Silica Morphogenesis by Alternative Processing of Silaffins in the Diatom Thalassiosira pseudonana

Nicole Poulsen and Nils Kröger‡

From the Lehrstuhl Biochemie I, Universitätstr. 31, Universität Regensburg, 93053 Regensburg, Germany

Received for publication, July 9, 2004, and in revised form, August 10, 2004
Published, JBC Papers in Press, August 10, 2004, DOI 10.1074/jbc.M407734200

For almost 200 years scientists have been fascinated by the ornate cell walls of the diatoms. These structures are made of amorphous silica, exhibiting species-specific, mostly porous patterns in the nano- to micrometer range. Recently, from the diatom *Cylindrotheca fusiformis* unusual phosphoproteins (termed silaffins) and long chain polyamines have been identified and implicated in biosilica formation. However, analysis of the role of silaffins in morphogenesis of species-specific silica structures has so far been hampered by the difficulty of obtaining structural data from these extremely complex proteins. In the present study, the five major silaffins from the diatom *Thalassiosira pseudonana* (tpSil1H, -1L, -2H, -2L, and -3) have been isolated, functionally analyzed, and structurally characterized, making use of the recently available genome data from this organism. Surprisingly, the silaffins of *T. pseudonana* and *C. fusiformis* share no sequence homology but are similar regarding amino acid composition and post-translational modifications. Silaffins tpSil1H and -2H are higher molecular mass isoforms of tpSil1L and -2L, respectively, generated in vivo by alternative processing of the same precursor polypeptides. Interestingly, only tpSil1H and -2H but not tpSil1L and -2L induce the formation of porous silica patterns in vitro, suggesting that the alternative processing event is an important step in morphogenesis of *T. pseudonana* biosilica.

During evolution many organisms (e.g. diatoms, sponges, radiolarians) have acquired the ability to use the ubiquitous monosilicic acid Si(OH)₄ for the formation of species-specifically structured, silica-based exo- or endoskeletons (1). This interesting biomineralization phenomenon is mediated by cellular organic (macro-) molecules that accelerate silicic acid polycondensation and control morphogenesis of the forming silica (2). Diatoms are an extremely large group (>10,000 species) of unicellular eukaryotic algae that play a major role in biological silica cycling. Within the last few years diatom biosilica-associated proteins (termed silaffins) and long chain polyamines (LCPA) have been identified and hypothesized to represent key components of the diatom biosilica-forming machinery. Silaffins and LCPA exhibit the remarkable ability to induce rapid silica deposition in vitro and to control the nanostructure of the forming silica (3). Therefore, unraveling the correlations between chemical structures, physical properties, and silica-forming activities of silaffins and LCPA will be important for understanding the molecular mechanism of species-specific biosilica nanopatterning. So far, silaffins have only been characterized from the diatom *Cylindrotheca fusiformis*. They are highly modified proteins/peptides rich in hydroxy amino acids (serine, threonine, hydroxyproline) and lysine residues. Silaffins natSil1A and -1B are O-phosphorylated at numerous sites and contain polyamine-modified lysine residues, features that enable these peptides to rapidly form silica nanospheres in vitro (4). Recently, natSil2, an acidic, glycosylated phosphoprotein of 40 kDa, has been characterized which represents a different type of silaffin. This protein lacks inherent silica-forming activity, yet it is able to regulate the silica forming activities of natSil1A and LCPA and strongly influences the nanostructure of *in vitro* formed silica (5). Unfortunately, because of the extremely high number of modified amino acids in natSil2, peptide mapping did not yield sufficient sequence information to clone the corresponding gene, and thus, the primary structures of regulatory silaffins have so far remained unknown.

It has been suggested that morphogenesis of nanopatterned silica in diatoms generally requires two components: first, molecules like LCPA or natSil1A that accelerate silica formation, and second, natSil2-like molecules that exhibit regulatory properties (5). However, because the biosilica architecture of *C. fusiformis* is rather unusual, being composed mainly of long, non-porous bands (Fig. 1, A and C), it has been unclear if general conclusions about the mechanism of diatom biosilica morphogenesis can be drawn from the properties of its silica-forming components. In the present study we have investigated the silaffins from the diatom *Thalassiosira pseudonana*, a species exhibiting porous biosilica nanopatterns (Fig. 1, B and D). It is demonstrated that the major silaffins of this organism represent natSil2-like proteins, with distinct capabilities of influencing silica morphogenesis by interactions with LCPA. In contrast, natSil1-like proteins appear to be absent in *T. pseudonana*. The complete genome sequence of *T. pseudonana* allowed identification of the silaffin genes from scarce peptide sequence information. Isolation of the corresponding cDNAs and biochemical analysis of the proteins enabled unprecedented insight into the chemical structures of regulatory silaffins.

