Correlating Solid-Binding Peptide Structure with Biomimetic Function

Brittney Hellner

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

François Baneyx, Chair

Guozhong Cao

Cole DeForest

Elizabeth Nance

Program Authorized to Offer Degree:

Department of Chemical Engineering

University of Washington

Abstract

Correlating Solid-Binding Peptide Structure with Biomimetic Function

Brittney Hellner

Chair of the Supervisory Committee:
François Baneyx
Department of Chemical Engineering

While much progress has been made in elucidating the mode of action of biomineralizing proteins and in emulating their function using combinatorially selected solid-binding peptides (SBPs), we remain far from duplicating the sophisticated architectures produced by biological systems. This is in part due to a poor understanding of the relationship between SBP conformation and inorganic adhesion and precipitation. This dissertation focuses on correlating the structure of Car9, a SiO_2/TiO_2-binding dodecapeptide originally isolated for its ability to recognize the edges of graphitic nanostructures, with silica binding affinity and titania precipitation activity. We start by quantifying the binding kinetics of a fusion protein between superfolder green fluorescent (sfGFP) and the Car9 SBP to silica using surface plasmon resonance (SPR) measurements, develop a two-step kinetic model that accurately captures the unique cooperative adhesion behavior of the fusion protein, and demonstrate that elimination of multiple basic residues converts the binding to a Langmuir process. Using a panel of mutants in the Car9 segment, we combine SPR characterization and molecular dynamics simulations initiated from Rosetta structural predictions to gain insights on the interplay of amino acid

composition, structure, and adhesion modality. We show that the high affinity binding of Car9 to silica stems from persistent electrostatic interaction with the surface that, along with high surface coverage, promotes cooperative interactions between neighboring SBPs. Additionally, we ascribe the transition to Langmuir adhesion to the loss of two or more such contacts due to mutagenesis. Finally, we show mutations in the Car9 extension biases the crystallinity of titania bioprecipitated by sfGFP-Car9 fusion proteins between anatase and TiO_2 (*B*) phases and harness our knowledge of SiO_2 adhesion to interpret these results. Collectively, our results provide a roadmap for developing a fundamental understanding of SBPs' inorganic adhesion and offers insight in how to intuitively design peptide-protein scaffolds to produce unique composite materials for applications ranging from biomedical engineering to catalysis.

Table of Contents

List of Figures

List of Tables

Acknowledgements

There are many people I would like to thank who have supported me over the last four years: My advisor, François Baneyx, for his guidance and mentorship during my PhD; My committee for providing their feedback and insight on progressing my project forward, and my lab-mates for their camaraderie and our mid-day coffee runs together.

I also want to thank my family for their continued love and support. Most of all, I would like to thank my husband Jon for his support from the beginning, starting from the application process, through every exam, and hurdle. You have always been my sounding board, rock, and together with Hank Floyd you two always manage to make me smile.

Chapter 1 Introduction

Nature routinely interfaces organic and inorganic materials to produce unique composites that fulfill a structural or protective role. Examples include bones and teeth made of hydroxyapatite calcium phosphate[1], mollusk shells composed of calcium carbonate[2], and diatoms skeletons with silica walls.[3] The mineralizing organisms that fabricate these materials do so with high specificity and nanoscale precision using proteins and other biomolecules[4] that have the ability to bind to inorganic materials and to act as scaffolds and building blocks to give highly ordered structures.[5] Understanding how proteins interact with inorganic materials and controlling protein-aided materials synthesis would open the door to a broad range of applications in electrochemistry[6], biosensing[7], drug delivery[7], and protein-templated device fabrication.[8]

This work focuses on: (1) gaining a predictive understanding of how solid-binding peptides (SBPs) fused to larger protein scaffolds adsorb to silica surfaces; (2) and on improving our understanding of the biomineralization process with a focus on titania precipitation. We ask these questions within the context of Car9, a solid binding peptide or SBP initially selected for its ability to bind to the negatively-charged edges of carbon nanostructures[9] and later found to interact with silica.[10] More specifically, we study how various mutations in the Car9 sequence and their insertion location within the superfolder green fluorescent protein[11] (sfGFP) scaffold, influences SBP structure and chemistry, and thus interactions with silica surfaces and titania precursors. Our long-term goal is to harness this understanding to predictively manipulate the interaction of Car9-related SBPs with silica interfaces, and to control the production of titania species with user-specified morphologies, phases and compositions.

1.1 Solid binding peptides (SBP)

A solid binding peptide (SBP) is a short stretch of amino acids with an affinity for inorganic materials.[12] SBPs have been used to precipitate, reduce, decorate, or assemble various materials[13]

and their genetic or chemical coupling to a protein can be used for controlled immobilization at inorganic interfaces.[14,15] SBPs are generally selected by cell or phage surface display in a process known as biopanning. For instance, in the FLiTrx bacterial flagellar display technology,[13,15] about 10^7 random dodecapeptides inserted within the active site loop of Thioredoxin 1 (TrxA), which is itself inserted within the flagellar FliC protein are expressed in *Escherichia coli* (*E. coli*). Each cell manufactures a defective flagella that contains a particular peptide, and the resulting library is contacted with an inorganic substrate of interest. The substrate is washed under conditions of various stringency to remove weak binders, strong binders are recovered by physically shearing off their flagella, and the biopanning process is repeated three to five times to enrich for high affinity binders. The plasmid DNA encoding the winning sequences reveal the amino acid composition of inorganic binders which are referred to as SBPs.[13,15]

1.2 The Car9 solid binding peptide

Car9 was originally isolated via FliTrX cell surface display by the Baneyx lab as carbon-binding, disulfide-constrained dodecapeptide of sequence *CGP*DSARGFKKPGKR*GPC* using the one letter amino acid code (the flanking CGP and GPC tripeptide are invariants in FliTrX display).[9] Car9 binds to the hydroxyl-, carbonyl- and carboxylate-rich edges of graphitic nanostructures, and exhibits a preference for sp3-hybrydized carbon.[9] Since its initial identification, Car9 has also been found to bind to silica,[10] enabling affinity purification,[10] immobilization in sol-gels,[16] and microcontact printing[17] of proteins to which it is fused. While these interactions likely implicate electrostatic contacts between the basic lysine (K) and arginine (R) residues of the Car9 sequence and the silanol-rich and negatively-charged surface of silica,[10] the precise binding mechanisms and peptide conformation at the silica interface remains unknown.

1.3 Silica-binding proteins and biosilicification

1.3.1 Formation of diatom frustules

For decades, there has been considerable interest in understanding how biomineralizing proteins control the formation of complex architectures such as the silica skeleton of single-celled algae called diatoms.[18,19] There are thousands of species of diatoms in nature, each with a unique cell wall called a frustule (**Figure 1.1**).[18] These porous structures which consist of silicic acid coated in a layer of organic material[20] are formed in intracellular compartments called silica deposition vesicles (SDVs).[21] SDVs contain a high concentration of silicic acid, $Si(OH)_4$, and are held at a more acidic pH than the surrounding cellular environment.[20,22] This promotes silicic acid hydrolysis and polycondensation to form silica.[18]

1.3.2 Silaffins

Silaffins are polypeptides isolated from the diatom *Cylindrotheca fusiformis*.[23] They have a high affinity for silica and template the formation of unique silica structures when mixed with a pre-hydrolyzed solution of silicic acid *in vitro.*[24] With a growing interest in biomimetic silica structures for enzyme immobilization, controlled release, biocatalysis, microfluidic reactors, biosensors, antifouling materials,[24] and even photonic devices,[25] there has been a drive to understand how silaffins control the deposition and templating of silica into unique architectures.

Three silaffins were isolated from *Cylindrotheca fusiformis;* silaffin-1A, silaffin-1B, and silaffin-2 with the corresponding masses 4, 8, and 17 kDa.[20] The first gene cloned was *sil1* from the *Cylindrotheca fusiformis* cDNA library, this gene encodes the precursor protein Sil1p (**Figure 1.2**) composed of 265 amino acids.[24] It has an acidic N-terminal domain that is removed during silica maturation and a basic C-terminus that is proteolytically cleaved into 7 peptide units (R1-R7) with repetitive amino acid sequences (**Figure 1.2**).[18,24] Silaffin-1A and silaffin-1B are a result of this proteolytic processing with silaffin-1A being divided into silaffin-$1A_1$ (R3-R7) and silaffin-$1A_2$ (R2) and silaffin-1B (R1).[26,27] A unique feature discovered about silaffins is the post-translational modification of certain amino acids. In silaffin-1A, silaffin-1B and silaffin-2,

hydroxy-amino acids are phosphorylated, the ϵ-amino groups of lysine are alkylated and some become hydroxylated or phosphorylated at the δ-C-atom.[26] Silaffin-2 also has hydroxyl amino acids that are sulfated or glycosylated.[26-28] The serine and lysine modifications increase the tendency of silaffins to form large aggregates to guide the nanopatterning of silica formation at ambient temperature and pH.[28] The lysine modifications are also thought to be one of the factors for the silica formation under acidic pH conditions.[23,25]

Silaffins also contain a unique KXXK motif (where X is any amino acid) which is repeated multiple times. Weineke and colleagues[29] have shown that this motif affects silica precipitation using three synthetic GKAXKG hexapeptides bearing 3 specific mutations in X (serine, glutamic acid, and alanine).[25] These mutations changed the overall charge, polarity, and hydrophobicity of the peptides. When used for silica formation the serine mutation promoted the production of more silica material, while the glutamic acid mutation produced the least. SEM images for each precipitate showed the varying morphologies for the three, the serine mutation consisted of spheres clustered together with a grainy precipitate, the glutamic acid mutation only had grainy silica, and the alanine mutation was mainly spheres clustered together.[29] These results demonstrate that the overall charge and the KXXK motif in silaffins affect the quantity and morphology of silica formed.[29]

1.3.3 R5 peptide

The R5 peptide is a 19 amino acids-long peptide of sequence SSKKSGSYSGSKGSKRRIL, which is repeated 19 times in Sil1p,[23] a Silaffin from *Cylindrotheca fusiformis.*[30] Post translationally modified R5 peptide in silaffin-1A has maximum silicification activity around pH 5.[23] Silica precipitation activity of the R5 peptide is said to be dependent on the peptides ability to aggregate and self-assemble forming a template for silica formation.[27] Synthetic R5 peptide's impact on biomimetic silica formation has been extensively studied; without the post translational

modifications this synthetic peptide only promotes silicification at pH values above 7.[23,31] Senoir et al. demonstrated that the amount of silica precipitated using synthetic R5 peptide is linearly dependent on the concentration of the silicic acid precursor.[31] Sprenger et al. used molecular dynamic simulations combined with a metadynamics enhanced sampling method to simulate three different R5 peptides with varying degrees of serine phosphorylation (control R5 peptide, S14 phosphorylated, all serines phosphorylated) at two pHs (5 and 7.5).[32] The results showed that the binding free energy of R5 to a silica surface decreases with decreasing pH and increasing number of serine phosphorylations.[32] Concluding that the control of the peptide surface interactions are dominated by the chemistry of the silica surface at pH 5 and the degree of phosphorylation dominates the R5 peptide interactions with the silica surface at pH 7.5.[32]

The RRIL motif has been proposed by Knecht et al. to drive peptide self-assembly, and produce a structure that influences the morphology of the precipitated silica.[33] This was done using truncated and mutated versions of the R5 peptide to examine the role of the RRIL motif. Truncated peptides without the RRIL motif and peptides with the RR sequence mutated showed significantly less silica precipitation activity, proving that it is more than just lysines and serines involved in silica precipitation.[33] However, a subsequent study by Lechner et al. argued that the RRIL motif has a minor impact on silica precipitation activity and more of an impact on the precipitate morphology.[34] This was demonstrated by using R5 peptide variants to precipitate silica, concluding that precipitation activity was more dependent on amino groups of the lysine residues due to activity loss when they were mutated. When the RRIL motif was moved to different locations in the peptides or broken up in the sequence, but still present, the peptide had similar precipitation activity, but the precipitate morphology changed.[34] Lechner et al. concluded that the RRIL motif in any form has to be present for silica precipitation, but the location of the motif contributes more to peptide assembly, therefore affecting the morphology.[34]

Beyond R5 composition and structure, processing conditions can greatly influence biosilica morphology. For instance, while spherical silica structures are normally observed with an excess of R5, bubbling nitrogen in the system changed the morphology to arches.[30] R5 has also been covalently linked to a cargo peptide via a disulfide linkage to encapsulate the R5/cargo peptide conjugate within silica during precipitation.[35] This cargo molecule was partially released from the silica precipitate after being exposed to TCEP, a phosphine that reduces disulfide linkages, demonstrating the possibility of controlled release of peptides from silica precipitated with the R5 peptide.[35] Luckarift et al. also demonstrated that using the R5 peptide to precipitate silica in the presence of the enzyme butyrylcholinesterase, allowed for the enzyme to be entrapped in the silica precipitate, while still retaining its' activity.[36]

Sewell and Wright, further demonstrated that R5 can precipitate TiO_2 under ambient conditions from a titanium lactate precursor, TiBALDH.[37] The precipitation of the TiO_2 material was dependent on R5 peptide concentration, as peptide concentration increased the amount precipitated increased. The composition of this TiO_2 material had anatase and rutile characteristics when heated above 600°C and nanoparticles in the size range of 13 to 50 nm.[37] Using the TiBALDH precursor and a phosphate buffer Cole et al. demonstrated that the R5 peptide can precipitate titanium phosphate with spherical and fused structures in the nanometer to micrometer size range.[38]

1.3.4 Silicateins

Diatoms are not the only organism with a biosilica skeletal structure, marine sponges (phylum Porifera) also have a skeletal structure, called the spicules, that is composed of biosilica formed using enzymatic silicateins.[39,40] Silicateins are the only major protein species found in the spicules of marine sponges.[41] Using the demosponge *Tethya aurantium* three silicatein subunits (silicatein-α, silicatein-β, silicatein-γ) ranging in size from 27 to 29 kDa were isolated.[41]

Silicateins are related to cathepsin, which is a hydrolytic enzyme, in the subfamily L of cysteine proteases, in marine sponges, specifically identified in *Geodia cydonium*.[42] The cysteine at the active site in cathepsin proteases is a serine in silicatein-α, but the six cysteine residues that form disulfide bridges in the proteases are also present in silicatein-α.[41] Sumerel et al. proposed that this serine replacement increases the nucleophilicity of silicatein for nucleophilic attack at the silicon substrate.[40,43]

Recombinant and native silicatein-α are active catalysts for the hydrolysis of silicon alkoxides to form biosilica at neutral pH and ambient temperature.[39,41] Also, similar to silaffins, silicateins undergo post translational modifications such as methylations of the N-terminal region and phosphorylations, which are thought to assist in the templating and biosilicification process.[44] Silicateins, like silaffins, have also been used to catalyze the hydrolysis and condensation of non-natural occurring metal oxides, such as gallium oxohydroxide (GaOOH), under mild conditions.[45]

1.3.5 The silica surface

Silica, *a.k.a.* silicon dioxide (SiO_2), is an earth-abundant material with numerous crystalline and amorphous allotropic forms,[46] found in the skeletons of diatoms, sand, and quartz.[47] Silica is tetrahedrally coordinated with four oxygen atoms around one silicon atom.[48] Silica surfaces contain variable amounts of siloxanes (O-Si-O) and silanol groups (Si-OH). The three types of silanol groups (**Figure 1.3**) on the surface are isolated silanol (≡Si-OH), geminal silanol ($=Si(OH)_2$), and vicinal silanol, two silanol groups participating in hydrogen bonding with each other.[49,50]

Proteins interact with silica surfaces through electrostatic, H-bonding and hydrophobic interactions with silanol groups, and through hydrophobic interactions with siloxane groups.[51,52] The process is not fully understood (see below) and further complicated by the fact that silica surfaces exhibit a varying distribution of silanol and siloxane groups, depending on surface

preparation and ionization state. (The latter parameter can be controlled experimentally by changing the solution pH.[53])

1.3.6 Peptide adsorption onto silica

Significant work has been directed at understanding peptide adsorption to silica. As mentioned above, electrostatic interactions, especially those involving the amine-rich side chains of lysine and arginine and negatively charged silanol groups, play an important role in the process.[54] Further work has however revealed that hydrogen bonds and polar interactions between silanols and the side chains of serine, histidine, and aspartic acid are important for the adsorption of non-cationic peptides to silica.[54] In addition, there is evidence that neutral peptides interact with silica through hydrophobic interactions.[55] Thus, although electrostatics likely dominate, adsorption is fine-tuned by a number of other interactions, raising the possibility that SBPs can be designed to bind to silica in a user-specified fashion.

Emami et al. demonstrated using MD simulation that the pH of solution in which silica surfaces are immersed is also an important parameter in the binding of peptides to silica.[56] Using a positive and negatively charged peptide at basic pH 9 (90% ionization of the surface) the peptide adsorption is dominated by electrostatic interactions in the positively charged peptide and at acidic pH 3 (0% ionization) the hydrophobic residues dominated the interactions with the silica surface.[56] The pH had little effect on a neutral peptides' binding, since the interaction with the silica surface was dominated by hydrogen bonds and hydrophobic interactions.[56]

Silica particle size also influences the binding of peptides to the surface.[57,58] Puddu and colleagues studied the effects silica particle size (28-500 nm) had on peptide adsorption using three peptides with different net charges.[58] All peptides had an increased adsorption per unit surface area of silica with an increase in silica particle size. The net positive peptide had the largest adsorption at 0.8 $mMgm^{-2}$ for the 500 nm particle, while the negative peptide had about 0.4 $mMgm^{-2}$.[58] These

results led to the conclusion that the larger particles have an increase in surface ionization causing a shift in the adsorption mechanism towards more electrostatic dependence.[58]

1.4 Experimental techniques for quantifying peptide adsorption

Experimental tools commonly used to study these SBP interactions include quartz crystal microbalance (QCM), surface plasmon resonance (SPR), sum frequency generation (SFG), solid state nuclear magnetic resonance (SSNMR), and atomic force microscopy (AFM).[59-66] Among these techniques, QCM, SPR, and AFM can provide kinetic and thermodynamic information, but SPR is particularly attractive because it rapidly and accurately quantifies binding kinetics by monitoring how the refractive index changes over time when an analyte binds to a gold surface coated with a target ligand or material.[67,68] While SPR has been extensively used to characterize the binding of biomacromolecules to immobilized ligands, lipid bilayers, polymer films, and other biomaterials,[69] it has more seldom been employed to monitor the interaction of SBPs with inorganic surfaces.[9,65,70-74] In most cases, adsorption isotherms are treated with a 1:1 Langmuir model and kinetic parameters and free energies of binding are extracted from the fit. The fact that QCM and SPR measurements are often noisy when conducted at low concentrations of synthetic peptides further complicates the recognition that a SBP adsorption process might not obey Langmuir kinetics. A significant amount of this work has been done by Sarikaya and coworkers who used a biexponential Langmuir model to capture the non-ideality of gold-binding peptides to either gold or platinum using both QCM and SPR.[65,71-73]

1.5 Biomimetic TiO_2 mineralization

Previous research has demonstrated that two family of proteins normally involved in biosilicification, silaffins and silicateins, are also capable of inducing titanium dioxide (titania, TiO_2) precipitation from soluble precursors.[37,75] This compound is technologically valuable and

has been used as a pigment, photocatalyst, UV blocker, and for the fabrication of oxygen sensors, lithium ion batteries, and solar cells.[76,77]

Titanium dioxide pigments are industrially produced through a sulfate or a chloride process. The sulfate process, which is used about 40% of the time, relies on concentrated sulfuric acid for the dissolution of the raw materials, usually ilmenite or titanium slag, to form titanyl sulfate.[78] The method yields anatase TiO_2 pigments and a high concentration of waste iron sulfate.[78] The chloride process relies on flame oxidation of $TiCl_4$ at temperatures of 900-1000ºC.[78] This method enables a finer control over the final anatase pigment product and is less labor intensive then the sulfate process. However, both methods have drawbacks including poor control over morphology, the need for high temperatures, and the production of environmental waste.[78]

1.5.1 Crystal phases of TiO_2

TiO_2 has many crystallographic forms with rutile, anatase, and brookite being most commonly found in nature and rutile, anatase, and TiO_2(*B*) forms most relevant to this work. TiO_2 is composed of TiO_6 octahedra in which a central titanium atom is surrounded by six oxygen atoms (**Figure 1.4**). How these octahedrons are coordinated determines the material's phase. Anatase and rutile are both tetragonal, while brookite is orthorhombic and TiO_2(*B*) is monoclinic.[79] Rutile is the most thermodynamically stable phase at all temperatures and pressures, and it is the most symmetric and compact polymorph, with each TiO_6 octahedron surrounded by ten octahedra (two of which are edge-sharing and eight corner-sharing; **Figure 1.5A**).[80,81] Both anatase and brookite are metastable, and transform irreversibly to rutile upon heating.[79] Anatase consists of four edge shared and four corner shared octahedra (**Figure 1.5B**).[82] The lattice structure of brookite is similar to anatase, but orthorhombic, with TiO_6 octahedra sharing three edges (**Figure 1.5C**).[80] Pure brookite is rarely studied experimentally because it is metastable, has low-symmetry, and often contains a secondary anatase or rutile phase. [79,83] TiO_2(*B*), is monoclinic and the least dense of

natural titania polymorphs.[84-86] It consists of both edge- and corner-shared octahedra, with a framework described as a stack of (011)-oriented sheets of double layers of edge-sharing TiO_2 octahedra. These sheets are connected through the octahedron's corners, leading to a relatively open structure (**Figure 1.5D**).[84,85,87]

The properties of TiO_2 depend on crystal structure, morphology, surface area, and porosity of the material.[88] In bulk solutions, rutile is the most stable phase, and anatase and brookite are metastable.[89] This is not necessarily true at the nanoscale. The phase stability TiO_2 nanoparticles depends on the environment (temperature and pH), on the amount of water present, and on how the particle interacts with water.[89] For instance, Zhang and coworkers have shown that particle size plays an important role in determining the transition temperature and nucleation behavior of the anatase to rutile phase change.[90] (Larger particles require higher temperatures to undergo phase transition.[90]) Understanding these properties is important for the development of TiO_2-based materials with tailored photoactivity.

1.5.2 Titanium precursors

Biological precipitation of titanium dioxide has been conducted with various titania precursors such as titanium isopropoxide[91] and titanium (IV) bis(ammonium lactato) dihydroxide (TiBALDH).[75] For this research, the focus is on TiBALDH, a water-soluble precursor that is stable at near neutral pH,[92] and consists of a titanium cation octahedrally coordinated with two bidentate lactate ligands in the x-y plane and two hydroxyl groups in the z-plane (**Figure 1.6A**).[80] The hydrolytically stable TiBALDH is however prone to nucleophilic attack in both the x-y plane at the lactate ligands, and the z-plane at the hydroxyl groups.[80] In the z-plane, a single hydrolysis and condensation event is required to remove the hydroxyl group and coordinate two TiO_6 octahedrons through a corner-sharing bond (**Figure 1.6B**), a structure most commonly found in rutile titania. However, in the x-y plane, two hydrolysis and condensation reactions are required to remove the

bidentate lactate, which leads to edge-sharing bonds (**Figure 1.6C**) that are more common in anatase titania. In peptide-aided precipitation reactions employing TiBALDH, both reactions may occur, which leads to the production of an amorphous TiO_2 material.[92]

1.5.3 Mineralization with a TiBALDH precursor

Expanding on the role of silicateins in biosilica formation, Sumerel and coworkers were first to show that titanium dioxide could be synthesized at low temperature, low pressure, and near neutral pH using TiBALDH as a precursor, and silicatein filaments as an inducer.[75] The TiO_2 thus formed was amorphous and exhibited characteristics of both rutile and anatase crystal phases.[75] Spherical TiO_2 particles with rutile characteristics and ranging in size between 200-400 nm were also obtained by adding recombinant silaffin (rSILC) to a solution of TiBALDH.[93]

In later work, Dickerson et al. found that peptides enriched in arginine, lysine, and histidine residues promoted the formation of amorphous TiO_2 containing small amounts of anatase and TiO_2(B) nanocrystals, and material yields increased with the amount of positive charge.[94] Consistent with an important role for basic residues, the cationic protein protamine was also found to precipitate amorphous TiO_2 from a TiBALDH precursor.[95] To investigate how peptide conformation would influence titania precipitation, the TiO_2 binding peptides, STB1 and RSTB1, mineralization ability was compared to a linear TiO_2 binding peptide LSTB1.[96] These results demonstrated that the constrained cyclic peptides mineralized titania more efficiently at pH 8, while the linear LSTB1 peptide did not, proving that the flexibility of the peptide plays a role in the condensation process of the titania precursor.[96]

Not only was peptide sequence and structure studied, so was the type of titania precursor used and other varying reaction conditions. Titania mineralization can also be done with titanium alkoxides as the precursor.[91] Puddu and colleagues showed that all reagents in the reaction play a role, by testing titania mineralization with known titania binding peptides, Ti-1 and Ti-2, and how

the mineralized titania results changed with varying buffers.[97] Mineralization efficiency and product morphology was compared in water, tris, and phosphate buffer. It was concluded that the phosphate ions in the phosphate buffer can co-precipitate titanium phosphate and titanium dioxide, yielding an unpure titanium dioxide with a different morphology than the TiO_2 produced in tris or water.[97]

Previous work demonstrates that peptide charge and structure change titania mineralization, but similar to silica formation, the reaction conditions and reagents have an effect. Since the initial TiO_2 mineralization work researchers have concentrated on making functional TiO_2 materials. Uniform and dispersed titania nanoparticles have been grown on micropatterned protein-polyelectrolyte surfaces at ambient conditions using a TiBALDH precursor.[76] Bifunctional peptides have also been designed to produce metal oxide-metal nanocomposites like TiO_2-Ag, TiO_2-Au, to make controlled fabricated materials.[77,98]

1.6 Reaction-diffusion processes

Reaction-diffusion processes are common in nature and produce beautiful patterns and structures,[99] (**Figure 1.7**), one specific example is rock patterns.[100,101] One aim of biomimetics research is to emulate the ability of nature to produce such unique patterns with graded materials. For instance, Gryzbowski and colleagues used a wet-stamping process (WETS) to create micro and nanostructures of defined geometry via reaction diffusion.[99,102] The approach relies on the use of a PDMS stamp to transfer a reactant such as $FeCl_3$ or $CoCl_2$, to a dry gelatin substrate infused with potassium hexacyanoferrate, $K_4Fe(CN)_6$. The reactants diffuse from the patterned area and react with hexacyanoferrate to form a thin precipitate with defined geometry.[99] Mathematically the process can be modeled, and the precipitation shapes predicted, by solving a set of partial differential equations of the form:

$$\frac{\partial c_i(r,t)}{\partial t} = \nabla \cdot (D_i \nabla c_i) + R_i(c_j, r, t) \quad (1.1)$$

where r is the characteristic dimension, t the time, c_i the concentration of species i, D_i the diffusion coefficient, and R_i the reaction rate.[99]

This dissertation focuses on the study of a specific SBP, Car9, and its' interactions with the inorganic materials SiO_2 and TiO_2. Chapter 2 discusses the development of an in-depth kinetic model that quantifies the binding kinetics of a protein-peptide fusion (sfGFP-Car9) to SiO_2 using SPR. Chapter 3 uses the kinetic model combined with Rosetta structural predictions and MD simulations to study the effects a SBP's structure has on its' function using a panel of Car9 variants. In Chapter 4 the SiO_2 knowledge from Chapters 2 and 3 is used to understand how a peptide sequence and structure influences the formation, crystallinity, and morphology of precipitated TiO_2. Also discussed is TiO_2 precipitation in a reaction-diffusion environment (Chapter 5) and the influence of peptide location in the sfGFP framework on SiO_2 binding (Chapter 6).

1.7 Figures

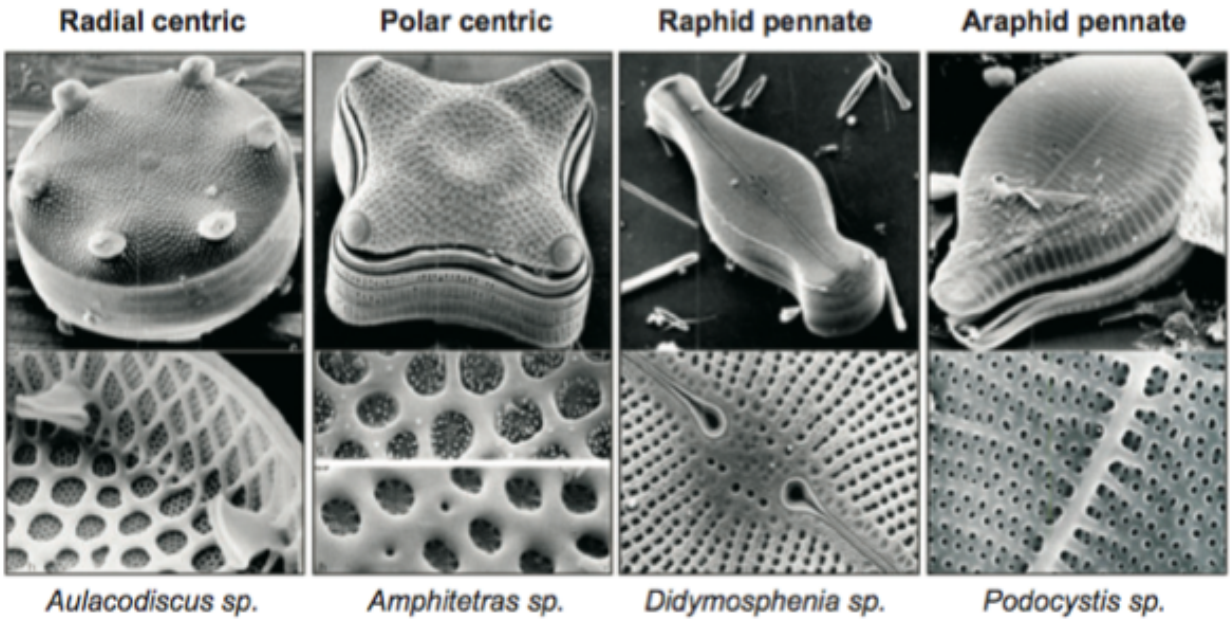


Figure 1.1 SEM images of the cell walls (frustules) of four different diatom species. [18]

Figure 1.2 The amino acid sequence of sil1p with acidic amino acids (green), basic amino acids (red) and the signal peptide (blue). The different R peptides, R1 to R7, (labeled on the left) that compose the C-terminal domain.[20]

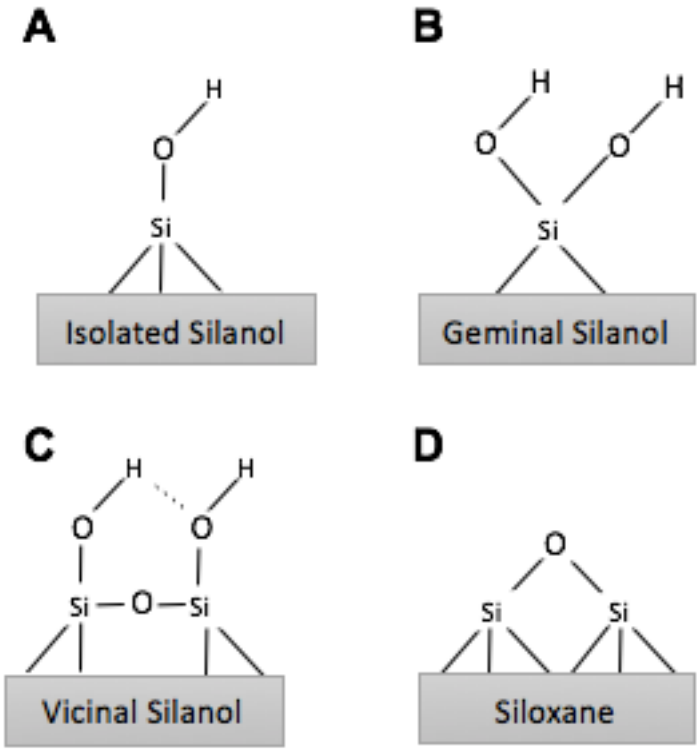


Figure 1.3 Surface chemistry of silica**.** Four types of groups, **(A)** isolated silanol Si(OH), **(B)** geminal silanol $Si(OH)_2$, **(C)** vicinal silanol, and **(D)** surface siloxane Si-O-Si, are found on the surface of silica.

