Designing protein-material interfaces

Shuai Zhang,[1,2] Harley Pyles,[3,4] David Baker[3,4,5,*] and James J. De Yoreo[1,2,*]

[1]Physical and Computational Sciences, Pacific Northwest National Laboratory, Richland, Washington, 99352, USA.
[2]Department of Materials Science and Engineering, University of Washington, Seattle, Washington, 98195, USA
[3]Department of Biochemistry, University of Washington, Seattle, Washington, 98195, USA.
[4]Institute for Protein Design, University of Washington, Seattle, WA 98105, USA.
[5]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98105, USA.

*Address correspondence to: james.deyoreo@pnnl.gov and dabaker@uw.edu

Abstract
This article addresses recent advances in using *de novo* protein design to create coherent interfaces between proteins and inorganic materials, either through protein self-assembly on crystal lattices or through directed nucleation and growth of crystals by protein scaffolds. Inspired by natural protein-crystal interfaces , we focus on a class of designed helical repeat proteins that present a repeating pattern of charged amino acid residues. We describe the use of in situ imaging and spectroscopic methods to investigate both the assembly of these proteins and their ability to direct crystal nucleation and growth. The findings reveal the importance of surface charge, facet-specific binding, solvent organization, and, more generally, the balance of protein-substrate-solvent interactions in determining how organized protein-materials interfaces emerge. Moreover, the results demonstrate the vast potential of protein design in materials science and elucidate the mechanisms by which interactions between biomolecules and inorganic surfaces lead to unique materials and morphologies.

Introduction

From harvesting solar energy, to purifying water, to capturing CO_2, living organisms have solved some of the most vexing challenges now facing humanity. They have done so by creating a vast library of proteins that can assemble into hierarchical structures to control the transport and transformation of matter and energy and direct the mineralization of inorganic components to form high-performance hybrid materials. While the high information content contained within the intricate sequences of proteins is crucial for accomplishing these tasks, assembly and mineralization are nonetheless constrained to proceed according to the physical laws that govern all such processes. An ability to design proteins at will, combined with an understanding of the mechanisms by which biological systems successfully manipulate those laws to create hierarchical materials, holds the promise of ushering in an era of materials design to address our most pressing technological challenges.

Principles from natural protein-material systems

Underlying many technological challenges that protein design can address[1, 2] is a need to create coherent interfaces between proteins and inorganic materials, either through protein assembly on the surfaces of materials or through the formation of materials on protein scaffolds. Two examples of natural protein systems illustrate principles that inspire approaches to designing proteins that carry out these processes. The first comes from the structures of certain classes of ice-binding proteins (IBPs) that present surfaces exhibiting an approximate epitaxial match to the ice lattice, which enables modulation of ice formation by binding ice nuclei through

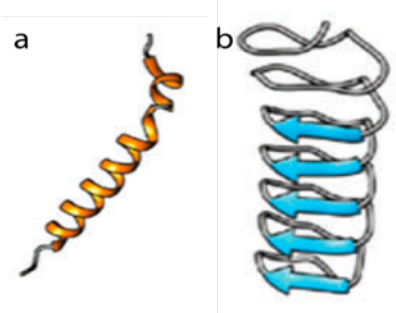

Figure 1 – (a) Type 1 and (b) insect antifreeze proteins consisting of α-helices and β-solenoids, respectively.

preorganized ice-like waters.[3] A number of these proteins are composed of α-helices or β-solenoids that present a flat interface with repeating patterns of amino-acid side-chains (**Figure 1**a, b), suggesting a basic strategy for *de novo* design of proteins that can similarly match the lattices of other inorganic crystals.

The second example comes from proteins of amelogenin that mineralize calcium phosphate to form the enamel of teeth, which is comprised of hydroxyapatite (Hap).[4] Amelogenin contains an amyloid region that drives its folding and assembly under appropriate solution conditions into a β-sheet nanoribbon of repeating amelogenin units, with the hydrophobic side of one β-sheet nanoribbon juxtaposed against the hydrophobic side of another to create a cross-β-sheet architecture (**Figure 2**a).[5] Moreover, when assembled on a hydrophobic surface like highly ordered pyrolytic graphite[6] (HOPG) or polystyrene,[7] the substrate provides a hydrophobic interface and thus a single β-sheet nanoribbon forms on the graphite surface, presenting the hydrophilic side to the solution. Imaging by atomic force microscopy (AFM) shows that these nanoribbons align on the graphite surface along three equivalent directions of the graphite lattice (**Figure 2**b).[6] When exposed to supersaturated calcium phosphate solution, amorphous calcium phosphate (ACP) nucleates on the nanoribbons (**Figure 2**c) at rates that far exceed those measured to date on any other substrate or protein scaffold, including collagen (col), which forms the scaffold on which Hap forms in bones, amelogenin prepared in the form of nanospheres (Amel sph), or the protein enamelin (Enam), which is also a protein involved in tooth formation (**Figure 2**d). Moreover, the creation of charged sites through phosphorylation

