Electronic Supporting Information (ESI)

for

Solid-state NMR spectroscopic studies of ^{13}C,^{15}N,^{29}Si-enriched biosilica from the marine diatom *Cyclotella cryptica*

Felicitas Kolbe,[1] Helena L. Ehren,[2] Simon Kohrs,[1] Daniel Butscher,[1] Lukas Reiß,[1#]

Marc Baldus,[2] Eike Brunner[1]

[1] Chair Bioanalytical Chemistry, Faculty of Chemistry and Food Chemistry, TU Dresden, 01062 Dresden, Germany

[2] NMR Spectroscopy, Bijvoet Center for Biomolecular Research Utrecht University, 3584 CH Utrecht, The Netherlands

[#] present address: Hochschule für Bildende Künste (HfBK) Dresden, Güntzstraße 34, 01307 Dresden, Germany

1. Additional solid-state NMR spectra of *C. cryptica*

Additional 1- and 2-dimensional NMR spectra

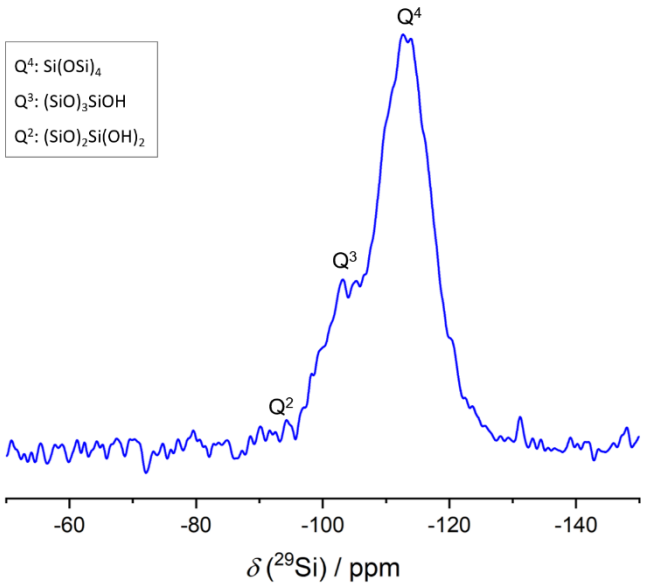

Figure S1. ^{29}Si MAS NMR spectrum (direct excitation) of ^{13}C,^{15}N,^{29}Si-enriched biosilica of *C. cryptica*. Recycle Delay: 1800 s.

Long relaxation times and recycle delays are required for quantitative detection of ^{29}Si NMR spectra.[1] Based on the long recycle delays applied during acquisition, quantitative evaluation of the spectra is possible here. The determined Q^4:Q^3 ratio of ca. 4:1 is typical for biosilica samples and shows the high condensation degree of the studied biosilica sample.

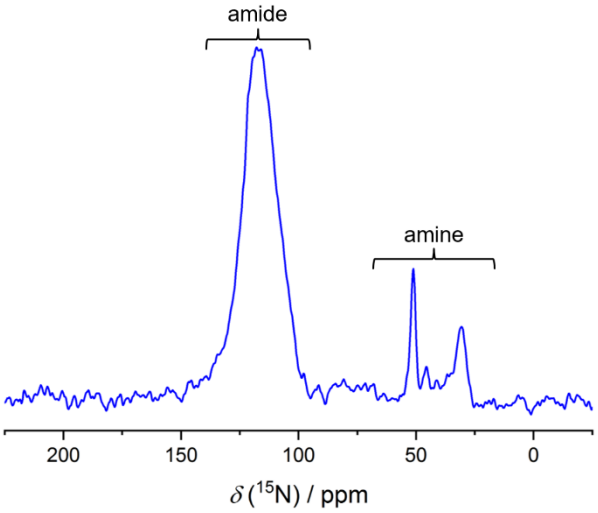

Figure S2. ^{15}N CP/MAS NMR spectrum of ^{13}C,^{15}N,^{29}Si-enriched biosilica of *C. cryptica*. CP contact time: 4 ms.

