ARTICLES
PUBLISHED ONLINE: 23 NOVEMBER 2010 | DOI: 10.1038/NCHEM.899
nature chemistry

Mineralization of the metre-long biosilica structures of glass sponges is templated on hydroxylated collagen

Hermann Ehrlich[1]*, Rainer Deutzmann[2], Eike Brunner[1], Enrico Cappellini[3], Hannah Koon[3], Caroline Solazzo[3], Yue Yang[3], David Ashford[4,5], Jane Thomas-Oates[5,6], Markus Lubeck[7], Carsten Baessmann[7], Tobias Langrock[8], Ralf Hoffmann[8], Gert Wörheide[9], Joachim Reitner[10], Paul Simon[11], Mikhail Tsurkan[12], Aleksander V. Ereskovsky[13,14], Denis Kurek[1,15], Vasily V. Bazhenov[1,16], Sebastian Hunoldt[1], Michael Mertig[17], Denis V. Vyalikh[18,19], Serguei L. Molodtsov[18,19], Kurt Kummer[18,19], Hartmut Worch[17], Victor Smetacek[20] and Matthew J. Collins[3]*

The minerals involved in the formation of metazoan skeletons principally comprise glassy silica, calcium phosphate or carbonate. Because of their ancient heritage, glass sponges (Hexactinellida) may shed light on fundamental questions such as molecular evolution, the unique chemistry and formation of the first skeletal silica-based structures, and the origin of multicellular animals. We have studied anchoring spicules from the metre-long stalk of the glass rope sponge (*Hyalonema sieboldi*; Porifera, Class Hexactinellida), which are remarkable for their size, durability, flexibility and optical properties. Using slow-alkali etching of biosilica, we isolated the organic fraction, which was revealed to be dominated by a hydroxylated fibrillar collagen that contains an unusual [Gly–3Hyp–4Hyp] motif. We speculate that this motif is predisposed for silica precipitation, and provides a novel template for biosilicification in nature.

Among the different biominerals, silica in its different amorphous forms is probably the most intriguing. It is probably the first and oldest natural bioskeleton, with unique mechanical properties and an extremely high specific surface area. Of the challenging topics that are receiving renewed attention today, the study of the mechanisms of biosilicification including the specificity of organic templates is among the most fascinating from chemical, biological and materials points of view.

Although it was first proposed by the groups of Morse[1,2] and Müller[3] that low molecular weight proteins—silicateins—play a pivotal role for the silification of spicules in the sponge class Demospongiae, the situation in the other siliceous spicule-producing sponge class, Hexactinellida, is less clear.

Hexactinellids are phylogenetically among the oldest metazoans, established in the Late Protoerozoic. Their skeleton is composed of silica-based spicules, the largest of which project from the body surface and serve as protective lateral spines or basal attachment roots. The basal twisted column of root tuft spicules in the 'glass rope sponge' (*Hyalonema sieboldi*; Porifera, Class Hexactinellida; Fig. 1a,b) can extend up to one metre in length and acts by anchoring the sponge in the soft bottom sediment. These spicules, which have remarkable optical properties[4], are both durable (deep-sea glass sponges live for centuries)[5] and flexible (they can be bent full circle)[6].

The presence of silicatein has been reported in the non-anchoring body microspicules of the hexactinellid *Crateromorpha meyeri*[7]. The metre-long anchoring spicules of *Hyalonema sieboldi* have been shown to be hierarchically structured[4,6], but the nature of the organic template on which silica is deposited has eluded identification. Using a novel, slow-etching method[8], we have previously reported collagen-like fibrillar proteins within both *H. sieboldi*[6,8] and *Monorhaphis chuni*[9,10] glass sponges. In this Article, we report the first detailed characterization of this.

Results and discussion

The C 1*s* near-edge X-ray absorption fine structure spectrum of *H. sieboldi* spicules shares characteristic features with a vertebrate collagen standard (Supplementary Fig. S1), as does ^{13}C solid-state nuclear magnetic resonance (NMR) spectroscopy of isolated fibrillar protein (Supplementary Fig. S2). In contrast to the identification of collagen in spicules of the glass sponge *Euplectella* sp.[11], which is probably a contaminant, we have isolated between 250 and 300 mg fibrillar protein per gram of glassy spicule from *H. sieboldi* (Fig. 1d). Polyclonal antibodies detected type I (but not type IV) collagen in the root spicules of *Hyalonema* sp. (Supplementary Fig. S3), but neither type I nor type IV collagens were detected in spicules of the demosponge *Petrosia* sp. or in the spicules supporting the body of two other hexactinellid species of the family Rosselidae. These results are consistent with an alternative macromolecular template to collagen for silicification, for example, chitin in the glass sponge *Rossella fibulata*[12].

A short digestion with papain of material solubilized by overnight digestion with trypsin, were purified by reversed-phase chromatography and subject to Edman degradation. All peptides that were sequenced had the characteristic $[Gly–Xaa–Yaa]_n$ repeat. 3-Hyp and 4-Hyp were found exclusively in positions Xaa and Yaa, respectively, and were identified unambiguously by comparison with authentic PTH amino-acid standards (Box 1).

The extracted collagen (Fig. 2a) was analysed by mass spectrometry using three different approaches (Supplementary Sections S8–S9d). To confidently assign the sequences and to differentiate

A full list of affiliations appears at the end of the paper.

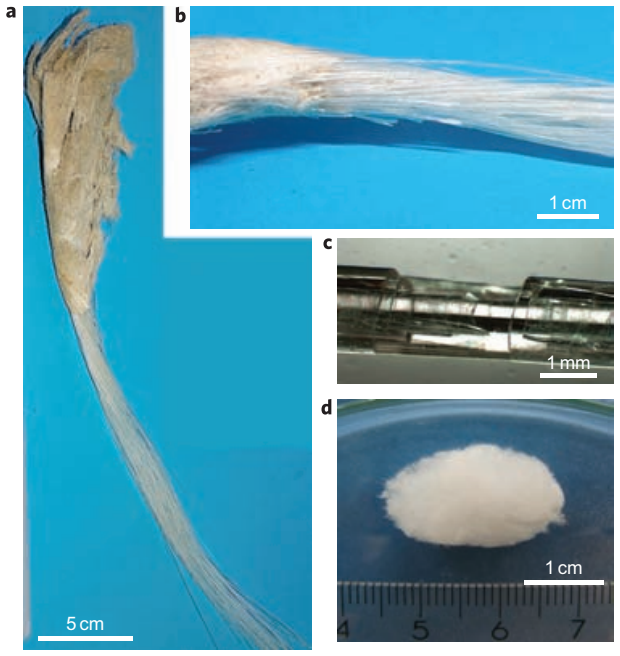


Figure 1 | Marine glass sponge *Hyalonema sieboldi*, a typical member of the Hyalonematidae family. a, Image of marine glass sponge *Hyalonema sieboldi*. Anchoring spicules of these sponges (**b**) have a multilayered structure and are organized according to the principle of 'cylinder in cylinder' (**c**). **d**, Fibrillar protein of a collagenous nature was isolated from the spicules using gentle desilicification in alkaline solution.

hydroxyproline (residue mass 113.04768 Da) from Leu/Ile (residue mass 113.08406 Da), a high-resolution, high mass accuracy mass spectrometer was used to carry out liquid chromatography-electrospray-tandem mass spectrometry (LC-ESI-MS/MS) analyses of the tryptic digest of the extracted collagen sample. The resulting data were searched against public and in-house protein sequence databases, and a range of peptides was identified as originating from collagen (Table 1). The spectra were also independently interpreted *de novo*, which resulted in the same assignments (Table 1).

Evidence for the peptide GAQGPLGPR identified from Edman sequencing (Box 1) was also obtained by mass spectrometry (Table 1), all other MS peptides were novel and most display the characteristic $[Gly–Xaa–Yaa]_n$ motif. For the longest peptide, the complete y_7–y_{13} ion series supports the existence of an unusual double hydroxylation 'Gly–Hyp–Hyp' (Supplementary Fig. S5). Spectra were also obtained that were confidently attributed to cytoskeletal actin (accession no. 3386376 from the ascidian *Molgula oculata*; Supplementary Table S1). Glass sponges lack contractile tissues, but thick actin microfilament bundles extending for hundreds of micrometres have been reported to form the core of the blunt giant rod-like extensions projecting from the edges of syncytial aggregates[13]. Actin microfilaments have been previously observed to be associated with silica-deposition vesicles in protists

Box 1 | Edman degradation

GAQ G(P)L GP
IGPDEPLKGKI
GGF GLO GR
G(V)D GNO GIX GAT GS
GGS GAO GLO GAI GNQ GAO
V GDO GLV GDL GAQ GPQ GSQ GLV G
GIO GPQ GFT GAI GVT GSO GEI GAO G
GSV **GOO** GNO GVQ GVS GO
GAT **GOO** GIS **GOO** GPQ GQO GTO GI
I GPA GPQ GQO **GOO** GPG GPX **GOO**

(Synurophyceae) and diatoms, where they are thought to be involved in shaping the cytoplasm. Their association with the organic spicule, entombed in silica, suggests a different role, perhaps associated with the maturation of the silica fibre.

Amino-acid analysis of the collagen isolated from the *H. sieboldi* spicule showed a Pro/Hyp ratio of 1.33, and a ratio of ~3:1 of trans-4-Hyp to trans-3-Hyp (Fig. 2c; Supplementary Table S2). These ratios are remarkably consistent with peptide sequence data, with hydroxylation of 33% of those Pro residues in the Xaa position and 100% in the Yaa position (as 3-Hyp and 4-Hyp, respectively) of the $[Gly–Xaa–Yaa]_n$ motif. This result is similar to those reported previously for *Geodia cydonium* sponge collagen[14]. The presence of both 4-hydroxyproline and 3,4-dihydroxyproline has been reported in siliceous cell walls of diatoms[15,16], and these authors suggested that hydroxylated amino acids could play a role in silicification of diatom cell walls. Hydroxy amino acids are known to be distributed in cell walls of diatoms[17] as well as in silicateins, the specific proteins responsible for silicification in demosponges[1–3].

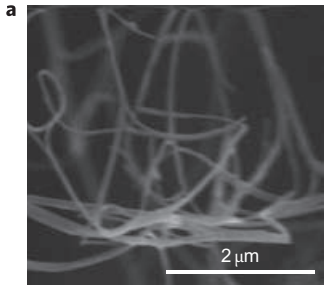
We were able to demonstrate the role of the hydroxylation state of collagen in silica polycondensation. The rate of silica formation was significantly higher in *H. sieboldi* spicular collagen than it was in two samples of collagen lacking significant trans-3-Hyp that were isolated from calf skin and *Chondrosia reniformis* (a non-spicular desmosponge) mesohyl (Supplementary Fig. S15). However, we were able to reduce silicification activity when the 3- and 4-hydroxyproline residues of *H. sieboldi* were protected by formation of a ketal group (Fig. 3). Functional recovery was restored when the ketal protecting group was removed (Fig. 3d, inset).

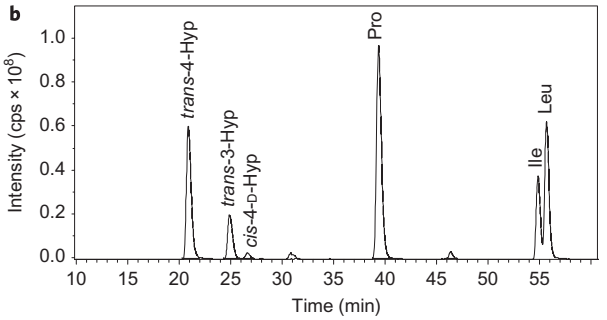
The collagen motif determined in *Hyalonema* is consistent with the model of Schumacher and colleagues[18], which describes 3(S)-hydroxyproline residues in the Xaa position of the collagen triple helix. This structure offers a plausible molecular model for the interaction between polysilicic acid and Gly–3Hyp–4Hyp polypeptides of isolated glass sponge collagen (Supplementary Fig. S12). It is established that the interaction between orthosilicic acid and hydroxyl groups is likely to be a hydrogen bond[19]. Our model shows the possibility of stable complex formation on the basis of hydrogen bonding between hydroxyl groups of polysilicic acid and surface exposed hydroxyls of 3-Hyp and 4-Hyp. Our model proposes a functional role for trans-3-Hyp in sponge collagen silicification. Collagen will

Table 1 | Collagen peptides identified by high-resolution mass spectrometry and manual *de novo* sequencing, or by Edman sequencing.

