Molecular Modelling for Tunable Sequence-Structure Design of Bioinspired Materials

Kaylyn Torkelson

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2024

Reading Committee:
Jim Pfaendtner, Chair
Lilo Pozzo
Elizabeth Nance

Program Authorized to Offer Degree:
Chemical Engineering

University of Washington

Abstract

Molecular Modelling for Tunable Sequence-Structure Design of Bioinspired Materials

Kaylyn Torkelson

Chair of the Supervisory Committee:
Jim Pfaendtner
Department of Chemical Engineering

Biological sequence-defined polymers, such as proteins, have inspired a wide range of synthetic sequence-defined polymers that can be exploited for applications in fields including drug delivery, biomedicine, catalysis, and novel materials development. These synthetic polymers show great opportunity for precise control of single-chain secondary and tertiary structure, multi-unit self-assembly, macromolecular hierarchical design, and overall function through careful tuning of sequence. However, the relationship between sequence, structure, and function in these molecules is often not well understood. Computational modelling in this sphere provides a great opportunity to gain nanoscale-level insight on these systems, understand the driving forces relating molecular sequence, structure, and larger function, and exploit these features for better sequence design. In this work, physics-based simulation models are used to analyze these sequence-structure-function relationships in peptoids and peptides. The knowledge gained from these simulations provides frameworks for understanding the folding of peptoids into helical hairpins as well as interfacial interactions to design better polymers for applications in biomimetic mineralization.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Graduate school can be pretty rough—it can be hard to look at any one person's isolated experience and say that one defines normal, but I definitely felt that my experience was unconventional. From a global pandemic that shut down the world in my first year to an advisor move in my last, there were many challenges that led to both personal and professional growth in this time period, for which I am deeply grateful. I was extremely lucky to have a support system that helped me through any challenges and colored my time with such positive memories in between. I would like to thank the people below that have left the biggest impact on me (and thus, my work) in the past few years.

To my family first, thank you for the YEARS of support it took to get me into graduate school in the first place, and all the support you have shown since. This accomplishment is for you.

To Jim, thank you for being the mentor I didn't know how to ask for. You told my cohort in 2019 that graduate school was the one chance you have to pick your boss, so choose who you want to work for wisely; you have since spent the last five years validating the choice I made back then. Thank you for teaching me everything you can about this field, being patient with my coding skills, running happy hours to keep the group together during COVID, encouraging me to apply for internships, handling my job-related meltdowns, standing up for me whenever I needed it, providing constant feedback, and encouraging my growth as a researcher and a person. I could not imagine going through this experience with anyone else.

To PRG 2020-2023 (and those left at UW that are PRG in spirit), thank you for being the best research group I could have possibly asked for. I learned so much from so many people in these five years, all while having the most fun doing it. I will forever be grateful for lunches after group meeting, walks to LSB for the good water, happy hours at Bad Jimmy's, and the countless times you all let me derail the entire office from working. It takes an army of great people to make a windowless basement feel like home, but I will definitely miss B17 because of you and the memories we shared there.

To Zach, thank you for being my partner in crime in Seattle, introducing me to every worthwhile brunch restaurant in the city, jamming in the car to Chappell Roan, sharing my love of cider, letting me tag along everywhere with your friends, drying my tears, and teaching me how fun trivia can be. I didn't know what a best friend was until I met you.

To Will, thank you for managing my chaos when needed and matching my energy the rest of the time. You have been a constant source of good advice, great friendship, and infectious laughter since we met, and I cannot wait for the future adventures we get ourselves into (or the future adventures I sign you up for that you begrudgingly agree to).

To Jaime, thank you for introducing me to the city in 2019, always being down for taquitos on the patio in the sun, showing me the best hikes and taco spots in the PNW, sharing my love of film photography, wiener dogs, and margaritas, and being such a reliable and supportive friend. I am grateful to have people in my life that make me feel so understood.

To Ben, thank you for going through this process with me, listening to my ridiculous stories, showing me around Austin whenever I show up there on a whim, and always being down for hiking trips in Sedona or camping trips in the PNW. I am so glad we will finally get to be in similar regions of the US going forward; we have certainly come a long way from UK ChemE and growing up in Kentucky.

To Jason (& Erika now too!), thank you for helping me through my first PhD recruitment weekend, frequenting half-price nacho Monday at College Inn enough that they recognize us when we come in, feeding me some of the best food I have ever had (and teaching me how to make it), providing advice on every aspect of life or graduate school or the job search, and for consistently showing me my worth. The capacity you both demonstrate for care, consideration, compassion, friendship, and support to all of the people in your life is inspiring and I cannot wait to buy you dinner sometime to thank you for that.

To Tucker, thank you for your professional mentorship, your much-used code-compiling skills (lol), and always being more impressed with me than I am with myself. You have encouraged so much self-confidence within me; your friendship has meant the world to me, and I am so excited to see what your future holds.

To everyone I have met, loved, left, cherished, or lost in the past five years: thanks for the serendipity, great memories, wild stories, challenges, support, growth, and encouragement. I wouldn’t be me without all of you.

Chapter 1: Introduction

Biological sequence-defined polymers, such as proteins, are critical for functions within natural environments including signaling, growth, providing structure, maintaining pH, transporting nutrients, and more [1]. Understanding the mechanism of a protein with one function from another relies on understanding the structure and fold of these biomolecules, which is defined from the sequence of each protein, even though these molecules are built from the same twenty candidates of amino acids [2,3]. Much work in the field of protein modelling and design has focused on understanding these sequence-structure-function relationships to develop synthetic proteins that can be highly specialized for improved stability, binding specificity, or interfacial interactions through modification of primary sequence and structure [4–10]. The field of synthetic sequence-defined polymers often relies on motifs common in these natural sequence-defined polymers, taking inspiration from protein or peptide sequences and structures with particular applications as a starting point for designing better catalysts, drug delivery systems, or novel materials [11–19]. One of the most recent additions to strengthen the field of bioinspired design is a class of materials known as peptoids, or N-substituted glycines, which are similar in backbone to naturally occurring peptides, but have sidechains connected to the backbone nitrogen instead of the backbone alpha carbon [20–27].

The primary draw of peptoids in this field is the diversity of available sidechains. While peptides and proteins are limited to the twenty native amino acids as building blocks, peptoids can be built from any monomer that can be synthesized from a tertiary amine without losing the biocompatibility or high immunogenicity we see in natural systems [28]. However, there are strong differences between peptoids and their peptide counterparts, namely in the relationship between sequence and structure. Peptoids are still capable of folding into complex secondary, tertiary, and quaternary structures based on sequence, but the driving forces behind peptoid folding are very different than those that govern peptide and protein-folding [29–33]. While peptides and proteins adopt secondary structures like alpha helices due to hydrogen bonding between the carbonyl oxygen and the hydrogens present in later sequence residues, the change in sidechain connectivity in peptoids prevents these hydrogen bonds from forming. Instead, peptoids are capable of adopting protein-like helices through the interactions between bulky, often aromatic sidechains [34–39]. This provides a great opportunity to develop synthetic sequence-defined polymers for specific applications based on careful tuning of sequence. The placement and order of monomers with particular secondary, tertiary, and quaternary structure driving traits could be the key to developing synthetic polymers with enzymatic functionality, or to better controlling the morphology of inorganic materials formed from organic template molecules.

The sequence-structure relationships that we are so interested in with peptides and peptoids have strong implications in numerous fields, including the field of biomimetic mineralization. Naturally, biomineralization is a process that occurs in the cell walls of diatoms, unicellular photosynthetic organisms that use silaffin as a primary protein to construct complex hierarchical silica nanostructures. This process can be mimicked in a laboratory setting through the use of the R5 peptide, one of the main unstructured repeat units in the silaffin protein, but without the morphological control of the natural process [40]. These biomimetic silica nanostructures have received interest lately for applications in catalysis, separations, biosensing, and drug delivery, but achieving a full mechanistic understanding of silica morphology necessary for these applications and the particular side chain chemistry, molecular structure, aggregate structure, concentration of precursor or molecule, and solution conditions required to create such a morphology proves difficult [41–45]. Computational modelling presents a great opportunity for understanding these nanoscale interactions in order to tune molecular sequence, structure, and inevitably function. This work seeks to advance the state of the art in molecular modelling to understand these sequence-structure relationships and provide frameworks for developing synthetic polymers with enzymatic-mimicking functions or peptides and peptoids of defined sequence that can create specific silica morphologies.

In chapter two, we develop forcefield parameters for peptoid monomers with varying levels of chiral, hydrophobic, or charged characteristics and devise a simulation method that allows us to relate polymer sequence to single-chain secondary and tertiary structure in helical hairpins that mimic native protein conformations. These first atomistic molecular dynamics simulations of peptoids are used to discuss the important structural features of these polymers driven by their primary sequence in both water and acetonitrile and provide a future-look at how we might design more complex peptoids with enzymatic-mimicking structures and functions in the future.

In chapter three, we highlight the use of bioinspired peptoids for modulating silica morphogenesis and compare the produced silica nanostructures between peptoids and peptides with similar sequences. We are able to predict a specific silica morphology by looking at a key computational marker established from simulation and can then use the structural information from these simulations to explain the characteristics of a sequence and its interfacial interactions with silica that may lead to this specific morphology. This provides us a basis for computational understanding of biomimetic mineralization using sequence defined polymers for the purposes of rational design of novel materials.

In chapter four, we build on our key computational marker identified in chapter three to provide a comprehensive study of silica binding peptides, their binding energy on amorphous silica or ordered quartz surfaces, and the silica morphologies they produce. We identify the experiments necessary to validate our method and improve understanding of the relationship between the concentration of biomolecule required and resulting silica morphology as well as offer key insight on how to speed these calculations in the future at a lower computational cost.

Contributions have also been made to additional in-progress work that builds on the foundations mentioned here, primarily forcefield construction for peptoid systems and enhanced sampling methods for understanding interfacial interactions pertinent to biomimetic mineralization. These projects include a collaboration with the Ferguson group at the University of Chicago focused on the collection of peptoid forcefield parameters to enable the high-throughput computational screening of peptoids to speed rational design of novel materials, as well as a project focused on peptide binding to facets of calcium phosphate to understand biomimetic mineralization outside of solely silica-based materials.

Chapter 2: Exploration of Tertiary Structure in Sequence-Defined Polymers Using Molecular Dynamics Simulations[1]

2.1 Abstract

Peptoids are a class of sequence-defined biomimetic polymers with peptide-like backbones and side chains located on backbone nitrogens rather than alpha carbons. These materials demonstrate strong ability for precise control of single-chain structure, multi-unit self-assembly, and macromolecular assembly through careful tuning of sequence due to the diversity of available sidechains, although the driving forces behind these assemblies are often not understood. Prior experimental work has shown that linked 15mer peptoids can mimic the protein helical hairpin structure by leveraging the chirality-inducing nature of bulky side chains and hydrophobicity, but there are still gaps in our understanding of the relationship between sequence, stability and particular secondary or tertiary structure. We present a molecular dynamics (MD) study on the folding behavior of these polymers into hairpins, discussing the differences in structure from sequences with various characteristics in water and acetonitrile and then compare the handedness preference of common helical motifs between solvents.

2.2 Introduction

Biological sequence-defined polymers, such as proteins or DNA, have inspired a wide range of synthetic sequence-defined polymers that can be exploited for applications in everything from biomedicine to novel materials development [46]. One such class of biomimetic sequence-defined polymers is peptoids, or N-substituted glycines, which mimic the backbone of peptides, but with the sidechain of these synthetic building blocks connecting to the backbone nitrogen instead of the backbone alpha carbon [20]. Peptoids have demonstrated a number of desirable properties, including high biocompatibility, immunogenicity, and antimicrobial and antifouling behavior [47–51]. Peptoids can also be synthesized from a larger collection of tertiary amines than the traditional peptide amino acids, allowing for more diverse sequences that will hopefully inspire more complex structures and functions. These materials, with this vast design space of monomer

---

[1] This chapter was reproduced from an under-review draft submitted to Biomacromolecules with permission from K. Torkelson and J. Pfaendtner.

building blocks, present great opportunity for the precise control of sequence, single chain secondary and tertiary structure, multi-unit self-assembly, and larger macromolecular assemblies [21,22,30,52–56].

While peptides and proteins are known to adopt secondary structures such as alpha-helices due to the presence of hydrogen bonding between oxygen and nitrogen atoms in the backbone, moving the functional group to the nitrogen as in peptoids removes the ability of this bond to form. However, these polymers are still capable of folding into a number of protein-like secondary structures, such as helices or beta sheets, due to the careful selection and placement of side chains [33,52,53]. Thus, a number of studies have focused on identifying side chains that induce stabilization of secondary structures in these materials. Early work in this field demonstrated that a bulky aromatic sidechain with a chiral center, such as Nspe, could induce helical folding in a peptoid that resembled a polyproline type-I helix, as reported by circular dichroism measurements and supported by molecular mechanics calculations [33,53]. Inverting the chiral center to obtain Nrpe was shown to induce helical folding in a peptoid to resemble a polyproline type-I helix of the opposite handedness compared to its Nspe counterpart [32], but further work has been needed to quantify the driving forces for the assembly of these peptoids into specific helices with a degree of homogeneity. The handedness preference of Nspe or Nrpe sidechains has been investigated more fully with DFT, molecular dynamics, and enhanced sampling methods [34–36], and such work has demonstrated that the free energy driving forces for the assembly of these peptoids into specific helices remains relatively small. However, new sidechain chemistries, including napthylethyl, n-aryl, and tert-butyl sequences, have also been identified as helical-forming sidechains and may show a much stronger driving force towards homogenous secondary structure in water [37–39].

While a large collection of work has focused on the creation of helical secondary structures, peptoid macrocycles, or larger macromolecular assemblies like nanosheets or nanotubes [21,22,30,55,57], far fewer researchers have focused on mimicking the common (yet more complex) protein-motif of a hairpin or helical bundle. Lee et al developed the first example of a peptoid helical bundle based on an amphiphilic 15mer peptoid that was shown to self-assemble in solution [54]. By chaining multiple 15mer units together with disulfide linkages, they were able to form structures that folded in aqueous environments due to stable hydrophobic cores and then confirmed the compactness of these structures through circular dichroism and the quenching of fluorescence resonance energy transfer (FRET) reporter groups. While Lee et. al continued to build upon the idea of helical peptoid bundles by introducing a zinc-binding functionality to previously identified sequences to demonstrate the ability of these sequences to mimic enzymatic behavior [58], little work has followed since on these specific secondary and tertiary structures as well as their possible biomimetic functions, especially from a computational standpoint. However, further work in the field of tertiary peptoid structures does exist, primarily focusing on the hydrophobic sequence patterning of monomers without any chiral or helix-inducing characteristic. This idea was first investigated through small angle X-ray scattering and dynamic light scattering data from the coil-to-globule transition of two polypeptoid sequences with different patterning, creating a model that can describe structural transitions based solely on hydrophobic forces [59]. Recently, we have seen renewed interest in understanding how the pattern of hydrophobic monomers in a larger polypeptoid sequence can change the compact structure that is formed in aqueous solution, with such studies utilizing coarse-grained simulations to probe the effect that hydrophobic placement has on the end-to-end distance of these polymers [60].

Compared to the advanced field of protein structure modeling, the field of peptoid modeling is still in relative infancy. Given the strong impact different monomers can have on the flexibility, structural characteristics, and self-assembly behavior of peptoid systems, further computational work would greatly aid the rational design of peptoid sequences with secondary and tertiary structure for tunable functions, as we are currently seeing in the investigation of larger macromolecular assemblies like nanotubes and nanosheets. A number of groups have used both atomistic and coarse-grained simulations to predict the properties, self-assembly pathways, structure, and stability of peptoid nanosheets or peptoid nanotubes in order to design more interesting hierarchical materials [29,30,61]. While recent advances in coarse-grained parameterization and simulation methods have allowed for the simulation of peptoids on the larger mesoscale [62,63], the investigation of secondary and tertiary structure in these systems must rely heavily on atomistic dynamics. Computational efforts in this field have largely been slowed down by the need to develop atomistic force field parameters for every new peptoid residue, due to the lack of off-the-shelf options limited to the amino acids so regularly used in the protein community. While much work has focused on properly capturing the flexibility of the peptoid backbone [36,64,65], papers that contain large numbers of simulated side chains at atomistic resolution are much less commonly found. With this work, we hope to leverage some of the recent advances in the development of peptoid force fields in order to add to the limited knowledge about the key driving forces that lead to higher order peptoid structures.

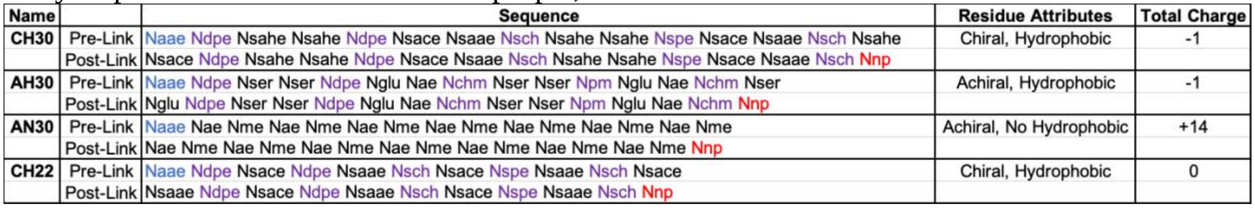
The rest of this manuscript is organized as follows. First we perform, to our knowledge, the first MD simulations of peptoid helical bundles or helical hairpins based on four sequences identified by Lee et al with various residue characteristics: a sequence with both chiral (helix-inducing) and hydrophobic residues, a sequence with achiral and hydrophobic residues, a sequence with no hydrophobic residues, and a sequence with chiral residues and smaller spacing between hydrophobic residues. We use steered molecular dynamics to establish initial configurations of each of these sequences, allow them to equilibrate, and then perform ensembles of long-time unrestrained MD simulations. To understand structural properties of these systems we use clustering, among other analyses, to discuss the features of each resulting structure in both water and acetonitrile. Next, building on our prior work analyzing thermodynamic stabilities of helical handedness with chiral sidechains, we perform an in-depth comparison of the handedness preference of one sequence with helix-inducing chiral side chains in both water and acetonitrile, exploring the stability and features of a helical bundle composed of left-handed/left-handed helices, left-handed/right-handed helices, right-handed/right-handed helices, or right-handed/left-handed helices in comparison to the same sequence with no helical character.

2.3 Methods

Simulation Setup

The sequences for each peptoid of interest in this study were taken from Lee et al [54] and are shown in Table 2.1 below, along with important attributes of the residues and the overall charge of each peptoid. The individual residues in each sequence are drawn in Figure 2.1. The four sequences chosen were selected for having a range of chiral or achiral residues, hydrophobic or not hydrophobic residues, and peptoid lengths. In the original study, fluorescence resonance energy transfer reporter groups were used to determine the distance between the N-terminal and C-terminal of each sequence, and these FRET labels (Naae and Nnp) have been included here in order to reproduce each system as accurately as possible.

Table 2.1: Table detailing the sequence of each peptoid of interest in this study before the linker and after the linker, important attributes of each residue, and the total charge of each sequence. Hydrophobic residues are shown in purple, while the FRET labels are shown in blue and red.



Name		Sequence	Residue Attributes	Total Charge
CH30	Pre-Link	Naae Ndpe Nsahe Nsahe Ndpe Nsace Nsaae Nsch Nsahe Nsahe Nspe Nsace Nsaae Nsch Nsahe	Chiral, Hydrophobic	-1
	Post-Link	Nsace Ndpe Nsahe Nsahe Ndpe Nsace Nsaae Nsch Nsahe Nsahe Nspe Nsace Nsaae Nsch Nnp		
AH30	Pre-Link	Naae Ndpe Nser Nser Ndpe Nglu Nae Nchm Nser Nser Npm Nglu Nae Nchm Nser	Achiral, Hydrophobic	-1
	Post-Link	Nglu Ndpe Nser Nser Ndpe Nglu Nae Nchm Nser Nser Npm Nglu Nae Nchm Nnp		
AN30	Pre-Link	Naae Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme	Achiral, No Hydrophobic	+14
	Post-Link	Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme Nae Nme Nnp		
CH22	Pre-Link	Naae Ndpe Nsace Ndpe Nsaae Nsch Nsace Nspe Nsaae Nsch Nsace	Chiral, Hydrophobic	0
	Post-Link	Nsaae Ndpe Nsace Ndpe Nsaae Nsch Nsace Nspe Nsaae Nsch Nnp		

Each peptoid residue was built in GaussView [66] and the GROMACS [67] insert-molecule function was used to create each peptoid chain. These peptoids were then solvated in a 10 nm box with either water or acetonitrile, and then with either sodium or chlorine ions to balance charge. Each residue is shown in Figure 2.1 below, including FRET labels and the linker residue.

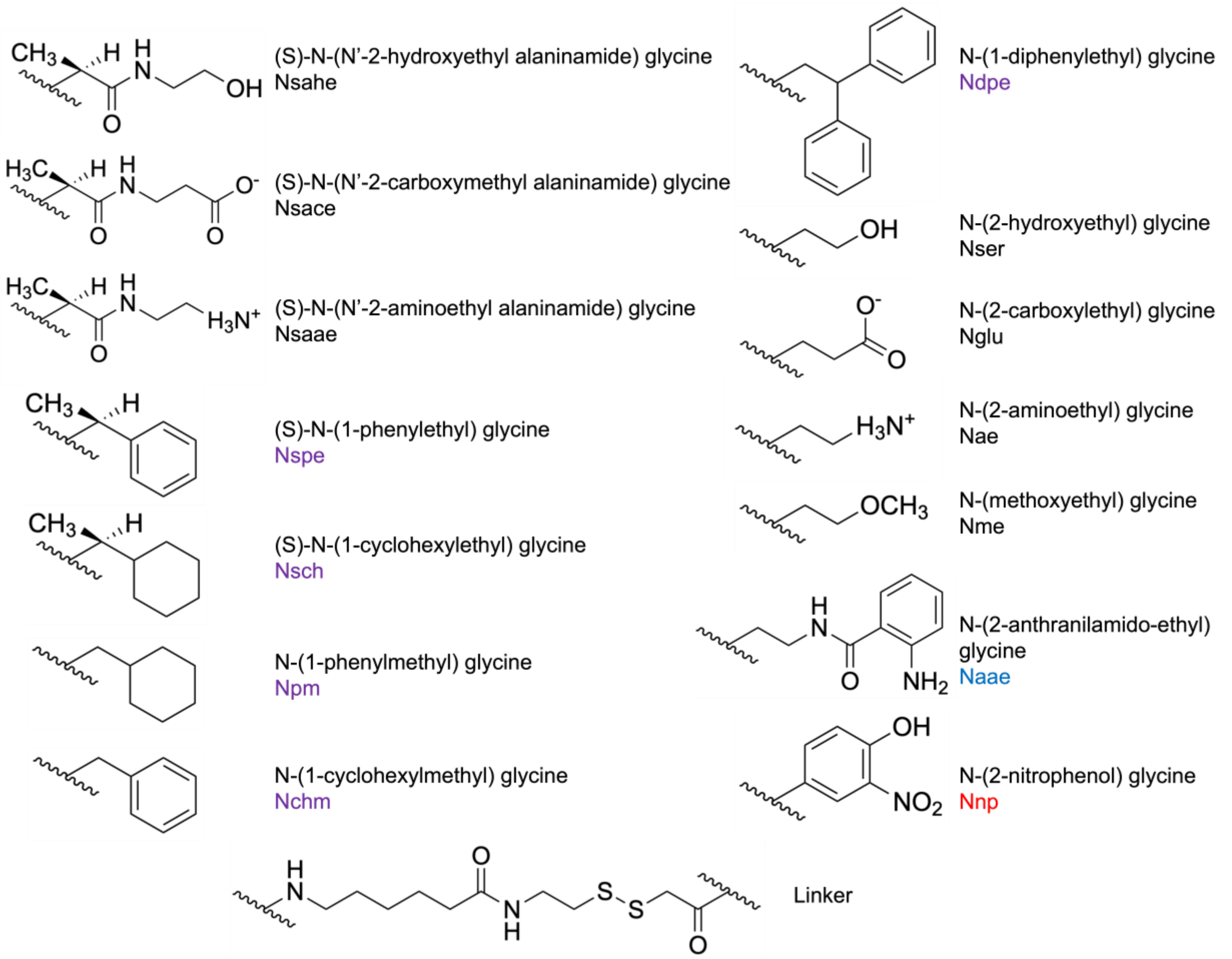


Figure 2.1: Reproduced from Lee et al, details the chemical structures of all side chains used with their full name and abbreviation. All hydrophobic sidechains have been abbreviated in purple, while the two FRET tags are shown in blue and red. The linker residue is shown at the bottom.

All systems were modeled using the CHARMM36 force field [68,69]. Parameters for the peptoid backbones were taken from Weiser and Santiso and the MFTOID forcefield [64,65], while the sidechain, linker, and FRET label parameters were taken from the CGenFF parameter generator [70] and comparison to existing peptide sidechain parameters, as mentioned in previous works [71]. The TIP3P water model was used for all water systems [72], while the acetonitrile forcefield parameters were taken from available off-the-shelf CHARMM forcefields for all systems in acetonitrile [68].