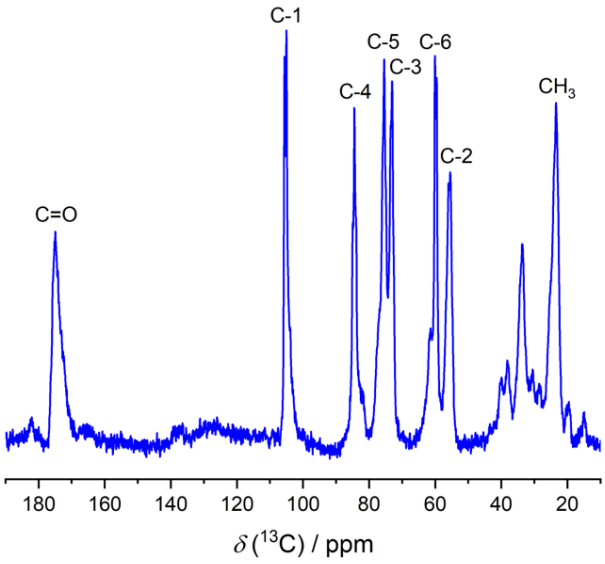

Figure S3. ^{13}C MAS NMR spectrum (direct excitation) of the NaOH-treated insoluble organic matrix of *C. cryptica*. The main signals can be assigned to different chitin conformations (cf. Table 1, main article).

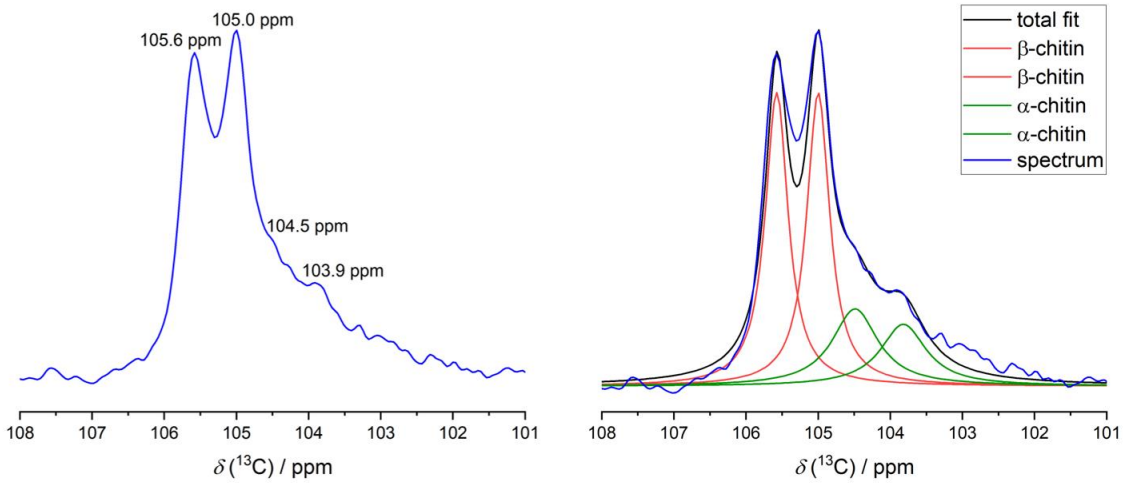

Figure S4. Zoom into the C-1 Signal at the ^{13}C MAS NMR spectrum (direct excitation) of the NaOH-treated insoluble organic matrix of *C. cryptica*. Right: Resulting fit of the C-1 signal.

A zoom into the characteristic C-1 signal confirms the presence of two different chitin conformations. The signal splits into two doublets due to the ^{13}C-^{13}C spin-spin coupling. The observed coupling constant is around 44 Hz for both C-1 signals. This signal with its two overlapping doublets can be fitted using Dmfit.[2] As result the contribution of both chitin signals to the total signal intensity can be estimated. For the C-1 signal which is used exemplary here, a α-like chitin : β-chitin ratio of 1:2 was determined. Thus, the amount of the β-chitin fibers is approximately twice as high as the amount of chitin in the α-like conformation.