MATERIALS AND METHODS

Culture Conditions—*T. pseudonana* clone CCMP1335 was grown in an artificial seawater medium according to the North East Pacific

**This work was supported by Deutsche Forschungsgemeinschaft Grant SFB 521-A2 and the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.**

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY706749, AY706750, and AY706751.

† To whom correspondence should be addressed. Tel.: 49-941-943-2837; Fax: 49-941-943-2866; E-mail: nils.kroeger@vkl.uni-regensburg.de.

‡ The abbreviations used are: LCPA, long chain polyamines; HPLC, high pressure liquid chromatography; PTC, phenylthiocarbamyl;
Silica Morphogenesis by T. pseudonana Silaffins

Culture Collection (ocgy.ubc.ca/projects/nepcc/media.htm) supplemented with 100 μg/ml penicillin and 100 μg/ml streptomycin at 18 °C under constant light at 5,000–10,000 lux.

Isolation of Silaffins and LCPA—Isolation of T. pseudonana biosilica was performed as described for another diatom species (7) with the addition of acetic acid to the wash fraction. After washing the column with 50 mM ammonium acetate and 0.5 mM ammonia, LCPA were eluted with 2 M NaCl in pH 10 buffer (100 mM ammonia, LCPA were eluted with 2 M NaCl in pH 10 buffer) and stored frozen at –85 °C for 12 h. The material was separated by reverse phase HPLC on a C18 column (Sephasil C18, 5 μm) and fractions containing seemingly homogenous components of 20 kDa were pooled, exhaustively dialyzed (cut off 7 kDa) against 10 mM ammonium acetate. After lyophilization and subjecting to fractionation using a Superdex200 HiLoad 16/60 column (Amersham Biosciences; running buffer 250 mM ammonium acetate, 0.25 mM/million flow rate). Fractions were analyzed by Tricine-SDS-PAGE (8) and staining with Stains All (SIGMA). Fractions eluting between 45 and 60 min (containing tpSil1/2L) were combined (Peak B). Peak A and Peak B were concentrated by ultrafiltration (cut off 10 kDa) and loaded separately onto a Mono Q HR-5/5 column (Amersham Biosciences) equilibrated with 50 mM Tris-HCl, pH 6.4. Elution was performed at a flow rate of 0.5 ml/min by linearly increasing the NaCl concentration to 2 M in 1 h. In runs with material from Peak A, tpSil3 were eluted between 22.5 and 28.5 min and tpSil1/2H eluted between 39 and 48.5 min. In runs with material from Peak B, tpSil1/2L eluted between 26.5 and 43.5 min. Fractions containing individual silaffins were analyzed by Tricine-SDS-PAGE and staining with Stains All (SIGMA). Fractions eluting between 45 and 60 min (containing tpSil1/2L) were combined (Peak B). Peak A and Peak B were concentrated by ultrafiltration (cut off 10 kDa) and loaded separately onto a Mono Q HR-5/5 column (Amersham Biosciences) equilibrated with 50 mM Tris-HCl, pH 6.4. Elution was performed at a flow rate of 0.5 ml/min by linearly increasing the NaCl concentration to 2 M in 1 h. In runs with material from Peak A, tpSil3 were eluted between 22.5 and 28.5 min and tpSil1/2H eluted between 39 and 48.5 min. In runs with material from Peak B, tpSil1/2L eluted between 26.5 and 43.5 min. Fractions containing individual silaffins were pooled, exhaustively dialyzed (cut off 7 kDa) against 10 mM ammonium acetate, and lyophilized. The dry residue was dissolved in water, adjusted to pH 5.5 with acetic acid/NaOH, and stored frozen at –20 °C until use.

Chemical Analysis of Silaffins—Phosphate (9), sulfate (10), and carbohydrate (11) contents were determined according standard protocols. Carbohydrate analysis was performed as described (5). Silica Formation in Vitro—Silica formation assays were performed in 100 mM silicic acid, 50 mM sodium acetate pH 5.5 as described (5).