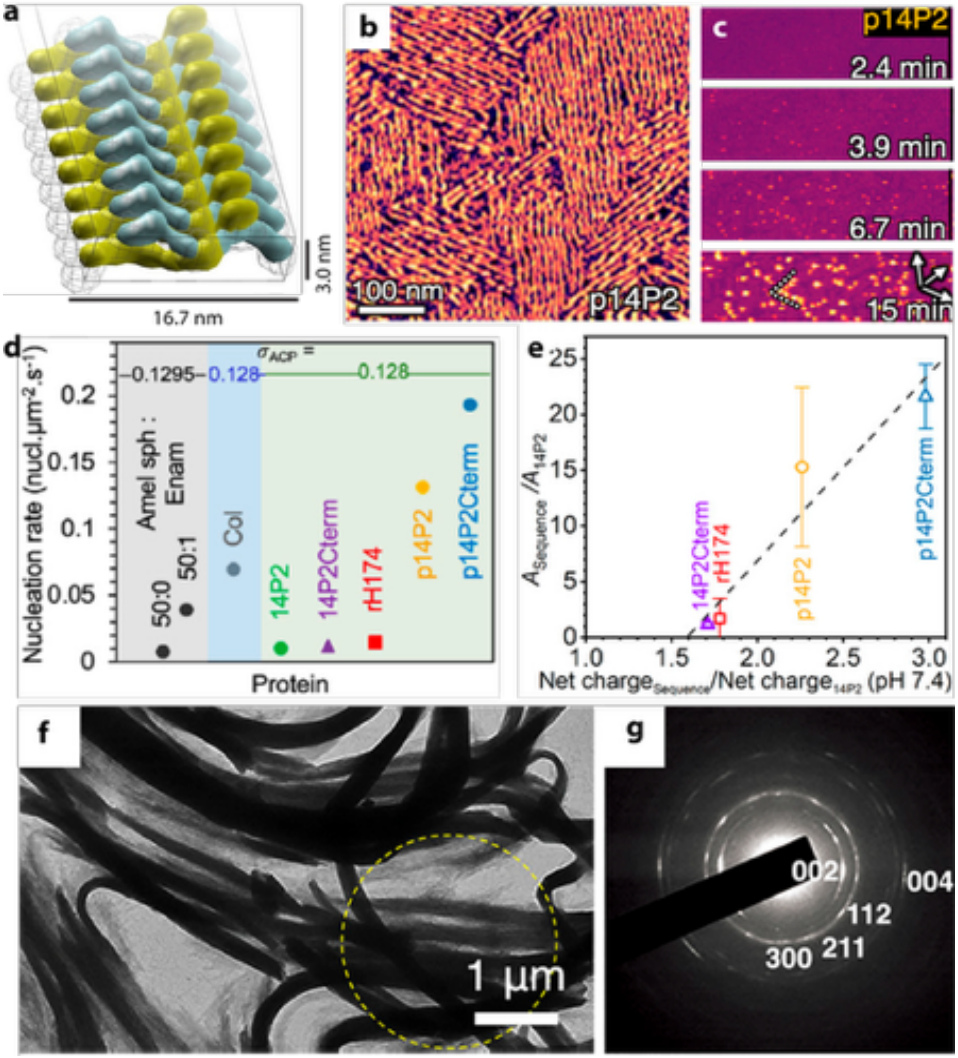

Figure 2-(a) CryoTEM reconstruction of nanoribbons of full-length human amelogenin (rH174). (b) AFM image of amelogenin peptide (p14P2) nanoribbons on HOPG. (c) In situ AFM images of ACP nucleation on p14P2 nanoribbons. Dashed lines indicate ACP nuclei aligning along nanoribbon directions (arrows). (d) ACP nucleation rates on recombinant human amelogenin (rH147), various amelogenin peptide segments (14P2, p14P2, 14P2Cterm, and p14P2Cterm), amelogenin nanospheres (Amel sph), enamelin (Enam), and collagen (Col). (e) kinetic pre-factor (A) vs. net charge at pH 7.4. (f, g) TEM images of mineralized amelogenin nanoribbons and the electron diffraction

of serine residues on the nanoribbons, which reproduces the phosphorylation seen in wild type amelogenin, further increases the nucleation rate.

Analysis of rate data for phosphorylated (denoted by prefix p) and unphosphorylated versions of multiple peptide fragments, including those with only the 14 amino acid residues that comprise the amyloid region of amelogenin (4P2 and p14P2), the same peptide fragments with the C-terminal HAP-binding segment appended onto it (14P2Cterm and p14P2Cterm), and a recombinant version of the full 174 amino acid long unphosphorylated human amelogenin protein (rH174) shows that the kinetic pre-factor in the expression for nucleation rate scales linearly with charge density and can span more than a factor of 20 from the least to the most effective sequence (**Figure 2**e). After nucleation, ACP then rapidly transforms to Hap, forming single crystals with the (001) axis aligned along the nanoribbons (**Figure 2**f,g).[8]

Formation of coherent protein-inorganic interfaces

The above examples illustrate two principles that have been exploited to interface proteins with inorganic materials through *de novo* design: 1) epitaxial lattice matching can be used to create a coherent interface between ordered assemblies of proteins and an inorganic crystal surface, and 2) a repeating pattern of charged residues on a flat interface can serve to drive formation of an inorganic phase.

Pyles et al.[9] used the first of these principles to create a designed helical repeat (DHR) protein that binds to the (001) face of muscovite (m-) mica (**Figure 3**a). Glutamate residues, which contain carboxylic side chains, were placed along an alpha-helix at positions separated by integer multiples of the 5.2 Å nearest-neighbor spacing between K^+ positions in the m-mica (001) lattice, and the dimension of the DHR backbone was selected to give a distance equal to that same spacing between helices in a sequence of N repeat units, thus ensuring that all successive rows of carboxylic groups fall on K^+ positions (**Figure 3**b). The resulting DHR10-mica-N (DHR-mica) proteins were indeed observed to bind along each of the three nearest-neighbor directions (**Figure 3**c, d). Moreover, at high protein and K^+ concentrations the resulting protein nanoribbons formed liquid crystal phases ranging from a high density disordered (HDD) phase, which is a kinetically trapped state, to a highly ordered smectic (Sm) phase (**Figure 3**e, f). The introduction of attractive dimeric interactions between the ends of the protein nanoribbons led to films of aligned arrays of continuous single-protein wide nanoribbons (**Figure 3**g), while designs that incorporated a trimeric interaction at one rod end enabled assembly of hexagonal pore networks with a pore diameter adjustable in increments of ~2 nm via the choice of the repeat number N (**Figure 3**h).

Experiments with DHR-mica protein assembly under a range of electrolyte concentrations and on different types of mica revealed the importance of interfacial solution structure on protein assembly.[9] When the proteins were assembled on fluorophlogopite (f-) mica under the same conditions that produced the smectic phase seen on m-mica (**Figure 4**a, left panel), the proteins remained in an HDD phase (**Figure 4**a, right panel). Changes in either protein or electrolyte concentrations by as much as an order of magnitude failed to alter this distinction. This difference in assembly behavior was unexpected because both the (001) surfaces of both micas consist of a 2D hexagonal network of tetrahedrally-coordinated Si-Al rings that surround cavities in which K^+ ions reside (Figure 3a). However, there is a subtle structural difference arising from ions that lie below the surface. In m-mica, sitting below the surface layer are OH^- ions that point towards the surface and are oriented along only two of the three nearest neighbor K^+ directions. The interaction of the OH^- ions with the surface layer causes two of the Si and Al atoms in the rings to be depressed from the surface by ~0.1Å. In f-mica, the OH molecules are replaced by F atoms with charge balance achieved through the replacement of Al with Mg in the tetrahedral rings. Consequently, the asymmetry is eliminated.