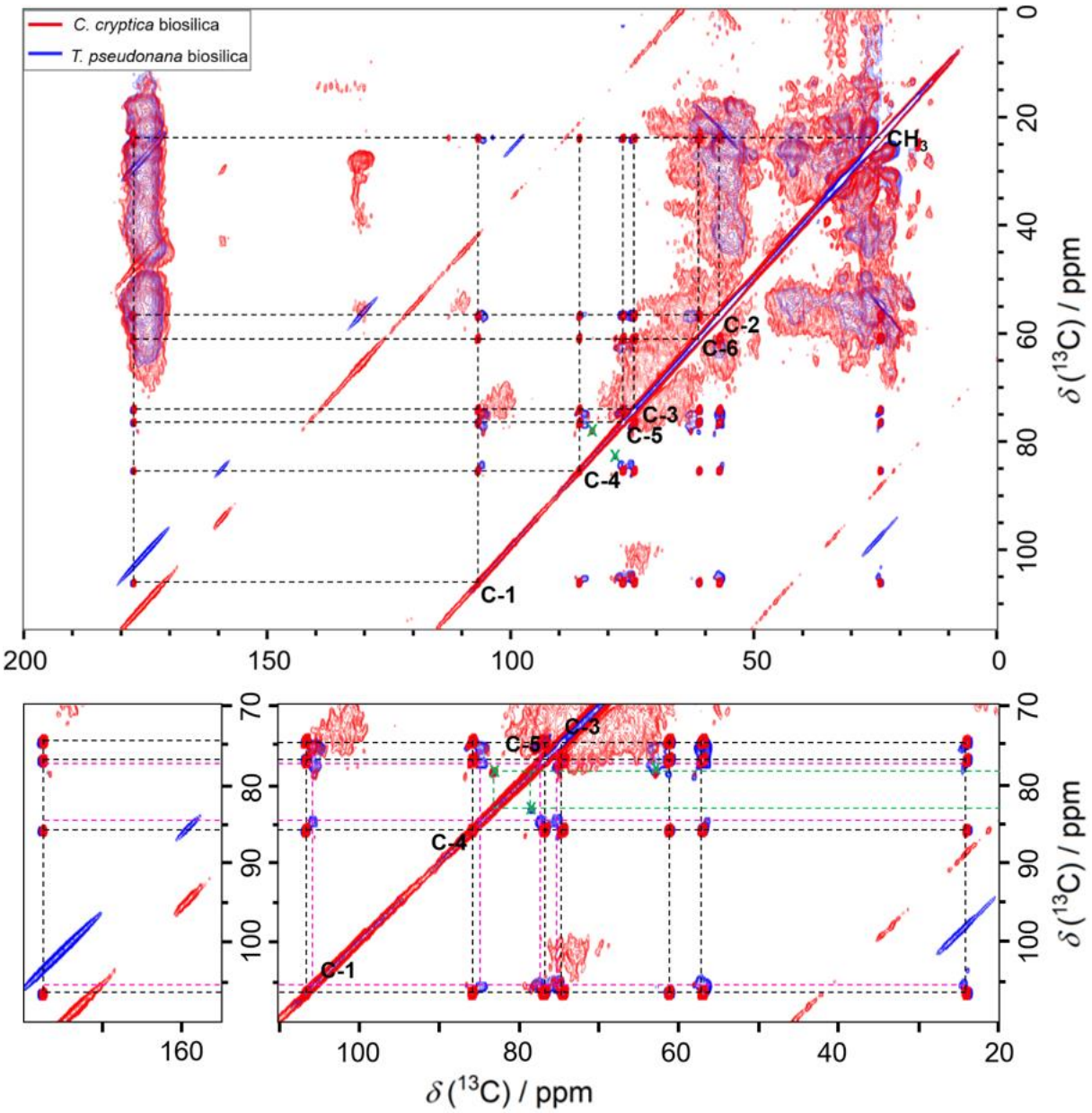

Figure S5. PDSD (Proton-Driven Spin-Diffusion) experiment of ^{13}C,^{15}N,^{29}Si-enriched *C. cryptica* and *T. pseudonana* biosilica. Top: The chitin correlation signals are highlighted. The correlation signals at 78 ppm and 83 ppm which are an indication for cross-linking occur in both samples (marked with green crosses). Bottom: A zoom into the carbohydrate region shows the presence of two different chitin forms. The amorphous chitin is marked in pink.

A comparison of the NMR spectra of *C. cryptica* and *T. pseudonana* reveals similar correlation signals especially in the carbohydrate region. Both species have two different forms of chitin, a crystalline β-chitin and an amorphous chitin, which was found to build a