Scanning Electron Microscopy—Samples were washed with water by repetitive centrifugation and resuspension, mounted onto a copper grid, and air dried. Images were obtained on a Leo 1530 field emission scanning electron microscope (Oberkochen, Germany).

RESULTS

Isolation of Silaffins and LCPA from T. pseudonana

A characteristic feature of C. fusiformis silaffins is their tight association with the cell wall, and therefore, they are only solubilized by completely dissolving the silica (4, 14). In search of silaffins in T. pseudonana, SDS/EDTA-extracted biosilica was dissolved in an ammonium fluoride solution, pH 4, and the extract was fractionated by ion exchange and gel permeation chromatography (Fig. 2, A and B). This procedure led to three fractions containing seemingly homogenous components of 20 and 85 kDa (Fig. 2C), which represent the major silaffins of...
concentrations had to be raised to /H11022 purified by chromatography on Mono Q (not shown) yielding fraction 3. T. pseudonana extract from as compared with LCPA from other diatom species, which may propyleneimine units attached to a putrescine residue (Fig. 3).

also contained LCPA ranging in molecular mass from 501.8 to as low as 20 able to precipitate silica already at tpSil1/2H concentrations that of tpSil3 because tpSil1/2H-LCPA assemblies were un-

tion (5). The inhibiting effect of tpSil1/2H was stronger than that of tpSil3-LCPA assemblies produced porous sheets of silica. The pores were irregularly arranged and exhibited non-uniform shapes, and their sizes ranged from 20 to 200 nm (Fig. 5B). Mixtures of tpSil3 and LCPA were able to simultaneously produce two very different silica structures; that is plates of densely packed, extremely small silica particles and large, polydisperse silica spheres ranging from 900 nm to 4.2 μm in diameter (Fig. 5, C and D). When tpSil3 was applied in concentrations up to 40 μm the silica plates displayed a relatively constant thickness of 500 nm, and the length of plate edges was on the order of tens of micrometers. The proportion of silica spheres in the precipitate decreased with increasing silaffin concentration and ceased at around 40 μm tpSil1/2L. At higher silaffin concentrations, when silica deposition was strongly reduced, very thin silica plates were produced (<50 nm thickness) that were easily damaged by the electron beam during scanning electron microscopy.

Previously, it has been demonstrated that mixtures of multivalent inorganic anions (e.g. phosphate) and LCPA mediate the formation of silica spheres with diameters ranging from 50 to 900 nm (7, 16). This is also the case with T. pseudonana LCPA. Therefore, morphogenesis of the spherical silica particles by tpSil1/2L-LCPA mixtures appears to be determined by the properties of LCPA, with tpSil1/2L playing a rather minor role, merely providing a high density of negative charges, allowing the formation of extremely large silica spheres at low polyanion concentration. In contrast, tpSil1/2H and tpSil3 are able to override the influence of LCPA on silica morphogenesis, inducing the formation of flat silica structures. A similar effect has previously been observed using mixtures of the polycationic R5 peptide (19 amino acids) and inorganic phosphate. In this system silica structures shifted from spherical to sheet-like, if high concentrations of polyhydroxy compounds (60% glycerol or 60% sucrose) were added (18). The molecular mechanism of this effect is unclear, yet it is in accordance with the fact that

T. pseudonana. Further analysis revealed that only fraction 1 contained a pure silaffin, which was termed tpSil3, whereas fractions 2 and 3 each represented mixtures of two highly homologous silaffins. For reasons that will be outlined below, the silaffins of fraction 3 were termed tpSil1L and tpSil2L (collectively tpSil1/2L; L denotes “low molecular mass isoform”), and the silaffins of fraction 2 were termed tpSil1H and tpSil2H (collectively tpSil1/2H; H denotes “high molecular mass isoform”). The T. pseudonana silaffins are highly acidic phosphoproteins that are glycosylated and sulfated (Table 1), thus resembling natSil2 from C. fusiformis (5).

The ammonium fluoride extract of T. pseudonana biosilica also contained LCPA ranging in molecular mass from 501.8 to 741.4 Da, which corresponds to methylated derivatives of 6–9 propyleneimine units attached to a putrescine residue (Fig. 3). The T. pseudonana LCPA are relatively short polyamine chains as compared with LCPA from other diatom species, which may carry up to 20 propyleneimine units (7, 15).

