

Sequence-structure-binding relationships reveal adhesion behavior of the Car9 solid-binding peptide: an integrated experimental and simulation study

Brittney Hellner, Sarah Alamdari, Harley Pyles,² Shuai Zhang,³ Arushi Prakash, Kayla G. Sprenger, James J. De Yoreo,³ David Baker,^{1,2,4} Jim Pfaendtner*, and François Baneyx*

Department of Chemical Engineering,¹Department of Biochemistry,²Institute for Protein Design, and ⁴Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA

³Physical Sciences Division, Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA

* jpfaendt@uw.edu

* baneyx@uw.edu

ABSTRACT

Solid-binding peptides (SBPs) recognizing inorganic and synthetic interfaces have enabled a broad range of materials science applications and hold promise as adhesive or morphogenetic control units that can be genetically encoded within desirable or designed protein frameworks. To date, the underlying relationships governing both SBP-surface and SBP-SBP interactions and how they give rise to different adsorption mechanisms remain unclear. Here, we combine protein engineering, SPR characterization, and molecular dynamics (MD) simulations initiated from Rosetta predictions to gain insights on the interplay of amino acid composition, structure, self-association and adhesion modality in a panel of variants of the Car9 silica-binding peptide (DSARGFKKPGKR) fused to the C-terminus of superfolder green fluorescent protein (sfGFP). Analysis of kinetics, energetics and MD-predicted structures shows that the high affinity binding of Car9 to the silanol-rich surface of silica is dominated by electrostatic contributions and a spectrum of several persistent interactions that, along with a high surface population of bound molecules, promote cooperative interactions between neighboring SBPs and higher order structure formation. Transition from cooperative to Langmuir adhesion in sfGFP-Car9 variants occurs in concert with a reduction of stable surface interactions and self-association, as confirmed by AFM imaging of proteins exhibiting the two different binding behaviors. We discuss the implications of these results for the *de novo* design of SBP-surface binding systems.

INTRODUCTION

Solid-binding peptides (SBPs) isolated from biomineralizing proteins¹ or selected by biopanning,²⁻⁴ have proven remarkably useful to modify surfaces, fabricate functional architectures and devices, and control the morphology, crystallography and properties of inorganic phases.^{3, 5-13} For these processes, binding and morphogenetic outcomes result from the complex interplay of solution conditions, surface chemistry and crystallography, and SBP composition, conformation, and dynamics at the biotic-abiotic interface. The local chemical and structural context experienced by an SBP may also be important if the peptide is genetically inserted within a larger protein or displayed on the surface of an organism. Additionally, host protein frameworks or displaying organisms may themselves contribute to inorganic adhesion, nucleation or growth. Gaining a holistic understanding of these factors could pave the way to the use of solid-binding proteins for the predictive design of materials with user-specified properties and under synthesis conditions that are both mild and environmentally friendly.

Within this context, the relationship between SBP interfacial structure and inorganic-binding properties is of particular interest. Models derived from experimental data and advanced simulations hold that most combinatorially-selected SBPs sample an ensemble of conformations in solution (i.e., they are intrinsically disordered),¹⁴ and that they achieve high-affinity binding either by assuming many different surface-adsorbed conformational states (entropic binders), or by making long-lasting contacts with the surface through one or more anchor residues (enthalpic binders).^{8, 15} For the latter class of binders, NMR studies of the titania-binding TBP hexapeptide¹⁶ (RKLPDA) in contact with TiO₂ and SiO₂ nanoparticles,¹⁷⁻¹⁸ suggest that acquisition of structure at the interface is important to orient anchor residues for optimal surface interactions. This view is consistent with the well-established role of all levels of structure in controlling the interaction of

biomineralizing proteins and peptides with inorganic interfaces.¹⁹⁻²³ To date, however, there is no experimental or computational path to identify those rare families of SBPs with the capacity to present both types of binding mechanisms.

The Car9 (DSARGFKKPGKR) SBP binds to the carboxyl- and hydroxyl-rich edges²⁴ and ends²⁵ of carbon nanostructures and to the silanol-rich surface of silica.²⁶ This rather promiscuous binding profile and the fact that Car9-silica interactions can be disrupted by addition of free lysine or arginine has enabled applications ranging from affinity protein purification to microcontact printing, and from controlled protein release to hybrid materials assembly.²⁵⁻³⁰ Using surface plasmon resonance (SPR), we found that proteins engineered with a Car9 extension exhibit distinctive sigmoidal adsorption sensorgrams on silica chips.³¹ To account for this behavior, we developed a rate law-based model in which an initial adsorption step involving cooperative interactions between nearest neighbors is followed by a second step associated with slower conformational rearrangements at the interface.³¹ This simple two-step model accurately captures adsorption kinetics, and the fitting parameters provide insights on the extent of cooperativity and the magnitude of the reaction rates. Remarkably, when four of the five basic residues present in the Car9 peptide are converted to glutamines (yielding Car9Q4 of sequence DSARGFQQPGQQ), silica adsorption is significantly reduced but not abolished, and it becomes possible to model adsorption kinetics with a classic Langmuir fit.³¹

In this study, we combine SPR measurements, Rosetta calculations and molecular dynamics (MD) simulations to gain insights on the interplay of amino acid composition, structure and silica adhesion modalities in a panel of Car9 variants fused to the C-terminus of superfolder green fluorescent protein (sfGFP), a thermostable variant of GFP that does not dimerize.³² We conduct an analysis of the kinetics, energetics and predicted structures from MD to understand the superior

silica-binding characteristics of the Car9 SBP, present additional examples of sfGFP-Car9 variants displaying cooperative or Langmuir binding mechanisms, and show through simulations and high-resolution AFM imaging that the two different adsorption modalities are due to shifts in the type of dominant interactions at the peptide-surface interface.

RESULTS

Predicting Car9 conformation at the silica interface. We used classical MD simulations to connect molecular scale structural details of the Car9-silica interface with insights from SPR experiments. To reduce the conformational space, we first modeled the Car9 peptide using the