chitin-based meshwork in *T. pseudonana*, previously. Interestingly, the correlation signals at 78 ppm and 83 ppm which were discussed for *C. cryptica* in the main article also occur in the spectrum of *T. pseudonana*. Since cross-linking between proteins and carbohydrates was already shown for *T. pseudonana*,[3,4] the occurrence of this correlation signals in both spectra is a very strong indication for cross-linking.

Additional $^{13}C\{^{29}Si\}$ REDOR data

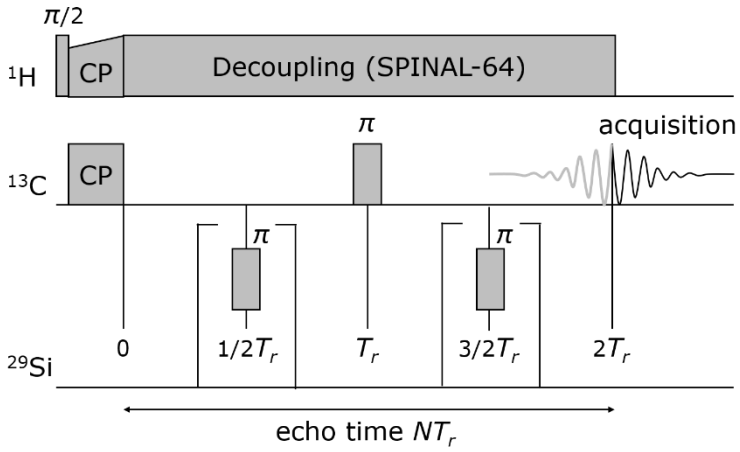

Figure S6. Pulse sequence of the $^{13}C\{^{29}Si\}$ REDOR experiment. A signal intensity S_r is measured when the π-pulses on the ^{29}Si channel cause partial echo dephasing. The echo intensity S_0 is measured without ^{29}Si irradiation. The scheme is drawn for N = 2.

