Erschienen in: Journal of Eukaryotic Microbiology ; 63 (2016), 4. - S. 536-546
https://dx.doi.org/10.1111/jeu.12298

Diatom Vacuolar 1,6-β-Transglycosylases can Functionally Complement the Respective Yeast Mutants

Weichao Huang, Carolina Río Bártulos & Peter G. Kroth

Pflanzliche Ökophysiologie, Department of Biology, Universität Konstanz, 78457 Konstanz, Germany

Keywords
Chrysolaminarin; functional complementation; Kre6; vacuolar sorting motif.

Correspondence
W. Huang, Department of Biology, Universität Konstanz, 78457 Konstanz, Germany
Telephone number: +49-7531-88-4230;
FAX number: +49-7531-88-4047;
e-mail: Weichao.Huang@uni-konstanz.de

ABSTRACT

Diatoms are unicellular photoautotrophic algae, which can be found in any aquatic habitat. The main storage carbohydrate of diatoms is chrysolaminarin, a nonlinear β-glucan, consisting of a linear 1,3-β-chain with 1,6-β-branches, which is stored in cytoplasmic vacuoles. The metabolic pathways of chrysolaminarin synthesis in diatoms are poorly investigated, therefore we studied two potential 1,6-β-transglycosylases (TGS) of the diatom *Phaeodactylum tricornutum* which are similar to yeast Kre6 proteins and which potentially are involved in the branching of 1,3-β-glucan chains by adding D-glucose as 1,6-side chains. We genetically fused the full-length diatom TGS proteins to GFP and expressed these constructs in *P. tricornutum*, demonstrating that the enzymes are apparently located in the vacuoles, which indicates that branching of chrysolaminarin may occur in these organelles. Furthermore, we demonstrated the functionality of the diatom enzymes by expressing TGS1 and 2 proteins in yeast, which resulted in a partial complementation of growth deficiencies of a transglycosylase-deficient Δ*kre6* yeast strain.

DIATOMS are unicellular photoautotrophic eukaryotes which contribute considerably to global net primary production (Falkowski et al. 1998). In addition to their ecological importance, diatoms are also very interesting regarding their cell biology for a number of reasons. They have evolved via secondary endosymbiosis, which is the uptake of a eukaryotic photosynthetic endosymbiont into a eukaryotic host cell. This process resulted in plastids with four surrounding membranes and a nuclear genome comprising genes of the former cyanobacterial endosymbiont, as well as those of the primary and the secondary host cells (Archibald 2009). Chrysolaminarin is a unique polysaccharide and, besides lipids, constitutes the principal energy storage compound of diatoms and related algae within the Heterokontophytes (Beattie et al. 1961; Ford and Percival 1965). The water soluble 1,3-β-glucan of diatoms may constitute 10–30% of the organic dry weight in the exponential growth phase of the algae, but it may accumulate up to 80% in the stationary phase (Myklestad 1974). Due to the relatively high contribution of chrysolaminarin to marine particulate matter, diatoms do play an essential role in the global and oceanic biogeochemical cycles (Falkowski et al. 1998; Hama and Handa 1992; Handa 1969; Van Oijen et al. 2004). Chrysolaminarin serves both as a short-term diurnal carbohydrate reserve compound and as a long-term storage polysaccharide. The degradation of the polysaccharides may deliver the carbon skeletons required for the synthesis of amino acids as well as the substrates for respiration (Granum and Myklestad 1999, 2001; Granum et al. 2002; Vårum and Myklestad 1984). Chrysolaminarin represents a branched 1,3-β-glucan with glucose units added at the C-6 position at a ratio of eleven 1,3-β-glucan units to one 1,6-β-glucan unit (Beattie et al. 1961; Ford and Percival 1965). In diatoms, chrysolaminarin is supposed to be stored in intracellular vacuoles in a noncrystalline form (Chiovitti et al. 2004).

The biosynthetic pathway of chrysolaminarin in diatoms as well as the respective enzymes are poorly investigated, therefore we have screened the JGI *Phaeodactylum tricornutum* genome database (Bowler et al. 2008) for genes encoding proteins that are potentially involved in chrysolaminarin synthesis or modification. We identified three gene models encoding putative 1,6-β-transglycosylases (TGS) in the diatom *P. tricornutum*, which we consider to be involved in the synthesis of 1,6-β-linked branches of chrysolaminarin. These three proteins share moderate similarity to conserved domains of Kre6 and Skn1 in *Saccharomyces cerevisiae* (Kroth et al. 2008). These two yeast proteins belong to family-16 glycosyl hydrolases that include proteins involved in the 1,6-β-branching of cell wall 1,3-β-glucans in yeast (Kurita et al. 2011; Nakamata et al. 2007; Roemer and Bussey 1991; Roemer et al. 1993). The

mode of formation of 1,6-β-linkages in side-chain branched 1,3-β-glucans from diatoms and related algae is mostly unknown. However, the modification of 1,3-β-glucan chains by transferring the D-glucose units to the C-6 position by transglycosylases as proposed for fungi is very likely (Bulone 2009).

As there had been reports that in diatoms polysaccharides may be involved in silica cell wall formation by forming an organic layer (Tesson and Hildebrand 2013), it is unclear yet whether the identified *P. tricornutum* TGSs may be involved in either chrysolaminarin modification, or the synthesis of the cell wall, or both. Therefore, the localization and the functional characterization of the TGSs are essential for the elucidation of the biosynthetic pathway of chrysolaminarin in diatoms. Moreover, the localization of TGS proteins may be indicative for the site of biosynthesis of mature chrysolaminarin within the cell.

We have chosen to study *P. tricornutum*, which is one of the model systems for diatoms because of the availability of both a genome sequence (Bowler et al. 2008) as well as tools like genetic transformation (Karas et al. 2015; Zaslavskaia et al. 2000), and gene silencing (De Riso et al. 2009; Lavaud et al. 2012). Recently the transcription activator-like effector nuclease technique for the targeted knocking out of genes has also been successfully implemented in *P. tricornutum* (Daboussi et al. 2014; Weyman et al. 2015). However, such knock down/out approaches may reveal a certain phenotype, but they may not directly demonstrate the nature of the enzymatic activity of the enzyme of interest. We therefore here demonstrate the functionality of the TGS enzymes by achieving functional complementation of *S. cerevisiae* deletion mutants by a homologous gene of interest from a diatom. This kind of complementation assays have been frequently used for genes from plants and humans (Hsu et al. 1993; Schild et al. 1990), but to our knowledge this is the first demonstration that diatom proteins can complement proteins in yeast.

The expression of GFP fusion proteins has been demonstrated to be a useful tool to study the intracellular location of proteins in a variety of organisms including diatoms (Gruber et al. 2007). In this report, we show that both TGS1 and TGS2 contain signal peptides as well as typical signal sorting motifs for endosomes in potential cytoplasmic tails (CTs). When expressed in *P. tricornutum*, GFP fusion proteins of TGS1::GFP and TGS2::GFP were located in or in vicinity of vacuolar membranes indicating that TGS1 and 2 are tonoplast-associated proteins.

MATERIALS AND METHODS

Strains, plasmids, and media

The wild-type and Δ*kre6* yeast strains BY4741 (MATa; his3Δ1 leu2Δ0 met15Δ0 ura3Δ0) and Y05574 (MATa; his3Δ1 leu2Δ0 met15Δ0 ura3Δ0; *kre6Δ*::kanMX4) were used for the complementation experiments. They were obtained from EUROSCARF (EUROpean *S. Cerevisiae* ARchive for Functional Analysis Frankfurt, Germany). BY4741 is the parent strain of Y05574 possessing different selectable markers (Brachmann et al. 1998). *Kre6* had been deleted by integrating *kanMX4* in Y05574. The yeast cDNA expression vector pAG300 (www.addgene.org) was used for complementation with the diatom genes (Horbach et al. 2009). This expression system is comprised of 2 μ origin of replication (2 μ ori), *ura3* selectable marker, and alcohol dehydrogenase 1 promoter and terminator (Vernet et al. 1987).

