UNIVERSITY of York

This is a repository copy of *Glycosylation Fosters Interactions between Model Sea Urchin Spicule Matrix Proteins. Implications for Embryonic Spiculogenesis and Biomineralization*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/131358/

Version: Accepted Version

Article:
Jain, Gaurav, Pendola, Martin, Koutsoumpeli, Eleni orcid.org/0000-0002-3070-3319 et al. (2 more authors) (2018) Glycosylation Fosters Interactions between Model Sea Urchin Spicule Matrix Proteins. Implications for Embryonic Spiculogenesis and Biomineralization. Biochemistry. ISSN 1520-4995

https://doi.org/10.1021/acs.biochem.8b00207

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Glycosylation fosters interactions between model sea urchin spicule matrix proteins. Implications for embryonic spiculogenesis and biomineralization.

Gaurav Jain,[1] Martin Pendola,[1] Eleni Koutsoumpeli,[2] Steven Johnson,[2] and John Spencer Evans[1*]

[1]Laboratory for Chemical Physics, Center for Skeletal and Craniofacial Biology, New York University, 345 E. 24th Street, NY, NY, 10010 USA.

[2]Department of Electronic Engineering, University of York, Heslington, York, YO105DD, United Kingdom.

KEYWORDS: Purple sea urchin, biomineralization, SpSM30, SpSM50, QCM-D, glycosylation

Supporting Information Placeholder

ABSTRACT: The formation of embryonic mineralized skeletal elements (spicules) in the sea urchin requires the participation of proteins, many of which may be interactive with one another and assist in the creation of an extracellular matrix wherein mineral formation takes place. To probe this we created a sea urchin spicule recombinant model protein pair system wherein we tested the interactions between two major spicule proteins, SpSM50 and the glycoprotein, SpSM30B/C. Both proteins are strong hydrogelators which manipulate early and later events in mineral formation. We discovered that the anionic glycan moieties of SpSM30B/C are required for interaction with the SpSM50 protein and that these interactions are Ca(II) – independent. Further, when these proteins complex together they form hybrid hydrogel particles that are physically distinct from their individual counterparts. Thus, glycan-mediated interactions play an important role in *in vitro* spicule protein assembly and most likely within the spicule itself.

Embryonic sea urchin mineralized spicules provide insights into the development and expression of both skeletal elements and biological materials at the nano- and mesoscales.[1] In the purple sea urchin embryo, *Strongylocentrotus purpuratus*, (Sp), the spicule consists of a single mesocrystal of magnesium-bearing calcite ($Ca_{.95}Mg_{.05}CO_3$ or MgC). The formation of MgC initiates with the assembly of amorphous calcium carbonate (ACC) nanoparticles directed by protein families designated as SpSM.[2-18] Subsequently the ACC phase transforms to crystalline MgC and the associated SpSM proteins become entrapped within the crystalline mineral phase and convey fracture resistance.[5,6,9-11,16-18] Because these integrated processes lead to skeletal formation and ultimately material enhancement, there is a compelling reason to study the spicule matrix and the spatial and functional relationships between SpSM proteins and mineral phase.

Two important trends are noted for SpSM proteins. All expressed SpSM proteins feature a folded C-type lectin-like domain (CTLL) at the N-terminus,[2-4,9,19-21] and this domain has been implicated in spicule matrix assembly and mineral formation.[9,19-21] Furthermore in many SpSM proteins there is also an intrinsically disordered Met/Asn/Gln/Pro/Gly-rich (MAQPG) repetitive C-terminal domain that is homologous to known elastomeric sequences. [2-4,9,19-21] We consider that both sequences are crucial for matrix protein assembly and material integrity and possibly direct processes such as protein – mineral phase interaction and matrix-mineral organization within the spicule.