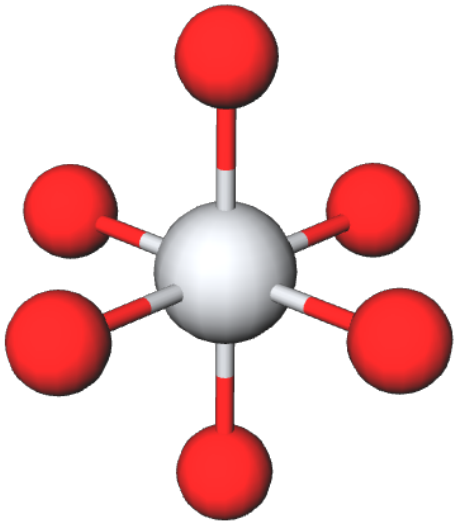


Figure 1.4 A TiO_6 octahedron with the titanium atom in silver and oxygen atoms in red. (Prepared using MolView)

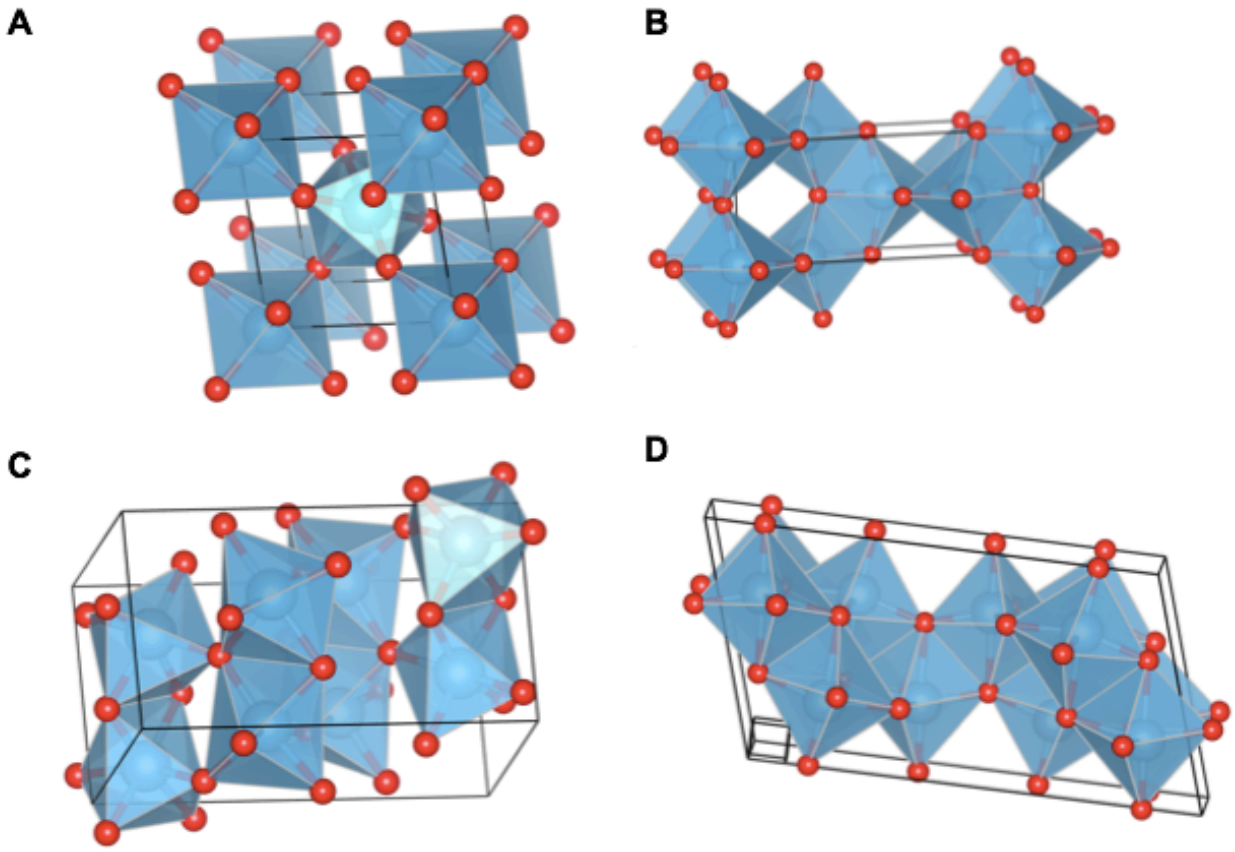


Figure 1.5 Organization of TiO_6 octahedra in various phases of titania: **(A)** rutile, **(B)** anatase, **(C)** brookite, and **(D)** TiO_2(B). Titanium atoms are represented by blue spheres and oxygen atoms by red ones. (Prepared using VESTA)

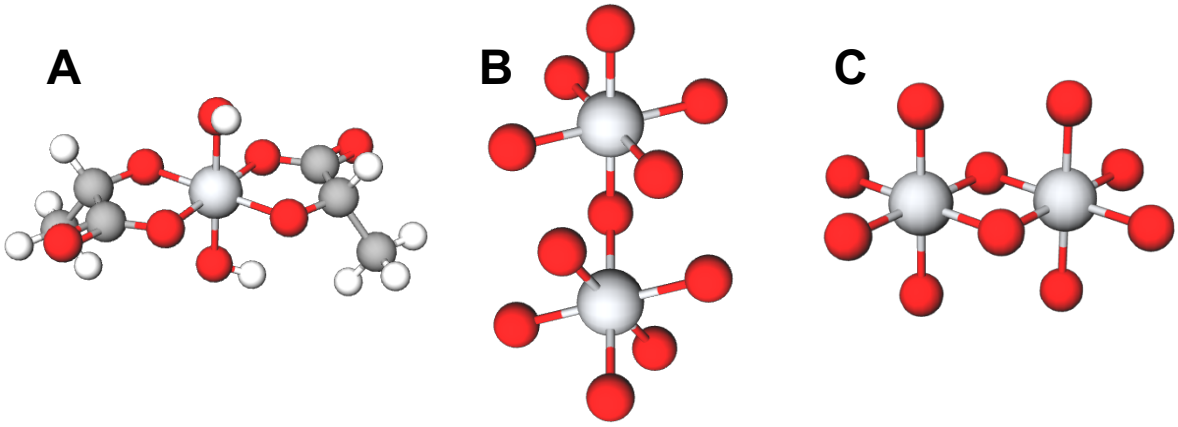


Figure 1.6 (A) Structure of the water-soluble titania precursor TiBALDH with titanium atoms in silver, oxygen atoms in red, carbon atoms in grey, and hydrogen atoms in white. Structures of **(B)** corner-shared TiO_6 octahedra, and **(C)** edge-shared octahedra. (Prepared using MolView)

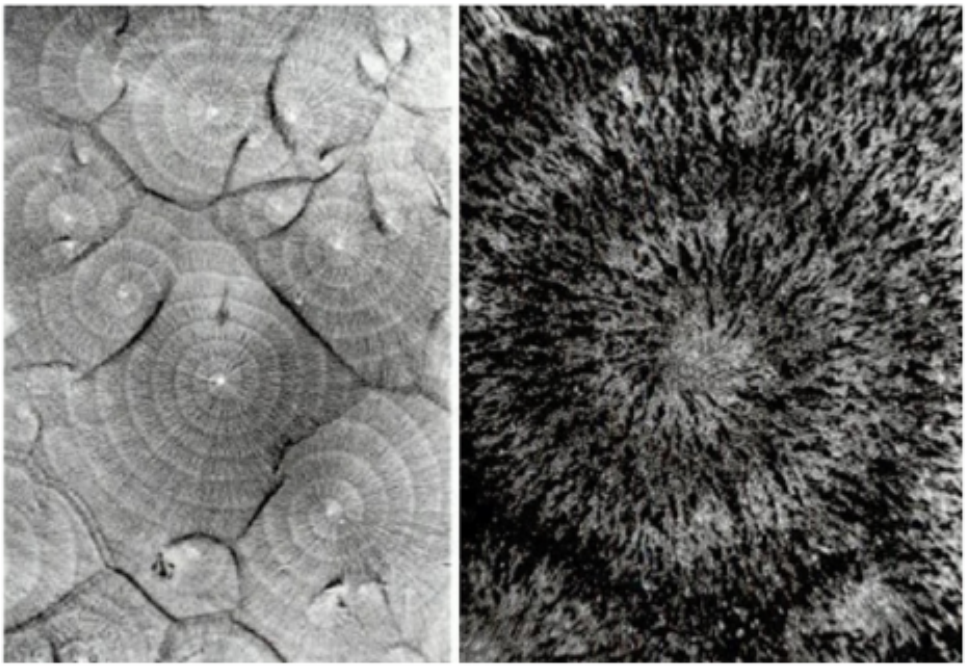


Figure 1.7 Reaction diffusion patterns formed in Dictyostelium cells.[100]

Chapter 2 Modeling the Cooperative Adsorption of Solid-Binding Proteins on Silica: Molecular Insights from SPR Measurements

2.1 Introduction

Whether derived from biomineralizing organisms or identified by phage or cell surface display, solid-binding peptides (SBPs, *a.k.a.* materials-binding peptides) provide a rich source of synthetic or genetically encodable chemistry to control the structure, composition, and properties of advanced materials and devices.[8,103-105] Applications are numerous and range from biomedical engineering[106-109] to catalysis,[110,111] and from energy materials[112-114] to opto-electronics.[59,115] However, despite many success stories and progress in characterization and simulation, we remain unable to replicate nature's exquisite degree of control over inorganic morphogenesis and assembly, in part because our understanding of how SBPs interact with surfaces remains limited.

While using SPR to quantify the binding of sfGFP-Car9 (a fusion between superfolder green fluorescent protein[11] and the Car9 SBP) to silica-coated gold chips, we observed an unusual sigmoidal adsorption behavior. Here, we describe the construction and validation of a two-step cooperative model that captures the silica binding kinetics of sfGFP-Car9 and other Car9 fusion proteins with high fidelity. We further show that extracted model parameters deliver insights on the complex interplay between surface, SBPs, fused framework and ions, which should improve our ability to design solid-binding proteins for practical applications and predictive materials synthesis.

2.2 Materials and Methods

2.2.1 DNA Manipulations and Protein Purification

Plasmid pET-24a(+)-sfGFP-Car9 was described elsewhere.[10] Plasmid pET-24a(+)-nh-mCherry-Car9 which encodes a mCherry-Car9 fusion protein with an unmodified N-terminus was constructed by PCR amplification of this segment from pET-24a(+)-mCherry-Car9[10] and ligation

into the *Nde*I and *Xho*I sites of pET-24a(+). Plasmid pET-24a(+)-sfGFP-Car9Q4 which encodes a fusion between sfGFP and Car9Q4 (DSARGFQQPGQQ) with an intervening GGGS linker was obtained by inserting a DNA cassette encoding these elements between the *Hin*dIII and *Xho*I sites of pET-24a(+)-sfGFP-Car9. To create plasmid pET-24a(+)-GK-Car9 which encodes a fusion between GK and Car9 with a GGGS linker, a DNA fragment coding for glucose hexokinase (GK) and flanked by *Nde*I and *Hin*dIII sites was PCR-amplified from *T. thermophilus* DNA and inserted into the same sites of pET-24a(+)-sfGFP-Car9. All constructs were verified by DNA sequencing. Proteins were expressed and purified by silica-affinity chromatography as described,[10,116] and the buffer exchanged to 20 mM Tris-HCl, pH 7.5 by dialysis. Concentrations were determined by BCA assay and samples were stored at -20°C.

2.2.2 SPR Experiments

Silica-coated surface plasmon resonance (SPR) chips were fabricated in house using a glass substrate coated with a 2 nm titanium adhesion layer, a 48 nm evaporated gold film, and a 4 nm silicon film deposited by plasma enhanced chemical vapor deposition (PEVCD).[9] Chips were cleaned by ethanol and UV-ozone treatment before use and mounted on the flow cell of a SPR sensor from the Institute of Photonics and Electronics (Prague, Czech Republic). All experiments were conducted at a flow rate of 50 µLmin^{-1} and a constant temperature of 25°C with protein concentrations varying between 0.0312 µmol L^{-1} and 1 µmol L^{-1} in 20 mM Tris-HCl at pH 7.5. Increasing the flow rate to 75 µL min^{-1} had no impact on the sigmoidicity of sfGFP-Car9 adsorption to silica, confirming kinetic control (Appendix A Note 1).

2.2.3 Model Parameters Estimation

Initial guesses for k_{a1}, k_{d1}, k_{a2}, k_{d2}, and u (Table A.1) were informed by the physical characteristics of the system and literature values. Rate constants for the first association reaction are typically on the order of 10^4 to 10^5 M^{-1} s^{-1} (classic Langmuir isotherm), while subsequent

association and dissociation rate constants are in the 10^{-3} to 10^{-4} s^{-1} range based on isotherm modeling of SPR data.[117] Bounds were adjusted in a 25% window based on initial guesses.

An objective function subject to governing equations, initial conditions, and parameter bounds was defined to minimize the Mean Square Error (MSE) of the SPR response. With the MSE defined as the sum of the squares of differences between model and experimental outputs divided by the total number of experimental data points, this function is expressed as:

$$Min\ \frac{1}{N}\sum_{j=1}^{N}[R_{exp,j} - R_{model,j}(p)]^2$$

$$s.t. \quad p^L \leq p \leq p^u$$

where N is the total number of experimental data points for the adsorption and desorption phases, $R_{exp,j}$ the experimental signal value for the j[th] data point, $R_{model,j}(p)$ the model-predicted value for the same point, p the vector of estimating parameters, and p^L and p^U its lower and upper bounds. To account for the large difference in the magnitude of the rate constants and perform a more accurate iterative optimization, the reference rate constant of the first association (k'_{a1}) was scaled by 10^5 while the other rate constants were expressed as an exponent of 10 (i.e., $k_{d1} = 10^{-k_{d1}}$, $k_{a2} = 10^{-k_{a2}}$, and $k_{d2} = 10^{-k_{d2}}$). Initial guesses, bounds, and converged parameters are summarized in Table A.1 with values rounded off to three-decimal place. Simulations and optimizations were carried out on a workstation with dual 8-core, 3.10GHz Intel Xeon processors, 32.0 GB RAM using the NLPsolve (optimization method: nonlinear simplex, and evaluation limit: maximum) in the optimization package of the Maple® software.

2.3 Results and Discussion

2.3.1 The Car9 Solid-Binding Peptide Confers Proteins to Which it is Fused Distinct Sigmoidal Adsorption Kinetics on Silica.

Silica (SiO_2) surfaces are terminated with siloxanes (O-Si-O) and silanol (Si-OH) groups which can be found as isolated silanols (≡Si-OH), geminal silanols (=$Si(OH)_2$), or vicinal silanols (two H-bonded silanol groups), at distributions and with concentrations that are modulated by method of preparation and surface ionization state.[53] Proteins that bind to silica in a nonspecific fashion do so via electrostatic and H-bonding interactions with silanol groups, and hydrophobic interactions with siloxane groups.[118] Although the same interactions control the adhesion of naturally-occurring and combinatorially-selected silica binding proteins and peptides to SiO_2, they are clearly deployed in a specific fashion.[54,55,58]

To better understand the molecular mechanisms that underpin the interaction of the Car9 SBP with silica surfaces, we quantified the binding kinetics of the sfGFP-Car9 fusion protein[17] using a gold SPR chip coated with a thin layer of SiO_2. Under conditions where wild type sfGFP has little affinity for silica (Figure A.1), all sensorgrams displayed a sigmoidal adsorption behavior typical of a cooperative process (**Figure 2.1A**). Indeed, the kinetics of the adsorption phase were poorly captured by a rate-law based 1:1 Langmuir binding model which assumes homogeneous surface, identical binding sites, monolayer coverage, and no interactions between adsorbate molecules on adjacent sites (**Figure 2.1B**, dashed lines).[119]

Car9 contains five basic amino acids whose positively charged sidechains might contact the negatively-charged silica surface. To probe the role of these interactions in the non-standard adsorption behavior, we constructed a Car9 variant called Car9Q4 in which Lys-7, Lys-8, Lys-11 and Arg-12 were converted to uncharged glutamines. To our surprise, the sfGFP-Car9Q4 fusion protein remained functional for silica binding and could be purified by silica affinity chromatography (Figure A.2). However, its affinity for the material was drastically reduced, the sensorgrams lost their sigmoidal character, and it was possible to accurately capture adsorption kinetics with a rate-based Langmuir fit (**Figure 2.1C**; also see Figure A.8). We conclude that, as

with the silaffin-derived R5 peptide.[33,34] and other combinatorially-selected silica-binding peptides,[54] basic residues play a key role in high-affinity binding to silica. For Car9, they are also critical in determining cooperative adsorption.

To rule out the possibility that sigmoidal adsorption is driven by GFP dimerization events,[11] and to explore the contribution of the fused protein framework, we appended Car9 to the C-terminus of the red fluorescent protein mCherry, and purified the resulting fusion by silica affinity chromatography (Figure A.2). **Figure 2.1D** shows that mCherry-Car9 followed similar sigmoidal kinetics. However, saturation coverage at equimolar protein concentration was about 20-30% lower for mCherry-Car9 compared to sfGFP-Car9 in spite of the fact that the two proteins have comparable molar masses, pIs, and architectures (Figure A.3).

In an effort to further generalize results, we fused Car9 to the C-terminus of a predicted glucose kinase (GK) gene amplified from the genome of *Geobacillus stearothermophilus*. To our knowledge, this enzyme has not previously been expressed, and because bacterial GKs can be monomeric, dimeric or tetrameric,[120] its quaternary structure is unknown. Thus, GK-Car9 provides an unbiased platform to test the influence of both the Car9 SBP and framework oligomeric status on silica adhesion.

We found that GK-Car9 expressed well in *E. coli* and that it could be purified to near homogeneity by silica affinity chromatography (Figure A.2). Native polyacrylamide gel analysis further revealed that, like *Thermus thermophilus* GK,[121] *G. stearothermophilus* glucokinase forms a homotetramer (**Figure 2.2A**). A BLAST alignment revealed that the two proteins share 58% homology and 40% identity at the amino acid level (Figure A.4), allowing us to model the structure of the *G. stearothermophilus* tetramer using SWISS-MODEL[122] and the *T. thermophilus* GK coordinates[121] (Figures A.4B and A.5). This exercise revealed that the C-termini are solvent-accessible and located in the projecting lobes of each protomer (**Figure 2.2B**).

SPR sensorgrams collected with GK-Car9 exhibited the characteristic sigmoidal adsorption behavior of Car9 fusions. Surface saturation was lower than with either sfGFP-Car9 or mCherry-Car9 (**Figure 2.2C**). Thus, at least when carboxyl termini are accessible, the oligomerization status of the fusion partner does not affect the cooperative nature of binding. To summarize, the Car9 SBP drives high affinity binding to silica through a unique cooperative adsorption process that relies on the presence of basic residues but is independent of the identity and oligomeric status of the fused protein scaffold.

2.3.2 Modeling the Kinetics of Car9-Silica Interactions

To model the sigmoidal character of Car9-mediated adsorption kinetics, we developed a rate law expression based on the Frumkin-Fowler-Guggenheim (FFG) isotherm,[123] which takes into consideration attractive or repulsive lateral interactions between neighboring adsorbate molecules and reduces to the Langmuir formalism when these interactions are absent.[124,125] Furthermore, because the kinetics of the desorption phase are poorly captured by a single exponential decay (Figure A.6), we included a second reaction step that can be conceptualized as a change in the conformation of the SBP once initial adsorption and nearest neighbor contacts have occurred at the silica interface. Diffusional limitations were neglected as justified in Appendix A Note 1. A schematic depiction of this process, and the governing equations of the resulting two-step model is presented in **Figure 2.3**.

The rate law expression for the initial adsorption step considers the product of the flux and a sticking probability which depends on the lateral interaction energy between adsorbed species. The equivalent rate constant for adsorption (k_{a1} in **Table 2.1**) is thus expressed as:

$$k_{a1} = \frac{N_A}{\sigma_{max}}\sqrt{\frac{RT}{2\pi M}}\exp\left(\frac{-E_0}{RT}\right)exp\left(\frac{-E_0\theta}{RT}\right) \tag{2.1}$$

where N_A is Avogadro's number, σ_{max} is the density of binding sites on the silica surface which is assumed to be identical under prescribed solution conditions, M is the molar mass of the adsorbing protein, R is the universal gas constant, E_0 is the molar intrinsic activation energy for bond formation between SBP and surface, and E_i is the molar interaction energy between nearest neighbors, a quantity that is negative for attractive interactions. We assume that the time-dependent SPR responses $R_1(t)$ and $R_2(t)$ are directly proportional to protein surface concentrations. This allows us to take θ, the surface coverage at time t, to be proportional to the ratio of the sum of the SPR signals in each adsorption phase divided by R_{max}, the maximum SPR signal at saturation:

$$\theta \propto \frac{R_1(t) + R_2(t)}{R_{max}} \tag{2.2}$$

At fixed temperature T, and for a given solid-binding protein and concentration of binding sites on the silica surface, the species and temperature-dependent properties in Equation (2.1) can be combined into a single parameter to give a rate constant of the form:

$$k_{a1} = k'_{a1} \exp(u\theta) \tag{2.3}$$

where:

$$k'_{a1}(T) = \frac{N_A}{\sigma_{max}} \sqrt{\frac{RT}{2\pi M}}\, exp\left(\frac{-E_0}{RT}\right) \tag{2.4}$$

and

$$u(T) = \frac{-E_i}{RT} \tag{2.5}$$

Thus, the entirety of protein-surface interactions is accounted for in the modified adsorption rate constant k'_{a1}. Of note, the overall rate constant k_{a1} also depends on the concentration of binding sites available on the surface through the exponential factor $\exp(u\theta)$ which changes with surface coverage. This factor accounts for cooperative SBP-SBP interactions and framework contributions. A detailed derivation of the governing equations can be found in Appendix A Note 2.

2.3.3 Model Validation

The solid lines in **Figure 2.1A** show that the proposed model closely fits the experimental data and that it accurately captures the distinct sigmoidal adsorption phase of sfGFP-Car9 on silica. Desorption kinetics, which are simultaneously fit by the optimization routine, are also captured with good fidelity. However, the match with experimental data is not as good, presumably because our model does not account for protein readsorption events. Examination of extracted parameters for sfGFP-Car9 (**Table 2.1**) reveal that the first adsorption step is kinetically dominant and that it occurs with positive cooperativity ($u > 0$), firmly implicating protein-protein interactions in high-affinity silica binding. The second step – an isomerization reaction – occurs with significantly slower kinetics (**Table 2.1**), as confirmed in Figure A.7 by a comparison of the simulated values of instantaneous adsorption and desorption rates for each reaction.

Our observation that sfGFP-Car9, mCherry-Car9 and GK-Car9 adsorb to silica with comparable u values (**Table 2.1**) indicates that interactions between nearest neighbor Car9 segments are the main driver for cooperativity and that framework-framework interactions only play a minor role in the process. The slight (~15%) increase in the value of u for GK-Car9 is not unexpected and consistent with a situation in which co-presentation of four SBPs per GK-Car9 oligomer increases the likelihood of lateral interactions with Car9 segments from neighboring tetramers.

Table 2.1 further shows that both mCherry-Car9 and GK-Car9 adsorb to silica with a higher k'_{a1} than sfGFP-Car9. In the case of GK-Car9, this result is well explained by the fact that four Car9 extensions per GK oligomer are more effective at reducing the energy barrier E_0 than a single SBP because they contribute combined attractive electrostatic interactions between protein and surface. The associated increase in k'_{a1} clearly dominates over the opposite effect brought about by the higher molar mass of GK-Car9 protomers (34.2-kDa compared to 28.6-kDa for

sfGFP-Car9 monomers; see Equation (2.4)). In the case of mCherry-Car9 – a monomeric protein comparable in mass and shape to sfGFP-Car9 – the increase in k'_{a1} likely results from more subtle framework effects. For instance, the "bottom" face of mCherry from which the Car9 extension projects has a higher electrostatic potential than the bottom face of sfGFP (Figure A.3). This might minimize short range electrostatic interactions between the mCherry framework and the positively charged SBP, allowing Car9 to more efficiently engage the silica surface.

It is finally of interest to note that both mCherry-Car9 and GK-Car9 desorb at higher rates than sfGFP-Car9 after initial adsorption. The higher k_{d1} values, along with a more complex interplay of association and dissociation events during the isomerization step account for the lower saturation coverage (e.g., compare **Figures 2.1A**, **2.1D** and **2.2C** at 0.1 μM). For comparison, fitting of sfGFP-Car9Q4 sensorgram data with a rate law-based Langmuir model that is appropriate for the dataset (**Figure 2.1C** and Figure A.8) yields values of k_a (3000 ± 100 $M^{-1}s^{-1}$) and k_d (0.18 × 10^{-3} ± 0.07 × 10^{-3} s^{-1}) that are typical of nonspecific protein adsorption processes.[126]

2.3.4 Influence of Ionic Strength

To assess the robustness of our model, we conducted additional SPR experiments using 0.125 μM sfGFP-Car9 and increasing concentrations of sodium chloride to screen charges on both silica surface and protein shell. All sensorgrams retained a sigmoidal character, establishing the validity of the cooperative two-step model under more physiological solution conditions (**Figure 2.4**). However, there was a ~40% decrease in saturation surface coverage in the presence of 50 mM NaCl, and an additional 10% decrease in θ at 100 mM NaCl. These results can be explained by the screening of silanol-rich Car9 binding sites by sodium ions, and consistent with the observation that the zeta potential of silica nanoparticles decreases from -55 mV to -27 mV when the NaCl concentration is raised from 0 to 100 mM at neutral pH.[127]

Parameters extracted from the model provide validation for the above hypothesis and insights on the molecular basis of ionic effects (**Table 2.1**). First, a decrease in the net number of Car9 binding sites (captured by parameter σ_{max} in Equation (2.4)) should cause k'_{a1} to increase with the salt concentration, a behavior that we observe experimentally. NaCl may also contribute to the increase in k'_{a1} by shielding repulsive interactions between sfGFP framework and silica surface, thus reducing the activation barrier E_0. Second, there is a 25% decline in the value of u in the presence of salt. This suggests that attractive electrostatic interactions between neighboring Car9 domains (and to a lower extent sfGFP monomers, see above) contribute to cooperativity and are shielded in the presence of salt. Third, the rise in the rate of protein desorption (k_{d1}) with salt concentration is consistent with such weakened lateral interaction and a lower protein coverage, much like we observe with GK-Car9. Finally, there is a progressive increase in the rate constant for the isomerization step (k_{a2}) with salt concentration (**Table 2.1**). While the underlying mechanism remains unclear, it is tempting to hypothesize that ionic screening of lysine, arginine and aspartate side chains facilitates conformational rearrangement of Car9 extensions at the silica interface.

2.4 Conclusions

SPR[67,68] and QCM[128] are mature characterization techniques that have proven remarkably useful for quantifying the binding of biomacromolecules to ligands immobilized on sensor chips. They have also been used for well over a decade to study the direct interaction of SBPs with inorganic and synthetic surfaces.[65,104] In most cases, adsorption isotherms are treated with a 1:1 Langmuir model and kinetic parameters and free energies of binding are extracted from the fit. At the exception of a handful of studies by Sarikaya and coworkers[65,71-73] in which a biexponential Langmuir model was used to capture non-ideality in the binding of gold-binding peptides– a phenomenon originally thought to be associated with the polycrystalline nature of the surface but

later attributed to surface diffusivity and time-dependent evolution of adsorbed peptides structures[74] – little attention has been paid to the appropriateness of the Langmuir treatment.[129] The fact that QCM and SPR measurements are often noisy when conducted at low concentrations of synthetic peptides further complicates the recognition that a SBP adsorption process might not obey Langmuir kinetics.

Here, we have used large fusion proteins to characterize the adsorption of the promiscuous and technologically useful Car9 peptide to silica-coated SPR chips. The approach yields high-quality sensorgrams, enables SBP compositional and functional analysis within a solubilizing context, and relevant to our long-term goal of installing SBPs within computationally designed scaffolds for controlled assembly of hierarchical structures, it provides molecular insights on how SBP, surface, fusion partner, and solution conditions modulate binding.

Our results indicate that Car9 endows protein scaffolds to which it is fused with a sigmoidal adsorption behavior on silica, irrespective of – but influenced by – the identity of the host framework and the SBP valency. This unique adsorption modality was accurately captured by a two-step kinetic reaction model in which cooperative binding driven by intermolecular contacts between Car9 extensions (and to a lesser extent by fused framework contributions) is followed by an isomerization step that "locks" SBP and fusion partner in a high-affinity interaction with the surface. The model is consistent with both the tendency of certain SBPs to self-assemble into ordered structures at interfaces,[64,74,115,130] and with the enthalpic binding mechanism proposed by Walsh and Knecht[104,131,132] in which anchor residues (amino acids that spend a significant amount of time at the interface) mediate strong enthalpic contacts between SBP and silica. By contrast, sfGFP-Car9Q4 exhibits a more traditional Langmuir behavior. It may therefore belong to the class of entropic SBPs which achieve high affinity binding by assuming multiple conformations at the interface.[104,131,132]

Such a conversion in binding modalities between Car9 isoforms suggests that it should be possible to gain a detailed understanding of how Car9 adsorbs to silica by generating a panel of Car9 mutants, pairing model parameter extraction from SPR experiments with structural predictions and advanced molecular dynamics simulations.

The approach described herein should be generalizable to other technologically useful SBPs exhibiting non-Langmuir adsorption. It should also provide a rich source of information on how SBP-SBP interactions, host framework and solution conditions modulate adhesion, which will be critical in enabling predictive material synthesis with solid-binding proteins.

2.5 Figures

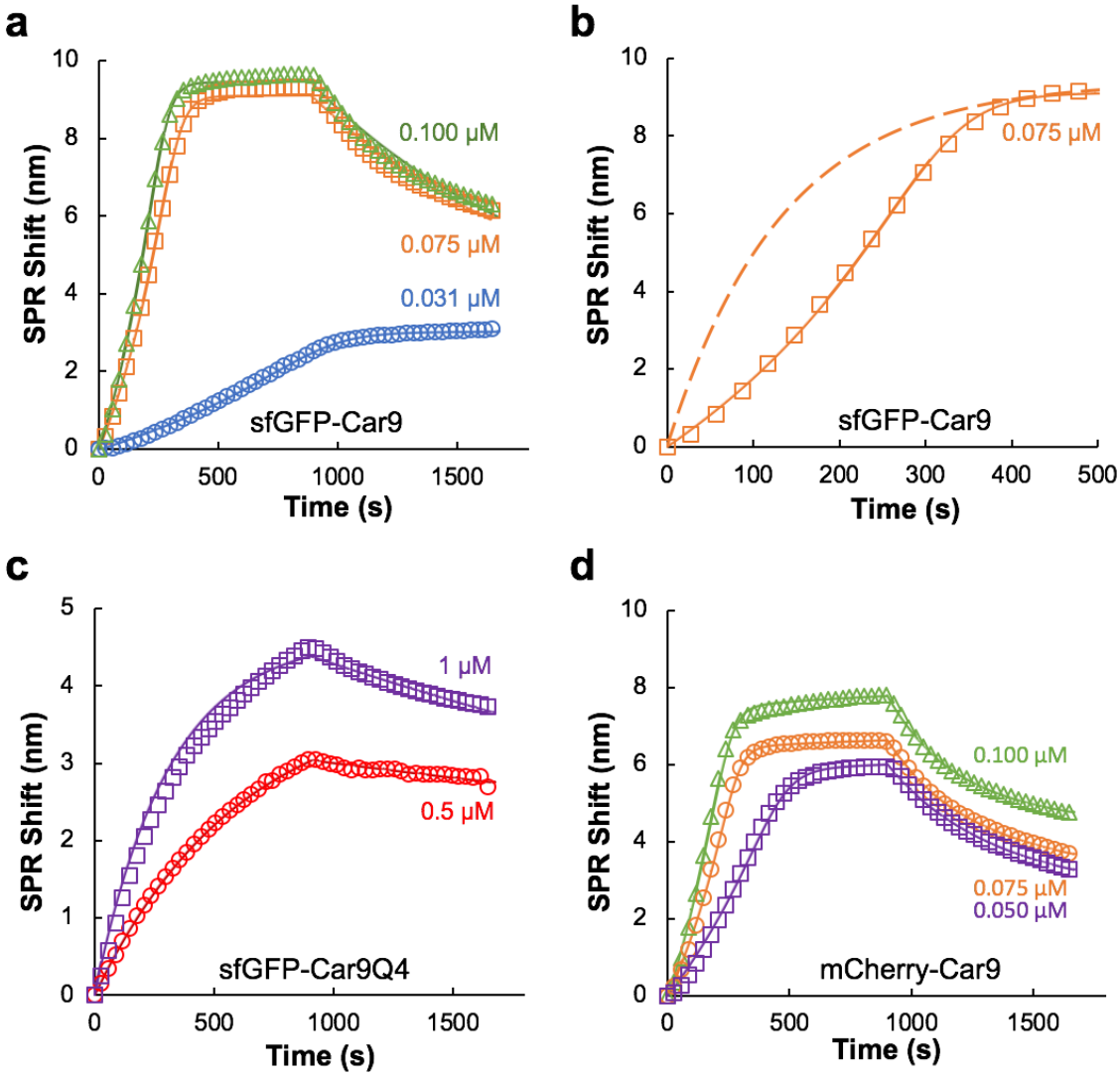


Figure 2.1 Adsorption kinetics of sfGFP-Car9 **(A-B)**, sfGFP-Car9Q4 **(C)**, and mCherry-Car9 **(D)** on silica-coated SPR chips. Sensorgrams were collected using the indicated protein concentrations. A wash phase in which protein-free buffer is flowed over the chip starts at t = 900 s and causes protein desorption. In panel **B**, fitting of the sfGFP-Car9 adsorption phase with a kinetic Langmuir model is denoted with a dashed orange line while the fit obtained with the two-step cooperative adsorption model is shown with a solid orange line. The SPR shift is plotted on a different scale and higher protein concentrations are used in panel **C**. All solid lines in panels **A** and **D** were obtained by fitting experimental data to the two-step cooperative adsorption model. All solid lines in panel **C** were obtained by fitting experimental data to a rate law-based Langmuir model.

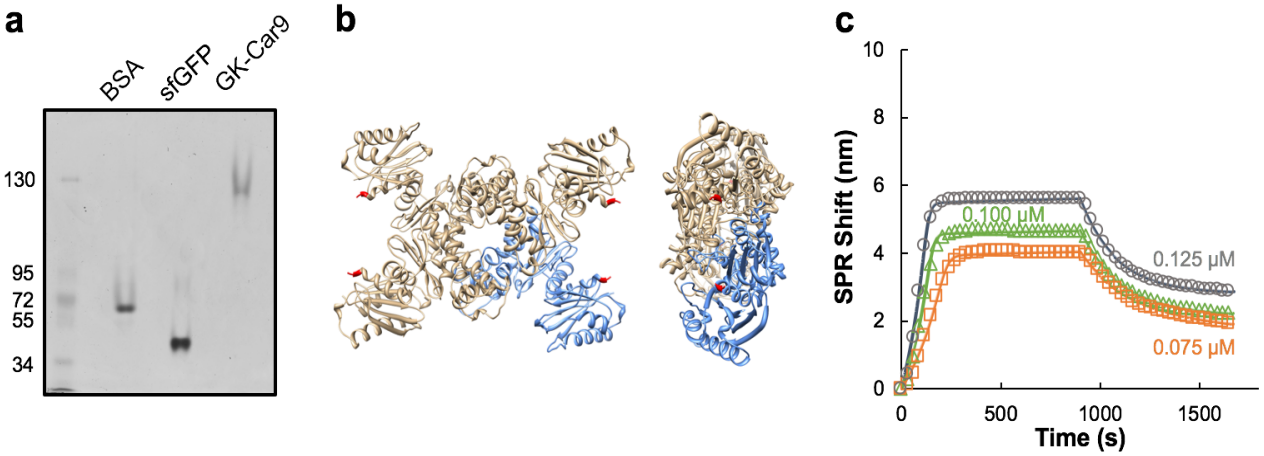


Figure 2.2 Oligomerization, predicted structure, and silica-binding properties of GK-Car9. **(A)** Native polyacrylamide gel analysis of GK-Car9 showing the migration positions of marker proteins and those of bovine serum albumin (BSA, 66-kDa) and sfGFP (26.8-kDa). **(B)** Predicted quaternary structure of the GK oligomer. In these front and side views, a monomer is shown in blue and the four C-termini are colored in red. **(C)** Adsorption kinetics of GK-Car9 on silica-coated SPR chips. Fits obtained with the two-step cooperative adsorption model are shown with solid lines.