MS/MS	Observed	M_r (expt.)	M_r (calc.)	ppm	Score	Expect
GPJ GPT GJQ GAR*	562.8033	1,123.592	1,123.5986	−5.87	77	0.000015
E GEJ GJO GET GPR	664.3041	1,326.5937	1,326.6052	−8.64	16	0.024
GJO GAO GJD GNO GPA GJR	832.9138	1,663.8130	1,663.8166	−2.13	86	8.9×10^{-9}
(MEGPT) GAP GAO GDA GVJ **GOO GOO** G(PQGPR)	896.4121	2,686.2146	2,686.2294	−5.51	66	1.1×10^{-6}
GAQ GPJ GPR	426.7374	851.4603	851.4613	−1.19	54	6.7×10^{-6}

Q/N, glutamine/asparagine deamidation; O, hydroxyproline; J indicates leucine or isoleucine; the presence of brackets indicates uncertain residue (P) or sequence order (for example, (XY) indicates either XY or YX). *A significant match to the starlet sea anemone *Nematostella vectensis* (phylum *Cnidaria*; gi|156394292).





c

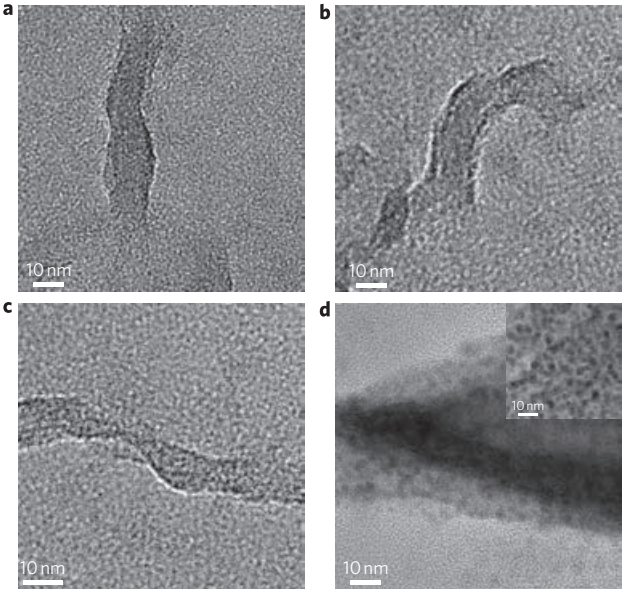
Amino acid	*Hyalonema sieboldi*
trans-4-Hyp	73.6 ± 0.8%
trans-3-Hyp	24.0 ± 0.4%
cis-4-D-Hyp	2.5 ± 0.4%
Ratio Pro/Hyp	1.33 ± 0.01

Figure 2 | Analysis of the isolated spicular organic matrix. a, SEM image of the nanofibrils observed in alkali extracts obtained after gentle demineralization over 14 days at 37 °C. **b**, Extracted ion chromatogram for the N_2-(5-fluoro-2,4-dinitrophenyl)-L-valine amide (FDVA) derivatives at m/z 412 (Hyp, Leu, Ile) and m/z 396 (Pro) from the hydrolysate of the organic matrix of demineralized *H. sieboldi* spiculae. cps, counts per second. **c**, Relative amounts of all three Hyp-isomers and the Pro/Hyp ratio based on the signal intensities shown in **b**. The table summarizes the data for the amino-acid analysis, which provides the amount of each amino acid (g or mol). As all amino acids represent isomers with identical mass, the percentage represents the content of each Hyp residue in collagen.

present a layer of hydroxyl groups that can undergo condensation reactions with silicic acid molecules with a consequent loss of water. As a result, the initial layer of condensed silicic acid will be held fixed to the collagenous template in a geometric arrangement that will favour further polymerization of silicic acid, similar to the model proposed by Hecky and colleagues[16]. It therefore appears that collagen was a novel template for biosilicification that emerged at an early stage during metazoan evolution, and that the occurrence of additional trans-3-Hyp plays a key role in stabilizing silicic acid molecules and initiating the precipitation of silica.

Hydroxylated collagen appears to form the basis for the extraordinary mechanical and optical properties of hexactinellid spicules[20]. The self-assembly properties of collagen and its templating activity with respect to silicification are consistent with recent ideas on the development of hierarchical silica-based architectures[21]. Macroscopic bundles of silica nanostructures result from the kinetic cross-coupling of two molecular processes: a dynamic supramolecular self-assembly and a stabilizing silica mineralization. The feedback interactions between template growth and inorganic deposition are driven non-enzymatically by means of hydrogen bonding. We speculate that the hydroxylated glass sponge collagen may change the nature of silica in aqueous solution by converting the distribution of oligomers into a more uniform and useful set of nanoparticle precursors for assembly into the growing solid (Supplementary Section S10).

Our findings suggest that in addition to the previously described silicatein-based biosilicification of sponge spicules, collagen has a key role to play in the formation of the long, flexible, optically pure anchoring spicules of the Hexactinellids. Increased atmospheric oxygen during the Proterozoic may have been linked to post-translational hydroxylation of proline and lysine residues[22], and it is tempting to speculate that the occurrence of silica- and hydroxylated collagen-based composites in skeletal structures of the first metazoan might be a co-evolutionary event. A reconstruction of the evolution of biocalcification as well as of biosilicification with respect to collagen may be a key way to obtain strong evidence



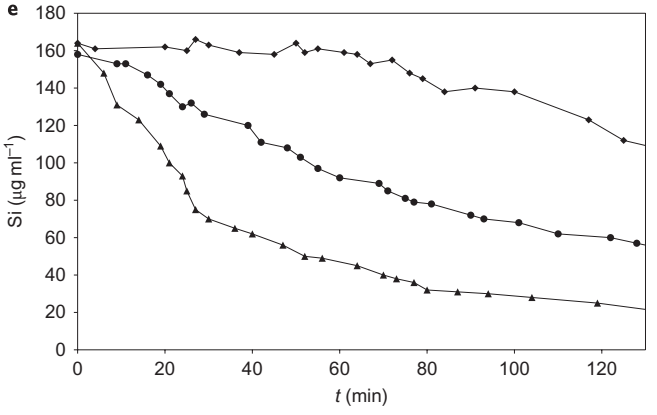


Figure 3 | HR-TEM images of silicification on *H. sieboldii* collagen. Silicification is apparent as nanoparticles after exposure of nanofibrillar *H. sieboldii* spicular collagen (**a**) to a solution of sodium methasilicate solution for 30 min. **b**, However, after protection of 3- and 4-hydroxyproline residues by ketal groups (Supplementary Figs S11 and S13), there is no visible silica deposition. **c**, Cleavage of the ketal protecting groups from collagen leads to a functional recovery with respect to silicification. **d**, The layer of silica nanoparticles is formed around the nanofibril of native spicular collagen during the first 30 min of silicification, as seen in the native collagen fibre (Supplementary Fig. S16). **e**, The results are in good agreement with measurements of activity (Supplementary Fig. S10) for non-protected collagen (filled triangle), which is lost following protection (filled diamonds), and partially restored when this protection is removed (filled circles).

of ancient, ancestral programs[23] of biomineralization based on this common template. The bioconstruction of the uniquely large siliceous structures (ten orders of magnitude longer than the spicules of demosponges) was probably enabled by the incorporation of collagen, which can play a role as a template and also provide structural support. This may mean a re-thinking of the role of collagen in the evolution of biomineralization, and almost certainly opens up new strategies for the biomimetic synthesis of silica-based materials.

Methods

Basal spicules of *Hyalonema sieboldi* (Hexactinellida: Porifera) Gray 1835, collected from a depth of 5,000 m in the Pacific (12 °N, 137 °E), as well as those from *H. sieboldi* collected by C. Eckert in Sagami Bay, Japan (FS 'Tansai Maru', St. TS 4–8, May 2004) were used in this study. Dried spicules were washed three times in distilled water, cut into pieces (2–5 cm long), and placed in a solution containing purified *Clostridium histolyticum* collagenase (Sigma) to digest any possible collagen contamination of an exogenous nature. After incubation for 24 h at 15 °C, the pieces of spicule were washed again, three times in distilled water, then dried and placed in 10 ml plastic vessels containing 5 ml of 2.5 M NaOH (Fluka) solution. The vessel was covered, placed under thermostatic conditions at 37 °C and shaken slowly for 14 days. Alkali extracts of *H. sieboldi* spicules containing fibrillar protein were dialysed against deionized water on Roth (Germany) membranes with a cut-off of 14 kDa. The dialysed material was dried under vacuum in a CHRIST lyophilizer (Germany) and used for collagen identification (described in detail in the Supplementary Information).

Analytical methods. The analytical methods used in this work include near-edge X-ray absorption fine structure spectroscopic method (NEXAFS), ^{13}C solid-state nuclear magnetic resonance (NMR), scanning electron microscopy (SEM), transmission electron microscopy (TEM), high-resolution transmission electron microscopy (HR-TEM), Fourier transform infrared spectroscopy (FTIR) and mass spectrometric methods, Edman degradation, the immunochemical method, ketal protection of the hydroxyl groups of the 3-hydroxyproline and 4-hydroxyproline of collagen, as well as measurements of the silica condensing activity of collagens). These are described in detail in the Supplementary Information.

Data deposition. All the raw mass spectrometric data associated with this manuscript may be downloaded from ProteomeCommons.org Tranche, http://tranche.proteomecommons.org, using the hash

VmBiB25HJz + NtxKBunAPvmvsR0Ax8ddFDBVdziCJKC9i4gLDSQeYQoqO/NP7RrTf89 RfZSMilHLn/OA0aWhn + Nyy8KIAAAAAAA4xYg = =

or the link

http://www.proteomecommons.org/data-downloader.jsp?fileName = VmBiB25HJz%2BNtxKBunAPvmvsR0Ax8ddFDBVdziCJKC9i4gLDSQeYQoqO/NP7RrTf89RfZSMilHLn/OA0aWhn%2BNyy8KIAAAAAAA4xYg = =

The hash may be used to demonstrate exactly which files were published as part of the data set of this manuscript, and also to check that the data have not changed since publication.

Received 2 February 2010; accepted 5 October 2010; published online 23 November 2010

References

1. Shimizu, K., Cha, J., Stucky, G. D. & Morse, D. E. Silicatein alpha: cathepsin L-like protein in sponge biosilica. *Proc. Natl Acad. Sci. USA* **95,** 6234–6238 (1998).
2. Cha, J. N. *et al.* Silicatein filaments and subunits from a marine sponge direct the polymerization of silica and silicones *in vitro*. *Proc. Natl Acad. Sci. USA* **96,** 361–365 (1999).
3. Müller, W. E. G. *et al.* Silicateins, the major biosilica forming enzymes present in demosponges: protein analysis and phylogenetic relationship. *Gene* **395,** 62–71 (2007).
4. Müller, W. E. G. *et al.* Unique transmission properties of the stalk spicules from the hexactinellid *Hyalonema sieboldi*. *Biosens. Bioelectron.* **21,** 1149–1155 (2006).
5. Dayton, P. K. Observations of growth, dispersal and population dynamics of some sponges in McMurdo Sound, Antarctica, in *Colloques internationaux du C.N.R.S. 291. Biologie des spongiaires* (eds Levi, C. & Boury-Esnault, N.), 271–282 (Editions du Centre National de la Recherche Scientifique, 1979).
6. Ehrlich, H. & Worch, H. Collagen, a huge matrix in glass-sponge flexible spicules of the meter-long *Hyalonema sieboldi*, in *Handbook of Biomineralization Vol. 1. The Biology of Biominerals Structure Formation* (ed. Bäuerlein, E.) 23–41 (Wiley VCH, 2007).
7. Müller, W. E. G. *et al.* Silicatein expression in the hexactinellid *Crateromorpha meyeri*: the lead marker gene restricted to siliceous sponges. *Cell Tissue Res.* **333,** 339–351 (2008).
8. Ehrlich, H. *et al.* A modern approach to demineralization of spicules in glass sponges (Porifera: Hexactinellida) for the purpose of extraction and examination of the protein matrix. *Russ. J. Marine Biol.* **32,** 186–193 (2006).
9. Ehrlich, H. *et al.* Nanostructural organization of naturally occurring composites. Part I. Silica-collagen-based biocomposites. *J. Nanomater.* doi:10.1155/2008/623838 (2008).
10. Ehrlich, H., Heinemann, S., Hanke, T. & Worch, H. Hybrid materials from a silicate-treated collagen matrix, methods for the production thereof and the use thereof. International patent WO2008/023025 (2008).
11. Travis, D. F., Francois, C. J., Bonar, L. C. & Glimcher, M. J. Comparative studies of the organic matrices of invertebrate mineralized tissues. *J. Ultrastruct. Res.* **18,** 519–550 (1967).
12. Ehrlich, H. *et al.* Nanostructural organization of naturally occurring composites. Part II. Silica-chitin-based biocomposites. *J. Nanomater.* doi:10.1155/2008/670235 (2008).
13. Leys, S. P. Cytoskeletal architecture and organelle transport in giant syncytia formed by fusion of hexactinellid sponge tissues. *Biol. Bull.* **188,** 241–254 (1995).
14. Diehl-Seifert, B. *et al.* Attachment of sponge cells to collagen substrata: effect of a collagen assembly factor. *J. Cell Sci.* **79,** 271–285 (1985).
15. Nakajima, T. & Volcani, B. E. 3,4-Dihydroxyproline: a new amino acid in diatom cell walls. *Science* **164,** 1400–1401 (1969).
16. Hecky, R. E., Mopper, K., Kilham, P. & Degens, E. T. The amino acid and sugar composition of diatom cell walls. *Mar. Biol.* **19,** 323–331 (1973).
17. Sadava, D. & Volcani, B. E. Studies on the biochemistry and fine structure of silica shell formation in diatoms. Formation of hydroxyproline and dihydroxyproliner in *Nitzschia angularis*. *Planta* **135,** 7–11 (1977).
18. Schumacher, M. A., Mizuno, K. & Bachinger, H. P. The crystal structure of a collagen-like polypeptide with 3(S)-hydroxyproline residues in the Xaa position forms a standard 7/2 collagen triple helix. *J. Biol. Chem.* **281,** 27566–27574 (2006).
19. Tilburey, G. E., Patwardhan, S. V., Huang, J., Kaplan, D. L. & Perry, C. C. Are hydroxyl-containing biomolecules important in biosilicification? A model study. *J. Phys. Chem. B* **111,** 4630–4638 (2007).
20. Kulchin, Y. N. *et al.* Optical fibres based on natural biological minerals—sea sponge spicules. *Quantum Electron.* **38,** 51–55 (2008).
21. Pouget, E. *et al.* Hierarchical architectures by synergy between dynamical template self-assembly and biomineralization. *Nature Mater.* **6,** 434–439 (2007).
22. Exposito, J.-Y., Cluzel, C., Garrone, R. & Lethias, C. Evolution of collagens. *Anat. Rec.* **268,** 302–316 (2002).
23. Livingston, B. T. *et al.* A genome-wide analysis of biomineralization-related proteins in the sea urchin *Strongylocentrotus purparatus*. *Dev. Biol.* **300,** 335–348 (2006).