To understand how the structural differences between m- and f-mica lead to these contrasting outcomes of DHR-mica-N assembly, a combination of molecular simulations and 3D atomic force microscopy was used to relate those outcomes to the interactions of the proteins with the mica surface in electrolyte solutions,[10] as well as the interfacial solution structure above the surface.[11] The results show that, while the distribution

of surface K^+ sites in the two micas is unaffected, the structure of the overlying hydration layers differs significantly. (**Figure 4**b-e) The lateral pattern of water density in the first hydration layer is predicted to be hexagonal for both micas (**Figure 4**b). However, the water dipoles, which point in three equivalent directions on f-mica, are co-aligned along a single direction in m-mica. Consequently, the second hydration layer on f-mica also has hexagonal symmetry, but that of m-mica consists of rows pointing along the unique direction, (**Figure 4**c) which is the same direction along which the DHR-mica proteins are aligned (Figure

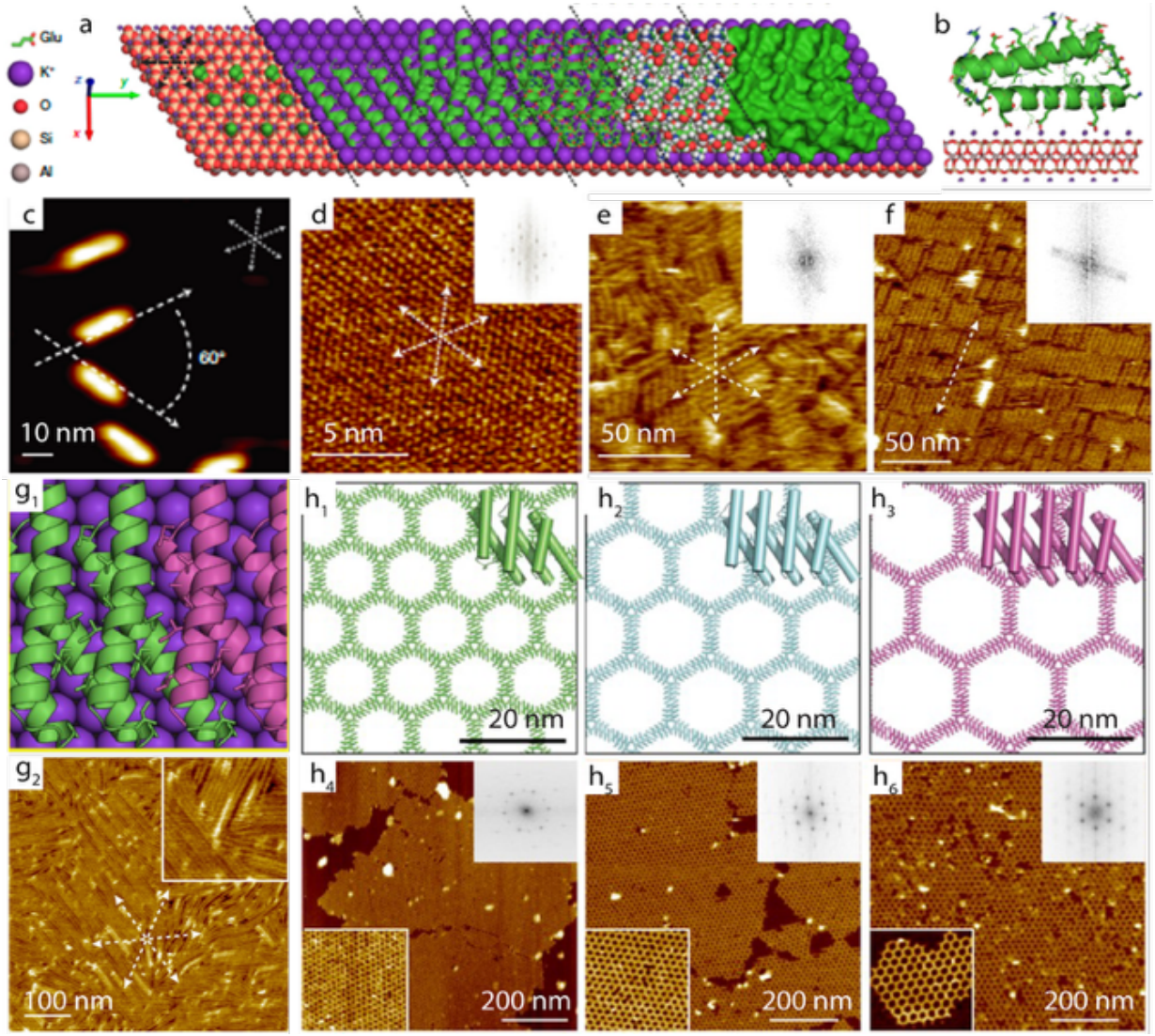

Figure 3 – (a) Model of DHR10-mica18 lattice-matched to the mica (001) surface through the K+ sublattice. (b) Projection view of DHR10-mica18 along the y direction. (c,d) AFM images of DHR10-mica18 adsorbed on mica and the lattice of the mica substrate in 10 mM KCl. The inset is the FFT of Panel d and the dashed arrows indicate the direction of the mica lattice. (e,f) AFM images of DHR10-mica18 assembled on mica in 100 mM and 3M KCl, respectively. The panels show HDD and Sm phases, respectively. The insets are the FFTs. (g) The top panel shows the intramolecular repeat–repeat interface and the intermolecular fibre interface of DHR10-mica18-non-cap (NC). The bottom panel shows AFM images of DHR10-mica18-NC in 3M KCl. (h) The top panel shows the computational models and monomers of DHR10-micaN-H hexagonal arrays for X=3, 4, 5. The bottom panel shows the AFM images of DHR10-micaN-H hexagonal arrays for N=3, 4, 5 in 3M KCl.