The best characterized SpSM proteins are SpSM50[2-4,9,11-15,19-20] (pI = 10.7, MW = 44.5 kDa, 428 AA, Uniprot: P11994, GenBank: AAA30071.1) and the A-F SpSM30 acidic glycoprotein isoforms[2-4,12-15,19,21] (pI = 5.7, MW ~ 43 kDa, 270 AA, Accession P28163)(Figure S1, Supporting Information). Intriguingly, knockdown of SpSM30 expression has little effect on spicule mineralization, formation or elongation, whereas knockdown of SpSM50 expression has a much more significant impact.[7,15] Recombinant model forms of both proteins (insect cell expressed, N-, O-linked glycosylated rSpSM30B/C-G, a hybrid of the "B" and "C" isoform sequences; bacteria expressed rSpSM50)[19-21] have been found to form protein hydrogel particles that act as "smart" hydrogels that dimensionally respond to pH and ions, bind and release water and ions, modify the surfaces and subsurfaces of calcite crystals, and regulate different aspects of the non-classical nucleation scheme.[19-21] Thus, the spicule matrix is a hydrogel environment consisting of multiple spicule matrix proteins, and it is within this hydrogel environment that ACC formation, assembly, stabilization, and transformation processes[16-18] take place during spiculogenesis. However, it is not clear if SpSM50 and SpSM30 interact with one another within this hydrogel matrix, and if so, whether the glycosylation of SpSM30 plays a role in this process.

Recently, to better understand mineral-associated ECM formation, we introduced model combinatorial recombinant biomineralization protein pair studies where we discovered a significant impact of defined protein molar ratios (1:1) on mineral formation.[22,23] We found that the protein pairs exhibited non-covalent binding to one another, forming hybrid protein hydrogels that promoted synergistic effects on either the early and/or later stages of nucleation and crystal growth.[22,23] In this new study, we applied this approach to learn whether or not SpSM50 and SpSM30 sequences interact with one another and if so, whether the glycan moieties of SpSM30 play a role in these interactions. Using a single recombinant model variant of SpSM50 (rSpSM50),[19,20] and two variants of SpSM30B/C (unglycosylated, bacteria expressed variant = rSpSM30B/C-NG; glycosylated, Sf9 insect cell expressed variant = rSpSM30B/C-G),[19,21] we find that only the glycosylated variant of the SpSM30B/C interacts with the SpSM50 sequence and does not require Ca^{2+}. This result indicates that the SpSM30B/C-G oligosaccharide moieties, and not the primary sequence itself, are important for spicule matrix protein-protein interaction. Furthermore, based on hydrogel particle size distributions, internal complexities (i.e., structure), and electrostatics, the interactive rSpSM30B/C-G : rSpSM50 complex forms protein hydrogel particles that differ from those formed by the individual proteins. Thus, using this model recombinant protein pair system, we conclude that post-translational glycosylation is critical for the assembly, formation, and mineralization of the spicule ECM.

The primary sequences of rSpSM30B/C-G and rSpSM30B/C-NG are identical, with the only difference being the presence of anionic glycans on the rSpSM30B/C-G insect cell variant.[21] Thus, a comparison of the two recombinant rSpSM30B/C proteins against the rSpSM50 sequence affords us an opportunity to test the contribution of insect cell glycosylation on SpSM30 – SpSM50 interactions. We conducted these studies at pH 8.0 since this is the relevant pH of the *in vitro* calcium carbonate mineralization assays conducted on these proteins,[19-21] with the actual *in situ* pH of the spicule matrix not known at present.

Using quartz crystal microbalance with dissipation (QCM-D) we examined the differences in interaction between the glycosylated rSpSM30B/C-G or the non-glycosylated