Acknowledgements

This work was partially supported by NE/C511148/1 grant to M.J.C. and J.T.-O., by a joint grant from Deutscher Akademischer Austauschdienst (DAAD) (grant ref. 325; A/08/72558) and Russian Ministry of Education and Science (RMES) (AVCP grant no. 8066) to V.V.B., and by the Erasmus Mundus External Co-operation Programme of the European Union 2009 to D.K. The Centre of Excellence in Mass Spectrometry (CoEMS) at York is supported by Science City York and Yorkshire Forward using funds from the Northern Way Initiative. The authors also thank H. Lichte for use of the facilities at the Special Electron Microscopy Laboratory for high-resolution and holography at Triebenberg, TU Dresden, Germany. The authors thank G. Bavestrello, K. Tabachnick, M. Hofreiter, H. Roempler and E. Bäeuerlein for their helpful discussions. Finally, the authors are very grateful to S. Paasch, H. Meissner, G. Richter, T. Hanke, O. Trommer, A. Mensch, S. Heinemann for excellent technical assistance. G.W. acknowledges funding from the German Science Foundation (DFG). Financial support by DFG (Graduiertenkolleg 378) to R.H. and T.L. and the European Fund for Regional Structure Development (EFRE, European Union and Free State Saxonia) to R.H. is gratefully acknowledged.

Author contributions

All authors contributed to the design or execution of experiments, or analysed data. H.E. supervised the experiments, carried out demineralization experiments, performed collagen isolation, and wrote the manuscript. P.S. performed SEM and HRTEM, and prepared figures. A.E. collected, prepared and identified sponge samples and contributed to writing the manuscript. M.M., D.V.V., K.K. and S.L.M. performed NEXAFS experiments and designed figures. M.T. and V.V.B. carried out collagen modification. S.H. performed FTIR and prepared figures. E.B. performed NMR. R.D. performed Edman degradation and R.H. and T.L. performed amino acid analysis and mass spectrometry. M.C., H.K., C.S., Y.Y., E.C., D.A., M.L., C.B. and J.T.-O. were involved in acquiring and interpreting the mass spectrometric data, and M.C., H.K., E.C., D.A. and J.T.-O contributed to the writing of the manuscript. H.W., M.C., H.E., G.W., J.R., V.S. and E.B. analysed the results with regard to evolutionary implications and mechanisms of biomineralization, designed concepts, and wrote the manuscript.

Additional information

The authors declare no competing financial interests. Supplementary information accompanies this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/. Correspondence and requests for materials should be addressed to H.E. and M.J.C.

[1]Institute of Bioanalytical Chemistry, Dresden University of Technology, D-01062 Dresden, Germany, [2]Institute for Biochemistry, Genetic and Microbiology, University of Regensburg, 93053 Regensburg, Germany, [3]BioArCh, Department of Biology, University of York, Heslington, York, YO10 5YW, UK, [4]Technology Facility, Department of Biology, University of York, Heslington, York YO10 5YW, UK, [5]Centre of Excellence in Mass Spectrometry, University of York, Heslington, York YO10 5DD, UK, [6]Department of Chemistry, University of York, Heslington, York YO10 5DD, UK, [7]Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany, [8]Institute of Bioanalytical Chemistry, Center for Biotechnology and Biomedicine (BBZ), Universität Leipzig, D-014103 Leipzig, Germany, [9]Department für Geo- und Umweltwissenschaften & GeoBio-Center^{LMU}, Ludwig-Maximilians-Universität, D-80333 München, Germany, [10]Department of Geobiology, University of Göttingen, D-37077 Göttingen, Germany, [11]Max Planck Institute of Chemical Physics of Solids, D-01187, Dresden, Germany, [12]Leibniz Institute of Polymer Research Dresden and Max Bergmann Center of Biomaterials, D-01005 Dresden, Germany, [13]Centre d'Oceanologie de Marseille, Station marine d'Endoume, Aix-Marseille Universite–CNRS UMR 6540-DIMAR, 13007 Marseille, France, [14]St. Petersburg State University, 199034 St. Petersburg, Russia, [15]Centre 'Bioengineering' Russian Academy of Sciences, 117312 Moscow, Russia, [16]Institute of Chemistry and Applied Ecology, Far Eastern National University, 690650 Vladivostok, Russia, [17]Max-Bergmann Center of Biomaterials and Institute of Materials Science, Dresden University of Technology, D- 01069 Dresden, Germany, [18]Institute of Solid State Physics, Dresden University of Technology, D- 01069 Dresden, Germany, [19]European XFEL GmbH, 22761 Hamburg, Germany, [20]Alfred Wegener Institute for Polar and Marine Research, D-27570 Bremerhaven, Germany.
*e-mail: hermann.ehrlich@tu-dresden.de; mc80@york.ac.uk

nature chemistry

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.899

Mineralization of the meter-long biosilica structures of glass sponges is templated on hydroxylated collagen

Hermann Ehrlich, Rainer Deutzmann, Eike Brunner, Enrico Cappellini, Hannah Koon, Caroline Solazzo, Yue Yang, David Ashford, Jane Thomas-Oates, Markus Lubeck, Carsten Baessmann, Tobias Langrock, Ralf Hoffmann, Gert Wörheide, Joachim Reitner, Paul Simon, Mikhail Tsurkan, Aleksandr V. Ereskovsky, Denis Kurek, Vasily V. Bazhenov, Sebastian Hunoldt, Michael Mertig, Denis V. Vyalikh, Serguei L. Molodtsov, Kurt Kummer, Hartmut Worch, Victor Smetacek and Matthew J. Collins.

Author Information. Correspondence and requests for materials should be addressed to H.E. (Hermann.Ehrlich@tu-dresden.de) and M.J.C. (mc80@yourk.ac.uk).

Supplementary information

Table of Contents

DOI: 10.1038/NCHEM.899

2. Supplementary figures

Supplementary methods and text

S1. *NEXAFS analysis*

Near-edge X-ray absorption fine structure (NEXAFS) spectroscopic methods were used as established at BESSY, Berlin[1,2]. All images chosen for this paper are representative examples of features observed several times for different samples.

S3. *Solid-state ^{13}C NMR*

Solid-state ^{13}C NMR experiments were performed on a Bruker Avance 300 spectrometer operating at 75.47 MHz for ^{13}C using a commercial double resonance 4 mm MAS NMR probe at 14 kHz sample spinning rate. Ramped ^{1}H-^{13}C cross-polarization (CP)[3,4] was used (contact time:

4 ms). SPINAL ^{1}H-decoupling[5] was applied during the signal acquisition. The spectra were referenced relative to tetramethylsilane (TMS). The spectral signature of two collagen samples is almost identical and agrees well with the spectra observed for collagen from other organisms[6].

S3. ***Immunchemical and Western blot analysis***

Basal spicules of *Hyalonema sieboldi* were overgrown with *Palythoa*-zooanthids, which were removed before analysing the spicule-related proteins. The spicules were cleaned using 2% NaClO solution. After treatment with NaClO the samples were carefully washed with distilled water 6 times until all remaining NaClO was washed away. The spicules were then freeze dried and stored at 4°C. For SDS gel analyses 100 mg spicule material was crushed using an agate stone mortar. The pulverised spicules were mixed with 100 mL Roti Load1 and agitated. Then the mixture was boiled at 95^{o}C for 5 min and treated for 15 min in an ultrasonic bath. The samples were centrifuged and the supernatant was used for gel electrophoresis. 13-14% SDS gels were run for 1 hour using 25 mA. For blotting an 0.1 µm PVDF membrane (Millipore) was used, with transfer buffer (25 mM Tris, 192 mM glycine-20% methanol) for 3 h at 150 mA.

The PVDF membrane was prepared 24 h at room temperature with Blotto buffer (TBS-Tween in skimmed milk). Collagen 1 antibody (Collagen Type I – Mammalian; ABIN97475 – antibodies online) was incubated for 3 h at room temperature (3x30 washing with Blotto buffer; 30 min incubation with anti-rabbit IgG Sigma A 3687 at room temperature; 3x20 min washing with Blotto buffer; staining with BCIP/NBT Blue (Sigma B3804)).

S4. ***Transmission electron microscopy (TEM).***

Conventional transmission electron microscopy (Fig. 3a-e and Fig. S11) was performed with a Philips CM200 FEG\ST Lorentz electron microscope at an acceleration voltage of 200 kV. For electron microscopy, a drop of the water suspension containing the sample was placed on the electron microscopy grid. After one minute, the excess was removed using blotting paper and thereafter dried in air. The electron microscopy grids (Plano, Germany) were covered with a perforated carbon film

S5. ***Scanning electron microscopy (SEM) and Energy dispersive x-ray microanalysis (EDX)***

The samples were fixed in a sample holder and covered with carbon, or with a gold layer for 1 min using an Edwards S150B sputter coater. The samples were then placed in an ESEM XL 30 Philips or LEO DSM 982 Gemini scanning electron microscope**.**

DOI: 10.1038/NCHEM.899

For EDX analysis air-dried samples were embedded in Epoxy resin without additional staining and cut on a Leica EM UCT ultramicrotome to obtain a flat block face. Samples were coated with carbon and analysed in an ESEM XL 30 Philips. EDX-analysis was done with an EDAX detecting unit and EDAX software.

S6. ***Cleavage of sponge collagen***

All experiments were carried out using lyophilized *H. sieboldi* sponge collagen that had been treated with 0.1 M sodium hydroxide to remove silica. The triple-helical conformation and the helicity of the collagen molecule are maintained on alkali treatment, up to a concentration of 2.5 M NaOH at 37°C *in vitro*[10,14] (Supplementary Fig.11).Collagen was only slightly soluble in the buffers used for proteolytic digestion (see below) and was therefore suspended by short sonication. Attempts to cleave the protein with cyanogen bromide in 70% formic acid or trypsin in ammonium hydrogen carbonate buffer were unsuccessful. Pepsin digestion in 0.5 M formic acid solubilized about half of the material, and almost complete solubilization was achieved using papain at 50°C (collagen/papain ratio 10:1, buffer: 0.1 M NH_4HCO_3, 2 mM mercaptoethanol, 0.1 mM EDTA). However the resulting peptide mixture was too complex for separation via HPLC. To obtain peptides for sequencing by Edman degradation the following protocol proved effective:

1 mg collagen was suspended in 500 µL papain buffer, sonicated for 10 s and digested with 2 µg papain at 50°C for 15 min. Digestion was stopped by adding 10 µL 1 M sodium iodoacetate and 10 µL 0.1 M $CaCl_2$. After incubation for 30 min, 10 µg trypsin was added and incubated overnight at 37°C. The peptides were separated using two-dimensional reversed-phase HPLC. The peptides were first separated by gradient elution on a Jupiter Proteo column (250 x 2 mm, 4 µm particle size, Phenomenex). Buffer A was 10 mM ammonium acetate, buffer B 70% acetonitrile in water. Selected peaks were lyophilized and separated on a Synergi Hydro-RP column (250 x 2 mm, 4 µm particle size, Phenomenex) using the standard trifluoroacetic acid/acetonitrile system.

The peptides were sequenced on a Procise 492A sequencer (Applied Biosystems) with on-line detection of the PTH amino acids, according to the manufacturer's instructions. Amino acids that could not be identified unambiguously are given in parentheses or denoted by X. 3Hyp and 4Hyp were identified by comparison with the retention times of the PTH amino acids prepared from the corresponding authentic amino acids obtained from Fluka.

S7. ***Amino acid analysis***

Amino acid analysis was performed as described elsewhere[7]. A collagen extract from *H. sieboldi* spicules was suspended in 0.05 mol/L acetic acid (2 mg/mL). 150 µL of this suspension were transferred into hydrolysis glass inlets and dried under vacuum. The samples were hydrolyzed in gas phase using 300 µL of 6 N hydrochloric acid (for hydrolysis, Sigma Aldrich, Steinheim) at 110°C for 24 h. The hydrolysates were dissolved in 15 µL of water and 1/3 was derivatized with N2-(5-fluoro-2,4-dinitrophenyl)-L-Valinamid (FDVA) and diluted to 500 µL. 100 µL were injected on an Aqua-C18 column (Phenomenex, Aschaffenburg) and the amino acid derivatives were eluted by a stepwise linear gradient from 20% to 65% acetonitrile in water, with 0.1% formic acid. Derivatized amino acids were detected with a Bruker esquire3000plus ion trap mass spectrometer, using an m/z scan from 300 to 800 or a product ion scan of *m/z* 412 (derivatized hydroxyproline, leucine and isoleucine).

S8. ***Protein extraction and digestion***

S8a. ***Tryptic Digestion***

5 mg of proteinaceous material extracted from a sample of *H. sieboldii* anchoring spicules were suspended in a 500 µL solution of 50 mM ammonium bicarbonate pH 8.0 and heated at 85˚C for 48 h and then allowed to cool at 37˚C. 10 µg of sequencing grade modified trypsin (Promega Corporation, Madison, WI) were added and digestion was allowed to proceed for 48 h at 37˚C; after 24 h, 5 µg of fresh trypsin were added. The sample was centrifuged at 20000 x g for 10 min at room temperature, and the supernatant collected and dried in a centrifugal evaporator. The pellet was resuspended in 15 µL of 0.1% aqueous trifluoroacetic acid (TFA).

S8b. ***In-Gel Digestion***

40 µL of Laemmli buffer (5 µL MCE, 95 µL LaemmLi) was added to 1 mg of the proteinaceous material. The solution was heated for 15 min at 95ºC and then briefly centrifuged. 30 µL of the supernatant, containing the solubilised material, was pipetted into the well of a NuPAGENovex 4-12% Bis-Tris Gel in an XCell *SureLock*™ Mini-Cell (Invitrogen, Paisley). A Broad Range (2–212 kDa) protein marker (BioLab, Hertfordshire) was also included for reference. Electrophoresis was run with MES-SDS Running Buffer (NuPAGE, Invitrogen) at 200 V for 35 min. The gel was then washed and stained overnight, at 4ºC, with 20 µL Simply blue safe stain (Invitrogen) and 2 mL of 20% NaCl (w/v in water). Following a water destain step, two faint bands between 116 and 158 kDa were visible.

