://merops.sanger.ac.uk>).

RESULTS AND DISCUSSION

Bacteria induced aggregation and morphotypic transition of P. tricornutum. In our study, we used the *P. tricornutum* strains Pt1 (CCAP 1055/1, originally isolated in 2003 from marine environments, Blackpool, UK) and Pt4 (CCAP 1052/6, isolated 1951 from brackish water, Baltic Sea, Finland) that both show a fusiform shape during bacteria-free, suspended incubation (De Martino et al. 2007). Growth experiments confirmed that cultivation without shaking, but with a continuous exchange of media, resulted in a transformation of the major part of the fusiform Pt1 population into the oval morphotype (up to nearly 100% oval cells), while Pt4 cells remained mostly fusiform (De Martino et al. 2007, 2011).

For a defined two-species diatom/bacteria model community, we selected the marine *Roseovarius* sp. strain 217 (R217) (Schäfer et al. 2005) that we found to grow in co-culture with *P. tricornutum* without added carbon source to the late exponential growth phase, implicating that strain R217 solely utilizes diatom-derived carbon (Fig. 1, C and D). Co-cultivation with strain R217 induced higher growth rates and growth yields of oval Pt1 biofilms in continuous-flow chambers compared to axenic cultures, based on biofilm turbidity measurements (Fig. 1A). By contrast, the growth of the fusiform Pt1 and Pt4 strains was not influenced neither under continuous flow nor under suspended stationary cultivation conditions (Fig. 1, B–D).

Besides affecting diatom growth, the presence of strain R217 also had a strong influence on the biofilm structure of Pt1, while Pt4 was not substantially affected (Fig. 2). In the continuous-flow chambers, axenic oval Pt1 formed thin biofilms with small cell aggregates (Fig. 2A), while the presence of R217 cells as well as of cell-free spent growth medium of strain R217 (2% v/v, data not shown), induced the formation of much larger cell aggregates (Fig. 2B). Furthermore, the biofilm co-cultures contained 90.0% ($\pm 1.3\%$) oval cells, compared to 66.6% ($\pm 5.8\%$) in axenic cultures. In suspended stationary cultures, axenic Pt1 grew only in fusiform shape, whereas Pt1/R217 co-cultures and those supplemented with 2% cell-free bacterial culture supernatant (data not shown) also contained oval, motile cells, as demonstrated by their swarming behavior on the surface of agar medium (Fig. 2, C and D).

Identification of extracellular diatom proteins. The extracellular space represents the primary contact zone between bacteria and diatoms. In order to correlate the bacteria-induced changes of the diatom biofilm growth, biofilm structure and of the

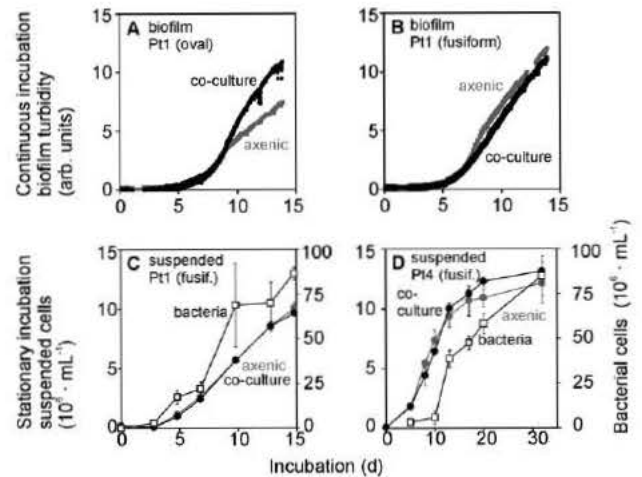


FIG. 1. *Roseovarius* sp. strain 217 in co culture with oval *Phaeodactylum tricornutum* increased biofilm growth. (A, B) The growth of axenic *P. tricornutum* biofilms (gray squares) and of biofilms co inoculated with *Roseovarius* sp. strain 217 (black circles) under continuous flow condition was followed by turbidity measurements (arbitrary units equal light absorption and, thus, biofilm thickness; see Buhmann et al. 2011). Growth curves depict data points that were acquired every 5 min. Corresponding growth curves were obtained in an independent experiment. (A) Representative biofilm growth experiment in continuous flow chambers inoculated with purely oval Pt1 cells, for which biofilms of diatom/bacteria co cultures always grew faster than axenic diatom biofilms. (B) Representative biofilm growth experiment in continuous flow chambers inoculated with purely fusiform Pt1 (or fusiform Pt4; not shown), for which the biofilm growth was not influenced by co cultivation with R217. (C, D) The growth of fusiform Pt1 and Pt4 in stationary suspended culture (shaken batch cultures), as determined with cell counts, was not influenced by a co cultivation with bacteria (black circles) in comparison to axenic diatom cultures (gray circles), as followed by determination of diatoms cell numbers; growth of bacteria was followed as colony forming units (open squares). Bars depict standard deviation, $n = 3$.

morphotypic transition of *P. tricornutum*, e.g., with previously reported changes in EPS secretion (Bruckner et al. 2011), we analyzed the extracellular proteome with a particular focus on the proteins that were secreted by the diatoms. In the first step, proteins were extracted from culture supernatant, desalted, and concentrated (see Materials and Methods), and analyzed by SDS-PAGE. This revealed different protein banding patterns between the extracellular protein extracts from diatom/bacteria co-cultures, or from axenic diatom cultures supplemented with 2% (v/v) cell-free bacterial culture supernatant, each in comparison to extracts from axenic diatom cultures (see Fig. S1 in the Supporting Information). These observations indicate changes in the expression of extracellular diatom proteins in dependence on the presence of bacteria or of bacterial factors present in spent bacterial culture supernatant. Furthermore, enzymatic deglycosylation of the protein samples strongly increased the performance of SDS-PAGE protein separation (see Fig. S1), indicating a high degree of glycosylation of the extracted extracellular proteins. For an analysis

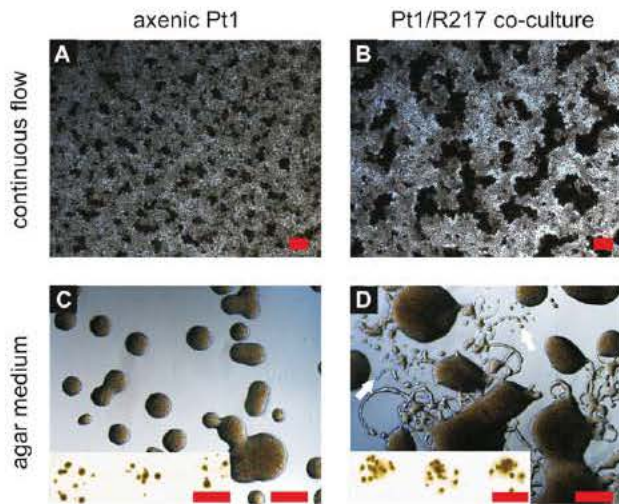


FIG. 2. Bacteria influence biofilm formation and morphotypic transition of *Phaeodactylum tricorutum* Pt1. (A) Under continuous flow conditions, axenic diatom biofilms matured into biofilms with small aggregates of oval cells that covered the glass surface completely. (B) Co cultivation of bacteria resulted in structured biofilms with large cell aggregates. (C) Axenic *P. tricorutum* formed compact colonies on agar medium, which is typical for the fusiform morphotype, after spotting of cell suspension taken from axenic shaken batch cultures. (D) *P. tricorutum* co cultured with R217 formed colonies with motile diatom cells (examples indicated by white arrows, typical for oval morphotype). Inserts (C, D): spotted cell suspensions in three replicates; scale bar 10 mm. All other scale bars (A–D), 100 μ m.

of the samples by peptide fingerprinting mass spectrometry, the extracellular proteins were extracted from supernatant of suspended (shaken) stationary diatom cultures that had been grown either axenically, or in diatom/bacteria co-culture, or axenically but supplemented with filtered bacterial culture supernatant. All cultures were inoculated with fusiform cells. In total, five shotgun proteomics experiments lead to the identification of 83 putatively secreted proteins (described further below, also see Appendix S2 in the Supporting Information). Furthermore, 52 proteins could be identified that most likely represented intracellular proteins, but these were typically identified with a much lower score and often only in single experiments, indicative of minor contaminations by intracellular proteins (see Appendix S2). Bioinformatic secretion predictions (SignalP, SecretomeP) indicated putative N-terminal signal peptides in 79% of the extracellular proteins. As a control, 33% putative intracellular proteins were predicted to have an N-terminal signal peptide.