All simulations were performed in GROMACS [67]. Hydrogen mass repartitioning was employed so that a timestep of 4 femtoseconds could be used for all production steps [73]. Energy minimization was carried out for a maximum of 10 picoseconds using the steepest descent algorithm, followed by NPT equilibration at 300 K and 1 bar for 500 picoseconds using the Berendsen barostat with a pressure damping parameter of 1.0 and a stochastic global thermostat with a temperature damping parameter of 0.1 [74]. Lennard Jones and Coulombic interactions were calculated with a cutoff of 1.0 nm and long-range electrostatic interactions were treated with the Particle Mesh Ewald method. After the NPT step, the final box size for all water simulations was 9.96 nm and for all acetonitrile simulations was 10.19 nm. All simulations underwent a 100 nanosecond NVT equilibration period, from which five initial starting configurations were chosen to be the starting structures for the five trials of each system.

Collective Variables for Describing and Equilibrating Structures

Multiple collective variables were defined in this study to compare resulting structures as well as equilibrate systems. The first collective variable defined, the simulated FRET distance, was used to measure the distance between the Naae and Nnp FRET labels on each system for each trial in order to be compared with experimental values of FRET distance listed in Lee et al. A high simulated FRET distance would indicate an extended structure, while a low simulated FRET distance would indicate a more collapsed structure, as shown in Figure 2.2 below. A coordination number collective variable was also implemented for each system as a proxy for determining how close certain residues were interacting (i.e., if two residues were close enough to be interacting, those residues would be defined as “coordinated”). For systems with hydrophobic residues (including CH30, AH30, and CH22), the center of mass of each hydrophobic residue was defined and then the coordination number between any of these residues on the pre-linker section of the sequence was calculated with any of these residues on the post-linker section of the sequence. For the system without hydrophobic residues (AN30), the center of mass of every third residue in the sequence was defined and the coordination number between any of these residues on the pre-linker section of the sequence was calculated with any of these residues on the post-linker section of the sequence. These residues were chosen to be consistent with the spacing of hydrophobic residues in the two sequences of the same length, CH30 and AH30. As shown in Figure 2.2 below, a collapsed system with a high degree of interaction between hydrophobic residues (necessary for the formation of a stable hydrophobic core) would have a high coordination number, while a system with a low degree of interaction (and thus likely extended) would have a low coordination number. The PLUMED plug-in for enhanced sampling was used to define these variables and further information about the specific formulas used can be found in the PLUMED manual [75].

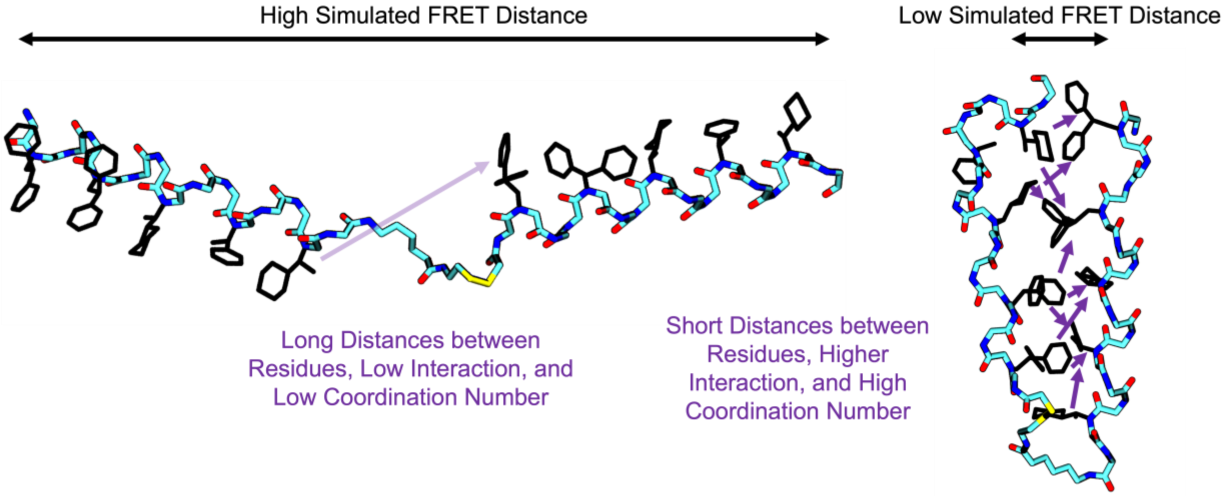


Figure 2.2: Simulation snapshots detailing a low versus high simulated FRET distance and a low versus high coordination number.

The ToidRMSD collective variable was used to equilibrate systems in the steering protocols mentioned in the next section. This collective variable was designed based on a previously established collective variable used to explore protein structure [76] but was adapted for peptoid structure and has been previously defined in Alamdari et al [35]. Within this collective variable, shown in Equation 2.1 below, the RMSD between the backbone nitrogen, the first carbon of the sidechain, the backbone alpha carbon, the carbonyl carbon, and the carbonyl oxygen of each

3-residue block of the peptoid of interest and a set reference structure is measured. Reference structures are constructed to match the dihedral angles and atom positions of each peptoid conformational minima; within this work, the helical minima alpha D plus, corresponding to a left-handed helix, was primarily used as the reference. A sigmoidal function ranging from 0 to 1 with values of n = 8, m = 12, and r0 = 0.05 nm is then used to determine the deviation of each 3-residue segment along the peptoid sequence with that chosen reference structure. This value is then summed over all the 3-residue blocks within the sequence. A larger value of ToidRMSD thus corresponds to a more helical system with more 3-residue blocks that closely resemble a helix, while a low value of ToidRMSD corresponds to a less helical system with fewer 3-residue blocks that closely resemble a helix. In this work, the pre-linker and post-linker sections of each sequence were defined separately so they could both be steered to the highest, most-helical value of peptoid RMSD without interference from the linker. In the case of the longer sequences (CH30, AH30, and AN30), the maximum value of ToidRMSD is 13 for a perfectly helical structure, while sequences that have a smaller number of 3-residue units with helical character will have a lower value of ToidRMSD.

$$ToidRMSD = \sum_i \frac{1 - \left(\frac{RMSD_i}{r_0}\right)^n}{1 - \left(\frac{RMSD_i}{r_0}\right)^m} \qquad (2.1)$$

Simulation Protocols for Structural Studies

This study combines steered MD simulations for establishing initial structures (e.g., two helical bundles connected by a linker) with classical MD simulations to study formation of hairpin structures and the long-term structural evolution of helical hairpins. Our overall approach is described below in Figure 2.3 with detailed explanation of each phase following:

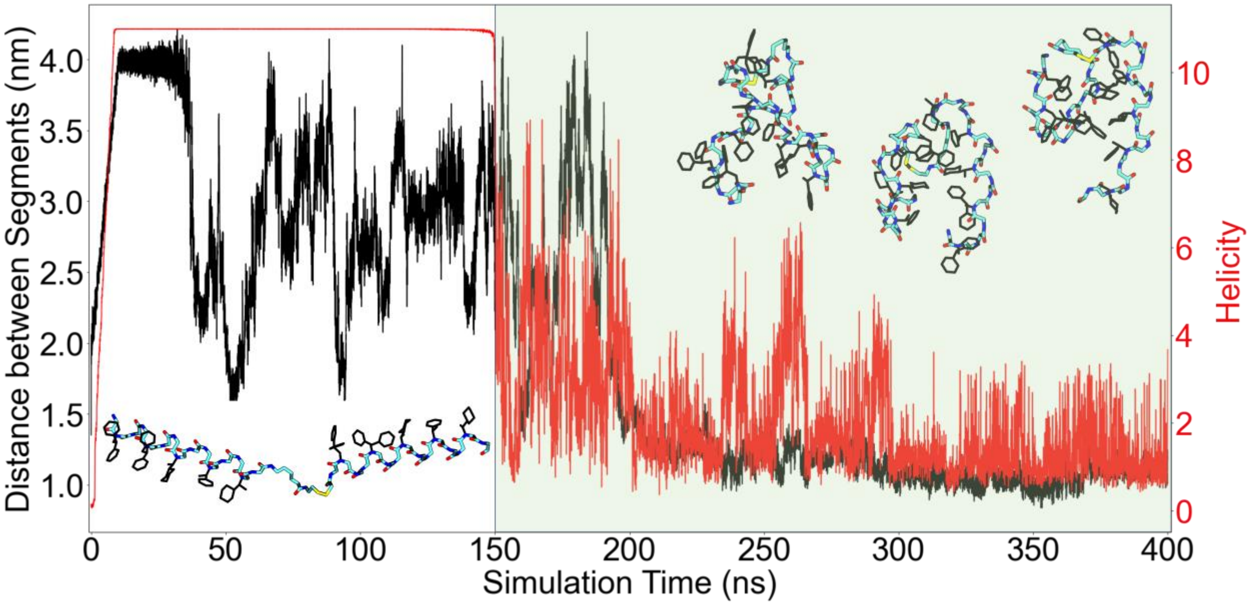


Figure 2.3: Detailed steering protocol diagram. The distance restraint is initialized early in the simulation to push the two halves of the peptoid apart while also steering each segment into a helix, leading to the representative structure on the bottom left. Around 50 nanoseconds, the distance restraint is slowly lowered, allowing the structure to collapse. At around 150 nanoseconds, the helical restraint is turned off, allowing the secondary structure to relax. There are then 250 nanoseconds of unrestrained production simulation (indicated with green overlay), with representative structures shown. This process is repeated for all five trials, during which average values are calculated for each trial and then averaged across trials. Error on these calculations is the standard error of the mean between the five trials. In the trial shown and most additional trials, the ToidRMSD helicity CV quickly approaches lower values after the restraint is removed, although structures still maintain helical character. This is likely due to slight deviations between the backbone dihedrals of the ToidRMSD "perfect helix" and the actual helical structures seen in simulation.

Two different steering protocols were implemented for each section of this work: one for the comparison of structures between different sequences that relied on the use of the ToidRMSD collective variable to define and steer to ideal helicity and another for the comparison of different handedness in the CH30 sequence that relied on the use of more specific dihedral angles to define and steer to ideal helicity.

For the first steering protocol, each trial was allowed to equilibrate for 200 picoseconds and then the distance between the center of mass of the pre-linker and post-linker sections of the sequence was pushed from the starting value to 4 nanometers over the next 10 nanoseconds, starting with a kappa value of 1000 and ending with a kappa value of 2500. During this timeframe for the systems that display helical character (CH30 and CH22), each helical section was then steered to the maximum value of the ToidRMSD CV, starting with a kappa value of 100 and ending with a kappa value of 2500. This helical restraint was held for the next 25 nanoseconds, over which the distance restraint was slowly removed and collapse (in the case of systems with hydrophobic residues) was witnessed. The system was allowed to equilibrate for another 65 nanoseconds and then the helical restraint was slowly removed over the next 50 nanoseconds, leading to a total simulation time for the steering portion of the protocol of 150 nanoseconds. Each simulation was then run with no restraints or steering for another 250 nanoseconds, leading to a total sampling time (sum for five trials) of 1.25 microseconds.

For the second steering protocol, each helix needed to be strictly confined to a more specific helical minima so that we could analyze structural differences between a hairpin with particular helical handedness. There are two primary helical minima present in peptoids: the Alpha D Plus minima, or the left-handed helix, and the Alpha D Minus minima, or the right-handed helix. To provide a full sampling of all possible helical states, we investigated the structure of a helical bundle composed of left-handed/left-handed helices, left-handed/right-handed helices, right-handed/right-handed helices, or right-handed/left-handed helices in comparison to the same sequence with no helical character at all. To accomplish this steering to particular handedness, every phi, psi, and omega dihedral was defined for the sequence. Then, four alphabeta variables were defined for each helical portion of the sequences: one for all the phi dihedrals, one for all the psi dihedrals, one for all the omega dihedrals, and one for all phi, psi, and omega dihedrals. The reference values for these alphabeta variables were then chosen to reflect the steering of each portion of the sequence to the preferred helical handedness (left-handed or right-handed). Each trial was equilibrated for 200 picoseconds, and then the distance between the COM of each helical section was pushed from the starting value to 4 nanometers over 10 nanoseconds, starting with a kappa value of 1000 and ending with a kappa value of 2500. During these 10 nanoseconds for all systems except that with no helical steering, the values of each alphabeta collective variable (four for each helix, or eight in total) were pushed to the maximum value, starting with a kappa of 100 and ending with a kappa of 1000. The helical restraint was held for the next 25 nanoseconds while the distance restraint was slowly removed so each system could experience hydrophobic collapse. The collapsed hairpin was allowed to equilibrate for another 65 nanoseconds and then the helical restraint was slowly removed over the next 50 nanoseconds, leading to a total simulation time for the steering portion of the protocol of 150 nanoseconds. Each simulation was then run with no restraints or steering for another 250 nanoseconds, leading to a total sampling time (sum of five trials) of 1.25 microseconds.

The PLUMED input files required to reproduce the steering protocol reported in this paper are available on PLUMED-NEST (www.plumed-nest.org) [77], the public repository of the PLUMED consortium, as plumID: 24.003. These files also contain the definitions of all collective variables discussed in this work.

Clustering and Calculation of RMSD

Data analysis was only performed on simulation trajectories after all restraints were removed (during the last 250 nanoseconds of each trial). Values for each of the collective variables previously highlighted (simulated FRET distance and coordination number) were averaged over this production period to provide an average for each trial. All five trials were then averaged; the error on each calculation was reported as the standard error of the mean between these five trials.

Clustering was performed on 3,125 frames from the last 250 nanoseconds of each trial's trajectory using the GROMOS algorithm [78] with a Cα RMSD cutoff of 0.2, allowing for a total of 15,625 structures to be analyzed per system by combining all 5 trials. This algorithm selects the first cluster based on the largest number of neighbors within that cutoff and then repeats that process with the remaining frames until each frame is assigned to a cluster number. For each system, the percentage of structures found in the top three clusters were calculated while the top three cluster centers for each system were visualized using VMD [79].

Three calculations of root mean square deviation (RMSD) were performed against two different reference structures; the reference structure for the RMSD of each helix was chosen as the structure from each trial at 100 ns (when the helical steering was still in place for the helical systems and when there was no helical steering for non-helical systems) and the reference structure for the RMSD for the entire structure was chosen as the largest cluster center from the above clustering analysis. Each reference structure was then loaded into VMD with the no-restraint trajectory from each trial and the RMSD Trajectory Tool was used to calculate the RMSD of the backbone (no linker included) for the first half of the sequence (before the linker), the second half of the sequence (after the linker), and then the entire sequence (before and after the linker) over all frames in the trajectory. For each trial, the RMSD was averaged over the 250 ns production period, and then averaged across those five trials to report a final value. Error on the final RMSD calculations is the standard error of the mean for all five trials.

2.4 Results and Discussion

Comparison of Sequences in Water and Acetonitrile

The four sequences chosen from Lee et. al contain a wide variety of residues with interesting characteristics. CH30 has both chiral residues, indicating that the system is helical, as well as hydrophobic residues, which indicate that the system likely collapses in water. AH30 has hydrophobic residues, indicating a compact structure, but no chiral residues, so the presence of helical sections of the sequence is highly unlikely. AN30 has no hydrophobic residues at all; instead, these hydrophobic residues have been replaced by a residue that is positively charged at neutral pH, leading to more extended structures. Lastly, CH22 has chiral (helix-inducing) and hydrophobic (compact) residues, but at a smaller spacing than CH30 due to the shortened sequence length.

When simulated in water, each of these residues imparts a different character to the structure of each sequence, which can be shown in Figure 2.4 below. The biggest three cluster centers for CH30 all show helical or partially helical structures of at least half of the sequence (pre-linker or post-linker), as well as hydrophobic cores indicated by the interaction between the hydrophobic residues shown in black, and the general compact nature of the hairpins. In contrast, the biggest three cluster centers for AH30 have no secondary structure but do still display compact hydrophobic cores shown by the interacting side chains in black. The top three cluster centers for AN30 make up a very small portion of the analyzed frames for that sequence, indicating a high level of dissimilarity between most of the structures here. The structures for this sequence are largely extended with a low level of rigidity. In particular, AN30 is shown to have 10-15x more clusters than the other systems at the 2 Å cutoff length. This significant increase compared to the other three structures is due to the fact that a long, extended system needs very little conformational change to affect a structural difference at this cutoff size. We verified the trends are insensitive to the cutoff size and selected 2 Å due to the reasonable volume of cluster weights found in the top 3 clusters of the other systems.

Lastly, CH22 shows helical character like that of CH30, but with a hydrophobic core that seems less well pronounced due to the shortness of the sequence and the different spacing of the hydrophobic residues. The sequences with hydrophobic residues (CH30, AH30, and CH22) all maintain a relatively similar FRET distance coinciding with collapsed structures as well as higher coordination numbers, indicating more interaction between those hydrophobic residues. The helical and hydrophobic sequences (CH30 and CH22) demonstrate lower RMSD of each helical section to a reference structure than the non-helical sequence (AH30) due to the rigidity that secondary structure imparts on the overall helical bundle, with the shortest sequence CH22 being noticeably less flexible and more stable. However, RMSD of each structure for all hydrophobic sequences is relatively similar and much lower than that of the non-hydrophobic and charged system AN30.

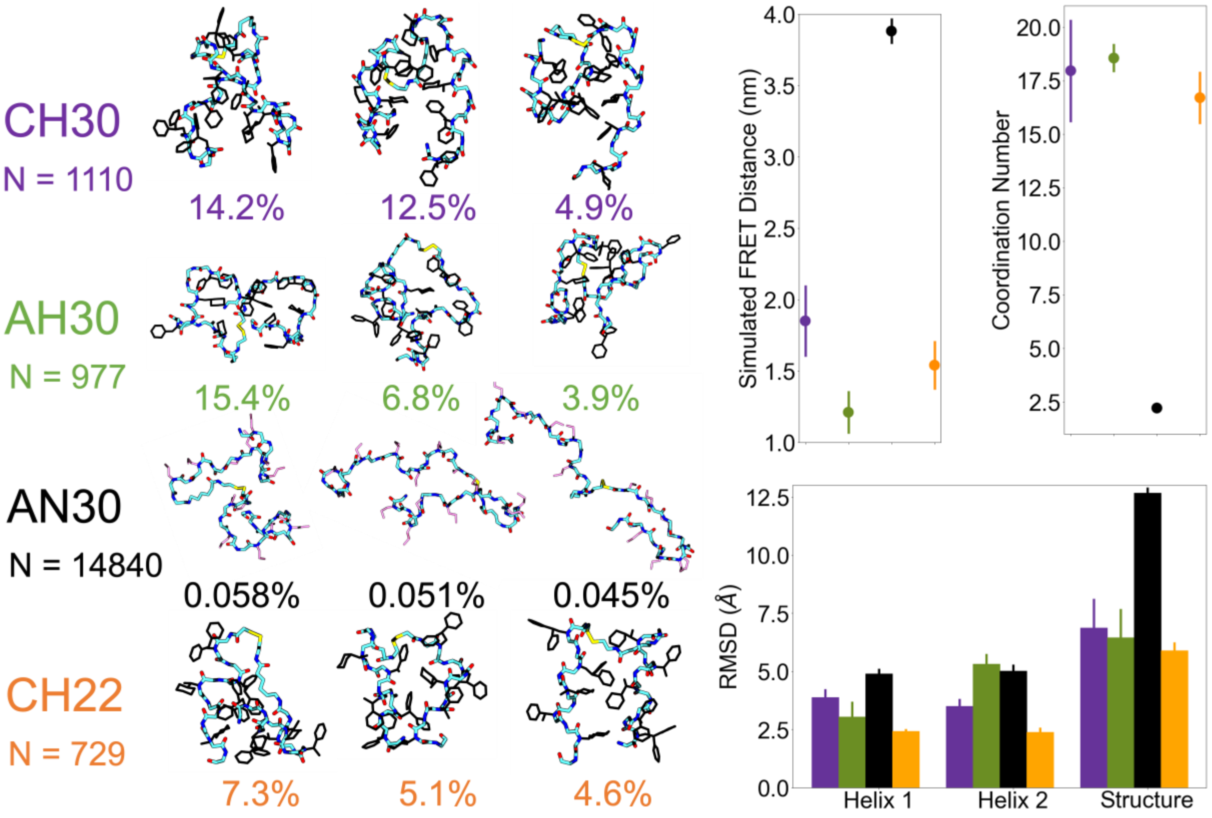


Figure 2.4: Three largest cluster centers for each sequence in water with the number of clusters for each sequence indicated below title. Simulated FRET distance, coordination number, and RMSD of section before the linker (Helix 1), section after the linker (Helix 2), and entire structure (CH30, AH30, AN30, and CH22 in order from left to right). CH30 is shown in purple, AH30 is shown in green, AN30 is shown in black, and CH22 is shown in orange.

These observations are consistent with our expectations of these sequences in water; those sequences with hydrophobic residues will collapse in order to hide those hydrophobic residues from the solvent, while the sequences with chiral residues will demonstrate more rigidity in their helical domains due to the added secondary structure elements. It is interesting to note here that neither CH22 or CH30 maintain a structure with two perfect helices; while there is a degree of helical character present in the major cluster centers for both, the driving forces towards a defined secondary structure for systems composed of residues like Nspe have been shown to be very small[34,35], so these sequences likely produce an ensemble of states with helical or partially helical structure.

We see a number of structural differences when the solvent is changed to acetonitrile. Lee et al. used acetonitrile to explore the cooperative unfolding of these exact peptoid systems, so it was expected that the solvent would lead to melting of the hydrophobic contacts responsible for a hydrophobic core, leading to much more extended states, but our simulation results find that the structural differences between these systems in each solvent are much more nuanced than which can be analyzed with experimental FRET distance alone. First, in acetonitrile, the hydrophobic residues in each system are no longer directed inwards to form the hydrophobic core of each structure and are instead pointing outwards into the solvent as shown in Figure 2.5 for sequences CH30, AH30, and CH22. This indicates sidechain-solvent interactions are energetically competitive with sidechain-sidechain interactions, in significant contrast to water. These three systems thus have a significantly lower coordination number between these residues and a higher simulated FRET distance than previously in water, although the top three cluster centers for the three hydrophobic sequences do show a certain degree of compact tertiary structure. The systems with helical character, CH30 and CH22, show stronger helical stability through lower helical RMSD (although they demonstrate very similar structural RMSD to the water simulations) as well as a higher percentage of structures in the top cluster centers and a lower overall number of clusters compared to these simulations in water. We believe that the hydrophobic solvent has a stabilizing effect on the secondary structure elements in CH30 and CH22, allowing for more helical homogeneity and thus more structural homogeneity in the sampled cluster frames. In contrast, sequence AH30 has no secondary structure elements to be stabilized in this way, resulting in a much higher number of clusters than seen previously in water because the driving force towards structural homogeneity is much smaller without the encouragement of the hydrophobic core.

Interestingly, the charged sequence, AN30, is able to collapse in acetonitrile due to the charge screening effect of the solvent on the positive charges that had previously led to significant electrostatic repulsion of these segments. This charge screening allows AN30 to maintain a compact structure, with a low simulated FRET distance and higher coordination number than previously seen in the water simulations. Remarkably, a change from water to acetonitrile leads to a 60-fold reduction in the number of unique 2 Å structural clusters at room temperature. This leads to a much lower helical and structural RMSD as well, significantly fewer clusters than in water, and a much higher percentage of structures in the top three cluster centers for this sequence.

While Lee et. al had indicated they saw cooperative unfolding of the tertiary structure of the hydrophobic sequences in acetonitrile demonstrated by higher FRET distances than seen in water, our simulations seem to indicate that these structures can still fold into collapsed structures in acetonitrile; while these collapsed structures in acetonitrile do have higher FRET distances than collapsed structures in water, the flexibility of these peptoid sequences means that they can still maintain hairpin-like structures even with these slightly larger FRET distances. Lee et al. also indicated that while they believed they saw cooperative unfolding of tertiary structure of the hydrophobic sequences in acetonitrile, they did not see unfolding of secondary structures in these systems, which is consistent with our results showing higher helical content of each chiral structure stabilized by acetonitrile compared to that seen previously in the aqueous water simulations.