DOI: 10.1038/NCHEM.899

The bands were excised from the gel, cut into 2 mm^3 cubes and placed in 'low bind' plastic microfuge tubes with 100 µL of 25 mM ammonium bicarbonate/ 50% acetonitrile. The samples were then vortexed for 30 min in this solution to remove the blue stain. This wash/dehydration step was repeated twice and followed by a further immersion in acetonitrile for 5 min. The shrunken gel pieces were then dried on a centrifugal concentrator for 20 min. 20 µL of prepared trypsin (0.4 µg) (proteomics grade enzyme, Sigma-Aldrich, Poole) was added to each tube containing the dried gel pieces. The gel pieces were allowed to rehydrate for 10-15 min on ice before 50 µL of trypsin reaction buffer (Sigma-Aldrich, Poole) was added and the digestion was allowed to proceed for 16 h at 37˚C. Following digestion, the samples were centrifuged for 30 s and the supernatant containing the extracted peptides was collected. 30 µL of extraction solution (50% (v/v) acetonitrile and 0.1% (v/v) TFA) was then added to the gel pieces and after 10 min of occasional gentle vortex mixing, the tubes were centrifuged for 30 s and the supernatant removed and combined with the first extract. To concentrate the peptides, the volume of the extract was reduced to 20 µL in a centrifugal concentrator.

S9. *Collagen peptide mass spectrometry sequencing*

S9a. *LC-MALDI-MS(/MS)*

3 µL of the digest solution were loaded onto an Ultimate micro-capillary HPLC system equipped with a monolithic capillary PS-DVB column (200 µm x 5 cm, Dionex Ltd., Leeds, UK), and separated using a gradient elution of two solvents: solvent A: 2% acetonitrile (ACN), 0.1% heptafluorobutyric acid (HFBA, Fluka, Gillingham, UK) in H_2O; solvent B: 0.1% HFBA in ACN with a flow of 3 µL/min adopting the following chromatographic profile: 2 min 1% solvent B, linear gradient to 50% solvent B in 18 min, linear gradient to 85% solvent B in 1.1 min, 4.9 min wash at 85% solvent B, linear gradient to 1% solvent B in 1.1 min and finally 7.9 min reconditioning in 1% solvent B. The matrix solution (7 mg/mL α-cyano-4-hydroxy cinnamic acid diluted in 60/40 v/v ACN/H_2O, 2% wt/wt ammonium citrate/matrix) was automatically added on-line, with a flow rate of 1.6 µL/min, while spotting nano LC fractions onto a blank stainless steel LC MALDI sample plate using a Probot microfraction collector (Dionex). 180, 6 s fractions were spotted starting 3 min after injection, and allowed to dry at room temperature. Plates were loaded into an AB 4700 Proteomics Analyzer TOF/TOF (Applied Biosystems, Foster City, CA), mass spectrometer operated in positive ion reflectron mode. Mass spectrometric data were acquired over the *m/z* range 650-4400. Default calibration was used. The ten most intense ions from each

fraction were automatically subjected to MS/MS analysis, using the 4000 Series Explorer V3.6 software (Applied Biosystems).

S9b. ***Medium resolution LC-ESI-MS(/MS)***

Aliquots of 3 µL of the collagen tryptic digest solution in ammonium bicarbonate buffer were loaded onto an Ultimate nanoLC system (Dionex), controlled by the Chromeleon 6.60 software and equipped with a PepMap C18 trap (300 µm x 0.5 cm, Dionex) and an Onyx C18 monolithic silica capillary column (100 µm x 15 cm, Phenomenex). The separation used a gradient elution of two solvents (solvent A: 2% acetonitrile, 0.1% formic acid in H_2O; solvent B: acetonitrile, 0.1% formic acid) with a flow of 1.1 µL/min, column temperature of 60°C and the following chromatographic profile: initial conditions 100% solvent A, linear gradient to 50% solvent B in 30 min, 5 min wash at 95% solvent B and finally 11 min reconditioning in 100% solvent A. The HPLC was interfaced with a QSTAR API Pulsar *i* LC-MS/MS System (Applied Biosystems) with a MicroIon Spray source (fitted with 20 µm I.D. capillary). Positive ESI- MS and MS/MS spectra were acquired using information-dependent acquisition (IDA). Instrument control, data acquisition and analysis were performed using the Analyst QS v1.1 software. Settings were: ion spray voltage: 5200 V, nebulising gas: 5, curtain gas: 35, ion acquisition range: *m/z* 350-1800. IDA settings were: MS: 1 s (acquisition of survey spectrum), MS/MS: 1 s on first and second most abundant ions that meet switch criteria, cycle time: 3 s, the collision energy setting (CE) was automatically calculated using the IDA CE Parameter Table. Switch Criteria were: ion acquisition range: *m/z* 350-1800, charge state: 2 to 4 exceeding 20 counts, switch after one spectrum. Former target ions were excluded for 60 s. The ion mass accuracy tolerance was 100 ppm.

S9c. ***LC-ESI- high resolution MS(/MS)***

Samples of 1 µL of the collagen tryptic digest solution in 0.1% aqueous TFA were loaded onto an Ultimate 3000 nanoLC system (Dionex) equipped with a PepMap100 5µm C_{18} trap (300 µm x 0.5 cm, Dionex) and a PepMap100 3µm C_{18} capillary column (75 µm x 15 cm, Dionex). The trap wash solvent was 0.1% aqueous TFA. The separation used a gradient elution of two solvents (solvent A: 0.1% formic acid in H_2O; solvent B: acetonitrile, 0.1% formic acid) with a flow of 0.3 µL/min, column temperature of 30°C and the following chromatographic profile: initial conditions 5% solvent B (5 min), linear gradient to 35% solvent B in 90 min, then a wash 90% solvent B for 10 min, immediately returning to starting conditions, and finally 20 min reconditioning in 5% solvent B. The LC was interfaced with a maXis LC-MS/MS system (Bruker

DOI: 10.1038/NCHEM.899

Daltonics, Bremen, Germany) with a nanospray source (fitted with Proxeon stainless steel needle 30 µm x 5 cm). Positive ESI-MS and MS/MS spectra were acquired using AutoMSMS mode. Instrument control, data acquisition and analysis were performed using the Compass 1.3 SR1 beta software. Settings were: ion spray voltage: 1450 V, dry gas: 8.5 l/min, 150°C, ion acquisition range: *m/z* 50-3000. AutoMS^n settings were: MS: 0.5 s (acquisition of survey spectrum), MS/MS: 0.6 s for precursor intensities above 10000 counts, for signals of lower intensities down to 1000 counts, acquisition time increased linearly to 2 s, ion acquisition range: *m/z* 400-1400, (lockmass *m/z* 1221.9906 excluded), five precursor ions exceeding 1000 counts, preferred charge states: 2 and 3. The collision energy and isolation width settings were automatically calculated using the AutoMSMS Fragmentation Table. Former target ions were excluded for 20 s.

S9d. ***Database searching***

The MS/MS peak list from the MALDI experiment was generated filtering at a minimum signal to noise ratio of 20. The MS/MS peak list from the QSTAR mass spectrometer was obtained directly from the IDA files using the vendor provided Mascot script (v1.6b23). The maXis MS/MS peak list was generated using the ProcessAutoMSn method in DataAnalysis (v4.0, Bruker) with the following parameters; peak width: 5, minimum intensity: 50 and s/n: 2. Peak lists were searched on a local server using the Mascot 2.2 algorithm (Matrix Science Ltd., London, UK), selecting trypsin as the proteolytic enzyme and allowing one missed cleavage. Oxidation (Met, Pro and Lys) and deamidation (Asn, Gln) were selected as variable modifications. Mass tolerance for precursor ions was set at 100 or 10 ppm and for fragment ions at 0.1 or 0.02 Da, to search medium and high resolution data respectively. The data were matched against the NCBI non-redundant protein database (NCBInr).

Searching against in-house 'Collagens' database

Mass spectra matching, even partially or with low scores, with peptide sequences containing the Gly-Xaa-Yaa collagen motif were selected for manual *de novo* sequencing. The new sequences obtained with this approach were uploaded onto an in-house database of Type I, II and III collagens, containing some 150 partial and complete sequences, either obtained from public databases, EST libraries or direct sequencing together with all possible tryptic peptide sequences generated from the permutation of all the positions occupied by Ile/Leu and hydroxyproline (Hyp). The high resolution maXis instrument resolution allowed discrimination of Ile/Leu from Hyp (Δmass = 0.03638 Da).

Supplementary Table 1. Actin peptides identified by high-resolution mass spectrometry and manual *de novo* sequencing.

Observed	Mr(expt)	Mr(calc)	ppm	Score	Expect	Peptide
566.7656	1131.5167	1131.5197	-2.57	65	0.00022	GYSFTTTAER
586.3206	1170.6266	1170.6318	-4.49	64	0.0004	EITALAPPTMK
895.9470	1789.8794	1789.8846	-2.93	94	6.9e-07	SYELPDGQVITIGNER
652.0248	1953.0527	1953.0571	-2.26	78	1.3e-05	VAPEEHPVLLTEAPLNPK
739.6848	2216.0327	2216.0419	-4.17	109	1.8e-08	DLYANTVLSGGSTMYPGIADR

Supplementary Table 2: Frequency of hydroxylation in peptides. 33% (7 of 21) prolines are hydroxylated in the Xaa position and 100% in the Yaa position. Numbers in brackets indicate assignment from MS data, from which the hydroxyproline isomer is not determined.

	Xaa			Yaa			Ratios from Amino acid analysis	
	Pro	3-Hyp	4-Hyp	Pro	3-Hyp	4-Hyp	$\frac{\text{Pro}}{\text{Hyp}}$	$\frac{\text{4-Hyp}}{\text{3-Hyp}}$
Gly	14	5(2)	-	-	-	18 (4)	1.33	3
	14	7				22	0.48	3

S.10. ***Silica condensing activity of collagens*****.**

To 0.1 mg mL^{-1} of corresponding collagen solution in 0.1 M sodium phosphate buffer pH7 was added 4.5 mM sodium silicate (ABCR GmbH, Germany), Samples were incubated at room temperature for 24h. Precipitated silica was collected by centrifugation and then assayed using the Silicate (Silicic Acid) Assay Kit (Merck, Darmstadt, Germany) as described previously (10). Fibrillar calf skin collagen (Fluka) and Chondrosia reniformis sponge collagen (KliniPharm, Frankfurt/Main, Germany) were prepared as described by us previously [11, 12] and used as references.

S11. ***Ketal protection of hydroxyproline groups of collagen***

In order to perform ketal protection of 3- and 4-hydroxyproline residues, a standard procedure was used[15-20]. *H. sieboldi* spicular collagen (8.2 mg, Supplementary Figure 13(**1**)) was soaked in 200 μL of dimethylformamide (DMF) (Sigma) for 30 min. Next, 200 μL of 2-methoxypropene (Fluka) were added, followed by the addition of a solution of 5 mg of camphorsulfonic acid (Sigma) in 150 μL of DMF. The reaction was kept at ambient temperature for 3 h, during which time the colourless solution became orange, and then the solution was extensively washed with DMF and isopropanol (Fluka). The resulting solid (Supplementary Figure 13(**2**)) was filtered off and dried under vacuum.

DOI: 10.1038/NCHEM.899

Half of the reacted spicular collagen (Supplementary Figure 13(**2**)) was kept for FTIR analysis and for established silicification activity measurements [10]. The other half was cleaved from its ketal protecting group with a 95:5 (v:v) trifluoroacetic acid (Sigma-Aldrich) (TFA)/H_2O mixture. In particular, the protected collagen (Supplementary Figure 13(**2**)) was flushed with 1 mL of isopropanol (10 s) while on a vacuum filter, followed by flushing with 2 mL of TFA/H_2O 95:5 (v:v) mixture (20 s) and an intensive isopropanol wash. Additionally, 0.5 mL of 2.5 M NaOH (Fluka) solution was used for neutralization and to remove any remaining TFA. The cleaved sample was then centrifuged and finally dialyzed against deionized water on Roth (Germany) VISKING membranes (Typ 27/32) with a MWCO of 14 kDa. The dialyzed material (Supplementary Figure 13(**3**)) was dried at room temperature and analyzed using FTIR spectroscopy. Corresponding silicification activity measurements[10] using 0.1 mg mL^{-1} of collagen solutions were also carried out.

S12. ***FTIR spectroscopy***

Infrared spectra were recorded on a Nicolet 210 FT-IR Spectrometer (Nicolet Analytical Instruments, Madison, WI, USA) using Golden Gate Single Reflection Diamond ATR System. 500 scans were recorded at a spectral resolution of 2 cm^{-1}. The spectra were offset corrected for comparison purposes and the maxima of the spectra were normalized.

References

1. Vyalikh, D.V., Danzenbächer, S., Mertig, M., Kirchner, A., Pompe, W., Dedkov, Y.S& Molodtsov, S.L. Electronic structure of regular bacterial surface layers. *Phys. Rev. Lett.* **92**, 238103-238104 (2004).
2. Vyalikh, D.V., Maslyuk, V.V., Blüher, A., Kade, A., Kummer, K., Bredow, T., Mertig, I., Mertig, M. & Molodtsov, S.L. Charge transport in proteins probed by resonant photoemission, *Phys. Rev. Lett.* **102**, 098101 (2009).
3. Pines, A., Gibby, M.G& Waugh, J.S. Proton-enhanced NMR of dilute spins in solids. *J. Chem. Phys.* **59,** 569-590 (1973).
4. Metz, G., Wu, X. & Smith, S.O. Ramped-Amplitude Cross Polarization in Magic-Angle-Spinning NMR, *J. Magn. Reson. A* **110**, 219-227 (1994).
5. Fung, B.M., Khitrin, A.K. & Ermolaev, K. An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* **142**, 97-101 (2000).
6. Huster, D. Solid-State NMR Studies of Collagen Structure and Dynamics in Isolated Fibrils and in Biological Tissues, *Ann. Rep. NMR Spectrosc.* **64**, 127-159 (2008).
7. Langrock, T., Garcia-Villar, N. & Hoffmann, R. Analysis of Hydroxyproline Isomers and Hydroxylysine by Reversed-Phase HPLC and Mass-Spectrometry. *J. Chromatogr. B* **847**, 282-288 (2007).
8. Heinemann, S., Ehrlich, H., Douglas, T., Heinemann, C., Worch, H., Schatton, W. & Hanke, T. Ultrastructural studies on the collagen of the marine sponge *Chondrosia reniformis* Nardo. *Biomacromolecules* **8**, 3452-3457 (2007).
9. Heezen, B.C., & Schneider, E.D. Current on the Bermuda Rise. *Nature* **211**, 611-612 (1966).
10. Müller W. E G; Krasko A.; Le Pennec G.; & Schröder H. C. Biochemistry and cell biology of silica formation in sponges**.** Microsc. Res. Tech **62(4),** 368-377 (2003).
11. Heinemann S., Ehrlich H., Knieb C., Hanke T. Biomimetic inspired hybrid materials based on silicified collagen. *Int. J. Mat. Res.* **98**, 603-608 (2007).
12. Heinemann S, Ehrlich H, Douglas T., Heinemann C., Worch H., Schatton W., Hanke T. Ultrastructural studies on the collagen of the marine sponge *Chondrosia reniformis* Nardo. *Biomacromolecules* **8**, 3452-3457 (2007).
13. DeLano, W. L. DeLano Scientific, San Carlos, USA (2002).

