Electronic Supplementary Material (ESI) for Journal of Materials Chemistry B.


# Supplementary Information

Functional mimicry of sea urchin biomineralization proteins with $CaCO_3$-binding peptides selected by phage display

Elke Völkle (nee Evgrafov), Fabian Schulz, Julia Maxi Kanold, Monika Michaelis, Kerstin Wissel, Franz Brümmer, Anna S. Schenk, Sabine Ludwigs, Joachim Bill, Dirk Rothenstein*

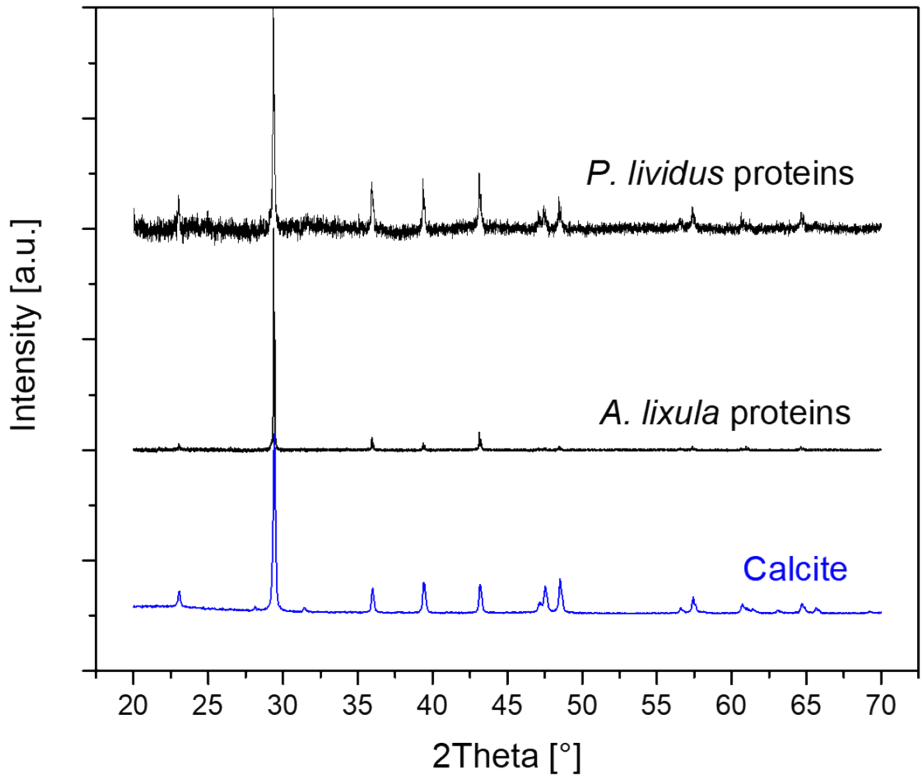
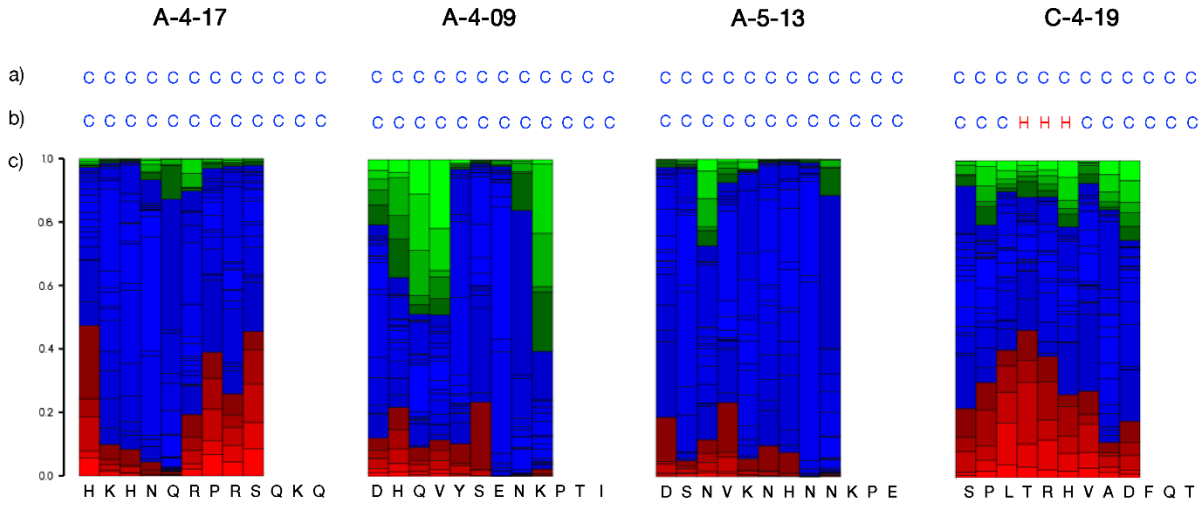
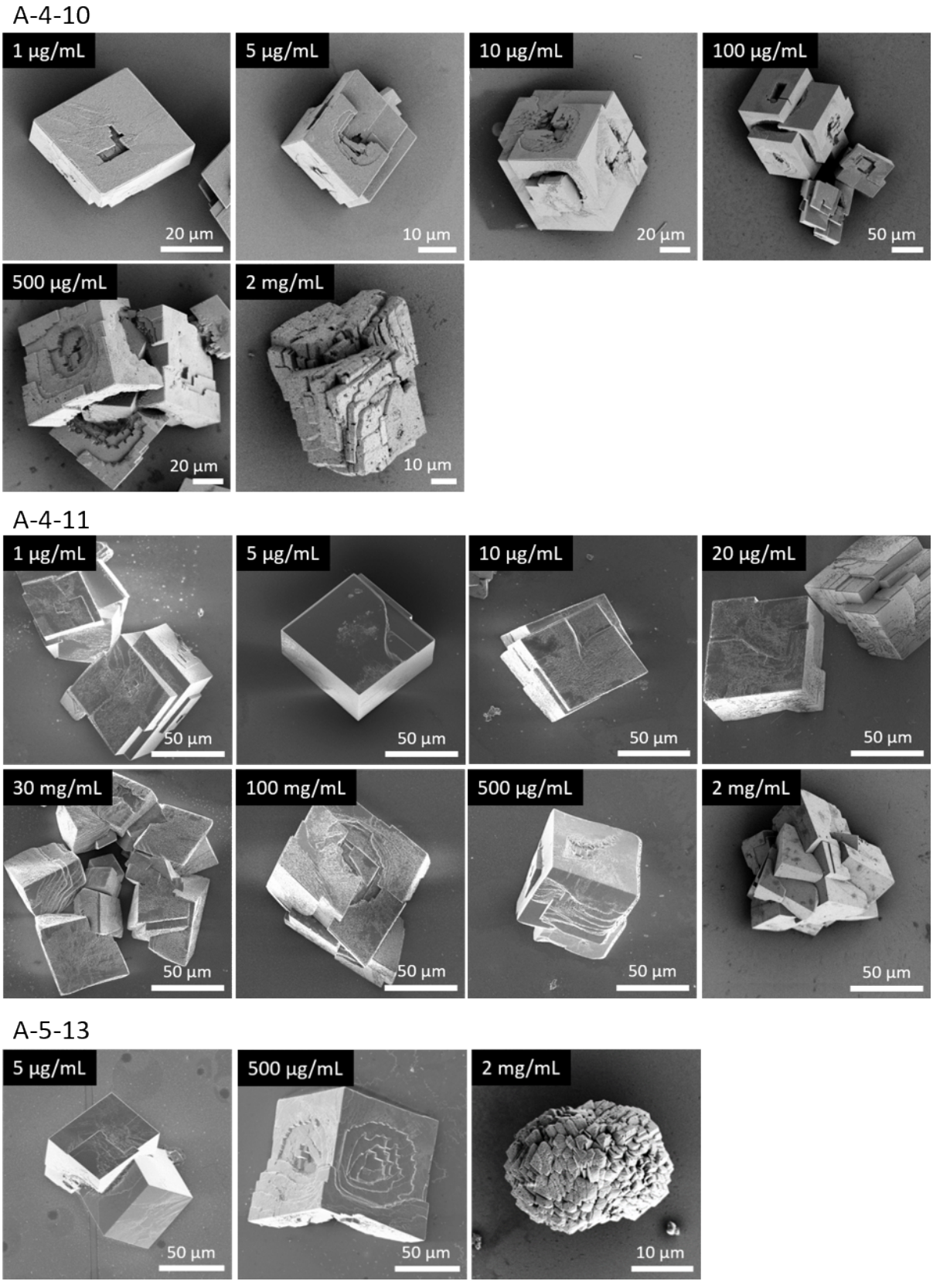