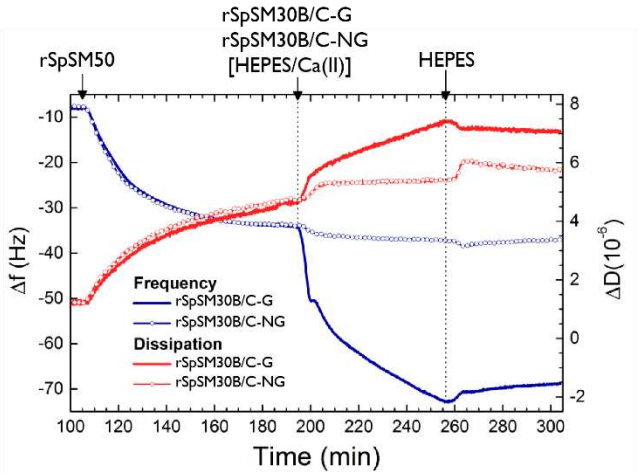


Figure 1. *QCM-D experiments of immobilized rSpSM50 (500 nM in 10 mM HEPES pH 8.0) exposed to either 500 nM rSpSM30B/C-G or 500 nM rSpSM30B/C-NG in 10 mM $CaCl_2$, 10 mM HEPES, pH 8.0. Plots show the third harmonic frequency (F3, blue) and dissipation (D3, red) observed under each scenario. The time-dependent introduction of rSpSM50, rSpSM30B/C-G (lines) or rSpSM30B/C-NG (circles) and HEPES washing solutions is noted on the plots by arrows and extended dashed lines. These experiments were repeated (Figure S1, Supporting Information) and found to be reproducible.*

rSpSM30B/C-NG with the surface-immobilized rSpSM50 protein (Figure 1). In each individual experiment, we created a layer of adsorbed rSpSM50 protein on poly(L-lysine) using a 500 nM protein solution in 10 mM HEPES buffer pH 8.0 (Figure 1, first arrow). Next, we introduced either rSpSM30B/C-G or rSpSM30B/C-NG (500 nM), both dissolved in a 10 mM $CaCl_2$ solution in HEPES buffer (second arrow), over the immobilized rSpSM50 layer. As seen in Figure 1, in the first scenario, the introduction of rSpSM30B/C-G induced a large shift in the frequency and dissipation of the functionalized sensor (which reflect mass deposition and viscoelasticity, respectively), indicative of a strong binding reaction between the two proteins. In the second scenario, the introduction of rSpSM30B/C-NG showed negligible shifts in frequency and dissipation, similar to a control measurement (Figure S2), where only 10 mM $CaCl_2$ solution in HEPES buffer was injected over the rSpSM50-functionalized sensor. This result suggests that the rSpSM50 and rSpSM30B/C-NG primary sequences are non-reactive. This result indicates that even though each protein sequence is electrostatically opposite (pI_{SpSM50} =10.7 ; $pI_{SpSM30B/C-G}$ =5.7), net protein charge is insufficient to promote protein-protein binding between rSPSM50 and rSpSM30B/C-NG.

The binding differences between rSpSM30B/C-G and rSpSM30B/C-NG were found to be reproducible (Figure S2, Supporting Information); thus, the presence of glycans play a key role in the interaction between the rSpSM30B/C and rSpSM50 proteins. Furthermore, control measurements of rSpSM30B/C-G and rSpSM30B/C-NG in HEPES buffer alone (Figure S3, Supporting Information) exhibited the same QCM-

D response as those in HEPES/Ca(II) buffer, indicating that the binding reaction is independent of Ca(II) ions.

Having confirmed interactions between rSpSM30B/C-G and rSpSM50, we chose to focus on this protein pairing for subsequent experiments at pH 8.0 sans Ca^{2+} to determine if these proteins co-hydrogelate. We achieved this using two methods: a) flow cytometry,[19-21] which probes the particle size distributions and internal granularities of the translucent protein hydrogel particles, and b) electrophoretic mobility / ζ potential measurements,[24,25] which determine particle sizes and electrostatic charges of protein hydrogel particles.