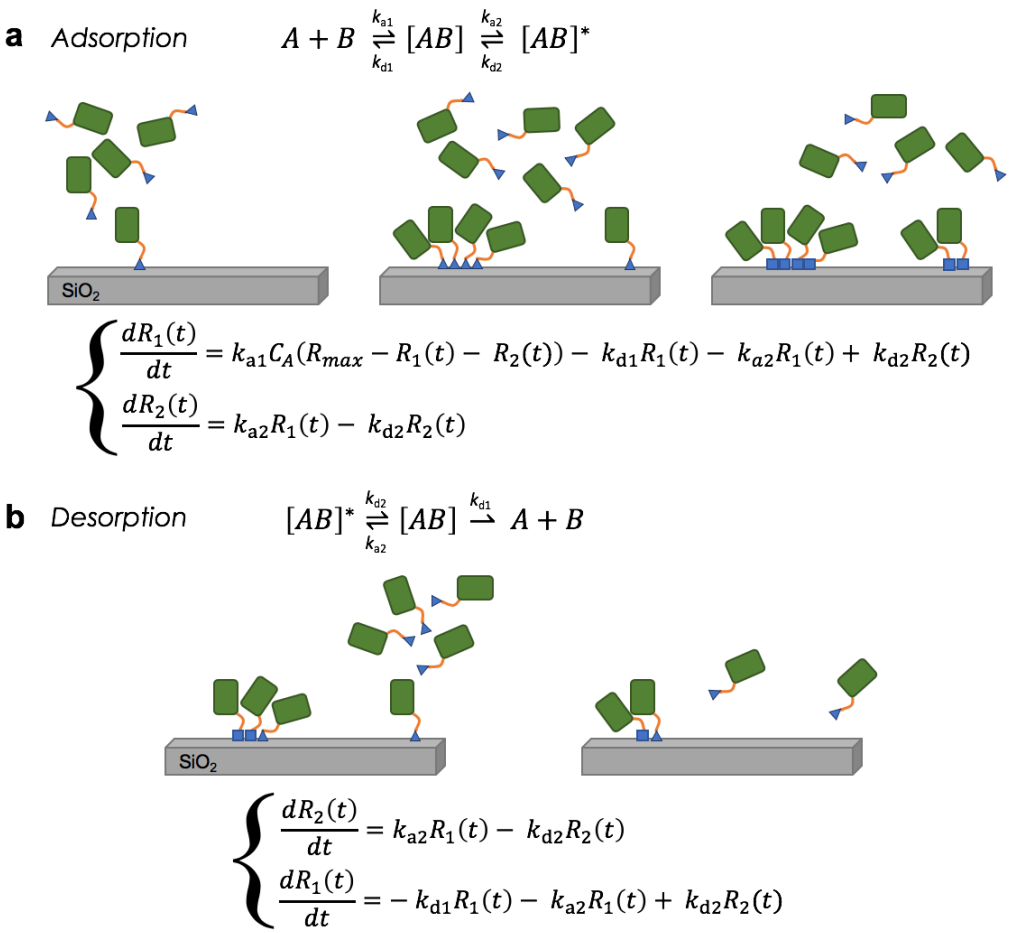


Figure 2.3 Proposed two-step cooperative binding model for the adsorption of Car9 fusion proteins to silica surfaces. The adsorption and desorption phases are schematically illustrated with proteins depicted in green, linkers in orange, and Car9 SBPs as blue triangles or blue squares to reflect the fact that they undergo a conformational change after they have bound to their nearest neighbors. (A similar conformation change occurs upon desorption). The governing equations which are simultaneously solved by the optimization routine are also shown. $R_1(t)$ and $R_2(t)$ correspond to the time-dependent SPR responses associated with the first and second adsorption/desorption phases, respectively. The Car9 SBP drives protein adsorption to the silica surface **(A)** at a rate dependent on the rate constant k_{a1}, the bulk protein concentration C_A, and the surface concentration of free binding sites. Interactions between neighboring Car9 moieties, and to a lesser extent between fused protein frameworks, engender cooperative behavior. Bound species next undergo a conformational rearrangement at the interface or desorb from the surface. These are parallel first-order processes with rate constants k_{d1} and k_{a2}, respectively. The isomerization step is reversible, and bound forms corresponding to signal $R_2(t)$ can convert back to the original isomeric form with first-order kinetics and a rate constant k_{d2}. The desorption phase **(B)** is described by the absence of analyte in the bulk liquid, (C_A=0). Adsorbed proteins switch between isomeric forms, with parallel desorption of the isomer corresponding to signal $R_1(t)$ back into the liquid phase. This causes the gradual decay in SPR signals observed in the desorption phase of the SPR sensorgrams. See Note 2 in Appendix A for a detailed explanation of model equations.

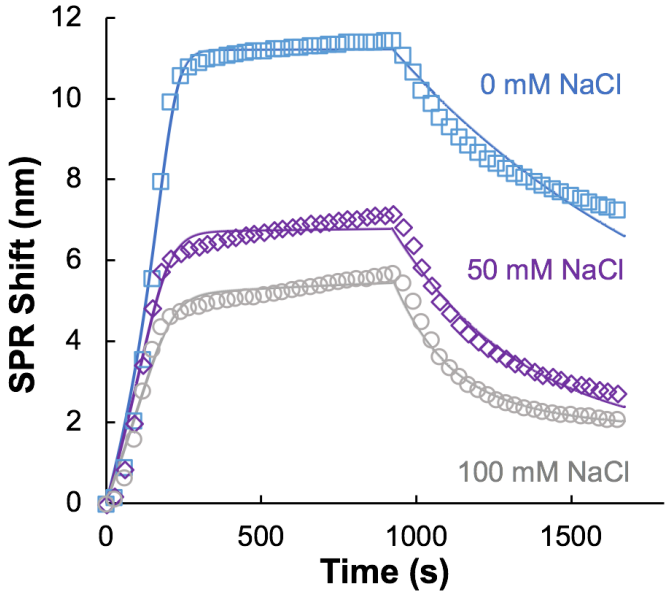


Figure 2.4 Influence of the NaCl concentration on the adsorption of sfGFP-Car9 (0.125 µM) to silica substrates.

2.6 Tables

Table 2.1 Kinetic parameters for the two-step adsorption of Car9 fusion proteins at pH 7.5

Protein[a]	**[NaCl]** (M)	**u**	k'_{a1} ($M^{-1}s^{-1}$)	k_{d1} $\times 10^{-3}$ (s^{-1})	k_{a2} $\times 10^{-3}$ (s^{-1})	k_{d2} $\times 10^{-3}$ (s^{-1})
sfGFP-Car9[a]	0	2.9±0.1	20700±1200	0.76±0.05	0.17±0.05	0.03±0.01
mCherry-Car9[a]	0	2.9±0.1	27100±1200	2.40±0.60	0.85±0.20	0.20±0.10
GK-Car9[a]	0	3.1±0.2	34100±900	4.32±0.90	0.67±0.10	0.05±0.01
sfGFP-Car9[b]	0	2.9	20000	0.84	0.12	0.060
sfGFP-Car9[b]	0.05	2.2	29000	2.15	0.20	0.020
sfGFP-Car9[b]	0.1	2.3	30000	4.71	0.51	0.100

[a]Mean and standard deviations were obtained from three fits of SPR data obtained on the same chip at protein concentrations of 0.031, 0.075 and 0.1 μM for sfGFP-Car9, 0.05, 0.075 and 0.1 μM for mCherry-Car9, and 0.075, 0.1 and 0.125 μM for GK-Car9.

[b]Model parameters were extracted from the experimental data of Fig. 4 which was conducted on a single chip and at sfGFP-Car9 concentration of 0.125 μM with the indicated concentrations of NaCl.

Chapter 3 Structure-function analysis of the Car9 silica-binding peptide

3.1 Introduction

Solid-binding peptides (SBPs) isolated from biomineralizing proteins[23] or selected by biopanning,[12,133,134] have proven remarkably useful to modify surfaces, fabricate functional architectures and devices, and control the morphology, crystallography and properties of inorganic phases.[8,64,104,133,135-140] 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 a 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),[141] 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).[104,132] For the latter class of binders, NMR studies of the titania-binding TBP hexapeptide[142] (RKLPDA) in contact with TiO_2 and SiO_2 nanoparticles,[143,144] 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.[32,145-148] 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.

In this chapter, we combine SPR measurements, the previously developed model, 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).[11] 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 propose that the crossover in adsorption modalities is due to shifts in the type of dominant interactions at the peptide-surface interface.

3.2 Materials and Methods

3.2.1 DNA manipulations and protein purification

Oligonucleotides encoding mutant Car9 sequences were purchased from Integrated DNA Technologies, hybridized, digested with *Hin*dIII and *Xho*I, and the resulting cassettes were inserted within the same sites of plasmid pET-24a(+)-sfGFP-Car9 which encodes a sfGFP-Car9 fusion protein under transcriptional control of the T7 promoter.[10] The presence of the correct mutations was verified by DNA sequencing. Proteins were purified as described[10,116] and stored at -20°C in 20 mM Tris-HCl, pH 7.5. Molecular masses of select mutants was verified by mass spectrometry.

3.2.2 SPR experiments, model parameter extraction and free energy calculations

Silica-coated surface plasmon resonance (SPR) chips consisting of a glass substrate coated with a 2 nm titanium adhesion layer, a 48 nm evaporated gold film, and a PECVD-deposited 4 nm silicon were fabricated as described.[149] Chips were cleaned by ethanol and UV-ozone treatment before being mounted on the flow cell of an SPR instrument (Institute of Photonics and Electronics,

Prague, Czech Republic). Sensorgrams were collected at 50 μLmin^{-1} and 25°C, and at protein concentrations varying between 0.063 and 1 μmol L^{-1}. Sensorgrams obtained with sfGFP-Car9 and its F6A and P9AG10A derivatives were fitted with a two-step cooperative adsorption model and the cooperativity parameters and rate constants were extracted from the fit as described.[149] Equilibrium dissociation constants were calculated for each reaction step ($K_{d1} = k_{d1}/k_{a1}$ and $K_{d2} = k_{d2}/k_{a2}$). Overall adsorption rate constants for the first reaction step (k_{a1}) were determined using Equation (3.1), the values of k_{a1}' and u extracted from the model, and monolayer surface coverages (θ) of 89% for sfGFP-Car9 and 92% for its F6A and P9AG10A variants. Monolayer coverages occurred at approximately 250 nM for all three proteins. The corresponding surface coverage for the R4QR12Q and K8AK11A variants were approximately 52% and 40%, respectively. Free energies of adsorption were calculated using $\Delta G = RT \ln(K_d)$ with R = 8.314 JK^{-1}mol^{-1} and T = 298K.

3.2.3 Rosetta calculations

Rosetta's *ab initio* fragment-assembly algorithm[150] was used to predict the peptide structures from their sequences. We used a modified version of algorithm that was found to better sample peptide structures,[151] that uses 4-mer and 6-mer fragments and fragment picking without secondary structure prediction. Around 79,000-95,000 structures were generated for each peptide sequence.

3.2.4 Molecular dynamics simulations

The Car9 dodecamer and its F6A, P9AG10A, K8AK11A, R4QR12Q variants were simulated using GROMACS 5.1.2[152] in a 6.95 nm x 6.85 nm x 7 nm box encompassing a silica surface, water molecules, a single peptide, and counterions. Three trials of Car9 and each mutant were simulated for 250ns (750 ns of MD per peptide in total) in the NVT ensemble, using the Rosetta predicted structures as the starting configuration, placed within 2 nm of the surface. The

peptide and counterions were modeled using CHARMM36,[153] the surface was obtained from INTERFACE[154], and SPC[155] water was used. The surface was adjusted to mimic pH 7 conditions, with 9% of surface hydroxyls ionized and neutralized with sodium ions. Further details surrounding MD simulations can be found in Appendix B.

The gromos[156] clustering algorithm (RMSD cut-off of 0.2 nm) in GROMACS was used to determine the top structure present in each trial, and the frequency of occurrence. To obtain an estimate of the structural differences between MD and Rosetta predictions, RMSD (using alpha carbon atoms) was calculated for each structure, using the Rosetta structure as a reference point. To visually compare Rosetta and molecular dynamics predictions of the peptide conformation, the Rosetta structure was superimposed on the top peptide cluster for each mutant using the visualization software VMD.[157] To better understand the affinity of each mutant towards the surface, we used GROMACS to tease out the short-range Coulombic and Lennard-Jones energies between the surface and the protein, and used time-averages across 3 trials for the last 100 ns of simulation, when the peptide was bound to the surface. Using PLUMED 2.4.2,[158] we tracked the z-distance between the center of mass (COM) of each sidechain in the peptide, to a reference point on the surface. These distances were time-averaged across the last 50ns for each 3 simulations; to understand the variance across the three trials, the mean and standard deviation from the time-averaged data was reported.

3.2.5 Molecular dynamics simulation details

All simulated peptides were capped with an acetyl group at the N-terminus. Unfavorable contacts in the initial configuration were removed using a steepest descent algorithm. The system was evolved in the NPT ensemble (T = 300 K, P = 1 atm) for 1 ns using the Donadio-Bussi-Parrinello thermostat[159] with a time constant of 0.1 ps, and the Berendsen barostat[160] with a time constant of 1 ps. Only the longest box dimension (z-direction) was allowed to change during the

NPT simulation. Production runs were conducted in the NVT ensemble (T = 300 K), with pressure maintained with the Donadio-Bussi-Parrinello thermostat (τ = 0.1 ps). Unfavorable contacts were removed as above and the pressure stabilized by simulating the system in the NPT ensemble for 1 ns using the Berendsen barostat.[160] Then equilibrated for 1 ns in NPT ensemble (Parrinello-Rahman barostat[161]). Following this, production runs were simulated in NVT for 250 ns for each peptide. All simulations utilized a 2 fs time step by constraining the bonds between hydrogen and other heavy atoms with the LINCS[160] algorithm. Electrostatic interactions were calculated with the particle mesh Ewald (PME)[162] summation method and a cutoff value of 1.0 nm. A van der Waals cutoff value of 1.0 nm was used.

3.3 Results and discussion

3.3.1 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 Rosetta software suite. *Ab initio* structure prediction was performed as described for the validation of designed peptide structures[151] From nearly 100,000 generated structures, Rosetta offered five low energy solutions (Figure B.1) sharing a common architecture (**Figure 3.1A**). 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 3.1A**). 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 CHARMM36[163] force field, and the INTERFACE[154] force field was used to represent the (100) surface of quartz under near neutral conditions (pH 7.5). We previously used this surface to simulate the binding of the silaffin-derived R5 peptide to silica.[32]

As shown in **Figure 3.1B** and **3.1C**, there was good agreement between the Cα 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 amino acid side chains and the surface across all three simulation trials revealed a significant involvement of R4, K7, and K8 in silica interactions (**Figure 3.1D**; also see Figure B.2): there 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).

3.3.2 Contribution of structure and electrostatics to silica binding

To understand the relationship between sequence and the observed binding motifs in Car9, we constructed a panel of four sfGFP-Car9 variants containing alanine or glutamine substitutions in residues proposed to play a structural role (F6A and P9AG10A) or to be involved in silica binding (K8AK11A and R4QR12Q). All proteins could be purified to near homogeneity by silica affinity chromatography[10,116] (Figure B.3), indicating that none of the mutations abolish silica binding. We next quantified the kinetics of adsorption by performing SPR measurements on silica-coated chips under conditions where mass transfer is not limiting and where unmodified sfGFP has little affinity for the surface.[164] We found that the F6A and P9AG10A variants exhibited the sigmoidal adsorption behavior that is characteristic of wild type sfGFP-Car9 and other Car9 fusion

proteins.[164] By contrast, the K8AK11A and R4QR12Q mutants bound to the surface in a more traditional (i.e., logarithmic) fashion (**Figure 3.2**).

The sigmoidal adsorption kinetics of F6A and P9AG10A mutants were well captured by a previously described model[164] in which an initial cooperative binding event modulated by nearest neighbor interactions between Car9 segments (and to a lesser extent fused passengers) is followed by a slow isomerization step that we ascribe to a conformational rearrangement at the silica interface (**Figure 3.2A-B**). The rate constant for the initial adsorption step can be written:

$$k_{a1} = \frac{N_A}{\sigma_{max}}\sqrt{\frac{RT}{2\pi M}}\,exp\left(\frac{-E_0}{RT}\right)exp\left(\frac{-E_i\theta}{RT}\right) = k'_{a1}exp(u\theta) \tag{3.1}$$

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 isomerization and desorption steps by simultaneously fitting all three sensorgram curves for both the adsorption and wash phases of the SPR experiments. **Table 3.1** 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 (3.1) 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. 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 3.1**).

For the K8AK11A and R4QR12Q variants, adsorption kinetics were well captured by a rate-law based Langmuir model[164] (**Figure 3.2C-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 3.1**). 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 silanol in the high-affinity binding of the R5 peptide,[33,34] synthetic LK peptides,[165,166] and certain combinatorially-selected peptides[54-56] to silica. Nevertheless, the binding of the K8AK11A and R4QR12Q mutants remained remarkably strong with ΔG values that were comparable and only about 25% lower than the wild type (**Table 3.1**). In short, a broad range of kinetic binding behaviors resides within a small band of free energies of binding.

3.3.3 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 B.1). **Figure 3.3** 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 simulations, while P9AG10A adopted a mix of folded and extended conformations (**Figure 3.3** and Figure B.4). 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 3.3**). 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), and further indicate that conversion of basic amino acids to uncharged residues can have either a structurally stabilizing (K8AK11A) or destabilizing (R4QR12Q) effect.

3.3.4 Role of electrostatic contacts

Consistent with the fact that sfGFP-Car9 has the largest k'_{a1} 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 3.4A-B**). 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 3.1D** for all four variants, averaging the distance of each residue's center or mass from the silica surface over the last 50 ns of the simulation for all three trials, and using the associated standard deviations as a proxy for side change fluctuations (**Figure 3.4C**). This analysis revealed that : (i) binders that follow Langmuir kinetics are characterized by having only 2 highly stable interactions of a side chain with the surface while sigmoidal binders have 3 or more such stable interactions; (ii) R4 is strongly stabilizing for sigmoidal binders; and (iii) uncharged residues S2 and (F/A)6 show notably closer stable interactions with the surface for sigmoidal binders compared to the Langmuir counterparts. As with the 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.

3.3.5 Discussion

Biomineralizing 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 or synthetic substrates.[8,12,104,135] 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-function relationship.

Here, we used protein engineering, SPR quantification of adsorption kinetics and MD simulations initiated from Rosetta predictions to gain such an understanding for the Car9 SBP. Our results indicated that Car9 has a high affinity for 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 3.1**). Although MD predictions 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 can lead to a high probability of Car9-Car9 cooperative interactions, likely driven by an optimization of electrostatic interactions between vicinal SBPs and 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 3.1**). Such rearrangements have been documents for large proteins adsorbed to silica[167,168] and likely include contributions from the fused framework interacting with the interface.

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 absorption, and therefore lower the initial adsorption rate (**Table 3.1**). Yet, surface occupancy remains high and the mechanism of cooperativity proposed above is 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 3.2** and **3.4**). 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 propose that the reduced level of surface interactions and occupancy explains the lack of SBP-SBP interaction that characterize sigmoidal binding and explains the decrease in desorption rates (**Figure 3.3**).

3.4 Conclusions

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. Thus, the *de novo* design of a 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 entropic binders whose interactions with interfaces are not driven by dominant surface-

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

3.5 Figures

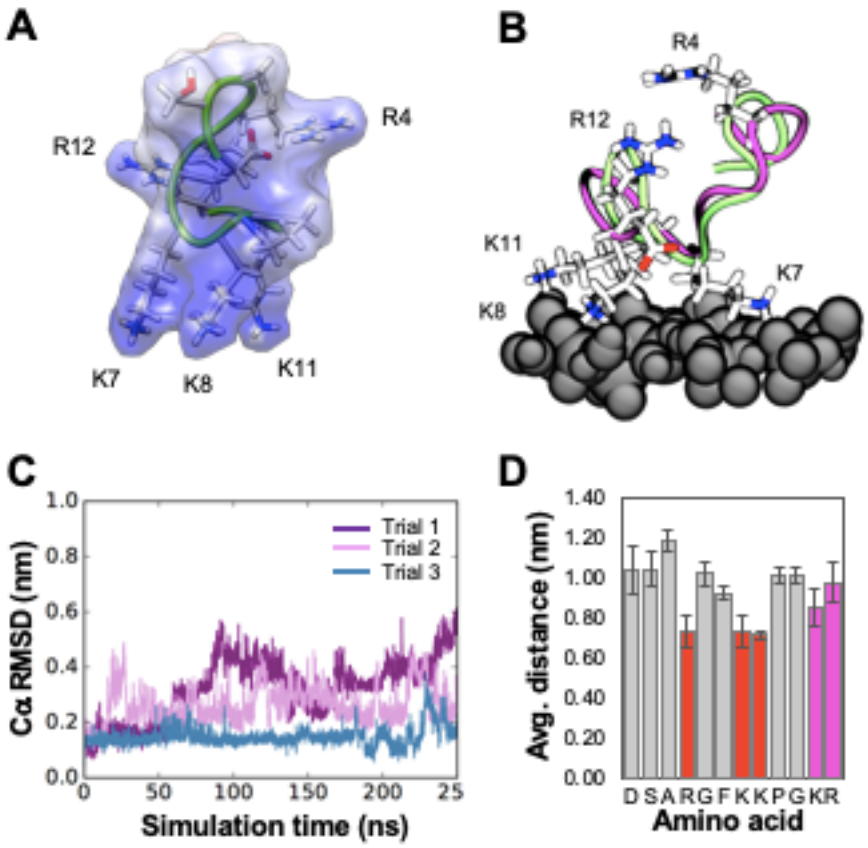


Figure 3.1 Predicted structure of the Car9 dodecapeptide at the silica interface. **(A)** Low energy structure proposed by Rosetta. The electrostatic surface is colored between -10 kTe (red, not visible) and +10 kTe (blue). **(B)** Superimposition of Rosetta structure (green) and top-ranked cluster determined from trial 1 of MD-simulated structure at the silica interface (pink). **(C)** Root mean square deviation (RMSD) traces of C∝ atoms 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.

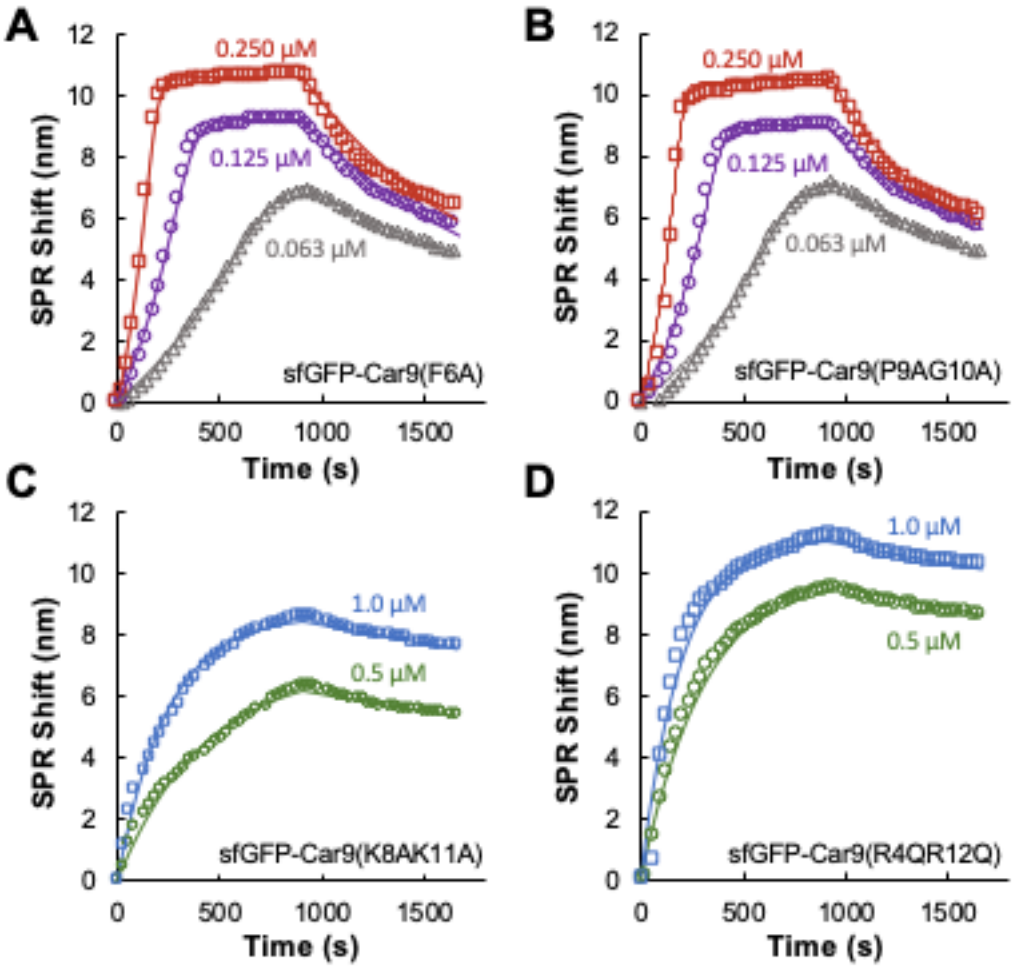


Figure 3.2 Adsorption kinetics of sfGFP-Car9 variants on silica-coated SPR chips. Sensorgrams were collected at the indicated concentrations of **(A)** sfGFP-Car9(F6A), **(B)** sfGFP-Car9(P9AG10A), **(C)** sfGFP-Car9(K8AK11A), or **(D)** sfGFP-Car9(R4QR12Q). Solid lines correspond to data fits obtained 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.

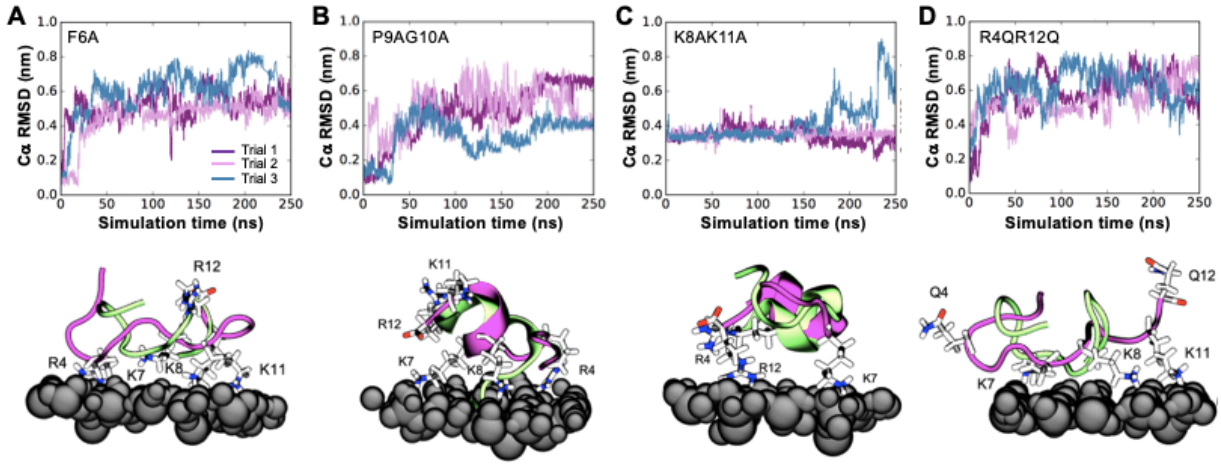


Figure 3.3 Root mean square deviation (RMSD) traces of C∝ atoms for three trials of the F6A **(A)**, P9AG10A **(B)**, K8AK11A **(C)**, and R4QR12Q **(D)** Car9 variants adsorbed to silica. Lower panels show the superimposition of Rosetta initial structures (green) and top-ranked clusters determined from one of the trials of the MD-simulated structure at the silica interface (pink).

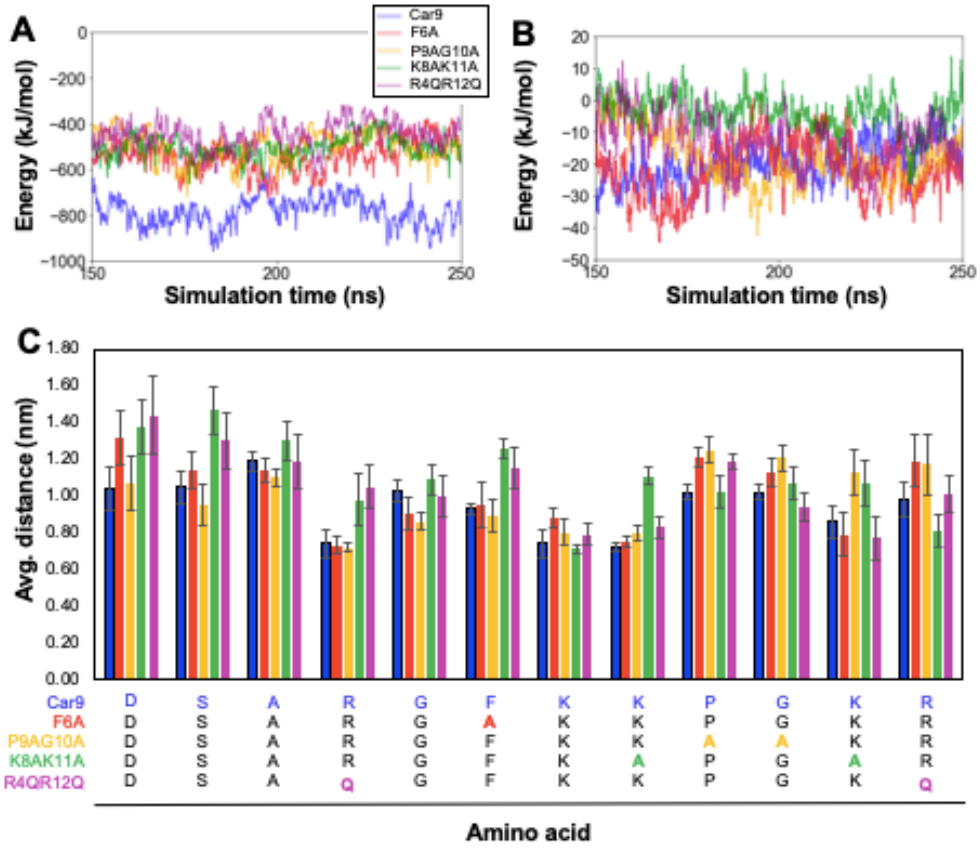


Figure 3.4 Decomposition of protein-surface potential energy into short-range components of the Coulombic **(A)** and Lennard-Jones **(B)** energies. **(C)** Average distance of the side chain center of mass of indicated residues to the silica surface (error bars are 1 standard deviation). The data are averaged across three trials over the last 50 ns.

3.6 Tables

Table 3.1 Kinetic and thermodynamic parameters describing the adsorption of sfGFP-Car9 and indicated variants to silica-coated SPR chips.

protein[a]	u	k'_{a1} or k_a ($M^{-1}s^{-1}$)	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 ($kJ.mol^{-1}$)	ΔG_2 ($kJ.mol^{-1}$)	ΔG ($kJ.mol^{-1}$)
sfGFP-Car9[b]	2.9 ± 0.1	20700 ± 1100	0.76 ± 0.02	0.17 ± 0.06	0.03 ± 0.01	-48.9	-4.3	-53.1
F6A	2.8 ± 0.1	12400 ± 500	0.79 ± 0.20	0.16 ± 0.04	0.02 ± 0.01	-47.3	-5.2	-52.6
P9AG10A	3.2 ± 0.1	10700 ± 700	1.06 ± 0.5	0.27 ± 0.18	0.03 ± 0.01	-47.1	-5.4	-52.4
K8AK11A	NA[c]	4200 ± 300	0.18 ± 0.04	NA	NA	NA	NA	-42.0
R4QR12Q	NA	5500 ± 700	0.13 ± 0.02	NA	NA	NA	NA	-43.5

[a] Adsorption 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).
[b] Data for sfGFP-Car9 is from reference 31.
[c] NA not applicable

Chapter 4 Influence of the Car9 peptide in TiO_2 precipitation

4.1 Introduction

Previous research has shown that silaffins and the R5 peptide, originally identified for their silica affinity, have the ability to mineralize TiO_2 from TiBALDH.[37] Our group previously demonstrated that a protein with a fused Car9 peptide has an affinity for carbon and silica.[9,10] In Chapter 3 we looked closer at the Car9 peptide's silica affinity and its' structure to function relationship by designing unique Car9 mutants (either removing positively charged amino acids or mutating key structural amino acids). Using surface plasmon resonance experiments and molecular dynamic simulations based on Rosetta structural predictions it was determined that Car9's binding to silica is highly dependent on electrostatic interactions. This work uses the sfGFP-Car9 and the sfGFP-mutant Car9 proteins and the knowledge about the Car9 peptide's structure to function relationship from the SiO_2 binding study and applies it to TiO_2 precipitation. The goal is to understand how a peptide's structure and charge influences TiO_2 formation by studying the effects it has on the yield, morphology, and crystallinity of the TiO_2 precipitate. In addition, this work gives insight into the long-term goal of controlling the formation of composite materials using specific protein-peptide fusions.

4.2 Materials and Methods

4.2.1 DNA manipulations and protein purification

Expression plasmids for wild type sfGFP, sfGFP-Car9, and for sfGFP-Car9 variants containing the F6A, P9AG10A, K8AK11A and R4QR12Q mutations in their Car9 segment were described in Chapter 3. A plasmid encoding the K8HK11H variant of sfGFP-Car9 was constructed by hybridizing oligonucleotides encoding the mutant sequence, digesting the double-stranded DNA with *Hind*III and *Xho*I and ligating the resulting cassette in the same sites of pET-24a(+)-sfGFP-Car9.[10] Proteins were purified as described[116] and stored at -20°C.

4.2.2 TiO_2 precipitation reactions

Titania precipitation reactions were conducted at room temperature in 37.5 mM citrate buffer using 25 mM of TiBALDH (Sigma-Aldrich 388165) at protein concentrations, pH, and incubation times specified in the text or figure legends. In typical experiments, 11.8 μL of 1.7M TiBALDH was added to 600 μL of 50 mM citrate buffer, pH 5.0 in a 1.5 mL microfuge tube and mixed five times with a pipette tip. Protein was then added at a 10 μM final concentration to bring the reaction volume to 800 μL. The solution was agitated on an orbital shaker (Invitrogen Dynabeads sample mixer) at 80 rpm for 120 min for most experiments and the precipitate was recovered by centrifugation at 8,000*g* for 5 min.