In an attempt to quantify the abundance of extracellular diatom proteins by quantitative proteomics, Pt1 cultures that were grown either as diatom/bacteria co-cultures, or axenically but supplemented with 2% cell-free bacterial culture supernatant, were labeled with ^{15}N nitrogen isotopes; the axenic control cultures were not isotope-labeled except for an internal control ($^{15}\text{N}/^{14}\text{N}$ axenic-axenic control).

After the cultivation, and before extracting the extracellular proteins, equal amounts of ^{14}N and ^{15}N isotope-labeled cell suspensions were mixed based on equal cell numbers. The proteomic analysis ($n = 3$) revealed that only three of the 83 putative extracellular proteins detected by shotgun proteomics (see above) could be detected in amounts that we considered to be sufficient for a quantification based on the ratio of $^{14}\text{N}:^{15}\text{N}$ -labeled peptides and a comparison between the axenic/bacteria treatments (Fig. 3; i.e., per our definition, at least 10 peptide-pairs for proteins ID52157, ID45679, ID47612). The first two quantified proteins (described further below) were found in about the same ratio (i.e., similar amounts of $^{14}\text{N}:^{15}\text{N}$ -labeled peptides) in axenic cultures and those co-cultivated with bacteria or bacterial culture supernatant (Fig. 3). Only an identified GPI-anchored alkaline phosphatase (AP; ID47612; described below) appeared to be significantly lower in abundance in diatom/bacteria co-cultures (Fig. 3). In two additional proteomic analyses, extracellular proteins from axenic and a 2% spent bacterial medium-treated Pt4 were extracted, revealing a similar protein composition compared to the identified Pt1 proteins.

The many peptides identified in the five shotgun proteomics experiments allowed for an estimation of at least the overall protein abundances. However, for the interpretation of these data, it is important to keep in mind that the observed peptide abundance may also be biased by the accessibility of proteolytic cleavage sites and/or the effectivity of ion transfer and MS/MS identification.

In total, 36 of the putatively secreted proteins were consistently identified by at least eight peptides with high scores (i.e. mostly >100; Fig. 4). The most

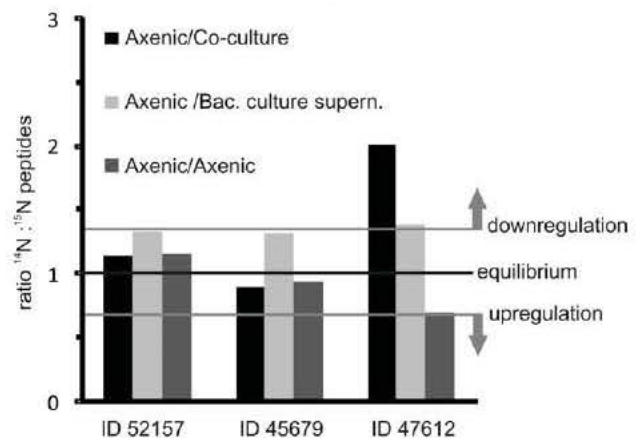


FIG. 3. Influence of co inoculated bacteria or of cell free bacterial culture supernatant on exemplary diatom protein expression as quantified by ^{15}N metabolic labeling. The two most highly abundant extracellular proteins (ID52157 and 45679) were not significantly differentially expressed in the presence/absence of bacteria. The extracellular phosphatase ID47612 appeared down regulated in the presence of bacteria or cell free bacterial medium. Bac. culture supern.: bacterial culture supernatant.

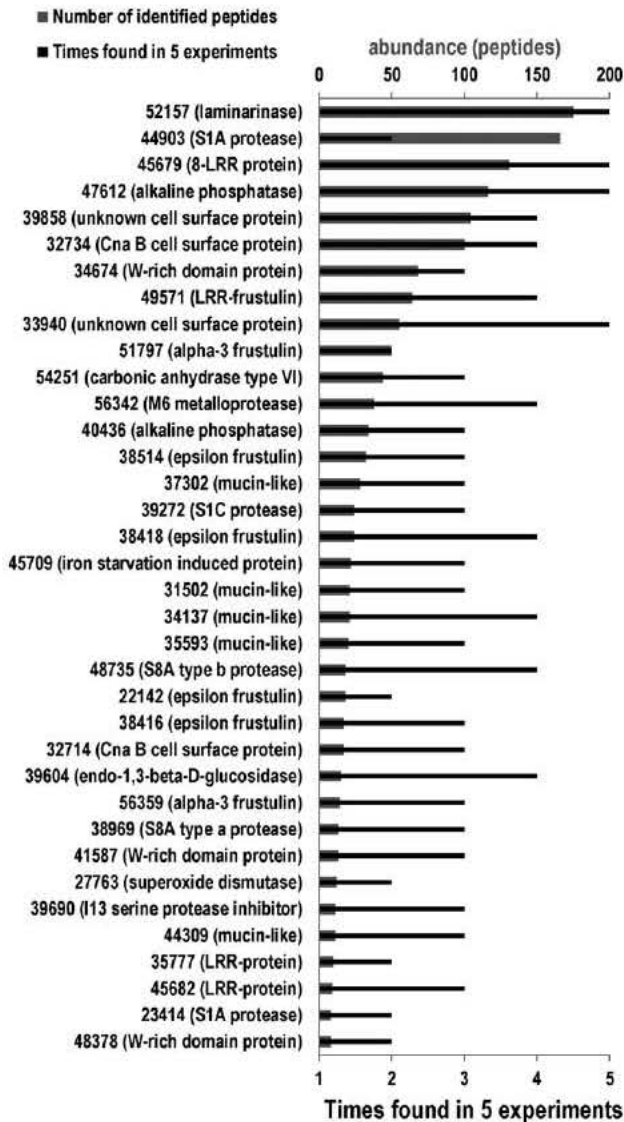


FIG. 4. The 36 most abundant extracellular *Phaeodactylum tricornutum* proteins identified. Grey, wider bars indicate the number of identified peptides found for each protein. Black bars indicate how often these proteins were identified within the five repetitive protein extractions and proteomics experiments. For each protein, the numbers represent protein identifiers in the JGI database of the *P. tricornutum* genome annotation. Their annotation based on this work is indicated in brackets.

abundant extracellular proteins were predicted glucosidases, proteases and proteins containing LRR domains, as well as predicted phosphatases and putative cell-surface proteins, whereas intracellular contaminations mostly represented putative nuclear, endoplasmic reticulum, or plastidic proteins (Fig. S2 in the Supporting Information). The sequences of identified proteins were aligned with their respective orthologues hits from BLASTp searches in order to identify functional domains (the alignments can be found in Appendix S3 in the Supporting Information). Bioinformatic sequence analyses of the amino acid composition of

the identified proteins revealed the dominance of cysteine-rich and proline-, serine- and threonine-rich (PST-) domains in several proteins (Appendix S4 in the Supporting Information). Based on domain analyses and the annotation of orthologues, the most abundantly identified extracellular diatom proteins were assigned to eight major groups, and are described and discussed in the following sections: (i) proteins involved in nutrient acquisition, (ii) PST-domain containing mucin-like proteins, (iii) proteases or protease inhibitors, (iv) sugar modifying enzymes, (v) cell-surface proteins (vi) LRR containing proteins, (vii) proteins involved in oxidative stress and defense, and (viii), signaling. Finally, 16 of the extracellular proteins identified could not be sorted into any of these groups (Fig. 5).

