Silica-precipitating Peptides from Diatoms

THE CHEMICAL STRUCTURE OF SILAFFIN-1A FROM CYLINDROTHECA FUSIFORMIS

Nils Kröger, Rainer Deutzmann, and Manfred Sumper

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

Two silica-precipitating peptides, silaffin-1A1 and -1A2, both encoded by the sil1 gene from the diatom Cylindrotheca fusiformis, were extracted from cell walls and purified to homogeneity. The chemical structures were determined by protein chemical methods combined with mass spectrometry. Silaffin-1A1 and -1A2 consist of 15 and 18 amino acid residues, respectively. Each peptide contains a total of four lysine residues, which are all clustered in two pairs in which the e-amino group of the first residue is linked to a linear polylamine consisting of 5 to 11 N-methylated propylamine units, whereas the second lysine is converted to ε-N,N-dimethyllysine. Silaffin-1A1 contains only a single lysine pair exhibiting the same structural features. One of the two remaining lysine residues was identified as ε-N,N,N-trimethyl-δ-hydroxylysine, a lysine derivative containing a quaternary ammonium group. The fourth lysine residue again is linked to a long-chain polylamine. Silaffin-1A1 is the first peptide shown to contain ε-N,N,N-trimethyl-δ-hydroxylysine. In vitro, both peptides precipitate silica nanospheres within seconds when added to a monosilicic acid solution.

Silicon oxide minerals, the main constituents of the earth's crust, are not exclusively formed by geological processes. In fact, hydrated silicon dioxide (silica), the second most abundant biogenic mineral (biomineral), is produced by a wide range of organisms including animals and higher plants (1). A large proportion of biogenic silica is formed by diatoms (2), which are unicellular algae that are ubiquitously present in marine and freshwater habitats (3). The main attribute of a diatom cell is its silica based cell wall. The intricate and ornate silicified cell walls of diatoms are one of the most outstanding examples of nanoscale-structured materials in nature. In the past, diatoms have been studied as model organisms to investigate the biochemical basis of biological silica formation (4–6). This has led to the discovery of silicic acid transporter proteins (7, 8) and chemical basis of biological silica formation (4–6). This has led to the discovery of silicic acid transporter proteins (7, 8) and putative components from different diatom species were identified that mediate the formation of silica nanospheres in vitro from a silicic acid solution (9, 10). These components are long-chain polylamines and polycationic polypeptides termed silaffins. The chemical structures of the polylamines have been completely elucidated. They are composed of linear chains of 8 to 20 N-methylated propylamine units that are attached to putrescine or a putrescine derivative (11). In contrast, there is only incomplete information about the chemical structure of the silaffins. Recently, a silaffin-encoding gene, termed sil1, has been cloned from the diatom Cylindrotheca fusiformis. The encoded polypeptide sil1p serves as a precursor molecule, which becomes proteolytically processed and post-translationally modified to produce the silica-precipitating peptides silaffin-1A and silaffin-1B. It has been demonstrated that silaffin-1A represents a mixture of peptide isoforms, and that their silica-precipitating activity depends on the presence of modified lysine residues (10). So far, two types of modified lysine residues (denoted Lysx and Lysy) have been characterized from the N-terminal octapeptide fragment SSKKKSGSY that is common to all silaffin-1A isoforms. Lysx represents a lysine residue that carries on its ε-amino group a linear polylamine consisting of 5 to 11 N-methylated propylamine units. Lysy has been shown to represent ε-N,N-dimethyllysine. However, no information was available for the chemical structures of the remaining parts of the silaffin-1A isoforms. In the present study we describe the silica-precipitating properties and complete the chemical structures of the peptides silaffin-1A1 and silaffin-1A2, which together account for all peptide isoforms of silaffin-1A.

EXPERIMENTAL PROCEDURES

Chemicals—Synthetic ε-N,N,N-trimethyl-δ-hydroxylysine (19) was a gift from B. E. Volcani (Scripps Institute for Oceanography, University of California, San Diego, La Jolla, CA).

Culture Conditions—Cylindrothea fusiformis was grown in artificial sea water medium as described previously (15).

Isolation of silaffin-1A and silaffin-1A2—The silaffin-1A fraction was isolated from purified cell walls of C. fusiformis as described previously (10) and subjected to high pressure liquid chromatography (HPLC) on a Sephasil C18 2.1/10 column using the SMART-System (Amersham Pharmacia Biotech). Elution of peptides was performed by increasing the concentration of buffer B from 0 to 25% in 35 min (buffer A: 0.1% trifluoroacetic acid in H2O; buffer B: 0.085% trifluoroacetic acid in acetonitrile). Fractions containing silaffin-1A1 and silaffin-1A2, respectively, were lyophilized and the residues were dissolved in H2O.

