ADVANCED FUNCTIONAL MATERIALS

Supporting Information

for *Adv. Funct. Mater.*, DOI: 10.1002/adfm.201802063

Stabilization of Mineral Precursors by Intrinsically Disordered Proteins

*Ashit Rao, Markus Drechsler, Stefan Schiller, Martin Scheffner, Denis Gebauer, and Helmut Cölfen**

Supplementary Information

Stabilization of Mineral Precursors by Intrinsically Disordered Proteins

Ashit Rao, Markus Drechsler, Stefan Schiller, Martin Scheffner, Denis Gebauer, Helmut Cölfen

Figure S1. Venn diagram representing the distributions of proteins containing putative low complexity sequences and disordered (>10 consecutive amino acids) regions from the spicule proteome identified using the SMART and DISOPRED servers, respectively [1].

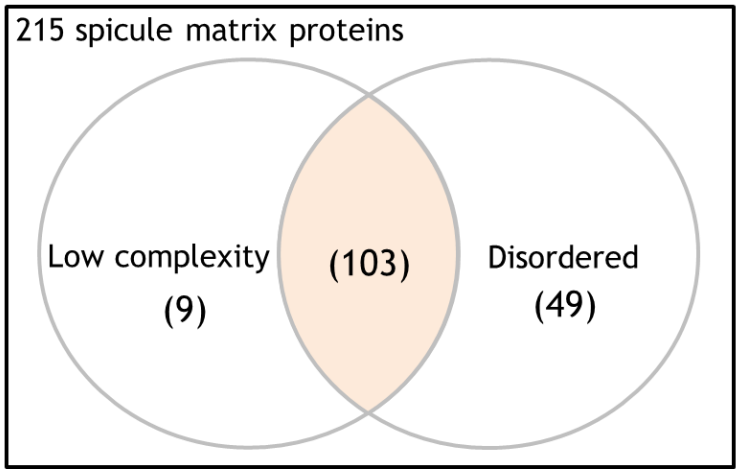


Figure S2. Primary sequences of recombinant protein additives containing LC-IDRs. The (6XHis)-SUMO tag is underscored. Polar and charged amino acid residues are marked in red. The effects of the SUMO tag are studied in conjugation with the disordered N-terminal transactivation domain (TAD) of a tumor suppressor protein[2] i.e. SUMO-p53TAD as a reference additive.

>SUMO-SM50GRR
MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
DLDMEDNDIIEAHREQIGGWPVNPQNPMSGPPGRAPVMKRQNPPVRPGQGGRQIPQGVGPQWEAVEVTAMRAFVCEVPAGR
NIPIGQQPGMGQGGFGNQQPGMGGRQPGFGNQPGMGGRQPGFGNQPGMGGRQPGWGNQPGVGGRQPGMGGQQPGW
GNQPGVGGRQPGMGGQPGVGGRQPGFGNQPGMVDNNQAWWTTTRLGNQPGVGGRQPGMGGQPGVGGRQPGVGGRQP
GFGNQPGVGGRQPGMGGQQPGMGGQPGVGGRQPGMGGRQPGFGNQPGVGGRQPGMGGQQ

>SUMO-LSM34GRR
MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
DLDMEDNDIIEAHREQIGGVAAGQNIPPGQQPGFGGQQPGFGGRQPGFGGQQPGFGQQPGFGGRQPGFGGRQPGFGGQQP
GFGGQQPGFGGQQPGFGGQQPGFGGQQPGFGGQQPGFGGQQPGFGGGPQRPGMGGQPNSPNPRFNRPRMLQEAETDVTG
S

>SUMO-MSP130
MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
DLDMEDNDIIEAHREQIGGQIVNTTFGAGGPMMTTPMAGFPRGTTWSPNAGAGGQGGQGQYPGQGGQGGQGGQGGQGQY
PGQGGQGGQGGQGGQGGQGGYPGQGGQGGPGYYPGQGGQGGQGGQGGWGQGGGQGGQGGGNNPQYPMIPTTPLPGTI
CNGTSTMSYVVSQINFHKFNAPAEVQRLKALHVRQPYTGQLGDPGP

>SUMO-Prisilkin39
MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
DLDMEDNDIIEAHREQIGGYSYGYPTGGYGGYSYGYPTGGYGGYSYGYPTGGYGGYSYGYPTGGYSGYSYGYPTGGYSGYSYGYPT
GGYSGYSYGYPTGGYSGYSYGYPTGGYSGYSYGYPTGGYSGYSYPTGGYSGYSYSSTPGYGYYGSGSGMGGMRSGYSYYSSPAPSYY
SSGSMTPGYGYYSSGSGIGGGMGSGYSYYSSPAPSYYSSSVSPGYGYYGSGSGMRGYGYYSSSTPMYYGSRSTGYGPFSSGLGG

>SUMO-p53TAD
MGHHHHHHGSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
DLDMEDNDIIEAHREQIGGMEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRM
PEAAPPVAPAPAAPTPAAPAPAPSWP

Table S1. Isoelectric points and abundant amino acids of the selected LC-IDRs.

LC-IDR	pI	Amino acid contents (%)
p53TAD	3.5	Pro (23.0), Ala (13.0)
Prisilkin-39	8.5	Gly (30.7), Tyr (27.7), Ser (21.4)
MSP130	9.2	Gly (34.4), Gln (16.4)
LSM34	11.9	Gly (34.0), Gln (20.6), Pro (15.6), Phe (10.6)
SM50	12.4	Gly (32.0), Gln (15.4), Pro (15.1)

Figure S3. Putative local disorder in (1) full length SM50 and selected LC–IDRs (Fig. S2) of (2) MSP130, (3) LSM34, (4) prisilkin and (5) p53 TAD predicted using DISOPRED (green), IUPRED (red) and RONN (blue) algorithms[1b, 3].