Silica Formation by Silaffin-LCPA Assemblies

The silica-forming properties of silaffins and LCPA from T. pseudonana were investigated in vitro using a freshly prepared silicic acid solution buffered to pH 5.5 by sodium acetate. In this system neither LCPA nor any of the three T. pseudonana silaffins alone was able to induce silica formation, yet mixtures of LCPA and silaffins exhibit silica-forming activity. This is consistent with previous observations demonstrating that supramolecular assemblies of polyanines and polyvalent anions lead to rapid silica precipitation (5, 16, 17). Interestingly, when the silica precipitation activities were analyzed as a function of the silaffin concentration in the presence of a constant amount of LCPA, the result was markedly different for each silaffin-LCPA assembly (Fig. 4). The activity of tpSil1/2L-LCPA assemblies steadily increased with increasing silaffin concentrations, reaching a plateau value (325 ± 19 mM silica at ≅60 μM tpSil1/2L). In contrast, tpSil1/2H and tpSil3 exhibited both activating and inhibiting effects, and thus, like silaffin natSil2 from C. fusiformis, represent regulators of silica formation (5). The inhibiting effect of tpSil1/2H was stronger than that of tpSil3 because tpSil1/2H-LCPA assemblies were unable to precipitate silica already at tpSil1/2H concentrations as low as 20 μM, whereas in the tpSil3-LCPA system tpSil3 concentrations had to be raised to >80 μM to completely inhibit silica precipitation. Furthermore, the T. pseudonana silaffins exhibited clearly distinct effects on silica morphogenesis (Fig. 5). Mixtures of tpSil1/2L and LCPA directed the formation of spherical silica particles at all silaffin concentrations (Fig. 5A). The diameters of the spheres increased with increasing silaffin concentrations yielding 230 ± 25 nm at 10 μM tpSil1/2L and up to 2.6 ± 1.7 μm at 100 μM tpSil1/2L. In contrast, tpSil1/2H-LCPA mixtures produced porous sheets of silica. The pores were irregularly arranged and exhibited non-uniform shapes, and their sizes ranged from 20 to 200 nm (Fig. 5B). Mixtures of tpSil3 and LCPA were able to simultaneously produce two very different silica structures; that is plates of densely packed, extremely small silica particles and large, polydisperse silica spheres ranging from 900 nm to 4.2 μm in diameter (Fig. 5, C and D). When tpSil3 was applied in concentrations up to 40 μm the silica plates displayed a relatively constant thickness of 500 nm, and the length of plate edges was on the order of tens of micrometers. The proportion of silica spheres in the precipitate decreased with increasing silaffin concentration and ceased at around 40 μM tpSil1/2L. At higher silaffin concentrations, when silica deposition was strongly reduced, very thin silica plates were produced (<50 nm thickness) that were easily damaged by the electron beam during scanning electron microscopy.

Previously, it has been demonstrated that mixtures of multivalent inorganic anions (e.g. phosphate) and LCPA mediate the formation of silica spheres with diameters ranging from 50 to 900 nm (7, 16). This is also the case with T. pseudonana LCPA. Therefore, morphogenesis of the spherical silica particles by tpSil1/2L-LCPA mixtures appears to be determined by the properties of LCPA, with tpSil1/2L playing a rather minor role, merely providing a high density of negative charges, allowing the formation of extremely large silica spheres at low polyanion concentration. In contrast, tpSil1/2H and tpSil3 are able to override the influence of LCPA on silica morphogenesis, inducing the formation of flat silica structures. A similar effect has previously been observed using mixtures of the polycationic R5 peptide (19 amino acids) and inorganic phosphate. In this system silica structures shifted from spherical to sheet-like, if high concentrations of polyhydroxy compounds (60% glycerol or 60% sucrose) were added (18). The molecular mechanism of this effect is unclear, yet it is in accordance with the fact that