DOI: 10.1038/NCHEM.899

14. Brunner, E., & Sternberg, U. Solid-state NMR investigations on the nature of hydrogen bonds. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 21-57 (1998).

15. Carey, F.A.; Sundbreg, R. J. Advanced Organic Chemistry. Part B: Reactions and Synthesis, 4th edn. Kluwer Academic/Plenum Publishers. New York, pp. 822-831 (2001).

16. Green, T. W.; Wuts, P. G. M. Protecting Groups in Organic Synthesis, 2nd ed.; John Wiley & Sons Inc: New York, 1991.

17. Kluge, A. F.; Untch, K. G.; Fried, J. H. Synthesis of Prostaglandin Models and Prostaglandins by Conjugate Addition of a Functionalized Organocopper Reagent. *J. Am. Chem. Soc.* **94**, 7827 (1972).

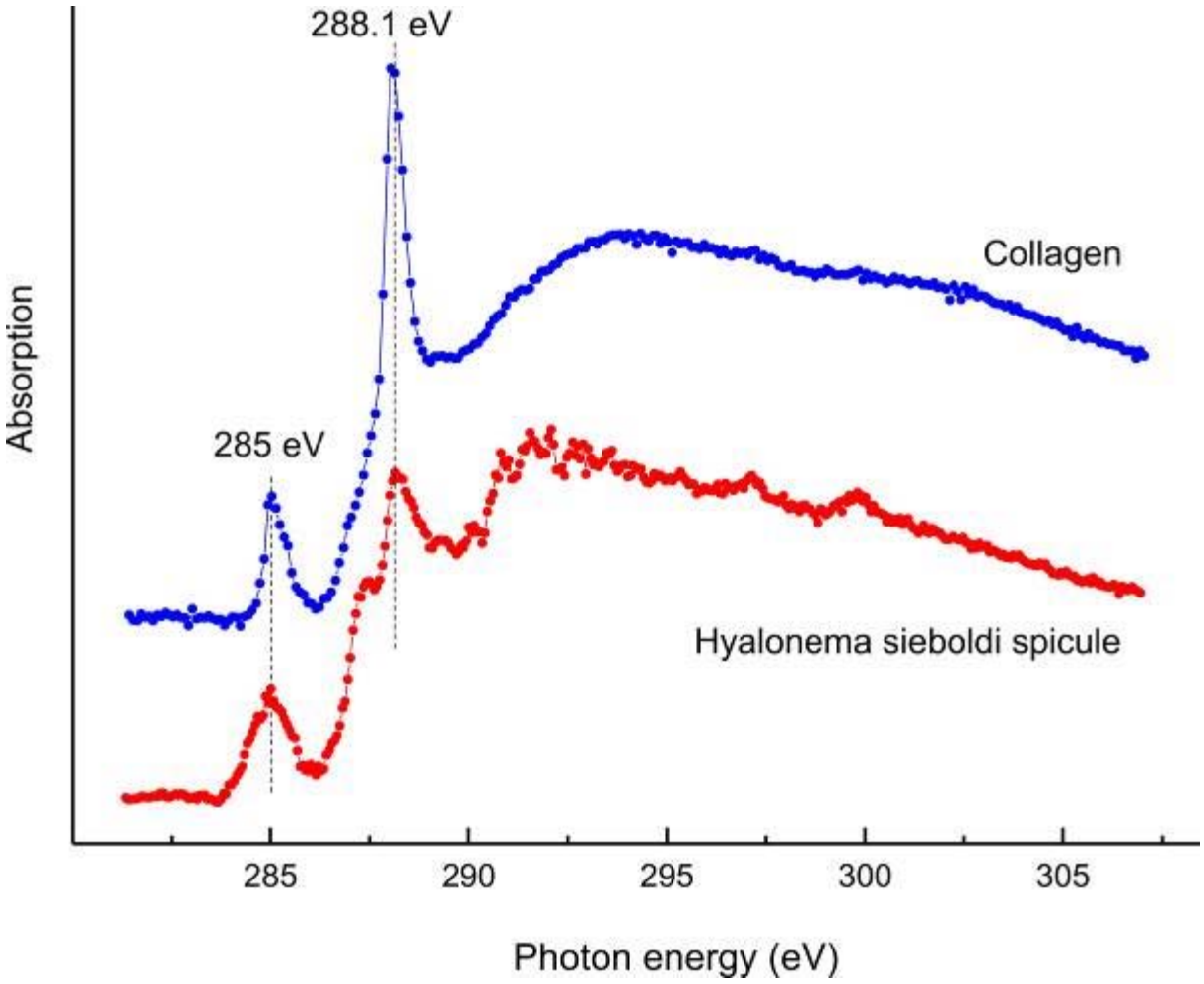
18. Effenberger, F. The Chemistry of Enol Ethers. *Angewandte Chemie (Int.Ed.)* **8(5)**, 295-312 (1969).

19. Orgueira, H. A., Bartolozzi A.B., Schell P., Litjens R.E.J.N., Palmacci E.R., Seeberger P.H. Modular synthesis of heparin oligosaccharides. *Chem. Eur. J.* **9**, 140-169 (2003).

20. Debost, J.L., Gelas, J., Horton, D., Selective preparation of mono-and diacetals of D-mannitol. *J. Org. Chem.* **48**, 1381-1382 (1983).

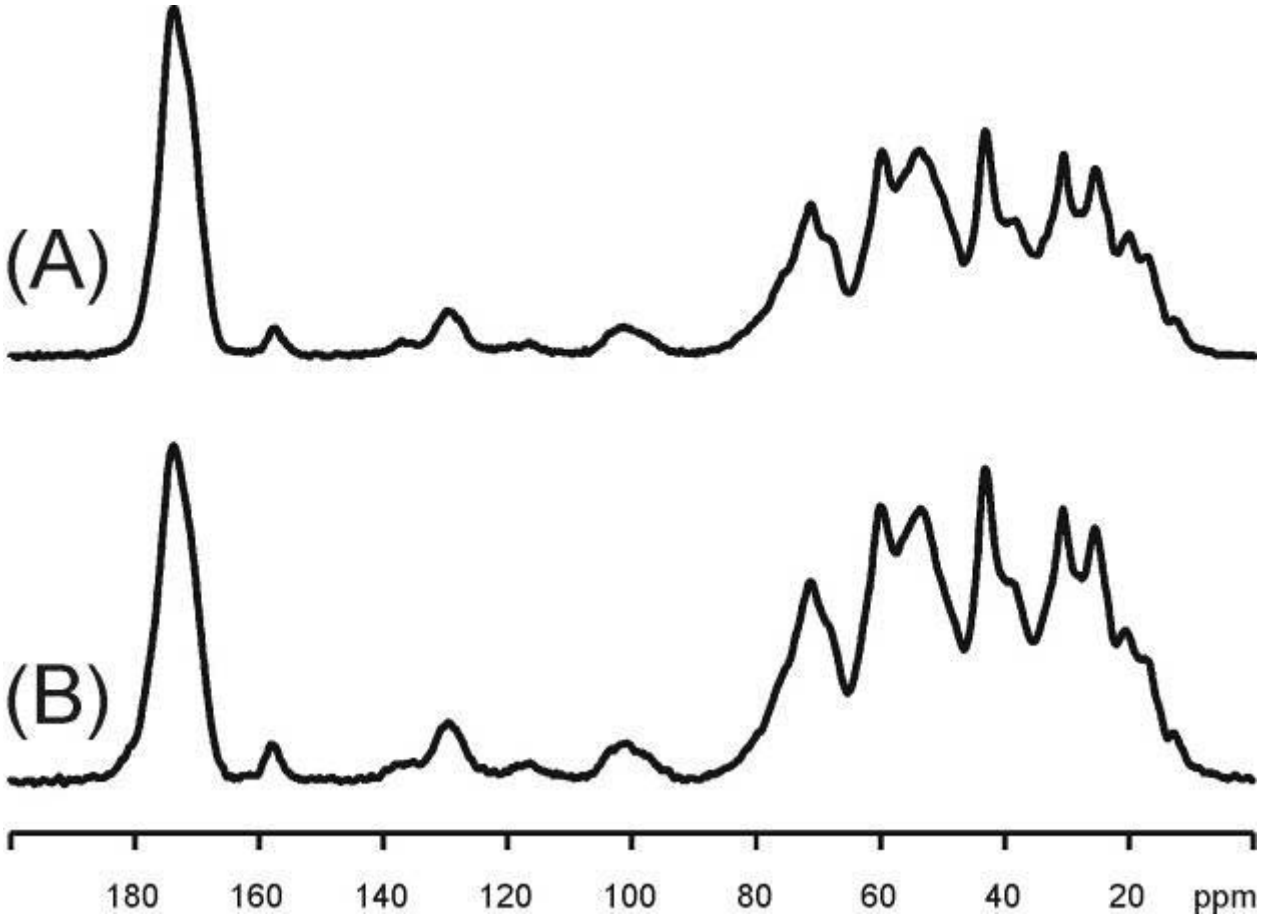
21. Socrates, G. Infrared and Raman Characteristic Group Frequencies: Tables and Charts. Wiley & Sons, p. 102 (2004).

Supplementary Figures

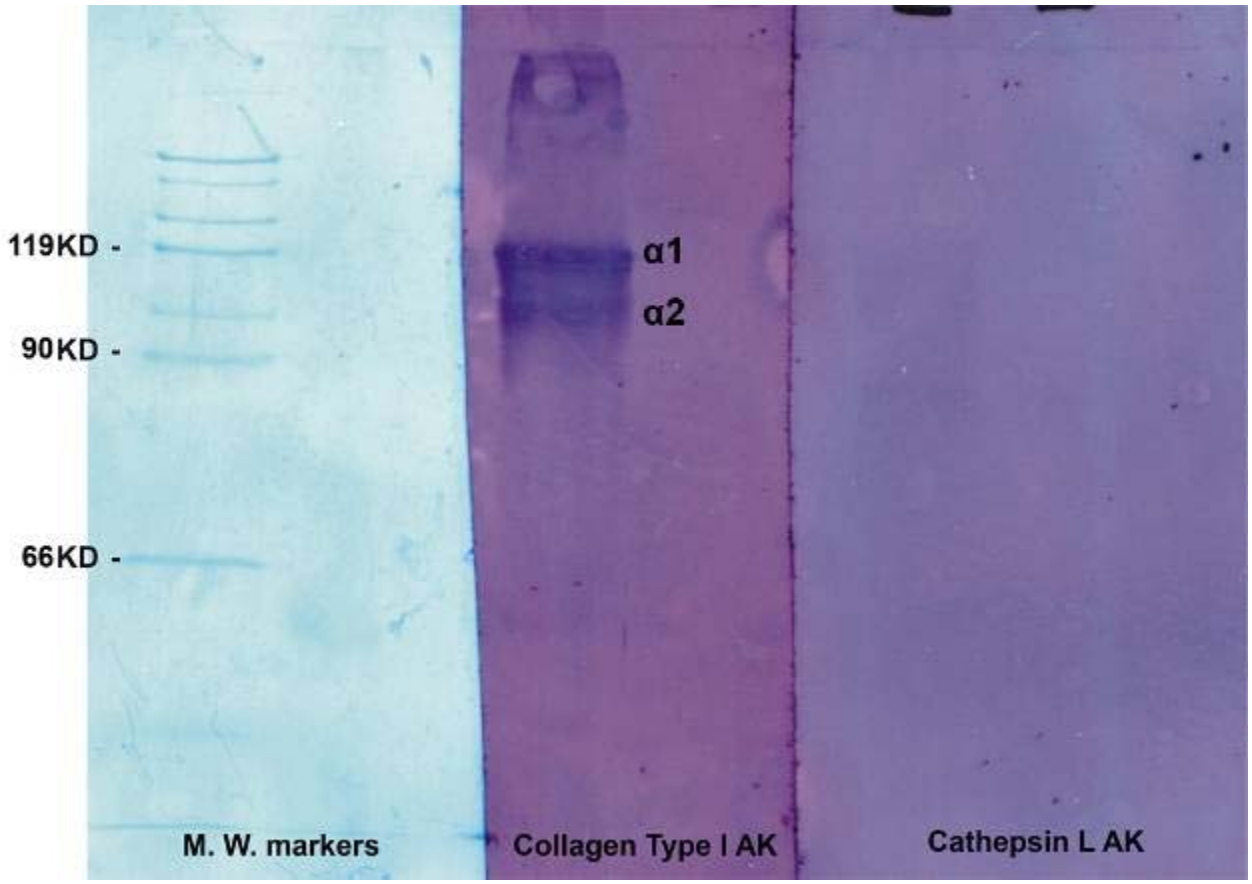


Supplementary Figure 1 | Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy experiments, performed at the C1s edge, provide evidence that the contribution of carbon is mainly due to organic components of *H. sieboldi* spicules. Moreover, the NEXAFS spectrum of the *H. sieboldi* spicule shows characteristic peaks similar to those obtained for calf collagen type I purchased from Sigma.