Proteins attributed to nutrient acquisition: Particularly during limited availability of dissolved inorganic phosphorus, aquatic microorganisms can access organically bound phosphorus through secretion of APs and 5'-nucleotidases, e.g., in order to utilize the phosphorus in extracellular phospholipids, proteins, or nucleic acids (Ammerman and Azam 1985). Several identified extracellular proteins of *P. tricornutum* are similar to phosphate-binding proteins or extracellular matrix 5'-nucleotidases (Fig. S3 in the Supporting Information; Appendix S3, Alignments 1, 2). Indeed, 5'-nucleotidase activity has previously been detected at the *P. tricornutum* plasma membrane, and AP activity in solution and at the cell walls (Flynn et al. 1986). AP activity was also found in EPS stalks of the diatom *Didymosphenia geminata* (Aboal et al. 2012). In each of our proteomic experiments, extracellular diatom APs were identified, particularly protein ID47612 (see above, Fig. 3). This protein has a putative C-terminal GPI-anchor (illustrated in Fig. 6) and a high sequence similarity to bacterial APs (see Appendix S3, Alignment 3). According to metabolic labeling, this AP was 40%–60% more abundant in axenic Pt1 cultures. This observation may reflect the previously reported growth-supporting effects of several bacteria on diatoms (Bruckner et al. 2008, 2011). Possibly, the

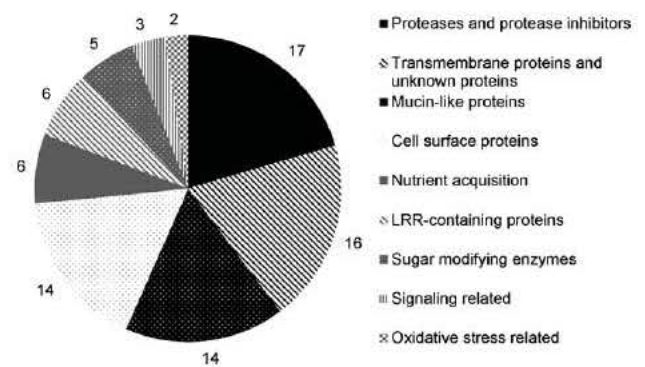


FIG. 5. Illustration of the identified extracellular *P. tricornutum* proteins assorted to functional classes.

added bacteria are contributing to the phosphorous acquisition; therefore, the alga does not need to release as much AP to achieve the same level of available phosphorous. Another phosphatase, ID40436, which is similar to the bacterial AP *phoA*, appeared to be present in higher abundance in axenic Pt1-cultures than in diatom/bacteria co-cultures (Fig. 6; Lin et al. 2013). This finding may indicate increased phosphorus availability in the presence of bacteria, possibly via recycling of nutrients by secreted bacterial phosphatases. Notably, besides the identified phosphatases, another abundant protein, ID54251, has similarities in the putative catalytic center to secreted human carbonic anhydrase VI (PDB: 3FE4; Appendix S3, Alignment 5). A secreted carbonic anhydrase could contribute to carbon dioxide acquisition of the alga and be relevant for growth.

Identification of mucin like proteins and extracellular proteins with PST rich domains: For many of the identified *P. tricornutum* proteins, we could not find known orthologues; however, some of these proteins showed similarities to extracellular mucins. Moreover, their amino acid sequences frequently contained repetitive elements and domains enriched in serine, threonine, and proline (up to 15%, 14% and 22%, respectively), often accompanied by alanine, glycine, and valine (Appendix S4). Such domains rich in the amino acids proline (P), serine (S), and threonine (T) are characteristic for eukaryotic mucins, and thus have been termed “PST-” and “mucin-domains” (Lang et al. 2007). Serine, threonine, and hydroxyproline are typical O-glycosylation sites in extracellular mucins and in structural proteins (Spiro 2002). Recently, also dihydroxyproline has been identified as a dominant amino acid in diatom adhesive trails (Poulsen et al. 2014). The

mucin-like proteins identified here frequently contain repetitive domains (Appendix S4; Fig. S4 in the Supporting Information). Most of the identified mucin-like proteins exhibited no conserved domain aside from regions with a remarkably high cysteine content (up to 21% C on a length of 207 amino acids), which is similar to mucins CysD and cysteine-knot domains. These have been proposed to function as multimerization domains, influencing the degree of cross-linking of the gelatinous extracellular matrix (Ambort et al. 2011). Therefore, the mucin-like proteins may function as structural proteins, providing a scaffold for the extracellular matrix, or function as a barrier, similar to metazoan lung mucins, which are thought to trap and remove bacteria (Knowles and Boucher 2002).

Proteases with PST domains and their inhibitors: Within the identified extracellular proteins containing the described PST- and cysteine-rich domains, these domains were frequently found adjacent to other protein domains, such as glucosidase-, LRR-domains or protease-domains, and particularly serine- and metalloprotease domains (illustrated in Fig. S5 in the Supporting Information). The S1A family serine peptidase ID44903 was among the most abundantly identified proteins (Fig. 4) in Pt1 cultures, while it could not be detected at all in Pt4 cultures. Elevated transcript numbers of this gene were previously found in cDNA libraries established from blue light-treated Pt1, urea-adapted Pt1, and oval morphotype Pt3 cultures (Maheswari et al. 2010). The ID44903 protein and another protein identified, ID23414, have similarities to trypsin-6 and comprise a conserved C-terminal Cx₁₈Cx₁₃Cx₈Cx₃Cx₂C motif, with an additional long C-terminal cysteine-rich domain in ID23414 (6.6% cysteines; illustrated in Fig. S5; Appendix S3, Alignment

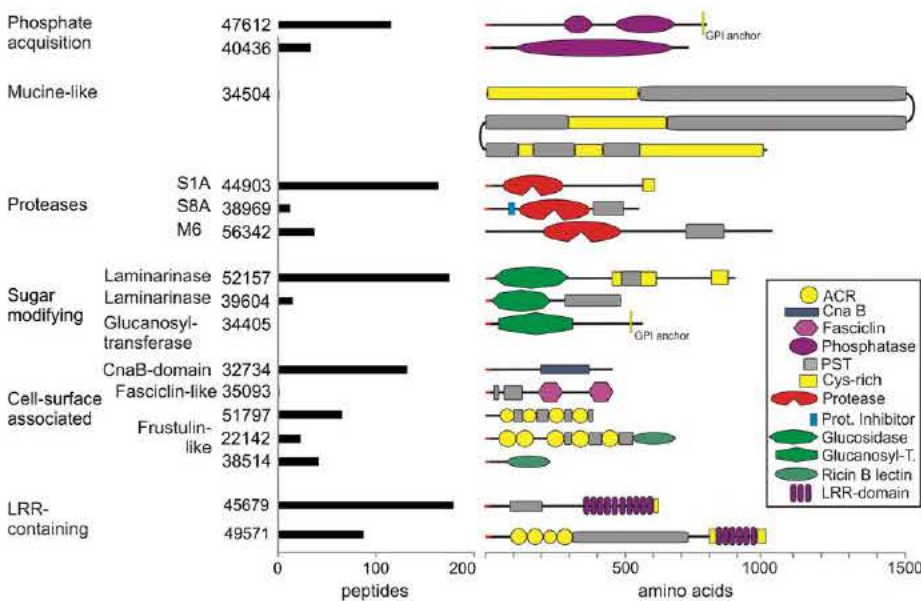


Fig. 6. Domain organization of selected identified proteins. Characteristic domains identified due to their homology to characterized protein domains were annotated. Proline, serine, and threonine rich regions and cysteine rich regions are marked as domains. Sketch: signal peptides, gray line; N terminal peptide backbone, black line; identified domains, sketches as indicated in the inset legend. Numbers represent protein identifiers in the JGI database 2.0 of the *Phaeodactylum tricornutum* genome annotation and the black bars the total number of identified peptides.