Flow cytometry provides light scattering parameters that one can monitor for particles under constant flow:[19-21] 1) forward scattered light component (FSC, x-axis), which determines particle size distributions; and 2) side-scattered light

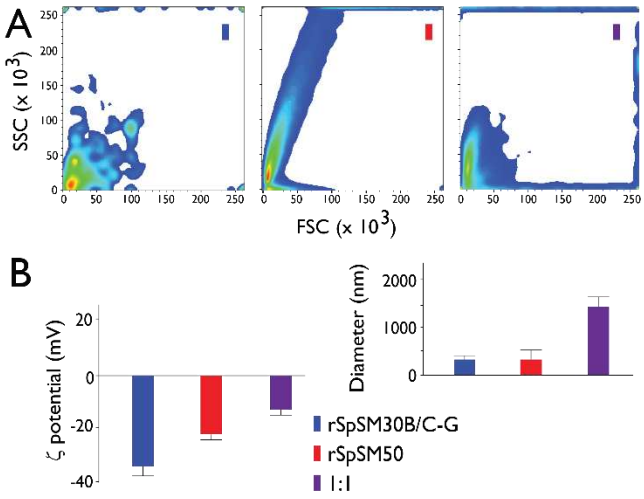


Figure 2. *(A) Flow cytometry 2-D density plots (FSC vs SSC) of 1.5 μM rSpSM30B/C-G (blue), rSpSM50 (red), and 1:1 rSpSM30B/C-G : rSpSM50 (purple) samples in 10 mM HEPES, pH 8.0. (B) Histogram plot of mean spicule matrix protein hydrogel particle zeta potentials and diameters in 10mM HEPES, pH 8.0. For (A), details regarding FSC and SSC parameters and other features can be found in Supporting Information. For (B), all measurements were done in triplicate. For both (A) and (B), individual protein concentrations were 500 nM.*

component (SSC, y-axis), which measures refracted and reflected light that occurs at any interface within the particles where there is a change in refractive index (RI) that results from variations in particle granularity or internal structure (see Supporting Information). As shown in Figure 2A, at pH 8.0 rSpSM50 and rSpSM30B/C-G generate hydrogel particles with unique particle size distributions and internal granularities. When present in a 1:1 mixture, both parameters change, indicating that the combination of the two spicule matrix proteins generates hydrogel particles that are non-equivalent to their individual counterparts. This suggests that the hydrogel particles that form from the 1:1 mixture are hybrid in nature.

We confirmed this finding by using Zetasizer instrumentation to assess particle sizes and ζ potentials (surface charges)[24,25] of hydrogel particles formed by rSpSM30B/C-G, rSpSM50, and 1:1 rSpSM30B/C-G : rSpSM50 (Figure 2B). Here, we note that each spicule matrix protein possesses approximately the same particle diameters at pH 8.0, but differ in their ζ potentials (30%), indicative of the differences in protein charge [i.e., $pI_{SpSM50} > pI_{SpSM30B/C}$] and associated ions. In the 1:1 scenario, the values for both the particle diameters and ζ potentials shift dramatically, with the particle diameters increasing ~4x and the ζ potentials decreasing ~2-3x from the individual protein values. We believe that the reduction in ζ potentials reflects the change in net electrostatics as the cationic rSpSM50 forms complexes with the anionic rSpSM30B/C-G molecules and creates hybrid protein hydrogel particles whose charges are lower in value due to ion pairing neutralization effects. Likewise, the increase in particle dimensions confirm that the two proteins are co-assembling together to form hydrogel particles that exceed the dimensions of the particles formed by the individual proteins themselves.

In this report, using model recombinant variants of *S. purpuratus* SpSM30B/C and SpSM50,[19-21] we confirm that these proteins are interactive and only the oligosaccharide-modified variant of rSM30B/C[19,21] exhibits non-covalent interactions with the major spicule matrix protein, rSpSM50 (Figure 1). Moreover, rSpSM30B/C-G interactions with rSpSM50 are Ca^{2+} - independent (Figure 1) and the resulting hydrogel particles that form possess physical properties (Figure 2) that are significantly different from hydrogel particles formed by the individual proteins, i.e., these are hybrid hydrogel particles. This could explain the role of SpSM30 proteins in spiculogenesis: to boost the activity of the predominant SpSM50 protein. This would explain why knockdown of SpSM30 demonstrated no observable changes in spicule formation,[7,15] since its absence would still permit SpSM50, a strong hydrdogelator and nucleation modulator, to function well enough to allow spicule matrix formation.