The amount of protein incorporated into titania precipitates was quantified by monitoring the decrease of sfGFP fluorescence in reaction supernatants over time or after 120 min of incubation. For these experiments, supernatant samples (150 μL) were aliquoted in triplicate in the wells of a black 96 well microplate (Genesse Scientific), and fluorescence emission was recorded at 510 nm on a SpectraMax M5 microplate reader (MolecularDevices) with excitation at 475 nm and a cutoff wavelength of 495 nm. Because the sfGFP chromophore exhibits reduced fluorescence under acidic conditions due to a change in the charge state of the chromophore[169,170] the concentration of free protein in reaction supernatants was calculated using a calibration curve constructed with sfGFP-Car9 in 37.5 mM citrate, pH 5.0 (Figure C.1). Readings were normalized to the fluorescence of a protein solution in 37.5 mM citrate buffer, pH 5.0.

The amount of unreacted TiBALDH was determined using a quantitative Tiron colorimetric assay.[97] For these experiments, 280 μL of 5mM Tiron (Honeywell Fluka 89460) in citrate buffer at pH 5, was mixed with 20 μL of sample (10 μL of supernatant diluted with 10 μL of water) in a clear 96 well plate (Genesee Scientific). The mixture was incubated for 1 hour on

the benchtop at room temperature and the absorbance was measured at 380 nm on a plate reader. The concentration of unreacted TiBALDH was calculated using the calibration curve of Figure C.2.

4.2.3 Characterization of precipitated titania

Titania precipitates were collected by centrifugation at 8,000 g for 5 min, washed 3 times with 200 μL of deionized water and lyophilized for 1.5 hours. The mass of precipitated titania was measured at different protein concentrations using tared 1.7 mL microcentrifuge tubes. Thermogravimetric analysis (TGA) was conducted on a TA Q50 instrument to determine the ratio of protein to titanium dioxide in the precipitated product. Samples where heated in an aluminum oxide pan from 20°C to 900°C, at a heating rate of 20°Cmin^{-1} under nitrogen gas flow.

For fluorescence imaging, precipitates were resuspended in 200 μL of deionized water by mixing with a pipette and a 10 μL aliquot was deposited on a microscope slide and imaged using a Nikon Eclipse TE200-U microscope with excitation at 485 nm. For SEM imaging, precipitates were mixed with 1mL of bleach (Ultra Bleach 94% Bleach) and heated at 70°C for 10 min to remove bound proteins. Samples were centrifuged at 10,000g for 5 min, washed twice with 1 mL of bleach solution, and washed three times with 1 mL of deionized water. Samples were resuspended in 100 μL of deionized water and 3 μL was spotted on a silicon wafer. Micrographs were collected on a FEI Sirion XL30 SEM. For x-ray diffraction measurements, lyophilized precipitates were deposited on a silicon wafer and spectra were collected on a Bruker D8 Discover XRD system. The size of the nanocrystallites was calculated using the XRD peak at $2\theta = 25°$ and the Debye-Scherrer equation

$$D = \frac{0.9\lambda}{\beta \cos\theta} \tag{4.1}$$

with D being the particle diameter (nm), λ the x-ray wavelength equal to 0.154 nm, β the full width at half the maximum of the desired peak in radians, and θ the Bragg angle in degrees (12.64º). The distribution of crystalline phases was determined by comparing the ratio of signal intensities at 2θ = 28.7° (a TiO_2(*B*)-specific reflection) and 25° (corresponding to TiO_2 anatase (25.3°) and TiO_2(*B*) (24.99°) reflections which overlap in our experiment) to the ratio of intensities at the same 2θ values for the pure TiO_2(*B*) phase (0.9; PDF 00-305-0088).

For TEM imaging, samples resuspended in 200 μL of water were subjected to 5 cycles of sonication (1 min each followed by 1 min rest) using a Fisherbrand Model 120 Sonic Dismembrator with a 3.2 mm probe and 30% amplitude. Aliquots (10 μL) were deposited onto a C-flat holey carbon grid (Electron Microscopy Sciences) and excess solution was wicked away with filter paper. Samples were stored in a desiccator for at least 24h before imaging at on a FEI Tecnai F20 SEM operated at an acceleration voltage of 200 kV.

4.3 Results and discussion

4.3.1 The Car9 extension promotes TiO_2 precipitation under acidic conditions

Initial attempts to induce TiO_2 precipitation using 10 μM sfGFP-Car9 and a solution of TiBALDH according to published protocols[93,94,97] failed. This result was not surprising in light of the fact that previous studies used synthetic peptides at molar concentrations exceeding that of our silica-binding protein by 20- to 500-fold. We reasoned that medium acidification would favor protein-mediated TiO_2 precipitation by promoting the hydrolysis of bound lactate ligands. Mindful of the fact that phosphate buffers can cause spontaneous titanium phosphate precipitation from TiBALDH solutions,[38,97] we made use of a citrate buffer that allowed us to tune the pH in 0.2 units increments while introducing citrate ions that are capable of modifying the reaction equilibrium by competitively chelating Ti(IV).[171,172]

Indeed, when the pH of the solution was dialed to 4.8, both wild type sfGFP and sfGFP-Car9 promoted the formation of a precipitate. However, only sfGFP-Car9 induced inorganic precipitation at pH 5.0, and the fusion protein became a rather ineffective inducer of mineralization when the pH was raised to 5.2 (**Figure 4.1A**). This sharp dependency on pH, together with the fact that no precipitation was observed in the absence of protein over the 4.0-5.8 pH range, provided us with an opportunity to decouple the influence of the sfGFP framework from that of the Car9 extension. Thus, all subsequent experiments were conducted in citrate buffer held at pH 5.0.

The ease of sfGFP detection by fluorescence spectroscopy (adjusted for its chromophore's emission loss under acidic conditions,[169,170] (Figure C.1) enabled us to accurately monitor the incorporation of sfGFP-Car9 within the mineralized phase as a function of protein concentration. As reported for the R5 peptide[37] and combinatorially-selected solid binding peptides,[94] the depletion of protein from supernatant fractions was linearly related to precipitate formation (**Figure 4.1B**) and the reaction occurred at a rate of 0.88 μM or protein incorporated per μM of protein supplied under our experimental conditions. In agreement with this result, both the mass of precipitated material (**Figure 4.1C**) and the amount of TiBALDH consumed (**Figure 4.1D**) increased linearly with the amount of sfGFP-Car9 added, although we started to observe a plateau in the amount of protein incorporated and TiBALDH consumed at sfGFP-Car9 concentrations exceeding 17 μM.

An analysis of the reaction kinetics further revealed that both sfGFP-Car9 and the wild type control were very rapidly incorporated in the mineral phase (**Figure 4.1E**). This burst occurred nearly instantaneously upon protein addition, and on a much faster timescale than we could measure experimentally. As expected, the presence of the Car9 extension at the C-terminus of sfGFP caused a significant increase in the initial kinetics of mineralization with 40% more sfGFP-Car9 than wild type protein partitioning in the insoluble fraction in the first 5 minutes of

reaction. Subsequent kinetics were slower and could be fitted with an exponential model with a rate constant that was two-fold higher for sfGFP-Car9 than for sfGFP (k = 8.2 x 10^{-3} s^{-1} vs. 3.8 x 10^{-3} s^{-1}). Finally, while 87% of the sfGFP-Car9 originally supplied was found in association with the mineral phase upon completion of the reaction ($t \geq 90$ min), 40% of the sfGFP remained soluble (**Figure 4.1E**).

Taken together, the above results suggest that protein-induced precipitation of solutions of TiBALDH acidified with citric acid results from the displacement of a solution equilibrium in which lactate or exchanged citrate ligands complexed with Ti(IV) species are displaced by the added protein which adsorb to small clusters and promote titania polycondensation. Our experiments provide several lines of evidence for such a solution equilibrium modulated by solution pH (and therefore proton-assisted reactions) play a key role.

4.3.2 Characterization of the sfGFP-Car9 TiO_2 precipitate

The sfGFP-Car9 TiO_2 precipitates were characterized using microscopy and XRD analysis. The dense fluorescent spots in the fluorescent microscope image, excitation of $\lambda = 485$ nm, confirm the presence of protein in the TiO_2 precipitate (**Figure 4.2A-B**). The SEM micrograph shows that the morphology of the TiO_2 precipitate is rough and porous (**Figure 4.2C**). The XRD data (**Figure 4.2D**) has a high background noise, a signature of amorphous material, with crystalline material approximately 2nm in size. These results give significant insight into the mechanism occurring during the TiO_2 precipitation when sfGFP-Car9 is present. The rutile unit cell has 20% edge sharing octahedra, the anatase unit cell consists of approximately 50% edge sharing titania octahedra, and the TiO_2(*B*) unit cell has even more edge-sharing octahedra.[79-81] The XRD data for sfGFP-Car9 has definitive signatures of both anatase TiO_2 and monoclinic TiO_2(*B*), meaning that the protein-peptide fusion promotes edge sharing titania octahedra. TEM (**Figure 4.2E**) confirms the XRD data and that the formed TiO_2 is highly amorphous with crystallites in the 2nm size range.

4.3.3 Influence of Car9 mutations on TiO_2 precipitation

Using sfGFP-Car9 we were able to successfully reproduce the results seen for TiO_2 precipitation using the R5 peptide, and that sfGFP-Car9 successfully precipitates TiO_2 with anatase and TiO_2(*B*) crystal characteristics. Our previous silica study showed that Car9 structure and electrostatics have a significant role in silica binding; less charge and a more flexible structure displayed weaker binding characteristics. This work uses the previously made Car9 mutants to understand how their mutations influence TiO_2 precipitation. The previous silica study found that the structural mutations, F6A and P9AG10A, are enthalpic silica binders with the positively charged side chains of the electrostatic amino acids acting as anchor residues in SiO_2 binding. These mutants differ from Car9 by lacking the peptide rigidity (F6A) and the hairpin (P9AG10A) that orient the peptide in the most efficient binding position. The electrostatic mutations, K8AK11A and R4QR12Q, have Langmuir binding kinetics and a weaker affinity to the silica surface. One additional mutant added to this study is K8HK11H, which is expected to show the effects of the weaker positively charged amino acid histidine.

Fluorescence depletion measurements (**Figure 4.3A**) collected after a 2-hour reaction time shows the effect of each mutation on TiO_2 precipitation, the higher the fluorescence depletion means the more protein in the TiO_2 precipitate and the larger the TiO_2 yield overall. All mutants precipitated more TiO_2 when compared to the wild type sfGFP. As expected, the two electrostatic mutants, K8AK11A and R4QR12Q, had a decreased amount precipitated, approximately 70%, compared to the 80% for Car9, confirming the positively charged side chains assist in the formation of TiO_2. The interesting results were in the structural mutations P9AG10A, F6A, and the histidine mutation K8HK11H. A slight decrease in amount of protein incorporated, from 80% to 75%, was determined for the hairpin mutation P9AG10A, showing that the orientation of the positively charged residues is important mechanistically for the formation of TiO_2. The mutant that

precipitated the least amount of TiO_2 material, therefore having the least amount of protein incorporated, was the F6A. This is unique because the peptide is missing the hydrophobic core that makes the Car9 peptide rigid, proving that peptide structure and flexibility has a significant impact on the hydrolysis of TiBALDH and formation of TiO_2. The K8HK11H results were also unique, this mutation mineralized more TiO_2 compared to the Car9 peptide. Many factors could have an influence on this, histidine has an imidazole ring in its side chain, this ring is protonated at the reaction pH and adds a more rigid side chain for TiO_2 precipitation, which could potentially increase its' efficiency. Histidine also has the capability to bind to TiO_2 not only through electrostatic interactions, but also through hydrogen bonding of the amine group and pi orbital interactions, all factors that could promote the formation of TiO_2.[173]

The XRD results, **Figure 4.3B**, gave significant insight into the mechanism of the TiO_2 mineralization and the effect of two amino acid mutations on the crystal phases of the TiO_2 formed. Similar to sfGFP-Car9 all mutants had anatase and TiO_2(*B*) crystal phase characteristics. The differences were in peak intensities and peak ratios used to determine the percentage of anatase and TiO_2(*B*) for each sample, **Figure 4.3C**. The crystalline material formed using sfGFP-Car9 is composed of approximately 80% TiO_2(*B*). The structural mutations, F6A and P9AG10A, formed approximately 70% TiO_2(*B*), showing that the peptide structure does not have a significant impact on the crystal phase and the reaction still promotes edge sharing octahedra. The structural mutations seem to have a larger effect on the amount of TiO_2 material formed compared to the crystal phase. The electrostatic mutations formed lower amounts of TiO_2(*B*) compared to Car9, K8AK11A approximately 60% and R4QR12Q approximately 40% demonstrating that the amine side chains of lysine and arginine promote TiO_2(*B*) crystal phase formation and mutating these affects both TiO_2 yield and the crystal phase formation. The histidine mutant, K8HK11H, formed

approximately 80% TiO_2(*B*), therefore concluding that the imidazole side chain in histidine promotes TiO_2(*B*) formation similar to arginine and lysine but has a larger influence on TiO_2 yield.

4.4 Conclusions

This work demonstrates that the Car9 peptide and its' mutants react with the destabilized TiBALDH forming a TiO_2/protein precipitate. The material is amorphous with some nanocrystalline structures in the 2 nm size range with both anatase and TiO_2(*B*) crystal phase characteristics. Mutations in the Car9 peptide show that simple two amino acid changes in a large protein-peptide fusion can influence the TiO_2 precipitation reaction, whether it is in terms of the yield or crystallinity. The highest TiO_2 yields required the positively charged amino acids, when mutated and the peptide structure was changed the yield decreased. TiO_2 mineralization is more sensitive to peptide structure compared to SiO_2 binding. Electrostatics not only affected the efficiency of mineralization, but also the nanocrystallinity. Positively charged amino acids promoted the formation of the more symmetric TiO_2(*B*) over the more common anatase, demonstrating how we can potentially bias crystallinity with just two amino acid mutations in a given peptide. This will be beneficial knowledge for our long-term goal of designing unique complex inorganic materials with specific functions.

4.5 Figures

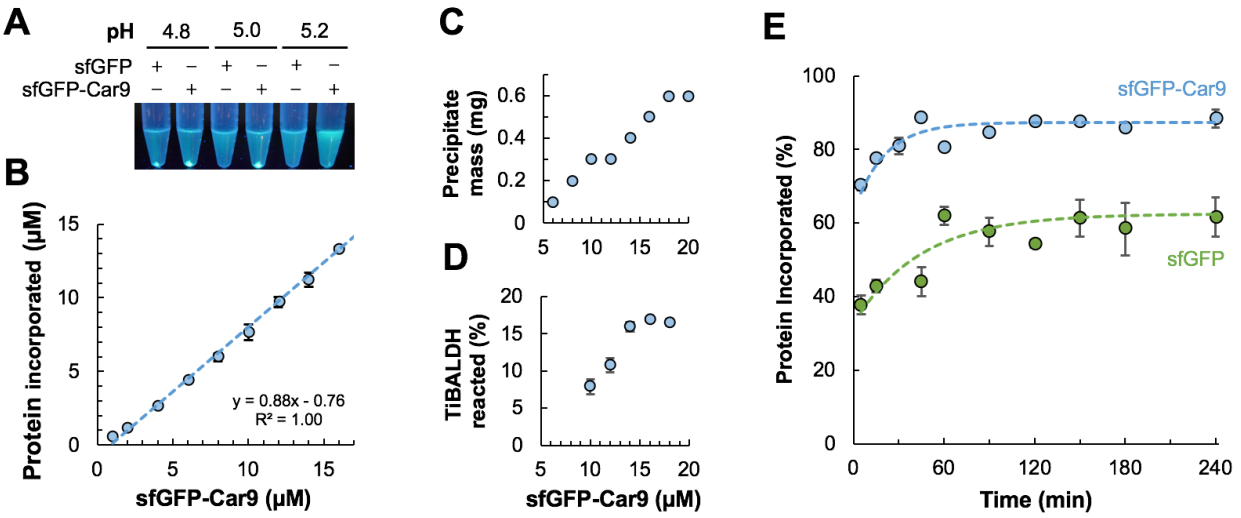


Figure 4.1 Precipitation of TiO_2 by sfGFP-Car9 and wild type sfGFP. **(A)** UV photograph of supernatant solutions and titania pellets obtained after centrifugation of TiBALDH solutions in citrate buffer supplemented with 10 µM of sfGFP or sfGFP-Car9. Solution pH are indicated. **(B)** Relationship between the concentration of sfGFP-Car9 used to induce titania precipitation and the incorporation of the protein into insoluble material. Relationships between the concentration of sfGFP-Car9 added and **(C)** the mass of the precipitated titania, or **(D)** the amount of unreacted Ti(IV). **(E)** Biphasic kinetics of sfGFP or sfGFP-Car9 (10 µM) incorporation into precipitated titania.

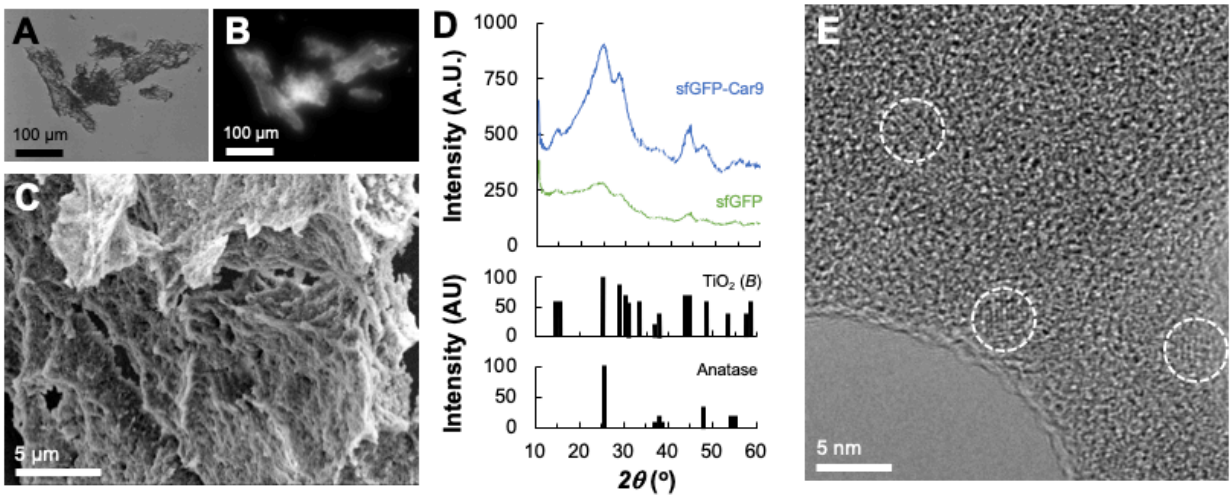


Figure 4.2 Characterization of the TiO_2 precipitate formed from sfGFP-Car9. Fluorescent microscope images at **(A)** phase contrast and **(B)** excitation λ_{ex}= 485 nm formed with 10 µM sfGFP-Car9 at pH 5. **(C)** SEM micrograph of TiO_2 precipitate formed with 5 µM sfGFP-Car9 at pH 5. **(D)** XRD analysis for a TiO_2 precipitate from 10 µM sfGFP-Car9 and 10 µM sfGFP at pH 5 having signature diffraction peaks for anatase and TiO_2(*B*). **(E)** TEM analysis of TiO_2 precipitate formed with 10 µM sfGFP-Car9 at pH 5.

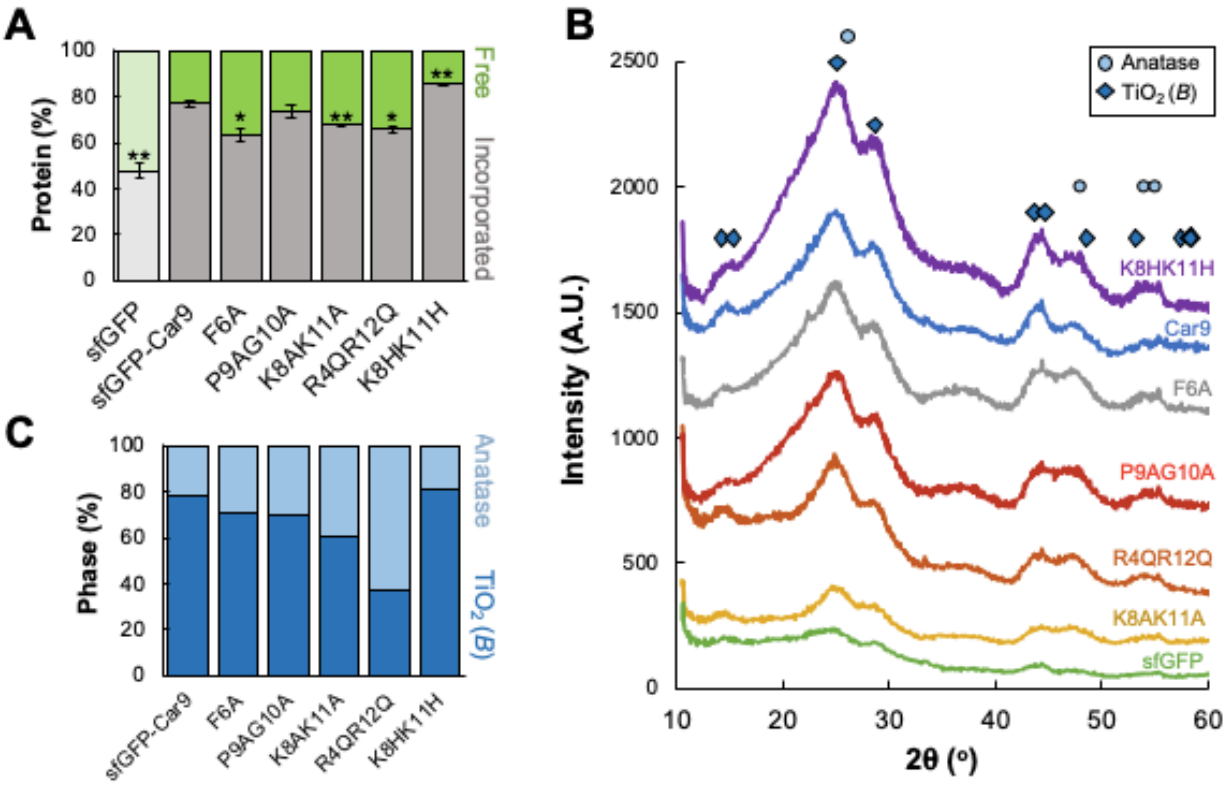


Figure 4.3 (A) TiO_2 precipitation reactions using 10 µM protein, 25 mM TiBALDH, 37.5 mM citrate buffer pH 5 and a reaction time of 2 hours with agitation at 80 rpm. Protein fluorescence depletion measurements with the calculated amount of free and incorporated protein percentage with a p-test (* $P < 0.05$ and ** $P < 0.01$) compared to sfGFP-Car9. **(B)** XRD profiles at room temperature for sfGFP, sfGFP-Car9 and Car9 mutants, with the signature anatase and TiO_2(*B*) peaks labeled. **(C)** Calculated percentages of anatase and TiO2(*B*) crystal phases in each sample. Calculations were done using the intensities of raw data compared to intensities of the signature crystal data.

Chapter 5 Diffusion-controlled TiO_2 precipitation

5.1 Introduction

Previously it was shown that TiO_2 can be precipitated using sfGFP-Car9 in an equilibrium process. This mineralization produced a mixture of anatase and TiO_2(B) using a titanium precursor, TiBALDH, mixed with citrate buffer at pH 5. When the Car9 tag was mutated, varying the structure or charge of the fused peptide, the morphology and crystallinity of the TiO_2 precipitate changed. Another approach to this method of protein induced TiO_2 precipitation is using a non-equilibrium reaction-diffusion process, taking advantage of the concentration gradient formed in a diffusion process and determining how this influences the TiO_2 morphology or crystallinity. The long-term goal would be to produce specifically designed graded materials with a controlled structure or crystallinity. The work discussed below demonstrates initial proof of concept experiments for this reaction-diffusion controlled TiO_2 precipitation.

5.2 Materials and Methods

5.2.1 DNA manipulation and protein purification

Oligonucleotides encoding mutant Car9 sequences were purchased from Integrated DNA Technologies, hybridized, digested with *Hind*III and *Xho*I and the resulting cassettes were inserted into pET-24a(+)-sfGFP-Car9 digested with the same enzymes. The presence of correct mutations was verified by DNA sequencing. BL21(DE3) cells containing wild type or mutant versions of sfGFP-Car9 mutants were grown to mid-exponential phase (A_{600}~ 0.5) at 37°C in 500 mL of LB medium supplemented with 50 μg/mL of kanamycin. Protein expression was induced by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) and cells were harvested by centrifugation at 6,000g for 10 min after 4 h of incubation at the same temperature. Cell paste was resuspended in 35 mL of 20 mM Tris-HCl, pH 7.5 supplemented with 2 mM EDTA and 1 mM PMSF and lysed using 6 cycles of sonication for 3 min at 30% duty power using a Branson Sonifier. Insoluble

material was removed by centrifugation at 10,000*g* for 15 min, and clarified extracts were incubated at 70°C for 10 minutes to cause thermolabile proteins to aggregate. Insoluble material was removed by centrifugation as above, and proteins were purified by silica-affinity chromatography using 5g of silica gel (35–60 mesh, 15 nm pore size, Sigma Aldrich) equilibrated in 20 mM Tris-HCl pH 7.5.[10,116] Proteins were eluted in 20 mM Tris-HCl supplemented with 0.5 M L-lysine pH 8.25 and subjected to dialysis against 3L of 20 mM Tris-HCl pH 7.5 with 3 cycles of buffer exchange. Protein concentrations were assayed using the BCA assay and molecular masses verified using mass spectrometry. Samples were stored at -20°C.

5.2.2 TiO_2 diffusion in agarose

For the diffusion experiments 4 mL of 1% agarose was poured into each round plate, with two stoppers approximately 0.5 cm in diameter spaced about 0.5 cm apart to form the wells, **Figure 5.1**. The agarose was allowed to solidify for 1 hour. At t = 0 hours 100 µL of 50 µM protein was added to one well. TiBALDH was mixed with citrate buffer pH 5 to give a resulting concentration of 25 mM TiBALDH and 37.5 mM citrate buffer pH 5. Depending on the experiment, short diffusion (t = 2hrs) or long diffusion (t = 20hrs), 100 µL of the TiBALDH citrate buffer mixture was added to the second well and allowed to diffuse and react with the protein. Each diffusion was done in triplicates.

5.2.3 Characterization by SEM

For SEM characterization, the TiO_2 precipitate was removed from the agarose. The sample was cut from the agarose and mixed with 1 mL bleach and heated for 30 minutes at 65ºC. After heating, the sample was centrifuged at 10,000g for 5 minutes and the bleach supernatant was removed. The sample was washed 3 times with 1 mL of dH_2O each time and centrifuging at 10,000g for 5 minutes in-between each wash. After the first wash the samples were mixed with 1mL of bleach and centrifuged at 10,000g for 5 minutes to remove any residual agarose. The

samples were washed again, 3 times with 1 mL dH_2O and centrifugation at 10,000g for 5 minutes. After the final centrifuging and removal of the supernatant the samples were resuspended in 100 µL of dH_2O. For SEM imaging 3 µL of the resuspended sample was spotted on a silica wafer and allowed to dry overnight. Samples were imaged on a FEI Sirion XL30 SEM.

5.3 Results and discussion

For diffusion controlled TiO_2 precipitation the effects of agarose on the reaction and the proper reaction conditions were determined. From initial diffusion in agarose experiments the TiBALDH/citrate buffer (precursor mixture) diffused approximately twice as fast as sfGFP-Car9. For proof of concept reaction-diffusion experiments sfGFP-Car9 was deposited in one well and diffused for 2 hours before the addition of the precursor mixture in the opposite well. **Figure 5.2** shows that the precursor mixture front reached the sfGFP-Car9 diffusion front after two hours. After 24 hours, a distinct TiO_2 precipitation zone was observed where the fronts met. This initial experiment demonstrates that the reaction diffusion process can be used for non-equilibrium synthesis of TiO_2.

Three Car9 mutants (F6A, K8AK11A, and K8HK11H) reaction-diffusion precipitation and protein concentration effects were studied by changing the diffusion time of the proteins before the addition of the precursor mixture. One short diffusion experiment (t = 2 hours protein diffusion before addition (**Figure 5.3**)) and a long diffusion experiment (t = 20 hours protein diffusion before addition (**Figure 5.4**)) were done. Both **Figure 5.3** and **Figure 5.4** show that the K8HK11H and wild type Car9 precipitate more material compared to F6A and K8AK11A. This correlates with the results from the equilibrium TiO_2 precipitation experiments from Chapter 4. **Figure 5.3** and **Figure 5.4** also demonstrate the differences between protein concentration. The density of the precipitate changes with the protein diffusion time. For the short diffusion, the TiO_2 precipitate is dense where the two fronts meet. The long diffusion produced a more visible gradient of

precipitation due to the concentration gradient of the protein diffusing for 20 hours before precursor addition. Both of these methods can be used to our advantage in designing specific controlled graded materials.

To give more insight into the TiO_2 produced in the reaction-diffusion experiments the precipitates were removed from the short diffusion experiments for Car9 and characterized using SEM (**Figure 5.5**). Two different morphologies were observed at the same magnification for the TiO_2 precipitate. One morphology was a connected thin structure (**Figure 5.5A**) and the other was single unconnected particles (**Figure 5.5B**). These results potentially mean that the concentration gradients of the protein influence the morphology of the precipitate formed.

5.4 Conclusions

These initial proof of concept experiments prove that protein induced TiO_2 mineralization can be done in a reaction-diffusion environment in 1% agarose. This shows that reaction-diffusion experiments can form different morphologies in TiO_2 precipitation. The next steps would be to produce more material with a finer control and developing a model to have a method of predicting how this material will form in varying locations.

5.5 Figures

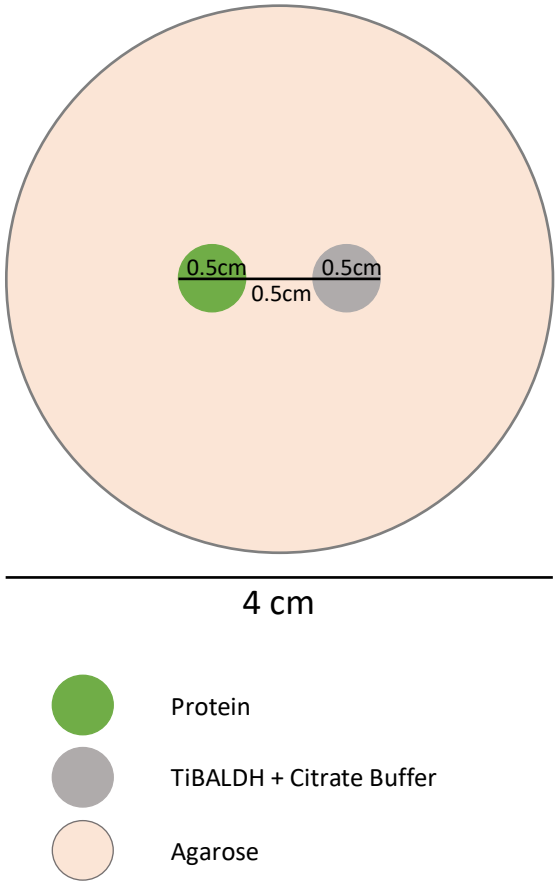


Figure 5.1 The non-equilibrium reaction-diffusion set up for the formation of TiO_2 from a protein and titanium precursor.

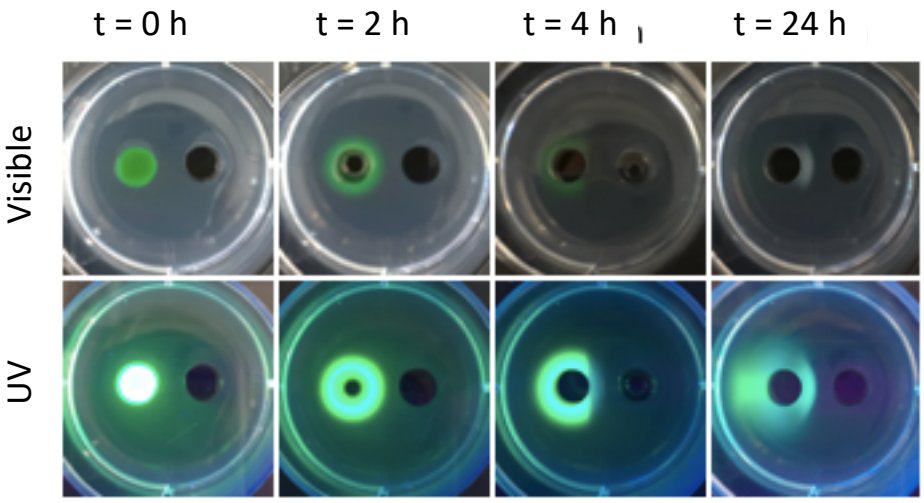


Figure 5.2 Non-equilibrium synthesis of TiO_2 using 50 μM sfGFP-Car9 and reaction-diffusion in 1% agarose. Addition of 100 μL of 25 mM TiBALDH in 37.5 mM citrate buffer pH 5 at 2 hours of protein diffusion.

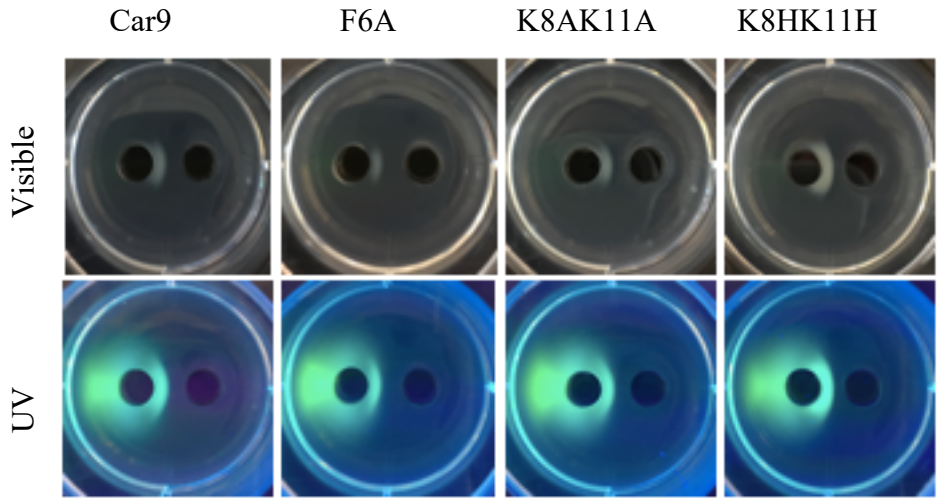


Figure 5.3 Non-equilibrium synthesis of TiO_2 using short reaction-diffusion in 1% agarose and 50 μM of protein. Showing the differences between the different sfGFP tagged proteins (Car9, F6A, K8AK11A, and K8HK11H) after a 24-hour reaction time. Addition of 100 μL of 25 mM TiBALDH in 37.5 mM citrate buffer pH 5 at 2 hours of protein diffusion.