Digestion of Silaffin-1A with Chymotrypsin and Separation of Peptides—Silaffin-1A (180 μg) was dissolved in 100 μl of 50 mM TrisHCl, pH 8, and 9 μg of chymotrypsin (N2-p-tosyl-L-lysine chloromethyl ketone-treated; Sigma) was added. Incubation was at 37 °C for 15 h. The resulting chymotryptic peptides were separated by HPLC using the same conditions as described above.

Acid Hydrolysis and Hydrazinolysis—Complete degradation of silaf-
The ratio of peak areas of silaffin-1A1 and silaffin-1A2 in the HPLC chromatogram is about 5 to 1. Therefore, it is reasonable to assume that each of the repeats R3–R7 of sil1p contributes to silaffin-1A1 production (Fig. 1B). A comparison of the sil1p sequence (10) with the results from amino acid sequencing of isolated silaffins reveals the following structural features of silaffin-1A1 and silaffin-1A2. 1) All of the lysine residues present are post-translationally modified. Three types of modified lysine residues (designated Lys<sup>x</sup>, Lys<sup>y</sup>, and Lys<sup>z</sup>) can be distinguished in silaffin-1A1. Lys<sup>z</sup> produces a signal between Arg and Tyr in the chromatogram of the amino acid sequencer and was previously identified as ε-N,N-dimethyllysine (10). Lys<sup>x</sup> is a so far unidentified lysine derivative exhibiting a characteristic peak between Ala and Arg. Lys<sup>y</sup> denotes a lysine derivative that does not show up at all in automated amino acid sequencing. For lysine derivative Lys<sup>y</sup> at position 4, a long-chain polypeptide was previously shown to be attached to the ε-lysine group (10). 2) Neither the C-terminal amino acid sequence RRIL predicted by the gene sequence of repeat R2 nor the corresponding sequences KRRNL and KRRIL predicted from repeats R3 to R7 showed up in amino acid sequencing of silaffin-1A2 and silaffin-1A1, respectively. Therefore, these residues may either have been removed or post-translationally modified during processing of the silaffin precursor polypeptide.

**Complete Amino Acid Sequence of Silaffin-1A—**To further analyze the chemical structure, silaffin-1A<sub>1</sub> was digested with chymotrypsin, which generates the previously described N-terminal octapeptide SSK<sup>K</sup>SGYSYSQ/SY (see the Introduction) and the C-terminal fragment SGSK/GS (Fig. 2). In reversed phase C<sub>18</sub> HPLC, the N-terminal octapeptide separates into five fractions (Fig. 2), which differ in masses by multiples of 71 Da (Table II) due to heterogeneity with respect to the chain length of the polyamine modification (10). Surprisingly, the same mass differences were observed in the subfractions derived from the C-terminal peptide SGSK/GS (Table II), suggesting that this peptide also contains a lysine residue carrying the polyamine modification. Because this type of lysine derivative is not detectable by automated amino acid sequencing, we hypothesized that it may constitute the C terminus, and therefore the sequence of the C-terminal peptide may rather be represented by SGSK/GSK<sup>K</sup>. To investigate this, the material from fraction 3 (C-terminal peptide) and fraction 8 (N-terminal octapeptide), respectively, was subjected to hydrazinolysis, and the resulting products were analyzed by ESI-MS. Hydrazinolysis leads to the breakdown of the peptide backbone, and all amino acids residues originally placed within the polypeptide chain become converted to the corresponding hydrazides. Only
Silica-precipitating Peptides from Diatoms

The amino acid sequences of the C-terminal peptides (fractions 1–5) and the N-terminal peptides (fractions 6–10) as determined by complete amino acid sequencing are indicated.