3 N. Poulsen and N. Kröger, unpublished observation.

Fig. 2. Purification and SDS-PAGE analysis of silaffins. A, gel-permeation chromatography (Superdex200) of the ammonium fluoride extract from T. pseudonana biosilica. B, anion exchange chromatography (Mono Q) of the material from peak A. The material from peak B was also purified by chromatography on Mono Q (not shown) yielding fraction 3. C, silaffins were analyzed before (−) or after (+) treatment with anhydrous HF (90 min, 0 °C). Lanes 3 and 5 (from the left) were stained with Stains All; the other lanes were stained with Coomassie Blue. Fr., fraction.
Silica Morphogenesis by *T. pseudonana* Silaffins

Chemical characteristics of *T. pseudonana* silaffins

| Carbohydrate/protein (µg/hmnl) | tpSil1/2L | tpSil1/2H | tpSil3 |
|--------------------------------|-----------|-----------|--------|
| Glc                            | 6         | 5         | 1      |
| Gal                            | 7         | 4         | 9      |
| Glc                            | 1         | 1         | 3      |
| Xyl                            | 5         | 3         | 3      |
| Sulfate/protein (mol/mol)       | 50        | 35        | 50     |
| Phosphate/protein (mol/mol)     | 40        | 440       | 110    |
| pI                              | =3        | 3-4.5     | =3     |

Table 1

Silica Concentrations were estimated according to GlcUA, glucuronic acid; Rhm, rhamnose. The number next to each monosaccharide indicates its relative abundance. In addition to the monosaccharides listed, each silaffin contains at least two unidentified monosaccharides. Silaffin concentrations were estimated according to complete acid hydrolysis (6 M HCl, 10 °C, 24 h). The isoelectric points (pI) were determined by two-dimensional PAGE.

Fig. 3. Schematic chemical structure of *T. pseudonana* LCPA.

Fig. 4. Silica-forming activities of silaffin-LCPA mixtures. The concentration of silica produced after 10 min of reaction time was measured at varying concentrations of tpSil1/2H (triangles), tpSil1/2L (dots), or tpSil3 (squares) in the presence of a constant concentration of *T. pseudonana* LCPA (0.75 µg/µl). Each data point represents the average value of three independent measurements.

Fig. 5. Silica morphogenesis by silaffin-LCPA mixtures (scanning electron microscopy analysis). Experiments were performed in the presence of a constant concentration of *T. pseudonana* LCPA (0.75 µg/µl). A, 20 µg tpSil1/2L (bar = 1 µm); B, 3.5 µg tpSil1/2H (bar = 0.5 µm); C, 20 µg tpSil3 (bar = 10 µm); D, 30 µg tpSil3 (bar = 1 µm).

Structural Characterization of Silaffins

The Polypeptide Sequences of tpSil3, tpSil1L, and tpSil2L—It is reasonable to assume that the distinct influence of silaffins on silica formation is a consequence of differences in their molecular architectures. To elucidate this correlation the complex chemical structures of silaffins need to be determined, and this difficult task requires knowledge of their polypeptide sequences. So far, the gene *sil1* from *C. fusiformis*, which encodes natSil1A and natSil1B peptides (14), has been the only known silaffin gene. Because of the extremely high number of modified amino acids, hardly any sequence information could be obtained from other silaffins, thus preventing cloning of the corresponding genes by PCR and DNA library screening (5). However, in the case of *T. pseudonana* the situation is greatly simplified by the availability of the complete genome sequence. To obtain amino acid sequence information tpSil1/2L, tpSil1/2H and tpSil3 were treated with anhydrous HF to remove all carbohydrate, phosphate, and sulfate moieties and then subjected to N-terminal amino acid sequencing. The sequence obtained from HF-treated tpSil3 (EGHGDHISIM) perfectly matched a gene model in the *T. pseudonana* genome database. Using the genome sequence information, the complete cDNA sequence (termed *tpSil3*) was obtained by overlapping 5'- and 3'-rapid amplification of cDNA ends PCR. The corresponding polypeptide tpSil3p (p denotes “precursor”) contains an N-terminal signal peptide (amino acids 1–17) for cotranslational import into the endoplasmic reticulum (19) followed by a nine-amino-acid pro-peptide preceding the N-terminal sequence of tpSil3 (Fig. 6A). The predicted molecular mass of the mature polypeptide is 21.2 kDa, which is considerably smaller than the 35-kDa apparent molecular mass of HF-treated tpSil3 (see Fig. 2). This difference results from the presence of HF-resistant posttranslational modifications in tpSil3 (see below). The polypeptide backbone exhibits no repetitive sequence domains, and the three most abundant amino acids, serine (19.0%), alanine (16.5%), and lysine (16.1%), are evenly distributed within the polypeptide chain.