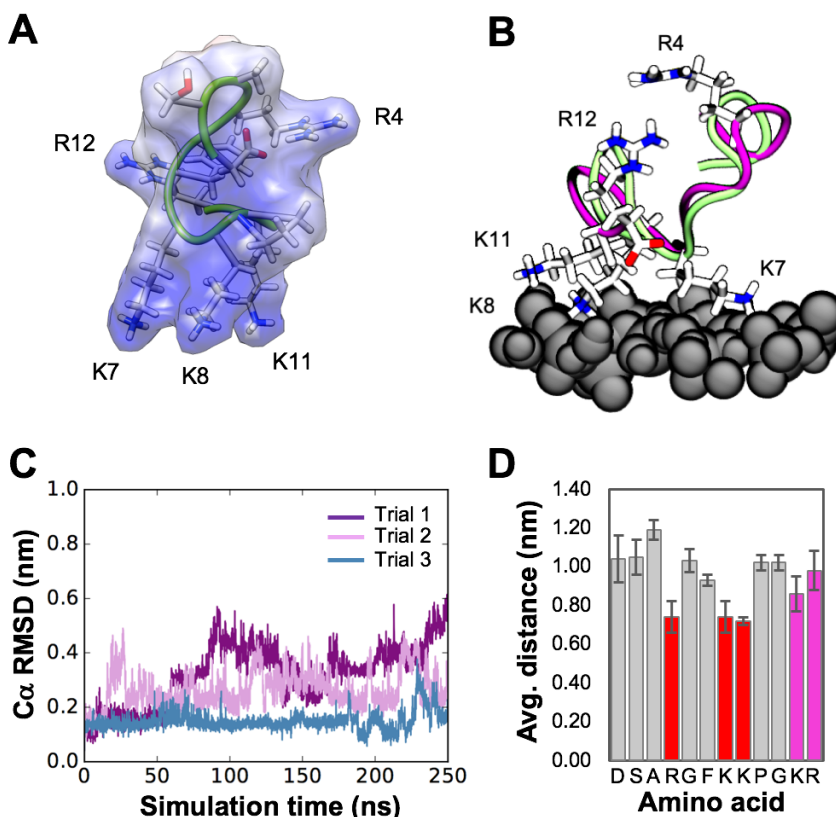


Figure 1. Predicted structure of the Car9 dodecapeptide at the silica interface. (A) Low energy structure proposed by Rosetta. The electrostatic surface is colored between $-10 kTe$ (red, not visible) and $+10 kTe$ (blue). (B) Superimposition of Rosetta structure (green) and top-ranked cluster determined from trial 1 of MD-simulated structure at the silica interface (pink). (C) Root mean square deviation (RMSD) traces of C α atoms between simulation and Rosetta-offered structure for three trials of Car9 adsorbed to silica. (D) Distance of the center of mass (COM) of each residue from the silica surface averaged over the last 50 ns of the three simulation trials. Primary and secondary anchor residues are shown in red and pink, respectively.

Rosetta software suite. *Ab initio* structure prediction was performed as described for the validation of designed peptide structures.³³ From nearly 100,000 generated structures, Rosetta offered five low energy solutions (**Figure S1** in Supporting Information) sharing a common architecture (**Figure 1A**). In it, K7, K8 and K11 form a tripod that protrudes from a bulbous region defined by hairpin P9-G10 and stabilized by a buried F6. The side chains of the two remaining basic residues (R4 and R12) project in opposite directions from the equator of the bulb, endowing the peptide with a uniformly positive charge (**Figure 1A**). Although these results are suggestive of a binding mode that involves the tripod, it is important to note that Rosetta has not been benchmarked on flexible peptides. Plans are under way to collect experimental constraints from solid state NMR to assess the accuracy of the model.

We used the Rosetta-proposed low energy configuration as a starting point to conduct MD simulations in the presence of an explicit silica interface. The peptide was parametrized with the CHARMM36³⁴ force field, and the INTERFACE³⁵⁻³⁶ force field was used to represent the (100) surface of quartz under near neutral conditions. We previously used this surface to simulate the binding of the silaffin-derived R5 peptide to silica.²³

As shown in **Figures 1B** and **1C**, there was good agreement between the C α traces calculated by Rosetta and those obtained in triplicate 250 ns simulation runs that included a silica interface. Inspection of average distances between the center of mass (COM) of individual peptide residues and the surface across all three simulation trials revealed a significant involvement of R4, K7 and K8 in silica interactions (**Figure 1D**; also see **Figure 4B** below and **Figure S2** in Supporting Information): there was very low variance in side chain distances once these residues were bound to the surface and binding was strong in all simulations. We refer to these amino acids as primary

anchor residues. The analysis further suggested that the side chains of K11 and R12 contribute to silica binding through more variability in the surface contacts (secondary anchor residues).

Contribution of structure and electrostatics to silica binding. To understand the relationship between amino acid sequence and binding motifs, we constructed a panel of sfGFP-Car9 variants containing substitutions in residues proposed to play a structural or surface anchoring role. Alanine substitutions were used to produce the F6A, P9AG10A, and K8AK11A mutants while glutamine substitutions were used to construct the R4QR12Q variant. The alternative choice

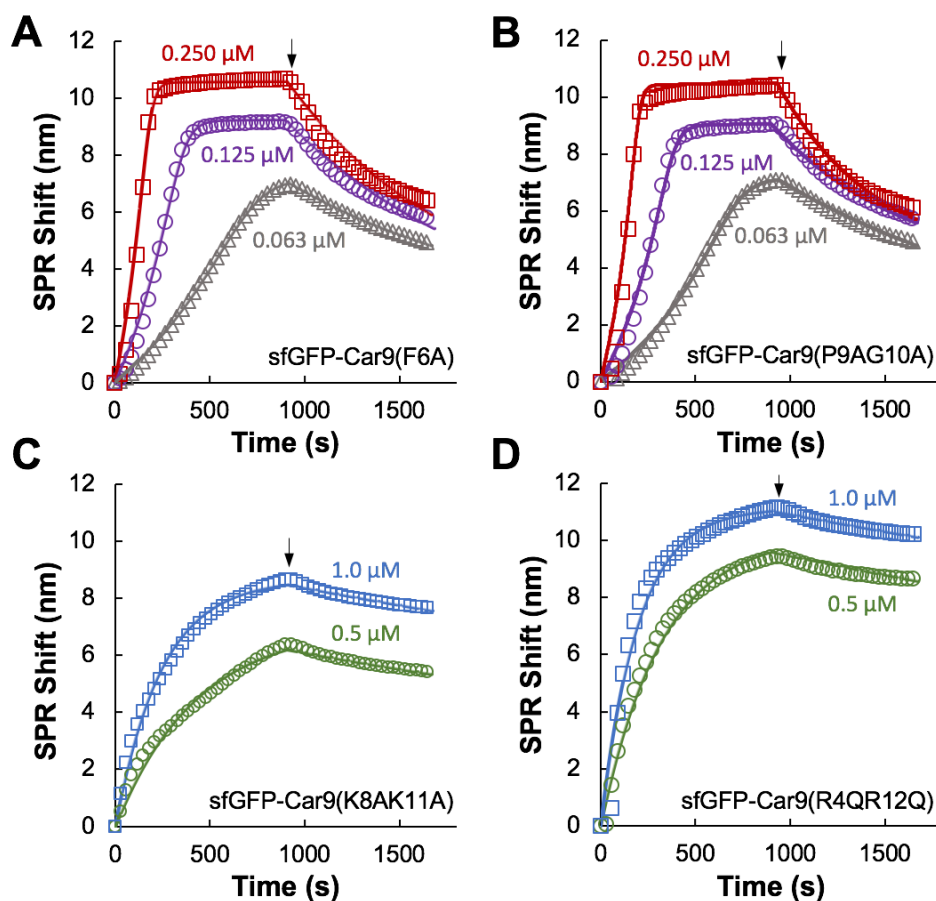


Figure 2. Adsorption kinetics of sfGFP-Car9 variants on silica-coated SPR chips. Sensorgrams were collected at the indicated concentrations of (A) sfGFP-Car9(F6A), (B) sfGFP-Car9(P9AG10A), (C) sfGFP-Car9(K8AK11A), or (D), sfGFP-Car9(R4QR12Q). Solid lines correspond to data fits obtained with a two-step cooperative adsorption model (A, B) or rate law-based Langmuir model (C, D). Arrows indicate the start of the wash cycle in which pure buffer is flowed over the chip.