Yeast cells were grown in YPD (1% Bacto yeast extract [Becton Dickinson and Company, Le Pont de Claix, France], 2% Bacto-peptone [Becton Dickinson and Company], 2% glucose) or in synthetic complete medium (SC) (0.67% yeast nitrogen base without amino acids [Becton Dickinson and Company], 0.01% amino acids without uracil [adenine, arginine, cysteine, leucine, lysine, threonine], 0.05% amino acids [aspartic acid, histidine, isoleucine, methionine, phenylalanine, proline, serine, tyrosine, valine], 2% glucose, 2% Bacto-agar [Difco laboratories, Detroit Michigan, USA]).

Synthesis of TGS1 and 2 cDNA

Cells of *P. tricornutum* were harvested by centrifugation at 5,000 *g* and the pellet was frozen in liquid nitrogen. Total RNA extraction and synthesis of cDNA was done as described previously (Vugrinec 2011).

Plasmid construction and transformation

Phaeodactylum tricornutum Bohlin (University of Texas Culture Collection, strain 646) was cultivated and transformed as described previously (Kroth 2007). The amplified *tgs1* and *2* genes have been cloned into the pPha-T1-HpaI-GFP and the pPha-T1-StuI-GFP vectors, respectively, as described in (Vugrinec 2011). These two vectors are derived from pPha-T1 vectors (GenBank accession AF219942.1) (Zaslavskaia et al. 2000). To clone the genes into the yeast vector, we used primers overlapping with pAG300 and amplified *tgs1* and *2 egfp* from the respective *P. tricornutum* vectors. The Gibson Assembly® Master Mix kit (New England Biolabs GmbH, Frankfurt, Germany) has been used to assemble genes into the vector pAG300 (Barnes 1994; Gibson et al. 2009, 2010). Oligos overlapping 15–16 nucleotide pairs with the vector pAG300 have been used for amplifying *tgs1* and *2*.

Tgs1, tgs2, as well as *tgs1*::*egfp* or *tgs2*::*egfp* gene fusions were assembled together with the linear vector pAG300 that had been digested with XbaI and XhoI using Gibson Assembly® Master Mix kit generating the vectors V_ptTGS1, V_ptTGS1::eGFP, V_ptTGS2, and V_ptTGS2::eGFP. Sequences of the fragments were confirmed by DNA sequencing. The plasmids were transformed into *S. cerevisiae* Δ*kre6* strain BY05574 using a protocol that was modified from the "Quick and Dirty" transformation protocol (Amberg et al. 2006) to yield the complemented yeast strains: T_ptTGS1, T_ptTGS1::eGFP, T_ptTGS2, and T_ptTGS2::eGFP. As a control,

BY4741 was transformed with the empty vector pAG300. After transformation, the yeast cells were grown on YSCM agar lacking uracil.

Fluorescence microscopy and the staining of vacuoles in *P. tricornutum*

The transformed cell lines have been examined with an epifluorescence microscope, Olympus BX51 (Olympus Europe, Hamburg, Germany), being equipped with a Zeiss AxioCam MRm digital camera system (Carl Zeiss MicroImaging GmbH, Göttingen, Germany). To view transmitted light images (100 × UplanFL objective; Olympus Europe), Normarski's differential interference contrast illumination was used.

For visualization of the structure of vacuoles in *P. tricornutum*, a green fluorescent vacuole membrane marker, MDY-64, was applied (Molecular Probes, Leiden, Netherlands). One microliter of 1 mM MDY-64 was diluted into 499 μl liquid samples (containing about 10^6 cells) of *P. tricornutum* in modified f/2 medium with artificial half-concentrated sea salts (16.6 g/l) and 0.09 μM $MnCl_2$ (Guillard 1975; Guillard and Ryther 1962). The final working concentration of MDY-64 was 2 μM, the cells were incubated at room temperature for 2 min. Green fluorescence was analyzed with the epifluorescence microscope.

In silico protein analyses

For signal peptides and transmembrane helices predictions, the software HMMER and SignalP were used (http://hmmer.janelia.org/, http://www.cbs.dtu.dk/services/SignalP-3.0/) (Bendtsen et al. 2004; Finn et al. 2011). For protein targeting prediction TargetP was utilized (www.cbs.dtu.dk/services/TargetP/) (Emanuelsson et al. 2000).

Phylogenetic analyses

Data Sampling was performed by blasting each *P. tricornutum* TGS protein sequence at the NCBI server (http://www.ncbi.nlm.nih.gov/) for nonredundant protein sequences (nr). Sequences from the following organisms were retrieved: metazoa, fungi, red algae, viridiplantae, alveolata, haptophyta, cryptophyta, stramenopiles, rhizaria, archaea, and prokarya. To obtain further stramenopile sequences, we screened the JGI genome database (www. http://genome.jgi.doe.gov/) for stramenopiles genomes. The data sets were reduced by discarding redundant sequences, too short sequences and sequences that were too divergent and not able to align to the *P. tricornutum* sequences. This way a data set of 66 fungal and stramenopile sequences was obtained. After aligning the amino acid sequences using ClustalW with the default settings for slow/accurate alignment and the output format Phylip (http://www.genome.jp/tools/clustalw/), the alignment was manually curated, gaps and ambiguously aligned regions were discarded, yielding 530 amino acid positions (see Supplement document 1). Phylogenetic trees were constructed based on maximum likelihood using the online program PhyML (Guindon and Gascuel 2003; Guindon et al. 2010). We used the PhyML version with automatic model selection by smart model selection and the Akaike Information Criterion (http://www.atgc-montpellier.fr/phyml-sms/). The substitution model WAG+G6+I was selected by the program (Whelan and Goldman 2001). Bootstrap analysis with 100 replicates was performed. The tree was displayed by the program Tree v 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS AND DISCUSSION

Sequence analyses

In order to identify potential transglycosylases, we screened the *P. tricornutum* genome database (JGI, version 2.0) by blasting Kre6 (Saccharomyces genome database [SGD] ID: S000006363) and Skn1 (SGD ID: S000003375) sequences from *S. cerevisiae* against *P. tricornutum* sequences. This way we identified three genes encoding putative 1,6-β-transglycosylases (gene models: 50238 [TGS1], 56509 [TGS2], 48300 [TGS3]). To confirm the similarity of the identified gene models to other transglycosylases, we performed alignments using the conserved TGS domains of known TGS protein sequences from yeast, showing that the *P. tricornutum* sequences cluster with other TGS sequences (Fig. S1). The deduced full protein sequences of the three putative *P. tricornutum* TGSs show a similar structure (Fig. 1A). They all contain an N-terminal signal peptide and a single transmembrane helix at the C-termini followed by a short domain that may represent a CT. This domain structure is clearly different to that of the respective yeast proteins which have a C-terminal transmembrane helix flanked by a long N-terminal CT domain without any signal peptide (Fig. 1A). Despite these structural domain differences between the *P. tricornutum* and *S. cerevisiae* TGS proteins, the conserved parts of TGS1, 2 and 3 still share a strong similarity of 37.5%, 35.5%, and 33.1% amino acid identity, respectively, to the Kre6 protein of *S. cerevisiae* (Fig. S1). Our phylogenetic analyses support these findings (Fig. 2). A deep blast search and a stringent removal of ambiguous sequences, resulted in a TGS tree with only stramenopile and fungal sequences. Interestingly, all TGS sequences in the obtained tree apparently have a common ancestor. The fungal sequences belong to the family 16 glycoside hydrolases (Barbeyron et al. 1998). The obtained tree shows Ascomycota and Basidomycota as sister groups (100 Bootstrap support), while both fungal clades are also close to the Oomycete sequences, which are nonphotosynthetic stramenopiles. The photosynthetic stramenopiles form a single branch, clearly separating TGS3 sequences from TGS1 and 2 sequences.