**Figure 1.** $CaCO_3$ polymorph induced by sea urchin proteins. The polymorph of mineralized $CaCO_3$ in the presence of *Arbacia lixula* proteins and *Paracentrotus lividus* proteins was determined by X-ray diffraction (XRD). The limited amount of mineralization products, led to the direct measurement of mineralization products on the chamber slides. XRD signals identified calcite as the predominant polymorph (see Figure 7 for comparison. For comparison the diffractogram of geological calcite (blue lines) (Alpha Aesar, Product no. 44520, LOT N04A030) is shown.

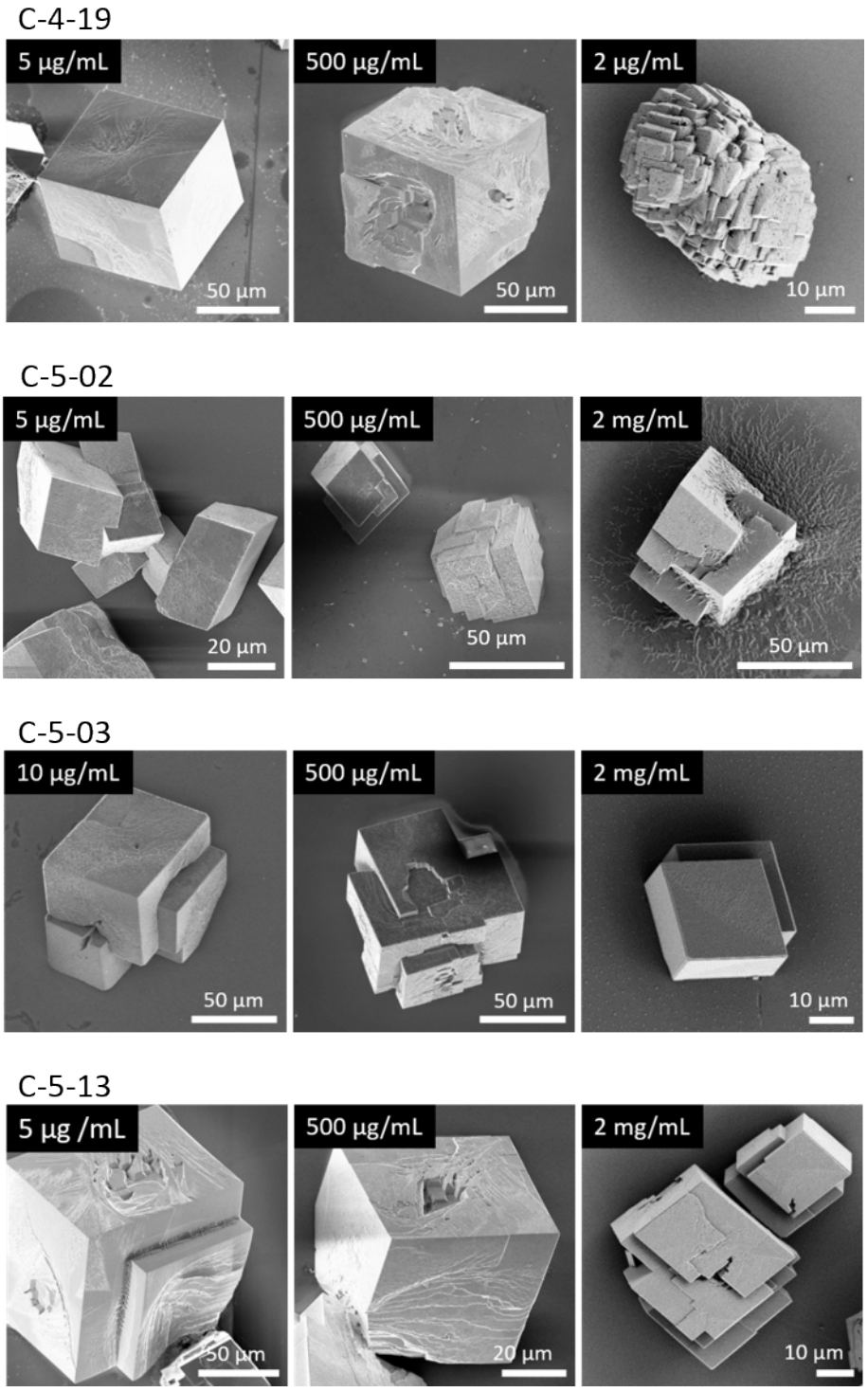



**SI Figure 2.** Secondary structure prediction of $CaCO_3$-binding peptides. The secondary structure of the four most potent peptides was predicted using various computational models. Notably, all peptides exhibit an unordered conformation. a) Protein Surface Accessibility and Secondary Structure Predictions (NetSurfP 3.0),[1] b) Protein Secondary Structure Prediction Server (PSSpred),[2] and c) De Novo Secondary Prediction with Pep-Fold [5-7].[3-5] The de novo prediction in Pep-Fold 3.0 is based on a four-residue-length structural alphabet (SA) letters (c.f. Methods), therefore the profile is of the size if the peptide length minus three residues. The results are represented using a one-letter code per amino acid position within the 12-mer peptides, with secondary structure elements indicated as Helix (H) in red, Sheet (E) in green, Turn (T) in black, and Coil (C) in blue. Pep-Fold results illustrate the likelihood of each amino acid adopting a secondary structure conformation using the same color code.