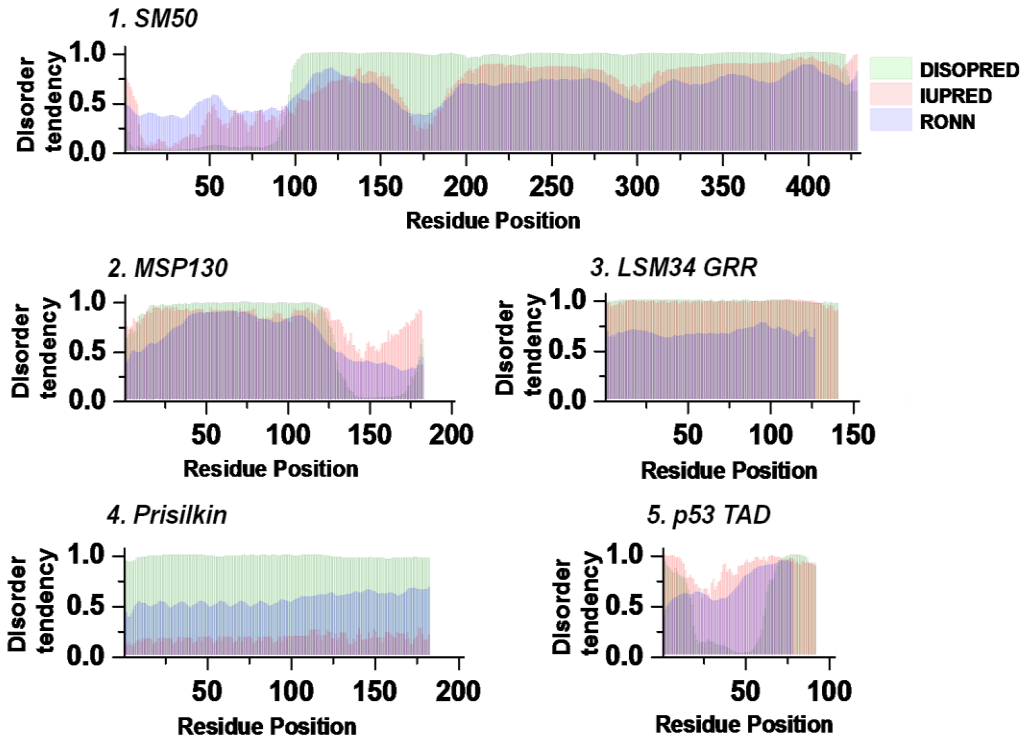


Figure S4. Representative time developments of free Ca^{2+} (left) and ion products (right) of calcium carbonate in presence of SUMO fusion products with LC-IDRs of (A) SM50, (B) LSM34, (C) MSP130, (D) prisilkin-39 and (E) p53TAD at pH 9.0 with 0 (black), 0.1 (blue) and 1 (red) mg/mL protein contents.