According to TargetP and SignalP presequence predictions, all three proteins (TGS1, 2 and 3) have a functional signal peptide, indicating that they are targeted to the secretory/ER system. In eukaryotic cells, the trafficking of proteins within the secretory pathways may rely on vari-

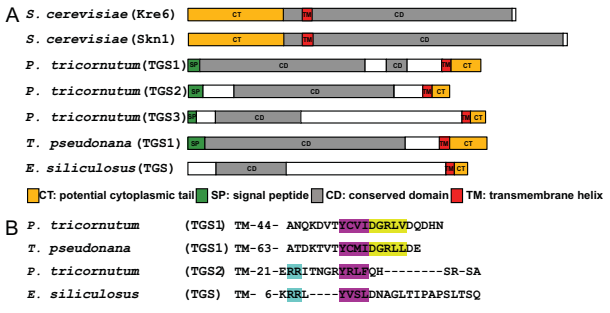

Figure 1 Predicted domain structures of TGS proteins from different organisms. (**A**) Putative structures of TGSs from the yeast *Saccharomyces cerevisiae* (Kre6 SGD ID: S000006363; Skn1 SGD ID: S000003375), the diatoms *Phaeodactylum tricornutum* (JGI ID: TGS1 50238, TGS2 56509, TGS3 48300), and *Thalassiosira pseudonana* (TGS1, XP_002288656) and the brown alga *Ectocarpus siliculosus* (NCBI: CBJ31299). CD, conserved TGS domain; CT, potential cytoplasmic tail; SP, signal peptide; TM, transmembrane helix. (**B**) Domains of the potential cytoplasmic tails of TGS1 and 2: Blue: di-basic RR motif; purple: Tyr-based motif; yellow: di-Leu-type motif.

ous additional sorting signals (Bonifacino and Traub 2003; Trowbridge et al. 1993). Most of the signals of proteins for the ER-Golgi-endosome network can be found in so-called CTs. Membrane proteins are initially synthesized in the ER, after correct folding and assembly they are delivered to the Golgi apparatus via the coat protein complex II (COPII) transport vesicles (Barlowe et al. 1994; Bi et al. 2002). Specific sorting signals are required for exporting the proteins from the ER to COPII. In yeast, mammalian, and plant cells, typical export signals from the ER to COPII of membrane cargo proteins have been identified, which include the di-acidic motif (DXE), the di-hydrophobic motif (LL), the di-aromatic motif (FF, YY), the di-basic motif (RR, RK), and the Tyr-based motif Y*XX*Ø (*X* is any amino acid, and Ø is a bulky hydrophobic group) (Barlowe 2003; Giraudo and Maccioni 2003; Hanton et al. 2005; Sato and Nakano 2002). Furthermore, for targeting to the trans-Golgi network, one more sorting motif is required at the CT, which can be recognized by adaptor protein (AP) complexes leading to transport vesicles. The major type of transport vesicles mediating post-Golgi traffic are clathrin-coated vesicles. The resulting transport vesicles are carrying proteins to the final destinations such as vacuolar membranes and plasma membranes (PM) (Hwang and Robinson 2009; Robinson and Pimpl 2014). Two major classes of signals recognized by AP complexes have been identified and characterized in plants, yeast and in animals. These are Tyr-based sorting signals (YXXØ, with *X* being any amino acid and Ø an amino acid with a bulky hydrophobic residue) and di-Leu-type sorting signals [D/E]XXXL[L/I]. There are also different adaptors which can recognize other Tyr-based signals (NPXY) and other di-Leu-type signals (DXXLL) (Bonifacino and Traub 2003; Pandey 2010; Robinson 2004; Wolfenstetter et al. 2012).

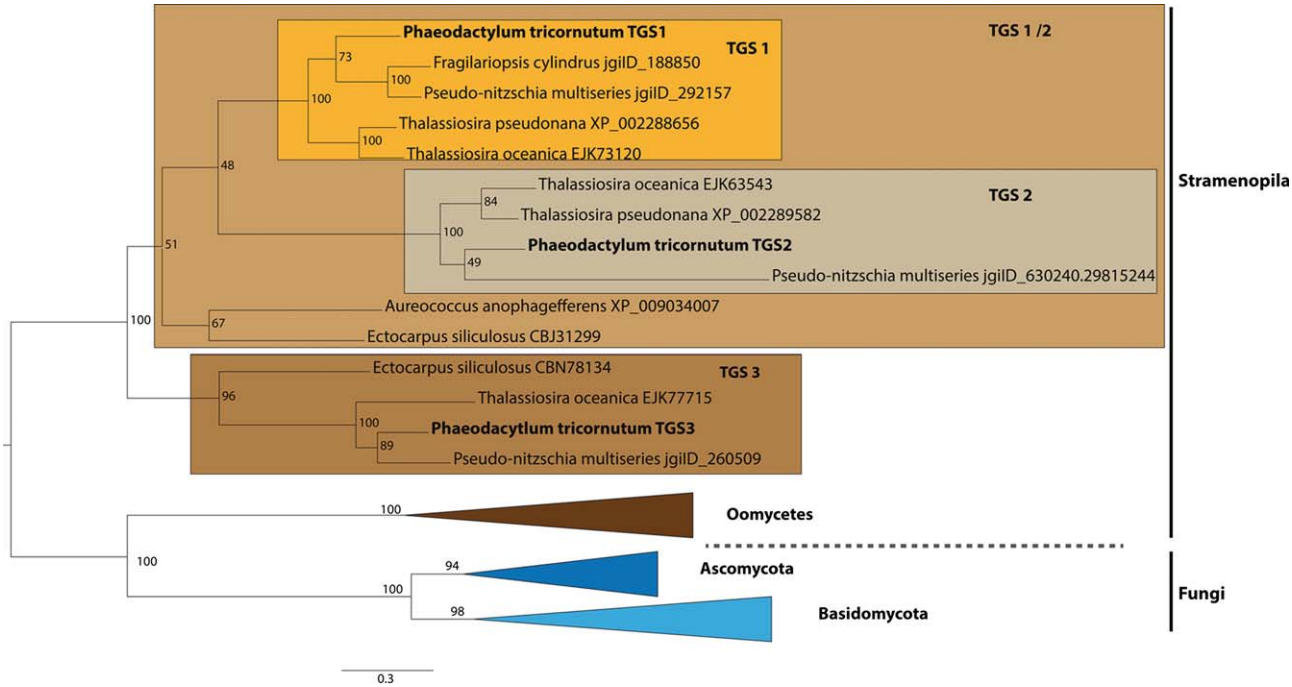

Figure 2 Maximum likelihood tree using PhyML, based on 66 TGS sequences and 530 amino acid positions. *Phaeodactylum tricornutum* TGS sequences are boldfaced. Numbers at nodes correspond to bootstrap values. The accession numbers are listed behind the strain names, and correspond to GenBank entries, unless otherwise noted. The used alignment for this tree is listed in supplement document 1. The tree was midpoint rooted. The Oomycete, Basidomycota, and Ascomycota subtrees are drawn in collapsed form.