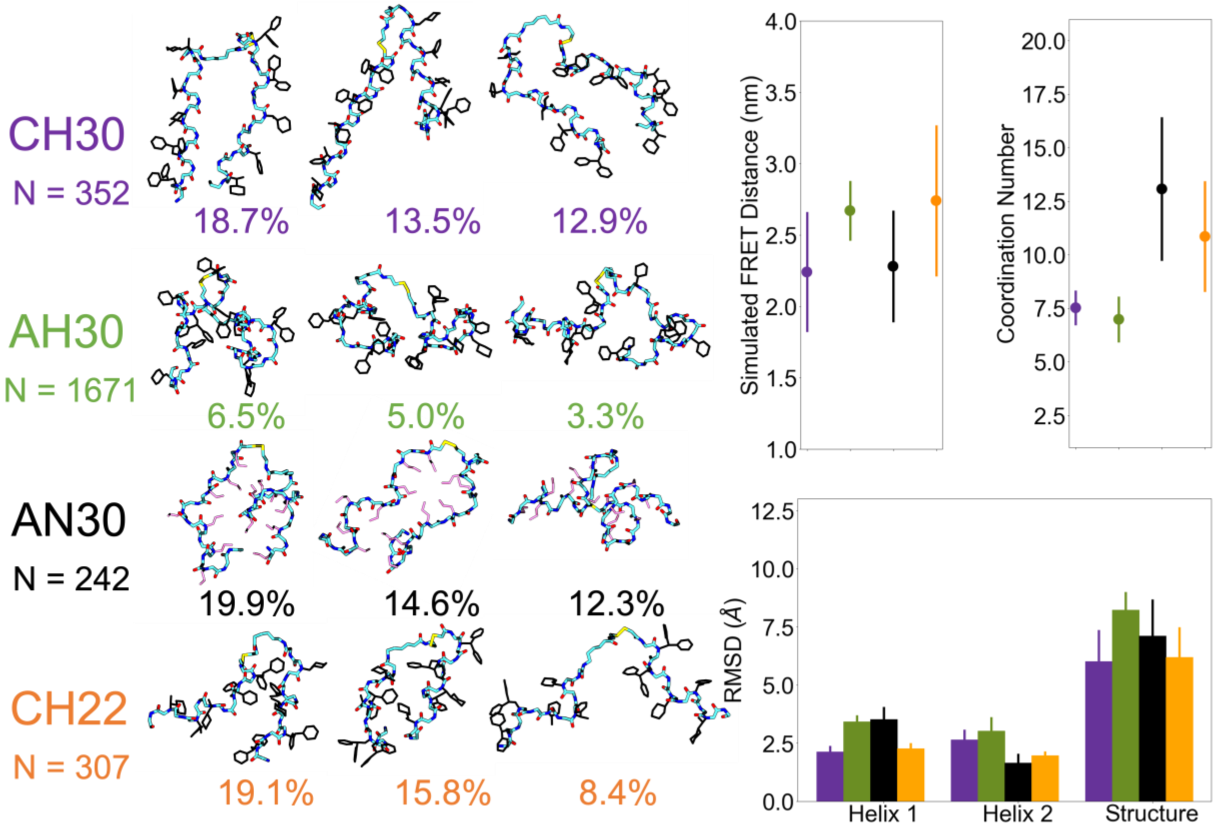


Figure 2.5: Three largest cluster centers for each sequence in acetonitrile with the number of clusters for each sequence indicated below title. Simulated FRET distance, coordination number, and RMSD of section before the linker (Helix 1), section after the linker (Helix 2), and entire structure (CH30, AH30, AN30, and CH22 in order from left to right). CH30 is shown in purple, AH30 is shown in green, AN30 is shown in black, and CH22 is shown in orange.

Comparison of Helical Structure in Water and Acetonitrile

After comparing the structure of the previous four sequences, we were interested in studying in more detail the subtle differences in structures of the CH30 sequence, which contains both helix-inducing chiral residues and hydrophobic residues responsible for compact structures. To do this, we used an updated steering protocol to specifically steer each pre-linker and post-linker section to the dihedral values of a defined peptoid helical minima of either left-handed or right-handed nature. The additional flexibility conferred to peptoid backbones leads to a significantly expanded thermally accessible set of conformations [35]and, unlike in peptides and proteins, the handedness of a particular helical structure is not definitive. Thus, we set out to use simulations to help shed light on the relative impact of helical handedness and solvent on stability of higher order peptoid structures.

This led to four main steered conformations: left-handed pre-linker section with left-handed post-linker section (referred to here as “LH/LH”), left-handed pre-linker section with right-handed post-linker section (“LH/RH”), right-handed pre-linker section with right-handed post-linker section (“RH/RH”), and right-handed pre-linker section with left-handed post-linker section (“RH/LH”). We then compared these with data from the CH30 sequence without helical steering (“None”).

Figure 2.6 below shows the three largest cluster centers for each of these different handed structures of this sequence in water, along with simulated FRET distance, coordination number of the hydrophobic residues, RMSD of the pre-link and post-link sections, and overall structural RMSD. There are collapsed hairpin structures evidenced by the similar coordination number for all systems produced from each steering protocol (including that without helical restraints), although the coordination number of hydrophobic residues is slightly lower for the structures with the non-helical steering protocol. On average, all structures show a similar FRET distance, indicating that any combination of left-handed or right-handed helices, or even no helices, can produce a compact structure. Each different steering has roughly the same number of clusters, again indicating similar stability. This is likely due to the small driving forces towards helical homogeneity in peptoids with Nspe sidechains in water [34,35]; this specific sequence in water likely produces an ensemble of states with varied helical percentages and left-or-right handed character.

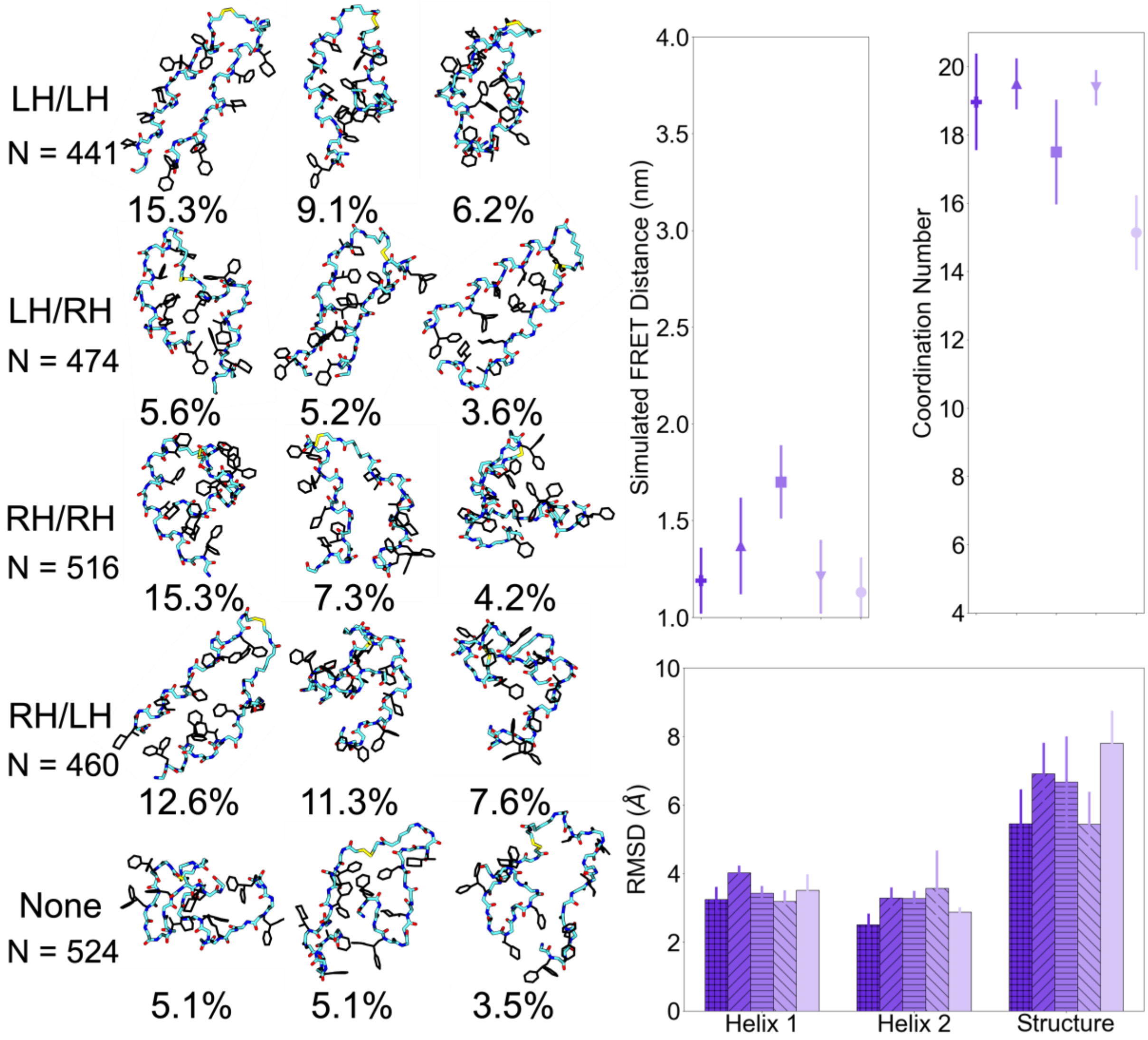


Figure 2.6: Three largest cluster centers for each handedness combination in water with the number of clusters for each indicated below title. Simulated FRET distance, coordination number, and RMSD of section before the linker (Helix 1), section after the linker (Helix 2), and entire structure. LH/LH, LH/RH, RH/RH, RH/LH, and no helical steering shown in order from left to right, with LH/LH being the darkest purple, the plus sign marker on the scatter plots, and the plus sign hatch marks on the bar plots.

The replacement of water with acetonitrile in these systems greatly changes these structures and their relative stability, as shown in Figure 2.7. All of the resulting structures show a much lower coordination number of hydrophobic residues in acetonitrile, as these residues are now pointing outwards into solution. The average simulated FRET distance for these systems is where we start seeing marked differences between the different handed systems with the RH/RH and RH/LH systems showing a much higher average FRET distance, indicative of structures that may not necessarily collapse with those FRET tags particularly close together. While the RMSD of each helical section for all types of handedness is lower in acetonitrile than in water, there aren't large differences between the five systems. However, the LH/LH steering has the lowest overall structural RMSD and the lowest number of clusters (194 v. 400-1000 clusters for all other systems). The presence of acetonitrile here stabilizes the interactions between Nspe sidechains, which has always been thought to adopt left-handed helices preferentially to any other handedness (due to being an S enantiomer). The hydrophobic solvent here likely causes a stronger degree of conformational homogeneity in the helical-inducing residues, which more strongly supports the LH/LH steering of the s-enantiomer residues than any other steering, leading to more defined configurations in this handed system than in any other structures produced from either steering protocol in water or acetonitrile.

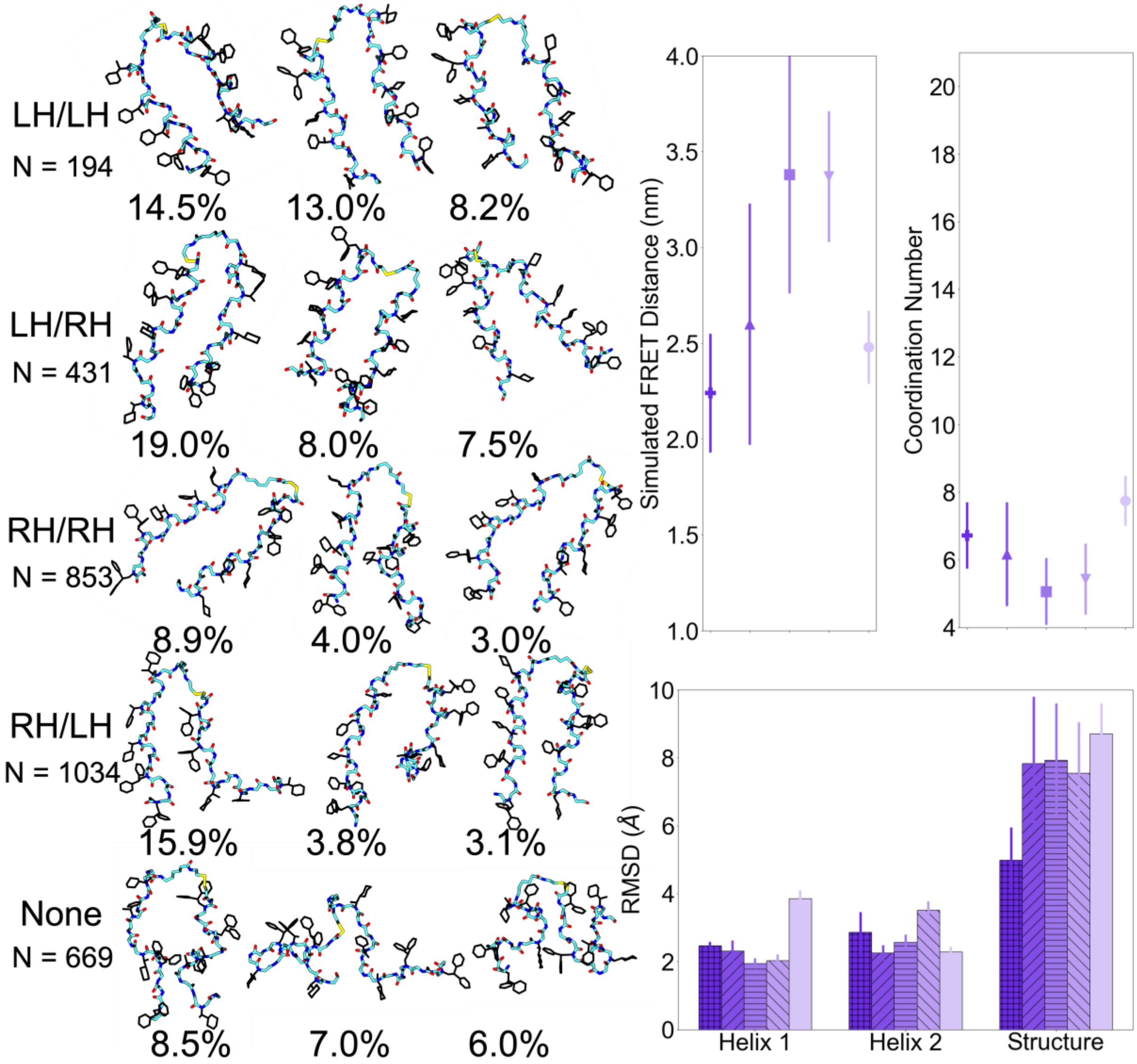


Figure 2.7: Three largest cluster centers for each handedness combination in acetonitrile with the number of clusters for each indicated below title. Simulated FRET distance, coordination number, and RMSD of section before the linker (Helix 1), section after the linker (Helix 2), and entire structure. LH/LH, LH/RH, RH/RH, RH/LH, and no helical steering shown in order from left to right, with LH/LH being the darkest purple, the plus sign marker on the scatter plots, and the plus sign hatch marks on the bar plots.

2.5 Conclusions

In this work, we shed new light on molecular scale driving forces that arise in the stabilization of secondary and tertiary structures in peptoids. Building off of seminal experiments that showed the feasibility of peptoid hairpins, we relate solvent, sequence, and structure through the study of unique chiral, non-chiral, hydrophobic, and charged residues. We show that the main drivers of the collapse of stable tertiary structures in water is the presence and placement of hydrophobic residues. We show that while chiral side chains can induce a level of helical structure in these hairpins, it is likely that the driving forces in water for helical homogeneity are fairly small and that sequences with chiral residues likely sample an ensemble of collapsed configurations with helical, partially helical, or non-helical sections. We also indicate that in acetonitrile, the secondary structure elements of peptoid sequences with chiral side chains are strengthened and we move closer to helical homogeneity, even without the stabilizing elements of a defined hydrophobic core.

In the future, this simulation method can be used to screen sequences computationally for desired structural elements in each solvent, including narrowing in on the differences in structures between sequences with small changes in hydrophobic or chiral residue placement. This work can also hopefully encourage further computational research into the applications of these secondary and tertiary structures in creating nanostructured peptoids that can eventually mimic the function of more complex proteins.

Chapter 3: Rational Design of Novel Biomimetic Sequence-Defined Polymers for Mineralization Applications[2]

3.1 Abstract

Silica biomineralization is a naturally occurring process, wherein organisms use proteins and other biological structures to direct the formation of complex, hierarchical nanostructures. Discovery and characterization of such proteins and their underlying mechanisms spurred significant efforts to identify routes for biomimetic mineralization that reproduce the exquisite shapes and size selectivities found in nature. A common strategy has been the use of short peptide sequences with chemistry mimicking those found in natural systems, such as the use of the silaffin-derived R5 peptide. While progress has been made using this approach, there are many limitations that have prevented breakthroughs in biomimicry. To advance our ability to use charged macromolecules for silica formation, we propose to use sequence-defined synthetic polymers known as peptoids, or N-substituted polyglycines, which present significant capability for the precise tuning of sequence and structure beyond what can often be achieved with peptides alone. This study presents a computationally predicted design of these polymers that leads to the controlled formation of silica nanomaterials. We investigate surface adsorption and the mineralization process through analysis of binding mechanisms and energetics of the R5 system. Next, we synthesized two R5-inspired peptoids and validated our prediction in the design of mineralization polymers through characterization using surface plasmon resonance and electron microscopy. This computationally guided study holds great promise for designing new sequences with unprecedented control of the placement of chemical functional groups, thus allowing for further unraveling of silicification mechanisms and the eventual design of sequence-defined synthetic polymers leading to the predictive synthesis of nanostructured functional materials.

3.2 Introduction

The natural process of biomineralization occurs in the cell walls of diatoms, photosynthetic single-celled algae known to produce complex hierarchical silica structures under ambient

---

[2] This chapter was reproduced with permission of K. Torkelson, N. Y. Naser, X. Qi, Z. Li, W. Yang, K. Pushpavanum, C-L. Chen, F. Baneyx, and J. Pfaendtner from *Chem. Mater.* 2024, 36, 2, 786–794. Copyright 2024 American Chemical Society. [71]

conditions. Synthetically, this process can be mimicked by using the R5 peptide (SSKKSGSYSGSKGSKRRIL) which is repeated in silaffin, a silica-mineralizing protein present in diatom cell walls. R5 supports the biomimetic mineralization of silica under simple reaction conditions by producing silica nanoparticles but does not deliver the structural complexity and morphological control of natural biomineralization processes; while control of the biomimetic silica process has made progress in recent years, the complex hierarchical silica structures we witness in nature continue to be the ultimate goal of works such as this.

Silica has classically been used in catalysis and separations but has gained interest in recent years for biotechnological and biomedical applications such as biosensors, drug delivery systems, and novel materials due to its defined structure and properties [41–45]. Multiple silica precipitating biomolecules have been used for the synthesis of biomimetic silica, including larger domains of silaffin [40], poly-L-lysine [80], R5 peptide [81,82], Car9 peptide [83], and LK repeat peptides [84], demonstrating the importance of cationic moieties and the ability to self-assemble in the mineralization process. The R5 peptide has become a specific focus in this area due to its short sequence, disordered backbone, biocompatibility, and ability to induce silica precipitation from silicic acid precursors at near-neutral pH. Furthermore, the R5 peptide precipitates silica morphology in vitro that is similar to those obtained with portions of the silaffin protein under ambient conditions, making the peptide a popular model system for biosilicification [41,85].

Our knowledge of key structure−function relationships within R5 and other mineralizing peptides has steadily progressed since the seminal paper by Kröger et al. [40,86,87] that first illustrated the mineralization capability of short repeat chain segments of silaffin under certain conditions [88] (e.g., phosphate buffer). Selected examples include detailed computational and experimental studies of individual post- translational modifications such as trimethylation or comprehensive screening studies [85,89–91]. However, unraveling a fully mechanistic understanding that quantitatively relates emergent silica structures and growth rates to key parameters such as side-chain chemistry, peptide single chain structure, peptide aggregate structure, relative concentrations of precursor and peptide, and buffer and pH effects has remained elusive. This can at least partially be attributed to a lack of experimental techniques that can probe the nascent stages of silica nucleation and growth concurrently with the properties of the peptide(s). Additional approaches are therefore needed to advance the state of the art.

Computational approaches such as molecular dynamics (MD) simulations represent a unique opportunity not only for improved understanding of existing peptide mineralizers but also to allow for the design of new sequences that synthesize silica with an expanded range of properties. For example, prior computational studies have shown that the phosphorylation of serine side chains is critical to the binding energy and interfacial conformation of R5 peptide with silica surfaces, providing a rationale for the pH-dependent behavior of R5 silica formation due to the strong relationship between net peptide charge, surface ionization state, and peptide conformation [90]. Molecular dynamics simulations have also been used to show the molecular structure of the R5 peptide responsible for the formation of stable nanometer-thin biosilica sheets [92]. Combined molecular dynamics and experimental approaches have been used to illustrate the structure of long-chain polyamines at silica interfaces as well [93].

While the R5 peptide and other proteins have been studied extensively for synthetic biosilicification, other biomolecules using sequence-specific chemistry present a new opportunity to understand, control, and design new silica morphologies based entirely on sequence and structure. Peptoids are a class of sequence-defined synthetic polymers with building blocks that are analogous to natural amino acids except that peptoid side chains are attached to backbone nitrogen atoms instead of alpha carbons [20,21]. This difference in backbone connectivity enables their increased stability and conformational flexibility compared to peptides. Peptoids can also be synthesized with a broad variety of side chains that can control both sequence and structure, making this class of materials an attractive focus for gaining finer control over silica morphology purely based on elements of the peptoid sequence. Recently, we found that short peptoid sequences self-assemble and facilitate the formation of various silica nanostructures, including nanofibers and nanosheets, in which the presence of silica-binding amino groups is important for the templated silica formation [94].

Herein, we seek to expand our understanding of peptoid- based biomimetic mineralization through the controlled study of peptoids with a side chain chemistry matching approach. We first use MD simulations to rationally design two peptoids that capture the essential features of the R5 peptide by comparing the biomolecule and silica binding free energies. Following this, the synthesis of a new biomineralizing peptoid is demonstrated with detailed characterization via surface plasmon resonance, microscopy, and MD simulations. Key descriptors for biomineralization efficiency are identified, namely, biomolecule/silica binding affinity and the interactions of silicate precursors with charged side chains. These findings are further validated with synthesis and simulation analysis of additional sequences, revealing key driving forces and suggesting new design rules for future studies.

3.3 Results

Noting that R5 peptide has been observed to be disordered in solution and at silica interfaces [82], we hypothesized that an exact R5 peptoid mimic would similarly synthesize silica nano- particles due to the lack of intrinsic backbone structure and the similarities in side chain chemistry. As described in the Materials and Methods section, we parameterized a force field for a peptoid mimic of R5 with side chains having identical chemical structure to the R5 peptide, with the only difference being the peptoid side chains connect to the backbone nitrogen instead of the backbone α carbon (n.b., for clarity, hereafter we refer to this peptoid as *toid*R5M or simply R5M for short as shown in Figure 3.1). A series of MD simulations were performed to study the structure of peptide R5 and *toid*R5M in solution and to study the adsorption behavior to amorphous silica interfaces (Figure 3.1). As shown in the figure, R5 and *toid*R5M have nearly identical behavior regarding the adsorption binding profile and overall binding energy. Recent work from Hamm et al. revealed that binding energy is a potentially useful predictor of the mineralization behavior of small molecules [95]. This finding alongside our prior work studying the role of serine phosphorylation in R5 binding to quartz [90] provides a strong basis for expecting a peptoid mimic of R5 to similarly produce silica nanoparticles of low dispersity like its peptide counterpart.

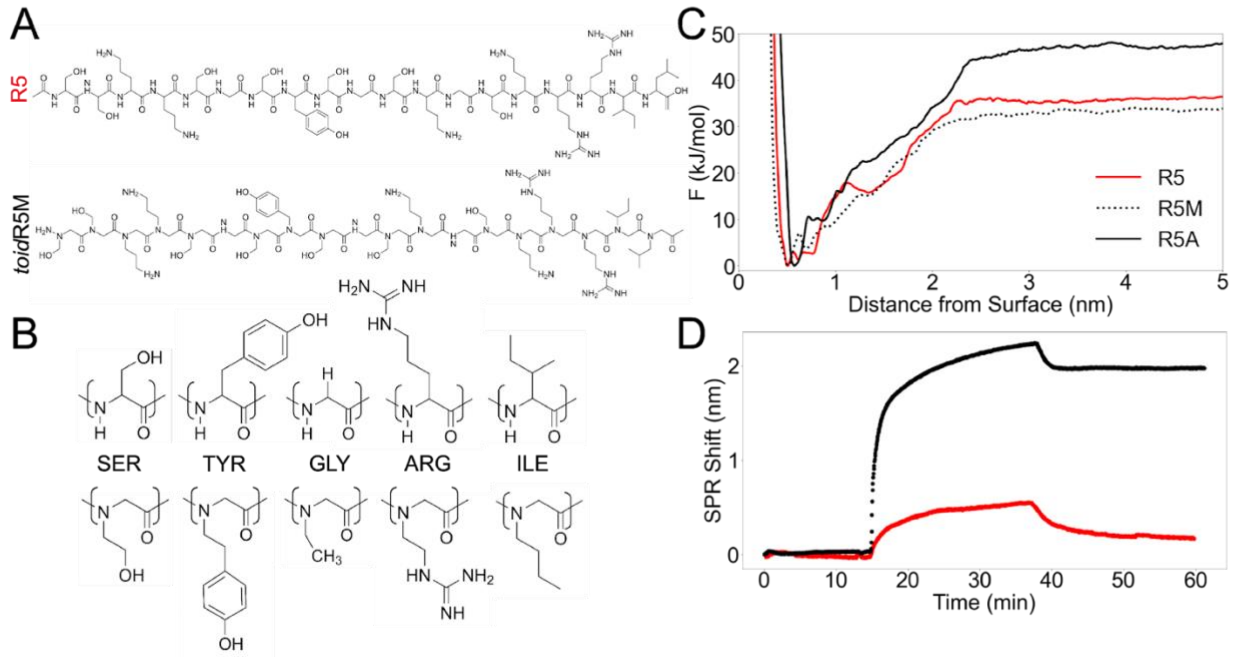


Figure 3.1: (A) Sequence of the R5 peptide (top) and the proposed R5 peptoid mimic toidR5M (bottom). (B) Comparison of side chain chemistry for serine, tyrosine, glycine, arginine, and isoleucine between the R5 peptide (top) and the synthesized R5 peptoid analogue or toidR5A (bottom). (C) Binding free energy of the center of mass (COM) of the R5 peptide (red), the R5 peptoid mimic or toidR5M (dotted black), and the R5 peptoid analogue or toidR5A (solid black) to an amorphous silica surface obtained through metadynamics calculations. (D) Surface plasmon resonance sensorgrams obtained by flowing 25 μM R5 peptide (red) or toidR5A (black) on a multichannel SPR chip. Samples were injected after 15 min of equilibration in pure buffer and the signal was allowed to reach a plateau corresponding to equilibrium binding. After an additional 23 min, pure buffer was injected to induce dissociation which occurred with similar kinetics for both species.