HF treatment of tpSil1/2L resulted in two bands on SDS-PAGE (see Fig. 2) that exhibited similar N-terminal sequences (upper band, LPGLXEMXXISXEDHYFFG; lower band,
Precursor polypeptides of *T. pseudonana* silaffins. Serine, proline, and threonine residues are depicted in blue. Lysine, arginine, and histidine residues are shown in red. Signal peptide sequences are shown in *black italics*. Amino acid sequences obtained from N-terminal sequencing of tryptic peptides are underlined. A, *tpSil3p* represents the precursor polypeptide for *tpSil3*. The N-terminal amino acid of *tpSil3* is marked by a red arrowhead. The pro-peptide sequence located between the signal peptide and the N terminus of *tpSil3* is depicted in green. B, alternative processing of *tpSil1p* gives rise to *tpSil1L* and *tpSil1H*, and alternative processing of *tpSil2p* yields *tpSil2L* and *tpSil2H*. The N- and C-terminal domains of *tpSil1* and *tpSil2* are denoted in the *right margin*. Identical amino acids in *tpSil1* and *tpSil2* are denoted by *black asterisks* in the *tpSil2* sequence. The N-terminal amino acids of *tpSil1L* and *tpSil2L* are marked by *red arrowheads*. The N-terminal amino acids of the 35-kDa fragments derived from the tryptic digest of *tpSil1H* are marked by *diamonds*.
23% proline. In reverse phase-HPLC of the PTC-dervatized amino acids a non-standard amino acid derivative was detected that eluted slightly earlier than PTC-hydroxyproline and represented 18% of the total amino acids, thus accounting for the missing amount of proline. The material from this peak has a molecular mass of 283 Da, which perfectly matches the mass of the PTC-dihydroxyproline. Indeed, in mass spectrometric analysis the 283-Da component exhibited the same fragmentation pattern as authentic PTC-dervatized 1,2,3-trans-3,4-cis-dihydroxyproline\(^2\) (20, 21). Dihydroxyproline is a very rare amino acid in biology (22), yet it was identified 35 years ago from acid hydrolysates of diatom biosilica (23). The silaffins tpSil1H and -2H are so far the only diatom proteins shown to contain this unusual amino acid.

**Modified Lysine Residues in \(T.\) pseudonana Silaffins**—A characteristic feature of \(C.\) fusiformis silaffins is the presence of alkylated lysines, which were hypothesized to be essential for supramolecular self-assembly and the interaction with silicic acid molecules (4, 5, 14). To search for modified lysines, \(T.\) pseudonana silaffins were separately subjected to exhaustive acid hydrolysis (6 M HCl, 110 °C, 24 h), and for each component amino acid analysis was performed by reverse phase-HPLC of the PTC derivatives as well as by electrospray ionization-mass spectroscopy without prior derivatization. Analysis of PTC-dervatized amino acids indicated that >80% of lysines are modified in \(T.\) pseudonana silaffins, because less than 20% of the predicted lysines were detected. Applying electrospray ionization-mass spectroscopy in the positive ion mode, the hydrolysates of each silaffin was shown to contain two dominant peak groups exhibiting (m+H\(^+\)) masses of 275 Da, 289 Da, 303 Da, 317 Da (peak group 1) and 319 Da, 333 Da, 347 Da (peak group 2). The molecules constituting these peak groups represent lysine derivatives, as was demonstrated by the following experiment. From the tryptic digest of tpSil1/2L a peptide was isolated exhibiting the sequence FFGXSHXSLX SXATXTL with \(X\) representing unidentified amino acid residues, and \(X’\) representing \(N,N\)-dimethyllysine, which was identified by its characteristic peak in the sequencing chromatogram (14). The peptide sequence matched amino acids 403–419 in tpSil2p (see Fig. 6), therefore demonstrating that the peptide was generated by the chymotryptic side activity of trypsin and that all four unidentified amino acid residues represented lysine derivatives. The mass spectrum of the hydrolyzed tryptic peptide contained \(N,N\)-dimethyllysine and the molecules of peak groups 1 and 2, which account for the unidentified lysine derivatives. Within each peak group, neighboring peaks differ by 14 Da in mass, suggesting that they represent methylation isomers of the same lysine derivative. These modified lysines appear to be highly positively charged, since they bind more strongly to cation exchange resins than lysine.\(^3\)

The chemical structures of the lysine derivatives are currently under investigation.