We have analyzed the CTs of the identified TGSs from diatoms (*P. tricornutum*, *Thalassiosira pseudonana*) and a brown alga (*Ectocarpus siliculosus*). The amino acid sequences of the potential CTs of *P. tricornutum* TGS1 and *T. pseudonana* TGS (protein ID: 3105) are very similar, they both contain the YXXØ as well as the [D/E]XXL[L/I]-type motifs (Fig. 1B). We could not find indications that this combination of signal motifs has been identified in other organisms before. YXXØ motifs can, besides binding to APs, also be recognized by COPII (Sato and Nakano 2002). Therefore, we propose that the YC[V/M]I motif of TGS1 is required for exporting proteins from the ER, while the DGRL[V/L] motif is required for targeting to the trans-Golgi network. The potential CTs of *P. tricornutum* TGS2 and *E. siliculosus* TGS have similar motifs; they both possess the di-basic motif RR next to the transmembrane helices (Fig. 1B). Di-basic motifs in the CT of animal glycosyltransferases, have been demonstrated to be important for exporting protein from the ER (Giraudo and Maccioni 2003). Furthermore, at the C-terminus of both sequences, a Tyr-based motif Y*XX*Ø has been identified (Fig. 1B). In conclusion, both, TGS1 and 2, contain ER-exiting-like and endosomes-like sorting motifs, indicating that TGS1 and 2 may be sorted to the vacuolar or the PM of *P. tricornutum*. As the sorting motifs of both TGS1 and 2 are different, they might be recognized by different APs. Within the respective domain of *P. tricornutum* TGS3, we could not detect any typical of such sorting signals. As furthermore the phylogenetic analyses indicate that TGS3 sequences form a clade separate from the TGS1/2 sequences, and screening of the *P. tricornutum* EST database (http://www.diatomics.biologie.ens.fr/EST3/index.php) (Maheswari et al. 2005, 2009) imply that TGS3 is only very poorly expressed, we decided to focus only on TGS1 and 2 in the following experiments.

Intracellular localization of TGS1 and 2 in *P. tricornutum*

In order to be able to address vacuolar structures in the diatom *P. tricornutum*, we established a staining procedure for diatom vacuoles, using the fluorescent vacuole membrane marker MDY-64. We found that incubation of the cells for 2 min at a concentration of 2 μM MDY-64 yielded the best results. Two different fluorescent staining patterns of vacuoles were observed. While most of the cells of *P. tricornutum* contain 2–3 big vacuoles associated with a few vesicles (Fig. 3B), some of the cells revealed a high number of vesicles associated with the vacuoles forming a continuous membraneous net (Fig. S2A). In a next step, we genetically fused the *egfp* gene to the 3′ end of both the genes PtTgs1 and PtTgs2 (Fig. 3A) and expressed the fusion proteins in *P. tricornutum*. We had decided to fuse the GFP protein to the C-termini of the TGS proteins because: (i) a fusion to the N-terminus would clearly have blocked the signal peptide, and (ii) there are reports from *Arabidopsis* and tobacco BY2 cells, that GFP, even when fused to the C-termini of the respective tonoplast proteins, does not affect the tonoplast localization/targeting (Czempinski et al. 2002; Wolfenstetter et al. 2012). Fluorescence microscopy images yielded similar fluorescent green morphology patterns as previously obtained with the vacuole dye, indicating that TGS1 and 2 may represent tonoplast-associated proteins (Figs 3C, D, S2B, C). This finding supports the prediction that the identified DGRLV and YRLF motifs of TGS1 and 2 indeed may serve as tonoplast sorting motifs in *P. tricornutum*. However, besides a location at the vacuole-like membranes, green fluorescent vesicle membranes associated with vacuoles were also observed in TGS2::GFP expressed cell lines. These vesicles may participate in transporting the proteins to vacuoles. Nevertheless, they were not obtained in TGS1::GFP expressed cell lines. These different fluorescent vacuolar membranes patterns may be explained by the different vacuolar sorting motifs of TGS1 and 2 (see above).

The Kre6 protein of yeast, which is involved in the synthesis of the cell wall 1,6-β-glucans, has been localized in yeast cell membranes of the endoplasmic reticulum (ER), Golgi apparatus and the PM (Kurita et al. 2011; Roemer et al. 1994). So far, there is no evidence that Kre6 may be associated with vacuolar membranes of yeast. These different localizations, together with the different domain structure may indicate that TGS1 and 2 in contrast to Kre6, may not be involved in the synthesis of the organic cell wall of *P. tricornutum*.

Expression of GFP fusion proteins in Δ*kre6* yeast strain BY05574

In order to demonstrate whether TGS1 and 2 from *P. tricornutum* can complement Kre6 deficient yeast cells, we expressed these genes in the respective yeast mutant using a yeast vector. The expression of heterologous genes can be hampered, if the codon usage of the recipient strain is not suitable. Therefore, to demonstrate successful expression, we attempted to visualize gene expression of TGS1 and 2 in the yeast cells by genetically fusing GFP as a reporter protein to the C-termini of TGS1 and 2 from *P. tricornutum*, respectively. We inserted the fusion constructs into the vector pAG300, and expressed these constructs in yeast cells. The obtained mutant cell lines showed green fluorescence (Fig. 4), indicating that indeed the diatom proteins can be expressed in required amounts in yeast. Interestingly, the GFP fluorescence of the fusion proteins was found to be present throughout the yeast cells, while not being evenly distributed. This could indicate that some further transport signals for protein sorting in yeast cells may be missing or misinterpreted, resulting in an ambiguous targeting of the diatom proteins. It also has been reported that Kre6 in yeast cells may be transported from the ER to the plasma membrane, but also to budding growth sites (Kurita et al. 2011). As Fig. 4 shows, the green fluorescence of overexpressed TGS1 and 2::GFP proteins in Δ*kre6* cells has been observed in the cytosol, the ER, possibly the tonoplast and some secondary pathway organelles, but never at the sites of developing budding daughter cells.