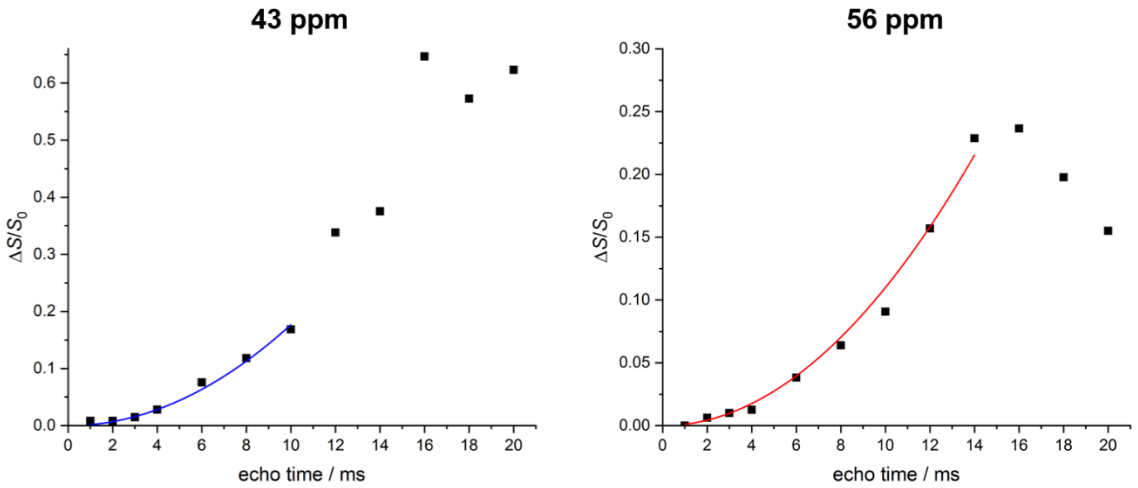

Figure S7. $^{13}C\{^{29}Si\}$ REDOR curves and the fitted first order approximation of the $^{13}C,^{15}N,^{29}Si$-enriched *C. cryptica* biosilia for the signals at 43 ppm and 56 ppm.

2. Scanning Electron Microscopy (SEM) images

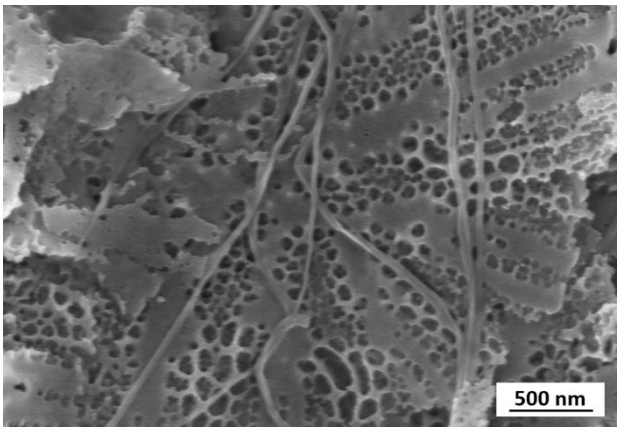

Figure S8. SEM image of the ^{13}C,^{15}N,^{29}Si-enriched *C. cryptica* biosilica prepared for the NMR experiments, originally. The fibers shown in this image are supposed to consist of β-chitin. Fibers like this were found in ca. 1 of 10 image sections only.

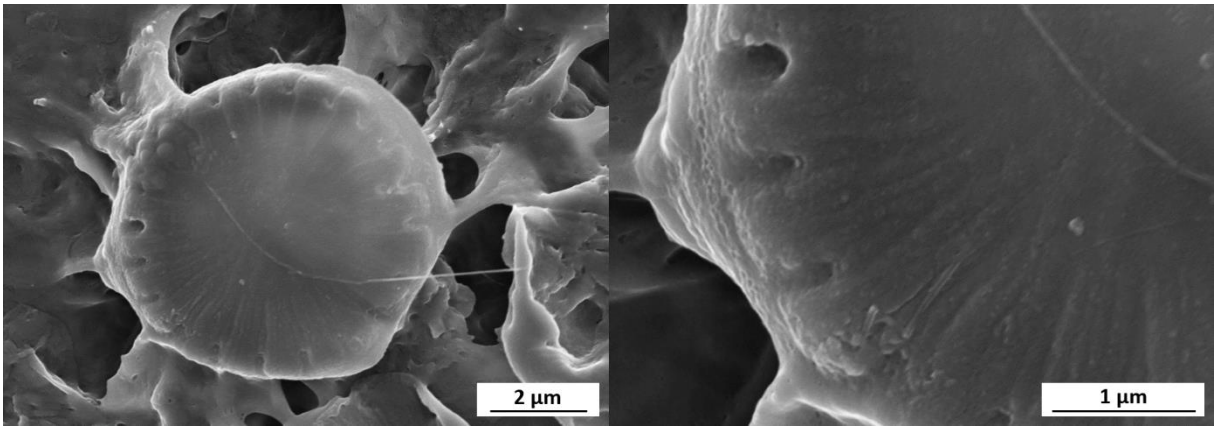

Figure S9. SEM images of the insoluble organic matrix of *C. cryptica*. Even after desilification the valve shape remains nearly unchanged. Various valve-shaped organic matrices were found. Additionally, some chitin fibers were also observed.