Although we do not know the non-bonding interactions between rSpSM50 – rSpSM30B/C-G, we speculate that protein-carbohydrate recognition is the key. Like the majority of SpSM proteins, rSpSM50 has a conserved CTLL carbohydrate recognition sequence near the N-terminus (Figure 3; Figure S1, Supporting Information) that features cat ionic regions.[2-4,19,20]

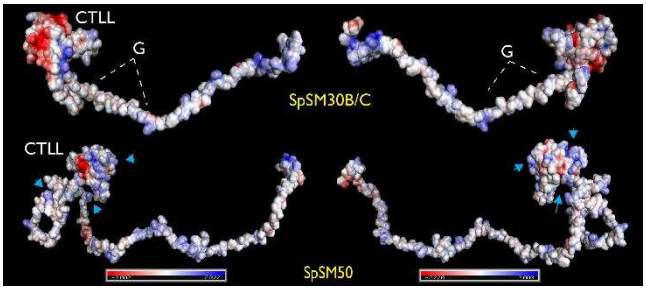


Figure 3*. PBEQ Solver/CHARMM-GUI solvation energy surface plots for DISOclust/INTFOLD2-predicted structures of SpSM50[20] and SpSM30B/C[21] primary sequences. Electrostatic surfaces are color-coded as: red = anionic; white = neutral; blue = cationic. CTLL domains are indicated; "G" = predicted region for SpSM30B/C O-glycosylation;[21] Blue arrows = cationic regions of SpSM50 CTLL sequence. Description of solvation modeling protocol can be found in Supporting Information.*

Since rSpSM30B/C-G anionic oligosaccharides are believed to exist within the extended MAQPG region[21] (Figure 3), we postulate that the accessible anionic glycan groups on rSpSM30B/C-G are recognized by the cationic CTLL domain of rSpSM50 and interact there. As mentioned elsewhere,[19-21] the intrinsically disordered MAQPG C-terminal region and the modified CTLL fold represent putative binding domains for spicule matrix protein-protein assembly and subsequent

hydrogelation, and thus are critical for the protein-mediated mineralization process to occur.[9,19-21] We suggest that subsequent studies should focus on native glycosylation, intrinsically disordered regions, and the role(s) that they play in embryonic sea urchin skeletal formation.

ASSOCIATED CONTENT

Supporting Information

Detailed experimental procedures; the primary amino acid sequences of SpSM30B/C and SpSM50 (Figure S1); Full QCM-D measurements of immobilized rSpSM50 exposed to either rSpSM30B/C-G or rSpSM30B/C-NG in 10 mM $CaCl_2$, 10 mM HEPES, pH 8.0 (Figure S2); QCM-D experiments of immobilized rSpSM50 exposed to either rSpSM30B/C-G or rSpSM30B/C-NG in 10 mM HEPES, or, to 10 mM $CaCl_2$ in 10 mM HEPES alone (Figure S3). The Supporting Information is available free of charge on the ACS Publications website as a PDF file.

AUTHOR INFORMATION

Dr. Gaurav Jain	jaing@nyu.edu
Dr. Martin Pendola	lmp461@nyu.edu
Dr. John Spencer Evans	jse1@nyu.edu
Eleni Koutsoumpeli	eleni.koutsoumpeli@york.ac.uk
Steven Johnson	steven.johnson@york.ac.uk

Corresponding Author

*To whom correspondence should be addressed. Tel: 347-753-1955; email jse1@nyu.edu

Author Contributions

The manuscript was written through contributions of all authors, and all authors have given approval to the final version of the manuscript.

Funding Sources

Portions of this research were supported by the Life Sciences Division, U.S. Army Research Office, under award W911NF-16-1-0262 (JSE). Research involving QCM-D measurements was supported by grants from the UK Engineering and Physical Sciences Research Council (EP/M02757/1, EP/P030017/1)(EK, SJ)

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This paper represents Contribution Number 92 from the Laboratory for Chemical Physics, New York University.