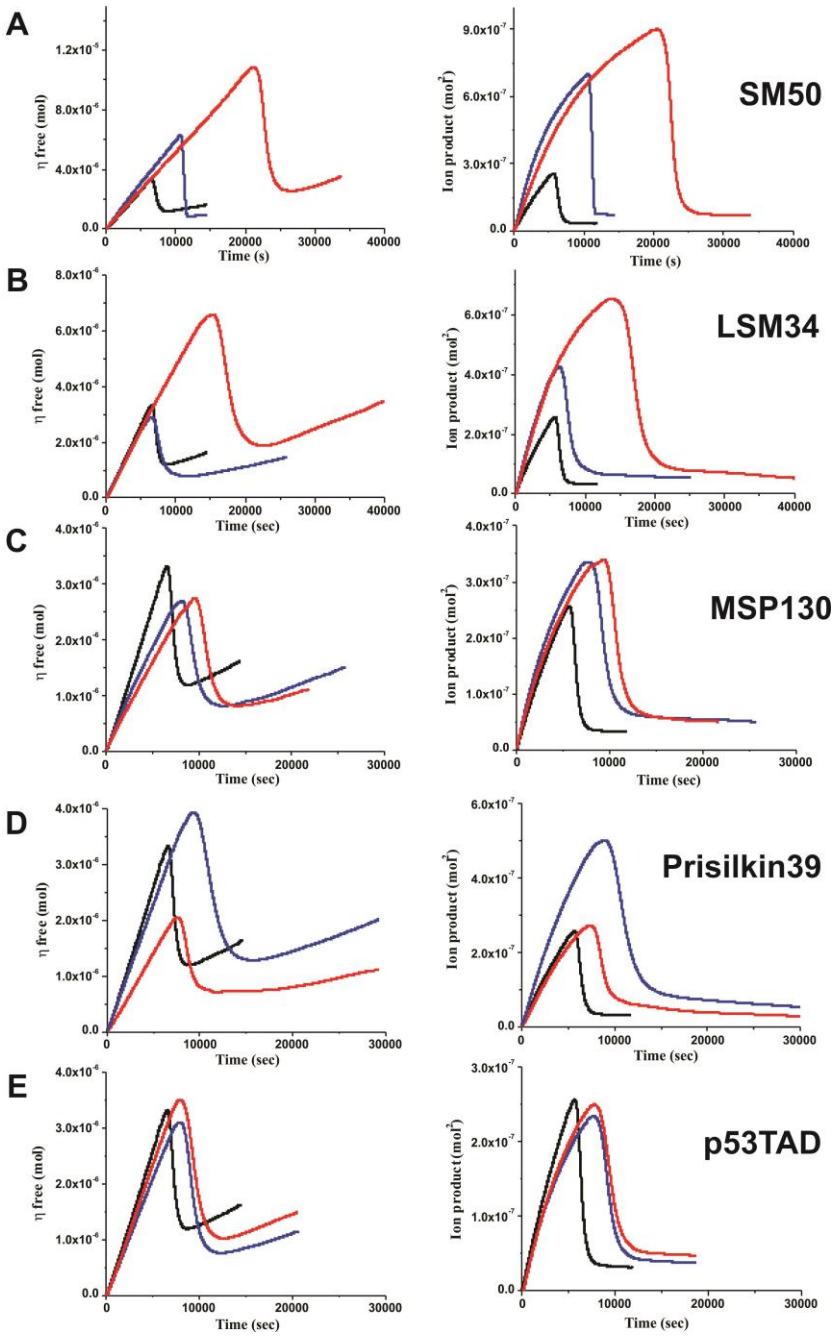


Figure S5. Representative development of free Ca^{2+} (left) and ion products (right) of calcium carbonate in presence of SUMO fusion products with LC-IDRs of (A) SM50, (B) LSM34, (C) MSP130, (D) prisilkin-39 and (E) p53TAD at pH 9.75 with 0 (black), 0.1 (blue) and 1 (red) mg/mL protein contents.

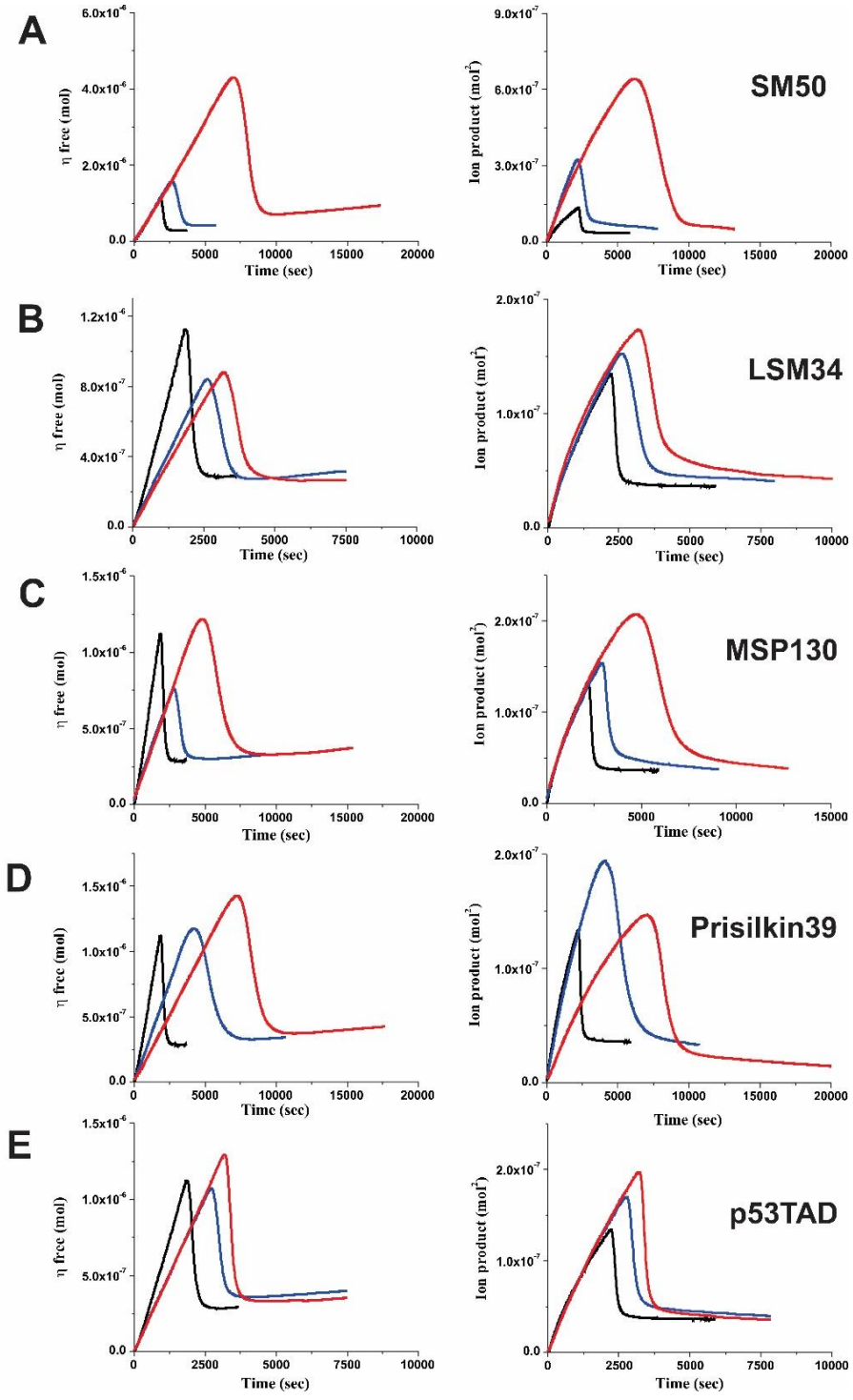


Figure S6. Simulated pH-dependent ionization of the LC–IDR sequences shown in Fig. S2 (Protein Calculator v3.4, C. Putnam, Scripps Research Institute, U.S.A), assuming that the residues exhibit pKa values identical to individual amino acids. The blue-shaded area represents the experimentally accessed pH range.