To test the predictions from MD simulations, we adopted a side chain similarity matching approach to account for two issues. First, the R5 peptide contains three glycine residues, which cannot typically be included in peptoid sequences. Second, several of the amino acid analogue residues were not readily available as precursors for peptoids. Thus, we identified a synthesizable peptoid analogue (hereafter, *toid*R5A or R5A) with several minor modifications that can be seen when compared to the canonical amino acid side chains. Note, however, that the sequence length is maintained in R5A as is the chemical character of each residue. Figure 3.1B highlights the differences between the analogue residues in *toid*R5A, with side chains that connect to the backbone nitrogen, and the canonical amino acid counterparts found in the R5 peptide, with side chains that connect to the backbone α carbon. In *toid*R5A, the serine and tyrosine analogues have an additional CH_2 in the alkyl chain, the glycine analogue replaces the hydrogen "side chain" with

$-CH_2CH_3$, the arginine analogue removes one CH_2 from the alkyl chain, and the isoleucine analogue replaces the branched isobutyl group with a butyl chain.

We parameterized a force field and carried out an identical series of simulations for *toid*R5A, which showed the binding predicted by metadynamics simulations to be substantially higher than the corresponding peptide and *toid*R5M systems (Figure 3.1C). Following this, both R5 and *toid*R5A were synthesized as described in the Materials and Methods and Supporting Information sections. To validate the trends from the binding energy predictions, we conducted surface plasmon resonance (SPR) experiments (Figure 3.1D) in which equimolar concentrations of the R5 peptide and the R5 peptoid analogue (*toid*R5A) were flowed over a silica-coated chip until the SPR shifts reached a plateau. This equilibrium shift, which is reached when the number of dissociation events equals the number of association events, approximates the macro- molecular coverage on the chip surface. It was approximately fourfold larger for *toid*R5A relative to R5, denoting a substantially higher binding affinity of the peptoid on silica relative to the peptide. While surface binding energy is likely not the sole predictor of mineralization efficiency, it is known to be a useful descriptor owing to the fact that strong binding lowers the net interfacial energy and thus reduces barriers to nucleation [95].

To understand the differences in binding energies between these systems, two-dimensional free energy profiles were constructed for each (Figure 3.2A) to identify the preferred backbone structure on the surface. The data show that *toid*R5A (Figure 3.2A, right) has one very clear binding mode to the surface, denoted by an extended radius of gyration around 1.4 nm. The other two systems have two binding modes: an extended structure at 1.6 nm for the R5 peptide and at 1.4 nm for the hypothetical *toid*R5M, and a collapsed structure that binds slightly further from the surface with a radius of gyration of 1.2 nm for the R5 peptide and of 0.8 nm for the *toid*R5M. We previously related the level of extension of a peptide to its surface binding energy [96], as well as a lower distribution of available states [97], and the trends appear to be repeated here. The systems with lower binding energy (R5 peptide and *toid*R5M) show two available states instead of one, including a collapsed state that is further separated from the surface by several angstroms. A more collapsed state at a lower radius of gyration is also present in the R5A system (Figure 3.2A, right) but at energy levels ~4*kT* higher than the dominant, extended binding mode we have identified at a higher radius of gyration. These energetic trends appear converged, as shown in Figure A-S8, indicating that while this local minima does exist, it does not compete with the global minima previously identified.

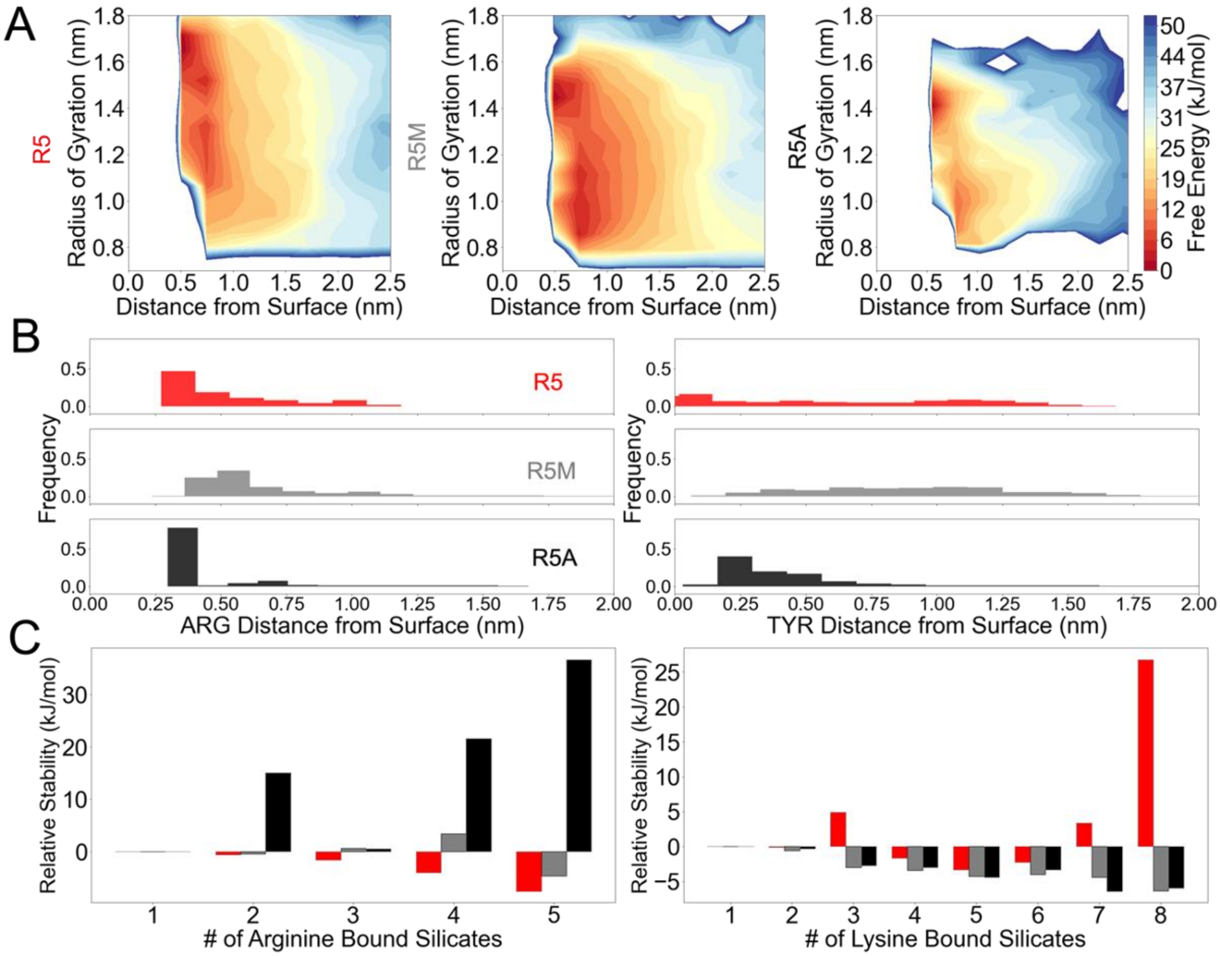


Figure 3.2: (A) Two-dimensional free energy surfaces for R5 peptide (left), *toid*R5M (middle), and *toid*R5A (right) showing the radius of gyration of each system against the distance of the center of mass (COM) of the backbone from the amorphous silica surface. (B) Histograms of the distance between the amorphous silica surface and the COM of both arginine residues (left) and the OH headgroup on the tyrosine side chain (right), indicating that both arginine residues and the tyrosine headgroup of *toid*R5A get closer to the surface, likely because of the modified side chain chemistry and a more flexible peptoid backbone. (C) Bar plots of the relative stability of interactions between silicate and arginine (left) and silicate and lysine (right) normalized to the binding of a single silicate anion. The R5 peptide is colored red, *toid*R5M gray, and the synthesized *toid*R5A is shown in black.

We next set out to specifically understand which residues contribute to the observed higher binding energy. The positively charged residues, lysine and arginine, have been shown to play the strongest role in silica binding [97,98], so we compared the relative propensity for binding among the three systems. Figure 3.2B shows a probability density for the side chain headgroup center of mass to the surface for the biomolecule structures we considered to be surface bound (backbone COM within 2 nm to the surface). While the lysine or lysine analogue residues nearer the N-terminus (residues 3 and 4) have roughly the same distance distributions for all three systems, larger differences are seen in the side chains near the end of the chain. Specifically, TYR8, LYS12, LYS15, and ARG16-17 (and their peptoid analogues) show marked differences, as seen in Figure A-S1. Here, we see that the structure of *toid*R5A allows for significantly closer distances of these charged and polar groups to the surface, which can be understood to be the causal factor for the stronger binding of the *toid*R5A compared to that of the R5 peptide and *toid*R5M. In the case of positively charged side chains (LYS and ARG) that are closer to the surface, visual inspection of trajectories shows a preference for these side chains to lie flat on the surface. It is also noteworthy that the headgroup of the tyrosine-like residue in the *toid*R5A has, on average, substantially closer and more narrow distance distributions although this side chain is one CH_2 group longer than tyrosine. The underlying reason for this is difficult to precisely attribute to a small number of attributes of *toid*R5A, however, it is likely that the tendency for the R5 peptide and *toid*R5M to fold into compact structures on the surface is a major contributor.

While this work and related studies [95] suggest that the biomolecule/surface binding energy may be a critical determinant of silicification outcomes, our prior work [85] also indicates that the strength of interaction between silica precursors in solution and charged side chains can also help understand and predict synthesis outcomes. To further investigate this, we performed simulations of single chains of each system in solution (R5 peptide, *toid*R5M, and *toid*R5A) with silica precursor ions and analyzed the free energies for association of silicates with lysine and arginine (or their respective side chains in the peptoids). Using the free-energy profiles directly from the simulations (Figure A-S2), we quantified the relative binding energy of the lysine and arginine headgroups with silicate anions in solution. From the resulting free energy profiles, we were able to quantitatively compare the binding energies between silicate and arginine or lysine headgroups in each system.

The data in Figure 3.2C show the stability of silicate anions binding to each headgroup relative to binding a single silicate anion, where positive values indicate stronger or more stable binding compared to one silicate anion and negative values indicate lower or less stable binding for silicate anions to that headgroup type. We see that *toid*R5A shows much stronger silicate binding to the headgroup of the arginine side chain compared to the R5 peptide and the *toid*R5M, as well as much lower silicate binding to the headgroups of the lysine side chains. Because the arginine heavy RRIL motif is generally thought to be a key driving force in the aggregation process of the R5 peptide [81,90,98,99], this increased interaction between silicates and arginine in *toid*R5A might influence both biomolecule−biomolecule interactions, as seen in Figure A-S2C, while potentially inhibiting more silicate interactions with the lysine active sites, thus making an impact on nucleation rates of new oligosilicates due to the impact on the structure and/or quantity of the charged side chains that act as nucleation sites.

To understand the impact of the differences we observed above in structure, binding energy, and affinity for silicate precursors, we compared the simulation results to the outcomes of silica mineralization reactions, which were conducted under standard mineralization conditions to allow for direct comparison of these outcomes between systems. These were conducted by mixing solutions of the R5 peptide or its peptoid analogue *toid*R5A at 1- or 3-mM final concentrations with an excess of orthosilicic acid produced by hydrolyzing the tetramethyl orthosilicate (TMOS) precursor in hydrochloric acid as described in the Materials and Methods section. Precipitation products were washed and characterized by scanning electron microscopy (SEM) (Figure 3.3). When used at a 3 mM concentration (Figure 3.3, top), both constructs yielded quasi-spherical silica particles in the 500− 600 nm size range. However, at the lower concentration, there was a marked difference in the synthesis outcomes: while the R5 peptoid analogue yielded similar 500 nm spherical particles, the peptide induced the formation of an extended structure consisting of interconnected, polycondensed nanostructures that lacked any morphological control.

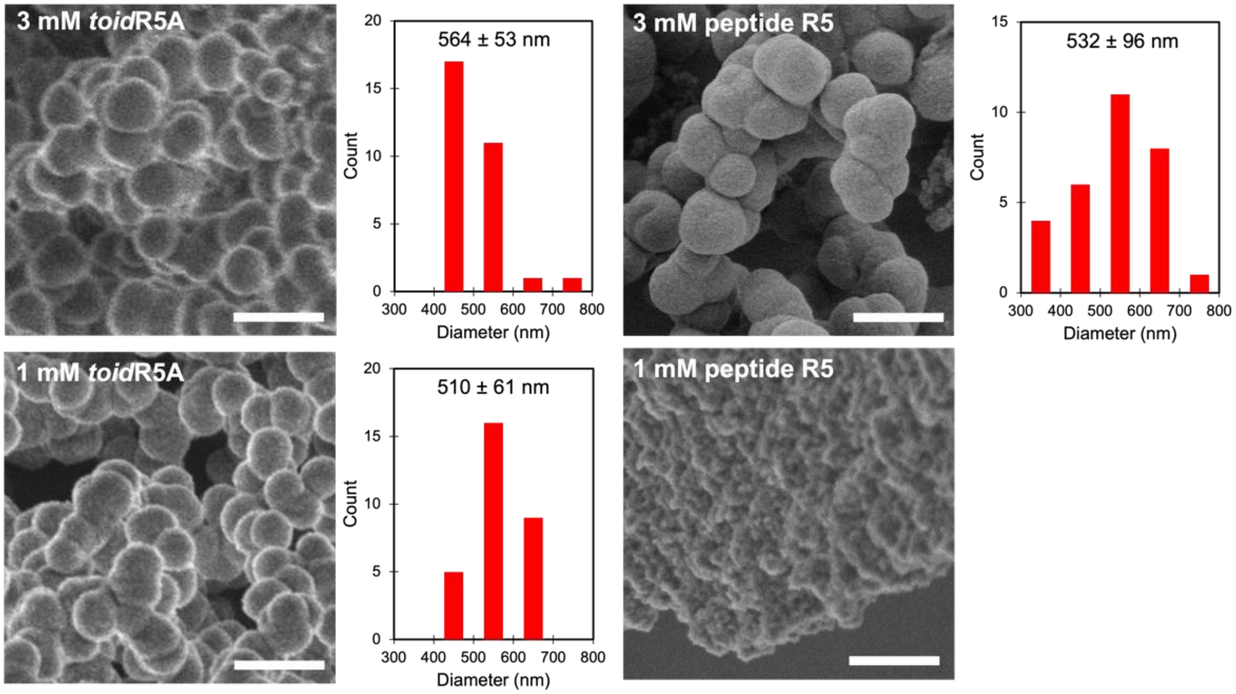


Figure 3.3: SEM images of silica mineralization products obtained in the presence of 3 mM (top) or 1 mM (bottom) *toid*R5A (left) or R5 peptide (right). Histograms show size distributions based on N = 30 particles and the scale bar is 1 µm.

The relationship between subtle structural differences between the peptide R5 and *toid*R5A and the notable differences in synthesis outcomes at low biomolecule concentrations can be rationalized in the context of the descriptors we identified from the simulations. Producing monodisperse silica nanoparticles from biomolecules requires a balance between the rate of nucleation of new sphere nuclei against the rate of growth of the individual spheres. Positively charged side chains act in a catalytic fashion in silicate polycondensation [100,101], thus the strength and persistence of interactions between the biomolecule and growing silica surfaces may be a crucial determinant of the fate of a particular synthesis experiment (i.e., positively charged side chains buried at a silica interface are not available to promote generation of new nuclei). The unique ability of *toid*R5A to preserve monodisperse silica nanoparticles at low concentrations of biomolecule precursor suggests that the binding energy is a critical signature of this synthesis outcome.

Finally, to further investigate the relationship between surface binding energy, precursor binding energy, and silica synthesis outcomes, we designed two additional molecules, reversing the peptide R5 and *toid*R5A sequences while keeping identical termini. The reverse peptide R5 sequence is thus LIRRKSGKSGSYSGSKKSS, and there is a corresponding reverse *toid*R5A peptoid that refers to identical side chain chemistries with an inverted sequence. As shown in Figure A-S3, the trends in key descriptors of binding affinity for amorphous silica and association with silicate precursors are preserved. Both reverse sequences nucleate silica nanoparticles when used at higher concentrations of 3 mM but fail to do so at lower concentrations (Figure A-S3A). The binding energy of both the reverse peptide and reverse *toid*R5A to amorphous silica is 15 kJ/mol lower compared to the forward *toid*R5A (Figure A-S3B), likely because forward *toid*R5A can maneuver side chain headgroups closer to the surface for stronger binding than the reverse peptide or its reverse peptoid counterpart, as demonstrated by narrower distance distributions (Figure A-S4). Likewise, the binding of silicate precursors to positively charged side chains follows similar trends (Figure A-S3C).

3.4 Conclusions

This research article reports a computational and experimental investigation of structure−function relationships for the design of new sequence-defined synthetic polymers to enable biomimetic mineralization. Using the archetypical R5 peptide as a model system, we designed and computationally characterized an exact peptoid mimic and subsequently synthesized an analogue of the R5 peptoid which was investigated computationally for its ability to bind silica and silicate ions and experimentally for its ability to induce the formation of quasi-spherical silica nanoparticles in the 500− 550 nm range. Through careful analysis of the differences and similarities in simulations and synthesis outcomes, we proposed several key features of biomolecule/silica interactions that should be targeted in future sequence design for the production of spherical nanoparticles and, potentially, additional types of morphologies with diverse shapes and sizes. Moreover, given the ability of peptoids to fold into secondary [31] and tertiary structures [54] as well as self-assemble into hierarchical nanostructures [22] , opportunities abound for future research studies using this platform of materials for the synthesis of more sophisticated silica structures as well as other metal oxides such as titanium dioxide. Further applications might involve conjugating peptoids into other polymers or using these strong nanoparticle-forming peptoids we have identified in conjunction with nanosheet or nanotube-forming peptoids to achieve more complex and hierarchical silica structures.

Future work should further investigate our finding that the single-chain binding energy is critically related to silicification outcomes. Given the extremely high computational cost of binding energy calculations that also explore backbone structure, new methods and advances are needed to improve the computational cost of these methods without sacrificing accuracy. Moreover, while the single chain−silicate precursor calculations seem to indicate a clear correlation between the synthesis outcomes and variance in the binding to the guanidinium groups of the peptoids and peptides, there is much to be done. For example, designing and testing new and short peptoid sequences that demonstrate strong surface binding with weak precursor binding to the arginine analogue (and vice versa) will do much to advance the state of art of mechanistic knowledge in these systems. Our work is a key first step in enabling significant improvements in the rational design of precursors for biomimetic mineralization through the lens of sequence/energetics/structure relationships. Future studies will build on this by using computers to screen potential mineralization candidates for their interactions with silica and precursors and surface and structural binding details.

3.5 Materials and Methods

Molecular Simulations

Systems. The R5 forward sequence is SSKKSGSYSGSKGSKRRIL, and the reverse sequence is LIRRKSGKSGSYSGSKKSS. We simulate five systems of interest in this study: (a) the R5 peptide, (b) the R5 peptoid mimic, (c) the R5 peptoid analogue, (d) the R5 reverse peptide, and (e) the R5 reverse peptoid analogue. The peptide version of R5 was constructed using the Visual Molecular Dynamics [79] Molefacture plug-in with an ACE N-terminus and a COOH C-terminus. Residues for all R5 peptoids (mimic and analogue, forward and reverse) were built within Gaussview [66]. The R5 peptoid mimic side chain chemistry is identical to that of the R5 peptide, just connected to the backbone at the nitrogen instead of the α carbon as shown in Figure 3.1. The R5 peptoid analogue side chain chemistry has been previously discussed in Figure 3.2. All peptoid systems have an NH_2 N-terminus and a CH_3 C-terminus. Each chain has a charge of +6 due to the protonation state of the lysine and arginine residues at neutral pH; this charge was neutralized in simulation with the addition of phosphate or silicate ions depending on the simulation type.

All systems were modeled using the CHARMM36 force field [68,69] and explicit TIP3P [72] water. Force field parameters for the peptoid backbones were taken from Weiser and Santiso [65] and the MFTOID force field [64] while the side chain parameters were taken from the CGenFF parameter generator [70] and compared to existing peptide side chain parameters. The phosphate ion was modeled based on the work of Mao et al. [102] using Antechamber [103,104] and RESP; a charge of −3 was used to mimic phosphoserine in a neutral pH environment and parameterization of phosphoserine used a combination of data from Steinbrecher et al. [105] and data contained within the CHARMM27 force field [106].

Silica Surface Simulations. The binding affinity of each system to silica particles was approximated by sampling the adsorption/ desorption of a single peptide/peptoid molecule on a semi-infinite silica slab in atomistic molecular dynamics simulations. We described the silica using the INTERFACE force field [107], which is compatible with the CHARMM force field by design. The silica surface was modeled using an amorphous structure (i.e., a suitable model for silica nanoparticles <200 nm) with the surface terminated with Q3 silanol groups. We tuned the ionization state of the silanol groups to 14.5% to reflect a solution pH of 7.5.

For all organic species, the simulations were set up by equilibrating a single molecule adjacent to an 8.06 × 8.29 nm amorphous silica slab solvated in water with two extra phosphate ions to neutralize the total charge. Simulations were performed using GROMACS [67] with the PLUMED [75] plugin for enhanced sampling. The LINCS [108] algorithm was used to constrain bonds between hydrogens and heavy atoms within the chain so that a 2 fs time step could be used for all simulations. Energy minimization was carried out for a maximum of 10 ps using the steepest descent algorithm. Lennard-Jones and Coulombic interactions were calculated with a cutoff of 1.0 nm, and long-range electrostatic interactions were treated with the particle mesh Ewald method. Energy minimization was followed by *NPT* equilibration at 298 K and 1 bar for 500 ps using the Berendsen barostat [74] with a pressure damping parameter of 1.0 ps and a *v*-rescale thermostat with a temperature damping parameter of 0.1 ps. Note that due to the semi-infinite structure of the silica slab, pressure coupling was applied only to the *z* dimension in the *NPT* equilibration. Moreover, the center of the silica slabs was frozen during all energy minimization, *NPT* and *NVT* steps, allowing only the silanol groups and the top layer of silica to move.

The binding free energy of each system was calculated using parallel-bias metadynamics (PBMetaD) [109] with multiple walkers [110] and adaptive Gaussians [111]. PBMetaD allows multiple collective variables to be biased in parallel to efficiently explore the free energy landscape on the time scale of a simulation, while multiple walkers and adaptive gaussians are employed to speed the convergence of each simulation. The initial configurations of eight walkers were selected from the last 50 ns of the *NVT* simulation, and each possessed a distinct configuration and location to ensure efficient sampling in the phase space. We biased 22 one-dimensional CVs in each PBMetaD calculation, including the orthogonal distance between the COM of the backbone and the surface, the orthogonal distance between the COM of each residue headgroup and the surface, and the radius of gyration. The total sampling time for each system was 12 μs (1.5 μs for each replica) or until convergence.

Silicate Simulations. We assessed the interactions between the R5 peptide/peptoid and silicates by solvating a single peptide or peptoid with six silicate ions, which corresponded to the number of positively charged residues on each chain. The structure and topology of the silicate were adopted from the work of Buckle et al [85]. The equilibration process was carried out in the same fashion as that of the surface simulations. All systems were relaxed in a cubic box with a length of 6.93 nm.

The affinity between peptide/peptoid and silicates was characterized by the coordination number of silicate ions around a specific group on the chain. Since the positively charged LYS and ARG residues are the most relevant, the biased CVs included the coordination number of silicate ions around all LYS groups and all ARG groups. We additionally biased the radius of gyration to maximize the sampling of the organic molecular configuration. The free energy calculations were again done using PBMetaD with multiple walkers and adaptive Gaussians and the initial configurations of the five walkers were chosen from the last 50 ns of the initial *NVT* simulation. The total sampling time for each system was 2.5 μs (500 ns for each replica) or until convergence.

Dimer Simulations. For the dimer simulations, two chains of the same type were randomly placed within an 8.7 nm cubic simulation box with phosphate ions. Each system underwent energy minimization and *NPT* equilibration similar to those of previous simulations mentioned here. The last frame of *NPT* equilibration was used to start a 100 ns unbiased *NVT* simulation at 300 K; five replicas were selected from this *NVT* simulation. Each of these replicas has been run for 1 μs per replica for a total of 5 μs of simulation time.

PBMetaD with multiple walkers and adaptive Gaussians was once again employed to calculate the dimerization free energy of each system. The production MetaD runs consisted of biasing three main sets of collective variables for a total of 20 CVs. For the first set of CVs, each chain was split up into groups of 4−5 residues, with each of these groups being assigned a bead number. The distance between the center of mass of each bead in the first chain and each bead in the second chain was then biased for a total of 16 bead−bead distances between the two chains. For the next set of CVs, the coordination number between the center of mass (COM) of the positively charged residue head groups for each chain (lysine and arginine) with the center of mass of each phosphate ion in solution was biased, for a total of two coordination numbers per simulation. Lastly, the radius of gyration of each peptide or peptoid chain within the box was biased for a total of two radii of gyration CVs per system.