ABBREVIATIONS

SpSM30B/C= B, C isoform hybrid spicule matrix glycoprotein; SpSM50 = spicule matrix protein SpSM50; rSpSM30B/C-G = recombinant insect cell expressed SpSM30B/C; rSpSM30B/C-NG = recombinant bacteria-expressed non-glycosylated SpSM30B/C; rSpSM50 = recombinant SpSM50; QCM-D = quartz crystal microbalance with dissipation monitoring; ACC = amorphous calcium carbonate; FSC-A = forward scattered component; SSC-A = side scattered component; PMC = primary mesenchymal cell; MgC = magnesium bearing calcite.

REFERENCES

1 Studart, A.R. (2012) Towards high-performance bioinspired composites. *Adv. Materials 24*, 5024-5044.

2 Cameron, R.A., Samanta, M., Yuan, A., He, D., Davidson, E. (2009) SpBase: the sea urchin genome database and web site. *Nucleic Acids Research.* http://www.spbase.org/

3 Samanta, M.P., Tongrasit, W., Istrail, S., Cameron, R.A., Tu, Q., Davidson, E.H., Stolc, V. (2006) The transcriptome of the sea urchin embryo. *Science* 314, 960-962.

4 Mann, K., Wilt, F.H., Poustka, A.J. (2010) Proteomic analysis of sea urchin (Strongylocentrotus purpuratus) spicule matrix. *Proteome Science 8*, 1-12.

5 Seto, J., Ma, Y., Davis, S.A., Meldrum, F., Schilde, U., Gourrier, A., Jaeger, C., Cölfen, H. (2012) Structure-property relationships of a biological mesocrystal in the adult sea urchin spine. *Proc. Natl. Acad. Sci USA 109*, 3699-3704.

6 Berman, A., Addadi, L., Kvick, A., Leiserowitz, L., Nelson, M., Weiner, S. (1990) Intercalation of sea urchin proteins in calcite: Study of a crystalline composite material. *Science 250*, 664-667.

7 Sharmankina, V.V., Kiselev, K.V. (2016) Expression of SM30(A-F) genes encoding spicule matrix proteins in intact and damaged sea urchin Strongylocentrotus intermedius. *Russ. J. Genetics* 52, 298-303.

8 Wilt, F., Killian, C.E., Croker, L., Hamilton, P. (2013) SM30 protein function during sea urchin larval spicule formation. *J. Struct. Biol. 183*, 199-204.

9 Rao, A., Seto, J., Berg, J.K., Kreft, S.G., Scheffner, M., Cölfen, H. (2013) Roles of larval sea urchin spicule SM50 domains in organic matrix self-assembly and calcium carbonate mineralization. *J. Struct. Biol. 183*, 205-215.

10 Wu, C.-H., Park, A., Joester, D. (2011) Bioengineering single crystal growth. *J. Am. Chem. Soc. 133*, 1658–1661.

11 Seto J., Zhang, Y., Hamilton, P., Wilt, F.H. (2004) The localization of occluded matrix proteins in calcareous spicules of sea urchin larvae. *J. Structural Biology 148*, 23-30.

12 Livingston, B.T., Killian, C.E., Wilt, F., Cameron, A., Landrum, M.J., Ermolaeva, O., Sapojnikov, V., Maglott, D.R., Buchanan, A.M., Ettensohn, C.A. (2006) Genome-wide analysis of biomineralization-related proteins in the sea urchin Strongylocentrotus purpuratus. *Devel. Biol. 300*, 335-348.

13 Killian, C.E., Wilt, F.H. (2008) Biomineralization of the echinoderm endoskeleton. *Chem. Rev. 108*, 4463–4474.

14 Kitajima, T., Urakami, H. (2000) Differential distribution of spicule matrix proteins in the sea urchin embryo skeleton. *Dev. Growth Differ. 42*, 295–306.