6). In metazoa, S1A proteases activate extracellular enzymes or protease-activated receptors that transduce signals for cytokine release (Chen et al. 2003, Cooper et al. 2011).

Eight of all identified proteins belong to the S8A family of unspecific subtilisin serine peptidases. When screening the *P. tricornutum* JGI database, we only found 16 further gene models for proteases that contain such an S8 domain (Pfam: PF00082). In contrast to S8 peptidases of most other organisms, the C-terminal domains of the identified diatom S8A peptidases feature PST domains or repetitive TRAP motifs (see Figs. 6 and S5). The identified protein ID38354 possesses a long C-terminal PST domain and resembles a bacterial cell-surface protease, while proteins ID38973 and ID50210 are similar to secreted thermo- and detergent stable bacterial alkaline serine proteases, like aqualysin (Fig. S5; Appendix S3, Alignment 7) and feature a C-terminal cysteine pattern. Most identified *P. tricornutum* S8 proteases feature inhibitory N-terminal I9 family predomains (Pfam: PF05922) and therefore are expressed as zymogens (see Appendix S3, Alignment 8). In other organisms, the S8 proteases can exhibit antimicrobial activity, permit proteolytic protein activation, degrade extracellular protein polymers, and alter cell adhesion (Burns et al. 2009, Suda et al. 2009).

Four metalloproteases were identified (Fig. S5), of which particularly protein ID56342 of the M6 family was highly abundant; also here, transcripts have previously been found to be abundant (Maheswari et al. 2010). Both identified M6 zinc metalloproteases, ID56342 and ID34191, feature conserved catalytic domains with a HEXXHXXGXXD motif (see Appendix S3, Alignment 9) and contain long C-terminal PST-domains, similar to *Volvox carteri* pheromone- and wounding-induced extracellular metalloproteases (Hallmann et al. 2001). Based on their similarity to toxic extracellular bacterial metalloproteases of the M4 and M6 family (see Appendix S3, Alignment 10A), the respective ortholog proteases might have a defensive function (Shinoda and Miyoshi 2011). The identified M11 peptidase ID49604 could, based on homology to matrix metalloproteases (Appendix S3, Alignment 10B), be involved in cell wall degradation during extracellular matrix remodeling (Heitzer and Hallmann 2002).

We also identified two protease inhibitors potentially involved in regulating extracellular protease activity. ID39690 is similar (Appendix S3, Alignment 11) to extracellular I13 family protease inhibitors (Pfam: PF00280) that inhibit S1 and S8 family serine peptidases and are released upon pathogen infection in plants (Plunkett et al. 1982, Heinz et al. 1991). ID35043 is similar (Appendix S3, Alignment 12) to family I25B of secreted peptidase inhibitors (e.g., antifungal plant phytochemicals) that inhibit unspecific papain-like C1 family cysteine proteases (Popovic et al. 2012).

During the interconversion from fusiform to oval morphology, the *P. tricornutum* cell wall and surface structures undergo a complex reorganization, which most likely involves also extracellular enzymes. Furthermore, pennate diatoms are capable of forming complex exopolymer structures, like tubes, stalks, fibers, and capsules that also can contain extracellular enzymes (Aboal et al. 2012). The organization of extracellular biopolymers in a specific manner may require extracellular matrix metalloproteases, similar to the extracellular matrix metalloproteases with PST domains of the multicellular green alga *V. carteri* (Heitzer and Hallmann 2002).

Sugar modifying enzymes with PST domains: Four proteins of the glucosyl hydrolases 16 (GH16) family were identified, including the two laminarinase-like endo-1,3- β -glucanases (EC 3.2.1.6) ID52157 and ID39604, and two licheninase-like endo-1,3(4)- β -glucanases (EC 3.2.1.73) ID38737 and ID46513 (Fig. S5). Particularly, the laminarinase-like protein ID52157 was found in high abundance. The catalytic domains of both laminarinase-like endo-1,3- β -glucanases are highly similar (see Appendix S3, Alignment 13) to a crystallized ortholog (PDB: 2HYK), as well as to an endo-1,3(4)- β -glucanase that binds to lichenan and rice cell walls (Asano et al. 2002, Fibriansah et al. 2007). Both proteins, ID52157 and ID39604, contain C-terminal PST-, GKS- or cysteine- rich domains, indicating their potential extracellular or cell-surface localization. The less abundant endo-1,3(4)- β -glucanases ID38737 and ID46513 have similarities (Appendix S3, Alignment 14) to bacterial extracellular licheninases, i.e., enzymes that cleave linear cell-surface β -glucanes (Varghese et al. 1994).

The identified protein ID34405 contains a GH72 domain, a predicted GPI-anchor (see Fig. 6; S526 with 99.6% specificity, according to PredGPI) and sequence similarities (Appendix S3, Alignment 15) to the yeast GPI-anchored cell-surface 1,3- β -glucanoyltransferases Gas2 and Gas4 (Pierleoni et al. 2008). Gas4 in fission yeast is involved into remodeling of the β -1,3-glucan containing cell wall, therefore the diatom ortholog may be involved into cell-surface carbohydrate modifications (De Medina-Redondo et al. 2008). In addition, the identified putative cell-surface GPI-anchored β -1,3-glycosyltransferase might contribute to defense by modifying cell-surface glucans, resulting in an enhanced antibacterial barrier. In plants, the deposition of the β -1,3-glucan callose on leaf surfaces is induced by bacterial pathogen-associated molecular patterns, such as flagellin (Luna et al. 2010).

Cell surface associated proteins: The identified putative cell surface-associated proteins are presumably only loosely attached and have been removed from the cells due to shear stress during shaking incubation. Abundant in Pt4 culture supernatant was protein ID32734. The central region of the protein has similarities to a CnaB-type domain (Pfam: PF05738)

containing bacterial protein, and high sequence similarity to two further identified proteins on the JGI database (Fig. S6 in the Supporting Information, Appendix S3, Alignment 16-18). In bacterial proteins, the CnaB-type domain functions as a stalk presenting collagen-binding domain on the cell surface.

An identified fasciclin (Pfam: PF02469) cell-adhesion domain containing protein, ID35093, is similar (11% identity) to a cell-adhesion protein CAA56621.1 of the multicellular green alga *V. carteri* (Appendix S3, Alignment 19). Monoclonal antibodies against this protein lowered the intercellular adhesion of the 4-cell *V. carteri* embryos (Huber and Sumper 1994). Interestingly, this ancient adhesion domain has recently been discussed as potential diatom cell-adhesion molecule (Willis et al. 2014).