**Table II**

| Peptide | Molecular mass | Δ |
|---------|----------------|---|
| 1       | 1135.3         |   |
| 2       | 1206.2         | 70.9 |
| 3       | 1277.3         | 71.1 |
| 4       | 1348.5         | 71.2 |
| 5       | 1419.7         | 71.2 |
| 6       | 1298.4         |   |
| 7       | 1369.6         | 71.2 |
| 8       | 1440.6         | 71.0 |
| 9       | 1511.7         | 71.1 |
| 10      | 1582.9         | 71.1 |

The C-terminal residue is released as free amino acid (18). As expected, hydrazinolysis of the material from fraction 8 (N-terminal octapeptide) generated a molecule of mass \((m + H)^+ = 729.8\) Da (Fig. 3A). This molecule corresponds to the hydrazide of a lysine derivative carrying eight methylated polyamine units, indicating the presence of a trimethylammonium group as well as an hydroxy group within this lysine derivative. Indeed, the obtained product ion spectrum (Fig. 3A) matches the spectrum obtained from authentic \(\epsilon\)-N,N,N-trimethyl-\(\delta\)-hydroxylysine (Fig. 3B). This lysine derivative had previously been found in acid hydrolysates obtained from total cell wall preparations of different diatom species including *C. fusiformis* (19).

**Methylation Pattern of the Long-chain Polyamine Modification**—Structural analysis of long-chain polypeptides by mass spectrometry revealed that their collision-induced fragmentation is caused exclusively by the cleavage of C–N bonds (11). According to this finding, the previously proposed structure of the polyamine moiety linked to lysine derivatives \(\text{Lys}^x\) in silaffin-1A1 (10) has to be reconsidered. A shift of two methyl groups within the polyamine chain enables the interpretation of all the fragment ions observed (10) by allowing C–N bond cleavages only. Therefore, a modified structural model is proposed for lysine derivative \(\text{Lys}^x\) (included in Fig. 5) in which the polyamine moiety is dimethylated at its terminal amino group, thus representing a methylation isomer of the previously proposed structure. The modified structural model was confirmed by reductive ethylation of lysines \(\text{Lys}^y\) and fragmentation analysis (by mass spectrometry) of the resulting ethylated derivatives. Reductive ethylation introduced exactly four ethyl groups into each \(\text{Lys}^y\) isoform, irrespective of chain-length variations in the polyamine moiety (data not shown).

The terminal amino group of the polyamine moiety is
dimethylated, because it is only the ε-amino group of the lysine core as well as the amino group of the very first propylamine unit that can be converted to an N-ethyl derivative (in addition to the α-amino group of the lysine core; Fig. 6B). This fact clearly indicates that both of these amino groups exist as secondary amines in the parent molecule, i.e. they are not methylated (these amino groups were assumed to be methylated in the previous structural model).

Silica Precipitation—It was previously shown that the silica precipitation activity of silaffin-1A peptides at pH values < 7 is dependent on the lysine modifications (10). Because the presence of ε-N,N,N-trimethyl-δ-hydroxylysine clearly distinguishes silaffin-1A1 from silaffin-1A2 (Fig. 5), it was investigated as to whether the two silaffin-1A isoforms have different pH-dependent silica-precipitating properties. In an in vitro assay, the amount of silica precipitated by silaffin-1A1 and silaffin-1A2, respectively, was found to be fairly constant at different pH values and almost identical for silaffin-1A1 (9.0–11.9 nmol of Si/nmol of peptide) and silaffin-1A2 (10.3–11.7 nmol of Si/nmol of peptide). Only at pH 5 did silaffin-1A1 show a slightly lower silica-precipitating activity as compared with silaffin-1A2 (Fig. 7A). The structures of the silica precipitates

![Silica Precipitation](image_url)

**FIG. 4.** Identification of ε-N,N,N-trimethyl-δ-hydroxylysine. Product ion spectra obtained by collision-induced fragmentation. A, product ions obtained from the (m + H)+ = 205.1 Da ion that is present in the amino acid hydrolysate of silaffin-1A1. B, product ion spectrum obtained from authentic ε-N,N,N-trimethyl-δ-hydroxylysine.

![Structural Analysis](image_url)

**FIG. 5.** Chemical structures of silaffin-1A1 and silaffin-1A2. The polypeptide backbones are depicted by the one-letter amino acid code. The chemical structures of the side chains of only the modified lysine residues are shown.

![Silica Precipitation](image_url)

**FIG. 6.** Structural analysis by ethylation of lysine derivative Lysε. Ethylated lysine residue Lysε of molecular mass (m + H)+ = 898.9 Da (molecular mass before ethylation was (m + H)+ = 786.9 Da, indicating the introduction of four ethyl groups) was isolated by the ion trap mode of the mass spectrometer and subjected to collision-induced fragmentation. A, product ion spectrum after fragmentation. B, proposed structure (schematic) of the (m + H)+ = 898.9 Da ion. Cleavage positions that lead to the observed fragment ion spectrum are depicted by arrows, and the corresponding molecular masses are indicated.