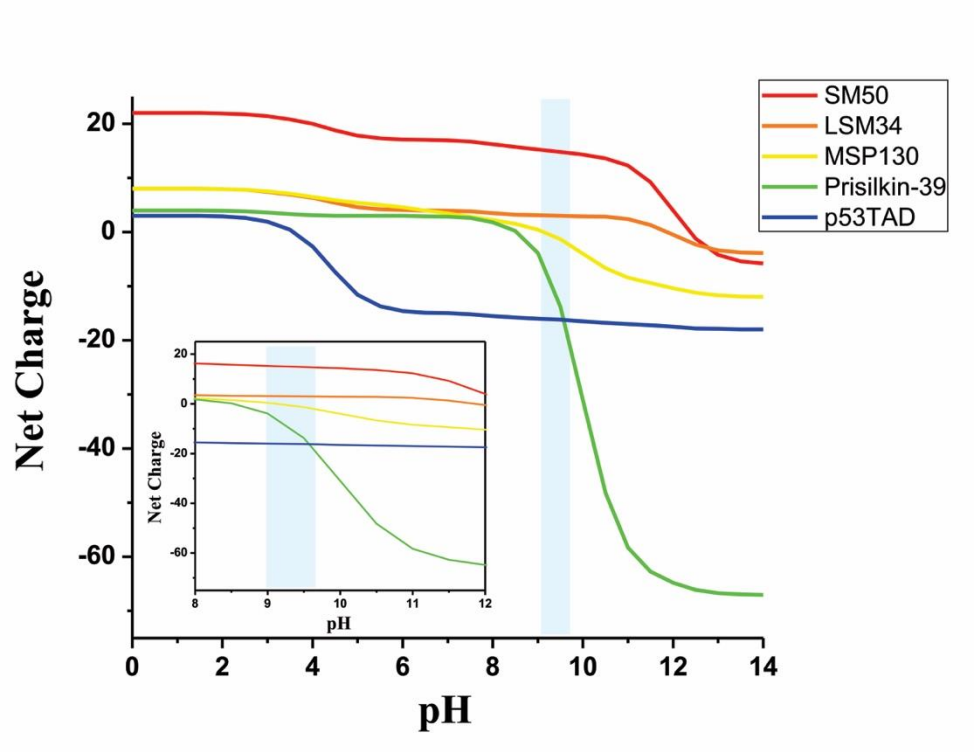


Figure S7. Representative bright field and polarization light microscopy of products formed in the presence of fusion proteins derived from (a) SM50, (b) LSM34, (c) MSP130 and (d) prisilkin-39 at pH 9.0, immediately after nucleation in fluid environments. Scale bars represent 50 µm. Arrows indicate liquid-like phases whereas the crystalline particles are encircled.

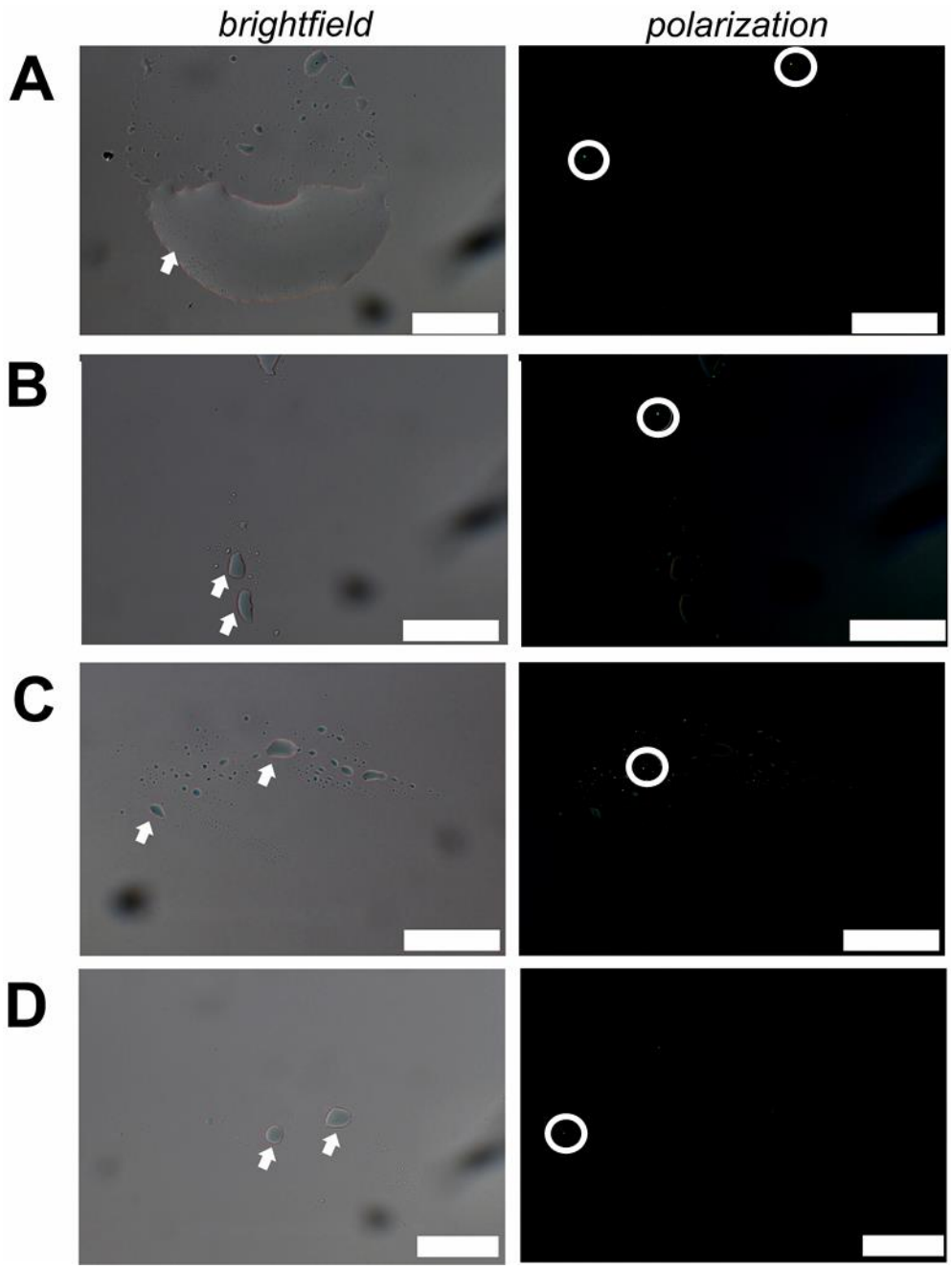


Figure S8. Time development of hydrodynamic radii of species from reference and additive-containing titrations performed at pH 9.0. Time points are normalized by using a scaling factor (F)[4] wherein a value of 1 corresponds to nucleation time i.e. the drop in free Ca^{2+} content in the respective titration run. Distinct developmental trends of particle sizes are represented by pink (prisilkin-39, p53 and reference) and teal (SM50, LSM34 and MSP130) zones.