Several identified proteins belong to the diatom-specific frustulins that cover the diatom valves and contain up to five acidic, cysteine-rich domains (ACR domains), which are interconnected by proline-rich stretches (Figs. 6 and S6; Kröger et al. 2004). Interestingly, the frustulin ACR domains from different diatom species share a conserved acidic amino acid pattern CxGxCx₅Cx₅₋₆Cx₁₀₋₁₃GC (i.e., *P. tricornutum*, *Fistulifera pelliculosa* and *Cylindrotheca fusiformis*, see Appendix S3, Alignment 20). Together with other identified proteins without ACR domain, some of the identified frustulins are similar (Fig. 6; Appendix S3, Alignment 21) to the *F. pelliculosa* epsilon frustulin, based on their C-terminal Ricin B lectin domain (PF00652, previously termed tryptophan-rich domain; Kröger and Poulsen 2008). Alignments with the Ricin B domain of the crystallized ebulin 1 protein (Appendix S3, Alignment 22) revealed the corresponding characteristic QxW amino acid pattern (Rutenber and Robertus 1991, Hazes 1996). The identified epsilon frustulin ID38422 had also been previously found to be highly transcribed in different *P. tricornutum* cDNA libraries (Maheswari et al. 2010). By contrast, the alpha-3 frustulin ID51797 lacks the Ricin B domain and was not identified in Pt1 cultures, but abundantly in Pt4 culture supernatant predominantly in the presence of bacteria (seven identified peptides in axenic and 43 in co-cultures). Correspondingly, ID51797 was found across 16 cDNA libraries to be one of the most abundantly expressed genes, and it appeared to be upregulated in oval morphotype libraries of Pt3, cold-stressed (oval) tropical accession Pt9, and dark-adapted Pt1 exposed to blue light (Maheswari et al. 2010). In *Nitzschia* species, frustulin expression was found to be induced by cadmium stress, suggesting their involvement in cell surface reinforcement or metal chelation (Santos et al. 2013). Furthermore, a novel LRR containing “LRR-frustulin” ID49571 was found in relatively low levels in Pt1 compared to Pt4 cultures (10 vs. 47 peptides; Fig. S7 in the Supporting Information).

Leucine rich repeat domain containing proteins with PST domain: Six identified proteins contain LRRs, a structural element of repeated 20–29 amino acids that are arranged in α -helices and β -turns. Arrays of such LRRs form a hydrophobic solenoid that acts as a binding or scaffolding element. LRR domains are present in diverse protein families, including extracellular receptors, binding- and inhibitory proteins (Wulff et al. 2009). The identified LRR containing proteins (“diatom-LRR”) are similar to plant-specific extracellular LRRs, featuring LxxLxxLxLxxNxLt/sGxIPxxLGx motifs, while the LRR domain is flanked by cysteine-rich regions (Fig. S7; Appendix S3, Alignment 23 and 24; Wulff et al. 2009). Typically, the identified LRR-proteins contain a C-terminal LRR domain and repetitive- or PST-domains; *P. tricornutum* genome analyses revealed several other representatives of such proteins with PST- and LRR-domains (Fig. S7B; Schulze et al. 2015). Two identified proteins, ID49571 and ID45941, contain six LRRs and a PST-domain, while the latter protein further features six long repetitive elements separated by a cysteine-rich region from the LRR domain (Appendix S3, Alignment 25). In contrast, ID49571 exhibits domain homologies to frustulins, as it contains four ACR domains that are interconnected with a long repetitive PDVST-rich spacer domain to the LRRs (Fig. 6, Appendix S5 in the Supporting Information). Its function might be similar to secreted plant inhibitory proteins such as the polygalacturonase-inhibiting LRR protein (PGIP), that are induced by cell wall degradation products (i.e., oligogalacturonides; Di Matteo et al. 2003, Davis et al. 1986). PGIP inhibits pathogen-derived cell wall lytic enzymes (i.e., polygalacturonases).

The most abundant LRR protein in Pt1 extracts was ID45679, which contains a PST domain, and in Pt4 extracts the ACR domain containing protein ID49571. This could reflect the requirement of cell surface defense for planktonic Pt4 strains in contrast to the benthic oval Pt1, which relies on the secretion of protective proteins that are retained in the extracellular matrix.

Indications for a role of reactive oxygen species in diatom defense. Aquatic organisms permanently have to cope with oxidative stress due to intracellular reactive oxygen species (ROS) as by-products of respiration and photosynthesis, and the exposure to extracellular ROS that are continuously produced in surface waters via ultraviolet irradiation (Vincent and Roy 1993). To lower the impact of ROS, superoxide dismutases (SODs) convert hydroxyl radicals to the less toxic hydrogen peroxide. The genome of *P. tricornutum* contains four SOD gene models, one of which (ID27763) has been identified in this study. This putative manganese SOD (89.01% fingerprint score, SODa prediction server) is similar to other algal SODs and to the extracellular bacterial

Mn-SOD A (Appendix S3, Alignment 27) (Fridovich 1995, Kwasigroch et al. 2008).

Several examples show that macroalgae utilize endogenously generated ROS for defense against pathogens. For instance, the red alga *Gracilaria chilensis* responds with H₂O₂ production upon cell wall-derived oligosaccharides, through a putative cell wall oligosaccharide oxidase (Weinberger et al. 2005, Potin 2008). Moreover, a red algal hexose oxidase inhibits growth of other algae via H₂O₂ production (Sullivan and Ikawa 1973). The identified *P. tricornutum* protein ID49706 has sequence homologies (Appendix S3, Alignment 28) to a bacterial galactose oxidase (GO), an enzyme that catalyzes the oxidation of galactose with dioxygen to D-galacto-hexodialdose and hydrogen peroxide. This protein further contains catalytic and substrate-binding sites similar to the crystallized fungal GO (Appendix S3, Alignment 28). In contrast to the fungal GO, the C-terminal regions of both *P. tricornutum* and bacterial GOs contain two Ricin B lectin domains that are separated by a PST-domain (Fig. S8 in the Supporting Information). Evidence for a ROS-based defense in stramenopiles is substantiated by a given example of multicellular stramenopiles that may react on bacterial flagellin in the micromolar range with H₂O₂ release and induction of antioxidant enzymes (Wang et al. 2012). To date, it remains unclear whether diatoms are able to discriminate between associated beneficial and pathogenic bacteria.

Extracellular signaling proteins: Our analyses indicated the presence of extracellular signaling proteins in diatoms (Fig. S8). For example, ID40414 is highly similar (Appendix S3, Alignment 29) to the cytokine macrophage migration inhibitory factor (MIF) that influences eukaryotic cell motility (Stark et al. 2013). BLASTp searches for a corresponding CD74-like MIF receptor in *P. tricornutum*, however, did not provide any significant hits. Similar to human MIF, the respective protein from the cyanobacterium *Prochlorococcus marinus* exhibits tautomerase activity, while lacking oxidoreductase activity, as the thioredoxin-like CxxC motif is absent. This is also the case for the *P. tricornutum* MIF (Wasiel et al. 2010).

Another identified potential cytokine, ID17701, is highly similar to human HSP70 (71% identical amino acids with P08107) that, besides functioning as chaperone, also functions as cytokine and induces intracellular calcium flux, upregulation of pro-inflammatory cytokine release and matrix metalloproteinase expression (Asea et al. 2000, Lee et al. 2006).

Furthermore, we identified a predicted single-pass transmembrane protein, ID47878, which harbors six extracellular epithelial growth factor (EGF) domains that are separated by a PST-domain from the transmembrane domain (Appendices S3 and S6 in the Supporting Information, Alignment 30 and 31). In

metazoa, the aggregation of cells is a coordinated event, involving conserved intercellular signaling proteins (Nichols et al. 2006). Therefore, the identified extracellular putative signal proteins, including proteases and their inhibitors, might transmit signals leading to cell aggregation.