All systems have reached convergence within the simulation time, as demonstrated by sufficient sampling of the variable of interest, convergence of the time evolution of each free energy profile, and convergence of the binding free energy calculation. The plots and a further discussion of this analysis can be found in the Supporting Information. The figures corresponding to the silica surface simulations are Figure A-S5 (sampling), Figure A-S6 (time evolution of free energy profile), Figure A-S7 (binding free energy), and Figure A-S8 (convergence of the 2D free energy surface), while the figures corresponding to the silicate simulations are Figure A-S9 (sampling) and Figure A-S10 (time evolution of free energy profile). Although it is difficult to obtain a precise error bar from well-tempered metadynamics simulations that use a transient bias potential, the long period of free-diffusion and convergence as demonstrated by Figure A-S5 through Figure A-S7 suggests a lower boundary of uncertainty of 0.41 kJ/mol for the R5 peptide, 0.25 kJ/mol for the R5 peptoid mimic, and 0.58 kJ/mol for the R5 peptoid analogue as determined from taking the standard deviation of the binding energy over the last 40% of the simulation. The PLUMED input files required to reproduce the results reported in this paper are available on PLUMED-NEST (http://www.plumed-nest.org/), the public repository of the PLUMED consortium [77], as plumID: 23.023. These input files also provide the precise definition of each collective variable that was biased, and the relevant parameters used to construct the metadynamics bias potential.

Experimental Methods

Materials. *N*,*N*-Diisopropylcarbodiimide (DIC), bromoacetic acid, and trifluoroacetic acid (TFA) were purchased from Chem-Impex International, Inc. and used as received. The primary amines were synthesized according to literature or commercially available. *N*,*N*-Diisopropylethylamine (DIPEA), All other reagents were obtained from commercial sources and used without further purification. Milli-Q water at 18 MΩ cm was used for all experiments.

Peptoid Material Synthesis. Rink amide resin (0.09 mmol) was used to generate the C-terminal amide peptoid R5. First, the Fmoc groups on the resin were removed in 2 mL of 20% (v/v) 4-methylpiperidine/*N*,*N*-dimethylformamide (DMF) by agitating for 40 min, filtering, and washing with DMF five times. For each DMF wash, 1 mL of DMF was added and then agitated for 1 min. An acylation reaction was then performed on the amino resin by adding 1.5 mL of bromoacetic acid in DMF (0.6 M), followed by addition of 0.30 mL of 50% (v/v) *N*,*N*-diisopropylcarbodiimide (DIC)/DMF. The mixture was agitated for 10 min at room temperature, filtered, and washed with DMF 5 times. Nucleophilic displacement of the bromine with different primary amines occurred by the addition of 1.5 mL of primary amine monomer (0.6 M) in *N*-methyl-2-pyrrolidone (NMP), followed by the agitation for 10 min at room temperature. The monomer solution was filtered from the resin and washed with DMF 5 times. The acylation and displacement steps were repeated until R5 was synthesized.

The final crude product was cleaved from the resin by addition of 2 mL of TFA/water (v/v = 95/5) and agitation for 30 min. The solution was collected by filtration followed by washing the resin with 95% TFA (1 mL, twice). The solvent was then evaporated off under a stream of N_2 gas, yielding the oily crude product, which was dissolved in H_2O/CH_3CN (v/v = 1:1) for HPLC purification. The crude products were purified by reverse-phase HPLC on an XBridge Prep C18 10 μm Optimum Bed Density (OBD) (10 μm, 19 mm × 100 mm), using an adaptable gradient of acetonitrile in H_2O with 0.1% TFA over 15 min. Purified peptoids were analyzed using Waters ACQUITY reverse-phase UPLC (corresponding gradient at 0.4 mL/ min over 7 min at 40 °C with an ACQUITYBEH C18, 1.7 μm, 2.1 mm × 50 mm column) that was connected with a Waters SQD2 mass spectrometry system. The purified peptoid R5 was lyophilized from its solution in a mixture of water and acetonitrile (v/v = 1:1). The peptoid powders were finally divided into small portions (1.0×10^{-6} mol) and stored at −80 °C. UPLC-MS and UPLC characterization for the R5 peptoid analogue and the reverse R5 peptoid analogue can be found in Figures B-S11 and B-S12.

Silica Mineralization and Scanning Electron Microscopy. Orthosilicic acid was freshly prepared before each experiment by vigorously mixing 0.15 mL of tetramethyl orthosilicate with 0.85 mL of 1 mM hydrochloric acid until dissolved. Samples were prepared by resuspending the preweighed lyophilized reagent (forward/reverse peptide/peptoid) in phosphate-citrate buffer ($[Na_2HPO_4]$ = 164.3 mM and [citric acid] = 7.1 mM at pH 7) to reach a final concentration ranging from 1 to 3 mM. To induce mineralization, orthosilicic acid was added to the reagent solution (1:10 volumetric ratio) and vortexed for 10 s. The reaction was allowed to progress at room temperature for 5 min before the vials were centrifuged for 10 min at 15,000*g*. The supernatant was carefully removed, and the pellets were washed three times using Milli-Q water and centrifugation for 10 min at 15,000*g*. The pellets were finally reconstituted in 100 μL of Milli-Q water, and 3 μL of that solution was deposited onto a silicon wafer and allowed to air-dry overnight before SEM imaging. Micrographs were acquired on FEI Sirion XL30 SEM operated at 2 kV and a spot size of 1 using through-lens detection (TLD) and the ultrahigh resolution (UHR) mode.

Surface Plasmon Resonance Experiments. 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. Chips were cleaned with ethanol and deionized water before use. They were mounted on the four-channel flow cell of an SPR sensor from the Institute of Photonics and Electronics (Prague, Czech Republic). The channels were allowed to equilibrate for 15 min by using a phosphate-citrate buffer.

Two channels were then used to run forward R5 peptide and forward R5 peptoid samples at 25 μM. The third was used to run phosphate− citrate buffer as a baseline/negative control. After 30 min, all channels were washed with a phosphate−citrate buffer until a plateau was reached. All experiments were conducted at a flow rate of 40 μL min^{-1} and at room temperature.

Chapter 4: A Comparative Study of Silica-Binding Peptides on Amorphous Silica and Ordered Quartz

4.1 Abstract

Organic-inorganic hybrid nanomaterials are of interest in applications including catalysis, energy storage, and biomedicine. A wide diversity of organic materials, such as proteins, peptides, and polymers, can impart unique properties on resulting material morphologies, including structural characteristics that can be used to tune function. However, this relationship between organic templates, ideal reaction conditions, and resulting nanomaterial is often hard to elucidate. In this work, we seek to evaluate the binding affinity of a number of silica-binding peptides on an amorphous silica and crystalline quartz surface to validate previous work that has demonstrate high binding affinity as a good indicator of monodisperse nanoparticle precipitation. We provide suggestions for future computational studies to speed calculations and explore new morphologies as well as future experimental studies to identify the relationship between biomolecule sequence, binding energy, and biomolecular concentration in mineralization experiments necessary for the production of monodisperse nanoparticles.

4.2 Introduction

Organic-inorganic hybrid nanomaterials have become a popular class of novel materials for their diverse range of functionalities, including applications in catalysis, electronics, biomedicine, and sensing [112–115]. Organic templates, or structure directing agents, such as polymers, proteins, or ligands, play a key role in determining the morphology, structure, and properties of the resulting hybrid materials, as has been seen in the formation of nanoparticles [116–119], generation of quantum dots [120–124], or synthesis of zeolites [125,126] or metal-organic frameworks [127–129]. However, the relationship between these organic templates, ideal reaction conditions, and the resulting nanomaterial morphologies and structures has often proven difficult to elucidate. One method of approaching this lack of understanding is to rely on bioinspired materials, or materials that mimic natural biological systems, such as those that are utilized in the process of silica biomineralization [130]. In the natural process of silica biomineralization, organisms, like marine diatoms or sponges for example, use proteins within their cell walls to direct the formation of complex hierarchical nanostructures. This process can be synthetically mimicked with the same proteins responsible for the natural process (silaffin in diatoms or silicatein in sponges), precipitating silica nanoparticles with relatively controlled shape and size [40,86,87,131,132]. The building blocks of proteins, smaller peptide units, also present promising opportunities for synthetic mineralization, as they can similarly be used to precipitate silica nanostructures with often easier synthesis, milder reaction conditions, and for more highly specific functions compared to the larger proteins [41,71,85,89,91,99,112,133].

Silica traditionally has been used for catalysis and separations, but more recently has seen adoption in biotechnological fields as part of therapeutic or diagnostic systems. Multiple peptides are known to precipitate silica, including the silaffin-derived R5 peptide and its mutants, as well as LK repeat peptides or poly-L-lysine, which all rely on the positively charged sidechains of lysine or arginine [80–83]. Shorter peptide sequences (<10 residues) with these common residues have also been highlighted for their ability to self-assemble into templates, like beta sheets or nanowires, that can then be used in the growth of more complex silica nanostructures including nanotubes or fibers [134–140]. Additionally, work has shown that silica-precipitating peptides can be conjugated with block polymers to create small monodisperse particles or in conjunction with silk proteins or silk films to create engineered silk-silica chimeras [141–144].

Detailed experimental and computational studies have begun to elucidate the key relationships between peptide sequence and resulting morphology, including studies of individual post-translational modifications on commonly identified silica precipitating peptides [85,89,91]. Molecular dynamics simulations have been increasingly utilized in this space as well, showing that the phosphorylation of serine side chains can explain the pH dependent behavior of silica formation in R5, clarifying the molecular structure of peptides responsible for stable biosilica sheets, or illustrating the structure of long-chain polyamines at silica interfaces [90,92,93]. Molecular dynamics simulations and enhanced sampling methods were also recently used to rationally design biomimetic polymers similar to the R5 peptide with experimentally validated higher mineralization efficiency and then define key descriptors for this higher efficiency, namely biomolecule/silica binding affinity [71].

In this work, we gather disparate experimental studies of silica biomorphogensis using peptides and perform the most comprehensive set of peptide/surface binding simulations of silica forming systems. We seek to validate this key descriptor for specific silica morphologies using a wider range of peptides with varying sequence length, structure, residue character, and charge in order to expand our understanding of biomimetic mineralization and the key factors relating sequence and morphology. We perform molecular dynamics simulations of these peptides on silica and quartz to understand the differences (if any) in binding to an amorphous v. crystalline surface and then seek to identify the particular threshold of binding affinity responsible for the development of monodisperse nanoparticles at different biomolecule concentrations. We highlight the key features of a strong binding sequence versus a weak binding sequence and provide suggestions on future computational studies in this area that might be able to speed necessary calculations, explore new morphologies, or provide data for future machine learning studies.

4.3 Methods

To provide a comprehensive analysis of mineralizing peptides in relation to silica and quartz, nine different short peptide sequences with known silica mineralization ability were chosen from the literature [40,101,112,138–140,145–148]. The sequences and expected silica morphologies, as well as charge in solution at pH 7.5 are listed in Table 4.1 below.

Table 4.1: Table depicting the silica-binding peptide name, amino acid sequence, a description of the silica morphology produced, the peptide charge at pH 7.5, and the reference from which each sequence was identified.

Name	Sequence	Description of Morphology	Peptide Charge	Reference
LK Alpha	LKKLLKLLKKLLKL	Spherical particles	+6	145, 148
Pep B	HKKPVPPKPKPEPKPQ	Spherical particles	+4	101
R5	SSKKSGSYSGSKGSKRRIL	Spherical particles	+6	40
V6K	VVVVVVK	Fiber or stack	+1	140
A6K	AAAAAAK	Flower like structures or fiber	+1	140
ALP	KFFAAK	Nanotube	+2	139
Si-4	MSPHPHPRHHHT	Network of partially fused spheres	+1	146, 147
Pep-1	KSLSRHDHIHHH	Network of partially fused spheres	+1	146
P11	QQRFQWQFQQQ	Nanotube or nanosheet	+1	138

All sequences were constructed with an acetyl N-terminus (ACE) and an N-methyl C-terminus (NME) for consistency. The structure of each sequence was built using tleap [149]. The silica surface was modeled as an amorphous structure with Q3 silanol group termination, as described in previous work [71]. The ionization state of the silanol groups was tuned to 14.5% to reflect a solution pH of 7.5. We used the (100) face of quartz surface, with a mix of Q2 and Q3 silanol group terminations. To approximate a pH of 7.5, the percentage of the deprotonated silanol groups was set to 16.5% with the deprotonated sites being randomly chosen [150]. The silica simulation box had x-y dimensions of 8.09 by 8.31 nm and the quartz simulation box had x-y dimensions of 6.68 by 6.97 nm; the z-dimension for each system varies slightly but is around 9.65 nm for all amorphous silica simulations and 13.2 nm for all quartz simulations. All systems were modeled using the CHARMM36 forcefield [68,69] and explicit TIP3P water [72]. The silica and quartz surfaces were modeled using the INTERFACE forcefield [107] due to its compatibility with this CHARMM forcefield.

Each peptide was placed in the simulation box above either the silica or the quartz surface depending on the simulation type and then solvated with approximately 17,500 water molecules using Packmol [151]. The charge of each peptide was neutralized using sodium (+1 charge) and/or phosphate ions (-2 charge), which were modelled on previous work from Mao et al [102]. All simulations were performed in GROMACS [67] with the PLUMED [75] plugin for enhanced sampling. The LINCS [108] algorithm was used to constrain all bonds between hydrogen and heavy atoms; all production NVT calculations were additionally performed with hydrogen mass repartitioning so that a timestep of 4 fs could be used [73].

Energy minimization was carried out on each system for a maximum of 10 picoseconds using the steepest descent algorithm and then followed by NPT equilibration at 300 K and 1 bar for 500 picoseconds using the Berendsen barostat with a pressure damping parameter of 1.0 and a stochastic global thermostat with a temperature damping parameter of 0.1 [74]. Lennard Jones and Coulombic interactions were calculated with a cutoff of 1.0 nm and long-range electrostatic interactions were treated with the Particle Mesh Ewald method. Due to the semi-infinite structure of the slabs, pressure coupling was applied only to the z dimension during the NPT equilibration. The center of each slab (both amorphous silica and quartz) was frozen during all simulation steps, allowing only the silanol groups and the top layer of the surfaces to move [71,150]. Following this NPT equilibration, all systems underwent a 100 ns NVT equilibration, from which 8 initial configurations of each system were chosen.

The binding free energy of each system was calculated using Parallel-Bias Metadynamics [109], with multiple walkers [110] and adaptive Gaussians [111] employed to speed the convergence of these simulations. Collective variables chosen include the orthogonal distance between the center of mass of the backbone of the peptide to the center of mass of the top of each silica or quartz surface, the orthogonal distance between the center of mass of each amino acid headgroup in the peptide and the center of mass of the top of either surface, and the radius of gyration of the peptide [71,150]. The total sampling time for each system was 6.4 microseconds (800 ns per replica), which we have deemed sufficient for convergence of the calculation.

4.4 Results and Discussion

Comparison of Sequence Binding on Silica and Quartz

Prior work in the field has shown that computationally predicted adsorption energies show correlation with particular morphological outcomes, namely that higher peptide binding energy to a silica surface corresponds to the production of monodisperse nanospheres through mineralization experiments [71]. In order to validate this hypothesis with new data, we performed a series of binding energy calculations for nine peptides on amorphous silica as well as quartz.

Unsurprisingly, the similarities between the amorphous silica and ordered quartz surfaces lead to very similar binding trends between peptides for each surface, as shown in Figure 4.1. The peptides that bind strongest to both surfaces are those with the highest overall charge: LK Alpha (+6 charge), Pep B (+4 charge), and the R5 peptide (+6 charge). However, LK Alpha binds almost 15 kJ/mol stronger than the next two peptides, despite the R5 peptide having the exact same charge. The peptides with lower charges bind roughly in the 0-20 kJ/mol range with fairly low driving forces towards adsorption, even though these systems have been experimentally proven to precipitate silica under mineralization conditions. We believe systems like this, composed of lower charged peptides with fewer residues, might need to form larger aggregates in order to template the growth of specific silica morphologies [135,137–140]. This claim is experimentally supported in literature by a number of works that highlight the more complex arrangements these peptides may form, such as the self-assembled antiparallel beta sheets necessary for mineralization using P11 [138] as well as templated nanofibers constructed using V6K or A6K [140]. We recommend that these morphologies be further investigated through coarse-grained methods, as the computational cost of investigating these much larger systems using atomistic molecular dynamics and the chosen enhanced sampling methods would be prohibitively high.

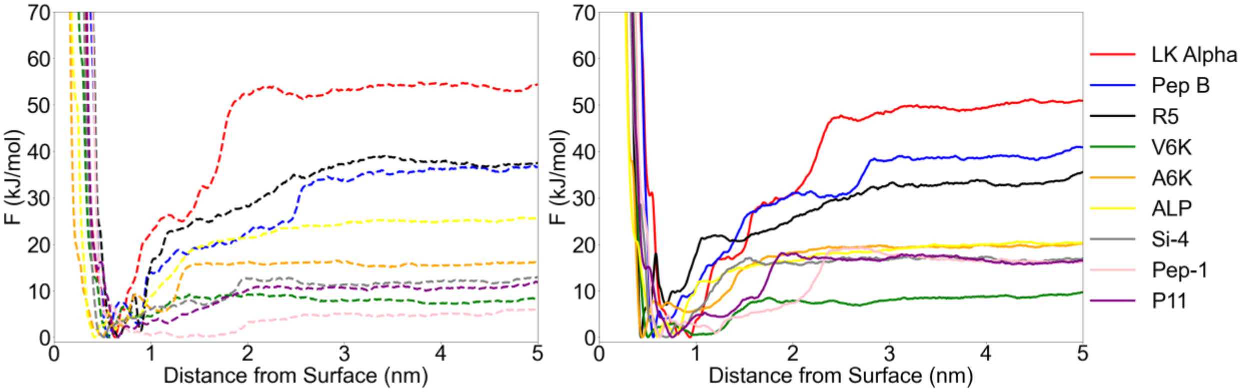


Figure 4.1: Binding free energy as a function of peptide distance from the surface for the silica surface (left, dashed lines) and for the quartz surface (right, solid lines). The legend is shown on the right; colors remain consistent across peptide name.

There are four peptides from the above figure that we identified as having similar morphologies to the monodisperse nanospheres or networks of partially fused spheres that we have focused on in previous work. LK Alpha, the strongest binding peptide, produces monodisperse nanoparticles which is consistent with prior hypotheses relating high adsorption energy to monodisperse particle production. R5, another highly charged and strong binding peptide, produces this morphology as well, although lower concentrations of the peptide see a transition from monodisperse nanoparticles to partially fused spheres. On the other hand, the Si-4 and Pep-1 peptides both have low binding energy at around 10 kJ/mol and produce networks of partially fused spheres, as shown in Figure 4.2 below.

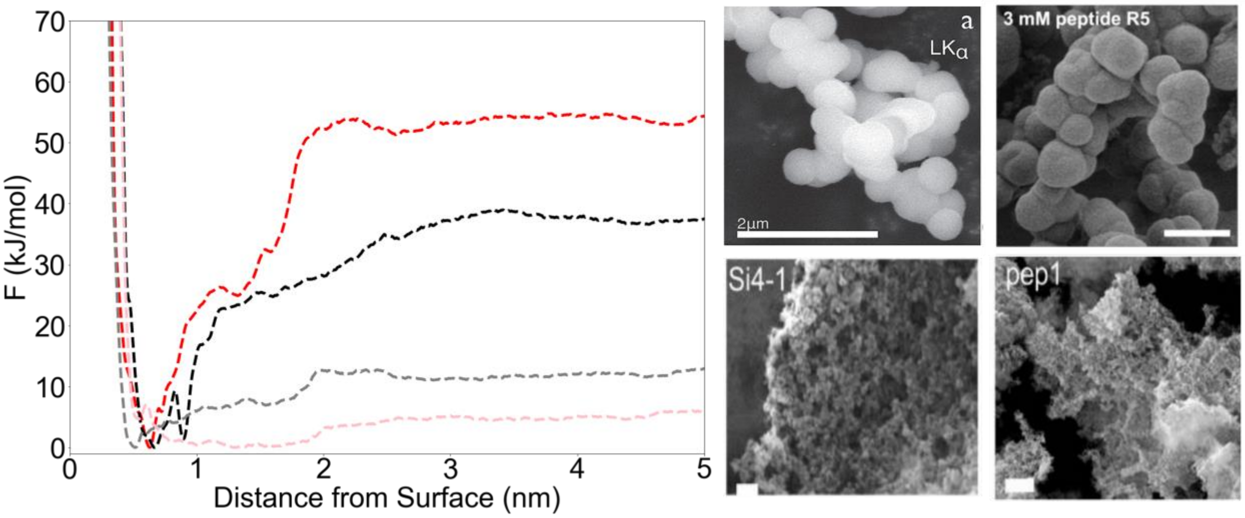


Figure 4.2: The binding free energy as a function of distance from the silica surface for LK Alpha (red), R5 (black), Si-4 (grey), and Pep-1 (pink). The SEM images of silica morphologies on the right have been reproduced from multiple sources and include LK Alpha[3] (top left), R5[4] (top right), Si-4[5] (bottom left), and Pep-1[8] (bottom right) [71,146,148].The two peptides with the highest binding energy, LK Alpha and R5, produce monodisperse nanoparticles while Si-4 and Pep-1, with significantly lower binding energies, produce a network of partially fused spheres.

While the SEM images shown above were taken under different mineralization conditions (i.e. silicate precursor used, concentration of biomolecule, buffer conditions, etc.) due to the number of literature sources these images were collected from, we believe that they provide a good validation metric for our hypothesis on the relationship between adsorption energy and morphology if we can perform a consistent panel of mineralization experiments. We also believe that these systems could also shed light on the concentration dependence of certain morphological constructs, which we will discuss in the proposed future work section.

Signatures of Strong Binding v. Weak Binding Sequences

We also use the data from these simulations to comprehensively validate what features make a strong binding peptide. In the future, this information can be leveraged for improved nanomaterials synthesis. Previous work has shown that strong binding is often attributed to peptides that can maintain either an extended state on the surface or a very specific preferred

---

[3] Adapted with permission from *J. Am. Chem. Soc.* 2014, 136, 43, 15134–15137. Copyright 2014 American Chemical Society.

[4] Adapted with permission from *Chem. Mater.* 2024, 36, 2, 786–794. Copyright 2024 American Chemical Society.

[5] Adapted with permission from *J. Am. Chem. Soc.* 2012, 134, 14, 6244–6256. Copyright 2012 American Chemical Society.

conformation [71,96], which we have explored below in Figure 4.3 through the use of two-dimensional free energy surfaces that relate the distance from the surface to the radius of gyration of each peptide. In Figure 4.3A, we highlight the three strongest binding peptides (LK Alpha, Pep B, and the R5 peptide) and demonstrate that these three peptides all achieve fairly close distances to the surface and maintain one clear binding mode to the surface. There is a very small range of ideal radii of gyration for each peptide to happily bind to the surface, as demonstrated by the very small regions of dark red on the plots. In contrast, the weaker binding peptides, Si-4 and Pep-1, shown in Figure 4.3B have a much wider range of equal energy conformations shown by the large collection of red in each figure. It is important to note here that not all of these peptides are of the same length, so natural differences in the range of radii of gyration possible or preferred are inherent. We are more specifically looking for a small range of possible strong-binding radii of gyration, not necessarily just a longer radius of gyration. It is also worth noting that the y-axis in Figure 4.3A and Figure 4.3B are different, as Si-4 and Pep-1 are both much shorter than R5 or Pep B. However, when you look at the range of radii of gyration with similar binding for all five systems you see that the preferred conformations for LK Alpha, R5, and Pep B are all within 0.05 nanometers of each other, while the range for Si-4 and Pep-1 is closer to 0.15-0.20 nanometers.

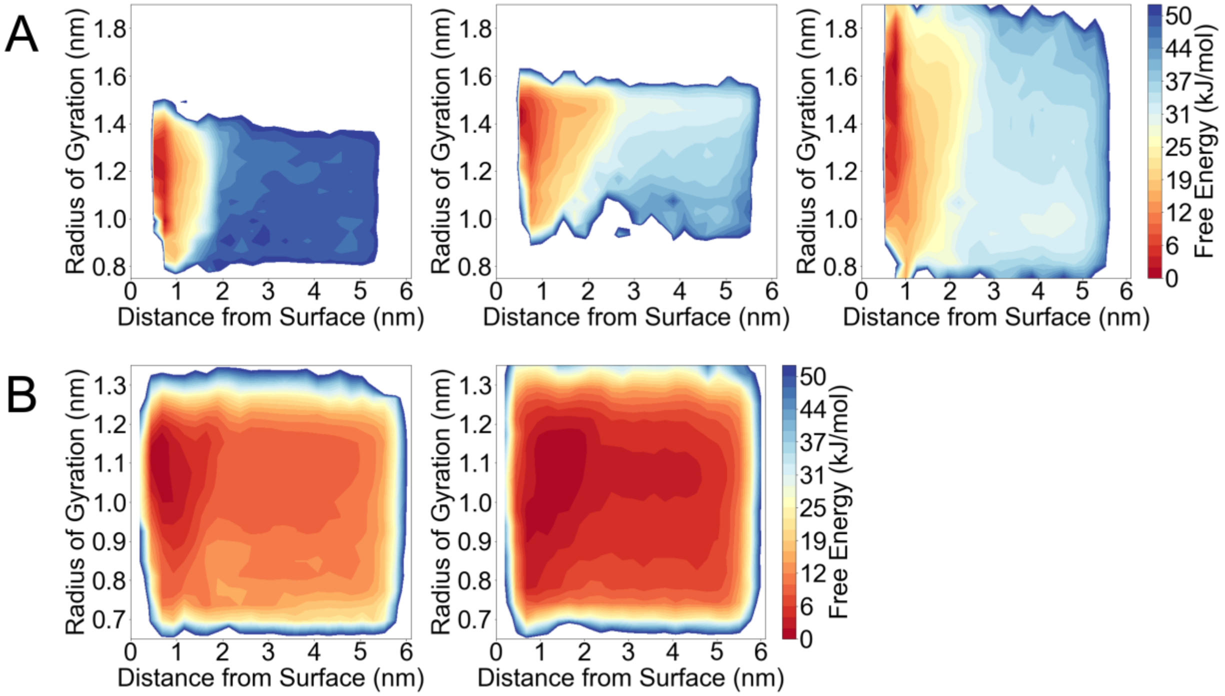


Figure 4.3: Two-dimensional free energy surfaces showing the radius of gyration of selected peptides as a function of the distance of the peptide from the silica surface for (A) stronger binding peptides and (B) weaker binding peptides. The stronger binding peptides from left to

right are LK Alpha, Pep B, and R5, while the weaker binding peptides from left to right are Si-4 and Pep-1.