4a, right panel).

The results obtained with the DHR-mica proteins demonstrate the ability to design protein-material interfaces that utilize epitaxial relationships to drive ordered assembly, as well as the versatility that *de novo* design provides in obtaining a wide range of self-assembled structures through the introduction of protein-protein interactions (Figure 3). However, the results obtained at high protein concentrations show that colloidal forces that reflect the collective behavior of both the proteins and the solvent, as well as the details of hydration structure, both of which are inherent to interfacial systems, need to be considered when designing protein-materials interfaces.

Protein-directed nucleation and growth of inorganic materials

Davila-Hernandez et al.[12] followed the 2nd principle discussed above by exploiting the repeating pattern of charged residues on flat DHR protein surfaces to drive nucleation of calcium carbonate. The conceptual approach exploited the fact that an interface which binds strongly to a crystal plane will create a low interfacial energy for crystals nucleated on that plane, thus lowering both the free energy barrier to nucleation, which scales with the cube of the overall interfacial energy, and the critical size above which growing particles reach thermodynamic stability (**Figure 5**a).

Starting with a series of DHR backbones with varying patterns of glutamate or aspartate residues, the pathways and products of calcium carbonate nucleation were compared to those observed in protein-free solutions as well as those containing bovine serum albumin (BSA), which served as a negative control presenting similar overall net-negative charge, but lacking either a repeating pattern or a flat interface (**Figure 5**b).

Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy and ex situ Transmission electron microscopy (TEM) showed that, both in the absence of any protein and in the presence of BSA, crystallization

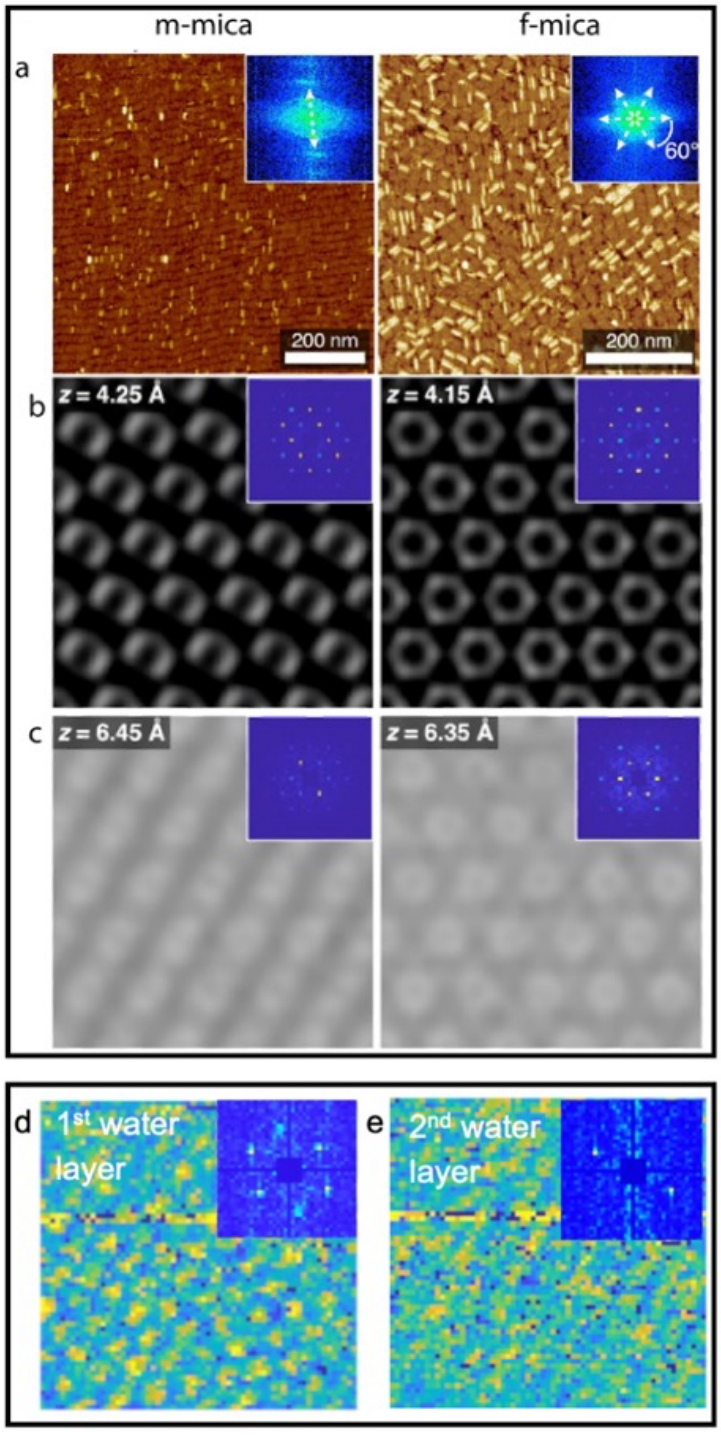

Figure 4 – (a) AFM images of DHR-mica18 assembly on m-mica and f-mica in 3M KCl. (b,c) 2D slices through predicated 3D water oxygen densities at stated height z above m-mica and f-mica in 3m KCl. The insets are FFTs. (d,e) 2D slices through the force map from 3D AFM on m-mica showing the first two hydration layers in 2M KCl, respectively. The insets are FFT showing the three-fold and two-fold symmetry of the first and second layers, respectively.