**DISCUSSION**

The analysis of the major biosilica-associated components has started to reveal the molecular principles that form the basis of silica morphogenesis in diatoms. This process appears to depend on supramolecular assemblies of LCPS and silaffins. So far LCPS from five diatom species have been characterized demonstrating that the LCPS structure is highly conserved, exhibiting variations solely regarding chain length and degree of methylation (see Fig. 3) (7, 15). In contrast, the chemical structures of silaffins are surprisingly diverse. None of the \(T.\) pseudonana silaffins identified in the present study exhibits sequence homology to \(C.\) fusiformis silaffins, yet certain structural features are clearly conserved among all known silaffins; they are rich in hydroxyamino acids as well as lysines, and both types of amino acid are posttranslationally modified. This suggests that conservation of the amino acid composition and the attached posttranslational modifications rather than conservation of the amino acid sequence is essential for silaffin function. To date all silaffins characterized are highly phosphorylated, and the regulatory silaffins (natSil2, tpSil1/2H, tpSil3), in addition, are glycosylated and sulfated. Although \(T.\) pseudonana silaffins lack the long chain polyamine-modified lysines that are characteristic for \(C.\) fusiformis silaffins, most of the lysines carry shorter, yet uncharacterized modifications. The functional implications of this difference in lysine modification are not yet clear but may relate to the fact that \(T.\) pseudonana silaffins are unable to form silica in vitro unless LCPS are added. This indicates that LCPS are essential components of the \(T.\) pseudonana biosilica-forming machinery. However, the morphogenesis of silica appears to be mainly determined by the silaffins. This assumption is strongly supported by the demonstration in vitro that tpSil1/2H and tpSil3 prevent LCPS from forming large, biologically irrelevant silica spheres and guide morphogenesis of flat silica plates and porous silica that rather resemble diatom biosilica structures. Therefore, alternative processing in vivo of the tpSil1p and tpSil2p polypeptides generating tpSil1/2H rather than tpSil1/2L proteins may be an important step to enable morphogenesis of the porous biosilica patterns in \(T.\) pseudonana.

Based on the data about the chemical structure of silaffins, the distinct influences on silica formation of tpSil1/2L proteins as compared with tpSil1/2H proteins may be explained as follows. The N-terminal domains of tpSil1H and -2H contain extremely high densities of hydroxyamino acids, which are likely to be glycosylated rather than phosphorylated, thus explaining the approximately 10 times higher content of carbohydrate in tpSil1/2H as compared with tpSil1/2L (see Table 1). As was previously demonstrated, the carbohydrate and sulfate moieties in the \(C.\) fusiformis silaffin natSil2 act as inhibitors in polyamine-dependent silica formation, whereas the protein-bound phosphate groups have an activating effect (5). Therefore, the clustering of carbohydrate and sulfate moieties in the \(T.\) pseudonana domains of tpSil1/2H proteins likely accounts for their strong inhibiting effect already at low protein concentrations (see Fig. 4). Removal of the N-terminal domains generates tpSil1/2L proteins, which exert a solely stimulating effect on LCPS-dependent silica formation, presumably because the influence of the phosphate residues outperforms the effect of the residual carbohydrate and sulfate moieties in these molecules.

Silicic acid polycondensation and silica deposition is generally promoted by polyamines (6, 24–25), indicating that this process requires cooperative interactions between amino groups and Si-OH groups, which may occur via multiple hydrogen bonds. Interestingly, tpSil1/2H, the most highly glycosylated \(T.\) pseudonana silaffin (Table 1) exerts the strongest inhibiting effect on LCPS-dependent silica formation (Fig. 4). This suggests that inhibition may be exerted by hydrogen bonding of LCPS molecules to the numerous hydroxyl groups of the silaffin, thereby impeding the interaction between LCPS and silicic acid molecules. The dihydroxyproline residues in tpSil1/2H may aid in this process directly by binding LCPS via their vicinal hydroxyl groups or indirectly by providing additional glycosylation sites. Further studies are required to analyze the mode of interaction between regulatory silaffins and LCPS.