We believe that as the presence of a limited number of available similar energy conformations is key to designing peptides that bind strongly to surfaces. This is likely due to the ability of these peptides, when in an extended or highly specific conformation on the surface, to maneuver positively charged sidechains closer to the negatively charged silica and quartz surfaces [71]. This idea is illustrated in Figure 4.4 below, in which histograms of the distance of each residue from the surface for the bound states are shown for the highest binding peptide and the lowest binding peptide on silica. Figure 4.4A shows that LK Alpha, the system that binds with the most specificity and the highest adsorption energy, is able to maneuver a number of the positively charged sidechains (shown in red) closer to the surface than the lowest binding system, shown in Figure 4.4B. The placement of positive charges throughout the LK Alpha sequence also holds the non-positively charged residues shown in black much closer to the surface, as opposed to Pep-1, which seems to lose the interaction of the last half of the sequence with the surface. This is shown by much wider distance distributions for residues 7 through 14 that seem to be mostly at 1.75-2 nm away from the surface COM.

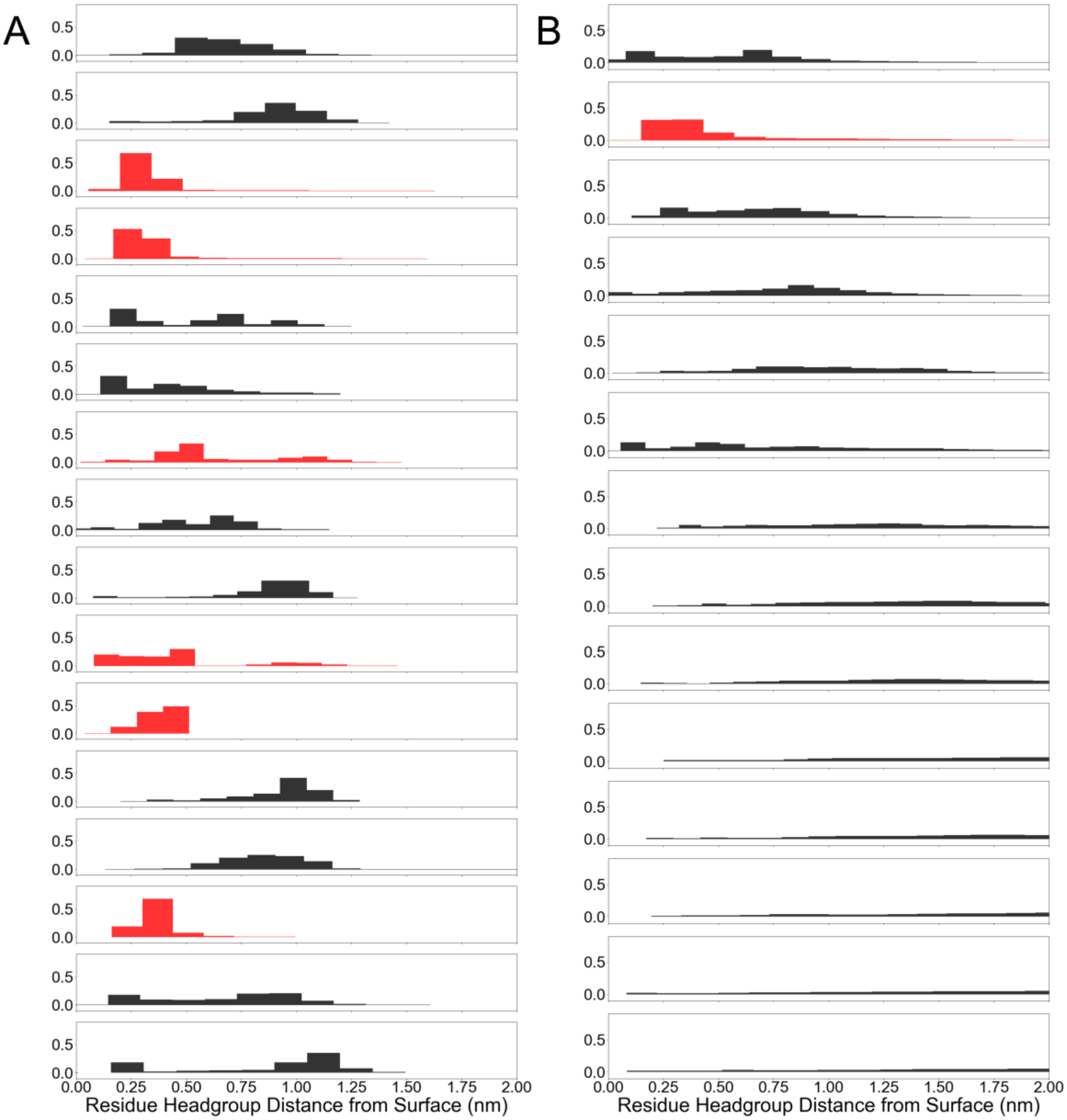


Figure 4.4: Histograms detailing the frequency of close binding residue headgroups for (A) LK Alpha, a strong binding peptide, and (B) Pep-1, a weak binding peptide. Each column shows one histogram for every residue in the sequence, starting with the ACE termini and ending with the NME termini for each peptide. The positively charged residue headgroups in each sequence are shown in red.

4.5 Conclusions and Future Work

The similar characteristics between amorphous silica and quartz lead to similar trends between the studied peptides on each surface, with most peptides showing a slight preference for the amorphous surface over the ordered one. The peptides with the highest overall positive charge see the strongest adsorption energy with the surface, while the other peptides may need to form larger aggregates to see noteworthy adsorption energies. The peptides with the highest surface binding energy to amorphous silica also are capable of producing monodisperse nanoparticles, while those with significantly lower binding energy (like Si-4 and Pep-1) how large networks of partially fused spheres, indicating that this could be an interesting avenue for future combined experimental and computational studies to pursue in more detail. The strongest binding sequences all have very specific, clearly defined binding modes that often demonstrate a high level of extension on the surface, and we believe this is due the ability of these sequences to maneuver their positively charged sidechains closer to the amorphous silica surface. This study validated conclusions drawn from previous work, but we hope to expand on this work with the proposed studies below.

Planned Experimental Validation

The SEM images for the LK Alpha, R5, Si-4, and Pep-1 peptides shown in Figure 4.2 are reproduced from a collection of literature sources to illustrate that stronger surface binding energy correlates with an ability to produce monodisperse nanoparticles as opposed to networks of partially fused spheres. While these literature sources likely maintained similar mineralization and SEM protocols, we are interested in validating our calculations with a complete panel of identical mineralization studies performed at multiple different concentrations of molecular precursor, hoping to narrow down the relationship between a specific computationally predicted binding energy and the concentration of molecular precursor that shifts expected morphology from monodisperse nanoparticles to partially fused spheres. These mineralization experiments and subsequent imaging are currently in progress with experimental collaborators in the Weidner Group, located at Aarhus University in Denmark.

Speeding Future Calculations

Additionally, the binding energy calculations for each of these peptides are computationally intensive, taking multiple months to reach convergence and hundreds of thousands of CPU hours to complete. Achieving a comprehensive analysis of multiple solid-binding peptides on different inorganic surfaces using PBMetaD, while accurate, will be too time and resource-consuming for most groups and research goals. Recent advances in computational methods have indicated that new methods are available for rapid estimation of such binding energies. The binding of a peptide to an inorganic surface can be estimated as a linear combination of the free energy contribution of each residue in the sequence, based on the statistics of a thermodynamically stable structure ensemble [150]. This method has been validated so far for five silica-binding peptides and we intend to provide validation for the peptides discussed in this work as well, on both the silica and quartz surfaces.

A model like this, which can achieve the binding energy calculation with roughly 10% of the computational cost as the PBMetaD calculation, will extend the opportunity for high-throughput simulation in this space to screen a number of sequences and identify those with binding energies suitable for specific research goals. This data could also be helpful in training machine learning algorithms in order to predict solid binding peptide effectiveness in the future.

Chapter 5: Conclusions and Future Work

The work presented here has advanced the state of the art in molecular modelling for understanding sequence-structure-function relationships in novel folding polymers and for biomimetic mineralization.

In chapter two, recent progress in the rapid development of atomistic forcefield parameters for sequence-defined polymers was leveraged to create a new library of peptoid residues, including those with varying levels of chiral, hydrophobic, or charged characteristics. This library was utilized along with the enhanced sampling technique of steered molecular dynamics to devise a simulation method that allows us to relate peptoid polymer sequence to single-chain secondary and tertiary structure in helical hairpins capable of mimicking native protein structures. We performed, to our knowledge, the first MD simulations of peptoid helical hairpins based on four sequences previously identified by experimentalists and compared the resulting structures using clustering to discuss the important features imparted by each residue type in water and acetonitrile. This was followed by an in-depth comparison of the handedness preference of a helical peptoid sequence that demonstrated how changing the solvent environment from water to acetonitrile can encourage more specific peptoid helical handedness and stabilize secondary structure characteristics in these materials, while still allowing tertiary structure to form despite the difference in hydrophobic driving forces. Here, we have provided a roadmap for future studies to utilize physics-based simulations to relate polymer sequence and structure. This work allows for nanoscale insight into peptoid structures that have previously only been probed through experimental methods like circular dichroism, providing a more accurate resolution to assess structural features and hopefully design more complex peptoids with enzymatic-mimicking functions.

In chapter three, we demonstrated the advantages of sequence-defined polymers in modulating silica morphogenesis by leveraging ease of synthesis with a wider set of possible sidechains. We achieve an unprecedented level of biomimicry in the creation of entirely synthetic polymer sequences that produce silica nanostructures with defined shape and size and identify the biomolecule/surface binding energy as a key computational marker in predicting sequences that might produce similar nanostructures. The developed simulation framework can be utilized with future biomolecular sequences to understand and predict possible silica morphologies prior to experimental synthesis and validation.

In chapter four, we utilized the simulation framework from chapter three on a wider range of peptide sequences and validated our ability to predict a specific morphology given biomolecule/surface binding energy. We identified necessary experiments to be performed by our experimental collaborator that could provide insight on relationships between computationally determined biomolecule/surface binding energy and the concentration of biomolecule necessary to produce defined nanospheres. We also narrowed in on the regions of parameter space in which this model is not successful, primarily on short sequences that must first self-assemble into a larger, more positively charged template in order to participate meaningfully in silica morphogenesis. We identify possible routes to understanding these more complex nanostructures using different modelling approaches, as well as highlighting how the calculations from the simulation framework identified originally in chapter three can be completed in smaller timeframes at a lower computational cost.

The work discussed here can also grow the field of molecular modelling; the expertise in peptoid forcefield development combined with the parameter library used in chapter two can be leveraged to speed high-throughput simulation efforts in the area of peptoid modelling, which has previously been held back by the lack of off-the-shelf forcefield options in this area. These high-throughput simulation efforts along with an understanding of secondary and tertiary structure characteristics in sequences like those in chapter two can now be applied to different systems, like metal chelating sequences, to move closer to the idea of developing peptoid enzymes for complex functions like catalysis. The enhanced sampling methods highlighted in chapter three and built upon in chapter four can be applied to different interfaces, like calcium phosphate surfaces, to understand the growth of teeth and bone from bioinspired peptides. The computational signatures identified within also present an opportunity for calculations derived from simulation to serve as data for machine learning algorithms to narrow down regions of parameter space that might benefit from more thorough modelling approaches, further speeding the rational design of bioinspired materials.

Appendix A: Supporting Information for Chapter 3

Reweighting Biased Simulations for Unbiased Ensemble Averages

Ensemble averages and reweighted probability distributions for relevant observables (e.g., sidechain distance distributions) were obtained from PBMetaD simulations using the quasi-static bias approximation [109].

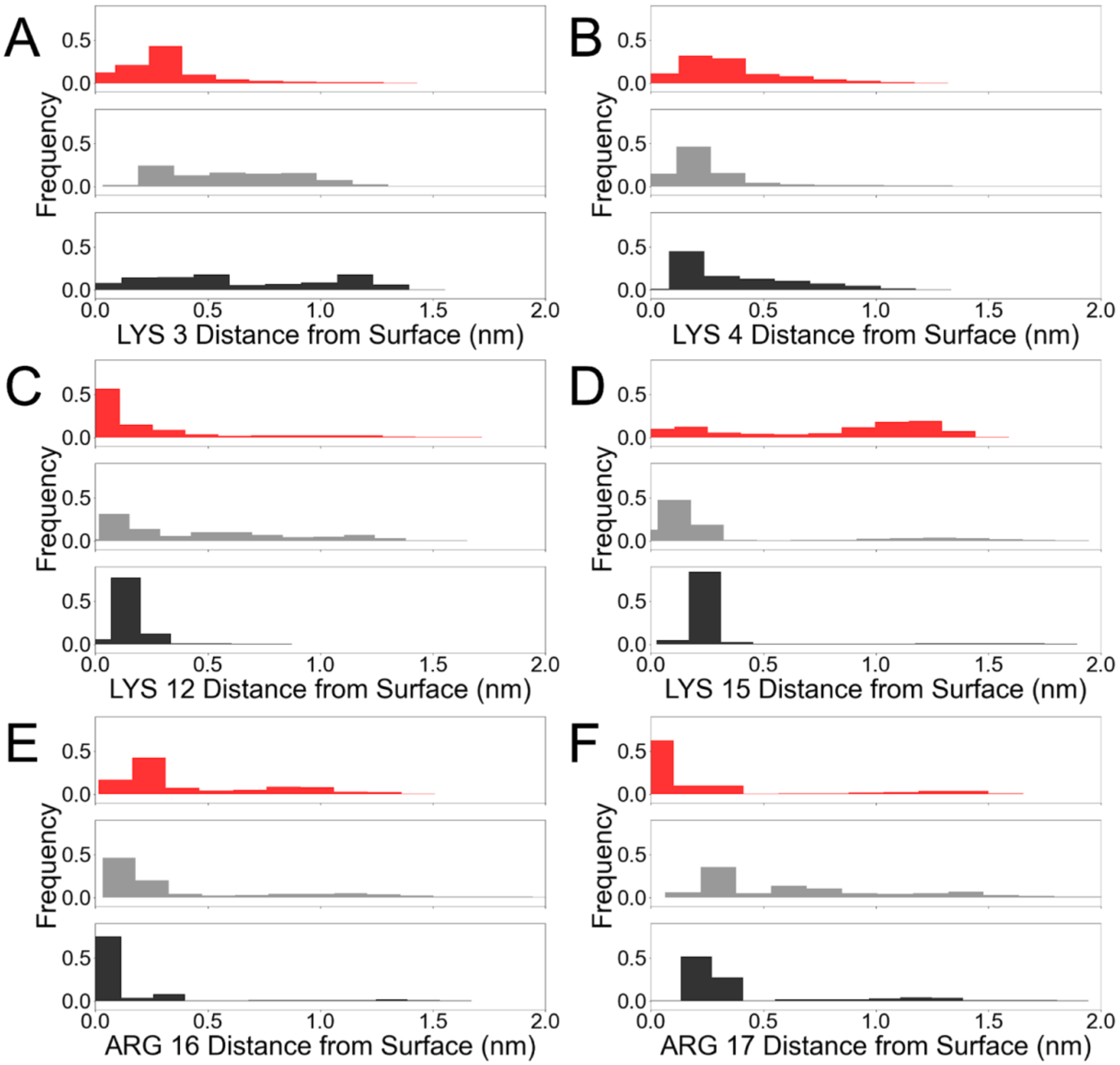


Figure A-S1: Histograms of the distance between the amorphous silica surface and the center of mass of the lysine headgroup in residue 3 (A), lysine headgroup in residue 4 (B), lysine headgroup in residue 12 (C), lysine headgroup in residue 15 (D), arginine headgroup in residue 16 (E), and arginine headgroup in residue 17 (F). The R5 peptide is shown in red, the R5 peptoid mimic in grey, and the synthesized R5 peptoid analog in black.

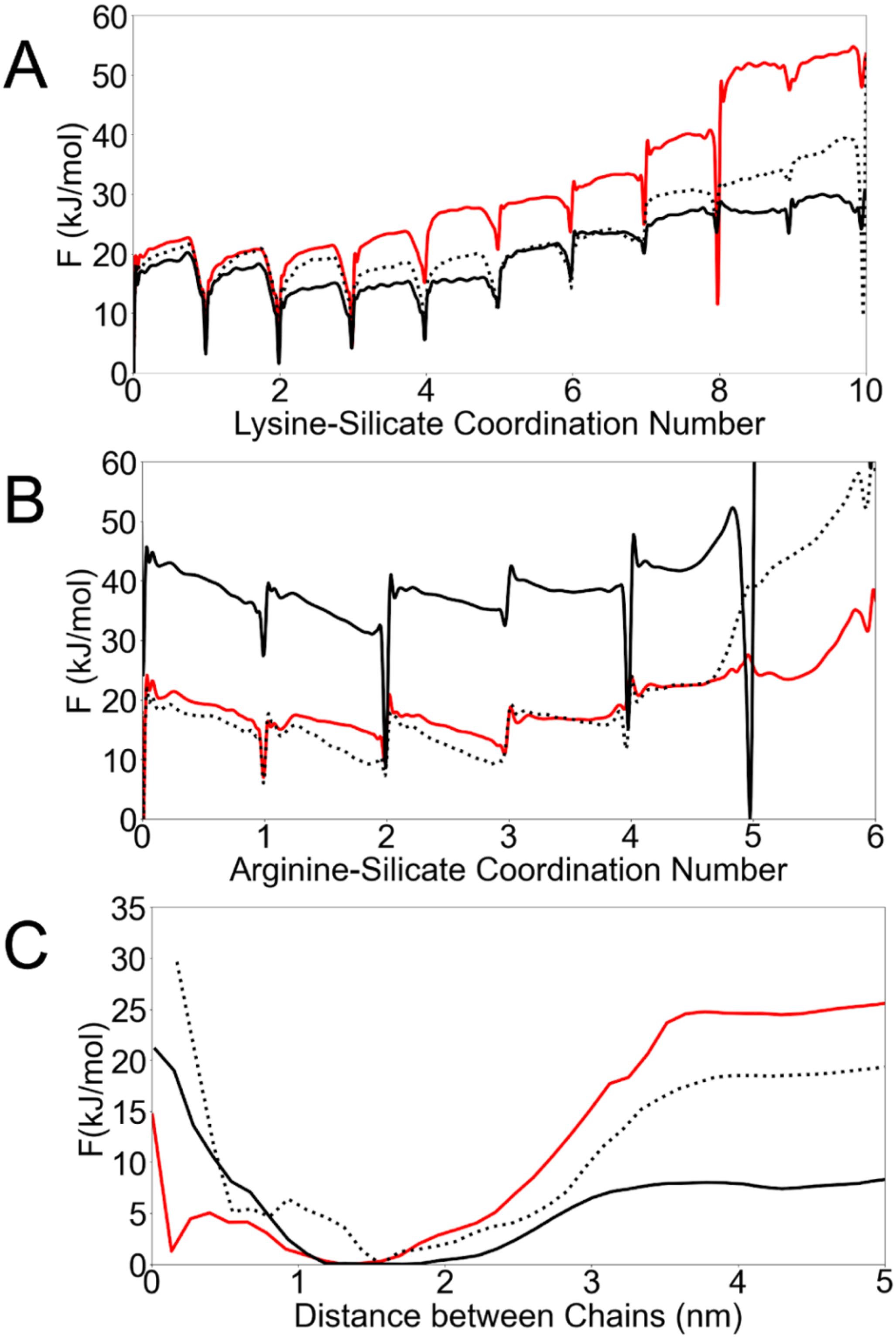


Figure A-S2: Free energy profiles of the interactions between lysine headgroups and silicate anions (A) and arginine headgroups and silicate anions (B) from the performed silicate simulations. The R5 peptide is shown in red, the R5 peptoid mimic is shown in dotted black, and the R5 peptoid analog is shown in solid black. Panel C shows the reweighted free energy profile of dimerization for each system, indicating that the R5 peptoid analog has the lowest propensity to aggregate with another molecule of the same type.

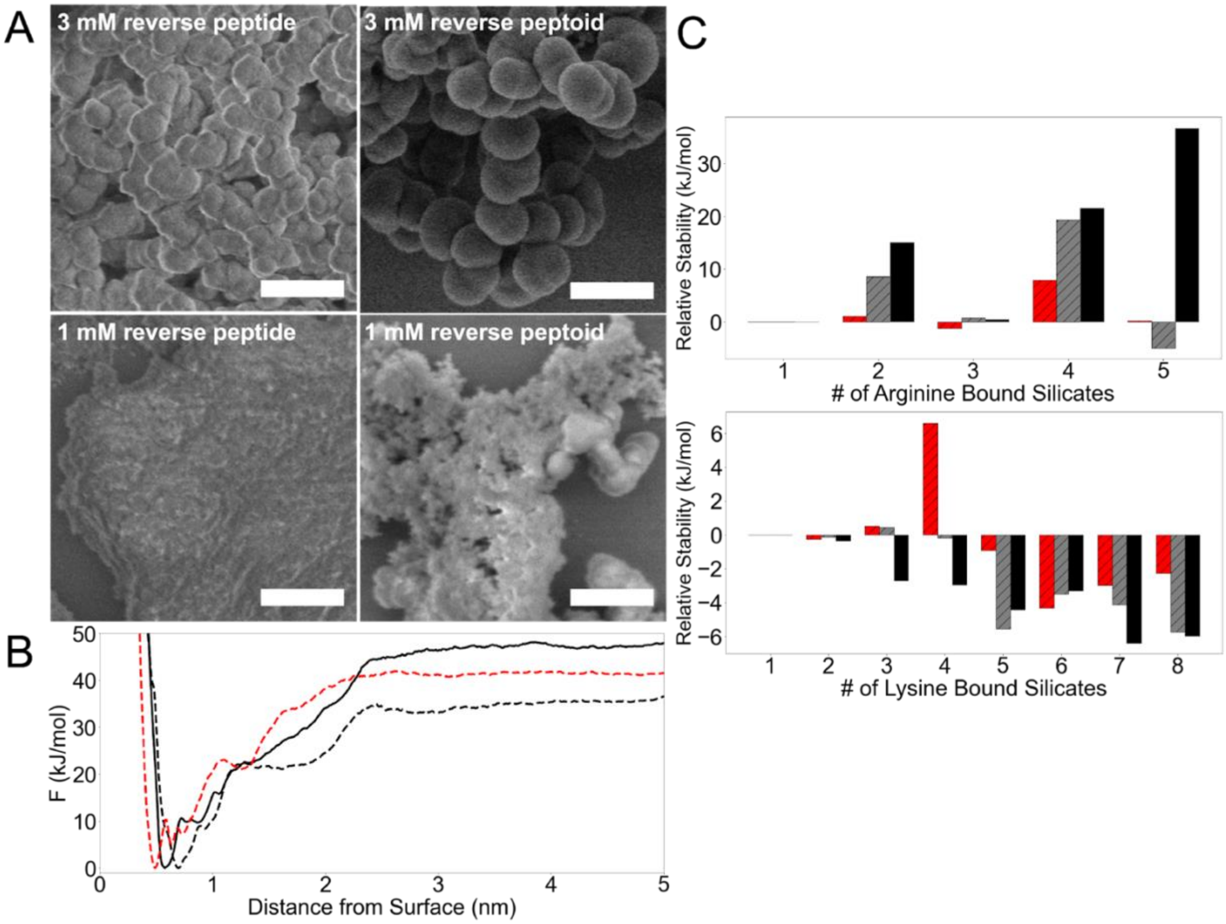


Figure A-S3: Panel A shows SEM images of resulting silica nanospheres for R5 peptide reverse (LEFT) and R5 peptoid analog reverse (RIGHT) at high concentrations of biomolecule (TOP) and low concentrations of biomolecule (BOTTOM) with all scale bars equal to 1 µm. Panel B shows a comparison of the binding free energy of the COM of each molecule backbone from an amorphous silica surface, with R5 peptoid analog in black, R5 peptide reverse in dashed red, and R5 peptoid analog reverse in dashed black. Panel C shows bar plots of the relative stability of interactions between the arginine headgroup and the silicate anion (TOP) and the lysine headgroup and the silicate anion (BOTTOM) normalized to one anion. Positive numbers indicate higher stability and negative numbers indicate lower stability.