![Silica Precipitation](image_url)

**FIG. 7.** Silica precipitation by silaffin-1A1 and silaffin-1A2. A, pH dependence of peptide-induced silica precipitation. The solid line (measured values in squares) shows the result for silaffin-1A1, and the dotted line (measured values in triangles) represents the results for silaffin-1A2. B and C, scanning electron microscopic images of silica precipitates induced at pH 6.4 by 3 mg/ml silaffin-1A1 (B) and 3 mg/ml silaffin-1A2 (C).
Silica-precipitating Peptides from Diatoms

were analyzed by scanning electron microscopy. At all pH values silaffin-1A peptides induced the formation of spherical silica nanoparticles, but at the level of scanning electron microscopy resolution, no morphological difference was noted between the precipitates induced by silaffin-1A₁ (Fig. 7B) and silaffin-1A₂ (Fig. 7C).

DISCUSSION

The present study describes for the first time the complete chemical structures of silica-precipitating peptides found in cell walls of the diatom *C. fusiformis*. These are silaffin-1A₁ and silaffin-1A₂, which consist of 15 and 18 amino acid residues, respectively. Both peptides contain a total of four lysine residues, and all of these are targets for post-translational modifications. In silaffin-1A₂, the lysine residues are clustered in two pairs with the first residue being linked to a long-chain polyamine and the second lysine being converted to e-Ν,N,N-trimethyllysine. In silaffin-1A₁, the same type of modified lysine pair is present only within the N-terminal part of the peptide. The remaining two lysine residues in the C-terminal part are separated by two intercalated amino acids; this motif appears to alter the strategy of post-translational modification. The lysine residue at position 12 becomes modified to e-Ν,N,N-trimethyl-δ-hydroxylysine, and it is now the C-terminal lysine residue that carries a long-chain polyamine modification. Remarkably, more than 30 years ago, Nakajima and Volcani (19) isolated and characterized for the first time e-Ν,N,N-trimethyl-δ-hydroxylysine in acid hydrolysates of total cell wall preparations from a number of diatoms. However, the corresponding proteins in diatoms containing this special type of modification remained elusive. Silaffin-1A₁ is (to our knowledge) the first polypeptide found in nature containing the e-Ν,N,N-trimethyl-δ-hydroxylysine residue.

Despite the structural differences of silaffin-1A₁ and silaffin-1A₂, both polycationic peptides show almost identical silica-precipitating activities and promote the formation of silica nanospheres in vitro (see Fig. 7). This result suggests that the silica-precipitating activities of silaffin-1A₁ and silaffin-1A₂ are dependent mainly on the polyamine modification attached to lysine residues. This is consistent with the finding that long-chain polyamines attached to putrescine that were isolated from diatom cell walls are also able to precipitate silica nanospheres (11), whereas synthetic silaffins lacking the lysine modifications are unable to precipitate silica at pH < 7 in vitro (10). In this respect it is interesting to note that silica formation in diatoms takes place in an acidic, intracellular compartment (20), and thus the polyamine moieties of the silaffin-1A peptides appear to be essential to mediate silica precipitation under physiological conditions. The e-Ν,N,N-trimethyl-δ-hydroxylysine present in silaffin-1A₁ is a structural element that might influence the ultrastructure of the precipitating silica. Remarkably, quaternary ammonium compounds are used in the technical production of zeolites for patterning of silicate structures in the nanometer size range (21). Possibly, the e-Ν,N,N-trimethyl-δ-hydroxylysine residue exerts a similar function in biosilica formation.

The role of the polypeptide backbones in silaffin-1A₁-mediated silica formation is much less clear. Isolation of silaffins from diatom biosilica involves treatment with anhydrous hydrogen fluoride that converts silica to volatile silicon tetrafluoride. Although this treatment does not attack peptide bonds, it does however specifically cleave O-glycosidic bonds (22). Silaffins contain a large number of hydroxyamino acid residues, which may be targets for O-glycosylation. However, a completely different technique for the extraction of silaffins from biosilica is required to investigate this possibility.