The cloning of the silaffin genes from \(T.\) pseudonana will now enable studies on the influence of silaffins on silica morphogenesis in vivo. It is possible to express \(T.\) pseudonana silaffins in the \(C.\) fusiformis because a genetic transformation system has
been established for this diatom (13). The resulting transformants are expected to exhibit altered biosilica patterns due to the presence of foreign silaffins with different silica formation activities. If so, it may even become possible to tailor the biosilica patterns by introducing into *C. fusiformis* different combinations or mutated variants of *T. pseudonana* silaffins.

Acknowledgments—We are grateful to Ginger Armbrust (University of Washington) for providing the *T. pseudonana* strain. We thank R. Deutzmann and E. Hochmuth for protein sequencing and mass spectrometry, G. Lehmann for technical assistance, and B. Gess and A. Kurz for two-dimensional PAGE analysis. We thank R. Deutzmann also for critically reading the manuscript. We are indebted to C. Taylor (Massey University, Palmerston North, New Zealand) for generously providing synthetic L-2,3-trans-3,4-cis-dihydroxyproline.

REFERENCES

1. Simpson, T. L. & Volcani, B. E. (1981) *Silicon and Siliceous Structures in Biological Systems*, Springer New York Inc., New York
2. Bauerlein, E. (2000) *Biomineralization: From Biology to Biotechnology and Medical Application*, Wiley-VCH, Weinheim, Germany
3. Kroger N. & Sumper, M. (2004) *J. Mater. Chem.* 14, 2059–2065
4. Kroger, N., Lorenz, S., Brunner, E. & Sumper, M. (2002) *Science* 298, 584–586
5. Poulsen, N., Sumper, M. & Kroger, N. (2003) *Proc. Natl. Acad. Sci. U. S. A.* 100, 12075–12080
6. Menzel, H., Horstmann, S., Behrens, P., Bärnreuther, P., Krueger, I. & Jahns, M. (2003) *Chem. Commun.* 24, 2994–2995
7. Kroger, N., Deutzmann, R., Bergsdorf, C. & Sumper, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* 97, 14133–14138
8. Schägger, H. & von Jagow, G. (1987) *Anal. Biochem.* 166, 368–379
9. Buss, J. E. & Stull, J. T. (1983) *Methods Enzymol.* 99, 7–14
10. Terho, T. T. & Hartiša, K. (1971) *Anal. Biochem.* 41, 471–476
11. Dubios, M., Gilles, K. A., Hamilton, J. K. & Reber, P. A. (1956)*Anal. Chem.* 28, 350–356
12. Lammli, U. K. (1970) *Nature* 227, 680–685
13. Fischer, H., Robl, Sumper, M. & Kroger, N. (1999) *J. Phycol.* 35, 113–120
14. Kroger, N., Deutzmann, R. & Sumper, M. (1999) *Science* 286, 1129–1132
15. Sumper, M. (2002) *Science* 295, 2430–2433
16. Sumper, M., Lorenz, S. & Brunner, E. (2003) *Angew. Chem. Int. Ed.* 42, 5192–5195
17. Brunner, E., Lutz, L. & Sumper, M. (2004) *Phys. Chem. Chem. Phys.* 6, 854–857
18. Naik, R. R., Whitlock, P. W., Rodriguez, F., Brott, L. L., Glawe, D. D., Clarson, S. J. & Stone, M. O. (2004) *Biomacromolecules* 5, 261–265
19. Nielsen, H., Engelbrecht, H., Brunak, S. & von Heijne, G. (1997) *Protein Eng.* 10, 1–6
20. Weir, C. A. & Taylor, C. M. (1999) *J. Org. Chem.* 64, 1554–1558
21. Fleet, G. W. & Son, J. C. (1988) *Tetrahedron* 44, 2637–2647
22. El-Ashry E.-S. H. & El Nemr, A. (2003) *J. Org. Chem.* 68, 10003–10009
23. Nakajima, N. & Volcani, B. E. (1969) *Science* 164, 1400–1402
24. Mitzu, T., Nagase, H., Fujiwara, N. & Ogoshi, H. (1998) *Bull. Chem. Soc. Jpn.* 71, 2017–2022
25. Patwardhan, S. V. & Clarson, S. J. (2001) *J. Inorg. Organomet. Polym.* 13, 193–203