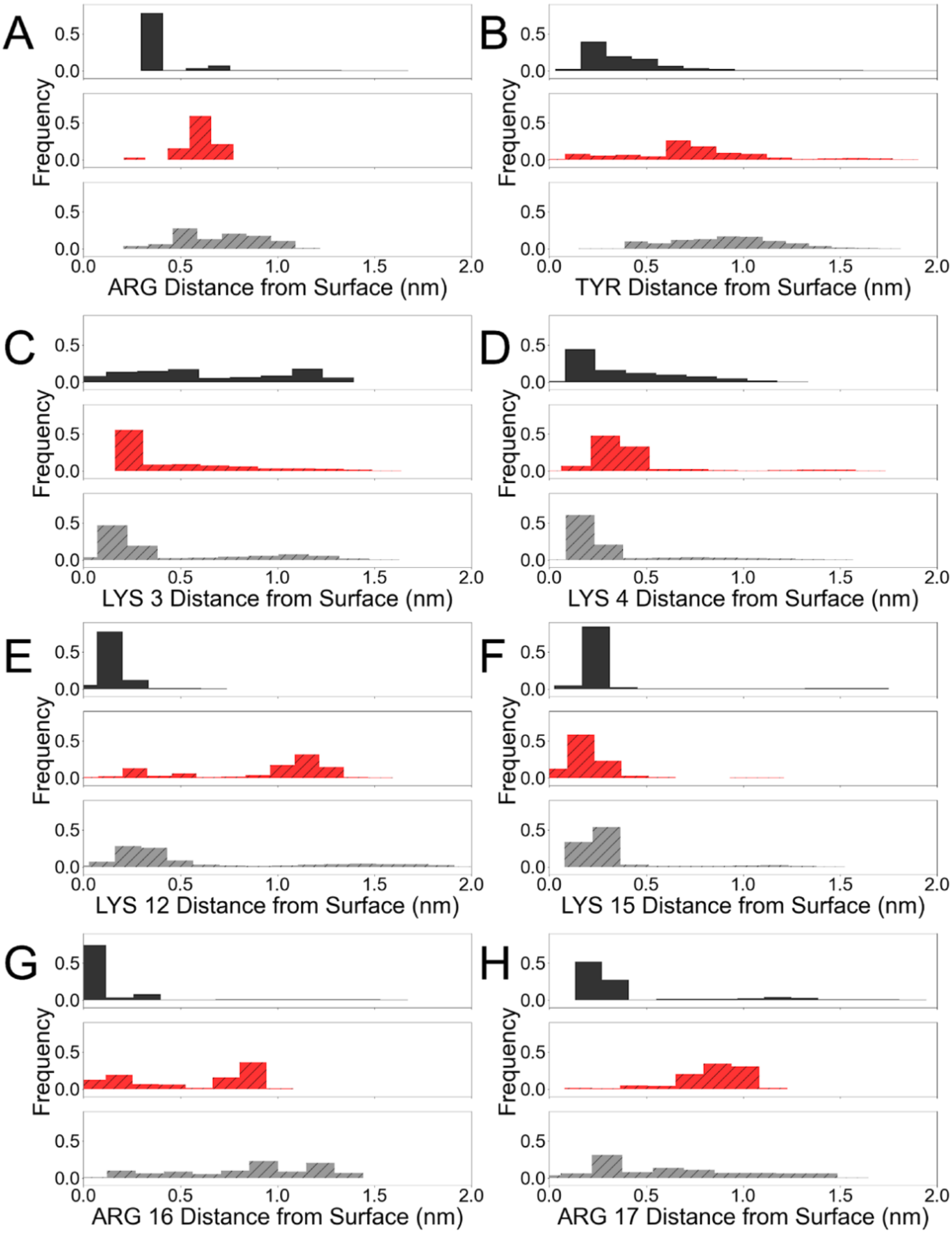


Figure A-S4: Histograms of the distance between the amorphous silica surface and the center of mass of both arginine residues (A), the tyrosine headgroup (B), the lysine headgroup in residue 3 (C), lysine headgroup in residue 4 (D), lysine headgroup in residue 12 (E), lysine headgroup in residue 15 (F), arginine headgroup in residue 16 (G), and arginine headgroup in residue 17 (H). The synthesized R5 peptoid analog is shown in black, while the R5 peptide reverse is shown in red with black hatch marks and the R5 peptoid analog reverse is shown in grey with black hatch marks.

Convergence of Metadynamics Calculations

Each metadynamics calculation was run sufficiently long to ensure convergence, which was assessed for the surface binding simulations through sufficient sampling of the backbone center of mass distance of each system from the amorphous silica surface as shown in Figure A-S5. The time evolution of each free energy profile has also been provided to assess convergence, as shown in Figure A-S6, and convergence of the binding free energy calculations have been shown in Figure A-S7. For the silicate simulations, sufficient sampling of the arginine-silicate and lysine-silicate interactions are shown in Figure A-S8, while the time evolution of the free energy profile of the arginine-silicate interactions for each system has been provided in Figure A-S9.

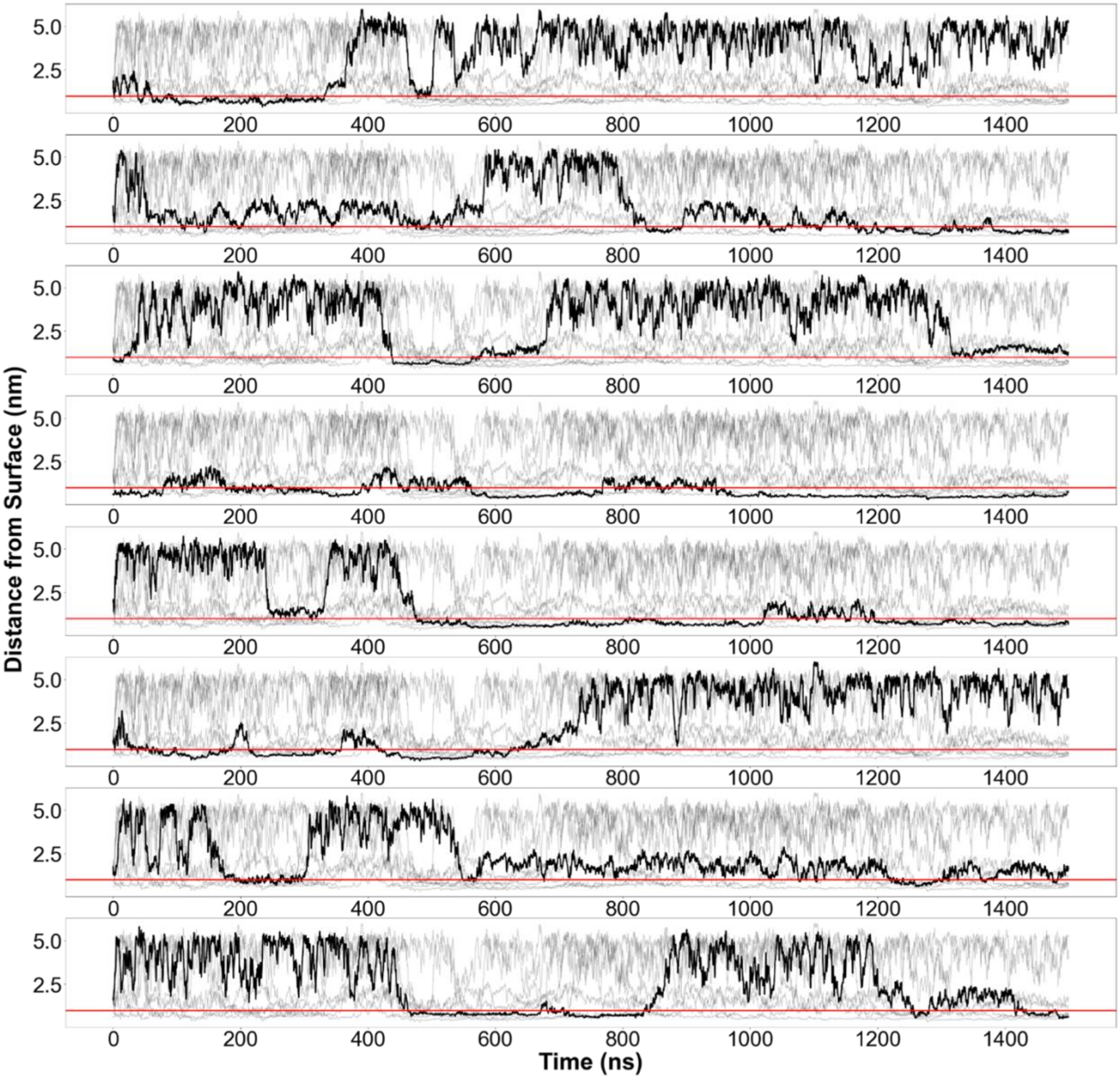


Figure A-S5: Sampling of the center of mass distance between the backbone atoms of the R5 peptide and the amorphous silica surface for each walker in each subplot over the total simulation time. The red line indicates the transition between surface bound and unbound states. Similar sampling and convergence was obtained for all systems.

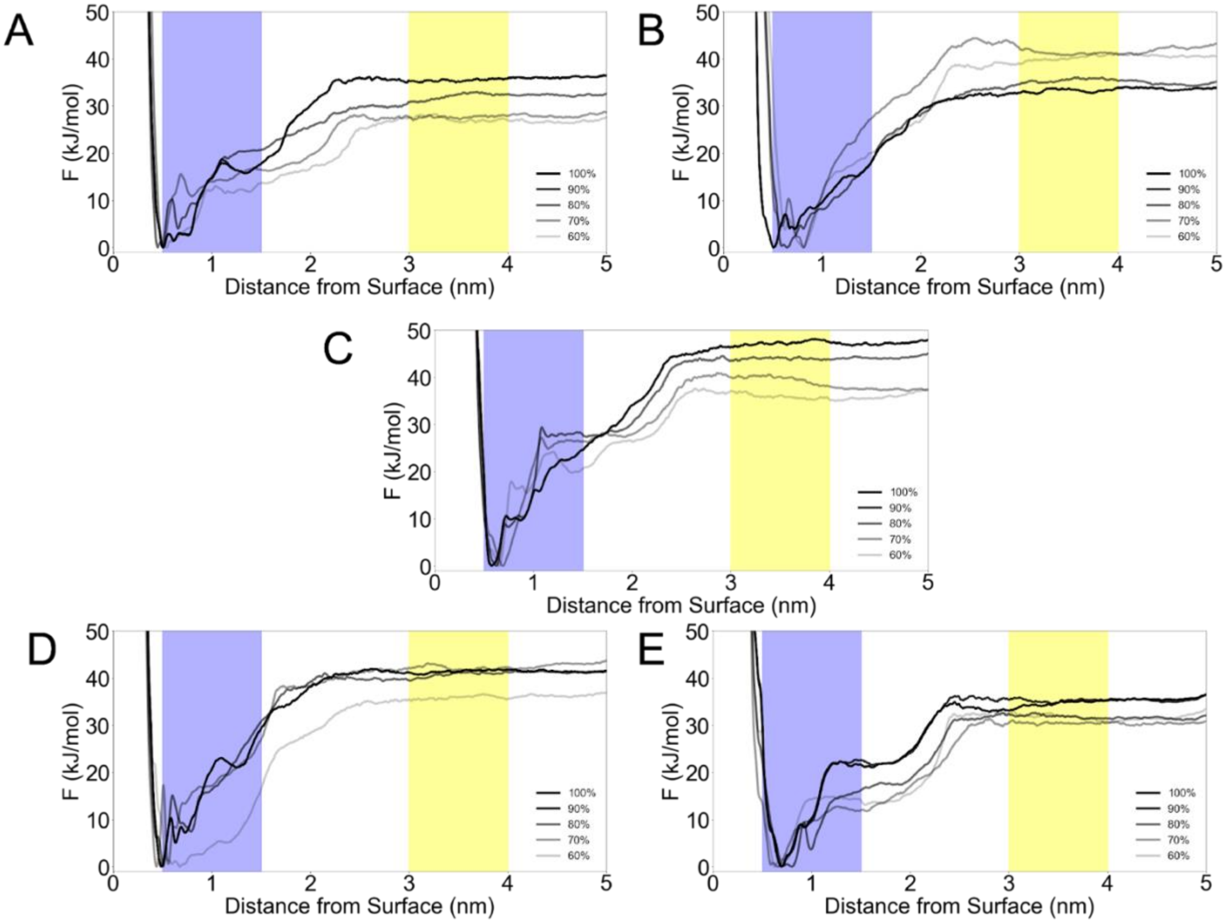


Figure A-S6: Time evolution of the free energy profiles of (A) R5 peptide, (B) R5 peptoid mimic, (C) R5 peptoid analog, (D) R5 peptide reverse, and (E) R5 peptoid analog reverse on amorphous silica. The x-axis is the orthogonal distance between the center of mass (COM) of each molecule backbone and the surface. Each legend indicates the percentage of simulation time relative to the total simulation time (1.5 microsecond/walker). The blue region has been highlighted to represent the surface bound state while the yellow region has been highlighted to represent the solution phase state for the calculation of binding free energy for each peptide on each surface.

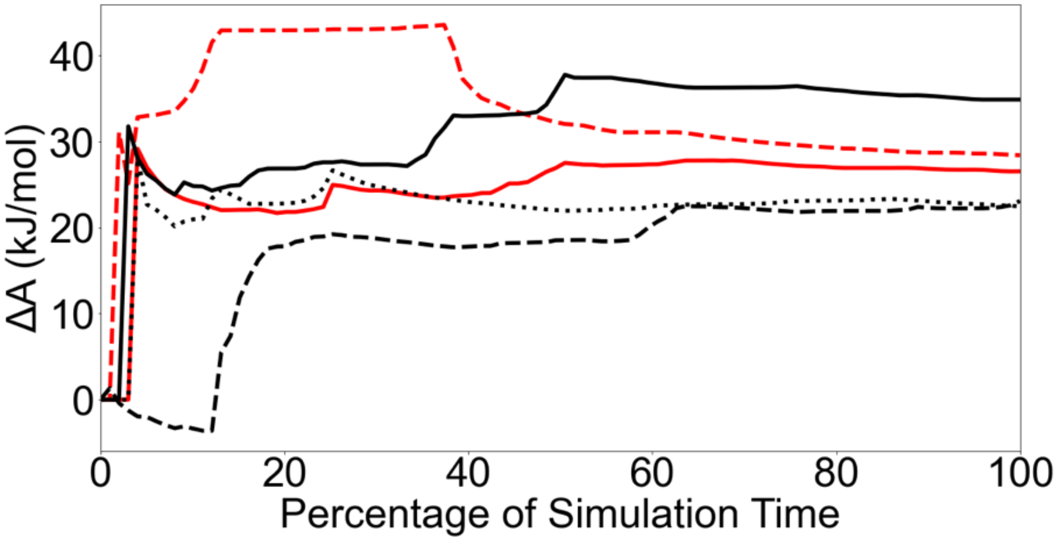


Figure A-S7: Free energy of binding calculated as a function of total simulation time (1.5 microseconds) for each molecule on the amorphous silica surface. The R5 peptide is shown in solid red, the R5 peptoid mimic in dotted black, the R5 peptoid analog in solid black, the R5 peptide reverse in dashed red, and the R5 peptoid analog reverse in dashed black. Convergence is indicated by a flattening of the curve to one consistent value.

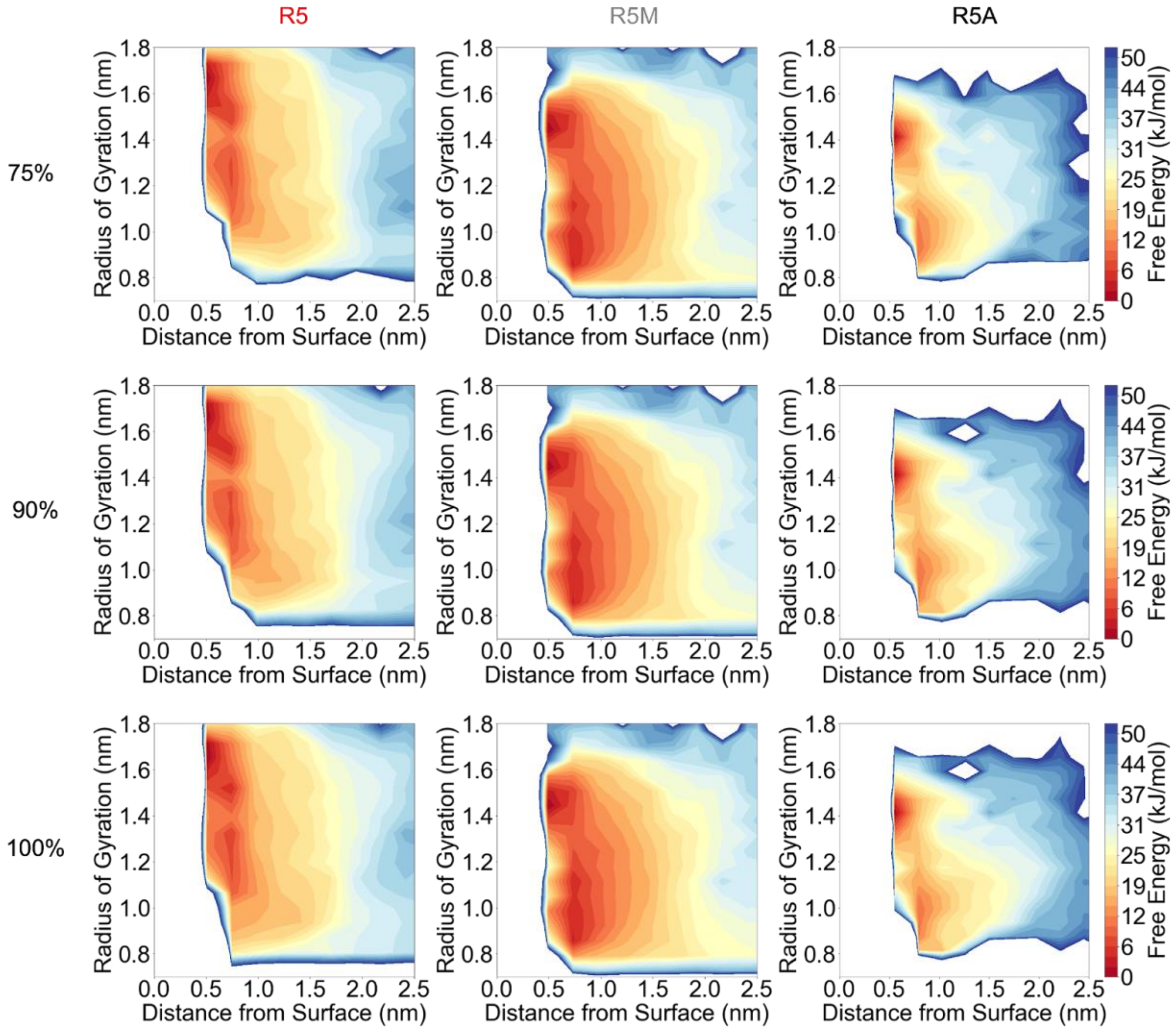


Figure A-S8: Two-dimensional free energy surfaces for the R5 peptide (LEFT), R5 peptoid mimic (MIDDLE), and the R5 peptoid analog (RIGHT) at 75% of simulation time (TOP), 90% of simulation time (MIDDLE), and 100% of simulation time (BOTTOM). The important minima shown are relatively the same across this section of simulation time, indicating convergence of our calculations and likely low uncertainty.

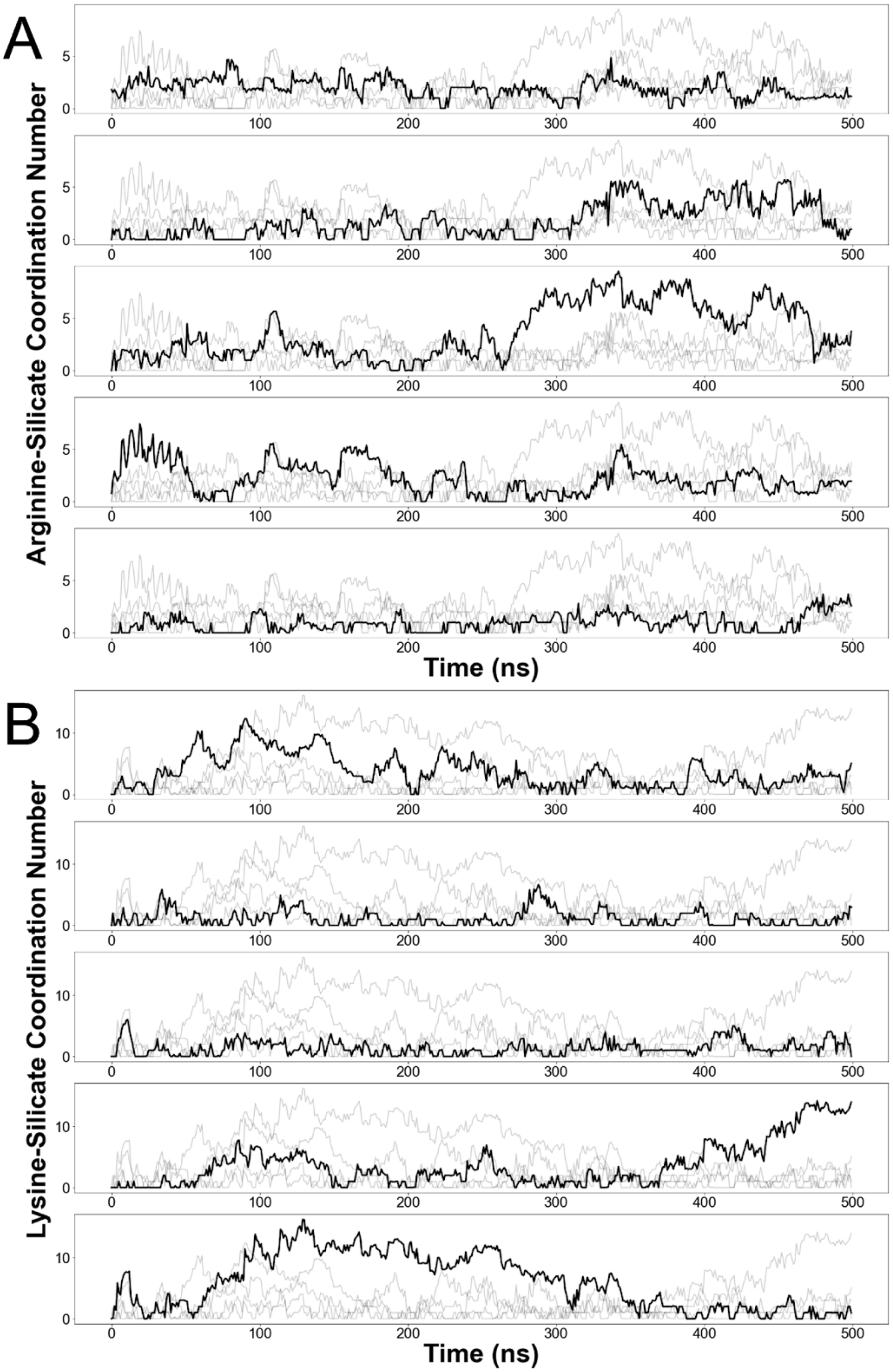

A
Arginine-Silicate Coordination Number
0
5
0
100
200
300
400
500
Time (ns)
B
Lysine-Silicate Coordination Number
0
10
0
100
200
300
400
500
Time (ns)

Figure A-S9: Sampling of the coordination number between the arginine headgroup and silicate anions (TOP) and the lysine headgroup and silicate anions (BOTTOM) for R5 peptide for each walker in each subplot over the total simulation time. Similar sampling and convergence was obtained for all systems.