of residues was motivated by the fact that the γ - and δ -carbons of the R4 and R12 side chains play a role in burying the hydrophobic F6 residue in the Rosetta-offered structure. By converting arginines to glutamines, we preserved these interactions while replacing the positively charged headgroups with neutral ones. All proteins could be purified to near homogeneity by silica affinity chromatography^{26, 28} (**Figure S3** in Supporting Information), indicating that none of the mutations abolish silica binding.

We next quantified the kinetics of adsorption by performing SPR measurements on silica-coated chips. These chips were manufactured by depositing a thin SiO₂ film onto a gold-coated glass substrate using plasma-enhanced chemical vapor deposition (PECVD).³¹ The presence of a continuous SiO₂ film was verified in X-ray photoelectron spectroscopy (XPS) analysis (**Figure S4** in Supporting Information). All chips were subjected to UV-ozone cleaning before use to remove adsorbed contaminants and reveal a silanol-rich surface that is likely composed of a combination of Q² [=Si(OH)₂] and Q³ [\equiv Si(OH)] environments based on contact angle measurements³⁶ (**Figure S5** in Supporting Information). SPR experiments conducted under conditions where mass transfer is not limiting and where unmodified sfGFP has little affinity for the surface,³¹ revealed that the F6A and P9AG10A variants exhibited the sigmoidal adsorption behavior that is characteristic of wild type sfGFP-Car9 and other Car9 fusion proteins.³¹ By contrast, the K8AK11A and R4QR12Q mutants bound to the surface in a more traditional (i.e., logarithmic) fashion (**Figure 2**).

The sigmoidal adsorption kinetics of F6A and P9AG10A mutants were well captured by a previously described model³¹ in which an initial cooperative binding event modulated by nearest neighbor interactions between Car9 segments (and to a lesser extent fused proteins) is followed by a slow second step that we ascribed to a conformational rearrangement at the silica interface (**Figure 2A-B**). In this model, which is valid under conditions where protein multilayers do not

form (**Figure S6** in Supporting Information), the rate constant for the initial adsorption step can be written:

$$k_{a1} = \frac{N_A}{\sigma_{max}} \sqrt{\frac{RT}{2\pi M}} \exp\left(\frac{-E_0}{RT}\right) \exp\left(\frac{-E_i}{RT}\theta\right) = k'_{a1} \exp(u\theta) \quad (1)$$

In this expression, the modified rate constant k'_{a1} captures the entirety of protein-surface interactions (with E_0 representing the molar intrinsic activation energy for protein adsorption and σ_{max} the density of binding sites), while an exponential term depending on protein coverage (θ) and nearest neighbor interactions (u) accounts for cooperativity.

We extracted k'_{a1} and u , as well as the rate constants for the subsequent rearrangement and desorption steps by simultaneously fitting all three sensorgram curves for both the adsorption and wash phases of the SPR experiments. **Table 1** shows that the most notable difference between the F6A and P9AG10A variants and the sfGFP-Car9 wild type was a 40-50% decrease in k'_{a1} . Considering that the three proteins have comparable molar masses (M) and that T , R and N_A are constant, equation (1) indicates that the decrease in k'_{a1} is caused by an increase in the activation energy for adsorption E_0 , likely to be associated with a conformational change that affects how primary and secondary anchor amino acids are presented to the interface (see below). Interestingly, while the F6A variant experienced a slight decrease in cooperativity, P9AG10A exhibited a small increase in u . Further highlighting the subtlety of these effects, the overall free energies of binding for the F6A and P9AG10A variants were virtually identical and within 1 kJ. mol⁻¹ of the ΔG for wild type sfGFP-Car9 (**Table 1**).

For the K8AK11A and R4QR12Q variants, adsorption kinetics were well captured by a rate-law based Langmuir model³¹ (**Figure 2C-D**). The proteins bound to the interface with rate constants (k_a) about fourfold lower than the wild type and desorbed from it with 4-to-8 fold lower rate constants (k_{d1} ; **Table 1**). These results are consistent with the fact that each mutant lacks a

predicted primary (K8 or R4) and secondary (K11 or R12) anchor residue. They are also in agreement with reports implicating interactions between positively charged side chains and negatively charged siloxides ($\equiv\text{SiO}^-\cdots\text{Na}^+$) in the high-affinity binding of the R5 peptide,³⁷⁻³⁸ synthetic LK peptides,³⁹⁻⁴⁰ and certain combinatorially-selected peptides^{36, 41-42} to silica.

To demonstrate that K7 also contributes to silica binding and rule out the possibility that glutamine substitutions exert a distinct impact on binding relative to alanine mutations, we constructed a K7AR12A variant of sfGFP-Car9. Except for its lower k_a , the protein behaved essentially like the K8AK11A and R4QR12Q mutants (**Figure S7** in Supporting Information), confirming both the assignment of anchor residues and the benign effect of glutamine substitutions. Of note, all variants with Langmuir adsorption behavior exhibited higher dissociation rate constants relative to cooperative binders, suggesting a less specific adhesion modality, and their ΔG values were comparable and only about 20% lower than the wild type (**Table 1, Figure S7**).