DOI: 10.1038/NCHEM.899



Supplementary Figure 2 | ^{13}C $\{^{1}H\}$ CP MAS NMR spectra of collagen: (a) Isolated according to[8] from the marine demosponge *Chondrosia reniformis*. (b) Isolated from the anchoring spicules of *H. sieboldi*.



Supplementary Figure 3 | Western blot analysis. Polyclonal antibodies detected type I collagen only in the basal spicules of *Hyalonema sieboldi* glass sponge.

DOI: 10.1038/NCHEM.899

G A Q G (P) L G P

G G F G L 4Hyp G R

G (V) D G N 4Hyp G I X G A T G S

G S V G 3Hyp 4Hyp G N 4Hyp G V Q G V S G 3Hyp

I G P D E P L K G K I

G I 4Hyp G P Q G F T G A I G V T G S 4Hyp G E I G A 4Hyp G

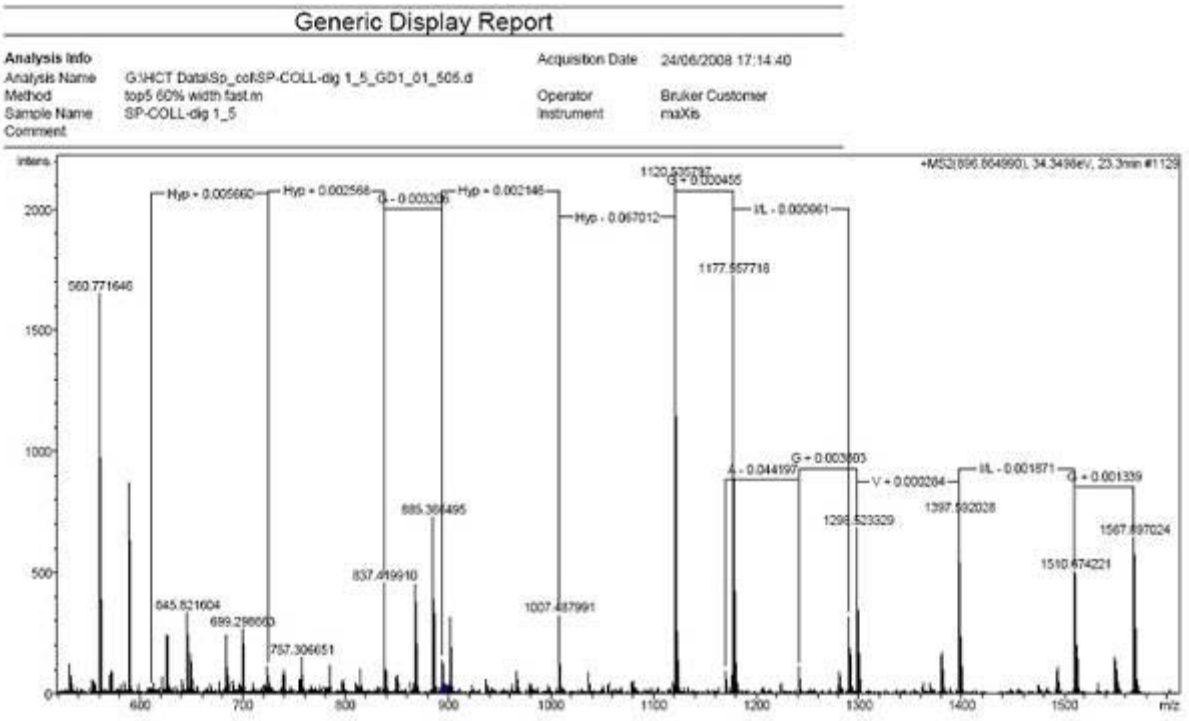
V G D 4Hyp G L V G D L G A Q G P Q G S Q G L V G

G A T G 3Hyp 4Hyp G I S G 3Hyp 4Hyp G P Q G Q 4Hyp G T 4Hyp G I

I G P A G P Q G Q 4Hyp G 3Hyp 4Hyp G P G G P X G 3Hyp 4Hyp

G G S G A 4Hyp G L 4Hyp G A I G N Q G A 4Hyp

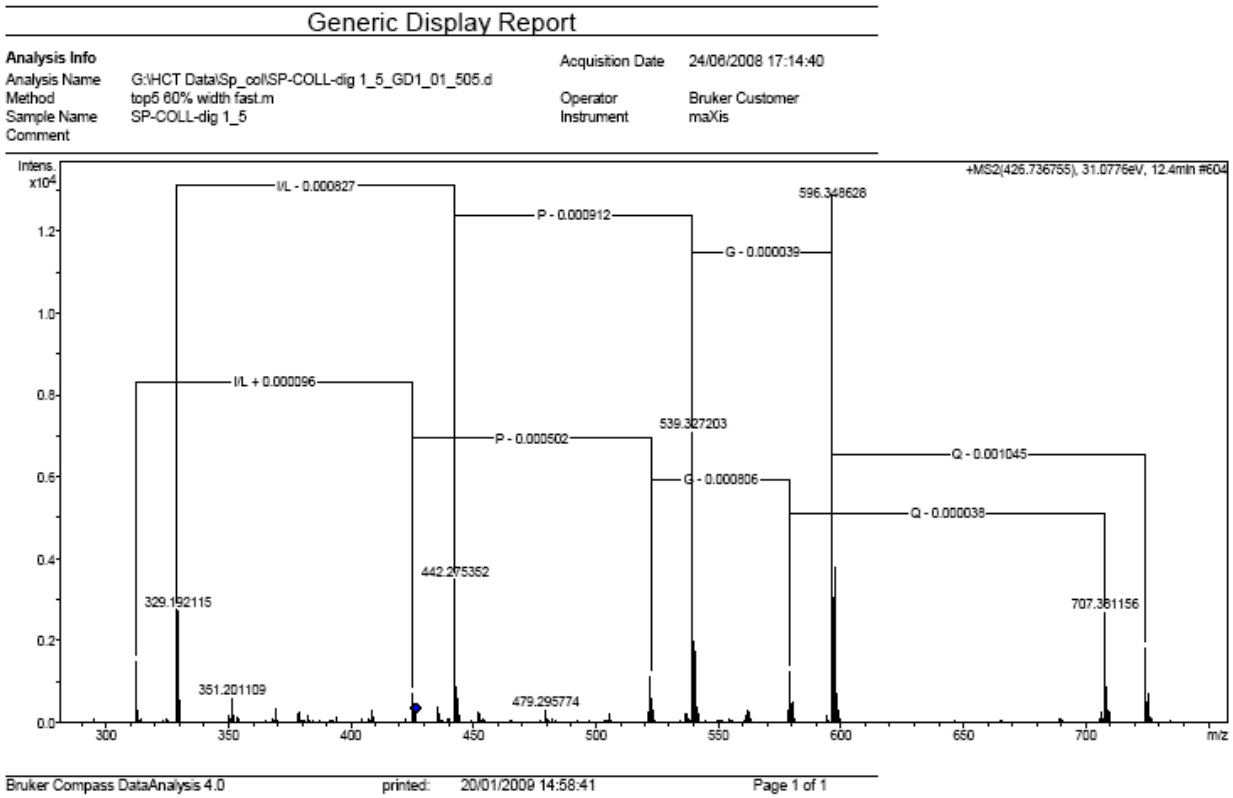
Supplementary Figure 4 | List of the peptides obtained on Edman degradative sequencing of peptides released and isolated from *H. sieboldi* spicule. Residues shown in parentheses are ambiguous.



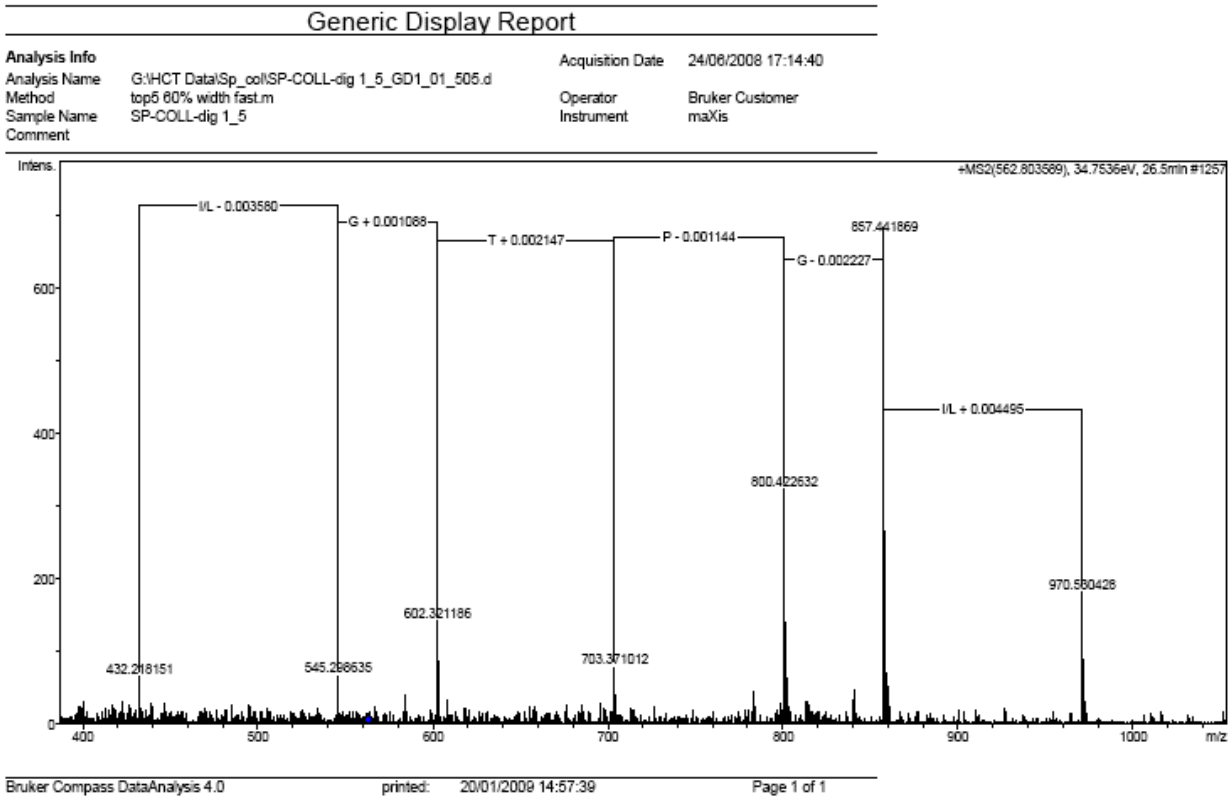
Database	Observed	Mr(expt)	Mr(calc)	ppm	Score	Expect	Peptide
Collagens	426.7374	851.4603	851.4613	-1.19	54	6.7e-06	GAQGPIGPR
NCBI	562.8033	1123.592	1123.5986	-5.87	77	0.000015	GPIGPTGLQGAR
Collagens	664.3041	1326.5937	1326.6052	-8.64	16	0.024	EGEJGUGETGPR
Collagens	832.9138	1663.8130	1663.8166	-2.13	86	8.9e-09	GUGAJGIDGNJGPAGIR
Collagens	896.4121	2686.2146	2686.2294	-5.51	66	1.1e-06	(MEGPT)GAPGA(JG)DAG VIGJJGJJG(PQGPR)

Supplementary figure 5 | Bruker maXis high-resolution product ion spectrum of *m/z* 896.4121 detected at retention time 23.3 min, showing partial amino acid residue assignments based on *m/z* increments between fragment ions. The y_7-y_{13} fragment ions support the confident identification of the 'Gly-Hyp-Hyp' motif.

DOI: 10.1038/NCHEM.899

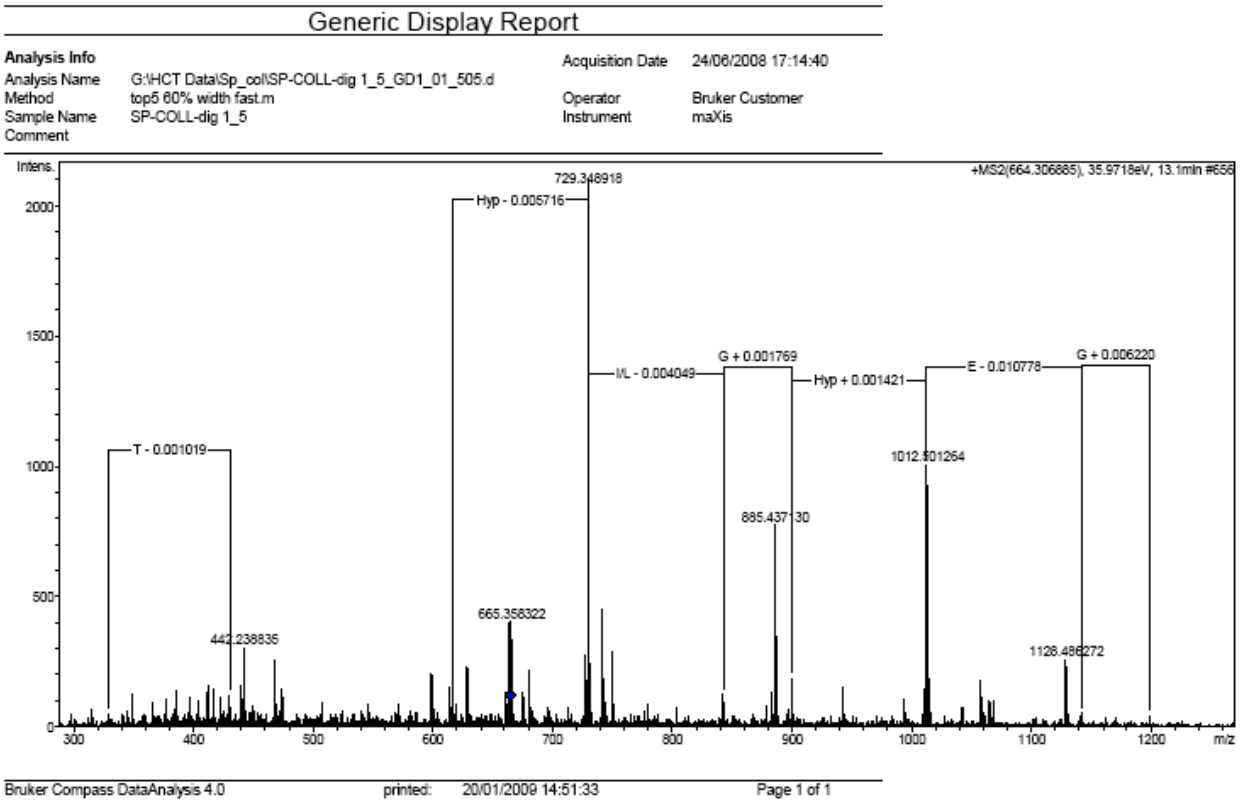


Supplementary Figure 6 | Bruker maXis high-resolution product ion spectrum of *m/z* 426.7374 detected at retention time 12.4 min, showing partial amino acid residue assignments based on *m/z* increments between fragment ions.