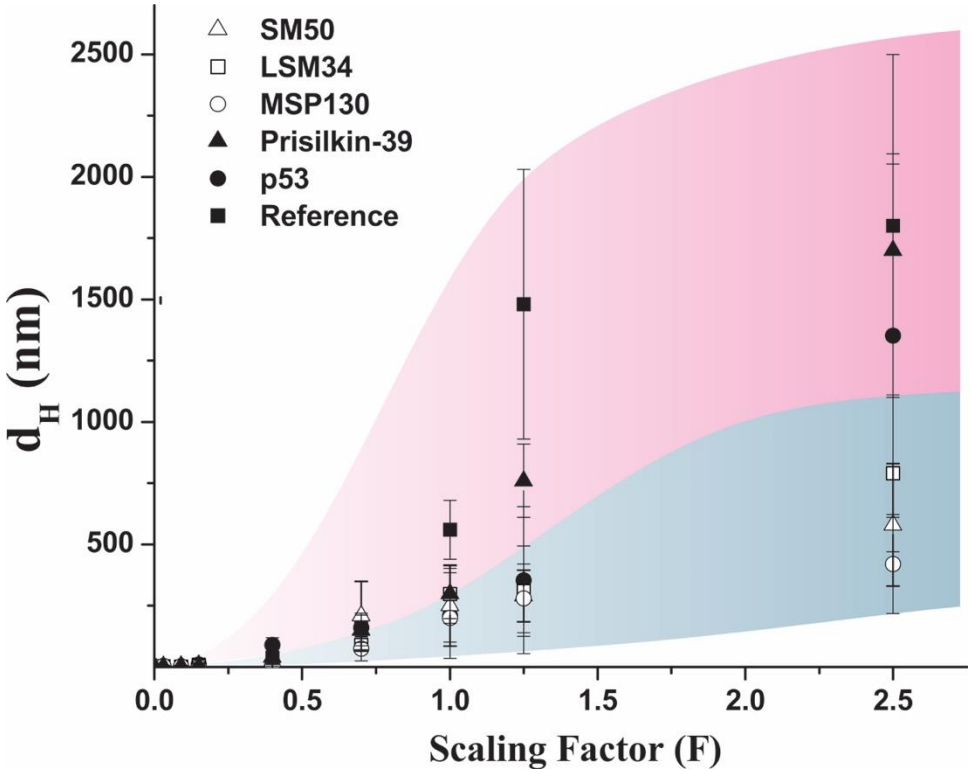


Figure S9. Representative TEM images of crystalline particles formed after the solvent evaporation of mineral products nucleated in the presence of SUMO-fusion LC-IDRs from (A, B) SM50, (C) LSM34, (D) MSP130, (E) prisilkin-39 and (F) p53TAD. The d spacing values match the calcite (104) and (113) planes. Scale bars represent 500 nm. For SUMO-p53TAD, note an outer thinner ring and the absence of vesicles under cryo conditions (Fig. 4), reflective of a poor and transient shell integrity.