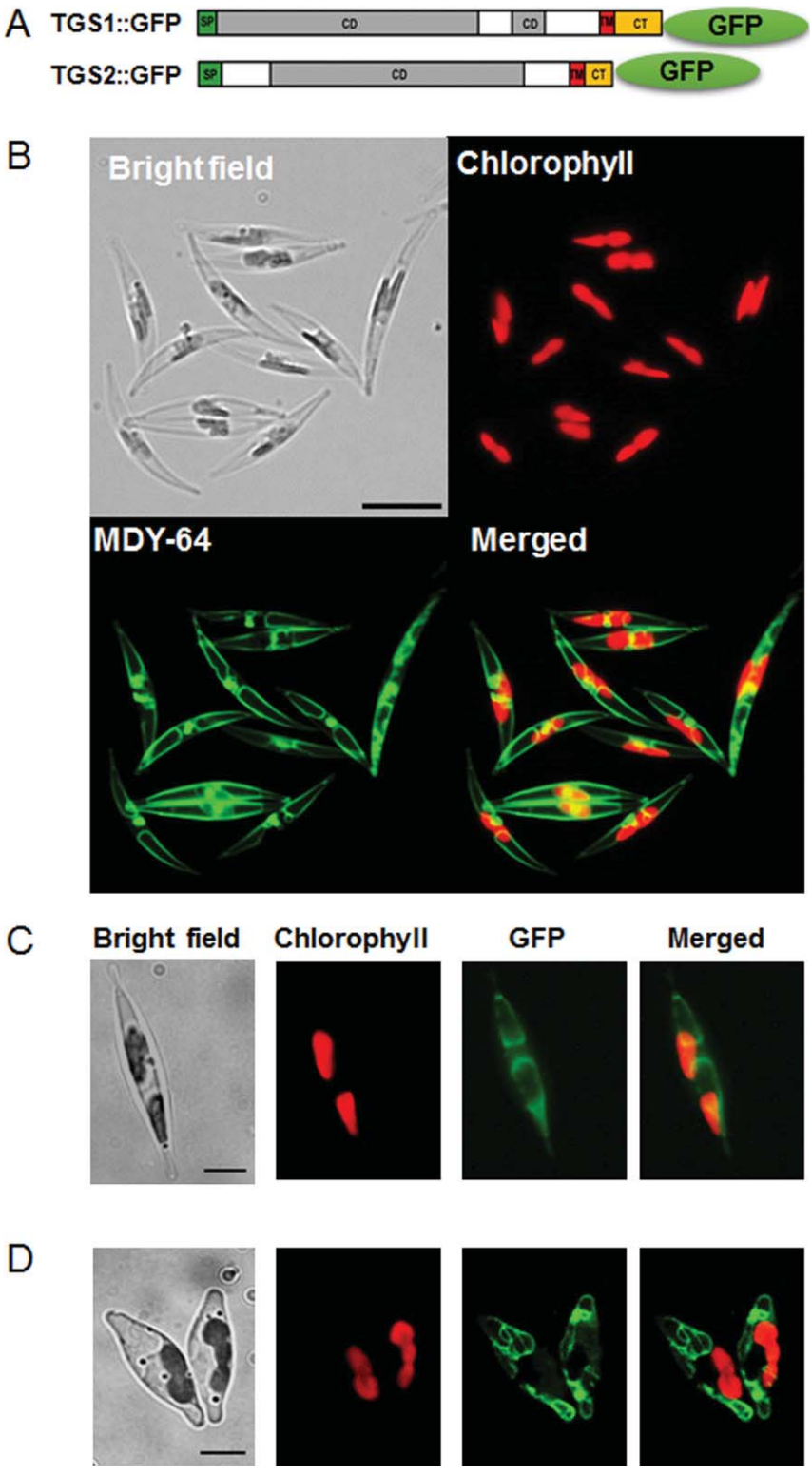

Figure 3 Analysis of the cellular localization of TGS1 and 2 in *Phaeodactylum tricornutum*. (**A**) Constructs of TGS1 and 2::GFP. TGS1 or 2 was C-terminally fused to GFP. (**B**) Staining of vacuoles with MDY-64. Upper left: bright filed image, upper right: chlorophyll fluorescence of the plastids; lower left: MDY-64 fluorescence; lower right, merged image. (**C, D**) Targeting of full-length TGS1::GFP and TGS2::GFP fusion proteins, respectively. From left to right: bright field image, chlorophyll fluorescence, GFP fluorescence, merged image. Bars indicate a length of 5 μm.

Determination of the function of TGS1 and 2: complementation of Δ*kre6* yeast strain with TGS1 and 2

In a next step, we tested the functionality of the diatom TGS proteins by expressing TGS1 and 2 (with and without attached GFP) in the kre6-deficient *S. cerevisiae* strain Δ*kre6* (Kurita et al. 2011; Roemer et al. 1993, 1994) and by screening for complementation. Deletion of *kre6* in yeast shows a clear phenotype due to distinct cell wall defects, resulting in a very retarded growth combined with a hypersensitivity to the stain Calcofluor White (CFW). The *kre6* mutant cells fail to grow in the presence of CFW, which is not lethal to wild-type cells (Kurita et al. 2011; Ram et al. 1994). The goal of this approach was the demonstration that the diatom TGS enzyme may rescue Δ*kre6* cells in the presence of CFW.

We performed complementation experiments by characterizing wild-type yeast cells, the Δ*kre6* strain and Δ*kre6* strains overexpressing TGS1 or 2 of *P. tricornutum*. When growing on agar plates, the wild-type cells (transformed with an empty vector) were not sensitive to a concentration of 30 μg CFW, while the Δ*kre6* cells showed a clear growth inhibition. Although the TGS1 or 2 expressing yeast mutants both showed some growth inhibition, they clearly were less affected by CFW than the Δ*kre6* strain alone, indicating that TGS1 and 2 can partially complement the missing Kre6 protein (Fig. 5A, B). Similar results were obtained for the TGS1::GFP and TGS2::GFP fusion proteins, indicating that the GFP protein, when fused to the C-termini, possibly does not affect enzymatic activities. We furthermore studied cellular growth of wild-type and mutant yeast cells in liquid culture. The cells were grown in SC medium including 15 μg/ml of CFW and growth was measured by following OD600 in a spectrophotometer. Figure 6A, B show the growth curves of the different yeast strains, demonstrating that Δ*kre6* cells transformed with empty vectors show a strong growth inhibition, while all strains complemented with TGS1, 2, TGS1::GFP, or TGS2::GFP were able to grow. Interestingly, TGS1 is apparently complementing Δ*kre6* better than TGS2, an effect that was also, although to a lesser extent, visible in the plate assays (Fig. 5A). The results indicate that cell wall defects of the Δ*kre6* strain can be partially rescued by overexpression of TGS1 or 2 in the mutant cell lines, confirming that both proteins are active 1,6-β-transglycosylases. TGS1 and 2 only partially rescued the phenotypes of the Δ*kre6* strain, which is consistent with the complementation by Kre6 itself in *S. cerevisiae* (Kurita et al. 2011; Nakamata et al. 2007). The complementation result is somewhat surprising, considering that Kre6 is lacking the signal peptide which is found in TGS1 and 2, indicating that they may reside in different compartments. However, this could either be explained by potentially different targeting mechanisms to the tonoplast in yeast cells and in diatoms (which possess a chloroplast ER), or by the rather ambiguous targeting of TGS1 and 2 in the yeast cells described above.

Recently, it has been reported that the amount of chrysolaminarin in diatoms was significantly reduced by knocking down UDP-glucose pyrophosphorylase (UGP) (Zhu et al. 2015). UDP-glucose has been demonstrated to supply activated glucose units for chrysolaminarin synthe-

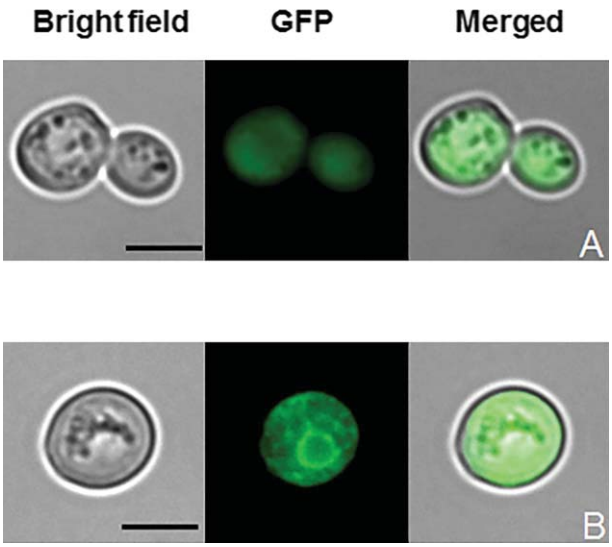

Figure 4 Images of TGS1 and 2::GFP fusion proteins expressed in the yeast strain BY05574. (**A**) Expression of TGS1::GFP. (**B**) Expression of TGS2::GFP. From left to right: bright field image, GFP fluorescence, merged images. Bars indicate length in 5 μm.