Table 1. Kinetic and thermodynamic parameters describing the adsorption of sfGFP-Car9 and indicated variants to silica-coated SPR chips.

protein ^a	Fit	u	k'_{a1} or k_a ($\text{M}^{-1}\text{s}^{-1}$)	k_{d1} or k_d $\times 10^{-3}$ (s^{-1})	k_{a2} $\times 10^{-3}$ (s^{-1})	k_{d2} $\times 10^{-3}$ (s^{-1})	ΔG_1 ($\text{kJ}\cdot\text{mol}^{-1}$)	ΔG_2 ($\text{kJ}\cdot\text{mol}^{-1}$)	ΔG ($\text{kJ}\cdot\text{mol}^{-1}$)
sfGFP-Car9 ^b	Sigmoidal	2.9 ± 0.1	20700 ± 1100	0.76 ± 0.02	0.17 ± 0.06	0.03 ± 0.01	-48.8 ± 0.3	-4.5 ± 2.0	-53.3 ± 2.0
F6A	Sigmoidal	2.8 ± 0.1	12400 ± 500	0.79 ± 0.20	0.16 ± 0.04	0.02 ± 0.01	-47.5 ± 0.8	-4.8 ± 0.4	-52.3 ± 0.9
P9AG10A	Sigmoidal	3.2 ± 0.1	10700 ± 700	1.06 ± 0.5	0.27 ± 0.18	0.03 ± 0.01	-47.1 ± 1.6	-5.4 ± 3.7	-52.3 ± 4.0
K8AK11A	Langmuir	NA ^c	4200 ± 300	0.18 ± 0.04	NA	NA	NA	NA	-42.1 ± 0.4
R4QR12Q	Langmuir	NA	5500 ± 700	0.13 ± 0.02	NA	NA	NA	NA	-43.5 ± 0.1

^aAdsorption kinetics of the sfGFP-Car9(K8AK11A) and R4QR12Q variants are captured with a rate-law based Langmuir model that only yields two rate constants (k_a and k_d).

^bData for sfGFP-Car9 is from reference 31.

^cNot applicable.

In short, a broad range of kinetic binding behaviors resides within a small band of free energies of binding.

Conformation of Car9 variants at the silica interface. To better understand the impact of the various mutations on interfacial conformation, we once again turned to MD simulations using low energy starting configurations provided by Rosetta (**Figure S1** in Supporting Information). **Figure 3** shows that all Car9 variants were flexible and sampled a large amount of conformational space. Inspection of RMSDs between simulations and Rosetta-proposed initial structures further revealed that F6A and R4QR12Q variants tended to adopt extended or random coil conformations over the duration of the simulations, while P9AG10A adopted a mix of folded conformations (**Figure 3** and **Figure S8** in Supporting Information). On the other hand, the K8AK11A mutant assumed a more folded configuration that exhibited the least amount of deviation from the original Rosetta structure (**Figure 3**). In short, our simulations confirm that both the F6A and P9AG10A mutations disrupt the compact conformation of wild type Car9 (with F6A being more destabilizing). They further indicate that conversion of basic amino acids to uncharged residues can have either a structurally stabilizing (K8AK11A) or destabilizing (R4QR12Q) effect, with

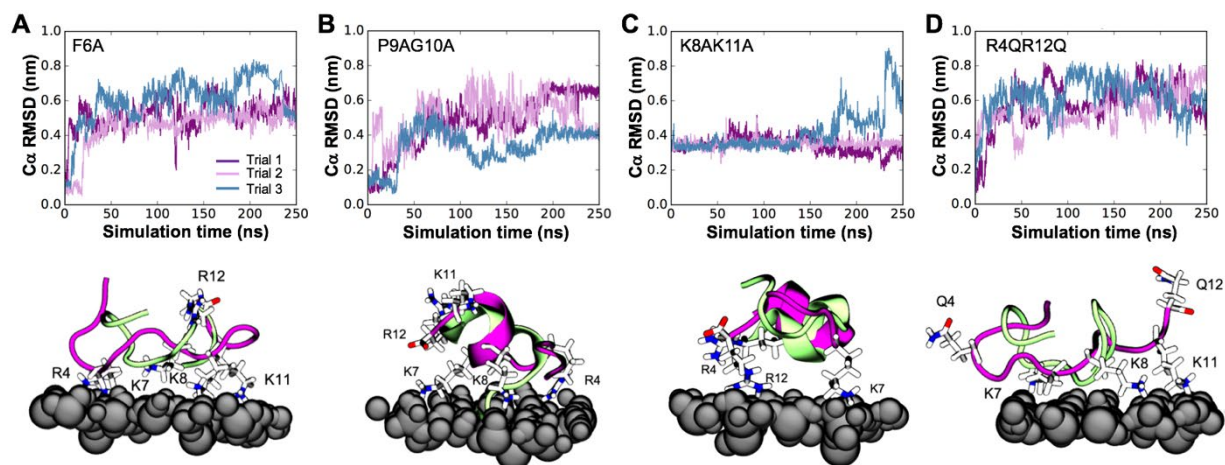


Figure 3. Root mean square deviation (RMSD) traces of C α atoms for three trials of the F6A (A), P9AG10A (B), K8AK11A (C), and R4QR12Q (D) Car9 variants adsorbed to silica. Lower panels show the superimposition of Rosetta initial structures (green) and top-ranked clusters determined from one of the trials of the MD-simulated structure at the silica interface (pink).

good self-consistency observed between Rosetta-predicted structures and the simulations of adsorbed peptides. In other words, the K8AK11A mutant retains compactness and helicity afforded by the removal of lysine while the R4QR12Q mutant, like the Car9 wild type, begins disordered and remain so during MD. The latter effect is presumably due to the sidechain of glutamine acting in a similar manner as lysine and preventing an energetically favorable compact fold within the Rosetta scoring function.

Role of electrostatic contacts. Consistent with the fact that sfGFP-Car9 has the largest k'_{off} and lowest ΔG in SPR measurements, an energetic analysis of the simulation trajectories unambiguously established that Car9 experiences the strongest interactions with the silica surface (**Figure S9** in Supporting Information). This interaction was almost entirely due to Coulombic terms, reiterating the importance of basic residues in the formation of contacts with the interface. However, there were no significant differences in the energy profiles of variants obeying cooperative (F6A and P9AG10A) or Langmuir (K8AK11A and R4QR12Q) adsorption kinetics.

To further explore the connection between electrostatics and adsorption modality, we repeated the calculations of **Figure 1D** for all four variants, averaging the distance of each residue's center of mass from the silica surface over the last 50 ns of the simulation for all three trials (**Figure 4**). This analysis revealed that: (i) cooperative binders are characterized by having substantially more stable interactions of a side chain with the surface compared to the Langmuir binders; (ii) R4 is strongly stabilizing for cooperative binders; and (iii) uncharged residues (F/A)6 show notably closer stable interactions with the surface for cooperative binders compared to their Langmuir counterparts, likely due to strong stabilization from neighboring charged residues. As with the