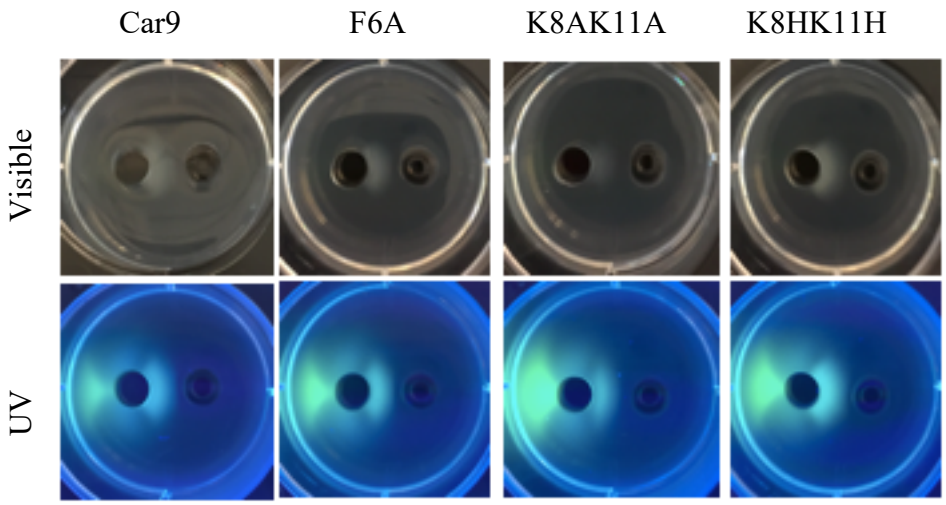


Figure 5.4 Non-equilibrium synthesis of TiO_2 using long reaction-diffusion in 1% agarose and 50 μM of protein. Showing the differences between the different sfGFP tagged proteins (Car9, F6A, K8AK11A, and K8HK11H) after a 40-hour reaction time. Addition of 100 μL of 25 mM TiBALDH in 37.5 mM citrate buffer pH 5 at 20 hours of protein diffusion.

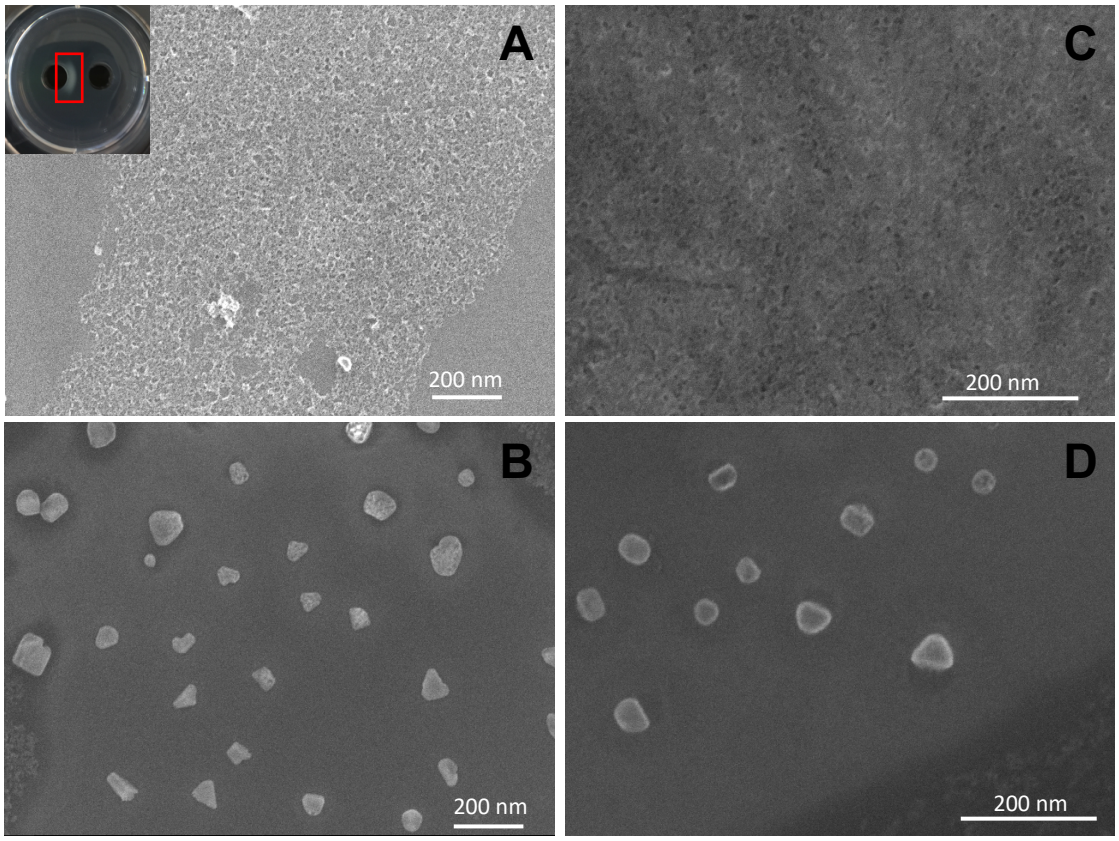


Figure 5.5 SEM photos from short diffusion experiments using 50 µM sfGFP-Car9 **(A)**, **(B)** and F6A **(C), (D)**. Proteins diffused for 2 hours before the addition of the TiBALDH/citrate buffer mixture and a reaction time of 24 hours.

Chapter 6 Influence of multiple Car9 peptides and location in sfGFP on SiO_2 binding affinity

6.1 Introduction

Previously we determined how Car9 peptide's structure and sequence affected its binding to silica and TiO_2 precipitation. In the future, we would like to make complex composite materials using proteins with fused SBPs as the binding agents to inorganic material. One approach is constructing a single protein with multiple peptide tags with the binding affinity for various inorganic materials. This specific work discusses the construction of varying sfGFP proteins that have individual or multiple Car9 tags inserted at different locations in the sfGFP framework and how these variations influence SiO_2 binding affinity.

6.2 Materials and Methods

6.2.1 DNA manipulation and protein purification

Genes for the proteins were inserted in a pET-24a (+) plasmid, between restriction enzyme sites *NdeI* at the N-terminus of sfGFP, *HindIII* at the C-terminus of sfGFP, and *XhoI* at the C-terminus of Car9. The Car9 tag was inserted into the loop between *BamHI* and *SpeI* digestive sites. BL21(DE3) cells containing sfGFP and the Car9 variants were grown to mid-exponential phase ($A_{600} \sim 0.5$) at 37°C in 500 mL of LB medium supplemented with 50 μg/mL of kanamycin. Protein expression was induced by addition of 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) and cells were harvested by centrifugation at 6,000g for 10 minutes after 4 h of incubation at the same temperature. All proteins with the Car9 peptide where purified using silica affinity chromatography, described previously,[10,116] and an additional 10-minute heat shock step at 70°C, since sfGFP is a thermostable protein. The control protein sfGFP was purified by fast liquid chromatography (FPLC) using ion exchange chromatography with a Whatman DE52 resin. The column was washed with 20mM Tris-HCl at pH 8 and the protein was loaded and washed with the same buffer. The washed protein was eluted using a gradient elution with 20mM Tris-HCl with

1M NaCl at pH 8. Most of the protein eluted at 85mM NaCl. Protein concentrations were determined using a BCA assay. Samples were stored at -20°C.

6.2.2 Protein expression and solubility study

Protein expression and solubility was determined using 25 mL cultures. Eight 25 mL cultures were inoculated with 0.5 mL of an overnight culture of BL21(DE3) cells containing the desired plasmids with the desired protein gene inserted. Cultures were grown at 37°C, protein expression was induced at $A_{600} = 0.5$. After 4 hours of expression 3 mL of each sample was centrifuged at 6,000 g for 10 minutes to be used in the solubility study. The soluble and insoluble fractions were separated using a French press. The protein in the soluble fraction was separated from other soluble material using chloroform-methanol extraction.

6.2.3 SPR experiments

Silica-coated surface plasmon resonance (SPR) chips were fabricated in house using a glass substrate coated with a 2 nm titanium adhesion layer, a 48 nm evaporated gold film, and a 4 nm silicon film deposited by plasma enhanced chemical vapor deposition (PEVCD). Chips were cleaned by ethanol and UV-ozone treatment before use and mounted on the flow cell of a SPR sensor from the Institute of Photonics and Electronics (Prague, Czech Republic). All experiments were conducted at a flow rate 70 $\mu L min^{-1}$ and a constant temperature of 25°C with protein concentrations at 0.1 $\mu mol\ L^{-1}$ in 20 mM Tris-HCl at pH 7.5.

6.3 Results and Discussion

6.3.1 Multi-tag influences on protein expression and purification

To understand the influence that multiple cationic Car9 peptides and their location have on silica binding sfGFP proteins with the Car9 peptide inserted in various locations were constructed and purified. **Figure 6.1A** shows the β-barrel sfGFP protein where the Car9 tag permutations are located. In total, there are 8 constructs, one being the control sfGFP. There are three versions of

single tagged sfGFP proteins, with a Car9 tag fused to the N-terminus (Car9-sfGFP) by a TEV protease cleavage site and a GGGS linker, in a permissible loop on the protein (sfGFP::Car9), and fused to the C-terminus (sfGFP-Car9) with a KLGGGS linker, **Figure 6.1B**. Car9-sfGFP and sfGFP-Car9 were constructed to test the location of the sequence of the tag and sfGFP::Car9 to test the influence of a constrained peptide. The three dual tagged proteins are: one with the N-terminal and in the loop Car9 tag (Car9-sfGFP::Car9), N-terminal and C-terminal Car9 tag (Car9-sfGFP-Car9), and in the loop and C-terminal Car9 tag (sfGFP::Car9-Car9). The final permutation of the tagged proteins is a tri tagged sfGFP protein (Car9-sfGFP::Car9-Car9) with Car9 at all three locations.

After protein construction the effects of multiple tags on protein expression, **Figure 6.2A**, and solubility, **Figure 6.2B**, were determined. The protein expression does not change by the addition of single or multiple Car9 tags, or the location of the Car9 tag in the sfGFP framework. When comparing the soluble and insoluble fraction of the expressed cells, it is apparent that the addition of the Car9 peptide increases the insolubility. It is unclear the difference in solubility for a single, dual, or triple tagged protein and a decent amount of protein can still be obtained from purification of the soluble fraction. All of the multivalent Car9 proteins where purified on a silica affinity column, with a relatively similar purity, **Figure 6.3**. One observation is that the single tagged proteins had the highest yield and adding multiple tags decreased the concentration of protein released from the silica column. This effect is most likely due to the strength of silica binding of the multi-tag proteins and the ability of being released from silica with 0.5 M Lysine elution buffer in 20 mM Tris-HCl at pH 8.25.

6.3.2 Influence of peptide location in sfGFP on SiO_2 affinity

SPR sensorgrams for the single tagged proteins, **Figure 6.4**, showed a significant difference in the adsorption phase when comparing the C-terminal Car9, sfGFP-Car9, to the loop

(sfGFP::Car9) and N-terminal (Car9-sfGFP). All three proteins displayed the similar sigmoidal cooperative two step adsorption with similar desorption rates. The significant difference was in the surface coverage and the rate of adsorption. The proteins sfGFP::Car9 and Car9-sfGFP had a faster on rate, but only covered about 70% of the surface that sfGFP-Car9 did, potentially due to the tag location and mode of adsorption. For sfGFP::Car9 the tag is located in a permissive loop, making it more rigid, and not nearly as flexible when binding to SiO_2 as the N- or C- terminal tags. This will also influence how the sfGFP backbone is positioned at the SiO_2 interface. The positioning of the backbone for sfGFP::Car9, **Figure 6.4**, could block open binding sites or cause electrostatic repulsion with other sfGFP::Car9 proteins therefore decreasing the overall surface coverage compared to Car9-sfGFP or sfGFP-Car9.

For the N-terminal tag, Car9-sfGFP, the increased adsorption rate and the decreased surface coverage could be due to the increased linker length, as well as the structure of the amino acids introduced to the SiO_2 surface. Previously in Chapter 3 we demonstrated that the Car9 tag on the C-terminus has tripod like structure, with the three lysine amino acid side chains protruding from the peptide and interacting with the silica surface, which is the optimal Car9 structure for electrostatic interactions with SiO_2 surface. When the Car9 tag is on the N-terminus it has a negatively charged aspartic acid at the N-terminus and the lysines are closer to the linker and are most likely not in the optimal structure for binding. This could cause a different mechanism of binding that is not as dependent on electrostatics changing the on rate and surface coverage. The decrease in surface coverage could also be due to the linker length, once the Car9 peptide is bound the longer linker can cause the sfGFP framework to have more movement near the surface that could potentially block open sites or repel other Car9-sfGFP proteins, decreasing the overall surface coverage.

6.3.3 Multi-tag influence on SiO_2 affinity

Three of the four multi-tag proteins had a significantly different mode of adsorption compared to the single tagged proteins, **Figure 6.4**. The proteins sfGFP::Car9-Car9, Car9-sfGFP-Car9, and Car9-sfGFP::Car9-Car9 all had a significantly lower rate of adsorption, so much lower that the adsorption time was increased from 15 minutes to 20 minutes in order to reach equilibrium surface coverage during the binding time. For these proteins the slower adsorption rate could be due to the Car9 peptide being on the same side of sfGFP. Car9 is a cationic peptide and when two are located near each other the electrostatic repulsion could interfere with the optimal binding structure causing the binding rate to decrease. This also could be the reason Car9-sfGFP::Car9 is the only multitag protein with a similar adsorption mechanism as sfGFP-Car9. This mutant has the Car9 on opposite sides of the protein so the peptides on the same protein most likely don't interfere with each other when binding to SiO_2. One interesting observation from the SPR data is that once the multi tagged proteins bind to SiO_2 they are not released from the surface and their desorption rates significantly decrease compared to the single tagged versions. This stronger affinity would explain why the yield decreases for multi tagged proteins in purification.

6.4 Conclusions and future direction

This work discusses initial steps in multi tag labeling of sfGFP with a Car9 peptide. Protein expression, solubility, and silica affinity were studied to understand the effects of multiple tags at various locations in the sfGFP framework. It was determined that the lower purification yields for multitag proteins were due to the proteins not being released efficiently from the silica with the elution buffer, 0.5 M Lysine in 20 mM Tris-HCl at pH 8.25. This lower yield problem could be addressed by doing an experiment varying concentration of lysine and/or the pH of the elution buffer to determine the optimal release conditions for the multitag proteins.

These initial experiments show the differences for the multitag proteins and help understand the next steps in order to fully understand the templating abilities of multitagged

proteins. Some further work would be to confirm how the number of tags and tag placement influence the binding by determining exactly the influence of increased flexibility with a longer linker, the rigidity of the peptide in the loop, or the optimal binding position depending on location. This could all be done, by more SPR experiments, confirming no mass transfer limitations, or with more advanced molecular dynamic simulations studying the protein-peptide fusion binding to SiO_2. Another approach that might be worth looking into is the influence on TiO_2 precipitation for these various protein-peptide fusions, and how the changes with the Car9 in the sfGFP framework influence the crystallinity or the morphology.

6.5 Figures

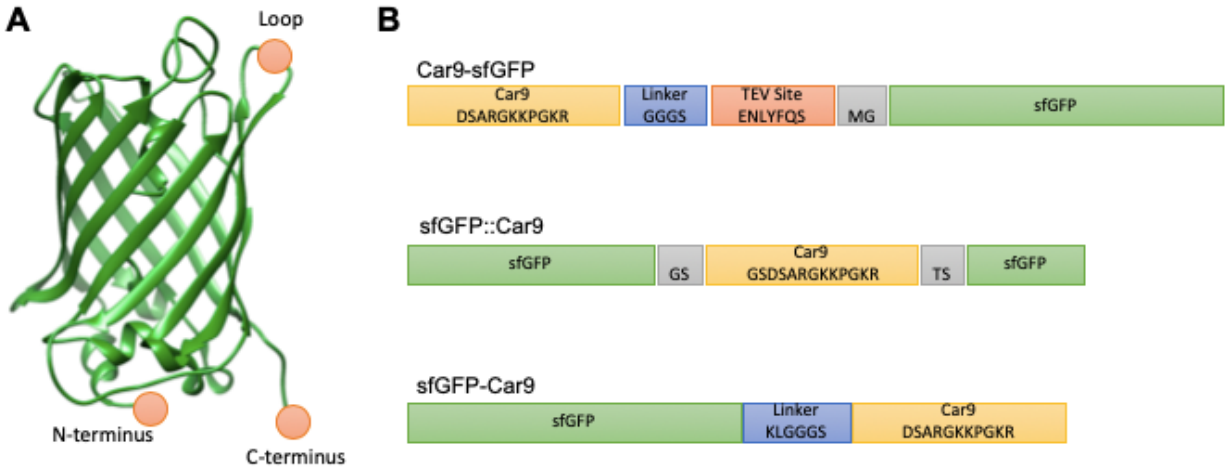


Figure 6.1 (A) The sfGFP protein with the three possible locations for the Car9 peptide (orange circle). **(B)** The added amino acids and linkers (single letter code) for the addition of the Car9 peptide on the N-terminus (Car9-sfGFP), in the loop (sfGFP::Car9), and on the C-terminus (sfGFP-Car9). For the multiple tagged proteins, the Car9 peptide is in the same location with the same sequences, there is just more than one.

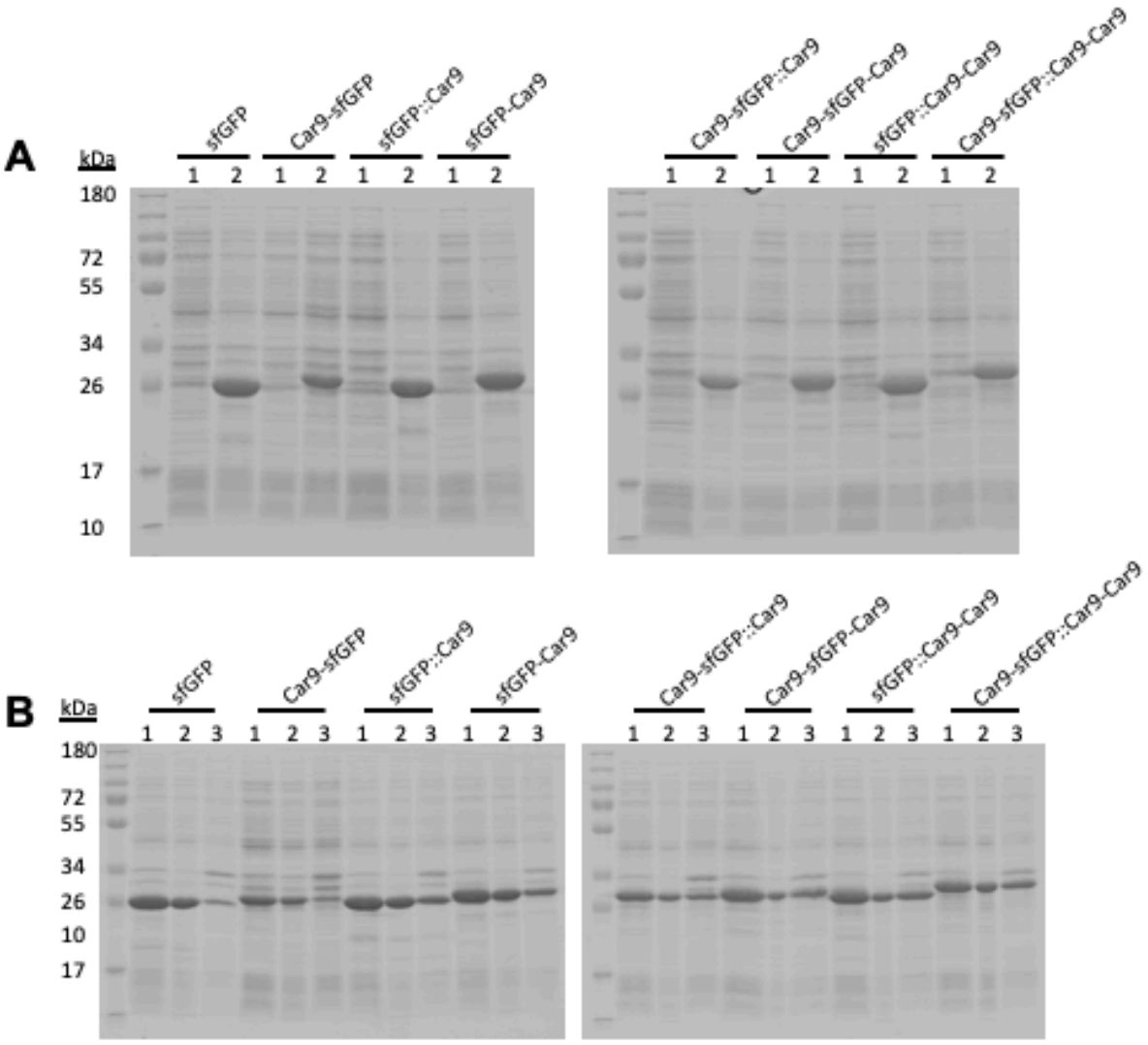


Figure 6.2 (A) SDS-PAGE gel displaying the effect of multiple Car9 tags and the placement of the Car9 tag on protein expression. (1) Pre induction (2) Post induction. **(B)** SDS-PAGE gel showing how protein solubility is affected by the different Car9 tagged proteins. (1) Whole cell (2) Soluble fraction (3) Insoluble fraction.

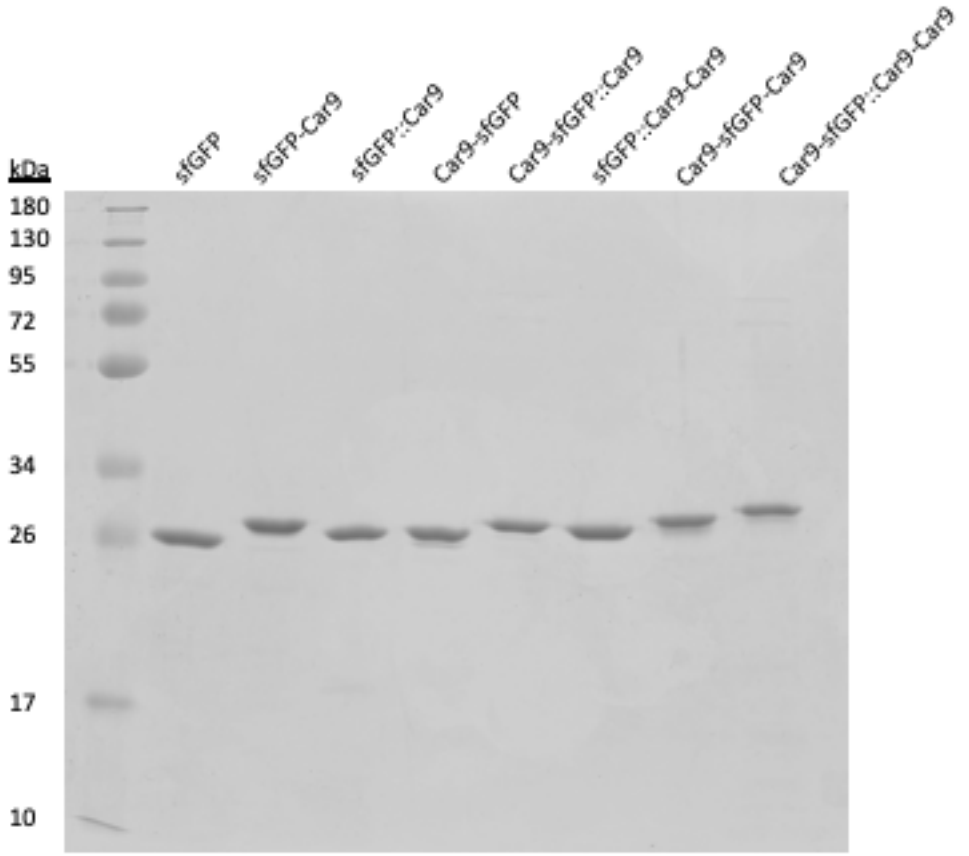


Figure 6.3 SDS-PAGE gel showing the purity of the Car9 tagged proteins purified by silica affinity chromatography..

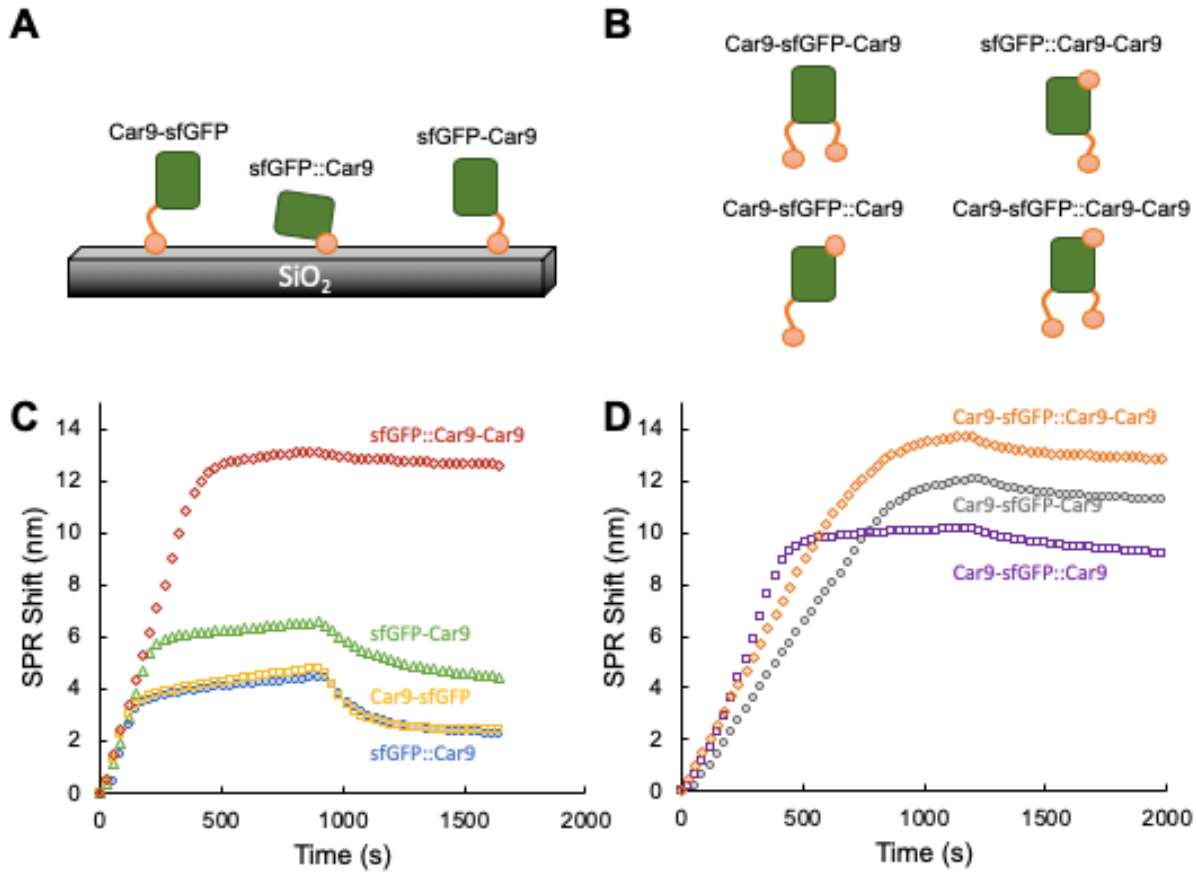


Figure 6.4 (A) Potential SiO_2 binding poses for the N-terminal Car9-sfGFP, with the longer linker, the rigid permissive loop sfGFP::Car9, and the C-terminal sfGFP-Car9. **(B)** Diagram of the multitagged Car9 proteins showing the location of the Car9 peptide and demonstrating how two tags could be "near" each other to have potential electrostatic repulsion. SPR results for the multivalent silica binding studies with the 7 various proteins at 0.1 μM in 20 mM Tris-HCl pH 7.5 with a flowrate of 70 μlmin^{-1} with an adsorption time of **(C)** 900 seconds and **(D)** 1200 seconds.

Chapter 7 Conclusions

In this work we have employed a combination of experiments and simulation to understand how silica binding kinetics and energetics correlate with amino acid composition and predicted conformation in a family of silica-binding peptides fused to the sfGFP framework.

In Chapter 2, we used SPR and silica-coated chips to characterize the adsorption behavior of these fusion proteins. We developed a cooperative two-step kinetic model that explains the high-affinity and sigmoidal binding of Car9 to the substrate through intermolecular contacts between nearest neighbor SBPs, and more accurately captures desorption kinetics through a slow isomerization step that stabilizes the fusion protein on the surface. In **C**hapter 3, we applied this model to a panel of Car9 mutants to assess how the energetic and kinetics of silica binding were influenced by mutations affecting the charge or structure of the Car9 SBP. We combined this information with molecular dynamic simulations initiated from *ab initio* Rosetta structure predictions to develop an in-depth understanding of how Car9 interacts with silica interfaces. In Chapter 4, we harnessed this knowledge to understand how peptide sequence and structure influenced the formation, crystallinity, and morphology of precipitated TiO_2. We discovered that Car9 and its K8HK11H mutant were most effective at promoting the formation of TiO_2(*B*) nanocrystallites, and that presentation of positively charged side chains in a compact conformation helped improve yields, and that conversion of basic residues to uncharged amino acids biased the phase of the crystallites towards TiO_2 anatase. In Chapter 5, we showed that the diffusion of solid-binding proteins and precursors in solid gels can be used to control the precipitation of TiO_2 in geometrically defined regions and to influence its crystallinity. Finally, in Chapter 6, we investigated how the local insertional context and the engineering of multiple Car9 segments within the sfGFP framework influenced silica binding kinetics and affinity.

The approaches and fundamental understanding developed in this work provide a roadmap that can be applied to other technologically useful SBPs. For instance, computationally inexpensive Rosetta predictions could be deployed to identify especially promising peptides from combinatorially selected populations, mutagenesis informed by MD simulations could be used to improve function (e.g., variable affinity for a material or precipitation of a defined crystallographic phase); insertion of SBPs at defined locations in the sfGFP framework and eventually other scaffolds could be used to modulate adhesion; and gel diffusion could be used to control morphology at larger scales. This work is also highly relevant to the broader goal of installing SBPs within computationally designed scaffolds for controlled assembly of hierarchal structures and the production of advanced materials and devices, with applications in fields as varied as biomedical engineering[105-108], catalysis[109,110], and opto-electronics.[114,115]

Appendix A Supplementary Information for Chapter 2

Note 1 Justifying the Assumption of Kinetic Control

In modelling the binding kinetics, we neglect any transport limitations and assume that the evolution of the SPR signal is determined by the attachment of the analyte molecule to the surface. The validity of this assumption is confirmed by estimating the Damkohler number for this system. The Damkohler number (Da) of a reaction-diffusion system is a dimensionless group that is essentially a measure of the relative rate of surface reaction to that of mass transport by diffusion. Alternatively, it may be regarded as the ratio of the characteristic timescales of diffusion to that of surface reaction. A low value of Da thus implies strong kinetic limitations relative to diffusion. For our system, Da is defined as:

$$Da = \frac{k_{a1} c_s l}{D} \tag{A.1}$$

where c_s is the surface concentration of binding sites, k_{a1} the rate constant for the first association reaction, and l and D the characteristic diffusion length scale and the diffusion coefficient of the protein in water, respectively.

For solid-binding proteins obeying the two-step cooperative model, k_{a1} is surface-coverage dependent and we take 100 $m^3\ mol^{-1}\ s^{-1}$ as a typical average value. Since rate laws are expressed in terms of the SPR signals $R_1(t)$ and $R_2(t)$, the exact concentration of surface binding sites is unknown, and we take c_s to have a typical value of 10^{-8} $molm^{-2}$ based on the method of Hansen *et al.*[174] The characteristic diffusion length perpendicular to the direction of bulk flow is approximately l = 5 μm while the diffusion coefficient of proteins in water is about 10^{-10} $m^2\ s^{-1}$ is used for D.[175,176]

Substituting these values in Equation (A.1) gives us a *Da* value of approximately 0.05 which implies that kinetic limitations are more severe than transport. The kinetically controlled regime is also indicated in the dissimilar SPR signal response signals for different analytes for

experiments at identical analyte flowrates. A transport-limited regime would likely mask the differences in association kinetics in the different proteins, all of which possess similar transport properties.

Note 2 Derivation of Model Equations

Let $\sigma_1(t)$ and $\sigma_2(t)$ denote the instantaneous surface concentrations of the two adsorbed protein isomers. We assume these quantities to be proportional to their respective contributions to the total SPR signal, giving us the following relations:

$$\sigma_1(t) = \kappa R_1(t) \tag{A.2}$$

$$\sigma_2(t) = \kappa R_2(t) \tag{A.3}$$

$$\sigma_{max}(t) = \kappa R_{max}(t) \tag{A.4}$$

where κ is a proportionality constant.

The total surface coverage may thus be written using Equations (A.2) – (A.4) as:

$$\theta = \frac{\sigma_1(t) + \sigma_2(t)}{\sigma_{max}} = \frac{R_1(t) + R_2(t)}{R_{max}} \tag{A.5}$$

recapitulating the chemical equations for the adsorption phase,

$$A + B \rightleftharpoons AB \rightleftharpoons AB^* \tag{A.6}$$

here, A denotes the adsorbing protein, and B denotes a surface binding site.

The rate of the initial adsorption step may be expressed as a product of a molar flux J_A and sticking probability p as follows:

$$r_{ads} = J_A p \tag{A.7}$$

For a constant bulk solution concentration c_A, the molar flux for adsorption from dilute solution may be found in terms of the Hertz-Knudsen equation as follows:[177,178]

$$J_A = c_A \sqrt{\frac{RT}{2\pi M}} \tag{A.8}$$

where the quantity $\sqrt{\frac{RT}{2\pi M}}$ is proportional to the average velocity obtained assuming a Maxwell-Boltzmann distribution for molecular velocities at a given temperature T. M denotes the molar mass of the adsorbing protein.

The sticking probability *p* represents the fraction of molecules arriving at the interface that get adsorbed, and thus corrects the bulk flux to give the net adsorption flux at the interface.[178] In the Frumkin-Fowler-Guggenheim model, *p* is a function of adsorbate–adsorbate interactions, in addition to the fraction of available sites and the activation energy barrier for bond formation. It is therefore denoted as:

$$p = (1-\theta) e^{\frac{-E_0}{RT}} e^{\frac{-E_i \theta}{RT}} \tag{A.9}$$

where E_0 the molar intrinsic activation energy for bond formation between the adsorbing molecule and the surface, and E_i the molar interaction energy considering nearest-neighbor interactions of the adsorbed species.[123] In the absence of adsorbate interactions, $E_i = 0$ and we recover the classic Langmuir dependence.