3. Monosaccharide analysis of *C. cryptica* biosilica and its insoluble organic matrix

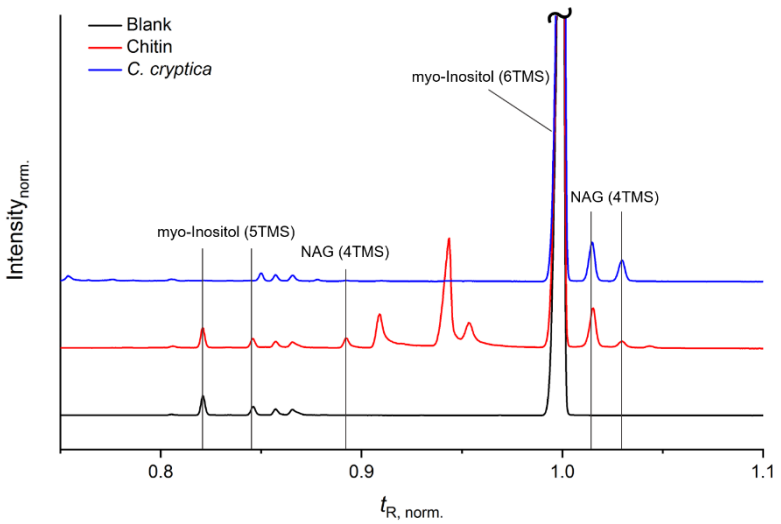

Figure S10. Presence of N-acetyl-glucosamine (NAG) in the isolated insoluble organic matrix of *C. cryptica* after chitinase treatment with chitinase from *Trichoderma viride* detected with GC-MS in comparison to a chitin reference sample.

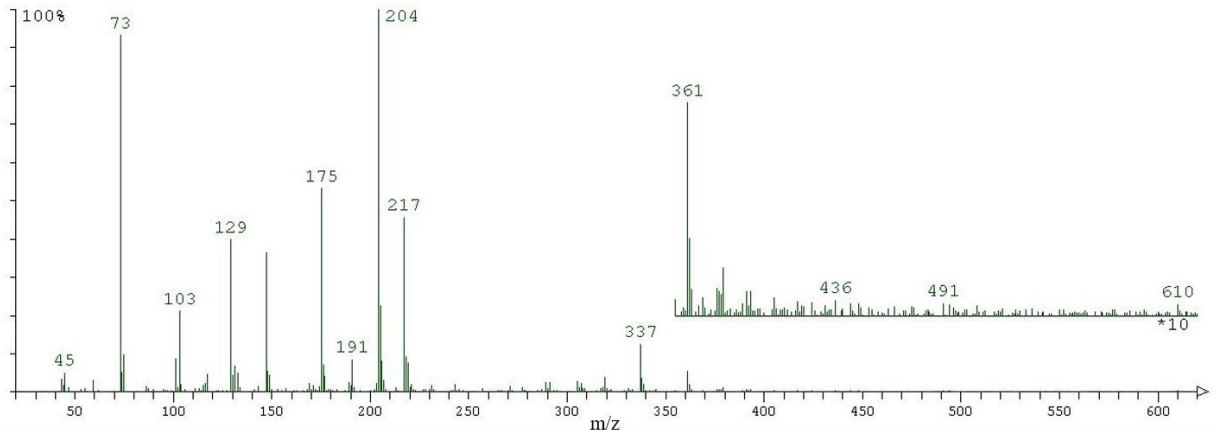

Figure S11. Exemplary mass spectrum of a non-assigned pyranose from the *C. cryptica* biosilica sample (normalized retention time $t_{R, norm.}$ = 1.240).