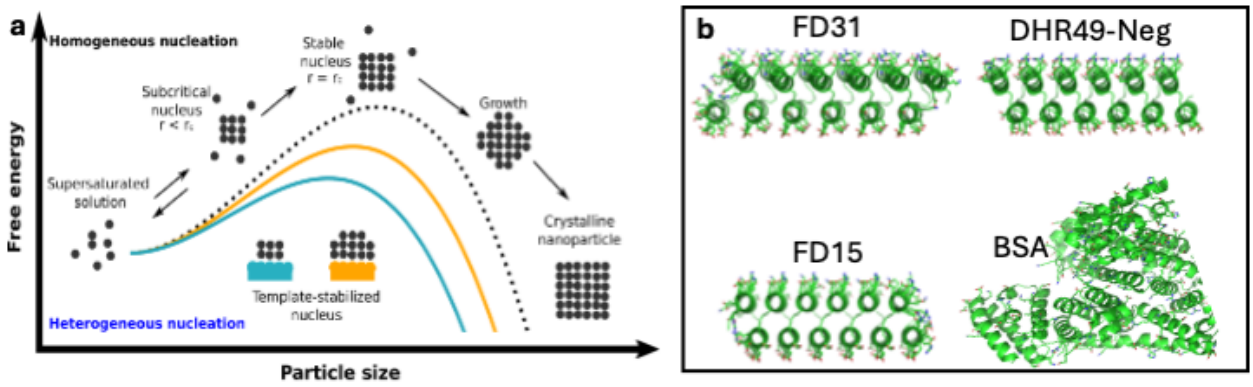

Figure 5 – (a) Concept behind templated nucleation of crystals. (b) Proteins FD31, FD15, DHR49-Neg, and BSA.

preceded first with the formation of amorphous calcium carbonate at high supersaturation or, at lower supersaturation. vaterite crystals that grew to 100s of nanometers in size before being replaced by calcite that grew to the micron-scale through dissolution-reprecipitation of the vaterite. The only effect of BSA was to slow the process, presumably due to complexation of Ca^{2+}. In contrast, in the presence of some of the DHR proteins (e.g., FD31 and DHR49-Neg), a high number density of nanocrystals on the order of 3 nm in diameter formed immediately, absent any precursor phase (**Figure 6**a). Moreover, in situ TEM revealed two distinct nucleation pathways. In the first, calcite nanocrystals nucleated throughout the solution in number densities that scaled with protein concentration, suggesting that individual proteins are capable of nucleating calcite particles. In the second, which occurred when the proteins were first mixed with the CaCl2 solution, highly crystalline, sheet-like supramolecular assemblies of the protein and calcium (**Figure 6**b) served as sacrificial templates, driving rapid nucleation of calcite upon introduction of a sodium bicarbonate and dissolving away as the calcite particles grew, presumably due to depletion of Ca^{2+} in the nearby solution (**Figure 6**c-e). Changes to the number of repeats in the proteins resulted in corresponding changes in the size of the primary calcite nanoparticles, with larger repeat numbers leading to smaller nanocrystals, presumably due to increased nucleation activity producing more, smaller nuclei or stronger crystal binding causing cessation of growth at smaller crystal size (**Figure 6**d-f).

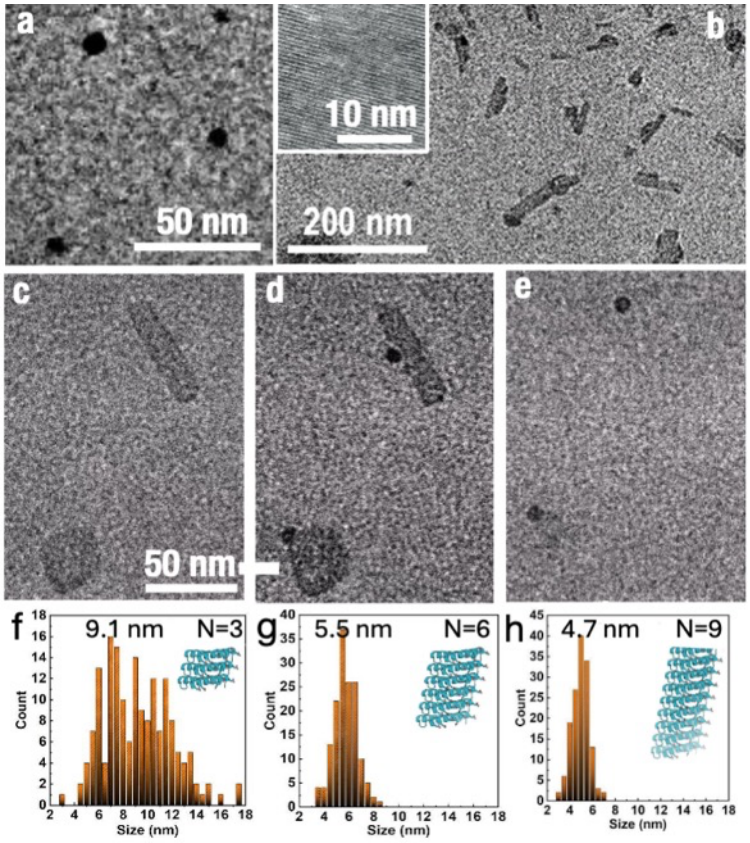

Figure 6 – (a) Calcite nanoparticles nucleated and stabilized by FD31. (b) Supramolecular assemblies of FD31 and Ca^{2+}. Inset shows their crystalline nature. (c-e) sequence of images showing calcite nucleation at supramolecular assemblies like those seen in (b). (f-g) Dependence of calcite crystal size distribution on number of repeats in the FD31 protein sequence.

Following nucleation, the proteins stabilized non-natural faces of the calcite nanocrystals — {110} in the case of FD31 and {202} in the case of DHR49-Neg. The larger nanocrystals were formed by the fusion of multiple nuclei and diffracted as single crystals. Thus, further growth occurred by repeated events of oriented attachment (OA) — i.e., attachment with crystallographic coalignment — of the nanocrystals (**Figure 7**a-c). As the crystals continued to grow by OA, the particle shapes gradually evolved to express the natural {104} facets but retained the particulate nature of the formation process (**Figure 7**d). In contrast to the behavior of FD31 and DHR49-Neg, other proteins that possessed different patterns of glutamic acid residues, had no measurable effect on either nucleation or growth. These results demonstrated that computationally designed proteins presenting repeating arrays on a flat interface can indeed be used to direct crystal nucleation and control growth, and that the specific pattern of charged residues can select for both the plane of nucleation and facet stabilization and the size of the resulting nanocrystals.