Equations (A.7) – (A.9) may now be combined to give the rate of the initial adsorption step:

$$r_{ads} = J_A p = c_A \sqrt{\frac{RT}{2\pi M}} \, (1-\theta) e^{\frac{-E_0}{RT}} e^{\frac{-E_i \theta}{RT}} \tag{A.10}$$

equivalently, we can use Equation A.5 to express the adsorption rate as follows:

$$r_{ads} = c_A \sqrt{\frac{RT}{2\pi M}} \, (1-\theta) e^{\frac{-E_0}{RT}} e^{\frac{-E_i \theta}{RT}} = \frac{1}{\sigma_{max}} \sqrt{\frac{RT}{2\pi M}} (\sigma_{max} - \sigma_1(t) - \sigma_2(t)) e^{\frac{-E_0}{RT}} e^{\frac{-E_i \theta}{RT}} \tag{A.11}$$

We also assume first-order kinetics for the desorption and isomerization reactions denoted in Equation (A.6). Thus,

$$r_{des} = k_{d1}\sigma_1(t) \tag{A.12}$$

and the net rate of the isomerization reaction is given by:

$$r_{is} = r_{isf} - r_{isb} = k_{a2}\sigma_1(t) - k_{d2}\sigma_2(t) \tag{A.13}$$

We can now use the rate law expressions given by Equations (A.11) – (A.13) to arrive at the governing equations for the surface concentrations $\sigma_1(t)$ and $\sigma_2(t)$, or equivalently, the SPR responses $R_1(t)$ and $R_2(t)$. Thus,

$$\frac{d\sigma_1(t)}{dt} = r_{ads} - r_{des1} - r_{is} \tag{A.14}$$

$$\frac{d\sigma_2(t)}{dt} = r_{is} \tag{A.15}$$

which can be written as,

$$\frac{d\sigma_1(t)}{dt} = \frac{1}{\sigma_{max}}\sqrt{\frac{RT}{2\pi M}}\left(\sigma_{max} - \sigma_1(t) - \sigma_2(t)\right)e^{\frac{-E_0}{RT}}e^{\frac{-E_i\theta}{RT}} - k_{d1}\sigma_1(t) - k_{a2}\sigma_1(t) + k_{d2}\sigma_2(t) \tag{A.16}$$

and,

$$\frac{d\sigma_2(t)}{dt} = k_{a2}\sigma_1(t) - k_{d2}\sigma_2(t) \tag{A.17}$$

Equations (A.16) and (A.17) are now connected to experimentally relevant quantities by substituting Equations (A.2) – (A.4), to give the final set of governing equations:

$$\frac{dR_1(t)}{dt} = k_{a1}\left(R_{max} - R_1(t) - R_2(t)\right) - k_{d1}R_1(t) - k_{a2}R_1(t) + k_{d2}R_2(t) \tag{A.18}$$

$$\frac{dR_2(t)}{dt} = k_{a2}R_1(t) - k_{d2}R_2(t) \tag{A.19}$$

where, for a given adsorbing species, site concentration and temperature, the constant parameters have been combined into an effective adsorption rate constant k_{a1}, given by:

$$k_{a1} = k'_{a1}e^{u\theta} \tag{A.20}$$

$$u = \frac{-E_i}{RT} \tag{A.21}$$

$$k'_{a1} = \frac{1}{\sigma_{max}}\frac{-E_0}{RT}\sqrt{\frac{RT}{2\pi M}} \tag{A.22}$$

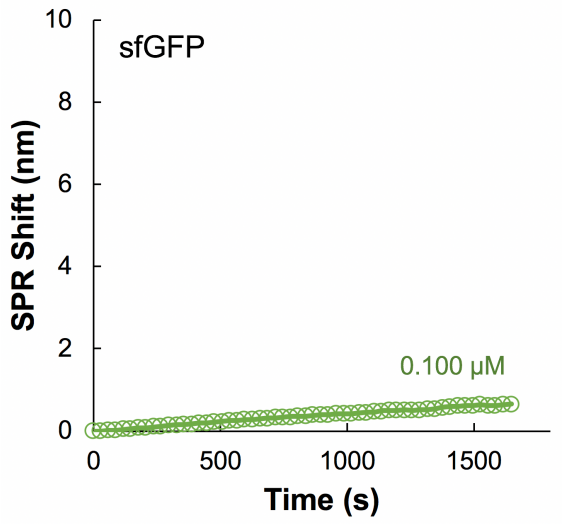


Figure A.1 Adsorption kinetics of sfGFP (0.1µM) on silica-coated SPR chips.

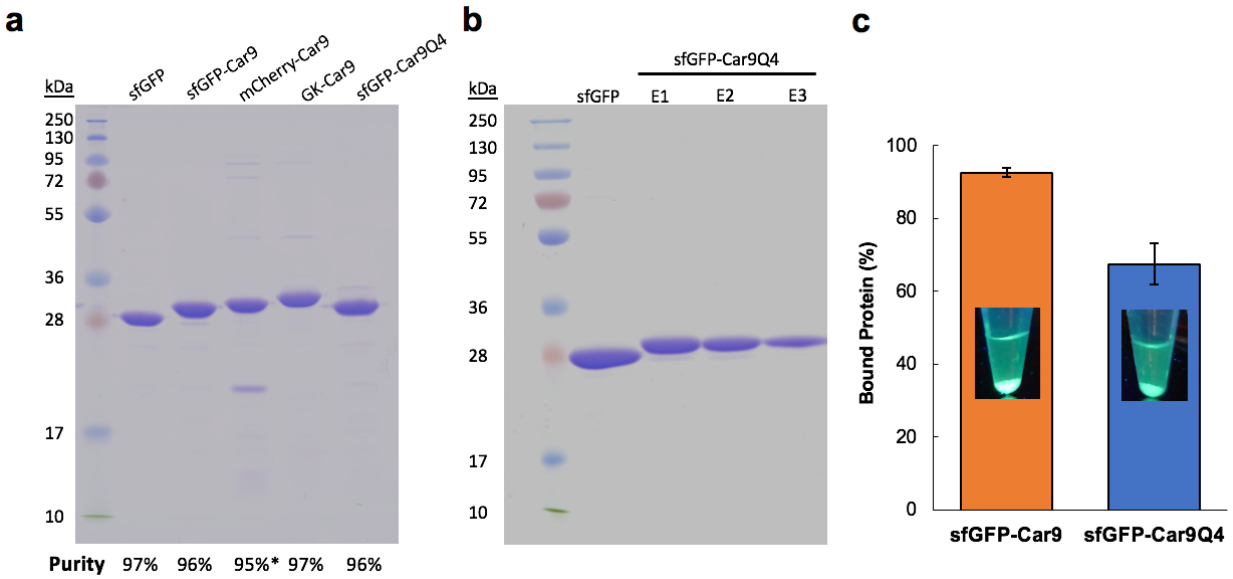
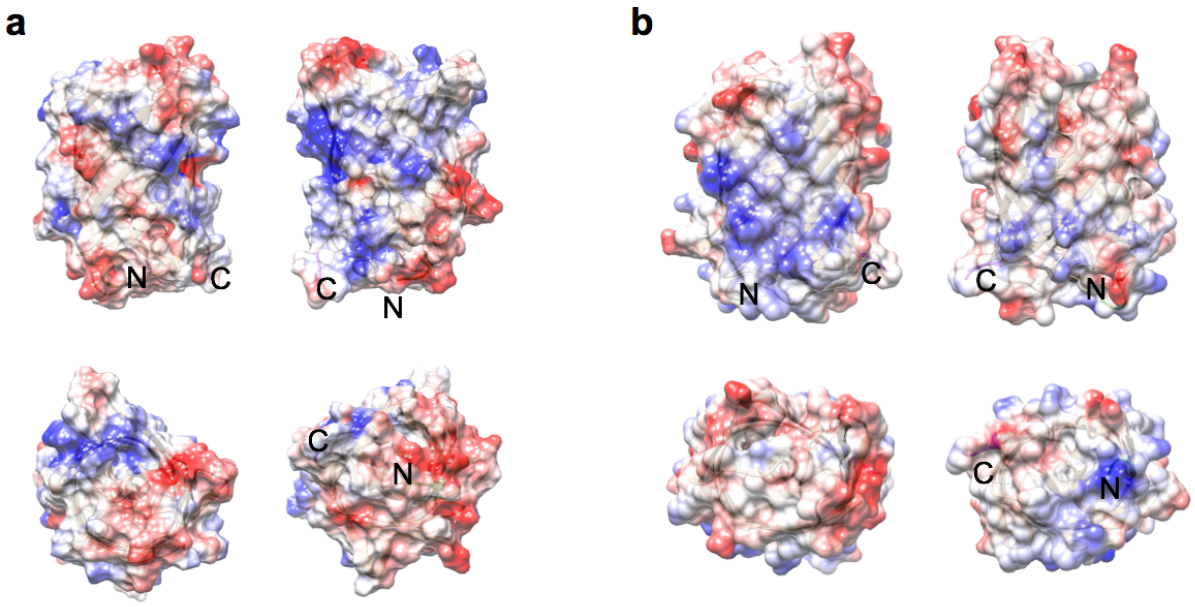


Figure A.2 Purification and silica-binding activity of proteins used in this study. (**A**) SDS-PAGE fractionation of purified proteins. At the exception of sfGFP which was isolated by conventional ion exchange chromatography,[17] all proteins were purified by silica affinity chromatography.[10,116] The lower molecular mass band in the mCherry-Car9 lane is a characteristic degradation fragment observed when mCherry is fractionated on SDS-PAGE gels.[179] This fragment is therefore included when calculating the purity of mCherry-Car9 (*). (**B**) SDS-PAGE analysis of successive sfGFP-Car9Q4 elution fractions from a silica column. Purified wild type sfGFP is included as a control. (**C**) Silica pull-down show that sfGFP-Car9Q4 retains a high level of silica-binding activity. For these experiments, 5 mg of silica gel was dispensed in triplicate microfuge tubes, and the washed powder was supplemented with 200 µL of 2 µM sfGFP-Car9 or sfGFP-Car9Q4. After 1 h incubation at room temperature, supernatant fluorescence was measured at 510 nm on a microplate reader with excitation at 475 nm and a cutoff wavelength of 495 nm. Values were normalized to untreated solutions of proteins (2 µM). Images were acquired on a transilluminator operated at 302 nm.



Protein	Mr (kDa)	pI
sfGFP	26.8	5.92
sfGFP-Car9	28.6	6.80

Protein	Mr (kDa)	pI
mCherry	26.9	5.62
mCherry-Car9	28.7	6.62

Figure A.3 Electrostatic surface of sfGFP **(A)** and mCherry **(B)** color-coded between -10 kcal/(mol*e) (red) and +10 kcal/(mol*e) (blue) were generated using Chimera. Images show clockwise the "front", "back", "bottom", and "top" of each protein. Tables list relative molar masses and pI.

a

Score	Expect	Method	Identities	Positives	Gaps
111 bits(277)	2e-32	Compositional matrix adjust.	76/190(40%)	111/190(58%)	12/190(6%)

```
Query  67   AIGIGAPGPVQEETGMLYEAVNL-GWKHYPLKRQLEEETGLPVAVDNDANIAALGEMWKG  125
            AIG+G PGP+    G++  A N+ G + +P++R LEE TG PV ++NDAN AAL E   G
Sbjct  60   AIGLGTPGPLDFRRGVIRFAPNIPGVQDFPIRRILEEATGRPVFLENDANAAALAEHHLG  119

Query  126  AGGGARHLLFVTLGTGVGGGVIANGAIVRGTNGAGGEIGHMTMVADGGAPCNCGKTGCLE  185
            A  G    L++T+ TG+GGGV+  G ++RG  G GGE+GH+T++  GG  C CG  GCLE
Sbjct  120  AAQGEESSLYLTVSTGIGGGVVLGGRVLRGERGQGGELGHLTLLP-GGPACGCGLEGCLE  178

Query  186  TIASATGIVRIAGEKLAASERPSALRGGDVTAKAVFDAAKTGDALALEVVEEVTRYLGLA  245
             +A+   + R A     A +RP       V  + +F   + GD  A  +V +  RY+G+
Sbjct  179  ALAAGRALERDA---TYAFQRP-------VDTRELFRLFQAGDPKAERLVLQAARYVGIG  228

Query  246  LANAANVTNP  255
            LA+     +P
Sbjct  229  LASLVKAFDP  238
```

b

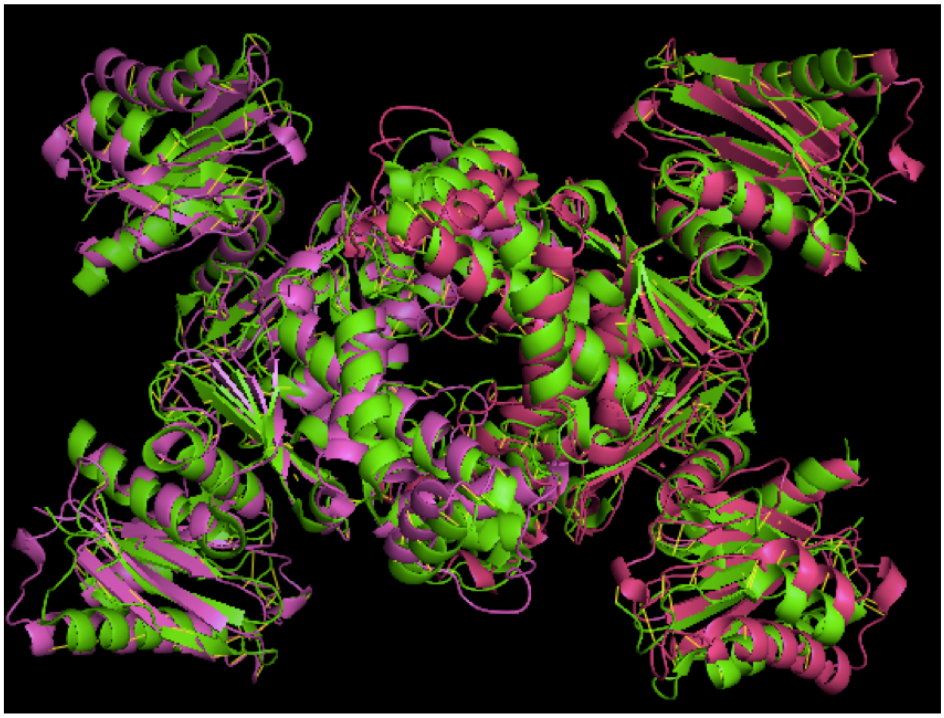
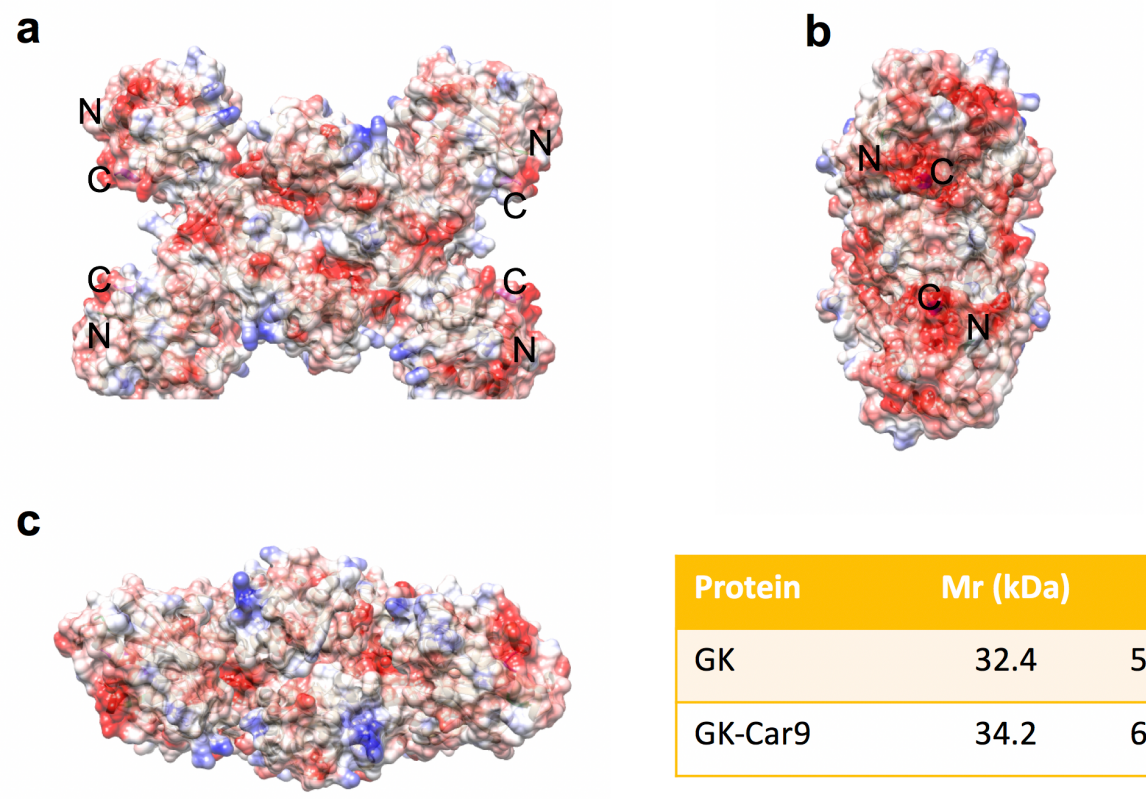


Figure A.4 (A) BLAST alignment of the *T. thermophilus* (3VOV, Query) and *G. stearothermophilus* (Sbjct) glucose kinase (GK) sequences over amino acids 60 to 238. **(B)** Dimeric homology models offered by SWISS-MODEL[122] (purple) were aligned to the 3VOV tetramer (green) in PyMol.



Protein	Mr (kDa)	pI
GK	32.4	5.56
GK-Car9	34.2	6.41

Figure A.5. Electrostatic surface of *G. stearothermophilus* glucose kinase (GK) color-coded between -10 kcal/(mol*e) (red) and +10 kcal/(mol*e) (blue) and generated using Chimera. Images show clockwise the "front", "right side", and "bottom" of the tetramer. The Table lists relative molar masses and pIs.

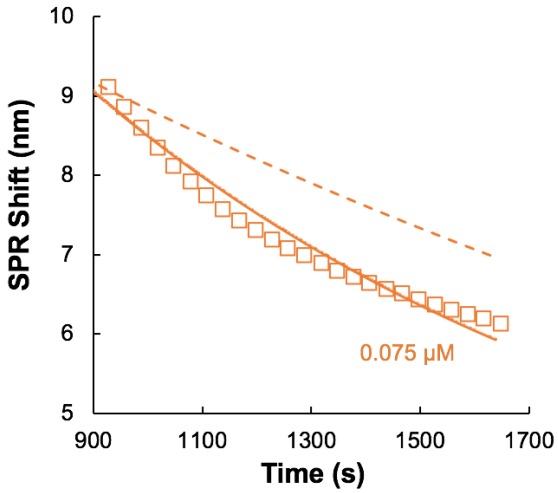


Figure A.6 Desorption kinetics of sfGFP-Car9 at 0.075 µM concentration modeled using a single exponential decay (dashed line) or the two-step cooperative model (solid lines).

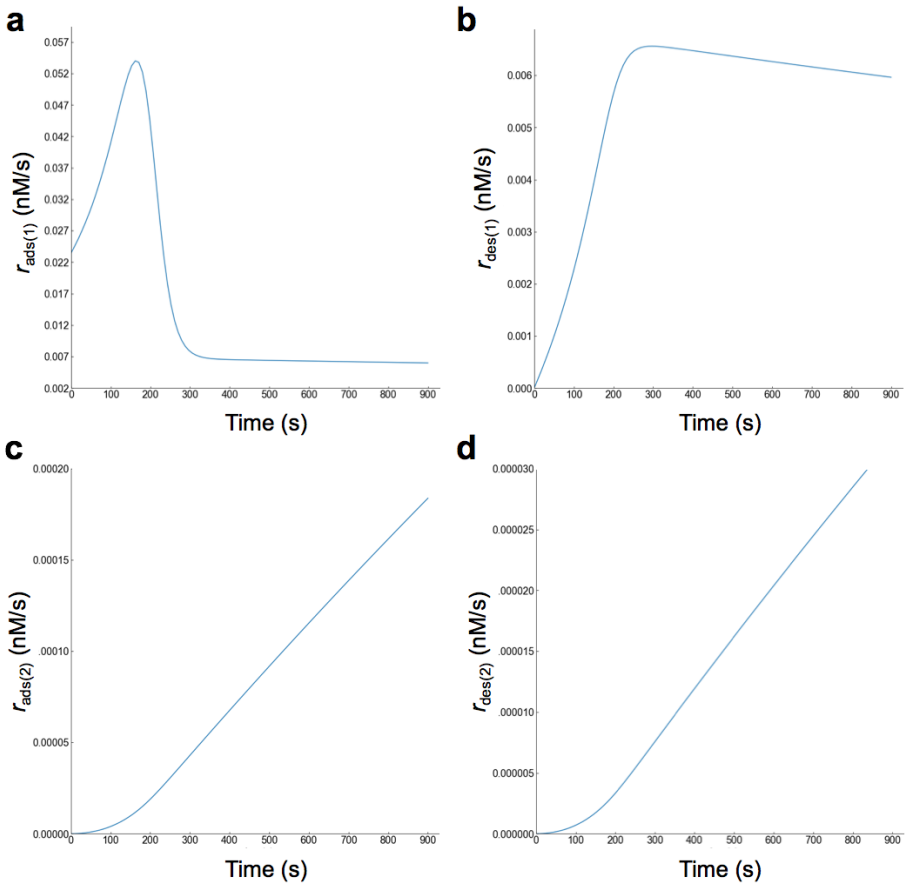


Figure A.7 Simulated instantaneous reaction rates for the initial cooperative adsorption **(A)** and desorption **(B)** from the silica surface, and for their second step adsorption **(C)** and desorption **(D)**. Note that the panels have different units. Simulations were generated using the fit parameters for 0.125 µM sfGFP-Car9 sensorgram data (Table 1 of the main text).

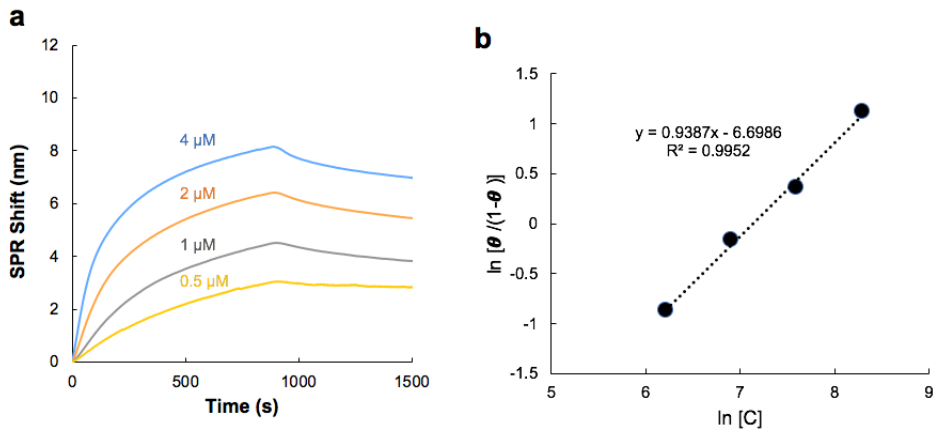


Figure A.8. SPR sensorgrams (**A**) and Langmuir adsorption isotherm (**B**) for the adsorption of sfGFP-Car9Q4 to silica. The slope of 0.94 is consistent with Langmuir behavior. Surface coverages (q) were calculated as the ratio of R_{eq} to R_{max} and a R_{max} value of 10.9 nm.

Table A.1: Summary of lower and upper bounds, and initial guesses for the fitting of sigmoidal sensorgrams using the proposed two step model.

Parameter	Protein	Condition	Lower Bound	Upper Bound	Initial Guess	Units
k'_{a1}	sfGFP-Car9	0 mM NaCl, pH 7.5	20000	25000	22000	$M^{-1}s^{-1}$
		50 mM NaCl, pH 7.5	20000	30000	25000	
		100 mM NaCl, pH 7.5	25000	35000	29000	
	GK-Car9	pH 7.5	15000	35000	18000	
	mCherry-Car9	pH 7.5	15000	33000	18000	
k_{d1}	sfGFP-Car9	0 mM NaCl, pH 7.5	0.32	1.0	0.55	$\times 10^{-3}$ s^{-1}
		50 mM NaCl, pH 7.5	0.32	3.2	1.26	
		100 mM NaCl, pH 7.5	1.0	10.0	5.0	
	GK-Car9	pH 7.5	0.32	1.0	0.63	
	mCherry-Car9	pH 7.5	0.32	1.0	0.63	
k_{a2}	sfGFP-Car9	0 mM NaCl, pH 7.5	0.06	0.32	0.13	$\times 10^{-3}$ s^{-1}
		50 mM NaCl, pH 7.5	0.03	0.20	0.06	
		100 mM NaCl, pH 7.5	0.16	1.0	0.56	
	GK-Car9	pH 7.5	0.06	1.0	0.1	
	mCherry-Car9	pH 7.5	0.06	1.0	0.1	
k_{d2}	sfGFP-Car9	0 mM NaCl, pH 7.5	0.02	0.06	0.03	$\times 10^{-3}$ s^{-1}
		50 mM NaCl, pH 7.5	0.02	0.06	0.03	
		100 mM NaCl, pH 7.5	0.001	0.01	0.03	
	GK-Car9	pH 7.5	0.02	0.32	0.03	
	mCherry-Car9	pH 7.5	0.02	0.32	0.1	
u	sfGFP-Car9	0 mM NaCl, pH 7.5	2.5	3.1	2.9	-
		50 mM NaCl, pH 7.5	1.8	2.8	2	
		100 mM NaCl, pH 7.5	1.5	3	2.4	
	GK-Car9	pH 7.5	2	3.5	2.5	
	mCherry-Car9	pH 7.5	2	3.5	2.5	

Appendix B Supplementary Information for Chapter 3

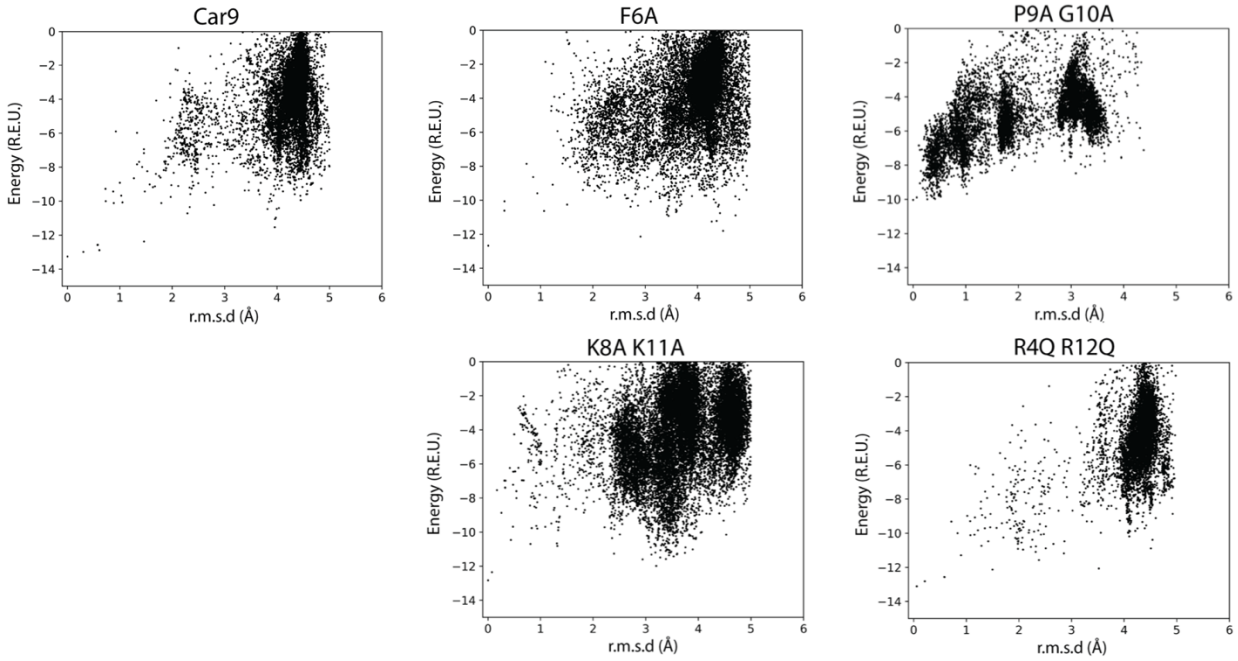


Figure B.1 Energy landscape sampled during Rosetta *ab initio* structure prediction of Car9 and its mutants. Rosetta energy in Rosetta Energy Units (R.E.U.) is plotted against C-α root mean square deviation (RMSD) to the lowest energy model. Rosetta energy was calculated with the Talaris 2014 score function.

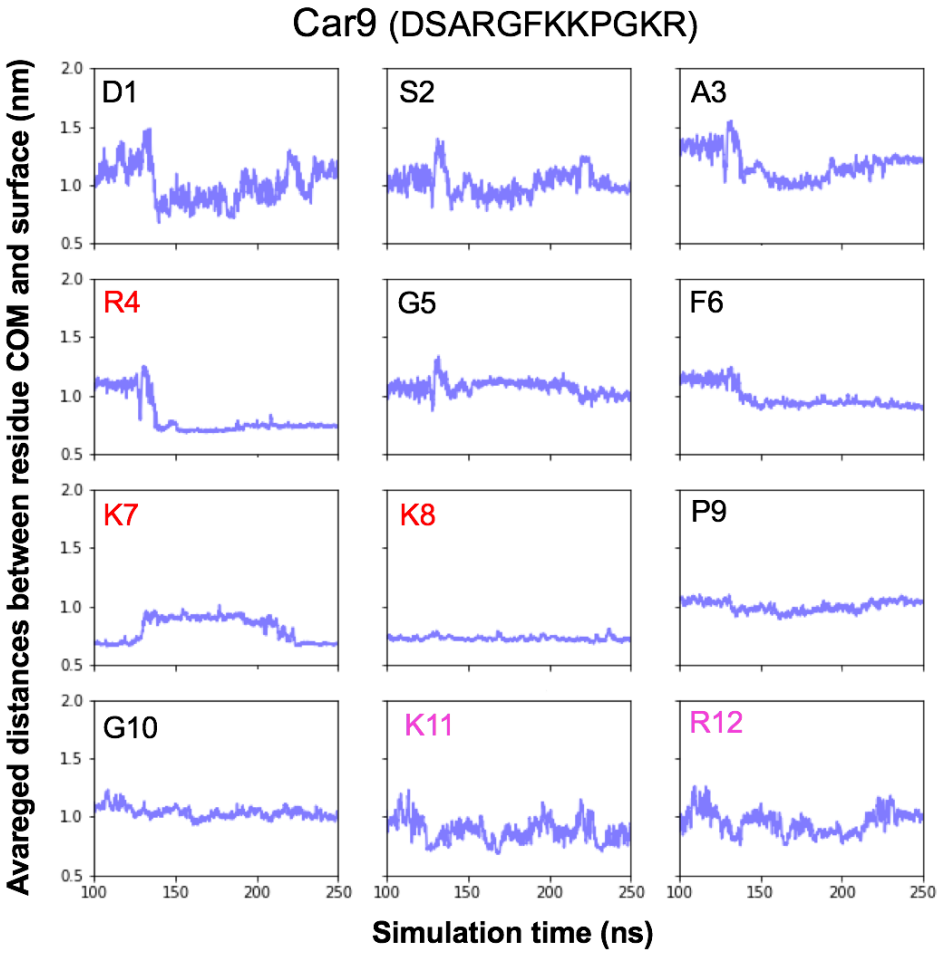


Figure B.2 Distance of each residue center of mass (COM) to the silica surface for the Car9 peptide. Distances were averaged for the last 150 ns of simulation over all three simulation trials. Anchor residues sit at short distances from the surface and exhibit low (primary anchor, red) or moderate (secondary anchor, pink) variance in the distances.

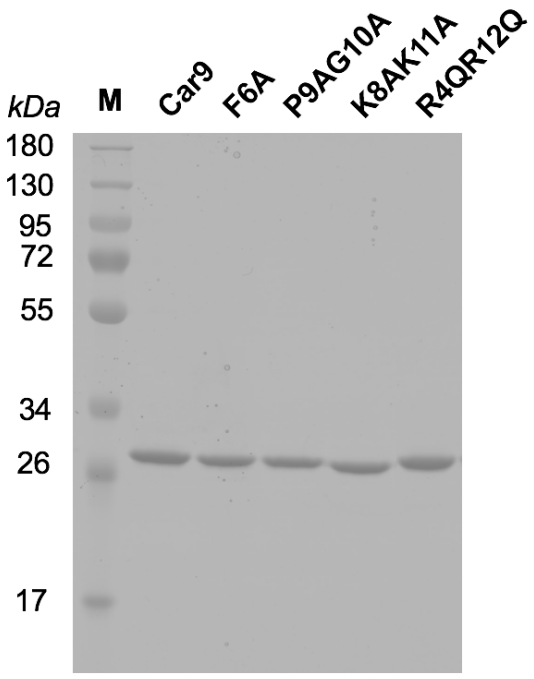


Figure B.3 SDS-PAGE fractionation of sfGFP-Car9 (Car9) and of its F6A, P9AG10A, K8AK11A, and R4QR12Q variants following silica affinity chromatography.