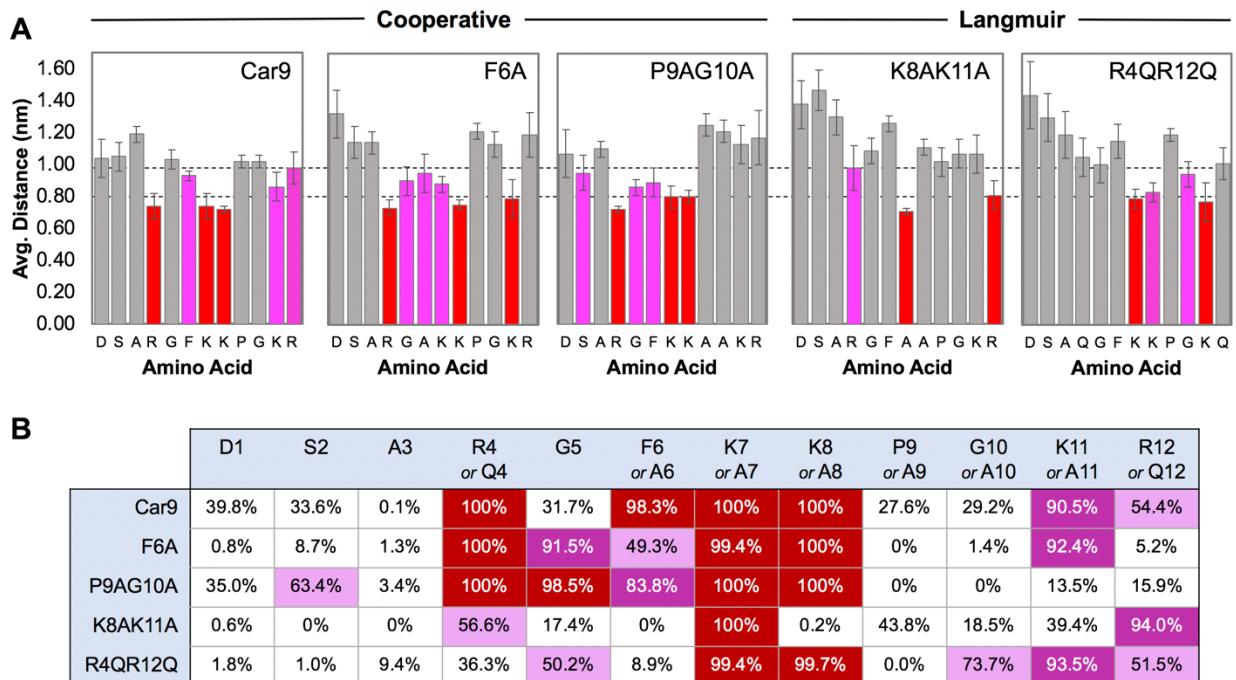


Figure 4. (A) Histograms showing the distance of the center of mass (COM) of each residue from the silica surface averaged over the last 50 ns of the three simulation trials. Primary and secondary anchor residues are shown in red and pink for both cooperative and Langmuir binders based on a distance cutoff of 0.8 nm and 1.0 nm, respectively (dashed lines). (B) Percent time spent by each residue at the silica surface during the last 50 ns of all three simulation trials. A red cell indicates that the residue was bound for over 98% of the simulation, a pink cell that it was bound for over 75% of the simulation, and a light pink cell, that it was bound for over 50% of the simulation.

Car9 analysis above, we equate stability of interaction with lower variance in the mean side chain – surface distances, indicating these average distances are found across multiple trials.

Self-association at the silica interface. To gain insights on how persistent interactions between cooperative binders and the silica interface might engender cooperative adsorption, we created an approximation of surface-bound Car9 peptides by capping the R4, K7 and K8 primary anchor residues with a harmonic restraint to trihydrogen orthosilicate (SiO_4H_3^-) groups and restraining the backbone so that the $\text{C}\alpha$ RMSD would fluctuate about the Rosetta-predicted structure. We placed three such peptides in a $\sim 6 \text{ nm}^3$ simulation box and conducted 5 independent simulations from different starting configurations. To quantify self-association events, we defined

a coordination number (CN) varying between 0 and 3 and capturing all possible oligomerization states (**Figure 5A**).

revealed a preference for dimers. The reason is that the compact helical structures of the mutant afford fewer opportunities for interpeptide contacts which are dominated by dimer-stabilizing lateral interactions of the form F6-F6 or F6-R4/R12⁴³ (**Figure 5D-E**). In short, while Car9 peptides are driven to form higher order structures when confined at high local concentrations at the silica interface, K8AK11A variants prefer at best to dimerize.

To validate these predictions and assess the role of the fused protein framework, we used atomic force microscopy (AFM) to visualize sfGFP, sfGFP-Car9 and its K8AK11A and R4QR12Q