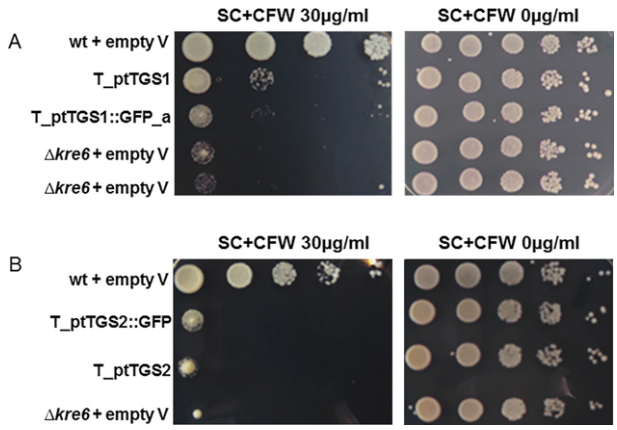

Figure 5 Complementaion of a Δ*kre6* yeast strain by overexpression of TGS1 or 2. (**A**) Sensitivity of strains overexpressing TGS1 to Calcofluor White (CFW). (**B**) CFW sensitivity of strains overexpressing TGS2. (**A, B**) Cells were grown in SC medium at 30 °C overnight until OD_{600nm} = 1, and then serially diluted four times by ten-fold. Eight microliters of these samples were spotted on an SC plate with or without 30 μg/ml CFW, and incubated at 25 °C overnight and then at 22 °C for 2 d. One T_ptTGS1 strain and two independent isolates of T_ptTGS1::GFP (T_ptTGS1::eGFP_a and T_ptTGS1::eGFP_b) have been chosen. One T_ptTGS2 strain and one T_ptTGS2::GFP strain have been selected. wt+empty V: BY4741 transformed with the empty pAG300 vector; Δ*kre6*+emptyV: Y05574 transformed with the vector.

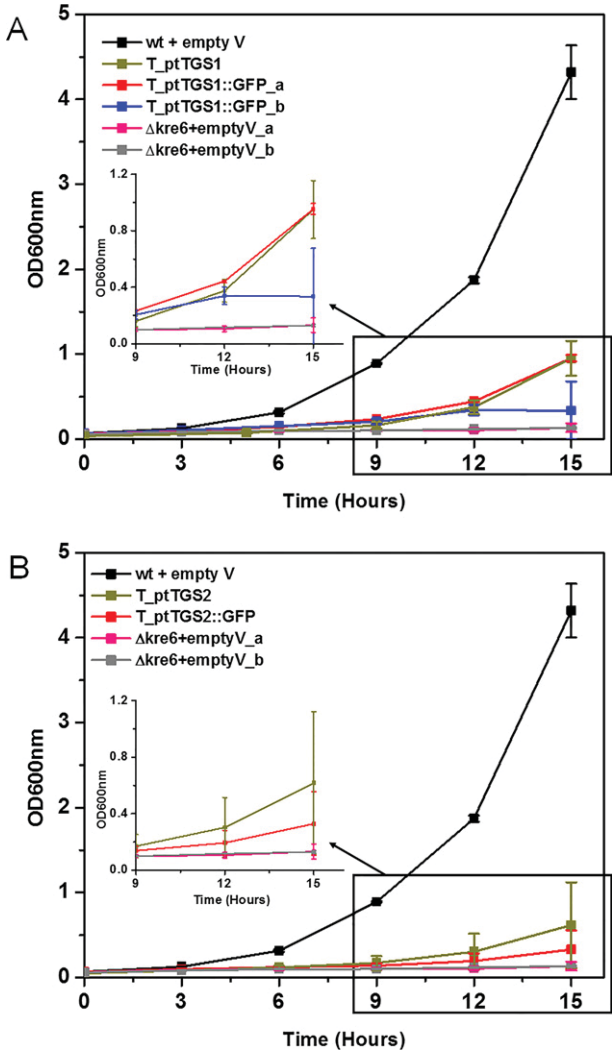

Figure 6 Growth of strains overexpressing TGS1 or 2 in liquid medium with CFW. (**A**) Growth of Δ*kre6*, wild-type, and strains expressing TGS1 in SC with CFW. (**B**) Growth of Δ*kre6*, wild-type, and strains expressing TGS2. Cells were grown in SC medium with 15 µg/ml CFW at 28 °C. Two independent cell lines of Δ*kre6* transformed with the empty vector (Δ*kre6*+emptyV_a, Δ*kre6*+emptyV_b) were used.

sis (Kroth et al. 2008; Roessler 1987). Two putative UDP-glucose pyrophosphorylase gene models (UGP1 [50444], UGP2 [23639]) have been described earlier (Kroth et al. 2008). Interestingly, UGP1 is a fusion protein that contains both UDP-glucose pyrophosphorylase and phosphoglucomutase (PGM) domains (Kroth et al. 2008). According to TargetP and SignalP predictions, UGP1 and 2 are targeted to the cytosol and the plastid, respectively. To elucidate the compartment in which chrysolaminarin is initially synthesized, it is essential to determine which UGP is involved in chrysolaminarin synthesis. Among the two UGPs, only the cytosolic UGP1 was studied so far (Daboussi et al. 2014; Zhu et al. 2015). After knocking down UGP1 in *P. tricornutum*, the cells apparently reallocated carbon flux from carbohydrate to lipid storage (Zhu et al. 2015). These results indicate that UGP1 may take part in the process of chrysolaminarin synthesis and that the synthesis of chrysolaminarin may be initiated in the cytosol, involving a potential glucan synthase located in the vacuolar membrane that may translocate the glucan backbone into the vacuole where it can be modified by the transglycosylases TGS1 and 2.

CONCLUSION

Diatoms, oomycetes, and brown algae possess 1,3-β-glucans, termed as chrysolaminarin, mycolaminarin, and laminarin, respectively. All these polysaccharides have similar structures based on linear 1,3-β-glucan chains branched infrequently with 1,6-β-linked residues. Laminarins furthermore contain mannitol residues which are not found in mycolaminarin and chrysolaminarin (Percival and Ross 1951; Read et al. 1996; Wang and Bartnicki-Garcia 1974). Therefore, the common occurrences of Kre6-like transglycosylases in diatoms, oomycetes, and brown algae as well as the functional complementation of Δ*kre6* by TGS1 and 2 support our hypothesis that TGS1 and 2 might be involved in the synthesis of 1,6-β-linked residues of chrysolaminarin in diatoms. The tonoplast localization of TGS1 and 2 in *P. tricornutum* is a further indication for this role, as within the tonoplast the enzymes may modify chrysolaminarin enzymatically. For the future investigation of chryolaminarin synthesis, it will be important to study the functionality of the TGS enzymes as well as the intracellular location of the key enzyme of this process, the glucan synthase.

ACKNOWLEDGMENTS

We thank Julia Reuther and Dr. Ulrike Mogg (University of Konstanz) for advice regarding the genetic transformation of yeast. We also thank our technician Doris Ballert for help with transformation of *P. tricornutum* cells. This work was supported by a stipend of the CSC (China Scholarship Council) to W.H., by the Graduate School Chemical Biology (KoRS-CB), and by the University of Konstanz.

LITERATURE CITED

Amberg, D. C., Burke, D. J. & Strathern, J. N. 2006. "Quick and Dirty" plasmid transformation of yeast colonies. *CSH Protoc*, 2006. pdb.prot4146. doi: 10.1101/pdb.prot4146.

Archibald, J. M. 2009. The puzzle of plastid evolution. *Curr. Biol.*, 19:R81–R88.

Barlowe, C. 2003. Signals for COPII-dependent export from the ER: what's the ticket out? *Trends Cell Biol.*, 13:295–300.

Barbeyron, T., Gerard, A., Potin, P., Henrissat, B. & Kloareg, B. 1998. The kappa-carrageenase of the marine bacterium Cytophaga drobachiensis. Structural and phylogenetic relationships within family-16 glycoside hydrolases. *Mol. Biol. Evol.*, 15 (5):528–537.