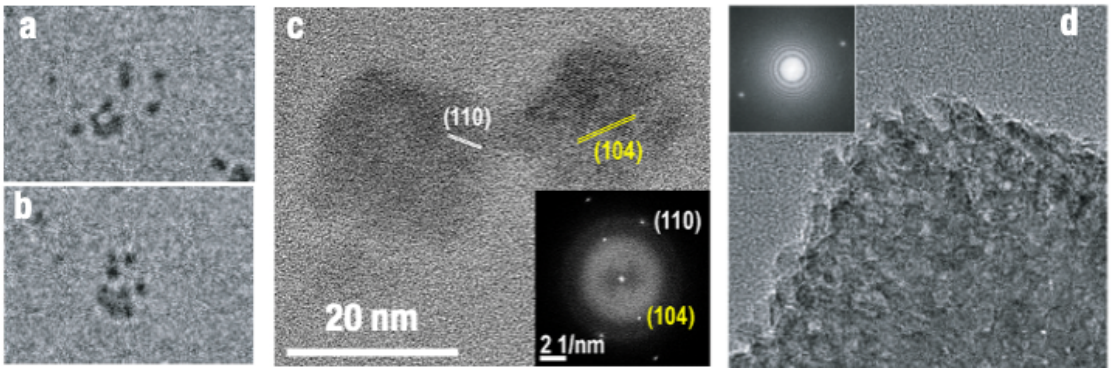

Figure 7 - (a) Sequence of images showing assembly of calcite nanocrystals like those seen in Figure 6a. (b) OA of two calcite nanocrystals exhibiting {110} facets. (c) Late-stage calcite formed by repeated OA events.

Summary and perspective

Using natural protein systems as an inspiration, we showed how proteins have been designed *de novo* to create an epitaxial match to an arbitrary crystal surface and how these designs can then be exploited to direct the nucleation of crystals and control their subsequent growth. We also show how *in situ* methods have been utilized to understand the mechanism of protein assembly and nanocrystal formation in these systems. The principle of assembly of a ordered protein-crystal interface through epitaxial matching of the protein to the crystal lattice has also been observed for a number of other systems of both proteins and peptides, including site-specifically modified enzymes (L-rhamnulose-1-phosphate aldolase (RhuA)) that assemble on mica through electrostatic interactions between the protein and the mica surface,[13] peptides selected by phage display for their binding affinity to MoS_2,[14] and silk fibroin proteins and de novo proteins on HOPG [15, 16] (**Figure 8**a-c). Moreover, the impact of subtle asymmetries in interfacial structure on assembly is not unique to the DHR-mica proteins. Other examples of symmetry breaking during assembly of proteins on mica were reported for both collagen[17, 18] and the RhuA (**Figure 8**d).[11] Thus the impact of epitaxy and interfacial structure on the architecture adopted by proteins at the materials interface seem broadly applicable across systems and are thus an important consideration in design. Similarly, the ability of repeating patterns of charged amino acid residues on a flat interface to template crystal formation both in the case of calcium phosphate nucleation on amelogenin and calcium carbonate on DHR proteins suggests a broadly applicable approach to design. Of course, for materials that differ dramatically from the ionic nature of salts, amino acid residues of a distinct character from acidic ones like glutamic acid can be expected to be more relevant. For example, in IBPs, the key amino acid often consists of repeating patterns of threonines, which structure ice-like waters on their surface to mediate ice-binding[3].

The research carried out using RhuA also highlights another approach to designing protein-material interfaces, which is to start with a natural protein as the basic building block and introduce site-specific modifications that enable assembly of the proteins[19] and then utilize inherent interactions between the

proteins and a substrate to drive that assembly to occur on a surface.[13] In the case of RhuA, electrostatic interactions are a key factor in driving adsorption and assembly on mica. RhuA is an electrostatically patchy particle, presenting a relatively flat negative C-terminal face and a four-legged positive N-terminal (**Figure 9**a). When chemically modified by introduction of cysteine residues at the four corners (to form ^{C98}RhuA), in bulk solution, the proteins form disulfide bonds between the cysteine residues, but the proteins adopt the alternating pattern of C-terminal up and N-terminal down due to avoid repulsive interactions between like-charged faces, forming a lattice that can be in an open or closed state. (**Figure 9**b). However, at low K^+ concentrations, mica surfaces are negatively charged; consequently, when mica is exposed to a solution of RhuA at low K^+ concentrations, it assembles primarily with the positive N-terminal side down (**Figure 9**c-e). In contrast, at high K^+ concentrations, the situation is reversed, and the proteins align with their negative C-terminal side down (Figure 9c-e). Moreover, the phase that forms at low K^+ is a closed lattice, but the phase that forms at high K^+ is an open lattice (Figure 9c-e).

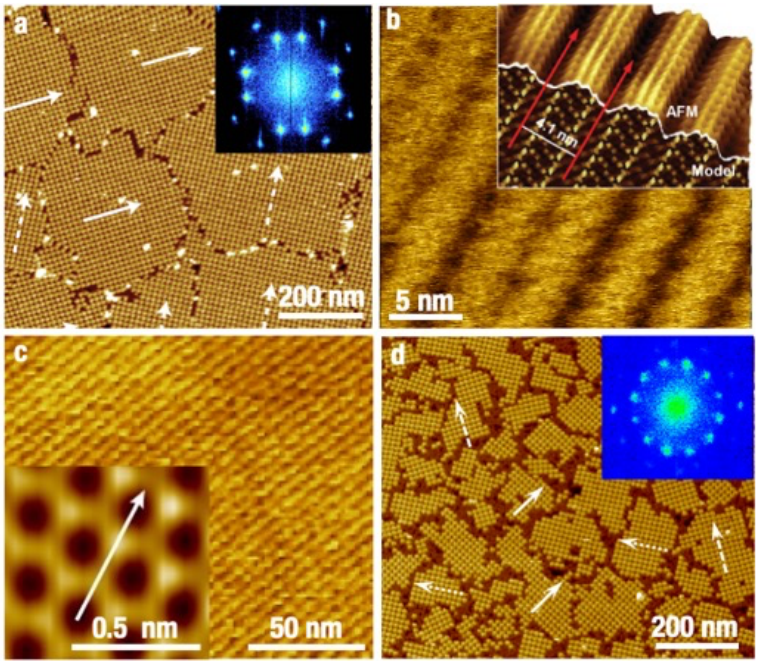

Figure 8 – (a) ^{C98}RhuA on m-mica (001). Inset is fast Fourier transform (FFT) of film showing that only two of the three equivalent directions of the nearest neighbor K^+ lattice are present. (b) Silk fibroin (SF) on HOPG. Inset shows high resolution image of HOPG lattice where arrow shows direction of SF rows. (c) MoS_2 binding peptides on MoS_2. Inset gives high resolution image and molecular model of the structure. (d) ^{C98}RhuA on f-mica. Inset is FFT of film showing