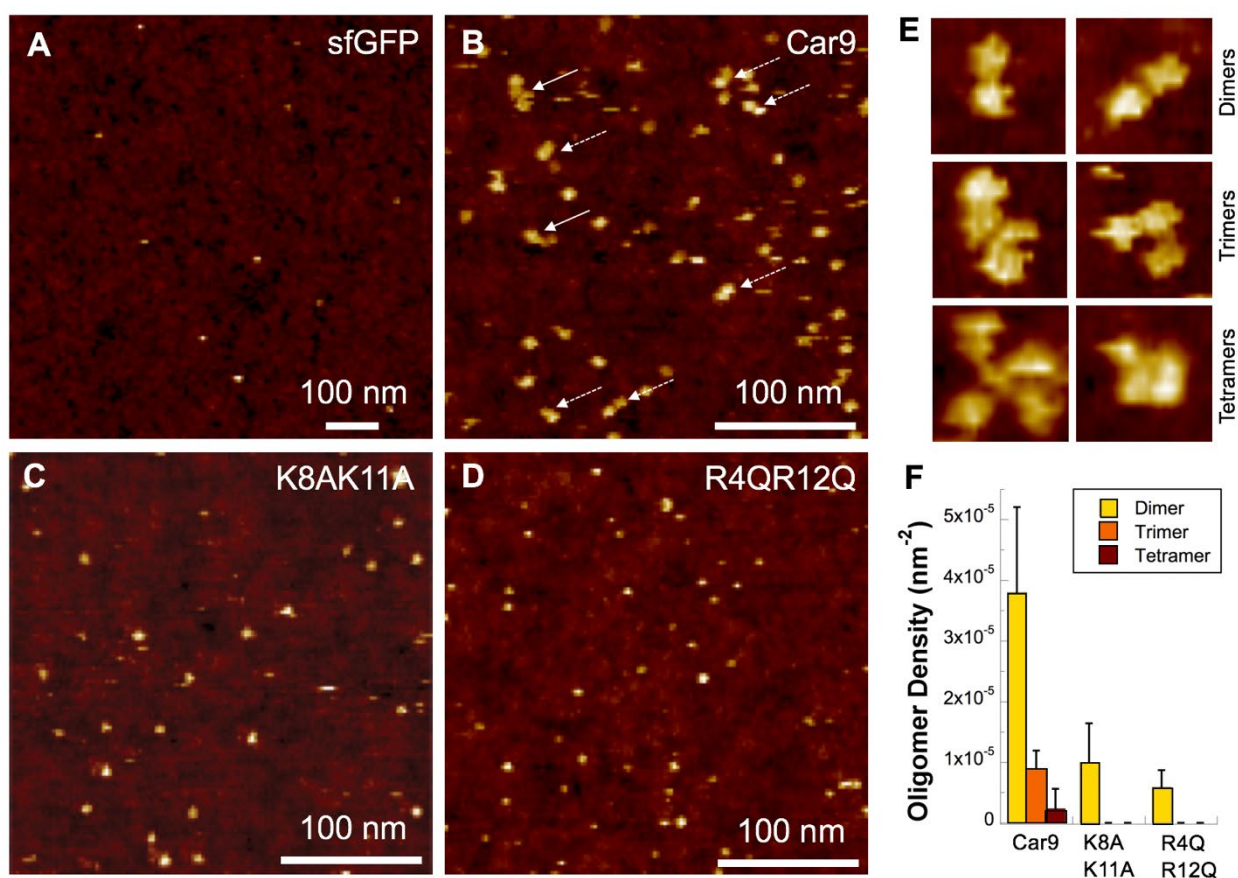


Figure 6. AFM images of wild type sfGFP (A), sfGFP-Car9 (B), and its K8AK11A (C) and R4QR12Q (D) variants adsorbed to Si wafers that had been subjected to UV-ozone treatment to produce a surface layer of silica. Solid and dashed arrows in panel B identify sfGFP-Car9 dimers and trimers, respectively. (E) Representative 30x30 nm images of sfGFP-Car9 dimers, trimers and tetramers. (F) Mean oligomer densities and standard deviations were obtained by counting at 5 imaging locations.

variants after 15 minutes of adsorption on Si wafers cleaned with UV-ozone-to produce a thin

surface layer of silica. Consistent with current and previous³¹ SPR data, unmodified sfGFP adsorbed sparsely to silica ($1.8 \times 10^{-5} \pm 5.0 \times 10^{-6}$ molecules per nm^2), and was only present as monomer the presence of a C-terminal Car9 extension increased the number of surface-bound molecules nearly 30-fold (to $5.0 \times 10^{-4} \pm 1.4 \times 10^{-4}$ molecules per nm^2), and the introduction of the K8AK11A and R4QR12Q substitutions in the Car9 sequence reduced surface coverage 2 to 4-fold relative to sfGFP-Car9 ($2.2 \times 10^{-4} \pm 7.8 \times 10^{-5}$ and $1.2 \times 10^{-4} \pm 4.3 \times 10^{-5}$ molecules per nm^2 for the K8AK11A and R4QR12Q variants, respectively) (**Figures 6A-D**). Furthermore, and in very good agreement with the simulations of **Figure 5**, we frequently observed dimers, trimers, and even tetramers of sfGFP-Car9 at the silica interface while higher order structures were rarer and limited to dimers for the K8AK11A and R4QR12Q variants (**Figure 5B-F**).

DISCUSSION

Biomining proteins such as amelogenin, silicateins and silaffins have benefited from millions of years of evolution to optimize their amino acid composition, post-translational modifications, and structure for superior function in niche environments. SBPs, on the other hand, are manually selected from combinatorial libraries that span a small fraction of the sequence space using a few rounds of biopanning against inorganic and synthetic substrates.^{2, 5-6, 8} As a result, what makes a given SBP more popular than another in subsequent characterization and technological efforts is typically based on history and some arbitrary measure of affinity or specificity rather than on a detailed understanding of the peptide's structure-binding relationship.^{37-38, 44-47}

Here, we used a deep integration of protein engineering, SPR quantification of adsorption kinetics, AFM imaging and MD simulations initiated from Rosetta predictions to gain such an understanding for the Car9 SBP. Our results indicate that Car9 has a high affinity for the silanol-

rich surface of silica, characterized by a low energy barrier for adsorption (E_0) and fast binding of the SBP and proteins fused to its C-terminus to the interface (**Table 1**). Although MD simulations initiated from the bound state show substantial structural variance, all Car9 conformers remain reasonably compact and show binding dominated by electrostatic contributions and a spectrum of several persistent, stably bound side chain residues. We propose that the high population of Car9 on the surface alongside these stable interactions lead to a high probability of Car9-Car9 cooperative interactions, between vicinal SBPs, as captured by the parameter u . The clusters are now in an energetic deep trap but conformational rearrangements still occur on a much slower time scale, contributing about $2kT$ of additional binding energy (**Table 1**). This prediction is supported by independent MD simulations and AFM experiments. First, Car9 peptides whose primary anchor residues have been capped with silicate groups readily form a variety of dimers and trimers in solution MD (**Figure 5**). Second, our AFM experiments show that the oligomerization behavior is not restricted to isolated Car9 peptides since sfGFP-Car9 fusion proteins also form higher order structures on the surface of silica (**Figure 6**).

Variants of sfGFP-Car9 that exhibit a binding modality similar to the wild type (F6A and P9AG10A) maintain similar features in the peptide-silica binding interface as determined by MD simulations. These mutations do increase the energy barrier for adsorption, and therefore lower the initial adsorption rate (**Table 1**). Yet, surface occupancy remains high and the mechanism of cooperativity proposed above is likely retained. By contrast, elimination of two of the positively charged anchor residues leads not only to a drastic reduction in adsorption rates and surface coverage, but also profoundly affects the nature of binding (**Figures 2 and 4**). The binding mechanisms that closely follow the Langmuir model are marked with a decreased likelihood of presenting the “stable” binding side chains (i.e., low variance in the binding distance) while also

generally placing the side chains further from the silica surface. We probed the relationship between SBP-surface binding behavior and peptide-peptide interactions using the K8AK11A mutant. Unlike Car9, this SBP failed to oligomerize beyond stable dimers in simulations (**Figure 5**), an observation that was confirmed by AFM imaging of the sfGFP-Car9(K8AK11A) variant. Thus, a lower surface coverage and an inability to form higher order structures beyond dimers lead to a Langmuir type of binding.