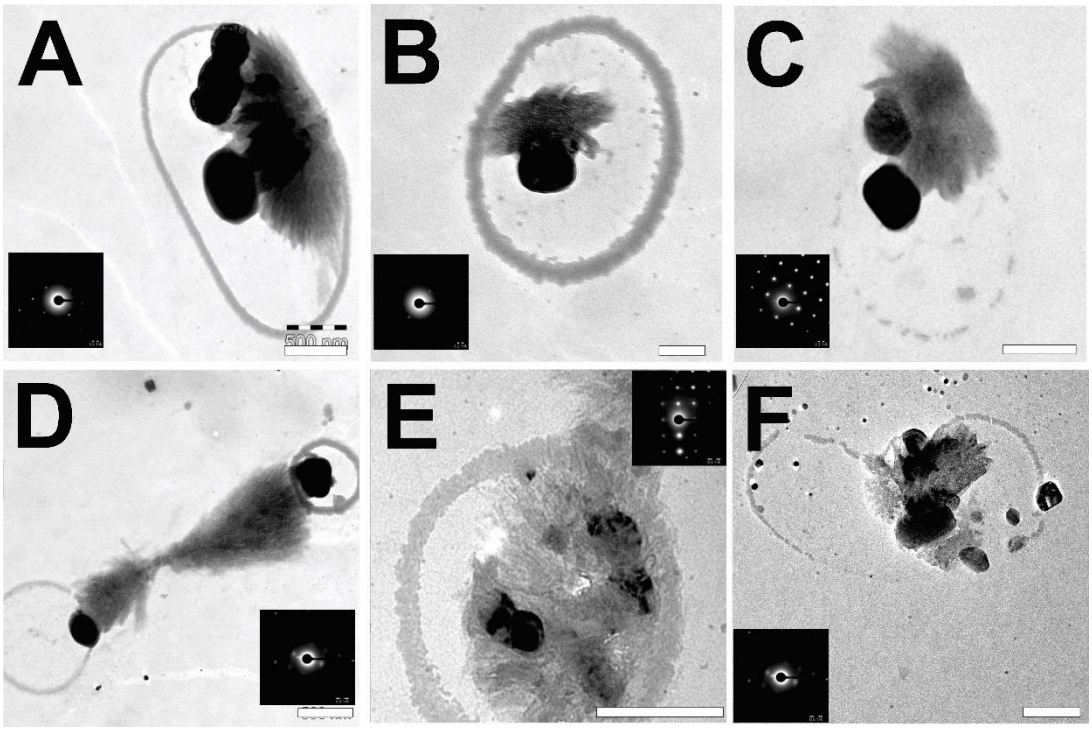


Figure S10. Representative cryo-TEM micrographs of the post-nucleation stage of calcium carbonate formation in the presence of SUMO-prisilkin-39 (1 mg/ml) at pH 9.75. Arrows indicate deformed vesicles associated with vaterite superstructures. Scale bars represent (A, B) 200 nm and (C) 100 nm. Nanostructured textures and single crystal-like diffraction pattern indicate certain crystallographic co-orientation with exposed (001) faces[5].

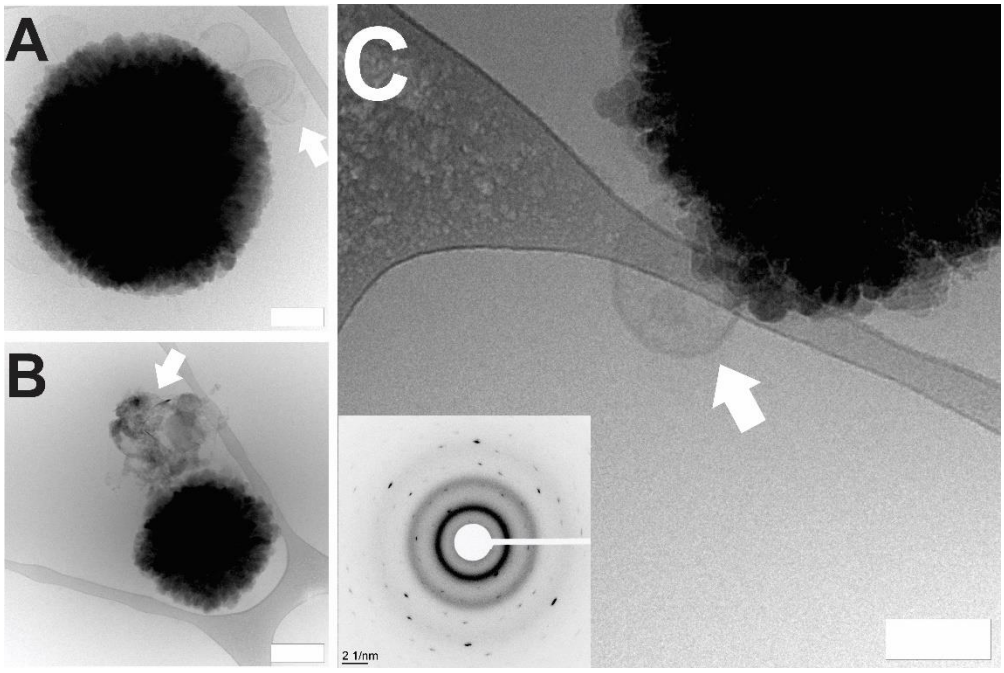


Figure S11. (A, B) ATR-FTIR spectra of calcium carbonate precipitated in the presence of different IDPs by direct mixing of precursors. An important observation is that an ethanol-assisted quenching of nucleated mineral does not yield an amorphous mineral phase, due to the protein denaturation by perturbations to the solvent composition.

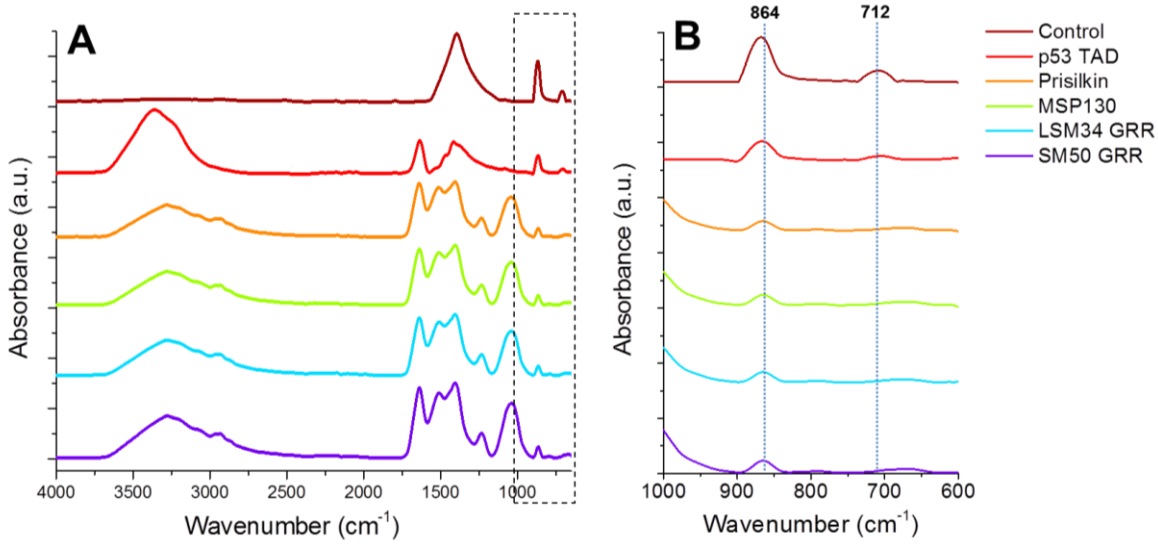


Table S2. Analyses of FTIR spectra depicted in Figure S11. The ratio of intensities of the respective ν_2 (854 cm^{-1}) and ν_4 (712 cm^{-1}) peaks of amorphous calcium carbonate and calcite are shown.