The results reported in the above studies demonstrate the versatility and fidelity of protein design as an approach to creating protein-inorganic interfaces and as a means for designing hybrid materials and directing the formation of inorganic phases. Moreover, the *in situ* studies and molecular simulations reveal the importance of surface charge, facet-specific binding, solvent organization, and, more generally, the balance of protein-substrate-solvent interactions in determining how organized protein-material interfaces emerge. These advances illustrate the vast potential of protein design in materials science and elucidate the mechanisms by which the interactions between biomolecules and materials lead to unique materials and morphologies.

Despite these advances, a major challenge yet to be resolved is the development of experimental methods to characterize the region between proteins and the materials on which they assemble. While a number of methods exist to probe the free surface of materials in solution, or the surface of proteins adsorbed to the interface with materials, including grazing incidence X-ray diffraction, three-dimensional AFM, and surface vibrational spectroscopy, probing the buried interface between proteins and surfaces has never been achieved. There is some hope that neutron scattering may offer a window into this region, but that has yet to be demonstrated. Similarly, defining the atomic-scale structure of the interface between protein templates and the minerals that form on them has proven challenging to achieve. Solid-state NMR can provide some information on the distances between amino acid side chains and ions in the crystal lattice and cryogenic TEM in principle can provide direct images of the interface, but the low number density of proteins at

mineral surfaces and the low contrast of the proteins, respectively, render these approaches difficult. If data on protein-material interfaces are ever to be used for protein design in the manner that the protein database has been used for design of proteins themselves, this problem will need to be surmounted.

Author contributions
All authors contributed to the writing of the manuscript.

Funding
The preparation of this review was supported at Pacific Northwest National Laboratory (PNNL) by the DOE Office of Science Basic Energy Sciences (BES), as part of the Energy Frontier Research Centers (EFRC) program: CSSAS, The Center for the Science of Synthesis Across Scales, under Award FWP 72448 and at the University of Washington by the Open Philanthropy Project Improving Protein Design Fund and The Howard Hughes Medical Institute. PNNL is operated by Battelle for the Department of Energy under contract no. DE-AC05-76RLO1830.

Competing Interests
The authors have no competing interests.

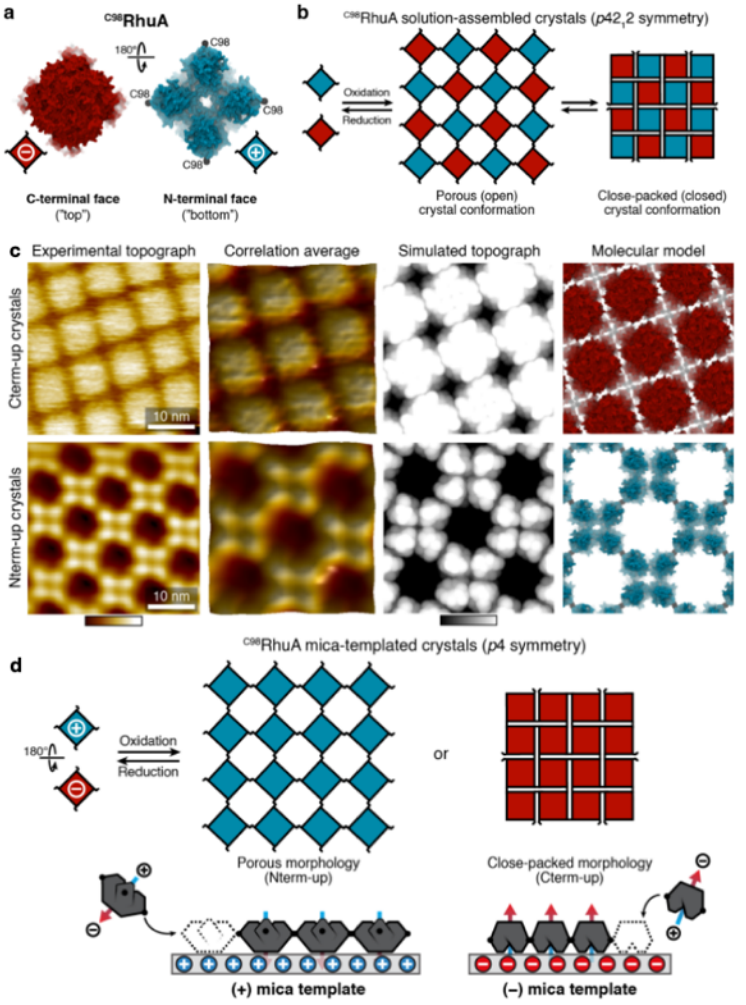

Figure 9 – (**a**) Solvent-excluded surfaces of ^{C98}RhuA termini faces, colored according to charge. C98 residues are shown as black spheres. (**b**) Schematic depicting C98RhuA solution self-assembly into 2D crystals and their ability to transition between open and closed lattices. (**c**) High-resolution AFM topographs and structural models of mica-templated crystals. Height ranges (gradients): 7 nm (experimental), 3 nm (simulated). (**d**) Cartoon overview of mica-templated self-assembly. Blue and red colors represent positive and negative charges, respectively.

References

(1) Brunette, T. J.; Parmeggiani, F.; Huang, P. S.; Habha, G. B.; Ekiert, D. C.; Tsutakawa, S. E.; Hura, G. L.; Tainer, J. A.; Baker, D. Exploring the repeat protein universe through computational protein design. *Nature* **2015**, *528* (7583), 580-+. DOI: 10.1038/nature16162.