CONCLUSION

Beyond providing a consistent molecular-scale interpretation for experimental results, this study suggests a potential strategy for accelerating the discovery and optimization of SBPs for materials science applications. We found that the initial structure prediction of Car9 by Rosetta in which a cluster of basic binding residues project from a central core persisted in MD simulations conducted in the presence of an explicit silica interface. We also found that the substantial structural variance predicted by Rosetta upon introduction of one or two substitutions in the Car9 sequence was observed in MD simulations. Finally, we showed that the knowledge of anchor residues could be used to construct simple model systems for using MD that relate SBP-SBP interactions to predictions of the emerging adsorption mechanisms of an entire fusion protein. Thus, the *de novo* design of an SBP-surface binding system could be first approached by rapidly screening candidate sequences with Rosetta to find similar structural motifs (e.g., cluster of sidechains with chemical compatibility to surfaces of interest), and by next subjecting these systems to MD simulations to provide guidance for mutagenesis, and to suggest a panel of substitutions across a family of binders that could be experimentally tested within the context of fusion proteins. In the future, the concept could be extended to materials with complex surface

chemistries, and include simulation components to probe and characterize the role of proteins fused to the SBPs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publication website at DOI:

Energy landscape sampled during Rosetta *ab initio* structure prediction of Car9 and its mutants, Distance of each residue center of mass to the silica surface for the Car9 peptide, SDS-PAGE fractionation of sfGFP-Car9 and its variants following silica affinity chromatography, XPS analysis of SPR chips, Contact angle measurements, SPR experiments at high protein concentrations, SPR characterization of the sfGFP-Car9(K7AR12A) mutant, Superimposition of Rosetta structures and top-ranked clusters, Decomposition of protein-surface potential energy, Additional snapshots of Car9 trimers, Justification of CN parameters, Materials and Methods.

AUTHOR INFORMATION

Corresponding Authors

[*jpfaendt@uw.edu](mailto:jpfaendt@uw.edu)

[*baneyx@uw.edu](mailto:baneyx@uw.edu)

ORCID

François Baneyx: 0000-0001-5596-7903

Notes

F.B. declares competing financial interest in Proteios Technology, Inc. which researches and commercializes Car9-based technologies.

ACKNOWLEDGMENTS

This material is based upon work supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, as part of the Energy Frontier Research Centers program: CSSAS, The Center for the Science of Synthesis Across Scales under Award Number DE-SC0019288. SPR chips were fabricated at the Washington Nanofabrication Facility, a member of the NSF National Nanotechnology Coordinated Infrastructure (NNCI). AFM experiment were conducted at Pacific Northwest National Laboratory (PNNL), a multi-program national laboratory operated for the US Department of Energy by Battelle under contract number DE-AC05-76RL01830. We gratefully acknowledge the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system and funded by the STF at the University of Washington

REFERENCES

- (1) Kroger, N.; Deutzmann, R.; Sumper, M. Polycationic peptides from diatom biosilica that direct silica nanosphere formation. *Science* **1999**, *286*, 1129-1132.
- (2) Baneyx, F.; Schwartz, D. T. Selection and analysis of solid-binding peptides. *Curr. Opin. Biotechnol.* **2007**, *18*, 312-317.
- (3) Brown, S. Metal-recognition by repeating polypeptides. *Nat. Biotechnol.* **1997**, *15*, 269-272.
- (4) Lee, S. W.; Mao, C.; Flynn, C. E.; Belcher, A. M. Ordering of quantum dots using genetically engineered viruses. *Science* **2002**, *296*, 892-895.
- (5) Dickerson, M. B.; Sandhage, K. H.; Naik, R. R. Protein- and peptide directed synthesis of inorganic materials. *Chem. Rev.* **2008**, *108*, 4935-4978.
- (6) Sarikaya, M.; Tamerler, C.; Jen, A. K.; Schulten, K.; Baneyx, F. Molecular biomimetics: Nanotechnology through biology. *Nat Mater* **2003**, *2*, 577-585.
- (7) Seeman, N. C.; Belcher, A. M. Emulating biology: Building nanostructures from the bottom up. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 6451-6455.
- (8) Walsh, T. R.; Knecht, M. R. Biointerface structural effects on the properties and applications of bioinspired peptide-based nanomaterials. *Chem. Rev.* **2017**, *117*, 12641-12704.

- (9) Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly. *Nature* **2000**, *405*, 665-668.
- (10) Naik, R. R.; Stringer, S. J.; Agarwal, G.; Jones, S. E.; Stone, M. O. Biomimetic synthesis and patterning of silver nanoparticles. *Nat Mater* **2002**, *1*, 169-172.
- (11) Nam, K. T.; Kim, D. W.; Yoo, P. J.; Chiang, C. Y.; Meethong, N.; Hammond, P. T.; Chiang, Y.-M.; Belcher, A. M. Virus-enabled synthesis and assembly of nanowires for lithium ion battery electrodes. *Science* **2006**, *312*, 885-888.
- (12) Chen, J.; Zhu, E.; Liu, J.; Zhang, S.; Lin, Z.; Duan, X.; Heinz, H.; Huang, Y.; De Yoreo, J. J. Building two-dimensional materials one row at a time: Avoiding the nucleation barrier. *Science* **2018**, *362*, 1135-1139.
- (13) Chiu, C. Y.; Li, Y.; Ruan, L.; Ye, X.; Murray, C. B.; Huang, Y. Platinum nanocrystals selectively shaped using facet-specific peptide sequences. *Nat. Chem.* **2011**, *3*, 393-399.
- (14) Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N. Introducing protein intrinsic disorder. *Chem. Rev.* **2014**, *114*, 6561-6588.
- (15) Tang, Z.; Palafox-Hernandez, J. P.; Law, W. C.; Hughes, Z. E.; Swihart, M. T.; Prasad, P. N.; Knecht, M. R.; Walsh, T. R. Biomolecular recognition principles for bionanocombinatorics: An integrated approach to elucidate enthalpic and entropic factors. *ACS Nano* **2013**, *7*, 9632-9646.
- (16) Sano, K.; Shiba, K. A hexapeptide motif that electrostatically binds to the surface of titanium. *J. Am. Chem. Soc.* **2003**, *125*, 14234-14235.
- (17) Mirau, P. A.; Naik, R. R.; Gehring, P. Structure of peptides on metal oxide surfaces probed by nmr. *J. Am. Chem. Soc.* **2011**, *133*, 18243-18248.
- (18) Suzuki, Y.; Shindo, H.; Asakura, T. Structure and dynamic properties of a ti-binding peptide bound to tio₂ nanoparticles as accessed by h-1 nmr spectroscopy. *J. Phys. Chem. B* **2016**, *120*, 4600-4607.
- (19) Goobes, G.; Goobes, R.; Schueler-Furman, O.; Baker, D.; Stayton, P. S.; Drobny, G. P. Folding of the c-terminal bacterial binding domain in statherin upon adsorption onto hydroxyapatite crystals. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 16083-16088.
- (20) Masica, D. L.; Ash, J. T.; Ndao, M.; Drobny, G. P.; Gray, J. J. Toward a structure determination method for biomineral-associated protein using combined solid- state nmr and computational structure prediction. *Structure* **2010**, *18*, 1678-1687.
- (21) Arachchige, R. J.; Burton, S. D.; Lu, J. X.; Ginovska, B.; Harding, L. K.; Taylor, M. E.; Tao, J.; Dohnalkova, A.; Tarasevich, B. J.; Buchko, G. W.; Shaw, W. J. Solid-state nmr identification of intermolecular interactions in amelogenin bound to hydroxyapatite. *Biophys. J.* **2018**, *115*, 1666-1672.
- (22) Shaw, W. J. Solid-state nmr studies of proteins immobilized on inorganic surfaces. *Solid State Nucl. Magn. Reson.* **2015**, *70*, 1-14.
- (23) Sprenger, K. G.; Prakash, A.; Drobny, G.; Pfaendtner, J. Investigating the role of phosphorylation in the binding of silaffin peptide r5 to silica with molecular dynamics simulations. *Langmuir* **2018**, *34*, 1199-1207.
- (24) Coyle, B. L.; Rolandi, M.; Baneyx, F. Carbon-binding designer proteins that discriminate between sp²- and sp³-hybridized carbon surfaces. *Langmuir* **2013**, *29*, 4839-4846.