No	*Additive*	*I ν_2 ACC/ I ν_4 Calcite*
1	Reference	3.8
2	p53TAD	3.3
3	Prisilkin-39	6.0
4	MSP130	5.7
5	LSM34	5.1
6	SM50	5.7

Figure S12. (A) IR spectra of mineral-protein composites in D_2O illustrating the amide I and II regions. (B) Changes in intensities of ThT fluorescence emission at 490 nm with different contents of Arg, a protein solubilizing agent[6].

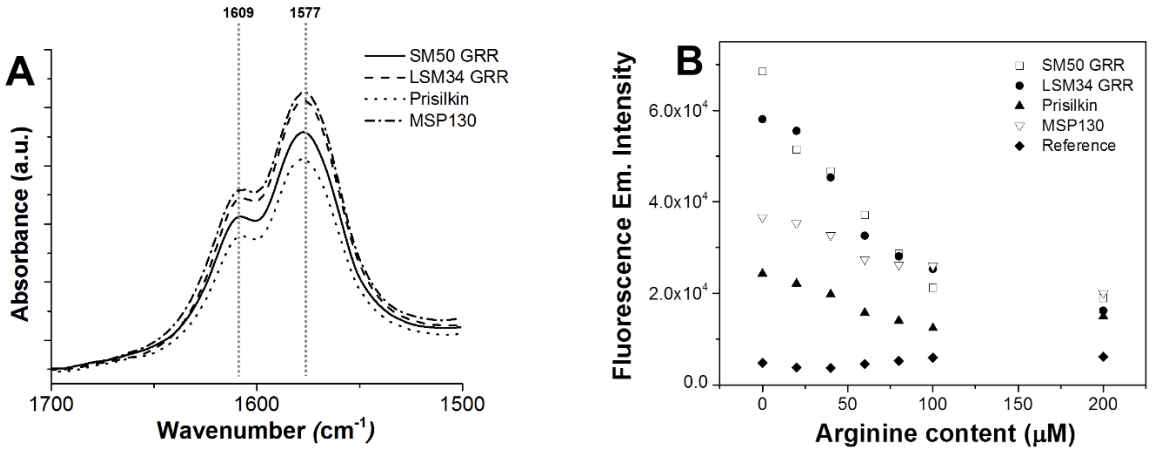


Figure S13. Fluorescence (left) and polarization (right) microscopy of SUMO fusions of LC-IDRs of (A) SM50, (B) LSM34, (C) MSP130 and (D) prisilkin-39 stained with thioflavin T (left) and congo red (right). Scale bars represent 50 μm.

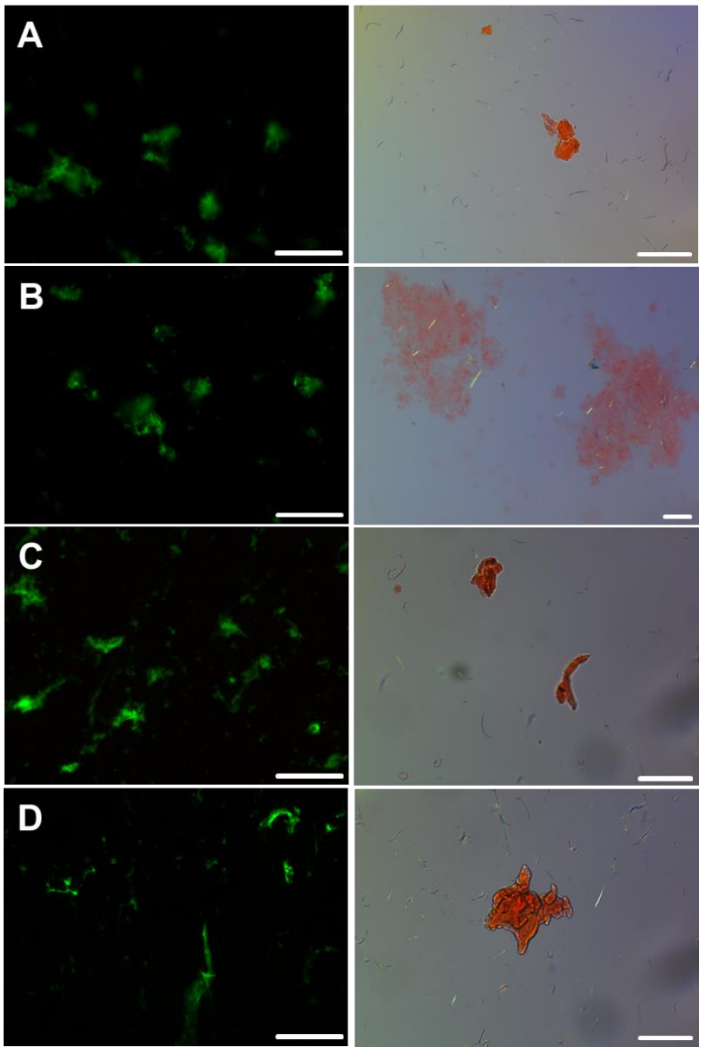


Figure S14. (A) Pseudo 3-dimensional plots for the 2-DSA analysis of the fusion protein containing the SM50 LC-IDR. . The smallest species (39 kDa) represents a protein monomer. The larger ones represent distinct multimers eg as marked in figure: 78 kDa (dimer), 120 kDa (trimer), 840 (21-mer). Certain data points are intermediate in molar mass, representing reversible self-association interactions, which are faster than the duration of the applied analytical ultracentrifugation experiment and are therefore averaged across the experimental duration. Representative TEM micrographs for the samples from (B) arginine solution (50 mM) and (C) carbonate buffer (20 mM).