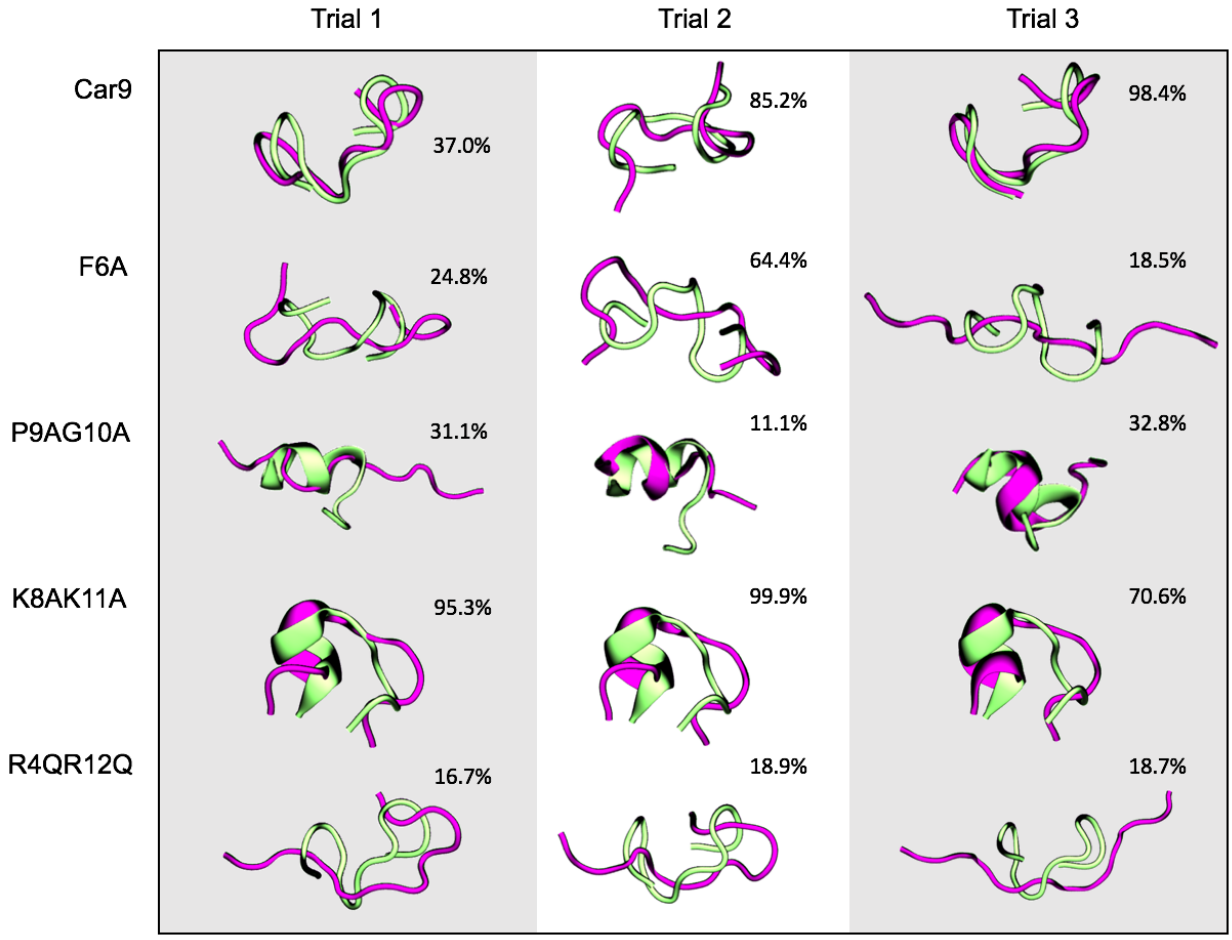


Figure B.4 Superimposition of Rosetta initial structures (green) and top-ranked clusters determined for each trial of MD-simulated structure at the silica interface (pink).

Appendix C Supplementary Information for Chapter 4

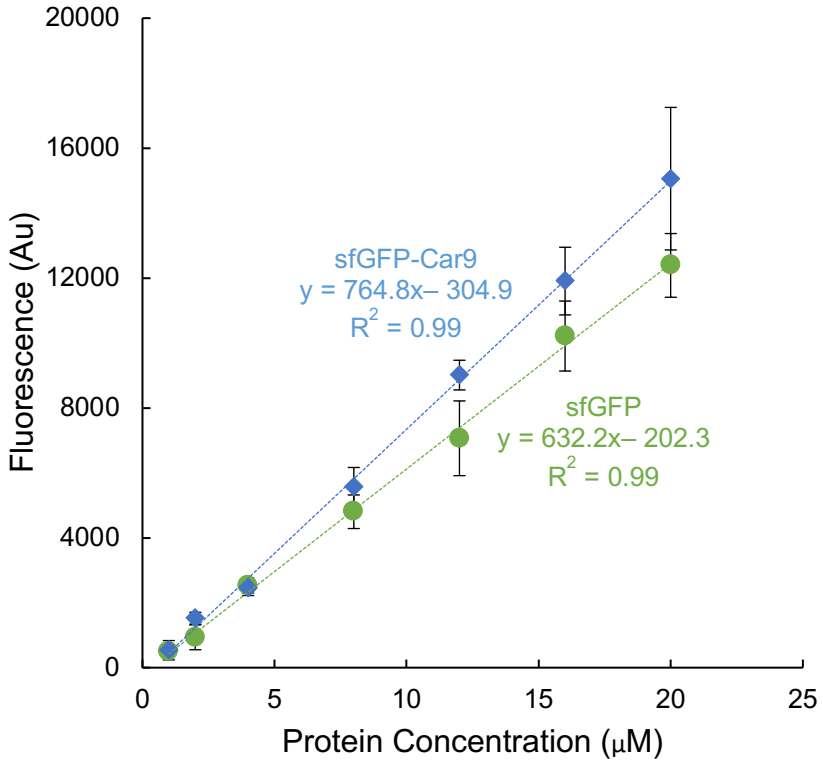


Figure C.1 Fluorescence calibration curve for sfGFP and sfGFP-Car9 in 37.5 mM citrate buffer at pH 5.

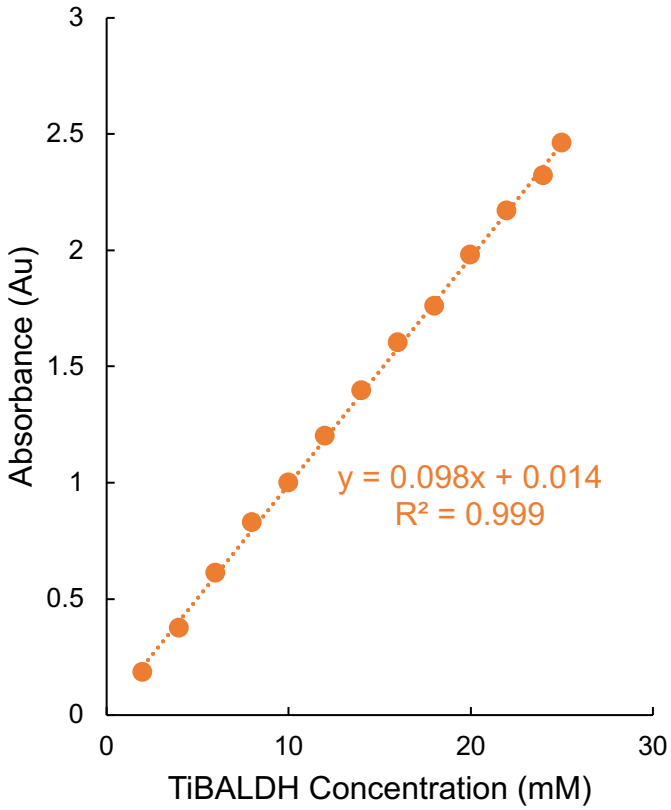


Figure C.2 Ti(IV) colorimetric assay calibration curve used to determine the concentration of unreacted TiBALDH that is remaining in solution after TiO_2 precipitation.

Appendix D Kinetic binding coefficients determined using the Langmuir and Hill Isotherm Fit method to SPR data

Estimating the kinetic binding coefficients (k_{on}, k_{off}, and K_D) from SPR sensorgram data and using these values to determine the free energy of binding (ΔG). First SPR sensorgram data was collected for each protein at varying concentrations (high affinity proteins 1000nM to 15.6 nM and low affinity 4000 nM to 62.5 nM). Two SPR chips (4 concentrations per chip) were used with one common concentration used to align the SPR data, Figure D.1. From this data a response equilibrium, R_{eq}, was estimated for each concentration, Figure D.2 and Table D.1. The R_{eq} is considered to be when the SPR sensorgram reaches a plateau for each concentration. For sensorgrams that seemed to have non-monolayer surface coverage at higher concentrations (sfGFP-Car9 at 1000 nM) the R_{eq} was assumed to be at the point the curve would level off and not considering the rise caused by non-monolayer surface coverage.

An isotherm plot was constructed from plotting the R_{eq} with concentration, Figure D.3. Each isotherm plot was fit with the Langmuir isotherm, Equation D.1, and the sigmoidal binders were also fit to the Hill isotherm, Equation D.2:

$$R_{eq} = \frac{[c]R_{max}}{K_D + [c]} \tag{D.1}$$

$$R_{eq} = \frac{[c]^n R_{max}}{K_D + [c]^b} \tag{D.2}$$

where R_{max} is the maximum SPR response, K_D the equilibrium dissociation constant, and *n* the Hill coefficient that denotes binding cooperativity. Dissociation constants and Hill coefficients were determined using the Chi-Square curve fitting method in Excel.

Kinetic off rate constants (k_{off}) were obtained by fitting raw SPR data in the wash cycle with an exponential decay, Equation D.3,

$$R = exp^{-k_{off}t} \quad \text{(D.3)}$$

where R is the SPR response and t the time. From these values the k_{on}, Equation D.4, and the Gibbs free energies of binding, Equation D.5, can be calculated.

$$k_{on} = \frac{k_{off}}{K_D} \quad \text{(D.4)}$$

$$\Delta G = -RT \ln(\frac{1}{K_D}) \quad \text{(D.5)}$$

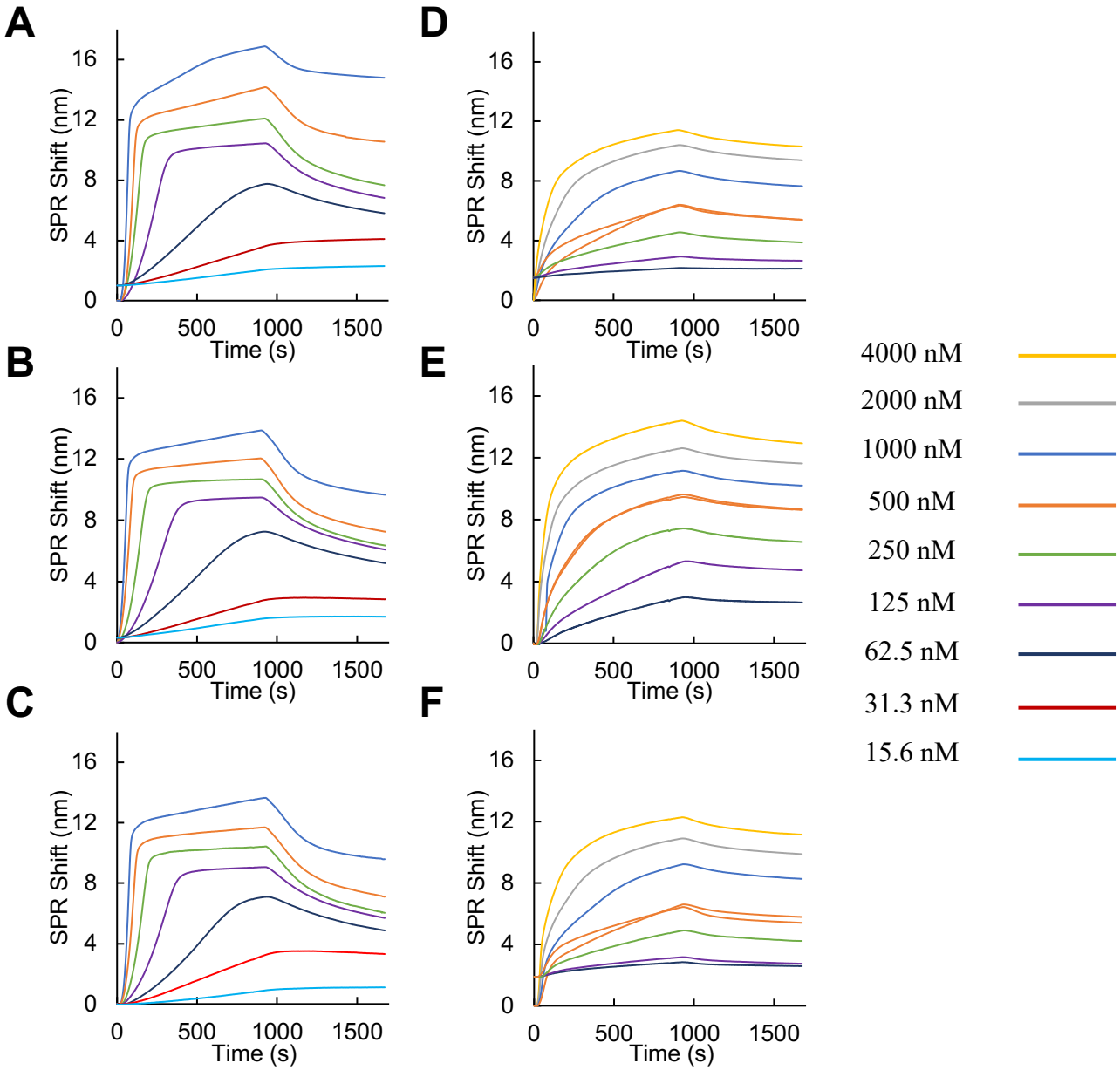


Figure D.1 Raw SPR sensorgrams for C-terminal sfGFP tagged proteins with **(A)** Car9, **(B)** F6A, **(C)** P9AG10A, **(D)** K8AK11A, **(E)** R4QR12Q, and **(F)** K7AR12A**.** Two SPR chips area aligned using the concentration 125 nM for sigmoidal binders and 500 nM for Langmuir binders.

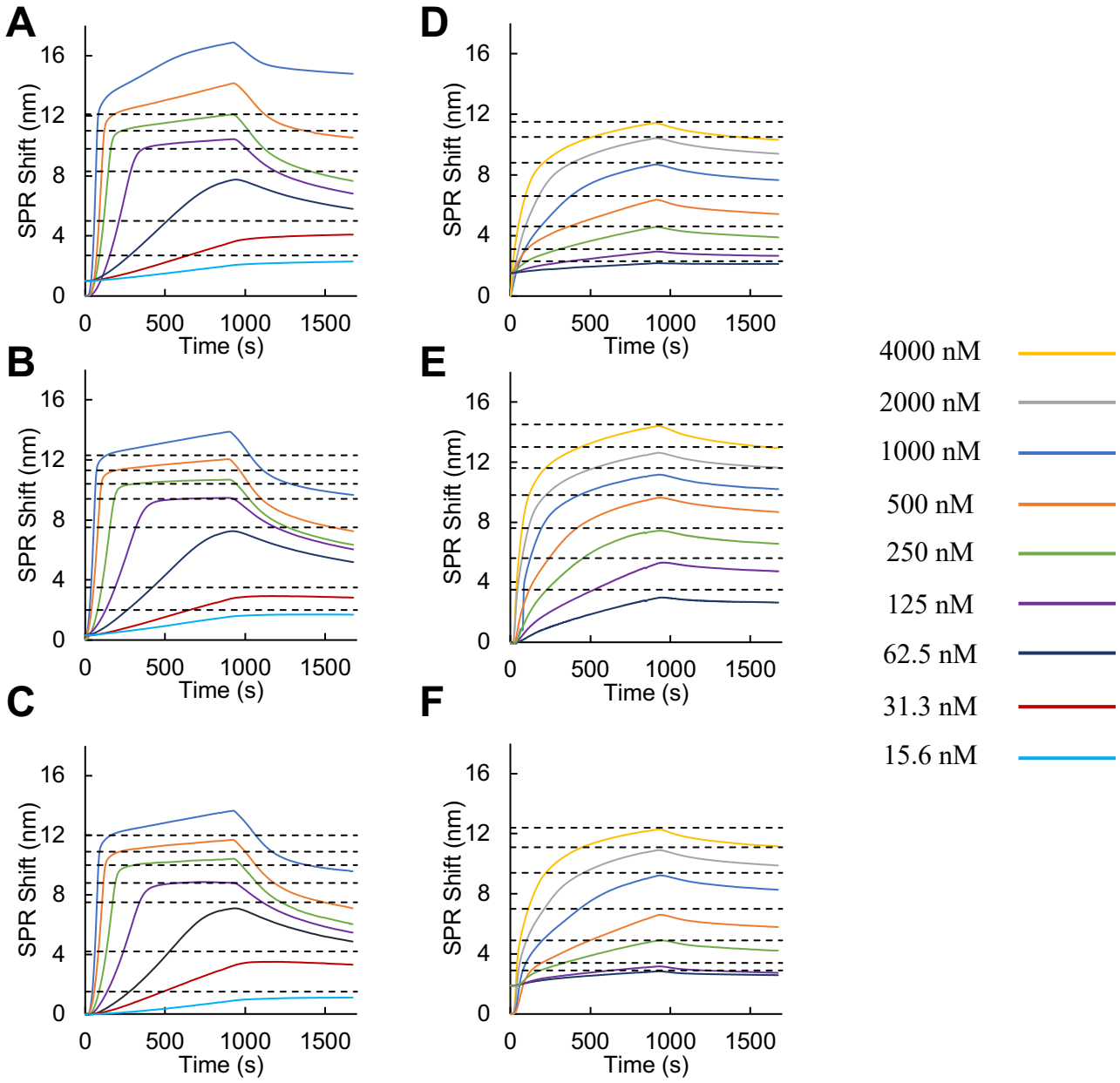


Figure D.2 R_{eq} estimates for each concentration for the raw SPR sensorgrams with the C-terminal sfGFP tagged proteins with **(A)** Car9, **(B)** F6A, **(C)** P9AG10A, **(D)** K8AK11A, **(E)** R4QR12Q, and **(F)** K7AR12A.

Table D.1 R_{eq} estimates for each concentration for the raw SPR sensorgrams with the C-terminal sfGFP tagged proteins.

Concentration (nM)	Car9 R_{eq} (nm)	F6A R_{eq} (nm)	P9AG10A R_{eq} (nm)	K8AK11A R_{eq} (nm)	R4QR12Q R_{eq} (nm)	K7AR12A R_{eq} (nm)
15.6	2.7	2.0	1.5	-	-	-
31.3	5.0	3.5	4.2	-	-	-
62.5	8.3	7.5	7.5	2.3	3.5	2.9
125	9.8	9.4	8.8	3.1	5.6	3.4
250	11.0	10.4	10.0	4.6	7.6	4.9
500	12.1	11.3	10.9	6.6	9.8	7.0
1000	13.5	12.3	12.0	8.8	11.6	9.4
2000	-	-	-	10.5	13.0	11.1
4000	-	-	-	11.5	14.5	12.4

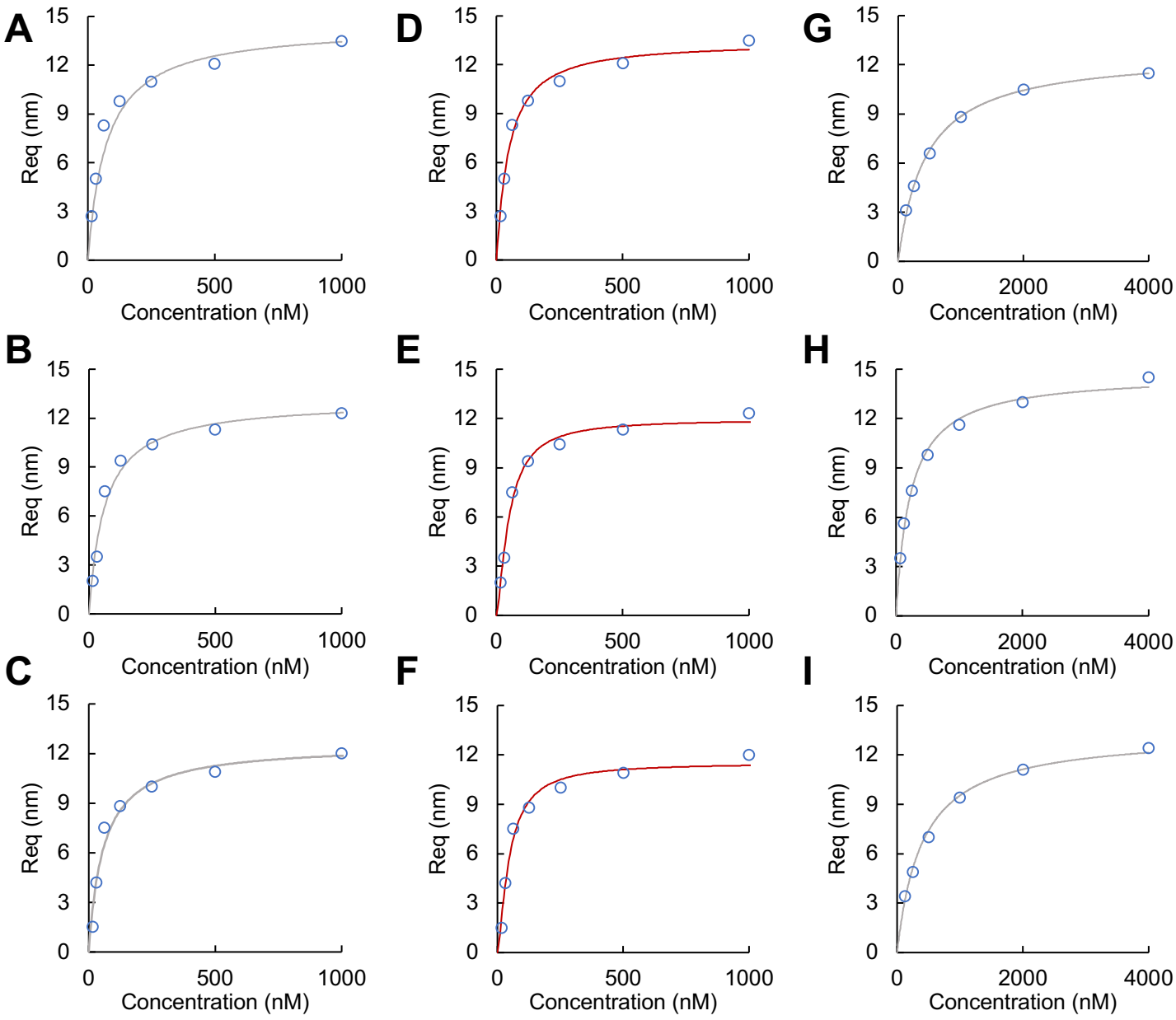


Figure D.3 Isotherm plots for SPR data of the sigmoidal binding proteins Car9, F6A, and P9QG10A. Each is fit to the Langmuir Isotherm (grey) or Hill Isotherm (red) for **(A, D)** Car9, **(B, E)** F6A, and **(C, F)** P9AG10A. Fits for the Langmuir binding proteins **(G)** K8AK11A, **(H)** R4QR12Q, and **(I)** K7AR12A.

Table D.2 Calculated kinetic binding constants by fitting the Langmuir equation to the isotherm data. K_D, k_{off}, and R_{max} (blue) where determined from fitting the data and k_{on}, ΔG, and Chi^2 where calculated from the values determined.

Langmuir Isotherm	K_D (μM)	k_{off} (s^{-1})	R_{max} (nm)	k_{on} ($\mu M^{-1}s^{-1}$)	ΔG ($kJmol^{-1}$)	Chi^2
Car9	0.0744	0.0035	14.4	0.0465	-40.7	0.85
F6A	0.0614	0.0031	13.1	0.0500	-41.1	0.52
P9AG10A	0.0580	0.0027	12.6	0.0469	-41.3	0.65
K8AK11A	0.4390	0.0027	12.7	0.0061	-36.3	0.02
R4QR12Q	0.2270	0.0024	14.7	0.0106	-37.9	0.22
K7AR12A	0.4009	0.0031	13.4	0.0078	-36.5	0.39

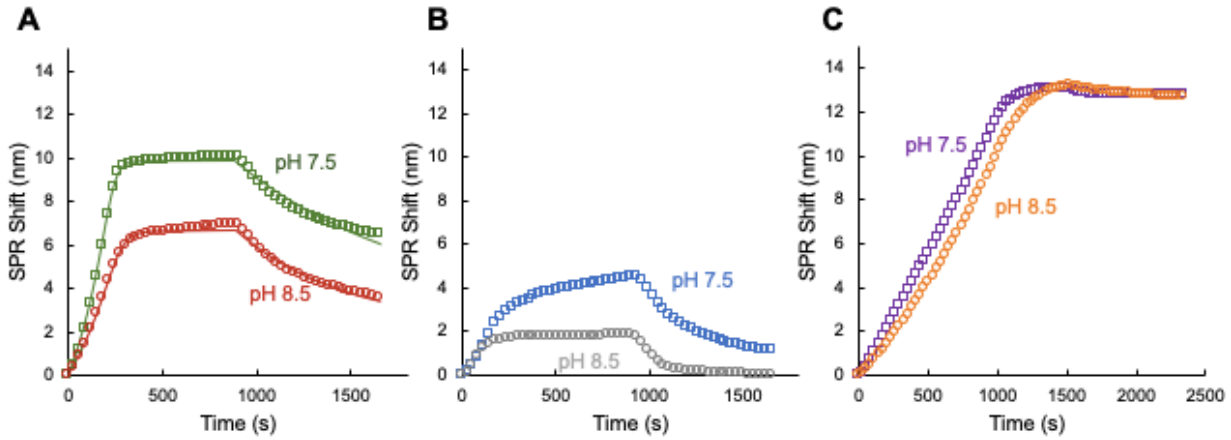
Table D.3 Calculated kinetic binding constants by fitting the Hill equation to the isotherm data. K_D, k_{off}, n, and R_{max} (blue) where determined from fitting the data and k_{on}, ΔG, and Chi^2 where calculated from the values determined.

Hill Isotherm	K_D (μM)	k_{off} (s^{-1})	n	R_{max} (nm)	k_{on} ($\mu M^{-1}s^{-1}$)	ΔG ($kJmol^{-1}$)	Chi^2
Car9	0.071	0.0035	1.10	13.4	0.0487	-40.8	0.31
F6A	0.284	0.0031	1.44	12.0	0.0108	-37.4	0.29
P9AG10A	0.225	0.0027	1.40	11.5	0.0121	-37.9	0.37

Table D.4 Average k_{off} values determined from fitting the exponential decay function, Equation D.3, to the off rate of the SPR sensorgram data for the five largest concentrations of each protein.

k_{off} error	k_{off} (s^{-1})
Car9	0.0035 ± 0.002
F6A	0.0031 ± 0.001
P9AG10A	0.0027 ± 0.001
K8AK11A	0.0027 ± 0.0004
R4QR12Q	0.0024 ± 0.0004
K7AR12A	0.0031 ± 0.0007

Appendix E Effects of pH on SiO_2 binding



D

Protein	pH	Concentration (μM)	u	k_{a1}' ($M^{-1}s^{-1}$)	k_{d1} x 10^{-3} (s^{-1})	k_{a2} x 10^{-3} (s^{-1})	k_{d2} x 10^{-3} (s^{-1})
sfGFP-Car9	7.5	0.1	2.9	22500	0.84	0.20	0.02
sfGFP-Car9	8.5	0.1	2.7	21700	1.45	0.32	0.04

Figure E.1 Comparison of effects of pH on silica binding, experiments were run in 20 mM Tris-HCl (pH 7.5 or pH 8.5) with a flow rate of 50 μLmin^{-1}. **(A)** 0.1 μM sfGFP-Car9 **(B)** 0.1 μM sfGFP::Car9 **(C)** 0.2 μM sfGFP::Car9-Car9. **(A)** and **(B)** adsorption time was 900 seconds and desorption time 900 seconds. **(C)** adsorption time was 1200 seconds and desorption time 900 seconds. **(D)** Kinetic values for sfGFP-Car9 at pH 7.5 and pH 8.5 fit with the two-step cooperative model.

References

1 Kim, H. M., Himeno, T., Kawashita, M., Kokubo, T. & Nakamura, T. The mechanism of biomineralization of bone-like apatite on synthetic hydroxyapatite: an in vitro assessment. *Journal of the Royal Society Interface* **1**, 17-22, doi:10.1098/rsif.2004.0003 (2004).

2 Dhami, N. K., Reddy, M. S. & Mukherjee, A. Biomineralization of calcium carbonates and their engineered applications: a review. *Frontiers in Microbiology* **4**, 314, doi:10.3389/fmicb.2013.00314 (2013).

3 Kröger, N., Deutzmann, R., Bergsdorf, C. & Sumper, M. Species-specific polyamines from diatoms control silica morphology. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 14133-14138 (2000).

4 Gröger, P., Poulsen, N., Klemm, J., Kröger, N. & Schlierf, M. Establishing super-resolution imaging for proteins in diatom biosilica. **6**, 36824, doi:10.1038/srep36824 https://www.nature.com/articles/srep36824#supplementary-information (2016).

5 Tong, Z. *et al.* Biomimetic and bioinspired synthesis of titania and titania-based materials. *RSC Advances* **4**, 12388-12403, doi:10.1039/C3RA47336H (2014).

6 Dorval Courchesne, N.-M., Steiner, S. A., Cantú, V. J., Hammond, P. T. & Belcher, A. M. Biotemplated Silica and Silicon Materials as Building Blocks for Micro- to Nanostructures. *Chemistry of Materials* **27**, 5361-5370, doi:10.1021/acs.chemmater.5b01844 (2015).

7 Dolatabadi, J. E. N. & de la Guardia, M. Applications of diatoms and silica nanotechnology in biosensing, drug and gene delivery, and formation of complex metal nanostructures. *TrAC Trends in Analytical Chemistry* **30**, 1538-1548, doi:https://doi.org/10.1016/j.trac.2011.04.015 (2011).

8 Dickerson, M. B., Sandhage, K. H. & Naik, R. R. Protein- and Peptide-Directed Syntheses of Inorganic Materials. *Chemical Reviews* **108**, 4935-4978, doi:10.1021/cr8002328 (2008).

9 Coyle, B. L., Rolandi, M. & Baneyx, F. Carbon-Binding Designer Proteins that Discriminate between sp2- and sp3-Hybridized Carbon Surfaces. *Langmuir* **29**, 4839-4846, doi:10.1021/la4000846 (2013).

10 Coyle, B. L. & Baneyx, F. A cleavable silica-binding affinity tag for rapid and inexpensive protein purification. *Biotechnology and Bioengineering* **111**, 2019-2026, doi:10.1002/bit.25257 (2014).

11 Pedelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotech* **24**, 79-88, doi:http://www.nature.com/nbt/journal/v24/n1/suppinfo/nbt1172_S1.html (2006).

12 Baneyx, F. & Schwartz, D. T. Selection and analysis of solid-binding peptides. *Current Opinion in Biotechnology* **18**, 312-317, doi:https://doi.org/10.1016/j.copbio.2007.04.008 (2007).

13 Tamerler, C. & Sarikaya, M. Molecular biomimetics: Utilizing nature's molecular ways in practical engineering. *Acta Biomaterialia* **3**, 289-299, doi:http://dx.doi.org/10.1016/j.actbio.2006.10.009 (2007).

14 Tamerler, C. & Sarikaya, M. Genetically Designed Peptide-Based Molecular Materials. *ACS Nano* **3**, 1606-1615, doi:10.1021/nn900720g (2009).

15 Coyle, B. L., Zhou, W. & Baneyx, F. *Bionanotechnology: Biological Self-assembly and its Applications Chapter 3: Protein-aided mineralization of inorganic nanostructures*. Chapter 3 (Caister Academic Press, 2013).
16 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. *Bioconjugate Chemistry* **27**, 2450-2459, doi:10.1021/acs.bioconjchem.6b00406 (2016).
17 Coyle, B. L. & Baneyx, F. Direct and reversible immobilization and microcontact printing of functional proteins on glass using a genetically appended silica-binding tag. *Chemical Communications* **52**, 7001-7004, doi:10.1039/C6CC02660E (2016).
18 Kröger, N. & Poulsen, N. Diatoms—From Cell Wall Biogenesis to Nanotechnology. *Annual Review of Genetics* **42**, 83-107, doi:doi:10.1146/annurev.genet.41.110306.130109 (2008).
19 Hildebrand, M. & Lerch, S. J. L. Diatom silica biomineralization: Parallel development of approaches and understanding. *Seminars in Cell & Developmental Biology* **46**, 27-35, doi:https://doi.org/10.1016/j.semcdb.2015.06.007 (2015).
20 Sumper, M. & Kroger, N. Silica formation in diatoms: the function of long-chain polyamines and silaffins. *Journal of Materials Chemistry* **14**, 2059-2065, doi:10.1039/B401028K (2004).
21 Hildebrand, M. & Lerch, S. J. L. Diatom silica biomineralization: Parallel development of approaches and understanding. *Seminars in Cell & Developmental Biology* **46**, 27-35, doi:http://dx.doi.org/10.1016/j.semcdb.2015.06.007 (2015).
22 Vrieling, E. G., Gieskes, W. W. C. & Beelen, T. P. M. SILICON DEPOSITION IN DIATOMS: CONTROL BY THE pH INSIDE THE SILICON DEPOSITION VESICLE. *Journal of Phycology* **35**, 548-559, doi:10.1046/j.1529-8817.1999.3530548.x (1999).
23 Kröger, N., Deutzmann, R. & Sumper, M. Polycationic Peptides from Diatom Biosilica That Direct Silica Nanosphere Formation. *Science* **286**, 1129-1132 (1999).
24 Lechner, C. C. & Becker, C. F. W. Silaffins in Silica Biomineralization and Biomimetic Silica Precipitation. *Marine Drugs* **13**, 5297-5333, doi:10.3390/md13085297 (2015).
25 Brott, L. L. *et al.* Ultrafast holographic nanopatterning of biocatalytically formed silica. *Nature* **413**, 291-293 (2001).
26 Kröger, N., Deutzmann, R. & Sumper, M. Silica-precipitating Peptides from Diatoms: THE CHEMICAL STRUCTURE OF SILAFFIN-1A FROM CYLINDROTHECA FUSIFORMIS. *Journal of Biological Chemistry* **276**, 26066-26070 (2001).
27 Kröger, N., Lorenz, S., Brunner, E. & Sumper, M. Self-Assembly of Highly Phosphorylated Silaffins and Their Function in Biosilica Morphogenesis. *Science* **298**, 584-586, doi:10.1126/science.1076221 (2002).
28 Poulsen, N., Sumper, M. & Kröger, N. Biosilica formation in diatoms: Characterization of native silaffin-2 and its role in silica morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12075-12080, doi:10.1073/pnas.2035131100 (2003).
29 Wieneke, R. *et al.* Silica precipitation with synthetic silaffin peptides. *Organic & Biomolecular Chemistry* **9**, 5482-5486, doi:10.1039/C1OB05406F (2011).
30 Naik, R. R. *et al.* Controlled formation of biosilica structures in vitro. *Chemical Communications*, 238-239, doi:10.1039/B210635C (2003).
31 Senior, L. *et al.* Structure and function of the silicifying peptide R5. *Journal of Materials Chemistry B* **3**, 2607-2614, doi:10.1039/C4TB01679C (2015).