- (25) Dunakey, S. J. G.; Coyle, B. L.; Thomas, A.; Xu, M.; Swift, B. J. F.; Baneyx, F. Selective labeling and decoration of the ends and sidewalls of single-walled carbon nanotubes using mono- and bispecific solid-binding fluorescent proteins. *Bioconjug. Chem.* **2019**, *30*, 959-965.
- (26) Coyle, B. L.; Baneyx, F. A cleavable silica-binding affinity tag for rapid and inexpensive protein purification. *Biotechnol. Bioeng.* **2014**, *111*, 2019-2026.
- (27) Coyle, B. L.; Baneyx, F. Direct and reversible immobilization and microcontact printing of functional proteins on glass using a genetically appended silica-binding tag. *Chem. Commun.* **2016**, *52*, 7001-7004.
- (28) Soto-Rodriguez, J.; Coyle, B. L.; Samuelson, A.; Aravagiri, K.; Baneyx, F. Affinity purification of car9-tagged proteins on silica matrices: Optimization of a rapid and inexpensive protein purification technology. *Protein Expr. Purif.* **2017**, *135*, 70-77.
- (29) Yang, W.; Hellner, B.; Baneyx, F. Self-immobilization of car9 fusion proteins within high surface area silica sol-gels and dynamic control of protein release. *Bioconjug. Chem.* **2016**, *27*, 2450-2459.
- (30) Swift, B. J.; Shadish, J. A.; DeForest, C. A.; Baneyx, F. Streamlined synthesis and assembly of a hybrid sensing architecture with solid binding proteins and click chemistry. *J. Am. Chem. Soc.* **2017**, *139*, 3958-3961.
- (31) Hellner, B.; Lee, S. B.; Subramaniam, A.; Subramanian, V. R.; Baneyx, F. Modeling the cooperative adsorption of solid-binding proteins on silica: Molecular insights from surface plasmon resonance measurements. *Langmuir* **2019**, *35*, 5013-5020.
- (32) Pedelacq, J. D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **2006**, *24*, 79-88.
- (33) Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V.; Kaas, Q.; Eletsky, A.; Huang, P. S.; Johnsen, W. A.; Greisen, P. J.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsiaris, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D. Accurate de novo design of hyperstable constrained peptides. *Nature* **2016**, *538*, 329-335.
- (34) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. Charmm: The biomolecular simulation program. *Journal of computational chemistry* **2009**, *30*, 1545-1614.
- (35) Heinz, H.; Lin, T.-J.; Kishore Mishra, R.; Emami, F. S. Thermodynamically consistent force fields for the assembly of inorganic, organic, and biological nanostructures: The interface force field. *Langmuir* **2013**, *29*, 1754-1765.
- (36) Emami, F. S.; Puddu, V.; Berry, R. J.; Varshney, V.; Patwardhan, S. V.; Perry, C. C.; Heinz, H. Prediction of specific biomolecule adsorption on silica surfaces as a function of pH and particle size. *Chem. Mater.* **2014**, *26*, 5725-5734.
- (37) Knecht, M. R.; Wright, D. W. Functional analysis of the biomimetic silica precipitating activity of the r5 peptide from *cylindrotheca fusiformis*. *Chem. Commun. (Camb.)* **2003**, 3038-3039.
- (38) Lechner, C. C.; Becker, C. F. A sequence-function analysis of the silica precipitating silaffin r5 peptide. *J. Pept. Sci.* **2014**, *20*, 152-158.

- (39) Baio, J. E.; Zane, A.; Jaeger, V.; Roehrich, A. M.; Lutz, H.; Pfaendtner, J.; Drobny, G. P.; Weidner, T. Diatom mimics: Directing the formation of biosilica nanoparticles by controlled folding of lysine-leucine peptides. *J. Am. Chem. Soc.* **2014**, *136*, 15134-15137.
- (40) Lutz, H.; Jaeger, V.; Berger, R.; Bonn, M.; Pfaendtner, J.; Weidner, T. Biomimetic growth of ultrathin silica sheets using artificial amphiphilic peptides. *Advanced Materials Interfaces* **2015**, *2*.
- (41) Puddu, V.; Perry, C. C. Peptide adsorption on silica nanoparticles: Evidence of hydrophobic interactions. *ACS Nano* **2012**, *6*, 6356-6363.
- (42) Patwardhan, S. V.; Emami, F. S.; Berry, R. J.; Jones, S. E.; Naik, R. R.; Deschaume, O.; Heinz, H.; Perry, C. C. Chemistry of aqueous silica nanoparticle surfaces and the mechanism of selective peptide adsorption. *J. Am. Chem. Soc.* **2012**, *134*, 6244-6256.
- (43) Flocco, M. M.; Mowbray, S. L. Planar stacking interactions of arginine and aromatic side-chains in proteins. *J. Mol. Biol.* **1994**, *235*, 709-717.
- (44) Geiger, Y.; Gottlieb, H. E.; Akbey, U.; Oschkinat, H.; Goobes, G. Studying the conformation of a silaffin-derived pentalyanine peptide embedded in bioinspired silica using solution and dynamic nuclear polarization magic-angle spinning nmr. *J. Am. Chem. Soc.* **2016**, *138*, 5561-5567.
- (45) Zane, A. C.; Michelet, C.; Roehrich, A.; Emani, P. S.; Drobny, G. P. Silica morphogenesis by lysine-leucine peptides with hydrophobic periodicity. *Langmuir* **2014**, *30*, 7152-7161.
- (46) Chen, H.; Su, X.; Neoh, K.-G.; Choe, W.-S. Qcm-d analysis of binding mechanism of phage particles displaying a constrained heptapeptide with specific affinity to sio₂ and tio₂. *Anal. Chem.* **2006**, *78*, 4872-4879.
- (47) Puddu, V.; Perry, C. C. Interactions at the silica-peptide interface: The influence of particle size and surface functionality. *Langmuir* **2014**, *30*, 227-233.

TOC GRAPHIC