(2) Huang, P. S.; Boyken, S. E.; Baker, D. The coming of age of de novo protein design. *Nature* **2016**, *537* (7620), 320-327. DOI: 10.1038/nature19946 From NLM Medline.

(3) Voets, I. K. From ice-binding proteins to bio-inspired antifreeze materials. *Soft Matter* **2017**, *13* (28), 4808-4823. DOI: 10.1039/c6sm02867e From NLM PubMed-not-MEDLINE.

(4) Gibson, C. W. The Amelogenin Proteins and Enamel Development in Humans and Mice. *J Oral Biosci* **2011**, *53* (3), 248-256. DOI: 10.2330/joralbiosci.53.248 From NLM PubMed-not-MEDLINE.
(5) Martinez-Avila, O.; Wu, S.; Kim, S. J.; Cheng, Y.; Khan, F.; Samudrala, R.; Sali, A.; Horst, J. A.; Habelitz, S. Self-assembly of filamentous amelogenin requires calcium and phosphate: from dimers via nanoribbons to fibrils. *Biomacromolecules* **2012**, *13* (11), 3494-3502. DOI: 10.1021/bm300942c From NLM Medline.
(6) Akkineni, S.; Zhu, C.; Chen, J. J.; Song, M.; Hoff, S. E.; Bonde, J.; Tao, J. H.; Heinz, H.; Habelitz, S.; De Yoreo, J. J. Amyloid-like amelogenin nanoribbons template mineralization via a low-energy interface of ion binding sites. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119* (19). DOI: ARTN e2106965119
10.1073/pnas.2106965119.
(7) Akkineni, S.; Doerk, G. S.; Shi, C.; Jin, B.; Zhang, S.; Habelitz, S.; De Yoreo, J. J. Biomimetic Mineral Synthesis by Nanopatterned Supramolecular-Block Copolymer Templates. *Nano Lett* **2023**, *23* (10), 4290-4297. DOI: 10.1021/acs.nanolett.3c00480 From NLM Medline.
(8) Bai, Y.; Yu, Z.; Ackerman, L.; Zhang, Y.; Bonde, J.; Li, W.; Cheng, Y.; Habelitz, S. Protein nanoribbons template enamel mineralization. *Proc Natl Acad Sci U S A* **2020**, *117* (32), 19201-19208. DOI: 10.1073/pnas.2007838117 From NLM Medline.
(9) Pyles, H.; Zhang, S.; De Yoreo, J. J.; Baker, D. Controlling protein assembly on inorganic crystals through designed protein interfaces. *Nature* **2019**, *571* (7764), 251-256. DOI: 10.1038/s41586-019-1361-6 From NLM Medline.
(10) Prelesnik, J. L.; Alberstein, R. G.; Zhang, S.; Pyles, H.; Baker, D.; Pfaendtner, J.; De Yoreo, J. J.; Tezcan, F. A.; Remsing, R. C.; Mundy, C. J. Ion-dependent protein-surface interactions from intrinsic solvent response. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (26). DOI: ARTN e2025121118
10.1073/pnas.2025121118.
(11) Alberstein, R. G.; Prelesnik, J. L.; Nakouzi, E.; Zhang, S.; De Yoreo, J. J.; Pfaendtner, J.; Tezcan, F. A.; Mundy, C. J. Discrete Orientations of Interfacial Waters Direct Crystallization of Mica-Binding Proteins. *J. Phys. Chem. Lett.* **2023**, *14* (1), 80-87. DOI: 10.1021/acs.jpclett.2c02948.
(12) Davila-Hernandez, F. A.; Jin, B.; Pyles, H.; Zhang, S.; Wang, Z. M.; Huddy, T. F.; Bera, A. K.; Kang, A.; Chen, C. L.; De Yoreo, J. J.; et al. Directing polymorph specific calcium carbonate formation with de novo protein templates. *Nat. Commun.* **2023**, *14* (1). DOI: ARTN 8191
10.1038/s41467-023-43608-1.
(13) Zhang, S.; Alberstein, R. G.; De Yoreo, J. J.; Tezcan, F. A. Assembly of a patchy protein into variable 2D lattices via tunable multiscale interactions. *Nat Commun* **2020**, *11* (1), 3770. DOI: 10.1038/s41467-020-17562-1 From NLM PubMed-not-MEDLINE.
(14) Chen, J. J.; Zhu, E. B.; Liu, J.; Zhang, S.; Lin, Z. Y.; Duan, X. F.; Heinz, H.; Huang, Y.; De Yoreo, J. J. Building two-dimensional materials one row at a time: Avoiding the nucleation barrier. *Science* **2018**, *362* (6419), 1135-+. DOI: 10.1126/science.aau4146.

(15) Shi, C.; Zorman, M.; Zhao, X.; Salmeron, M. B.; Pfaendtner, J.; Liu, X. Y.; Zhang, S.; De Yoreo, J. J. Two-dimensional silk. *Sci Adv* **2024**, *10* (38), eado4142. DOI: 10.1126/sciadv.ado4142 From NLM PubMed-not-MEDLINE.
(16) Brown, C. L.; Aksay, I. A.; Saville, D. A.; Hecht, M. H. Template-Directed Assembly of a de Novo Designed Protein. *J. Am. Chem. Soc*. 2002 ,124(24), 6846-68488. DOI: 10.1021/ja026127
(17) Narayanan, B.; Gilmer, G. H.; Tao, J.; De Yoreo, J. J.; Ciobanu, C. V. Self-assembly of collagen on flat surfaces: the interplay of collagen-collagen and collagen-substrate interactions. *Langmuir* **2014**, *30* (5), 1343-1350. DOI: 10.1021/la4043364 From NLM Medline.
(18) Gisbert, V. G.; Benaglia, S.; Uhlig, M. R.; Proksch, R.; Garcia, R. High-Speed Nanomechanical Mapping of the Early Stages of Collagen Growth by Bimodal Force Microscopy. *ACS Nano* **2021**, *15* (1), 1850-1857. DOI: 10.1021/acsnano.0c10159 From NLM PubMed-not-MEDLINE.
(19) Brodin, J. D.; Ambroggio, X.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chem.* **2012**, *4*, 375–382. DOI: 10.1038/nchem.1290