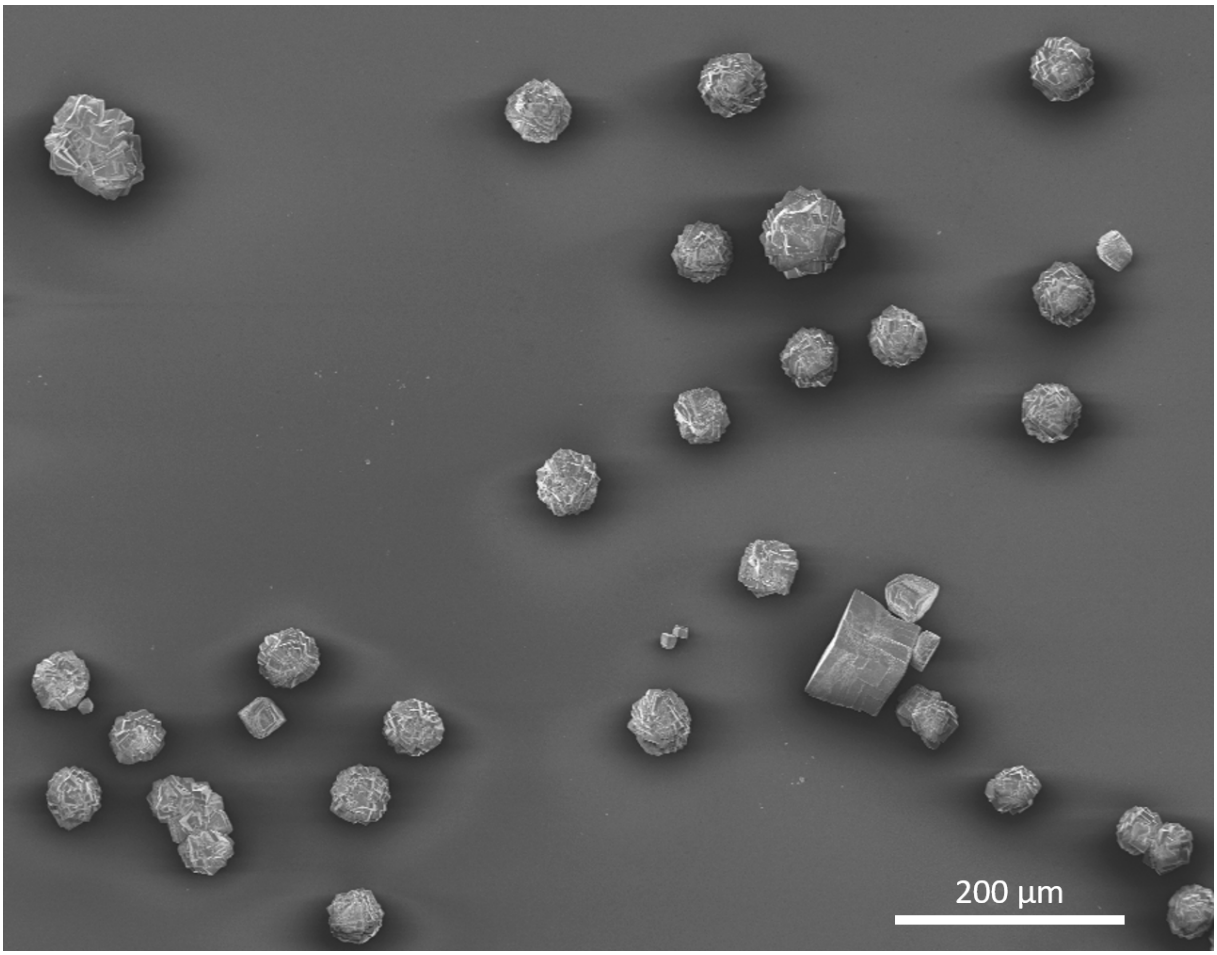

1. M. H. Høie, E. N. Kiehl, B. Petersen, M. Nielsen, O. Winther, H. Nielsen, J. Hallgren and P. Marcatili, *Nucleic Acids Res.*, 2022, **50**, W510-W515.
2. R. Yan, D. Xu, J. Yang, S. Walker and Y. Zhang, *Sci Rep-Uk*, 2013, **3**, 2619.
3. P. Thévenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux and P. Tufféry, *Nucleic Acids Res.*, 2012, **40**, W288-W293.
4. Y. Shen, J. Maupetit, P. Derreumaux and P. Tufféry, *J Chem Theory Comput*, 2014, **10**, 4745-4758.
5. A. Lamiable, P. Thévenet, J. Rey, M. Vavrusa, P. Derreumaux and P. Tufféry, *Nucleic Acids Res.*, 2016, **44**, W449-W454.