Roseovarius sp. strain 217 and P. tricornutum as a diatom/bacteria model community. *P. tricornutum* cultures with natural bacterial assemblages were not available from culture collections, as they have mostly been treated with antibiotics. Therefore, we decided for a designed diatom/bacteria model community based on the ability of the bacteria to grow with secreted diatom substances as a sole source of carbon. The selected *Roseobacter* strain R217 had been isolated from seawater and not directly from a *P. tricornutum* culture, thus the observed effects may represent rather a general, non-specific response to the presence of non-associated bacteria. Similarly, previous studies have demonstrated that many different bacteria in a non-specific manner may affect the secretion of diatom exopolysaccharides (Bruckner et al. 2008, 2011). Changes in diatom aggregation pattern were also observed in diatom/bacteria co-cultures of the diatom *Planorhynchium frequentissimum* and the satellite *Pseudomonas* strain (Buhmann et al. 2011). Therefore, the observed changes in diatom aggregation behavior and the identified constitutively expressed defense-involved proteins may be related to a general antibacterial defense towards non-symbiotic bacteria.

In additional experiments (data not shown), we have investigated the response of *P. tricornutum* to the *Roseobacter* strain *Ruegeria pomeroyi* DSS-3 and found it also to increase the amounts of oval *P. tricornutum* cells. In contrast to R217, *Ruegeria pomeroyi* DSS-3 secreted a substance that apparently is algicidal against *P. tricornutum* on solid medium.

The bacterium used in this study, R217, is able to perform aerobic anoxygenic photoheterotrophy; thus, light may contribute to bacterial energy production, whereas the strain does not possess enzymes for carbon fixation (reviewed in Luo and Moran 2014). Microscopy revealed that the bacteria grew well distributed as well as attached to the diatoms, presumably in order to consume secreted diatom polysaccharides, whereas a competition for light in terms of light-directed taxis was not observed (see Fig. S9 in the Supporting Information). Accordingly, the photoheterotrophy of the bacterial strain is presumably not relevant for the observed interactions.

CONCLUSIONS

This study provides a newly established bacteria/diatom model community that reveals a bacteria-induced conversion of planktonic (fusiform) towards aggregating (oval) *P. tricornutum* cells, which may represent a stress- or defense-related response, as previously discussed (De Martino et al. 2011).

The proteomic identification and bioinformatic characterization of novel soluble extracellular diatom proteins revealed mucin- (or PST-rich) protein domains that so far are unknown in diatoms. The abundant PST-domains appear to be typical for extracellular soluble diatom proteins, which, due to their putative glycosylation, may influence the diffusion dynamics and activity of extracellular proteins.

The proteomics data did not reveal proteins that were significantly upregulated in the presence of bacteria, aside from an AP, of which more peptides were found in axenic cultures compared to bacteria-/diatom co-cultures, for both strains Pt1 and Pt4. Overall, no major differences were found between the extracellular proteomes of *P. tricornutum* strains Pt1 and Pt4, therefore the observed induction of the oval morphotype could not be correlated with any of the identified proteins. The identified proteins, however, were reproducibly found and hint to the importance of PTS-rich domains in extracellular diatom proteins. For investigating whether any of these proteins are involved in the morphotypic interconversion from fusiform to oval *P. tricornutum* diatom cells, it might be more adequate to study their gene expression via quantitative real-time PCR.

Interestingly, a large number of the identified secreted proteins show striking similarities to proteins that are involved in higher plant defense (e.g., secreted hydrolytic enzymes, like endo- β -1,3 glucanases, or S8 serine proteases, or ROS-generating enzymes), but less to proteins that are assumed to be involved in symbiotic interactions (Misas-Villamil and van der Hoorn 2008, Wang et al. 2012). As these defense-related proteins have also been identified in axenic cultures, they are possibly constitutively expressed, e.g., by diatoms in natural biofilms, in which diatoms are always surrounded by distinct bacterial assemblages.

Overall, this work provides access to basic mechanisms of interspecies interactions and the sessile life in multicellular associations, and demonstrates the demand for a further characterization of particularly the extracellular protein fraction.

The authors acknowledge the excellent support by Andreas Marquardt and the Proteomics Facility of the University of Konstanz, as well as Annette Ramsperger and Doris Ballert for their technical support. We further would like to thank Karin Denger, University of Konstanz, for the gift of the *Roseovarius* sp. strain 217 and Michael Schroda for very helpful discussions on quantitative proteomics experiments. This work was supported by the University of Konstanz, as well as the Deutsche Forschungsgemeinschaft, DFG (SFB454, project B11).

Abdullahi, A. S., Underwood, G. J. C. & Gretz, M. R. 2006. Extracellular matrix assembly in diatoms (*Bacillariophyceae*). Environmental effects on polysaccharide synthesis in the model diatom *Phaeodactylum tricornutum*. *J. Phycol.* 42:363–78.

- Aboal, M., Marco, S., Chaves, E., Mulero, I. & García Ayala, A. 2012. Ultrastructure and function of stalks of the diatom *Didymosphenia geminata*. *Hydrobiologia* 695:17–24.
- Ambort, D., van der Post, S., Johansson, M. E., Mackenzie, J., Thomsson, E., Krengel, U. & Hansson, G. C. 2011. Function of the CysD domain of the gel forming MUC2 mucin. *Biochem. J.* 436:61–70.
- Ammerman, J. W. & Azam, F. 1985. Bacterial 5 nucleotidase in aquatic ecosystems: a novel mechanism of phosphorus regeneration. *Science* 227:1338–40.
- Asano, T., Taki, J., Yamamoto, M. & Aono, R. 2002. Cloning and structural analysis of bglM gene coding for the fungal cell wall lytic β 1,3 glucan hydrolase BglM of *Bacillus circulans* IAM1165. *Biosci. Biotechnol. Biochem.* 66:1246–55.
- Asea, A., Kraeft, S. K., Kurt Jones, E. A., Stevenson, M. A., Chen, L. B., Finberg, R. W., Koo, G. C. & Calderwood, S. K. 2000. HSP70 stimulates cytokine production through a CD14 dependant pathway, demonstrating its dual role as a chaperone and cytokine. *Nat. Med.* 6:435–42.
- Borowitzka, M. A. & Volcani, B. E. 1978. The polymorphic diatom *Phaeodactylum tricornutum*: ultrastructure of its morphotypes. *J. Phycol.* 14:10–21.
- Bruckner, C. G., Bahuliker, R., Rahalkar, M., Schink, B. & Kroth, P. G. 2008. Bacteria associated with benthic diatoms from Lake Constance: phylogeny and influences on diatom growth and secretion of extracellular polymeric substances. *Appl. Environ. Microbiol.* 74:7740–9.
- Bruckner, C. G., Rehm, C., Grossart, H. P. & Kroth, P. G. 2011. Growth and release of extracellular organic compounds by benthic diatoms depend on interactions with bacteria. *Environ. Microbiol.* 13:1–8.
- Buchan, A., González, J. & Moran, M. 2005. Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* 71:5665–77.
- Buhmann, M., Kroth, P. G. & Schleheck, D. 2011. Photoautotrophic heterotrophic biofilm communities: a laboratory incubator designed for growing axenic diatoms and bacteria in defined mixed species biofilms. *Environ. Microbiol. Reports* 4:133–40.
- Burns, J. L., Ginn, B. R., Bates, D. J., Dublin, S. N., Taylor, J. V., Apkarian, R. P., Amaro Garcia, S., Neal, A. L. & DiChristina, T. J. 2009. Outer membrane associated serine protease involved in adhesion of *Shewanella oneidensis* to Fe(III) oxides. *Environ. Sci. Technol.* 44:68–73.
- Chen, C., Darrow, A. L., Qi, J. S., D'Andrea, M. R. & Andrade-Gordon, P. 2003. A novel serine protease predominately expressed in macrophages. *Biochem. J.* 374:97–107.
- Chen, Z., Zhang, J., Lei, X., Lai, Q., Yang, L., Zhang, H., Li, Y., Zheng, W., Tian, Y. & Yu, Z. 2015. *Mameliella phaeodactyli* sp. nov., a member of the family *Rhodobacteraceae* isolated from the marine algae *Phaeodactylum tricornutum*. *Int. J. Syst. Evol. Microbiol.* 65:1617–21.
- Cooper, D. M., Pechkovsky, D. V., Hackett, T. L., Knight, D. A. & Granville, D. J. 2011. Granzyme K activates protease activated receptor 1. *PLoS ONE* 6:e21484.
- Davis, K. R., Darvill, A. G., Albersheim, P. & Dell, A. 1986. Host pathogen interactions XXIX. Oligogalacturonides released from sodium pectate by endopolygalacturonic acid lyase are elicitors of phytoalexins in soybean. *Plant Physiol.* 80:568–77.
- De Martino, A., Bartual, A., Willis, A., Meichenin, A., Villazan, B., Maheswari, U. & Bowler, C. 2011. Physiological and molecular evidence that environmental changes elicit morphological interconversion in the model diatom *Phaeodactylum tricornutum*. *Protist* 162:462–81.
- De Martino, A., Meichenin, A., Shi, J., Pan, K. & Bowler, C. 2007. Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. *J. Phycol.* 43:992–1009.
- De Medina Redondo, M., Arnáiz Pita, Y., Fontaine, T., Del Rey, F., Latgé, J. P. & De Aldana, C. R. V. 2008. The β 1,3 glucanoyl transferase gas4p is essential for ascospore wall maturation and spore viability in *Schizosaccharomyces pombe*. *Mol. Microbiol.* 68:1283–99.