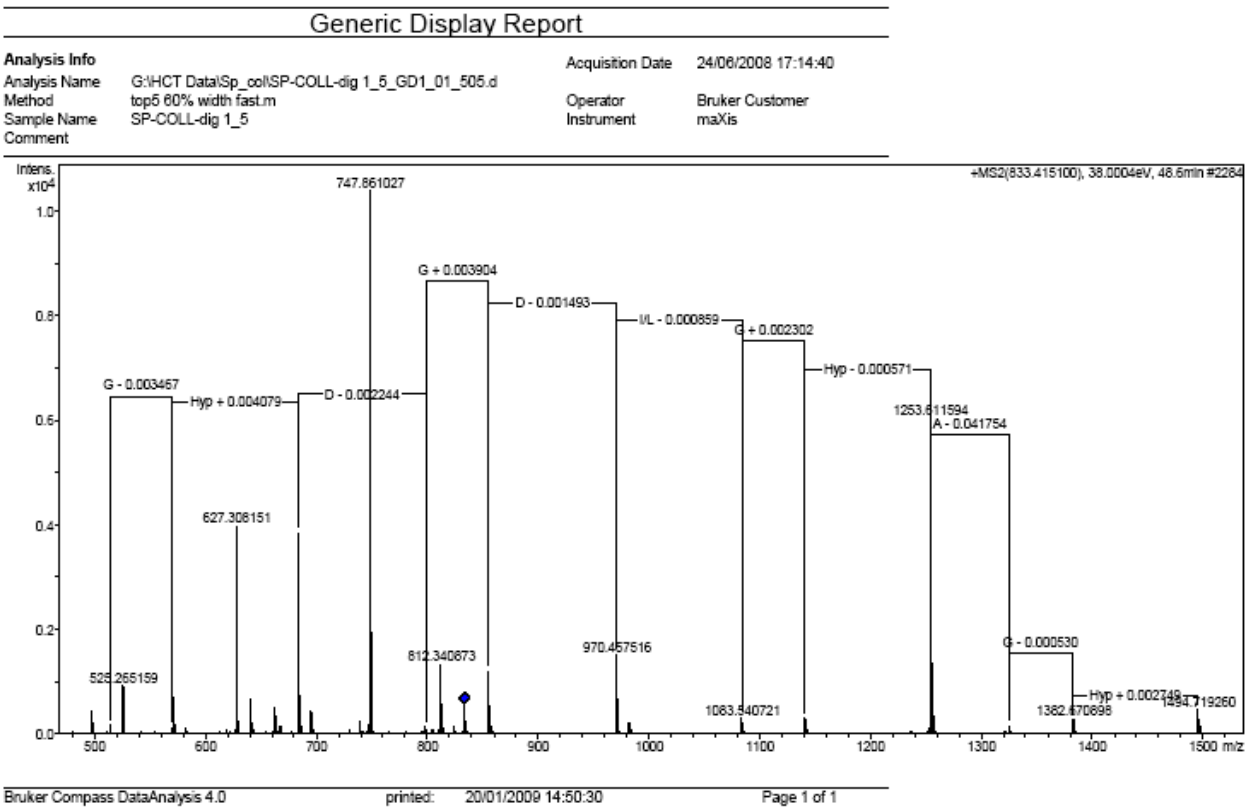


Supplementary Figure 7 | Bruker maXis high-resolution product ion spectrum of *m/z* 562.8033 detected at retention time 26.5 min, showing partial amino acid residue assignments based on *m/z* increments between fragment ions.

DOI: 10.1038/NCHEM.899

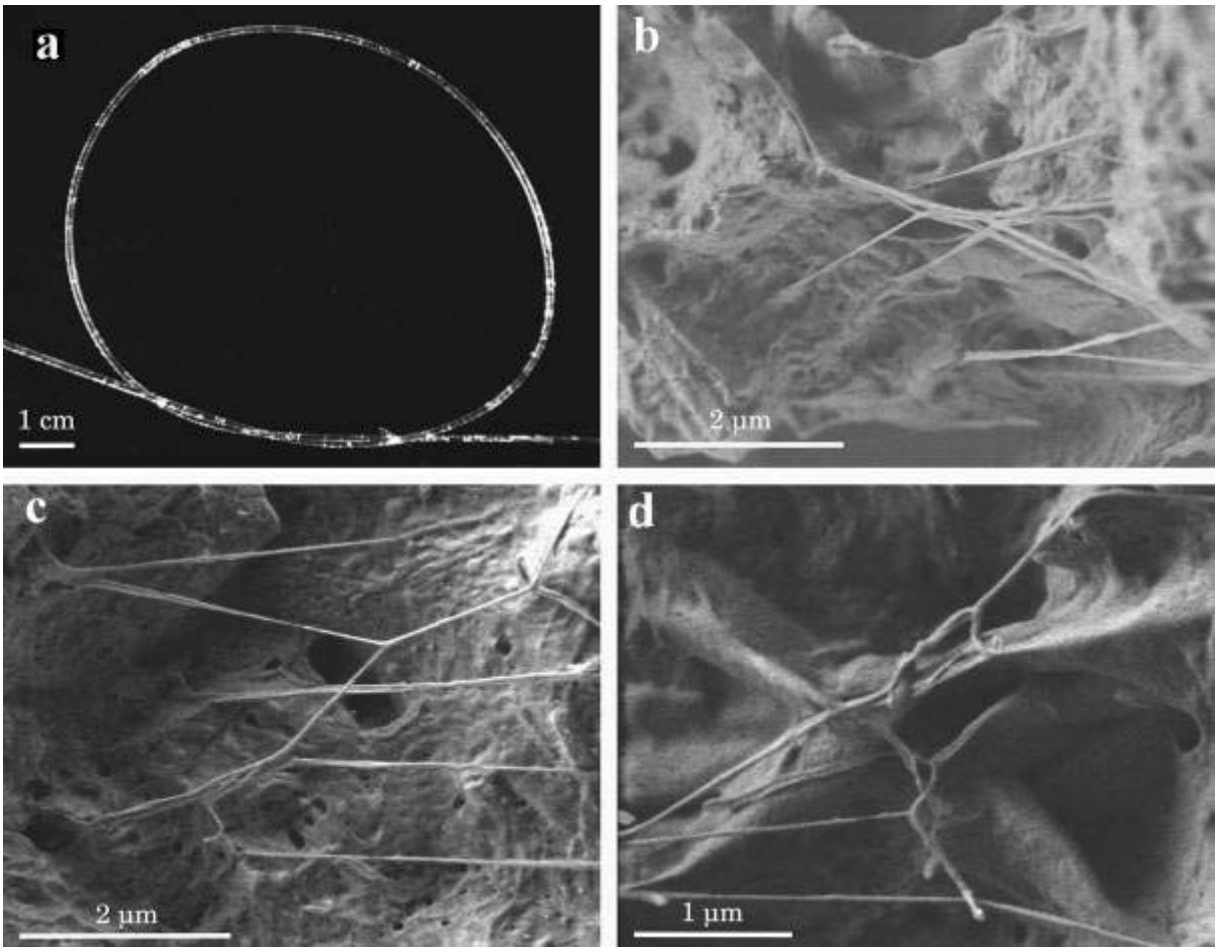


Supplementary Figure 8 | Bruker maXis high-resolution product ion spectrum of *m/z* 664.3041 detected at retention time 13.1 min, showing partial amino acid residue assignments based on *m/z* increments between fragment ions.

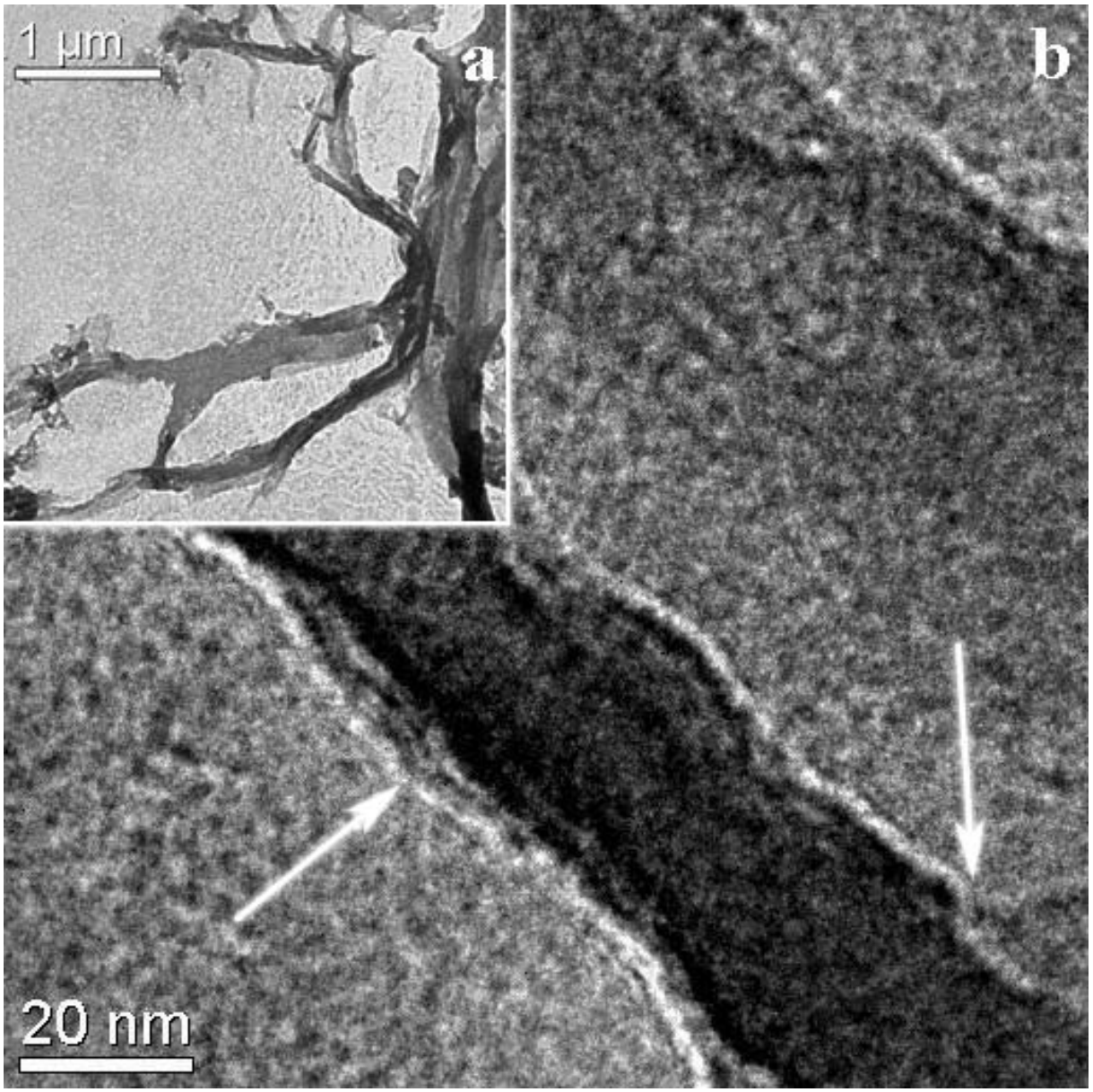


Supplementary Figure 9 | Bruker maXis high-resolution product ion spectrum of *m/z* 832.9138 detected at retention time 48.5 min, showing partial amino acid residue assignments based on *m/z* increments between fragment ions.

DOI: 10.1038/NCHEM.899

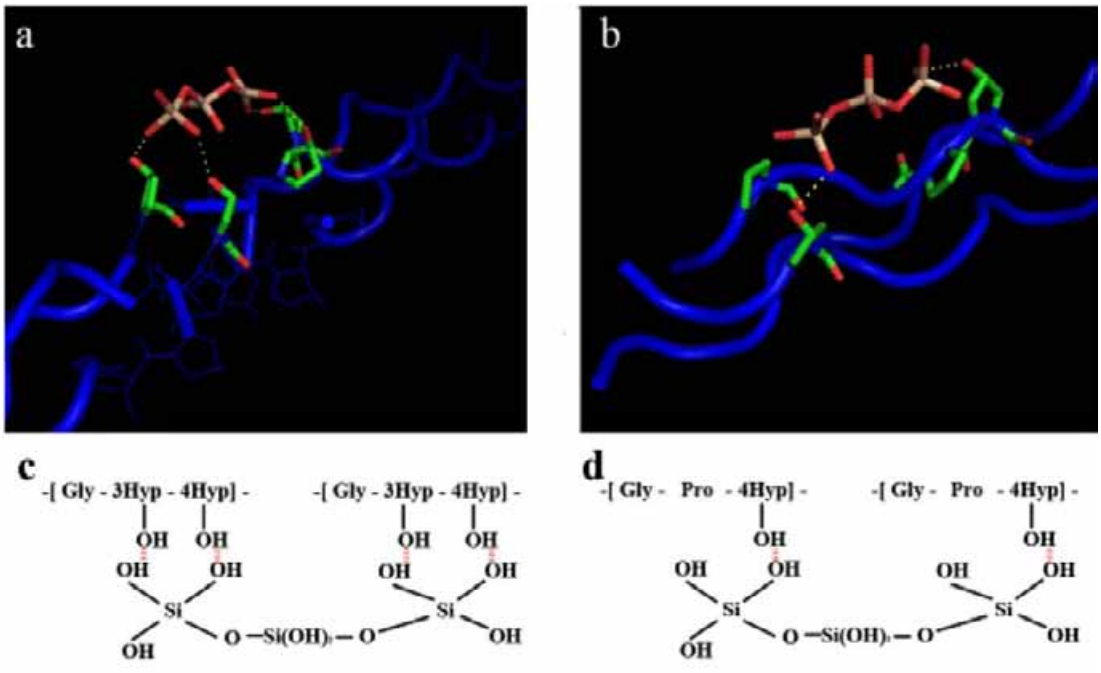


Supplementary Figure 10 | Unique flexibility of anchoring spicules of *H. sieboldi* glass sponge (**a**) is proposed to be possible due to plywood organization of collagen nanofibrils (**b, c, d**) observed within these partially demineralised silica-based biocomposites.