15 Urry, L.A., Hamilton, P.C., Killian, C.E., Wilt, F.H. (2000) Expression of spicule matrix proteins in the sea urchin embryo during normal and experimentally altered spiculogenesis. *Dev. Biol. 225*, 201–213.

16 Tester, C.C., Wu, C.-H., Krejci, M. R., Mueller, L., Park, A., Lai, B., Chen, S., Sun, C., Joester, D. (2013) Time-resolved evolution of short- and long-

range order during the transformation of ACC to calcite in the sea urchin embryo. *Adv. Fun. Mat. 23*, 4185-4194.

17 Politi, Y, Metzler, R.A., Abrecht, M., Gilbert, B., Wilt F.H., Sagi, I, Addadi, L., Weiner, S., Gilbert, P. (2008) Transformation mechanism of amorphous calcium carbonate into calcite in the sea urchin larval spicule. *Proc Natl Acad Sci U S A. 105*, 17362-17366.

18 Gong, Y.U.T., Killian, C.E., Olson, I.C., Appathurai, N.P., Amasino, A.L., Martin, M.C., Holt, L.J., Wilt, F.H., Gilbert, P. (2012) Phase transitions in biogenic amorphous calcium carbonate. *Proc Natl Acad Sci USA. 109*, 6088-6093.

19 Pendola, M., Davidyants, A., Jung, Y.S., Evans, J.S. (2017) Sea urchin spicule matrix proteins form mesoscale hydrogels that exhibit selective ion interactions. *ACS Omega*, *2*, 6151-6158.

20 Jain, G., Pendola, M., Huang, Y.C., Gebauer, D., Evans, J.S. (2017) A model sea urchin spicule matrix protein, rSpSM50, is a hydrogelator that modifies and organizes the mineralization process. *Biochemistry*, *56*, 2663-2675.

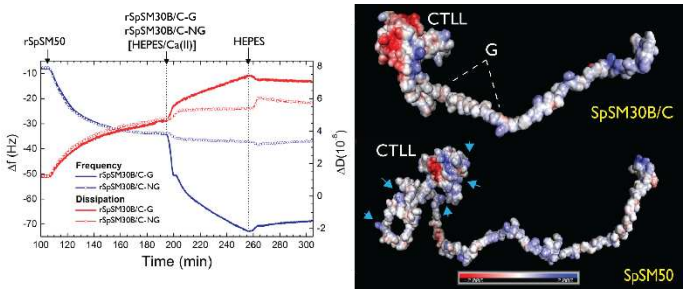
21 Jain, G., Pendola, M., Rao, A., Cölfen, H., Evans, J.S. (2016) A model sea urchin spicule matrix protein self-associates to form mineral-modifying protein hydrogels. Biochemistry, 55, 4410-4421.

22 Jain, G., Pendola, M., Huang, Y.C., Colas, J.J., Gebauer, D., Johnson, S., Evans, J.S. (2017) Functional prioritization and hydrogel regulation phenomena created by a combinatorial pearl-associated 2-protein biomineralization model system. Biochemistry, 56, 3607-3618.

23 Chang, E.P., Roncal-Herrero, T., Morgan, T., Dunn, K.E., Rao, A., Kunitake, J.A.M.R., Lui, S., Bilton, M., Estroff, L.A., Kröger, R., Johnson, S., Cölfen, H., Evans, J.S. (2016) Synergistic biomineralization phenomena created by a nacre protein model system. Biochemistry 55, 2401-2410.

24 Khan, F., Li, W., Habelitz, S. (2012) Biophysical characterization of synthetic amelogenin C-terminal peptides. Eur. J. Oral Sci. 120, 113-122.

25 Uskokovic, V., Castiglione, Z., Cubas, P., Zhu, L., Li, W., Habelitz, S. (2010) Zeta-potential and particle size analysis of human amelogenins. J. Dent. Res. 89, 149-153.



Insert Table of Contents artwork here