- Di Matteo, A., Federici, L., Mattei, B., Salvi, G., Johnson, K., Savino, C., De Lorenzo, G., Tsernoglou, D. & Cervone, F. 2003. The crystal structure of polygalacturonase inhibiting protein (PGIP), a leucine rich repeat protein involved in plant defense. *Proc. Natl. Acad. Sci. U S A* 100:10124–8.
- Edgar, L. A. 1983. Mucilage secretions of moving diatoms. *Proto plasma* 118:44–8.
- Fibriansah, G., Masuda, S., Koizumi, N., Nakamura, S. & Kumasaka, T. 2007. The 1.3 Å crystal structure of a novel endo β 1,3 glucanase of glycoside hydrolase family 16 from alkaliphilic *Nocardiopsis* sp. strain F96. *Proteins: Struct., Funct., Bioinf.* 69:683–90.
- Flynn, K., Opik, H. & Syrett, P. 1986. Localization of the alkaline phosphatase and 5' nucleotidase activities of the diatom *Phaeodactylum tricornutum*. *J. Gen. Microbiol.* 132:289–98.
- Fridovich, I. 1995. Superoxide radical and superoxide dismutases. *Ann. Rev. Biochem.* 64:97–112.
- Gardes, A., Iversen, M. H., Grossart, H. P., Passow, U. & Ullrich, M. S. 2010. Diatom associated bacteria are required for aggregation of *Thalassiosira weissflogii*. *ISME J.* 5:436–45.
- Guillard, R. R. L. & Ryther, J. H. 1962. Studies of marine planktonic diatoms I. *Cyclotella nana* Hustedt and *Detonula confervacea* Cleve. *Can. J. Microbiol.* 8:229–39.
- Hallmann, A., Amon, P., Godl, K., Heitzer, M. & Sumper, M. 2001. Transcriptional activation by the sexual pheromone and wounding: a new gene family from *Volvox* encoding modular proteins with (hydroxy)proline rich and metalloproteinase homology domains. *Plant J.* 26:583–93.
- Hazes, B. 1996. The (QxW) 3 domain: a flexible lectin scaffold. *Prot. Sci.* 5:1490–501.
- Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S. & Grutter, M. G. 1991. Refined crystal structures of subtilisin Novo in complex with wild type and two mutant eglins: comparison with other serine proteinase inhibitor complexes. *J. Mol. Biol.* 217:353–71.
- Heitzer, M. & Hallmann, A. 2002. An extracellular matrix localized metalloproteinase with an exceptional QEXXH metal binding site prefers copper for catalytic activity. *J. Biol. Chem.* 277:28280–6.
- Hoagland, K. D., Rosowski, J. R., Gretz, M. R. & Roemer, S. C. 1993. Diatom extracellular polymeric substances: function, fine structure, chemistry, and physiology. *J. Phycol.* 29:537–66.
- Huber, O. & Sumper, M. 1994. Algal CAMs: isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila* fasciclin I. *EMBO J.* 13:4212–22.
- Knowles, M. R. & Boucher, R. C. 2002. Mucus clearance as a primary innate defense mechanism for mammalian airways. *J. Clin. Invest.* 109:571–7.
- Kroger, N., Bergsdorf, C. & Sumper, M. 2004. Frustulins: domain conservation in a protein family associated with diatom cell walls. *Eur. J. Biochem.* 239:259–64.
- Kroger, N. & Poulsen, N. 2008. Diatoms from cell wall biogenesis to nanotechnology. *Ann. Rev. Gen.* 42:83–107.
- Kwasigroch, J. M., Wintjens, R., Gilis, D. & Rooman, M. 2008. SODA: an Mn/Fe superoxide dismutase prediction and design server. *BMC Bioinformatics* 9:257.
- Lang, T., Hansson, G. C. & Samuelsson, T. 2007. Gel forming mucins appeared early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* 104:16209–14.
- Lee, K. J., Kim, Y. M., Kim, D. Y., Jeoung, D., Han, K., Lee, S. T., Lee, Y. S., Park, K. H., Park, J. H. & Kim, D. J. 2006. Release of heat shock protein 70 (Hsp70) and the effects of extracellular Hsp70 on matrix metalloproteinase 9 expression in human monocytic U937 cells. *Exp. Mol. Med.* 38:364–74.
- Lin, H. Y., Shih, C. Y., Liu, H. C., Chang, J., Chen, Y. L., Chen, Y. R., Lin, H. T., Chang, Y. Y., Hsu, C. H. & Lin, H. J. 2013. Identification and characterization of an extracellular alkaline phosphatase in the marine diatom *Phaeodactylum tricornutum*. *Mar. Biotechnol.* 15:425–36.
- Luna, E., Pastor, V., Robert, J., Flors, V., Mauch Mani, B. & Ton, J. 2010. Callose deposition: a multifaceted plant defense response. *Mol. Plant Microbe Interact.* 24:183–93.
- Luo, H. & Moran, M. A. 2014. Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78:573–87.
- Maheswari, U., Jabbari, K., Petit, J. L., Porcel, B., Allen, A., Cadoret, J. P., De Martino, A. et al. 2010. Digital expression profiling of novel diatom transcripts provides insight into their biological functions. *Genome Biol.* 11:R85.
- Misas Villamil, J. C. & van der Hoorn, R. A. L. 2008. Enzyme inhibitor interactions at the plant pathogen interface. *Curr. Op. Plant Biol.* 11:380–8.
- Nalewajko, C., Lee, K. & Fay, P. 1980. Significance of algal extracellular products to bacteria in lakes and in cultures. *Microb. Ecol.* 6:199–207.
- Nichols, S., Dirks, W., Pearse, J. & King, N. 2006. Early evolution of animal cell signaling and adhesion genes. *Proc. Natl. Acad. Sci. USA* 103:12451–6.
- Pierleoni, A., Martelli, P. L. & Casadio, R. 2008. PredGPI: a GPI anchor predictor. *BMC Bioinformatics* 9:392.
- Plunkett, G., Senear, D. F., Zuroske, G. & Ryan, C. A. 1982. Proteinase inhibitors I and II from leaves of wounded tomato plants: purification and properties. *Arch. Biochem. Biophys.* 213:463–72.
- Popovic, M. M., Bulajic, A., Ristic, D., Krstic, B., Jankov, R. M. & Gavrovic Jankulovic, M. 2012. In vitro and in vivo antifungal properties of cysteine proteinase inhibitor from green kiwi fruit. *J. Sci. Food Agric.* 92:3072–8.
- Potin, P. 2008. Oxidative burst and related responses in biotic interactions of algae. In Amsler, C. [Ed.] *Algal Chemical Ecology*. Springer Verlag, Berlin Heidelberg, pp. 245–71.
- Poulsen, N., Kroger, N., Harrington, M. J., Brunner, E., Paasch, S. & Buhmann, M. T. 2014. Isolation and biochemical characterization of underwater adhesives from diatoms. *Biofouling* 30:513–23.
- Rutenber, E. & Robertus, J. D. 1991. Structure of ricin B chain at 2.5 Å resolution. *Proteins: Struct., Funct., Bioinf.* 10:260–9.
- Santos, J., Almeida, S. P. & Figueira, E. 2013. Cadmium chelation by frustulins: a novel metal tolerance mechanism in *Nitzschia palea* (Kützinger) W. Smith. *Ecotoxicology* 22:166–73.
- Schafer, H., Abbas, B., Witte, H. & Muyzer, G. 2002. Genetic diversity of 'satellite' bacteria present in cultures of marine diatoms. *FEMS Microbiol. Ecol.* 42:25–35.
- Schafer, H., McDonald, I. R., Nightingale, P. D. & Murrell, J. C. 2005. Evidence for the presence of a CmuA methyltransferase pathway in novel marine methyl halide oxidizing bacteria. *Environ. Microbiol.* 7:839–52.
- Schulze, B., Buhmann, M. T., Bártulos, C. R. & Kroth, P. G. 2015. Comprehensive computational analysis of leucine rich repeat (LRR) proteins encoded in the genome of the diatom *Phaeodactylum tricornutum*. *Mar. Genomics* 21:43–51.
- Shimoda, S. & Miyoshi, S. I. 2011. Proteases produced by vibrios. *Biocontrol Sci.* 16:1–11.
- Smith, D. & Underwood, G. 1998. Exopolymer production by intertidal epipellic diatoms. *Limnol. Oceanogr.* 43:1578–91.
- Spiro, R. G. 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12:43R–56R.
- Stanley, M. S. & Callow, J. A. 2007. Whole cell adhesion strength of morphotypes and isolates of *Phaeodactylum tricornutum* (Bacillariophyceae). *Eur. J. Phycol.* 42:191–7.
- Stark, K., Eckart, A., Haidari, S., Tirniceriu, A., Lorenz, M., von Bruhl, M. L., Gartner, F. et al. 2013. Capillary and arteriolar pericytes attract innate leukocytes exiting through venules and 'instruct' them with pattern recognition and motility programs. *Nat. Immunol.* 14:41–51.
- Suda, Y., Rodriguez, R. K., Coluccio, A. E. & Neiman, A. M. 2009. A screen for spore wall permeability mutants identifies a secreted protease required for proper spore wall assembly. *PLoS ONE* 4:e7184.
- Sullivan, J. D. & Ikawa, M. 1973. Purification and characterization of hexose oxidase from the red alga *Chondrus crispus*. *BBA Protein Struct. M.* 309:11–22.
- Varghese, J. N., Garrett, T. P., Colman, P. M., Chen, L., Hoj, P. B. & Fincher, G. B. 1994. Three dimensional structures of two plant beta glucan endohydrolases with distinct substrate specificities. *Proc. Natl. Acad. Sci. USA* 91:2785–9.