4. Amino acid analysis of *C. cryptica* biosilica

Experimental setup

The amino acid analysis was performed using a slightly modified protocol for LC-MS (liquid chromatography coupled with mass spectrometry) measurements.[5]

1. Hydrolysis

For the amino acid analysis, biosilica samples were hydrolyzed. 5 mg of a sample were sealed under vacuum in a glass ampoule after adding 0.5 mL 6 N HCl. The glass ampoules were put into a thermomixer and hydrolysis was performed at 110°C and shaking of 250 rpm for 24 h. Afterwards, the ampoules were opened and the hydrolysates were dried under vacuum at room temperature.

2. Sample preparation

The dried hydrolysates were extracted with 0.5 mL of ultrapure water six times (20800*g*, 5 min). The supernatants of the first three extraction steps were collected in one vial as main fraction. The supernatants of the three later extraction steps were used for a second extraction of the residue from the main fraction. The enriched rest fraction was collected in another vial as rest fraction. 100 µL of an internal standard solution (300 mg/L L-3,4-Dihydroxyphenylalanine (L-DOPA, Alfa Aesar) in ultrapure water) were added to the main fraction. All main and rest fractions were dried under vacuum at room temperature. The residue of the rest fraction was dissolved in 1 mL ultrapure water and transferred to the vial of the main fraction. The sample was diluted 1:10 for the measurements.

Calibration was performed using a standard solution that contains all 21 proteinogenic amino acids and the internal Standard L-DOPA.

3. LC-MS analysis

For liquid chromatography, a Polaris 3 C18 ether column (150x3 mm, Agilent Technologies) was used in a HPLC system Agilent 1260 Infinity (Agilent Technologies). For detection, an Agilent 6538 UHD Accurate-Mass Q-TOF-mass spectrometer with electrospray ionization (ESI) ion source was used. 5 µL of each sample were injected. The flow rate was set to 0.3 mL/min. The temperature of the column oven was 25°C. For the elution a mixture of 0.1% HCOOH in H_2O and acetonitrile is used. The solvent gradient is described in Table S1. The drying gas in the coupled mass spectrometer was set to a temperature of 250°C and a flow rate of 12 L/min. Ionization occurs using ESI (positive) with an ionization voltage of 2500 V. The fragmentor voltage was set to 80 V, the voltage of the skimmer was set to 45 V and the voltage of the octopole was set to 250 V.

Two samples of *C. cryptica* biosilica were prepared and measured parallel for validation of the results.

Table S1. Solvent gradient during LC-MS measurements used for amino acid analysis.

t / min	0.1% HCOOH in H_2O / %	acetonitrile / %
0	99	1
10	99	1
15	0	100
25	0	100
27	99	1
40	99	1

Results

The two parallel measurements of the hydrolysates of *C. cryptica* biosilica show similar results. The averaged results are shown in Figure S12. Note, that asparagine and glutamine react under the chosen hydrolysis conditions (24 h, 110°C, 6 N HCl) to aspartic acid and glutamic acid and cannot be determined separately. The disulfide bridges in cystine were not broken during hydrolysis. Thus, the concentration of cysteine should be twice the concentration determined for cystine. Tryptophan is destroyed completely during hydrolysis.