Barlowe, C., Orci, L., Yeung, T., Hosobuchi, M., Hamamoto, S., Salama, N., Rexach, M. F., Ravazzola, M., Amherdt, M. & Schekman, R. 1994. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell*, 77:895–907.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proc. Natl Acad. Sci. USA*, 91:2216–2220.

Beattie, A., Hirst, E. L. & Percival, E. 1961. Studies on the metabolism of the *Chrysophyceae*. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. *Biochem. J.*, 79:531–537.

Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.*, 340:783–795.

Bi, X., Corpina, R. A. & Goldberg, J. 2002. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature*, 419:271–277.

Bonifacino, J. S. & Traub, L. M. 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annu. Rev. Biochem.*, 72:395–447.

Bowler, C., Allen, A. E., Badger, J. H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otillar, R. P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J. A., Brownlee, C., Cadoret, J. P., Chiovitti, A., Choi, C. J., Coesel, S., De Martino, A., Detter, J. C., Durkin, C., Falciatore, A., Fournet, J., Haruta, M., Huysman, M. J. J., Jenkins, B. D., Jiroutova, K., Jorgensen, R. E., Joubert, Y., Kaplan, A., Kroger, N., Kroth, P. G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jezequel, V., Lopez, P. J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L. K., Montsant, A., Oudot-Le Secq, M. P., Napoli, C., Obornik, M., Parker, M. S., Petit, J. L., Porcel, B. M., Poulsen, N., Robison, M., Rychlewski, L., Rynearson, T. A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M. R., Taylor, A. R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L. S., Rokhsar, D. S., Weissenbach, J., Armbrust, E. V., Green, B. R., Van De Peer, Y. & Grigoriev, I. V. 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*, 456:239–244.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P. & Boeke, J. D. 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14:115–132.

Bulone, V. 2009. Biosynthetic enzymes for (1,3)-ß-glucans and (1,3;1,6)-ß-glucans in protozoans and chromistans. Biochemical characterization and molecular biology. *In:* Bacic, A., Stone, B. A. & Fincher, G. B. (eds), Chemistry, biochemistry and biology of 1,3-beta-glucans and related polysaccharides. Academic Press, San Diego. p. 233–258.

Chiovitti, A., Molino, P., Crawford, S. A., Teng, R., Spurck, T. & Wetherbee, R. 2004. The glucans extracted with warm water from diatoms are mainly derived from intracellular chrysolaminaran and not extracellular polysaccharides. *Eur. J. Phycol.*, 39:117–128.

Czempinski, K., Frachisse, J. M., Maurel, C., Barbier-Brygoo, H. & Mueller-Roeber, B. 2002. Vacuolar membrane localization of the Arabidopsis 'two-pore' K+ channel KCO1. *Plant J.*, 29:809–820.

Daboussi, F., Leduc, S., Maréchal, A., Dubois, G., Guyot, V., Perez-Michaut, C., Amato, A., Falciatore, A., Juillerat, A., Beurdeley, M., Voytas, D. F., Cavarec, L. & Duchateau, P. 2014. Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nat. Commun.*, 5. doi: 10.1038/ncomms4831.

De Riso, V., Raniello, R., Maumus, F., Rogato, A., Bowler, C. & Falciatore, A. 2009. Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res.*, 37:e96. doi: 10.1093/nar/gkp448.

Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.*, 300:1005–1016.

Falkowski, P. G., Barber, R. T. & Smetacek, V. V. 1998. Biogeochemical controls and feedbacks on ocean primary production. *Science*, 281:200–207.

Finn, R. D., Clements, J. & Eddy, S. R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.*, 39(Suppl. 2):W29–W37.

Ford, C. W. & Percival, E. 1965. Carbohydrates of *Phaeodactylum tricornutum*. Part II. A sulphated glucuronomannan. *J. Chem. Soc.*, 7042–7046.

Gibson, D. G., Smith, H. O., Hutchison, C. A., Venter, J. C. & Merryman, C. 2010. Chemical synthesis of the mouse mitochondrial genome. *Nat. Methods*, 7:901–U905.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. & Smith, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, 6:343–345.

Giraudo, C. G. & Maccioni, H. J. F. 2003. Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1. *Mol. Biol. Cell*, 14:3753–3766.

Granum, E., Kirkvold, S. & Myklestad, S. M. 2002. Cellular and extracellular production of carbohydrates and amino acids by the marine diatom *Skeletonema costatum*: diel variations and effects of N depletion. *Anglais*, 242:12.

Granum, E. & Myklestad, S. M. 1999. Effects of $NH4^+$ assimilation on dark carbon fixation and β-1,3-glucan metabolism in the marine diatom *Skeletonema costatum* (Bacillariophyceae). *J. Phycol.*, 35:1191–1199.

Granum, E. & Myklestad, S. M. 2001. Mobilization of β-1,3-glucan and biosynthesis of amino acids induced by $NH4^+$ addition to N-limited cells of the marine diatom *skeletonema Costatum* (Bacillariophyceae). *J. Phycol.*, 37:772–782.

Gruber, A., Vugrinec, S., Hempel, F., Gould, S. B., Maier, U. G. & Kroth, P. G. 2007. Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol. Biol.*, 64:519–530.

Guillard, R.R.L. 1975. Culture of phytoplankton for feeding marine invertebrates. *In:* Smith, W. L. & Chanley, M. H. (eds), Culture of marine invertebrate animals. Springer, US. p. 29–60.

Guillard, R. R. L. & Ryther, J. H. 1962. Studies of marine planktonic diatoms: I. *Cyclotella Nana Hustedt*, and *Detonula Confervacea* (Cleve) Gran. *Can. J. Microbiol.*, 8:229–239.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.*, 59:307–321.

Guindon, S. & Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.*, 52:696–704.

Hama, J. & Handa, N. 1992. Diel variation of water-extractable carbohydrate composition of natural phytoplankton populations in Kinu-ura Bay. *J. Exp. Mar. Biol. Ecol.*, 162:159–176.

Handa, N. 1969. Carbohydrate metabolism in the marine diatom *Skeletonema costatum*. *Mar. Biol.*, 4:208–214.

Hanton, S. L., Bortolotti, L. E., Renna, L., Stefano, G. & Brandizzi, F. 2005. Crossing the divide – transport between the endoplasmic reticulum and Golgi apparatus in plants. *Traffic*, 6:267–277.

Horbach, R., Graf, A., Weihmann, F., Antelo, L., Mathea, S., Liermann, J. C., Opatz, T., Thines, E., Aguirre, J. & Deising, H. B. 2009. Sfp-type 4'-phosphopantetheinyl transferase is indispensable for fungal pathogenicity. *Plant Cell*, 21:3379–3396.

Hsu, L. C., Chiou, T. J., Chen, L. & Bush, D. R. 1993. Cloning a plant amino acid transporter by functional complementation of a yeast amino acid transport mutant. *Proc. Natl Acad. Sci. USA*, 90:7441–7445.

Hwang, I. & Robinson, D. G. 2009. Transport vesicle formation in plant cells. *Curr. Opin. Plant Biol.*, 12:660–669.

Karas, B.J., Diner, R.E., Lefebvre, S.C., McQuaid, J., Phillips, A.P.R., Noddings, C.M., Brunson, J.K., Valas, R.E., Deerinck, T.J., Jablanovic, J., Gillard, J.T.F., Beeri, K., Ellisman, M.H., Glass, J.I., .Hutchison, C.A., Smith, H.O., Venter, J.C., Allen, A.E., Dupont, C.L. & Weyman, P.D. 2015. Designer diatom episomes delivered by bacterial conjugation. *Nat. Commun.*, 6. doi: 10.1038/ncomms7925.