Comparison of the silaffin-1A₁ and silaffin-1A₂ sequences with the sequences deduced from the *sil1* gene revealed that during maturation of the silaffins, the C-terminal tetrapeptides RRIL and RRNL, respectively, become cleaved off. This processing step completely removes all arginine residues that are originally present in the silaffin precursor polypeptide *sil1p* (see Fig. 1B). Remarkably, arginine is the biosynthetic precursor of putrescine (23), and the latter has been shown to serve as the attachment site for long-chain polyamines in *C. fusiformis* and other diatoms (11). Therefore, it is intriguing to speculate that *sil1p* is also the precursor for the putrescine-linked polyamines. After conversion of the arginine residues in the silaffin precursor to ornithine residues, the latter may become modified by the same enzymatic machinery that attaches propylamine units to the appropriate lysine residues in silaffins. Subsequently, silaffin peptides and putrescine-based polyamines could be produced simultaneously by proteolytic processing and decarboxylation of the polyamine-modified ornithine residues. If so, *sil1p* of *C. fusiformis* would give rise to two different sets of silica-precipitating molecular species.

Acknowledgments—We thank E. Hochmuth for expert technical assistance and LEO Elektronenmikroskopie (Oberkochen, Germany) for help with the scanning electron microscopy.

REFERENCES

1. Lowenstam, H. A. (1981) *Science* **211**, 1126
2. Trüger, P., Nelson, D. M., van Bennekom, A. J., DeMaster, D. J., Leynaert, A., and Queziner, B. (1995) *Science* **268**, 375–379
3. Round, F., Crawford, R., and Mann D. (1990) *The Diatoms*, Cambridge University Press, Cambridge
4. Volcani, B. E. (1978) in *Biochemistry of Silicon and Related Problems* (Bendz, G., and Lindquist, I., eds) pp. 177–204, Plenum Press, New York
5. Robinson, D. H., and Sullivan, C. W. (1987) *Trends. Biochem. Sci.* **12**, 151–154
6. Kroger, N., and Sumper, M. (2000) in *Biomineralization* (Bauerlein, E., ed) pp. 151–170, Wiley-VCH, Weinheim, Germany
7. Hildebrand, M., Volcani, B. E., Gassmann, W., and Schroeder, J. I. (1997) *Nature* **385**, 688–689
8. Hildebrand, M., Dahlin, K., Volcani, B. E. (1998) *Mol. Gen. Genet.* **260**, 480–486
9. Kroger, N., Lehmann, G., Rachel, R., and Sumper, M. (1997) *Eur. J. Biochem.* **239**, 259–264
10. Kroger, N., Deutzmann, R., and Sumper, M. (1999) *Science* **286**, 1129–1132
11. Kroger, N., Deutzmann, R., Bergdorff, C., and Sumper, M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14133–14138
12. Mann, S., and Ozin, G. (1996) *Nature* **382**, 313–318
13. Parkinson, J., and Gordon, R. (1999) *Trends Biotechnol.* **17**, 190–196
14. Cha, J. N., Stucky, G. D., Morse, D. E., and Deming, T. J. (2000) *Nature* **403**, 289–292
15. Kroger, N., Bergdorff, C., and Sumper, M. (1994) *EMBO J.* **13**, 4676–4683
16. Nachbau, E., and Leiseder, G. (1971) *Monatsh. Chem.* **102**, 1718–1723
17. Jentoft, N., and Dearborn, D. G. (1983) *Methods Enzymol.* **91**, 570
18. Narita, K., Matsuo, H., and Nakajima, T. (1975) in *Protein Sequence Determination* (Needleman, S., ed) pp. 30–103, Springer, Heidelberg
19. Nakajima, T., and Volcani, B. E. (1970) *Biochem. Biophys. Res. Commun.* **39**, 28–33
20. Vrielings, E. G., Gieskes, W. C. W., and Beelen, T. P. M. (1999) *J. Physiol.* **35**, 548–555
21. Burkett, S. L., and Davies, M. E. (1996) in *Comprehensive Supramolecular Chemistry* (Alberti, G., and Bein, T., eds) Vol. 7, pp. 465–483, Pergamon, Oxford
22. Mort, A. J., Lampert, D. T. A. (1977) *Anal. Biochem.* **82**, 289–309
23. Tabor, C. W., and Tabor, H. (1984) *Annu. Rev. Biochem.* **53**, 749–790
Silica-precipitating Peptides from Diatoms: THE CHEMICAL STRUCTURE OF SILAFFIN-1A FROM CYLINDROTHECA FUSIFORMIS
Nils Kröger, Rainer Deutzmann and Manfred Sumper

J. Biol. Chem. 2001, 276:26066-26070.
doi: 10.1074/jbc.M102093200 originally published online May 10, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M102093200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 19 references, 4 of which can be accessed free at http://www.jbc.org/content/276/28/26066.full.html#ref-list-1