The amino acid analysis shows that the silica-associated proteins mainly consist of few main proteinogenic amino acids. Especially the amino acids serine, glycine, aspartic and glutamic acid as well as alanine and leucine should be mentioned. This observation is in good agreement to former amino acid analysis of diatom biosilica and its insoluble organic matrices.[5,6]

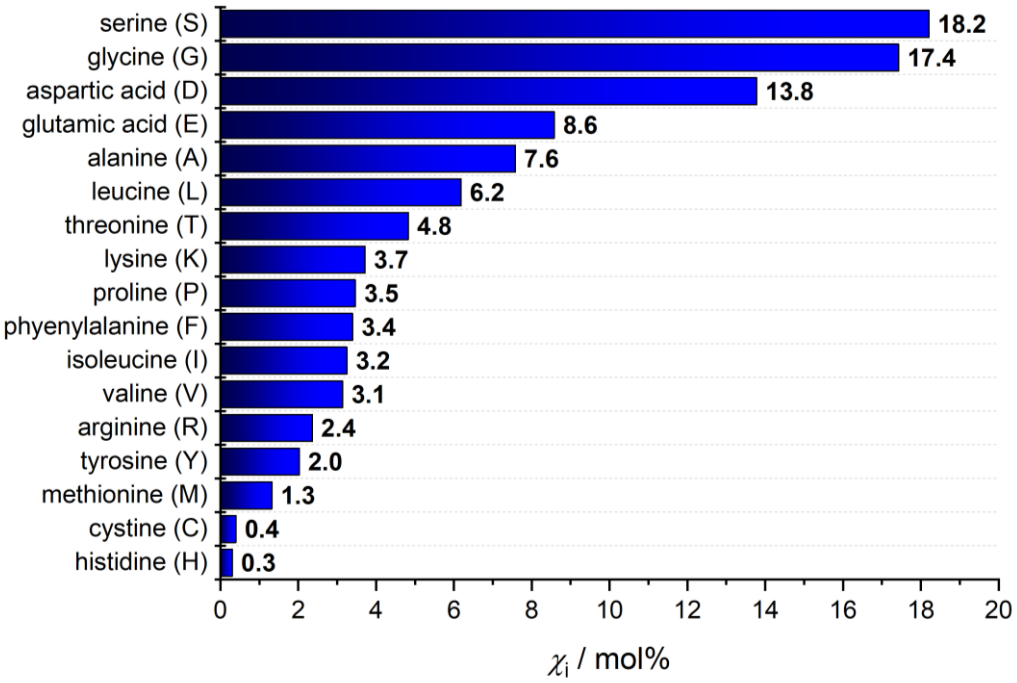

Figure S12. Results of the amino acid analysis of *C. cryptica* biosilica (mole fraction χ_i in mol%) using LC-MS. The amino acids asparagine and glutamine as well as cysteine and tryptophan are not listed due to the described influence of the chosen hydrolysis conditions (see above).

5. References

(1) Kaden, J.; Brückner, S. I.; Machill, S.; Krafft, C.; Pöppl, A.; Brunner, E. Iron Incorporation in Biosilica of the Marine Diatom Stephanopyxis Turris: Dispersed or Clustered? *BioMetals* **2017**, *30* (1), 71–82.

(2) Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calvé, S.; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z.; Hoatson, G. Modelling One- and Two-Dimensional Solid-State NMR Spectra. *Magn. Reson. Chem.* **2002**, *40* (1), 70–76.

(3) Poulsen, N.; Sumper, M.; Kröger, N. Biosilica Formation in Diatoms: Characterization of Native Silaffin-2 and Its Role in Silica Morphogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (21), 12075–12080.

(4) Poulsen, N.; Kröger, N. Silica Morphogenesis by Alternative Processing of Silaffins in the Diatom Thalassiosira Pseudonana. *J. Biol. Chem.* **2004**, *279* (41), 42993–42999.

(5) Jantschke, A.; Koers, E.; Mance, D.; Weingarth, M.; Brunner, E.; Baldus, M. Insight into the Supramolecular Architecture of Intact Diatom Biosilica from DNP-Supported Solid-State NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2015**, *54* (50), 15069–15073.

(6) Pawolski, D.; Heintze, C.; Mey, I.; Steinem, C.; Kröger, N. Reconstituting the Formation of Hierarchically Porous Silica Patterns Using Diatom Biomolecules. *J. Struct. Biol.* **2018**, *204* (1), 64–74.