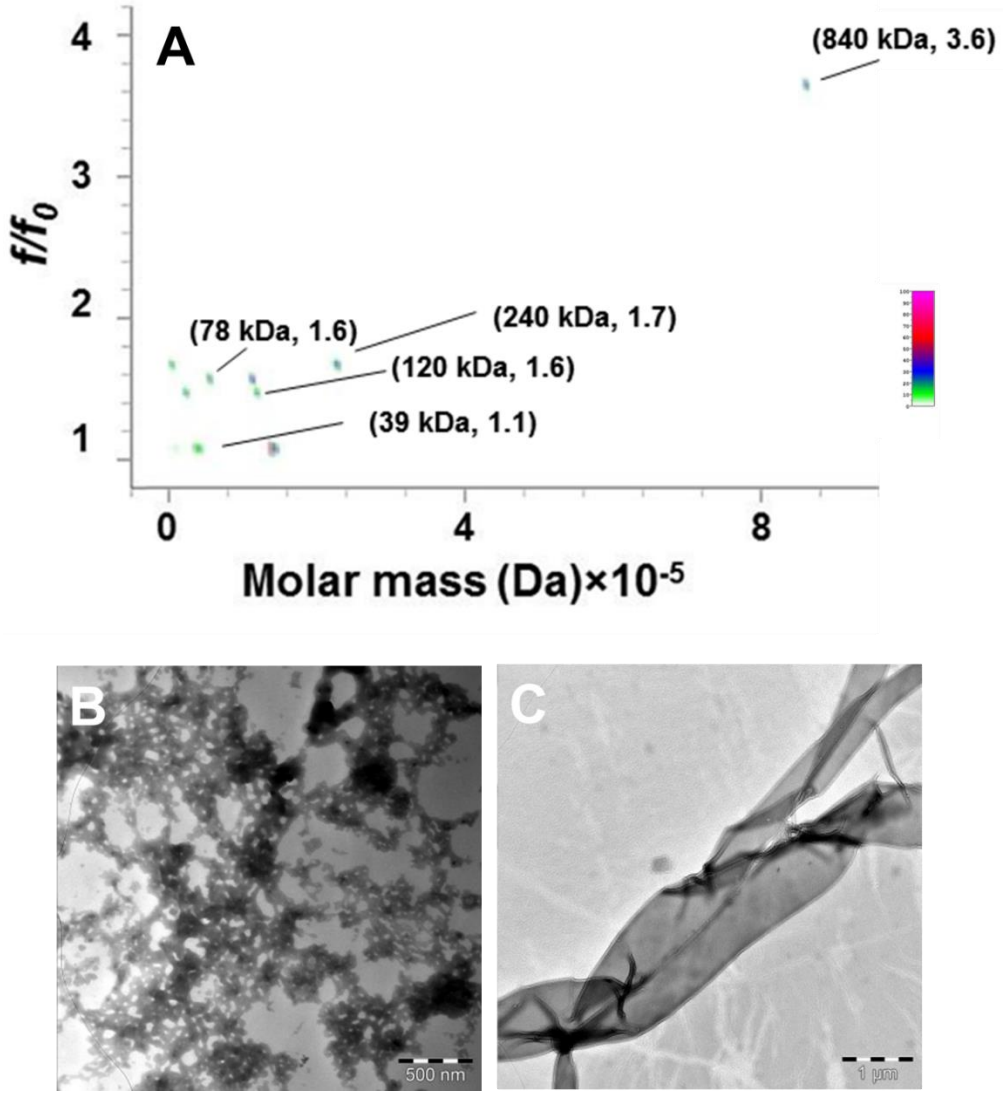
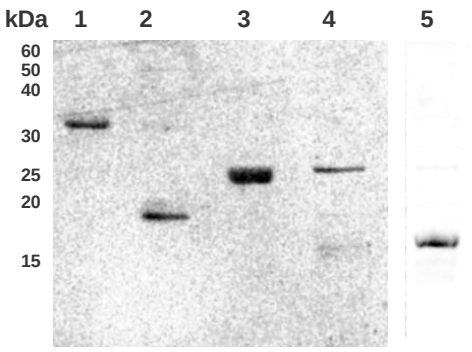


Figure S15. Representative SDS-PAGE of purified SUMO tagged proteins containing LC-IDRs of *Lane 1* SM50 *Lane 2* LSM34 *Lane 3* MSP130, *Lane 4* prisilkin and *Lane 5* p53TAD.



References

[1] a) J. Schultz, Milpetz, F., Bork, P., Ponting, C.P., *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 5857−5864; b) J. J. Ward, L. J. McGuffin, K. Bryson, B. F. Buxton, D. T. Jones, *Bioinformatics* **2004**, 20, 2138; c) K. Mann, F. H. Wilt, A. J. Poustka, *Proteome Science* **2010**, 8, 33; d) J. C. Wootton, S. Federhen, in *Methods in enzymology*, Vol. 266, Elsevier **1996**, p. 554.

[2] M. Wells, H. Tidow, T. J. Rutherford, P. Markwick, M. R. Jensen, E. Mylonas, D. I. Svergun, M. Blackledge, A. R. Fersht, *Proceedings of the National academy of Sciences* **2008**, 105, 5762.

[3] a) Z. Dosztányi, V. Csizmok, P. Tompa, I. Simon, *Bioinformatics* **2005**, 21, 3433; b) Z. R. Yang, R. Thomson, P. Mcneil, R. M. Esnouf, *Bioinformatics* **2005**, 21, 3369.

[4] R. Ashit, J. K. Berg, M. Kellermeier, D. Gebauer, *European Journal of Mineralogy* **2014**, 26, 537.

[5] A. W. Xu, M. Antonietti, H. Cölfen, Y. P. Fang, *Advanced Functional Materials* **2006**, 16, 903.

[6] K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J. S. Philo, T. Arakawa, *Biotechnology progress* **2004**, 20, 1301.