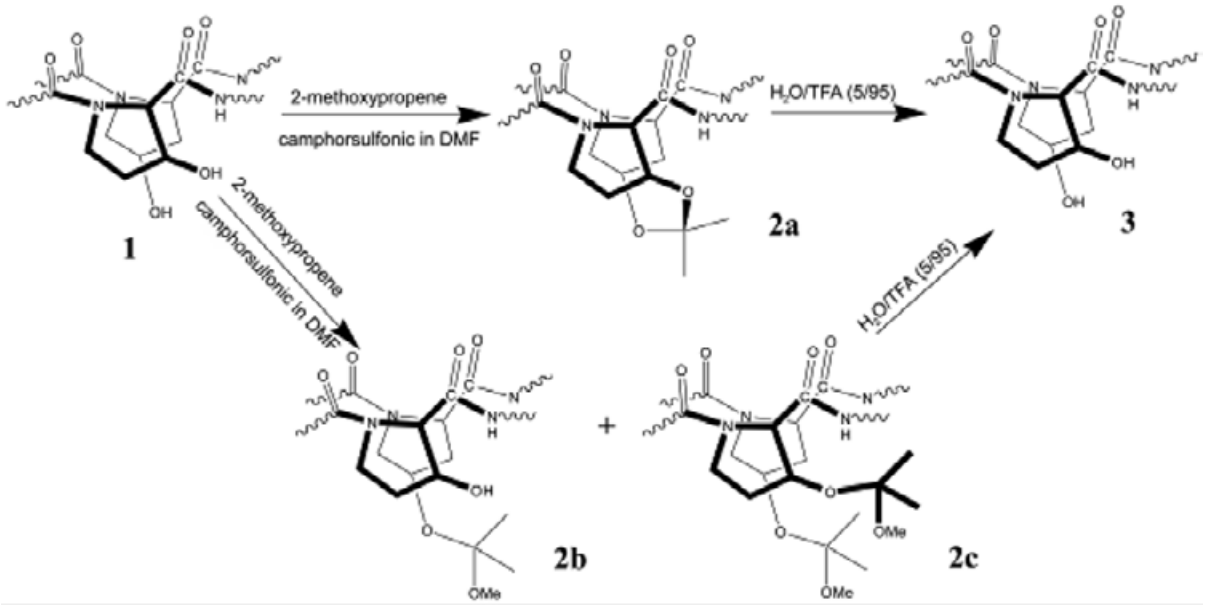


Supplementary Figure 11 | High-resolution transmission electron microscopy image of the fragment of *H. sieboldi* spicular microfibril (**a**); the arrows indicate the presence of nanofibrillar structures (**b**) the diameter of which corresponds to that of collagen triple helices (1.5 nm) which may indicate the presence of collagen.

DOI: 10.1038/NCHEM.899

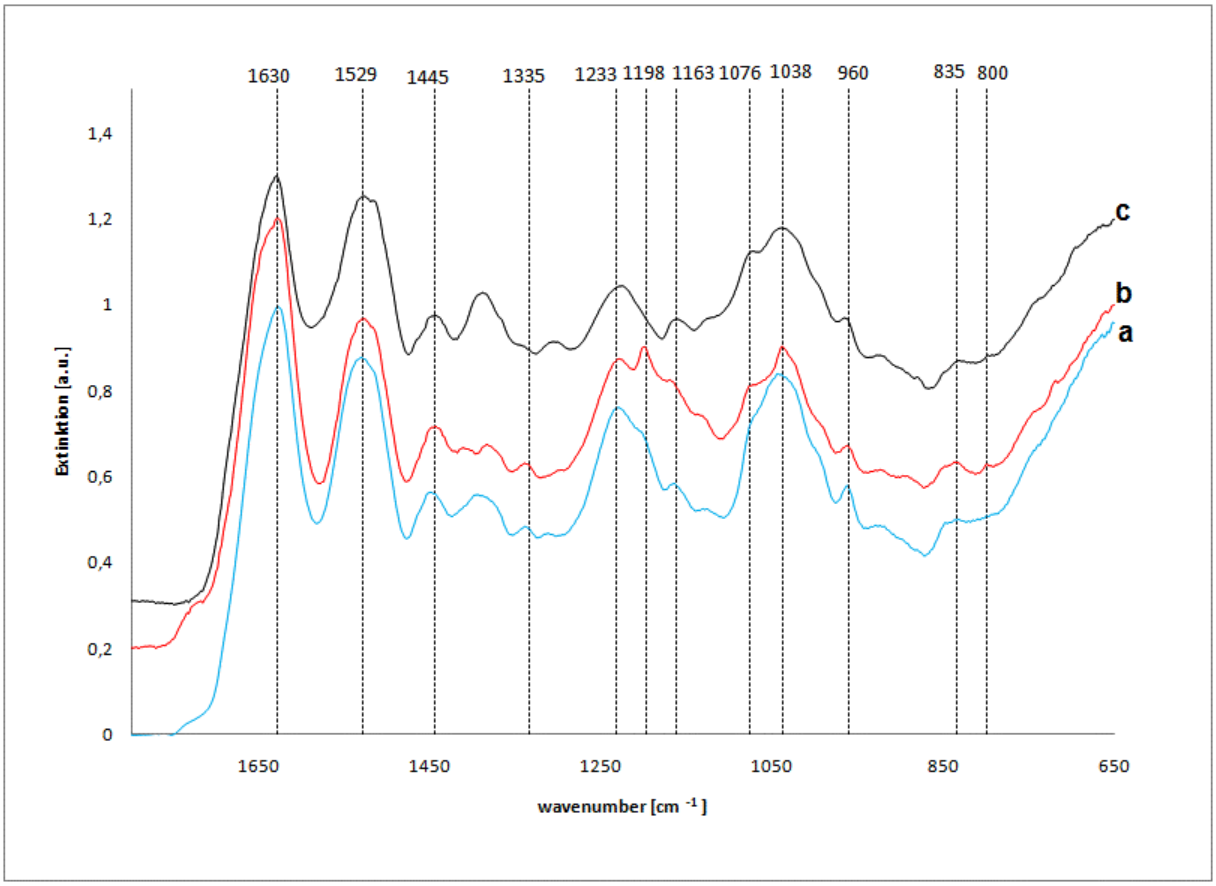


Supplementary Figure 12 | Model of interaction between polysilicic acid and Gly-3Hyp-4-Hyp polypeptide (**a**) and/or with Gly-Pro-4-Hyp polypeptide (**b**) located in collagen triple helix (blue). In the selected amino residues, carbon atoms are colored green, oxygen red, nitrogen blue, and silica orange. Mechanisms for these interactions are schematically represented in (**c**) and (**d**). Models were produced in PYMOL[13] on the basis of collagen-like polypeptide taken from Protein Data Bank (ID 2G66). Interactions were modeled in accordance with optimal distances for hydrogen bonding[14].

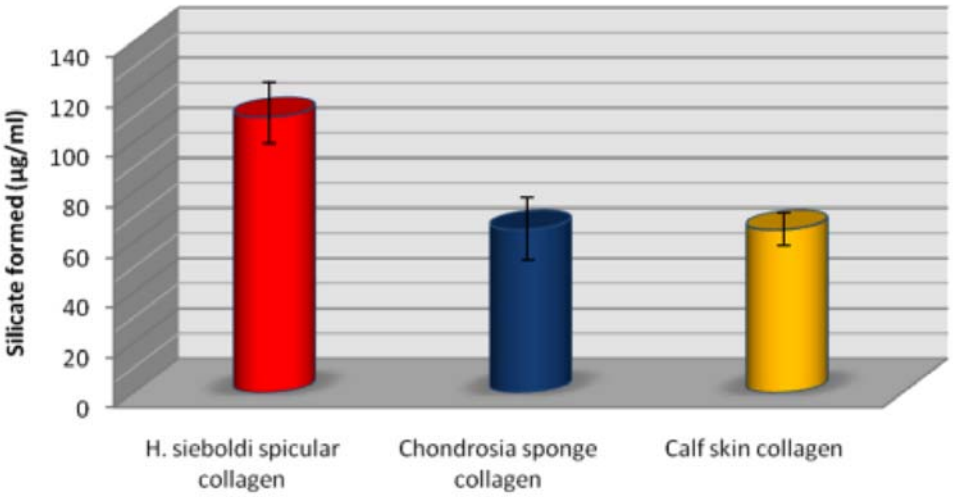


Supplementary Figure 13 | Schematic view of the reaction pathways described in **S.11**.

DOI: 10.1038/NCHEM.899

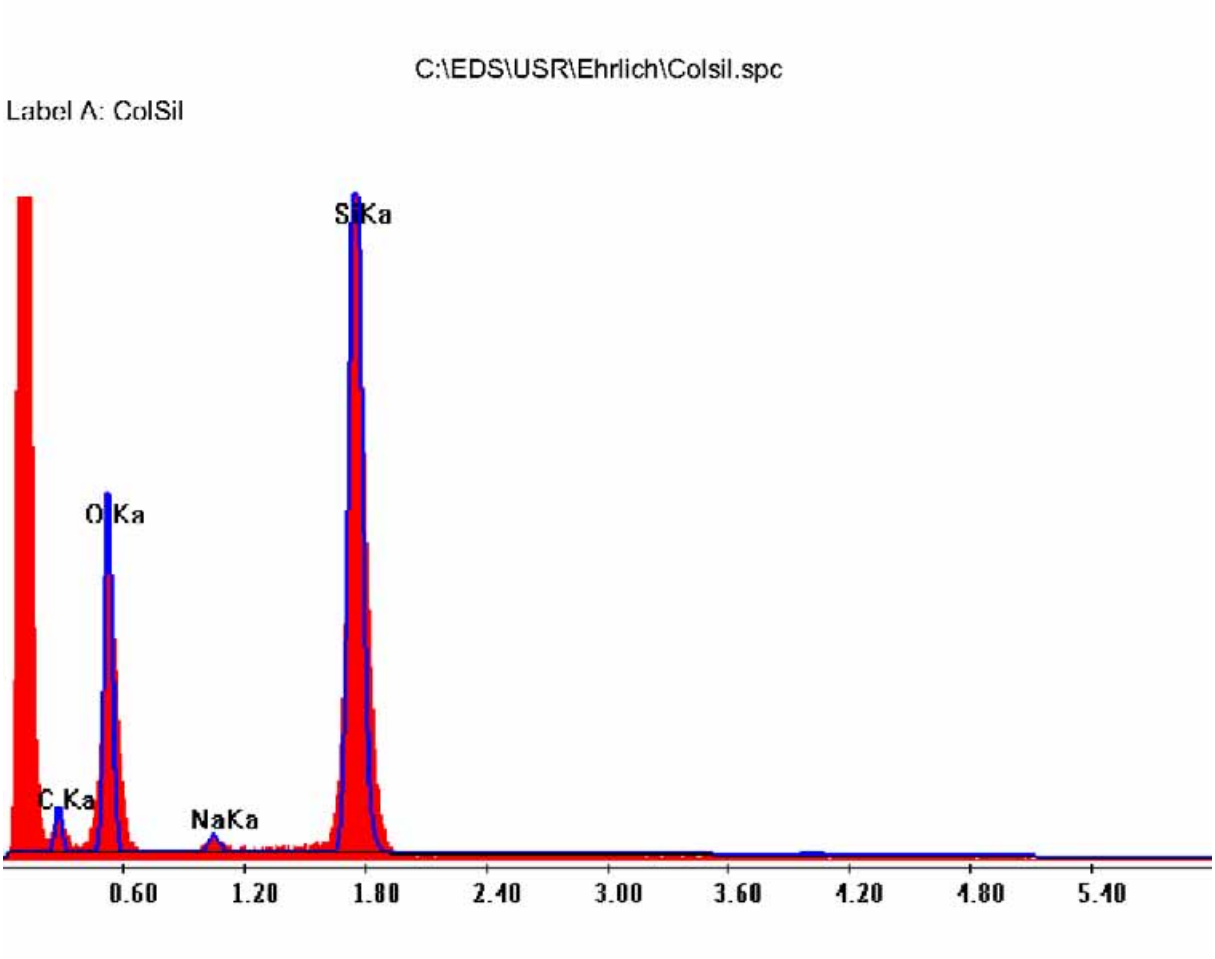


Supplementary Figure 14 | FTIR spectroscopy of *H. sieboldi* spicular collagen: a – un-treated spicular collagen; **b** – protected spicular collagen; **c**- spicular collagen cleaved from ketal protection group. The appearance of the ketal group with characteristic peak vibrations[21] at 1038 and more strongly at 1198 cm^{-1} in the IR spectrum of protected collagen (**b**), gives strong evidence for the occurence of this reaction. The removal of the ketal groups using the TFA-based procedure is confirmed by the disappearance of the corresponding IR peaks (**c**).



Supplementary Figure 15 | Silica condensing activity of *H. sieboldi* spicular collagen in comparison with mesohylar collagen[12] from demosponge *Chondrosia reniformis* and calf skin collagen. The average of ten separate incubations, with error bars, is shown.

DOI: 10.1038/NCHEM.899



Supplementary Figure 16 | EDX analysis of the sample represented in Fig.3d, e.