- Vincent, W. F. & Roy, S. 1993. Solar ultraviolet B radiation and aquatic primary production: damage, protection, and recovery. *Environ. Rev.* 1:1 12.
- Wagner Dobler, I. & Biebl, H. 2006. Environmental biology of the marine *Roseobacter* lineage. *Ann. Rev. Microbiol.* 60:255 80.
- Wang, S., Zhao, F., Wei, X., Lu, B., Duan, D. & Wang, G. 2012. Preliminary study on flg22 induced defense responses in female gametophytes of *Saccharina japonica* (Phaeophyta). *J. Appl. Phycol.* 25:1215 23.
- Wasiel, A. A., Rozeboom, H. T. J., Hauke, D., Baas, B. J., Zandvoort, E., Quax, W. J., Thunnissen, A. M. W. & Poelarends, G. J. 2010. Structural and functional characterization of a macrophage migration inhibitory factor homologue from the marine cyanobacterium *Prochlorococcus marinus*. *Biochemistry* 49:7572 81.
- Weinberger, F., Leonardi, P., Miravalles, A., Correa, J. A., Lion, U., Kloareg, B. & Potin, P. 2005. Dissection of two distinct defense related responses to agar oligosaccharides in *Gracilaria Chilensis* (Rhodophyta) and *Gracilaria Conferta* (Rhodophyta). *J. Phycol.* 41:863 73.
- Widdel, F. & Pfennig, N. 1981. Studies on dissimilatory sulfate reducing bacteria that decompose fatty acids. *Arch. Microbiol.* 129:395 400.
- Willis, A., Eason Hubbard, M., Hodson, O., Maheswari, U., Bowler, C. & Wetherbee, R. 2014. Adhesion molecules from the diatom *Phaeodactylum tricornutum* (Bacillariophyceae): genomic identification by amino acid profiling and in vivo analysis. *J. Phycol.* 50:837 49.
- Wulff, B. B. H., Chakrabarti, A. & Jones, D. A. 2009. Recognition specificity and evolution in the tomato *Cladosporium fulvum* pathosystem. *Mol. Plant Microbe Interact.* 22:1191 202.

Figure S1. Enzymatic deglycosylation of *Phaeodactylum tricornutum* extracellular proteins revealed differences in banding pattern on silver stained polyacrylamide gels between extracts from axenic or bacteria-induced cultures.

Figure S2. Overview over the most abundantly identified potentially intracellular proteins.

Figure S3. Identified proteins possibly involved in nutrient uptake.

Figure S4. Domain organization of the identified mucin-like proteins.

Figure S5. Identified proteases, protease-inhibitors and sugar-modifying enzymes.

Figure S6. Domain organization of the identified *P. tricornutum* cell surface and frustulin proteins.

Figure S7. Identified extracellular proteins with leucine-rich repeat domain and PST-domain (A) and similar proteins found in *P. tricornutum* genome analyses (B).

Figure S8. Signaling and oxidative stress related proteins.

Figure S9. Micrographs of *Roseovarius* sp. strain 217 in co-culture with *P. tricornutum*.

Appendix S1. Supplementary methods file for protein extraction and mass spectrometry.

Appendix S2. Tabular overview over identified extracellular and intracellular protein hits.

Appendix S3. Alignments of selected identified proteins with characterized orthologs. The following alignments were performed with Clustal W. Substrate binding sites and catalytic residues of crystallized enzymes refer to the pdbsum-database (<http://www.ebi.ac.uk/pdbsum/>) or the MEROPS database (<http://merops.sanger.ac.uk/>).

Appendix S4. Domain organization of selected identified putative mucin-like extracellular *P. tricornutum* proteins. Proteins with more than 50 mol-% AGPSTV amino acids are marked with asterisks.

Appendix S5. Amino acid sequence of LRR-frustulin ID49571. Underlined: predicted signal peptide. Bold, underlined: frustulin ACR domain. Bold: PTSAVD rich repeat. Italics: LRR-domain.

Appendix S6. Amino acid sequence of ID47878. For visualization of EGF domains (cysteine-rich extracellular epithelial growth factor domain, frequently found in extracellular proteins), cysteines are marked black with white letters. The predicted signal peptide is underlined, a putative transmembrane domain is marked with bold italics.