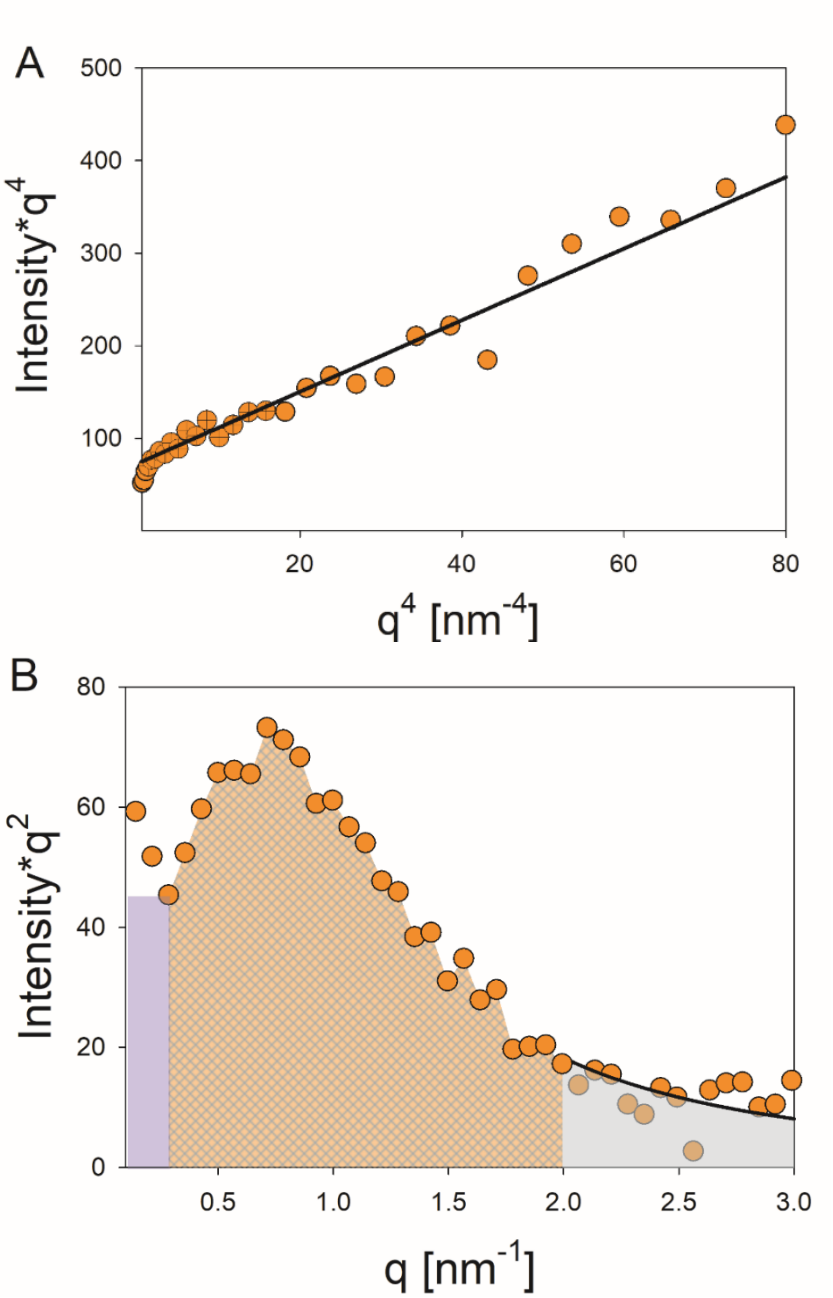

**SI Figure 3.** Effect of aragonite-binding peptides on the particle morphology in bioinspired mineralization approaches. Peptides were applied in concentrations between 1 µg/mL and 2 mg/mL.




**SI Figure 4.** Effect of calcite-binding peptides on the particle morphology in bio-inspired mineralization approaches. Peptides were applied in concentrations between 1 µg/mL and 2 mg/mL.




**SI Figure 5.** SEM overview image of $CaCO_3$ particles mineralized in the presence of peptide A-04-09. The peptide was added to the reaction mixture at a concentration of 10 µg/ml. The mineralized calcite particles were highly homogeneous in shape and size. Scale bar 200 µm.