32 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*, doi:10.1021/acs.langmuir.7b02868 (2017).
33 Knecht, M. R. & Wright, D. W. Functional analysis of the biomimetic silica precipitating activity of the R5 peptide from Cylindrotheca fusiformis. *Chemical Communications*, 3038-3039, doi:10.1039/B309074D (2003).
34 Lechner, C. C. & Becker, C. F. W. A sequence-function analysis of the silica precipitating silaffin R5 peptide. *Journal of Peptide Science* **20**, 152-158, doi:10.1002/psc.2577 (2014).
35 Lechner, C. C. & Becker, C. F. W. Modified silaffin R5 peptides enable encapsulation and release of cargo molecules from biomimetic silica particles. *Bioorganic & Medicinal Chemistry* **21**, 3533-3541, doi:http://dx.doi.org/10.1016/j.bmc.2013.04.006 (2013).
36 Luckarift, H. R., Spain, J. C., Naik, R. R. & Stone, M. O. Enzyme immobilization in a biomimetic silica support. *Nat Biotech* **22**, 211-213 (2004).
37 Sewell, S. L. & Wright, D. W. Biomimetic Synthesis of Titanium Dioxide Utilizing the R5 Peptide Derived from Cylindrotheca fusiformis. *Chemistry of Materials* **18**, 3108-3113, doi:10.1021/cm060342p (2006).
38 Cole, K. E., Ortiz, A. N., Schoonen, M. A. & Valentine, A. M. Peptide- and Long-Chain Polyamine- Induced Synthesis of Micro- and Nanostructured Titanium Phosphate and Protein Encapsulation. *Chemistry of Materials* **18**, 4592-4599, doi:10.1021/cm060807b (2006).
39 Andre, R., Tahir, M. N., Natalio, F. & Tremel, W. Bioinspired synthesis of multifunctional inorganic and bio-organic hybrid materials. *FEBS Journal* **279**, 1737-1749, doi:10.1111/j.1742-4658.2012.08584.x (2012).
40 Müller, W. E. G. *et al.* Silicateins, the major biosilica forming enzymes present in demosponges: Protein analysis and phylogenetic relationship. *Gene* **395**, 62-71, doi:https://doi.org/10.1016/j.gene.2007.02.014 (2007).
41 Shimizu, K., Cha, J., Stucky, G. D. & Morse, D. E. Silicatein α: Cathepsin L-like protein in sponge biosilica. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6234-6238 (1998).
42 Krasko, A. *et al.* Cathepsin, a major protease of the marine sponge Geodia cydonium: purification of the enzyme and molecular cloning of cDNA. *Mol Mar Biol Biotechnol* **6**, 296-307 (1997).
43 Sumerel, J. L. & Morse, D. E. in *Silicon Biomineralization: Biology — Biochemistry — Molecular Biology — Biotechnology* (ed Werner E. G. Müller) 225-247 (Springer Berlin Heidelberg, 2003).
44 Armirotti, A. *et al.* Primary Structure and Post-Translational Modifications of Silicatein Beta from the Marine Sponge Petrosia ficiformis (Poiret, 1789). *Journal of Proteome Research* **8**, 3995-4004, doi:10.1021/pr900342y (2009).
45 Kisailus, D., Truong, Q., Amemiya, Y., Weaver, J. C. & Morse, D. E. Self-assembled bifunctional surface mimics an enzymatic and templating protein for the synthesis of a metal oxide semiconductor. *Proceedings of the National Academy of Sciences* **103**, 5652-5657, doi:10.1073/pnas.0508488103 (2006).
46 Rimola, A., Costa, D., Sodupe, M., Lambert, J.-F. & Ugliengo, P. Silica Surface Features and Their Role in the Adsorption of Biomolecules: Computational Modeling and Experiments. *Chemical Reviews* **113**, 4216-4313, doi:10.1021/cr3003054 (2013).

47 Belton, D. J., Deschaume, O. & Perry, C. C. An overview of the fundamentals of the chemistry of silica with relevance to biosilicification and technological advances. *The Febs Journal* **279**, 1710-1720, doi:10.1111/j.1742-4658.2012.08531.x (2012).
48 Peri, J. B. & Hensley, A. L. The surface structure of silica gel. *The Journal of Physical Chemistry* **72**, 2926-2933, doi:10.1021/j100854a041 (1968).
49 Zhuravlev, L. T. The surface chemistry of amorphous silica. Zhuravlev model. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **173**, 1-38, doi:https://doi.org/10.1016/S0927-7757(00)00556-2 (2000).
50 RK, I. Chapter 1 General chemistry of silica. *Journal of Chromatography Library* **16**, 1-14, doi:http://dx.doi.org/10.1016/S0301-4770(08)60805-2 (1979).
51 Sulpizi, M., Gaigeot, M.-P. & Sprik, M. The Silica–Water Interface: How the Silanols Determine the Surface Acidity and Modulate the Water Properties. *Journal of Chemical Theory and Computation* **8**, 1037-1047, doi:10.1021/ct2007154 (2012).
52 Gröger, C., Lutz, K. & Brunner, E. Biomolecular Self-assembly and its Relevance in Silica Biomineralization. *Cell Biochemistry and Biophysics* **50**, 23-39, doi:10.1007/s12013-007-9003-2 (2008).
53 Parida, S. K., Dash, S., Patel, S. & Mishra, B. K. Adsorption of organic molecules on silica surface. *Advances in Colloid and Interface Science* **121**, 77-110, doi:https://doi.org/10.1016/j.cis.2006.05.028 (2006).
54 Patwardhan, S. V. *et al.* Chemistry of Aqueous Silica Nanoparticle Surfaces and the Mechanism of Selective Peptide Adsorption. *Journal of the American Chemical Society* **134**, 6244-6256, doi:10.1021/ja211307u (2012).
55 Puddu, V. & Perry, C. C. Peptide Adsorption on Silica Nanoparticles: Evidence of Hydrophobic Interactions. *ACS Nano* **6**, 6356-6363, doi:10.1021/nn301866q (2012).
56 Emami, F. S. *et al.* Prediction of Specific Biomolecule Adsorption on Silica Surfaces as a Function of pH and Particle Size. *Chemistry of Materials* **26**, 5725-5734, doi:10.1021/cm5026987 (2014).
57 Shrivastava, S., Nuffer, J. H., Siegel, R. W. & Dordick, J. S. Position-Specific Chemical Modification and Quantitative Proteomics Disclose Protein Orientation Adsorbed on Silica Nanoparticles. *Nano Letters* **12**, 1583-1587, doi:10.1021/nl2044524 (2012).
58 Puddu, V. & Perry, C. C. Interactions at the Silica–Peptide Interface: The Influence of Particle Size and Surface Functionality. *Langmuir* **30**, 227-233, doi:10.1021/la403242f (2014).
59 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. *Analytical Chemistry* **78**, 4872-4879, doi:10.1021/ac0603025 (2006).
60 Phillips, D. C. *et al.* Side Chain, Chain Length, and Sequence Effects on Amphiphilic Peptide Adsorption at Hydrophobic and Hydrophilic Surfaces Studied by Sum-Frequency Generation Vibrational Spectroscopy and Quartz Crystal Microbalance. *The Journal of Physical Chemistry C* **111**, 255-261, doi:10.1021/jp0645263 (2007).
61 Weidner, T., Breen, N. F., Li, K., Drobny, G. P. & Castner, D. G. Sum frequency generation and solid-state NMR study of the structure, orientation, and dynamics of polystyrene-adsorbed peptides. *Proceedings of the National Academy of Sciences* **107**, 13288-13293, doi:10.1073/pnas.1003832107 (2010).
62 Thyparambil, A. A., Wei, Y. & Latour, R. A. Determination of Peptide–Surface Adsorption Free Energy for Material Surfaces Not Conducive to SPR or QCM using AFM. *Langmuir* **28**, 5687-5694, doi:10.1021/la300315r (2012).

63 Russo, C. J. & Passmore, L. A. Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. *Nature methods* **11**, 649-652, doi:10.1038/nmeth.2931 (2014).
64 Chen, J. *et al.* Building two-dimensional materials one row at a time: Avoiding the nucleation barrier. *Science* **362**, 1135, doi:10.1126/science.aau4146 (2018).
65 Tamerler, C., Oren, E. E., Duman, M., Venkatasubramanian, E. & Sarikaya, M. Adsorption Kinetics of an Engineered Gold Binding Peptide by Surface Plasmon Resonance Spectroscopy and a Quartz Crystal Microbalance. *Langmuir* **22**, 7712-7718, doi:10.1021/la0606897 (2006).
66 Nergiz, S. Z., Slocik, J. M., Naik, R. R. & Singamaneni, S. Surface defect sites facilitate fibrillation: an insight into adsorption of gold-binding peptides on Au(111). *Physical Chemistry Chemical Physics* **15**, 11629-11633, doi:10.1039/C3CP50972A (2013).
67 Guo, X. Surface plasmon resonance based biosensor technique: A review. *Journal of Biophotonics* **5**, 483-501, doi:doi:10.1002/jbio.201200015 (2012).
68 Homola, J. Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species. *Chemical Reviews* **108**, 462-493, doi:10.1021/cr068107d (2008).
69 Green, R. J. *et al.* Surface plasmon resonance analysis of dynamic biological interactions with biomaterials. *Biomaterials* **21**, 1823-1835, doi:https://doi.org/10.1016/S0142-9612(00)00077-6 (2000).
70 Hnilova, M. *et al.* Effect of Molecular Conformations on the Adsorption Behavior of Gold-Binding Peptides. *Langmuir* **24**, 12440-12445, doi:10.1021/la801468c (2008).
71 Seker, U. O. S. *et al.* Thermodynamics of Engineered Gold Binding Peptides: Establishing the Structure–Activity Relationships. *Biomacromolecules* **15**, 2369-2377, doi:10.1021/bm4019006 (2014).
72 Seker, U. O. S. *et al.* Adsorption Behavior of Linear and Cyclic Genetically Engineered Platinum Binding Peptides. *Langmuir* **23**, 7895-7900, doi:10.1021/la700446g (2007).
73 Seker, U. O. S., Wilson, B., Sahin, D., Tamerler, C. & Sarikaya, M. Quantitative Affinity of Genetically Engineered Repeating Polypeptides to Inorganic Surfaces. *Biomacromolecules* **10**, 250-257, doi:10.1021/bm8009895 (2009).
74 So, C. R., Tamerler, C. & Sarikaya, M. Adsorption, Diffusion, and Self-Assembly of an Engineered Gold-Binding Peptide on Au(111) Investigated by Atomic Force Microscopy. *Angewandte Chemie International Edition* **48**, 5174-5177, doi:doi:10.1002/anie.200805259 (2009).
75 Sumerel, J. L. *et al.* Biocatalytically Templated Synthesis of Titanium Dioxide. *Chemistry of Materials* **15**, 4804-4809, doi:10.1021/cm030254u (2003).
76 Kharlampieva, E. *et al.* Bioenabled Surface-Mediated Growth of Titania Nanoparticles. *Advanced Materials* **20**, 3274-3279, doi:10.1002/adma.200800054 (2008).
77 Bavykin, D. V., Friedrich, J. M. & Walsh, F. C. Protonated Titanates and TiO2 Nanostructured Materials: Synthesis, Properties, and Applications. *Advanced Materials* **18**, 2807-2824, doi:10.1002/adma.200502696 (2006).
78 G®¢zquez, M. J. s., Bol®™var, J. P., Garcia-Tenorio, R. & Vaca, F. A Review of the Production Cycle of Titanium Dioxide Pigment. *Materials Sciences and Applications* **Vol.05No.07**, 18, doi:10.4236/msa.2014.57048 (2014).
79 Sang, L., Zhao, Y. & Burda, C. TiO2 Nanoparticles as Functional Building Blocks. *Chemical Reviews* **114**, 9283-9318, doi:10.1021/cr400629p (2014).
80 Kumar, S. G. & Rao, K. S. R. K. Polymorphic phase transition among the titania crystal structures using a solution-based approach: from precursor chemistry to nucleation process. *Nanoscale* **6**, 11574-11632, doi:10.1039/C4NR01657B (2014).

81 Dylla, A. G., Henkelman, G. & Stevenson, K. J. Lithium Insertion in Nanostructured TiO2(B) Architectures. *Accounts of Chemical Research* **46**, 1104-1112, doi:10.1021/ar300176y (2013).
82 Mo, S.-D. & Ching, W. Y. Electronic and optical properties of three phases of titanium dioxide: Rutile, anatase, and brookite. *Physical Review B* **51**, 13023-13032 (1995).
83 Bakardjieva, S. *et al.* Transformation of brookite-type TiO2 nanocrystals to rutile: correlation between microstructure and photoactivity. *Journal of Materials Chemistry* **16**, 1709-1716, doi:10.1039/B514632A (2006).
84 Liu, Z. *et al.* Nanostructured TiO2(B): the effect of size and shape on anode properties for Li-ion batteries. *Progress in Natural Science: Materials International* **23**, 235-244, doi:http://dx.doi.org/10.1016/j.pnsc.2013.05.001 (2013).
85 Vittadini, A., Casarin, M. & Selloni, A. Structure and Stability of TiO2-B Surfaces: A Density Functional Study. *The Journal of Physical Chemistry C* **113**, 18973-18977, doi:10.1021/jp9073009 (2009).
86 Li, W. *et al.* Highly Crystalline Mesoporous TiO2(B) Nanofibers. *The Journal of Physical Chemistry C* **118**, 3049-3055, doi:10.1021/jp408112z (2014).
87 Armstrong, A. R., Armstrong, G., Canales, J. & Bruce, P. G. TiO2-B Nanowires. *Angewandte Chemie International Edition* **43**, 2286-2288, doi:10.1002/anie.200353571 (2004).
88 Chen, X. & Mao, S. S. Titanium Dioxide Nanomaterials: Synthesis, Properties, Modifications, and Applications. *Chemical Reviews* **107**, 2891-2959, doi:10.1021/cr0500535 (2007).
89 Carp, O., Huisman, C. L. & Reller, A. Photoinduced reactivity of titanium dioxide. *Progress in Solid State Chemistry* **32**, 33-177, doi:http://dx.doi.org/10.1016/j.progsolidstchem.2004.08.001 (2004).
90 Zhang, J., Li, M., Feng, Z., Chen, J. & Li, C. UV Raman Spectroscopic Study on TiO2. I. Phase Transformation at the Surface and in the Bulk. *The Journal of Physical Chemistry B* **110**, 927-935, doi:10.1021/jp0552473 (2006).
91 Hernández-Gordillo, A., Hernández-Arana, A., Campero, A. & Vera-Robles, L. I. Biomimetic Sol–Gel Synthesis of TiO2 and SiO2 Nanostructures. *Langmuir* **30**, 4084-4093, doi:10.1021/la500203k (2014).
92 Kinsinger, N. M., Wong, A., Li, D., Villalobos, F. & Kisailus, D. Nucleation and Crystal Growth of Nanocrystalline Anatase and Rutile Phase TiO2 from a Water-Soluble Precursor. *Crystal Growth & Design* **10**, 5254-5261, doi:10.1021/cg101105t (2010).
93 Kröger, N. *et al.* Bioenabled Synthesis of Rutile (TiO2) at Ambient Temperature and Neutral pH. *Angewandte Chemie International Edition* **45**, 7239-7243, doi:10.1002/anie.200601871 (2006).
94 Dickerson, M. B. *et al.* Identification and Design of Peptides for the Rapid, High-Yield Formation of Nanoparticulate TiO2 from Aqueous Solutions at Room Temperature. *Chemistry of Materials* **20**, 1578-1584, doi:10.1021/cm071515t (2008).
95 Jiang, Y. *et al.* Biomimetic synthesis of titania nanoparticles induced by protamine. *Dalton Transactions*, 4165-4171, doi:10.1039/B802745E (2008).
96 Choi, N. *et al.* The interplay of peptide sequence and local structure in TiO2 biomineralization. *Journal of Inorganic Biochemistry* **115**, 20-27, doi:http://dx.doi.org/10.1016/j.jinorgbio.2012.05.011 (2012).
97 Puddu, V., Slocik, J. M., Naik, R. R. & Perry, C. C. Titania Binding Peptides as Templates in the Biomimetic Synthesis of Stable Titania Nanosols: Insight into the Role

of Buffers in Peptide-Mediated Mineralization. *Langmuir* **29**, 9464-9472, doi:10.1021/la401777x (2013).
98 Liu, C., Jiang, Z., Tong, Z., Li, Y. & Yang, D. Biomimetic synthesis of inorganic nanocomposites by a de novo designed peptide. *RSC Advances* **4**, 434-441, doi:10.1039/C3RA44630A (2014).
99 Grzybowski, B. A., Bishop, K. J. M., Campbell, C. J., Fialkowski, M. & Smoukov, S. K. Micro- and nanotechnology via reaction-diffusion. *Soft Matter* **1**, 114-128, doi:10.1039/B501769F (2005).
100 Chen, S.-Y. *et al.* Effects of Salinity on Oil Recovery (the "Dilution Effect"): Experimental and Theoretical Studies of Crude Oil/Brine/Carbonate Surface Restructuring and Associated Physicochemical Interactions. *Energy & Fuels* **31**, 8925-8941, doi:10.1021/acs.energyfuels.7b00869 (2017).
101 Keim, B. Alan Turing's Patterns in Nature, and Beyond. *Science* (2011).
102 Fialkowski, M., Campbell, C. J., Bensemann, I. T. & Grzybowski, B. A. Absorption of Water by Thin, Ionic Films of Gelatin. *Langmuir* **20**, 3513-3516, doi:10.1021/la036298z (2004).
103 Sarikaya, M., Tamerler, C., Schwartz, D. T. & Baneyx, F. Materials Assembly and formation using engineered polypeptides. *Annual Review of Materials Research* **34**, 373-408, doi:10.1146/annurev.matsci.34.040203.121025 (2004).
104 Walsh, T. R. & Knecht, M. R. Biointerface Structural Effects on the Properties and Applications of Bioinspired Peptide-Based Nanomaterials. *Chemical Reviews* **117**, 12641-12704, doi:10.1021/acs.chemrev.7b00139 (2017).
105 Kashiwagi, K., Tsuji, T. & Shiba, K. Directional BMP-2 for functionalization of titanium surfaces. *Biomaterials* **30**, 1166-1175, doi:https://doi.org/10.1016/j.biomaterials.2008.10.040 (2009).
106 Mieszawska, A. J., Nadkarni, L. D., Perry, C. C. & Kaplan, D. L. Nanoscale Control of Silica Particle Formation via Silk−Silica Fusion Proteins for Bone Regeneration. *Chemistry of Materials* **22**, 5780-5785, doi:10.1021/cm101940u (2010).
107 Huang, Z.-b., Shi, X., Mao, J. & Gong, S.-q. Design of a hydroxyapatite-binding antimicrobial peptide with improved retention and antibacterial efficacy for oral pathogen control. *Scientific Reports* **6**, 38410, doi:10.1038/srep38410 (2016).
108 Yi, H. *et al.* M13 Phage-Functionalized Single-Walled Carbon Nanotubes As Nanoprobes for Second Near-Infrared Window Fluorescence Imaging of Targeted Tumors. *Nano Letters* **12**, 1176-1183, doi:10.1021/nl2031663 (2012).
109 Bedford, N. M. *et al.* Sequence-Dependent Structure/Function Relationships of Catalytic Peptide-Enabled Gold Nanoparticles Generated under Ambient Synthetic Conditions. *Journal of the American Chemical Society* **138**, 540-548, doi:10.1021/jacs.5b09529 (2016).
110 Bedford, N. M. *et al.* Elucidation of Peptide-Directed Palladium Surface Structure for Biologically Tunable Nanocatalysts. *ACS Nano* **9**, 5082-5092, doi:10.1021/acsnano.5b00168 (2015).
111 Lee, Y. J. *et al.* Fabricating Genetically Engineered High-Power Lithium-Ion Batteries Using Multiple Virus Genes. *Science* **324**, 1051, doi:10.1126/science.1171541 (2009).
112 Oh, D. *et al.* Biologically enhanced cathode design for improved capacity and cycle life for lithium-oxygen batteries. *Nature Communications* **4**, 2756, doi:10.1038/ncomms3756 https://www.nature.com/articles/ncomms3756#supplementary-information (2013).
113 Nam, Y. S. *et al.* Biologically templated photocatalytic nanostructures for sustained light-driven water oxidation. *Nature Nanotechnology* **5**, 340, doi:10.1038/nnano.2010.57

https://www.nature.com/articles/nnano.2010.57#supplementary-information (2010).
114 Soto-Rodríguez, J., Hemmatian, Z., Josberger, E. E., Rolandi, M. & Baneyx, F. A Palladium-Binding Deltarhodopsin for Light-Activated Conversion of Protonic to Electronic Currents. *Advanced Materials* **28**, 6581-6585, doi:doi:10.1002/adma.201600222 (2016).
115 Hayamizu, Y. *et al.* Bioelectronic interfaces by spontaneously organized peptides on 2D atomic single layer materials. *Scientific Reports* **6**, 33778, doi:10.1038/srep33778 https://www.nature.com/articles/srep33778#supplementary-information (2016).
116 Soto-Rodríguez, 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 Expression and Purification* **135**, 70-77, doi:http://dx.doi.org/10.1016/j.pep.2017.05.003 (2017).
117 Morton, T. A., Myszka, D. G. & Chaiken, I. M. Interpreting Complex Binding Kinetics from Optical Biosensors: A Comparison of Analysis by Linearization, the Integrated Rate Equation, and Numerical Integration. *Analytical Biochemistry* **227**, 176-185, doi:https://doi.org/10.1006/abio.1995.1268 (1995).
118 Ghose, S., McNerney, T. M. & Hubbard, B. Preparative protein purification on underivatized silica. *Biotechnology and Bioengineering* **87**, 413-423, doi:doi:10.1002/bit.20125 (2004).
119 Lan, Q., Bassi, A. S., Zhu, J.-X. & Margaritis, A. A modified Langmuir model for the prediction of the effects of ionic strength on the equilibrium characteristics of protein adsorption onto ion exchange/affinity adsorbents. *Chemical Engineering Journal* **81**, 179-186, doi:https://doi.org/10.1016/S1385-8947(00)00197-2 (2001).
120 Romero-Rodríguez, A., Ruiz-Villafán, B., Rocha-Mendoza, D., Manzo-Ruiz, M. & Sánchez, S. Biochemistry and regulatory functions of bacterial glucose kinases. *Archives of Biochemistry and Biophysics* **577-578**, 1-10, doi:https://doi.org/10.1016/j.abb.2015.05.001 (2015).
121 Nakamura, T., Kashima, Y., Mine, S., Oku, T. & Uegaki, K. Characterization and crystal structure of the thermophilic ROK hexokinase from Thermus thermophilus. *J Biosci Bioeng* **114**, 150-154, doi:10.1016/j.jbiosc.2012.03.018 (2012).
122 Waterhouse, A. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**, W252-W258, doi:10.1093/nar/gku340 (2014).
123 R. Fowler, E. A. G. *Statistical Thermodynamics*. (Cambridge University Press, 1949).
124 de Keizer, A. & Fokkink, L. G. J. Specific adsorption of organic cations at the silver iodide—electrolyte interface. *Colloids and Surfaces* **51**, 323-337, doi:https://doi.org/10.1016/0166-6622(90)80150-3 (1990).
125 Al-Abadleh, H. A., Mifflin, A. L., Bertin, P. A., Nguyen, S. T. & Geiger, F. M. Control of Carboxylic Acid and Ester Groups on Chromium (VI) Binding to Functionalized Silica/Water Interfaces Studied by Second Harmonic Generation. *The Journal of Physical Chemistry B* **109**, 9691-9702, doi:10.1021/jp050782o (2005).
126 Brewer, S. H., Glomm, W. R., Johnson, M. C., Knag, M. K. & Franzen, S. Probing BSA Binding to Citrate-Coated Gold Nanoparticles and Surfaces. *Langmuir* **21**, 9303-9307, doi:10.1021/la050588t (2005).
127 Meissner, J., Prause, A., Bharti, B. & Findenegg, G. H. Characterization of protein adsorption onto silica nanoparticles: influence of pH and ionic strength. *Colloid and Polymer Science* **293**, 3381-3391, doi:10.1007/s00396-015-3754-x (2015).

128 Marx, K. A. Quartz Crystal Microbalance: A Useful Tool for Studying Thin Polymer Films and Complex Biomolecular Systems at the Solution−Surface Interface. *Biomacromolecules* **4**, 1099-1120, doi:10.1021/bm020116i (2003).
129 Latour, R. A. The langmuir isotherm: A commonly applied but misleading approach for the analysis of protein adsorption behavior. *Journal of Biomedical Materials Research Part A* **103**, 949-958, doi:doi:10.1002/jbm.a.35235 (2015).
130 So, C. R. *et al.* Controlling Self-Assembly of Engineered Peptides on Graphite by Rational Mutation. *ACS Nano* **6**, 1648-1656, doi:10.1021/nn204631x (2012).
131 Palafox-Hernandez, J. P. *et al.* Comparative Study of Materials-Binding Peptide Interactions with Gold and Silver Surfaces and Nanostructures: A Thermodynamic Basis for Biological Selectivity of Inorganic Materials. *Chemistry of Materials* **26**, 4960-4969, doi:10.1021/cm501529u (2014).
132 Tang, Z. *et al.* Biomolecular Recognition Principles for Bionanocombinatorics: An Integrated Approach To Elucidate Enthalpic and Entropic Factors. *ACS Nano* **7**, 9632-9646, doi:10.1021/nn404427y (2013).
133 Brown, S. Protein-Mediated Particle Assembly. *Nano Letters* **1**, 391-394, doi:10.1021/nl0155173 (2001).
134 Lee, S.-W., Mao, C., Flynn, C. E. & Belcher, A. M. Ordering of Quantum Dots Using Genetically Engineered Viruses. *Science* **296**, 892-895, doi:10.1126/science.1068054 (2002).
135 Sarikaya, M., Tamerler, C., Jen, A. K. Y., Schulten, K. & Baneyx, F. Molecular biomimetics: nanotechnology through biology. *Nat Mater* **2**, 577-585 (2003).
136 Seeman, N. C. & Belcher, A. M. Emulating biology: Building nanostructures from the bottom up. *Proceedings of the National Academy of Sciences* **99**, 6451-6455, doi:10.1073/pnas.221458298 (2002).
137 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* **405**, 665-668, doi:10.1038/35015043 (2000).
138 Naik, R. R., Stringer, S. J., Agarwal, G., Jones, S. E. & Stone, M. O. Biomimetic synthesis and patterning of silver nanoparticles. *Nature Materials* **1**, 169-172, doi:10.1038/nmat758 (2002).
139 Nam, K. T. *et al.* Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. *Science* **312**, 885-888, doi:10.1126/science.1122716 (2006).
140 Chiu, C.-Y. *et al.* Platinum nanocrystals selectively shaped using facet-specific peptide sequences. *Nature Chemistry* **3**, 393, doi:10.1038/nchem.1025 https://www.nature.com/articles/nchem.1025#supplementary-information (2011).
141 Habchi, J., Tompa, P., Longhi, S. & Uversky, V. N. Introducing Protein Intrinsic Disorder. *Chemical Reviews* **114**, 6561-6588, doi:10.1021/cr400514h (2014).
142 Sano, K.-I. & Shiba, K. A Hexapeptide Motif that Electrostatically Binds to the Surface of Titanium. *Journal of the American Chemical Society* **125**, 14234-14235, doi:10.1021/ja038414q (2003).
143 Mirau, P. A., Naik, R. R. & Gehring, P. Structure of Peptides on Metal Oxide Surfaces Probed by NMR. *Journal of the American Chemical Society* **133**, 18243-18248, doi:10.1021/ja205454t (2011).
144 Suzuki, Y., Shindo, H. & Asakura, T. Structure and Dynamic Properties of a Ti-Binding Peptide Bound to TiO2 Nanoparticles As Accessed by 1H NMR Spectroscopy. *The Journal of Physical Chemistry B* **120**, 4600-4607, doi:10.1021/acs.jpcb.6b03260 (2016).

145 Goobes, G. *et al.* Folding of the C-terminal bacterial binding domain in statherin upon adsorption onto hydroxyapatite crystals. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16083-16088, doi:10.1073/pnas.0607193103 (2006).
146 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 (London, England : 1993)* **18**, 1678-1687, doi:10.1016/j.str.2010.09.013 (2010).
147 Arachchige, R. J. *et al.* Solid-State NMR Identification of Intermolecular Interactions in Amelogenin Bound to Hydroxyapatite. *Biophysical Journal* **115**, 1666-1672, doi:https://doi.org/10.1016/j.bpj.2018.08.027 (2018).
148 Shaw, W. J. Solid-state NMR studies of proteins immobilized on inorganic surfaces. *Solid State Nuclear Magnetic Resonance* **70**, 1-14, doi:https://doi.org/10.1016/j.ssnmr.2014.10.003 (2015).
149 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* **35**, 5013-5020, doi:10.1021/acs.langmuir.9b00283 (2019).
150 Bradley, P., Misura, K. M. S. & Baker, D. Toward High-Resolution de Novo Structure Prediction for Small Proteins. *Science* **309**, 1868-1871, doi:10.1126/science.1113801 (2005).
151 Bhardwaj, G. *et al.* Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329, doi:10.1038/nature19791 https://www.nature.com/articles/nature19791#supplementary-information (2016).
152 Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19-25, doi:https://doi.org/10.1016/j.softx.2015.06.001 (2015).
153 Mackerell Jr., A. D. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry* **25**, 1584-1604, doi:10.1002/jcc.20082 (2004).
154 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* **29**, 1754-1765, doi:10.1021/la3038846 (2013).
155 Berendsen H. J. C.; Postma, J. P. M. v. G., W. F.; Hermans, J. . Interaction Models for Water in Relation To Protein Hydration. *Intermolecular Forces* (1981).
156 Daura, X. *et al.* Peptide Folding: When Simulation Meets Experiment. *Angewandte Chemie International Edition* **38**, 236-240, doi:10.1002/(sici)1521-3773(19990115)38:1/2<236::Aid-anie236>3.0.Co;2-m (1999).
157 Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33-38, doi:https://doi.org/10.1016/0263-7855(96)00018-5 (1996).
158 Bonomi, M. *et al.* PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961-1972, doi:https://doi.org/10.1016/j.cpc.2009.05.011 (2009).
159 Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126**, 014101, doi:10.1063/1.2408420 (2007).
160 Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **18**,

1463-1472, doi:10.1002/(sici)1096-987x(199709)18:12<1463::Aid-jcc4>3.0.Co;2-h (1997).
161 Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **52**, 7182-7190, doi:10.1063/1.328693 (1981).
162 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, 10089-10092, doi:10.1063/1.464397 (1993).
163 Brooks, B. R. *et al.* CHARMM: The Biomolecular Simulation Program. *Journal of computational chemistry* **30**, 1545-1614, doi:10.1002/jcc.21287 (2009).
164 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*, doi:10.1021/acs.langmuir.9b00283 (2019).
165 Baio, J. E. *et al.* Diatom Mimics: Directing the Formation of Biosilica Nanoparticles by Controlled Folding of Lysine-Leucine Peptides. *Journal of the American Chemical Society* **136**, 15134-15137, doi:10.1021/ja5078238 (2014).
166 Lutz, H. *et al.* Biomimetic Growth of Ultrathin Silica Sheets Using Artificial Amphiphilic Peptides. *Advanced Materials Interfaces* **2**, 1500282, doi:10.1002/admi.201500282 (2015).
167 Daly, S. M., Przybycien, T. M. & Tilton, R. D. Coverage-Dependent Orientation of Lysozyme Adsorbed on Silica. *Langmuir* **19**, 3848-3857, doi:10.1021/la026690x (2003).
168 Asanov, A., Delucas, L., B. Oldham, P. & W. Wilson, W. *Heteroenergetics of Bovine Serum Albumin Adsorption from Good Solvents Related to Crystallization Conditions*. Vol. 191 (1997).
169 Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I. M., Verkhusha, V. V. & Turoverov, K. K. Beta-Barrel Scaffold of Fluorescent Proteins: Folding, Stability and Role in Chromophore Formation. *International review of cell and molecular biology* **302**, 221-278, doi:10.1016/B978-0-12-407699-0.00004-2 (2013).
170 Stepanenko, O. V., Stepanenko, O. V., Kuznetsova, I. M., Verkhusha, V. V. & Turoverov, K. K. Sensitivity of Superfolder GFP to Ionic Agents. *PLoS ONE* **9**, e110750, doi:10.1371/journal.pone.0110750 (2014).
171 Collins, J. M., Uppal, R., Incarvito, C. D. & Valentine, A. M. Titanium(IV) Citrate Speciation and Structure under Environmentally and Biologically Relevant Conditions. *Inorganic Chemistry* **44**, 3431-3440, doi:10.1021/ic048158y (2005).
172 Kefalas, E. T. *et al.* Mononuclear Titanium(IV)−Citrate Complexes from Aqueous Solutions: pH-Specific Synthesis and Structural and Spectroscopic Studies in Relevance to Aqueous Titanium(IV)−Citrate Speciation. *Inorganic Chemistry* **44**, 2596-2605, doi:10.1021/ic049276o (2005).
173 Mudunkotuwa, I. A. & Grassian, V. H. Histidine Adsorption on TiO2 Nanoparticles: An Integrated Spectroscopic, Thermodynamic, and Molecular-Based Approach toward Understanding Nano–Bio Interactions. *Langmuir* **30**, 8751-8760, doi:10.1021/la500722n (2014).
174 Hansen, R., Bruus, H., Callisen, T. H. & Hassager, O. Transient Convection, Diffusion, and Adsorption in Surface-Based Biosensors. *Langmuir* **28**, 7557-7563, doi:10.1021/la3000763 (2012).
175 Lebedev, K., Mafé, S. & Stroeve, P. Convection, diffusion and reaction in a surface-based biosensor: Modeling of cooperativity and binding site competition on the surface

and in the hydrogel. *Journal of Colloid and Interface Science* **296**, 527-537, doi:https://doi.org/10.1016/j.jcis.2005.09.032 (2006).
176 Sadana, A. & Sii, D. Binding kinetics of antigen by immobilized antibody: influence of reaction order and external diffusional limitations. *Biosensors and Bioelectronics* **7**, 559-568, doi:https://doi.org/10.1016/0956-5663(92)85007-W (1992).
177 He, S., Lam, A. T., Jeune-Smith, Y. & Hess, H. Modeling Negative Cooperativity in Streptavidin Adsorption onto Biotinylated Microtubules. *Langmuir* **28**, 10635-10639, doi:10.1021/la302034h (2012).
178 Persad, A. H. & Ward, C. A. Expressions for the Evaporation and Condensation Coefficients in the Hertz-Knudsen Relation. *Chemical Reviews* **116**, 7727-7767, doi:10.1021/acs.chemrev.5b00511 (2016).
179 Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K. & Tsien, R. Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral. *Proceedings of the National Academy of Sciences* **97**, 11990, doi:10.1073/pnas.97.22.11990 (2000).