Kroth, P. G. 2007. Genetic transformation: a tool to study protein targeting in diatoms. *Methods Mol. Biol.*, 390:257–267.

Kroth, P. G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M. S., Stanley, M. S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E. V. & Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *PLoS One*, 3:e1426.

Kurita, T., Noda, Y., Takagi, T., Osumi, M. & Yoda, K. 2011. Kre6 protein essential for yeast cell wall beta-1,6-glucan synthesis accumulates at sites of polarized growth. *J. Biol. Chem.*, 286:7429–7438.

Lavaud, J., Materna, A. C., Sturm, S., Vugrinec, S. & Kroth, P. G. 2012. Silencing of the violaxanthin de-epoxidase gene in the diatom *Phaeodactylum tricornutum* reduces diatoxanthin synthesis and non-photochemical quenching. *PLoS One*, 7. doi: 10.1371/journal.pone.0036806.

Maheswari, U., Mock, T., Armbrust, E. V. & Bowler, C. 2009. Update of the Diatom EST Database: a new tool for digital transcriptomics. *Nucleic Acids Res.*, 37(Database issue):1–5.

Maheswari, U., Montsant, A., Goll, J., Krishnasamy, S., Rajyashri, K. R., Patell, V. M. & Bowler, C. 2005. The diatom EST database. *Nucleic Acids Res.*, 33(Database issue):4–7.

Myklestad, S. 1974. Production of carbohydrates by marine planktonic diatoms. I. Comparison of nine different species in culture. *J. Exp. Mar. Biol. Ecol.*, 15:261–274.

Nakamata, K., Kurita, T., Bhuiyan, M. S., Sato, K., Noda, Y. & Yoda, K. 2007. KEG1/YFR042w encodes a novel Kre6-binding endoplasmic reticulum membrane protein responsible for beta-1,6-glucan synthesis in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, 282:34315–34324.

Pandey, K. N. 2010. Small peptide recognition sequence for intracellular sorting. *Curr. Opin. Biotechnol.*, 21:611–620.

Percival, E. G. V. & Ross, A. G. 1951. 156. The constitution of laminarin. Part II. The soluble laminarin of Laminaria digitata. *J. Chem. Soc.*, 720–726. doi:10.1039/JR9510000720.

Ram, A. F. J., Wolters, A., Hoopen, R. T. & Klis, F. M. 1994. A new approach for isolating cell wall mutants in *Saccharomyces cerevisiae* by screening for hypersensitivity to calcofluor white. *Yeast*, 10:1019–1030.

Read, S. M., Currie, G. & Bacic, A. 1996. Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydr. Res.*, 281:187–201.

Robinson, M. S. 2004. Adaptable adaptors for coated vesicles. *Trends Cell Biol.*, 14:167–174.

Robinson, D. G. & Pimpl, P. 2014. Clathrin and post-Golgi trafficking: a very complicated issue. *Trends Plant Sci.*, 19:134–139.

Roemer, T. & Bussey, H. 1991. Yeast beta-glucan synthesis: KRE6 encodes a predicted type II membrane protein required for glucan synthesis in vivo and for glucan synthase activity in vitro. *Proc. Natl Acad. Sci. USA*, 88:11295–11299.

Roemer, T., Delaney, S. & Bussey, H. 1993. SKN1 and KRE6 define a pair of functional homologs encoding putative membrane proteins involved in beta-glucan synthesis. *Mol. Cell. Biol.*, 13:4039–4048.

Roemer, T., Paravicini, G., Payton, M. A. & Bussey, H. 1994. Characterization of the yeast (1–>6)-beta-glucan biosynthetic components, Kre6p and Skn1p, and genetic interactions between the PKC1 pathway and extracellular matrix assembly. *J. Cell Biol.*, 127:567–579.

Roessler, P. G. 1987. UDPGlucose pyrophosphorylase activity in the diatom *Cyclotella cryptica*. Pathway of chrysolaminarin biosynthesis. *J. Phycol.*, 23:494–498.

Sato, K. & Nakano, A. 2002. Emp47p and its close homolog Emp46p have a tyrosine-containing endoplasmic reticulum exit signal and function in glycoprotein secretion in *Saccharomyces cerevisiae*. *Mol. Biol. Cell*, 13:2518–2532.

Schild, D., Brake, A. J., Kiefer, M. C., Young, D. & Barr, P. J. 1990. Cloning of three human multifunctional de novo purine biosynthetic genes by functional complementation of yeast mutations. *Proc. Natl Acad. Sci. USA*, 87:2916–2920.

Tesson, B. & Hildebrand, M. 2013. Characterization and localization of insoluble organic matrices associated with diatom cell walls: Insight into their roles during cell wall formation. *PLoS One*, 8:e61675.

Trowbridge, I. S., Collawn, J. F. & Hopkins, C. R. 1993. Signal-dependent membrane protein trafficking in the endocytic pathway. *Annu. Rev. Biochem.*, 9:129–161.

Van Oijen, T., Van Leeuwe, M. A., Gieskes, W. W. C., De Baar, H. J. W. & Baar, H. J. W. D. 2004. Effects of iron limitation on photosynthesis and carbohydrate metabolism in the Antarctic diatom *Chaetoceros brevis* (Bacillariophyceae). *Eur. J. Phycol.*, 39:161–171.

Vårum, K. M. & Myklestad, S. 1984. Effects of light, salinity and nutrient limitation on the production of β-1,3-d-glucan and exo-d-glucanase activity in *Skeletonema costatum* (Grev.) Cleve. *J. Exp. Mar. Biol. Ecol.*, 83:13–25.

Vernet, T., Dignard, D. & Thomas, D. Y. 1987. A family of yeast expression vectors containing the phage f1 intergenic region1. *Gene*, 52:225–233.

Vugrinec, S. 2011. *Protein targeting into diatom plastids*. Dissertation, phd thesis. http://nbn-resolving.de/urn:nbn:de:bsz:352-133680.

Wang, M. C. & Bartnicki-Garcia, S. 1974. Mycolaminarans: storage (1→3)-β-D-glucans from the cytoplasm of the fungus *phytophthora palmivora*. *Carbohydr. Res.*, 37:331–338.

Weyman, P. D., Beeri, K., Lefebvre, S. C., Rivera, J., McCarthy, J. K., Heuberger, A. L., Peers, G., Allen, A. E. & Dupont, C. L. 2015. Inactivation of *Phaeodactylum tricornutum* urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnol. J.*, 13:460–470.

Whelan, S. & Goldman, N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.*, 18:691–699.

Wolfenstetter, S., Wirsching, P., Dotzauer, D., Schneider, S. & Sauer, N. 2012. Routes to the tonoplast: the sorting of tonoplast transporters in *Arabidopsis* mesophyll protoplasts. *Plant Cell*, 24:215–232.

Zaslavskaia, L. A., Lippmeier, J. C., Kroth, P. G., Grossman, A. R. & Apt, K. E. 2000. Transformation of the diatom *Phaeodactylum tricornutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J. Phycol.*, 36:379–386.

Zhu, B. H., Shi, H. P., Yang, G. P., Lv, N. N., Yang, M. & Pan, K. H. 2015. Silencing UDP-glucose pyrophosphorylase gene in

Phaeodactylum tricornutum affects carbon allocation. *New Biotechnol.*, doi: 10.1016/j.nbt.2015.06.003.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Conserved domain alignment of TGS1, TGS2, TGS3, Kre6 and Skn1.

Figure S2. Analysis of the cellular localization of TGS1 and 2 in *P. tricornutum*.

Data S1. Manually curated ClustalW alignment used for the calculation of the tree.