**SI Figure 6.** T-parameter analysis based on SAXS data obtained from $CaCO_3$ particles mineralized in the presence of peptide A-04-09. A) Porod plot ($I(q)\ q^4$ vs. $q^4$) of the data points (orange symbols), where linear regression (black solid line; data points used for regression indicated by crossed symbols) yields the Porod constant P as the intercept (with $I(q) = \frac{P}{q^4} + I_{Laue}$). B) Kratky plot ($I(q)\ q^2$ vs. $q$) of the data points after subtraction of the Laue background $I_{Laue.}$ The area under the Kratky curve, which corresponds to the integral intensity $\tilde{I}$ was determined as follows: The data points with q < 0.3 $nm^{-1}$ (where the signal is dominated by the contribution of the external facets of the crystals) were approximated by $I(q)\ q^2 = I(0.3\ nm^{-1})\ (0.3\ nm^{-1})^2$ (rectangular area indicated in blue); the area under the curve segment with 0.3 $nm^{-1} \leq q \leq 2\ nm^{-1}$ was determined by trapezoidal integration (area indicated in orange); the area under the curve in the large q limit with q > 2 $nm^{-1}$ was extrapolated by the Porod law (black solid line, area indicated in grey). Based on these considerations, a T-parameter

of 2.28 nm was calculated corresponding to a platelet thickness of D = 1.14 nm when assuming dilute (non-interacting) disc-shaped scattering objects (i.e. pores filled with organic occlusions).

**SI Table 1.** Aragonite-binding peptides selected by phage display. In total, 33 different peptide sequences were selected. Peptide sequences are sorted by the calculated isoelectric point (p*I*) (Vector NTI software, Invitrogen). The amino acids of the peptide sequence are denoted in single letter abbreviations.

| Peptide | Frequency | Sequence | pI |
|---|---|---|---|
| A-4-16 | 1/44 | DFLPNAIGDRQY | 4,47 |
| A-4-09 | 1/44 | DHQVYSENKPTI | 5,69 |
| A-4-14 | 1/44 | FPTLTRHEAVEI | 5,69 |
| A-5-11 | 1/44 | WIDNKQDANSHA | 5,69 |
| A-5-06 | 1/44 | RLLNTETEHSVY | 5,69 |
| A-5-17 | 1/44 | PNGQPTQGSEKG | 6,25 |
| A-4-01 | 1/44 | GIQVSNLARLHD | 7,19 |
| A-4-06 | 1/44 | IPASSRHNPPIE | 7,19 |
| A-4-27 | 1/44 | TMHGSISSEPKN | 7,19 |
| A-5-13 | 1/44 | DSNVKNHNNKPE | 7,19 |
| A-4-30 | 1/44 | VASPHASWHKDL | 7,38 |
| A-4-03, A-4-05, A-4-18 | 3/44 | YPGIMKTPSTPS | 8,88 |
| A-4-23 | 1/44 | TSTSSLRTMPLY | 8,88 |
| A-4-24 | 1/44 | WPSWKNPAYSLT | 8,88 |
| A-4-26, A-4-29, A-5-02, A-5-05, A-5-15 | 5/44 | YPGIMKTPSTPS | 8,88 |
| A-4-10 | 1/44 | KGTSPQVTSTHN | 9,06 |
| A-4-21 | 1/44 | RNHFEPLKHAIP | 9,06 |
| A-5-20 | 1/44 | HHKEMHSRMPAP | 9,06 |
| A-4-04 | 1/44 | YSLSAGYRKLMV | 9,72 |
| A-4-25 | 1/44 | YSRPAYKLPFGM | 9,72 |
| A-4-02 | 1/44 | LVGRVLLMLLFL | 10,00 |
| A-4-07 | 1/44 | HHKILHTGPYRL | 10,00 |
| A-4-11, A-4-12, A-5-10 | 3/44 | LPKWKAPHPFFS | 10,00 |
| A-4-13 | 1/44 | APKWLPHSKHWS | 10,00 |
| A-4-19, A-4-20, A-5-09 | 3/44 | HHKILHTGPYRL | 10,00 |
| A-4-22 | 1/44 | HPLWKATQMHKF | 10,00 |
| A-4-28 | 1/44 | WNAPGRYNASKT | 10,00 |
| A-5-07, A-5-14 | 2/44 | HKKPSPNMWPLS | 10,00 |
| A-5-03 | 1/44 | HPIHKPPHKRWN | 11,12 |
| A-5-12 | 1/44 | KPNHPHLSLWRW | 11,12 |
| A-4-15 | 1/44 | TKNPPRTQRPTK | 12,02 |
| A-4-17 | 1/44 | HKHNQRPRSQKQ | 12,02 |
| A-4-08 | 1/44 | QHNPILRPQPRR | 12,30 |

**SI Table 2.** Calcite-binding peptides selected by phage display. In total, 39 different peptide sequences were selected. Peptide sequences are sorted by the calculated isoelectric point (p*I*) (Vector NTI software, Invitrogen). The amino acids of the peptide sequence are denoted in single letter abbreviations.

| Peptide | Frequency | Sequence | pI |
|---|---|---|---|
| C-5-13 | 1/39 | QYTFENPQTNVA | 3,81 |
| C-5-15 | 1/39 | WLEQSIPAPQNP | 3,81 |
| C-4-18 | 1/39 | SFAEDTPYRPPS | 4,47 |
| C-4-03 | 1/39 | SSTDVLSNDPSR | 4,47 |
| C-5-19 | 1/39 | AEYTYPAHQATF | 5,50 |
| C-4-13 | 1/39 | YGLPTQANSMQL | 5,50 |
| C-5-01 | 1/39 | VASSPGYNANTW | 5,50 |
| C-4-20 | 1/39 | GDYDSLTRAHYL | 5,69 |
| C-4-01 | 1/39 | DTYSIMNLTSPR | 6,25 |
| C-4-08 | 1/39 | EQTMRYFLSSTP | 6,25 |
| C-4-29 | 1/39 | GQGKEYPIPNPI | 6,25 |
| C-4-24 | 1/39 | AMRQLPLGQDYA | 6,25 |
| C-4-07 | 1/39 | INASSDAPSRGS | 6,25 |
| C-5-08 | 1/39 | SLQITMQQMELR | 6,25 |
| C-5-18 | 1/39 | SYWPFNALYKEG | 6,25 |
| C-5-05 | 1/39 | ADPYHIRAAAQP | 7,19 |
| C-4-25 | 1/39 | TLELPTRQAQHP | 7,19 |
| C-4-19 | 1/39 | SPLTRHVADFQT | 7,19 |
| C-5-06 | 1/39 | WGSSTWRPEHTP | 7,19 |
| C-5-12 | 1/39 | SAHGTSTGVPWP | 7,19 |
| C-5-09 | 1/39 | DTLPSRPAVTLH | 7,19 |
| C-4-15 | 1/39 | HNTTQWHLIALS | 7,38 |
| C-4-26 | 1/39 | TIHNHINGSSSV | 7,38 |
| C-5-16 | 1/39 | HYTHTIMASANN | 7,38 |
| C-5-17 | 1/39 | HRIHYTDVVPNQ | 7,38 |
| C-4-30 | 1/39 | AVARHDTLHTGH | 7,56 |
| C-4-27 | 1/39 | DLKGFWIKGPSA | 8,88 |
| C-4-06 | 1/39 | TYSGQITAVNSR | 8,88 |
| C-4-10 | 1/39 | LPKYASDRNWGT | 8,88 |
| C-4-28 | 1/39 | SYLWNVNREVQK | 8,88 |
| C-5-02 | 1/39 | GSLVYRSGSHAP | 9,06 |
| C-4-11 | 1/39 | HVPSTSWNKMQH | 9,06 |
| C-4-04 | 1/39 | QPPWLPGSATTR | 10,00 |
| C-4-14 | 1/39 | TQNAVRLDRWLT | 10,00 |
| C-4-17 | 1/39 | GNQFNAQTRQGS | 10,00 |
| C-4-21 | 1/39 | QHTPRVLSHEGR | 10,00 |
| C-4-12 | 1/39 | SGGPRTSPAQPH | 10,00 |
| C-5-20 | 1/39 | SWMPHPRWSPQH | 10,00 |
| C-5-03 | 1/39 | SVSVGMKPSPRP | 11,12 |