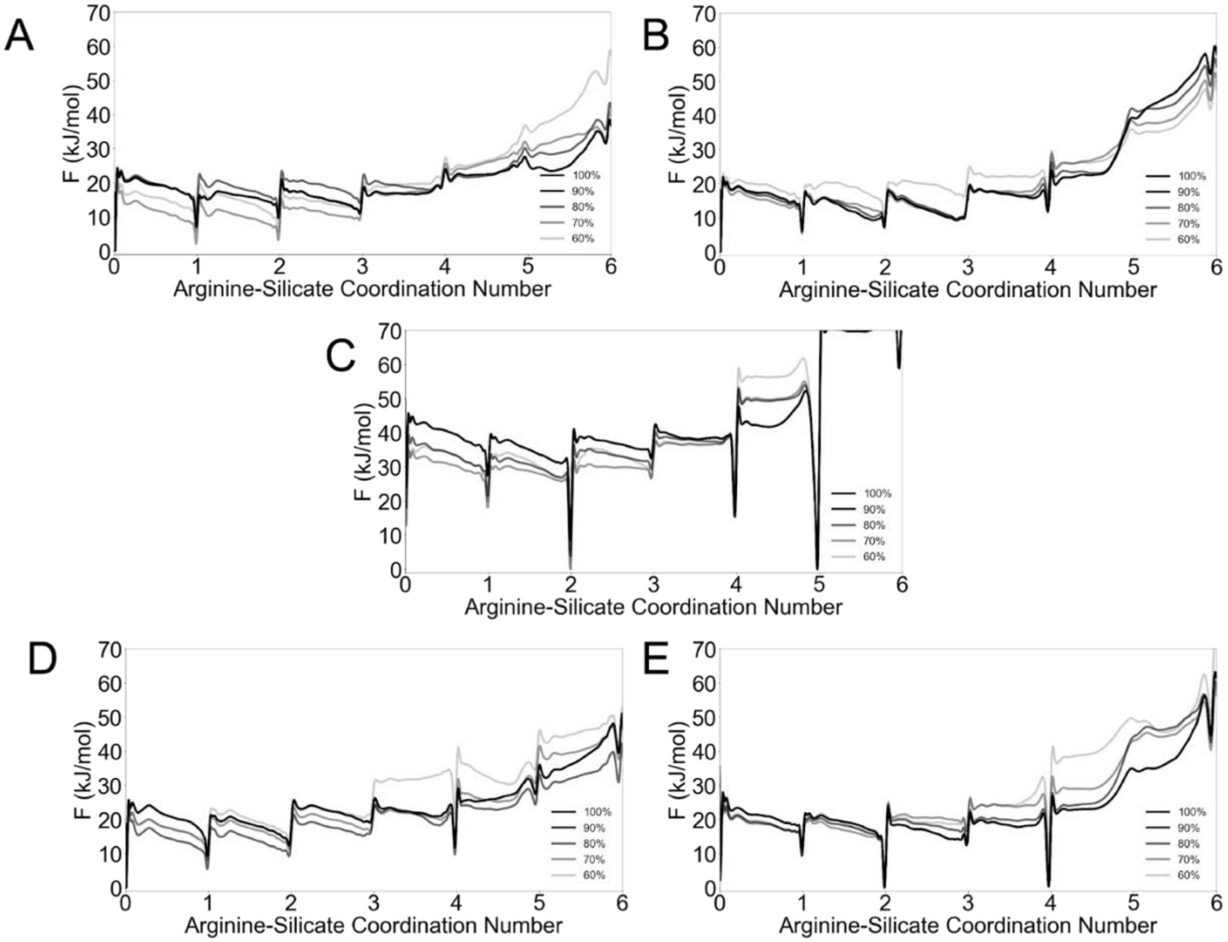


Figure A-S10: Time evolution of the free energy profiles of the arginine-silicate interactions in (A) R5 peptide, (B) R5 peptoid mimic, (C) R5 peptoid analog, (D) R5 peptide reverse, and (E) R5 peptoid analog reverse. The x-axis is the coordination number between arginine headgroups and silicate anions. Each legend indicates the percentage of simulation time relative to the total simulation time (500 ns/walker). Similar convergence was seen in the lysine-silicate coordination number profiles.

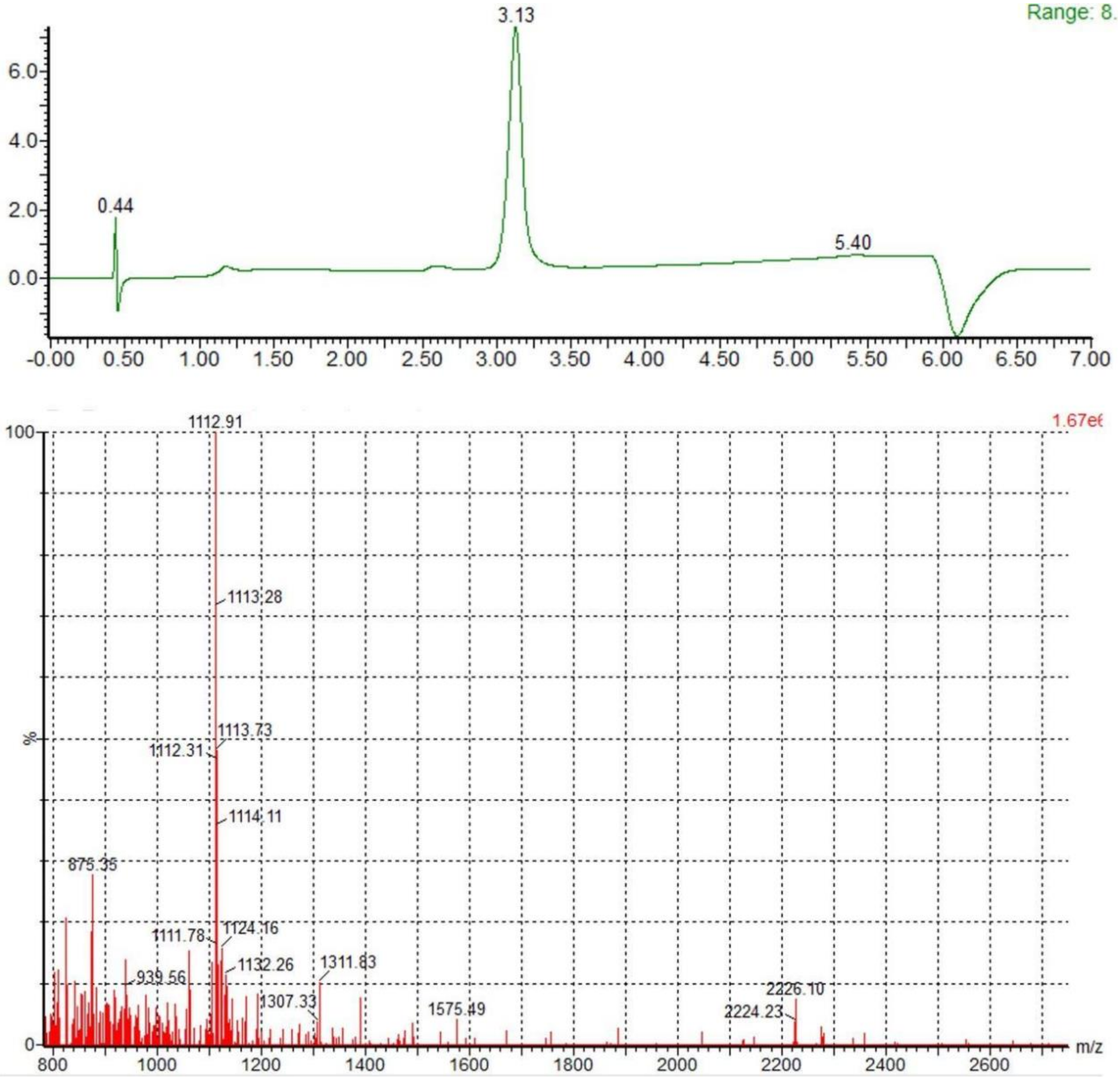


Figure A-S11: UPLC-MS characterization of toidR5A (TOP) and UPLC characterization of toidR5A with the gradient of 0 - 40% CH$_3$CN in H$_2$O (BOTTOM). MS characterization of R5. 2222.7 (Molecular weight), 1112.3 (Found: [M+2H]$^{2+}$)

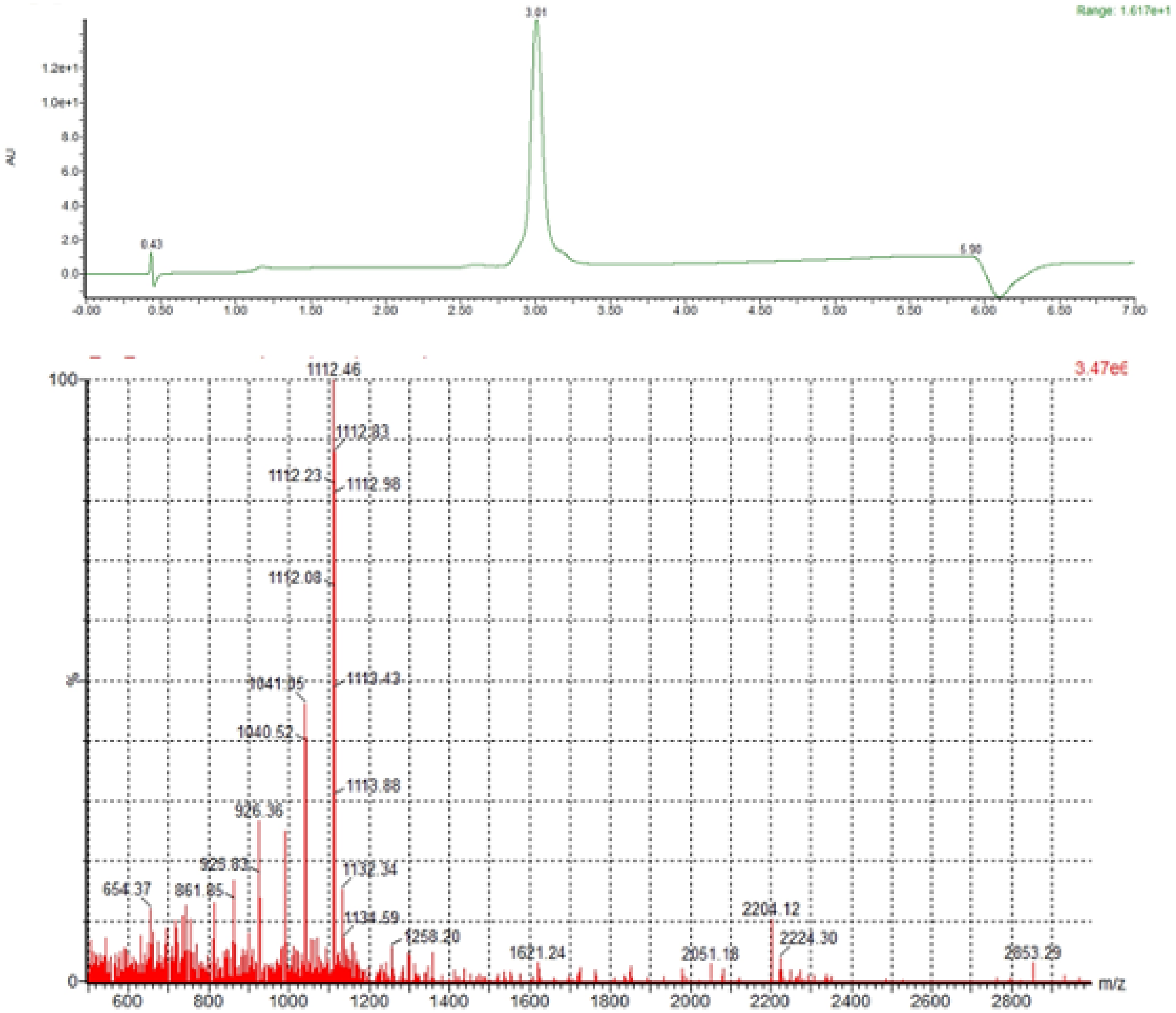


Figure A-S12: UPLC-MS characterization of reverse toidR5A (TOP) and UPLC characterization of reverse toidR5A with the gradient of 0 - 40% CH_3CN in H_2O (BOTTOM). MS characterization of r-R5. 2222.7 (Molecular weight), 1112.5 (Found: $[M+2H]^{2+}$) The UPLC shows a small shoulder around the major peak at t=3.01 minutes. The shoulder has nearly identical MS with the center major peak at t=3.01 min, indicating these are not significant impurities but more likely different protonation states of the guanidinium and amino groups on the arginine and lysine analog sidechains.

References

(1) Morris, R.; Black, K. A.; Stollar, E. J. Uncovering Protein Function: From Classification to Complexes. *Essays Biochem* **2022**, *66* (3), 255–285. https://doi.org/10.1042/EBC20200108.

(2) Serçinoğlu, O.; Ozbek, P. Sequence-Structure-Function Relationships in Class I MHC: A Local Frustration Perspective. *PLoS One* **2020**, *15* (5), e0232849. https://doi.org/10.1371/journal.pone.0232849.

(3) Koehler Leman, J.; Szczerbiak, P.; Renfrew, P. D.; Gligorijevic, V.; Berenberg, D.; Vatanen, T.; Taylor, B. C.; Chandler, C.; Janssen, S.; Pataki, A.; Carriero, N.; Fisk, I.; Xavier, R. J.; Knight, R.; Bonneau, R.; Kosciolek, T. Sequence-Structure-Function Relationships in the Microbial Protein Universe. *Nat Commun* **2023**, *14* (1), 2351. https://doi.org/10.1038/s41467-023-37896-w.

(4) Hussain, A.; Brooks III, C. L. Guiding Discovery of Protein Sequence-Structure-Function Modeling. *Bioinformatics* **2024**, *40* (1). https://doi.org/10.1093/bioinformatics/btae002.

(5) Sadowski, M. I.; Jones, D. T. The Sequence–Structure Relationship and Protein Function Prediction. *Curr Opin Struct Biol* **2009**, *19* (3), 357–362. https://doi.org/10.1016/j.sbi.2009.03.008.

(6) Ferruz, N.; Heinzinger, M.; Akdel, M.; Goncearenco, A.; Naef, L.; Dallago, C. From Sequence to Function through Structure: Deep Learning for Protein Design. *Comput Struct Biotechnol J* **2023**, *21*, 238–250. https://doi.org/10.1016/j.csbj.2022.11.014.

(7) Chowdhury, R. Editorial: Advances in Protein Structure, Function, and Design. *Front Bioeng Biotechnol* **2023**, *10*. https://doi.org/10.3389/fbioe.2022.1108962.

(8) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat Rev Mol Cell Biol* **2019**, *20* (11), 681–697. https://doi.org/10.1038/s41580-019-0163-x.

(9) Gao, W.; Mahajan, S. P.; Sulam, J.; Gray, J. J. Deep Learning in Protein Structural Modeling and Design. *Patterns* **2020**, *1* (9), 100142. https://doi.org/10.1016/j.patter.2020.100142.

(10) Watson, J. L.; Juergens, D.; Bennett, N. R.; Trippe, B. L.; Yim, J.; Eisenach, H. E.; Ahern, W.; Borst, A. J.; Ragotte, R. J.; Milles, L. F.; Wicky, B. I. M.; Hanikel, N.; Pellock, S. J.; Courbet, A.; Sheffler, W.; Wang, J.; Venkatesh, P.; Sappington, I.; Torres, S. V.; Lauko, A.;

De Bortoli, V.; Mathieu, E.; Ovchinnikov, S.; Barzilay, R.; Jaakkola, T. S.; DiMaio, F.; Baek, M.; Baker, D. De Novo Design of Protein Structure and Function with RFdiffusion. *Nature* **2023**, *620* (7976), 1089–1100. https://doi.org/10.1038/s41586-023-06415-8.

(11) Sandak, A.; Butina Ogorelec, K. Bioinspired Building Materials—Lessons from Nature. *Front Mater* **2023**, *10*. https://doi.org/10.3389/fmats.2023.1283163.

(12) Wang, Y.; Naleway, S. E.; Wang, B. Biological and Bioinspired Materials: Structure Leading to Functional and Mechanical Performance. *Bioact Mater* **2020**, *5* (4), 745–757. https://doi.org/10.1016/j.bioactmat.2020.06.003.

(13) Lampel, A. Biology-Inspired Supramolecular Peptide Systems. *Chem* **2020**, *6* (6), 1222–1236. https://doi.org/10.1016/j.chempr.2020.03.005.

(14) Ongey, E.; Pflugmacher, S.; Neubauer, P. Bioinspired Designs, Molecular Premise and Tools for Evaluating the Ecological Importance of Antimicrobial Peptides. *Pharmaceuticals* **2018**, *11* (3), 68. https://doi.org/10.3390/ph11030068.

(15) Yu, X.; Wang, Z.; Su, Z.; Wei, G. Design, Fabrication, and Biomedical Applications of Bioinspired Peptide–Inorganic Nanomaterial Hybrids. *J Mater Chem B* **2017**, *5* (6), 1130–1142. https://doi.org/10.1039/C6TB02659A.

(16) Scornet, N.; Delarue-Cochin, S.; Azoury, M. E.; Le Mignon, M.; Chemelle, J.-A.; Nony, E.; Maillère, B.; Terreux, R.; Pallardy, M.; Joseph, D. Bioinspired Design and Oriented Synthesis of Immunogenic Site-Specifically Penicilloylated Peptides. *Bioconjug Chem* **2016**, *27* (11), 2629–2645. https://doi.org/10.1021/acs.bioconjchem.6b00393.

(17) Kong, H.; Yang, G.; He, P.; Zhu, D.; Luan, X.; Xu, Y.; Mu, R.; Wei, G. Self-Assembly of Bioinspired Peptides for Biomimetic Synthesis of Advanced Peptide-Based Nanomaterials: A Mini-Review. *Nano Futures* **2023**, *7* (1), 012001. https://doi.org/10.1088/2399-1984/acafbe.

(18) Kim, S.; Kim, D.; Jung, H. H.; Lee, I.; Kim, J. IL; Suh, J.; Jon, S. Bio-Inspired Design and Potential Biomedical Applications of a Novel Class of High-Affinity Peptides. *Angewandte Chemie International Edition* **2012**, *51* (8), 1890–1894. https://doi.org/10.1002/anie.201107894.

(19) Yu, J.; Horsley, J. R.; Abell, A. D. Peptides as Bio-Inspired Electronic Materials: An Electrochemical and First-Principles Perspective. *Acc Chem Res* **2018**, *51* (9), 2237–2246. https://doi.org/10.1021/acs.accounts.8b00198.

(20) Sun, J.; Zuckermann, R. N. Peptoid Polymers: A Highly Designable Bioinspired Material. *ACS Nano* **2013**, *7* (6), 4715–4732. https://doi.org/10.1021/nn4015714.

(21) Li, Z.; Cai, B.; Yang, W.; Chen, C.-L. Hierarchical Nanomaterials Assembled from Peptoids and Other Sequence-Defined Synthetic Polymers. *Chem Rev* **2021**, *121* (22), 14031–14087. https://doi.org/10.1021/acs.chemrev.1c00024.

(22) Shao, L.; Ma, J.; Prelesnik, J. L.; Zhou, Y.; Nguyen, M.; Zhao, M.; Jenekhe, S. A.; Kalinin, S. V.; Ferguson, A. L.; Pfaendtner, J.; Mundy, C. J.; De Yoreo, J. J.; Baneyx, F.; Chen, C.-L. Hierarchical Materials from High Information Content Macromolecular Building Blocks: Construction, Dynamic Interventions, and Prediction. *Chem Rev* **2022**, *122* (24), 17397–17478. https://doi.org/10.1021/acs.chemrev.2c00220.

(23) Chen, C.-L.; Qi, J.; Zuckermann, R. N.; DeYoreo, J. J. Engineered Biomimetic Polymers as Tunable Agents for Controlling $CaCO_3$ Mineralization. *J Am Chem Soc* **2011**, *133* (14), 5214–5217. https://doi.org/10.1021/ja200595f.

(24) Yang, W.; Yin, Q.; Chen, C.-L. Designing Sequence-Defined Peptoids for Biomimetic Control over Inorganic Crystallization. *Chemistry of Materials* **2021**, *33* (9), 3047–3065. https://doi.org/10.1021/acs.chemmater.1c00243.

(25) Ma, J.; Cai, B.; Zhang, S.; Jian, T.; De Yoreo, J. J.; Chen, C.-L.; Baneyx, F. Nanoparticle-Mediated Assembly of Peptoid Nanosheets Functionalized with Solid-Binding Proteins: Designing Heterostructures for Hierarchy. *Nano Lett* **2021**, *21* (4), 1636–1642. https://doi.org/10.1021/acs.nanolett.0c04285.

(26) Kim, J. H.; Kim, S. C.; Kline, M. A.; Grzincic, E. M.; Tresca, B. W.; Cardiel, J.; Karbaschi, M.; Dehigaspitiya, D. C.; Chen, Y.; Udumula, V.; Jian, T.; Murray, D. J.; Yun, L.; Connolly, M. D.; Liu, J.; Ren, G.; Chen, C.-L.; Kirshenbaum, K.; Abate, A. R.; Zuckermann, R. N. Discovery of Stable and Selective Antibody Mimetics from Combinatorial Libraries of Polyvalent, Loop-Functionalized Peptoid Nanosheets. *ACS Nano* **2020**, *14* (1), 185–195. https://doi.org/10.1021/acsnano.9b07498.

(27) Luo, Y.; Song, Y.; Wang, M.; Jian, T.; Ding, S.; Mu, P.; Liao, Z.; Shi, Q.; Cai, X.; Jin, H.; Du, D.; Dong, W.; Chen, C.; Lin, Y. Bioinspired Peptoid Nanotubes for Targeted Tumor Cell Imaging and Chemo-Photodynamic Therapy. *Small* **2019**, *15* (43). https://doi.org/10.1002/smll.201902485.

(28) Connolly, M. D.; Xuan, S.; Molchanova, N.; Zuckermann, R. N. Submonomer Synthesis of Sequence Defined Peptoids with Diverse Side-Chains; 2021; pp 241–270. https://doi.org/10.1016/bs.mie.2021.04.022.

(29) Zhao, M.; Zhang, S.; Zheng, R.; Alamdari, S.; Mundy, C. J.; Pfaendtner, J.; Pozzo, L. D.; Chen, C.-L.; De Yoreo, J. J.; Ferguson, A. L. Computational and Experimental Determination of the Properties, Structure, and Stability of Peptoid Nanosheets and Nanotubes. *Biomacromolecules* **2023**, *24* (6), 2618–2632. https://doi.org/10.1021/acs.biomac.3c00107.

(30) Zhao, M.; Lachowski, K. J.; Zhang, S.; Alamdari, S.; Sampath, J.; Mu, P.; Mundy, C. J.; Pfaendtner, J.; De Yoreo, J. J.; Chen, C.-L.; Pozzo, L. D.; Ferguson, A. L. Hierarchical Self-Assembly Pathways of Peptoid Helices and Sheets. *Biomacromolecules* **2022**, *23* (3), 992–1008. https://doi.org/10.1021/acs.biomac.1c01385.

(31) Fowler, S. A.; Blackwell, H. E. Structure–Function Relationships in Peptoids: Recent Advances toward Deciphering the Structural Requirements for Biological Function. *Org Biomol Chem* **2009**, *7* (8), 1508. https://doi.org/10.1039/b817980h.

(32) Wu, C. W.; Sanborn, T. J.; Zuckermann, R. N.; Barron, A. E. Peptoid Oligomers with α-Chiral, Aromatic Side Chains: Effects of Chain Length on Secondary Structure. *J Am Chem Soc* **2001**, *123* (13), 2958–2963. https://doi.org/10.1021/ja003153v.

(33) Armand, P.; Kirshenbaum, K.; Falicov, A.; Dunbrack, R. L.; Dill, K. A.; Zuckermann, R. N.; Cohen, F. E. Chiral N-Substituted Glycines Can Form Stable Helical Conformations. *Fold Des* **1997**, *2* (6), 369–375. https://doi.org/10.1016/S1359-0278(97)00051-5.

(34) Alamdari, S.; Pfaendtner, J. Origins of Conformational Heterogeneity in Peptoid Helices Formed by Chiral *N* -1-Phenylethyl Sidechains. *J Phys Chem B* **2023**, *127* (27), 6163–6170. https://doi.org/10.1021/acs.jpcb.3c02576.

(35) Alamdari, S.; Torkelson, K.; Wang, X.; Chen, C.-L.; Ferguson, A. L.; Pfaendtner, J. Thermodynamic Basis for the Stabilization of Helical Peptoids by Chiral Sidechains. *J Phys Chem B* **2023**, *127* (27), 6171–6183. https://doi.org/10.1021/acs.jpcb.3c01913.

(36) Jain, R. K.; Hall, C. K.; Santiso, E. E. Using Enhanced Sampling Simulations to Study the Conformational Space of Chiral Aromatic Peptoid Monomers. *J Chem Theory Comput* **2023**, *19* (24), 9457–9467. https://doi.org/10.1021/acs.jctc.3c00803.

(37) Shah, N. H.; Butterfoss, G. L.; Nguyen, K.; Yoo, B.; Bonneau, R.; Rabenstein, D. L.; Kirshenbaum, K. Oligo(N-Aryl Glycines): A New Twist on Structured Peptoids. *J Am Chem Soc* **2008**, *130* (49), 16622–16632. https://doi.org/10.1021/ja804580n.

(38) Roy, O.; Dumonteil, G.; Faure, S.; Jouffret, L.; Kriznik, A.; Taillefumier, C. Homogeneous and Robust Polyproline Type I Helices from Peptoids with Nonaromatic α-Chiral Side Chains. *J Am Chem Soc* **2017**, *139* (38), 13533–13540. https://doi.org/10.1021/jacs.7b07475.

(39) Stringer, J. R.; Crapster, J. A.; Guzei, I. A.; Blackwell, H. E. Extraordinarily Robust Polyproline Type I Peptoid Helices Generated via the Incorporation of α-Chiral Aromatic N-1-Naphthylethyl Side Chains. *J Am Chem Soc* **2011**, *133* (39), 15559–15567. https://doi.org/10.1021/ja204755p.

(40) Kröger, N.; Deutzmann, R.; Sumper, M. Polycationic Peptides from Diatom Biosilica That Direct Silica Nanosphere Formation. *Science (1979)* **1999**, *286* (5442), 1129–1132. https://doi.org/10.1126/science.286.5442.1129.

(41) Lechner, C.; Becker, C. Silaffins in Silica Biomineralization and Biomimetic Silica Precipitation. *Mar Drugs* **2015**, *13* (8), 5297–5333. https://doi.org/10.3390/md13085297.

(42) Choi, O.; Kim, B.-C.; An, J.-H.; Min, K.; Kim, Y. H.; Um, Y.; Oh, M.-K.; Sang, B.-I. A Biosensor Based on the Self-Entrapment of Glucose Oxidase within Biomimetic Silica Nanoparticles Induced by a Fusion Enzyme. *Enzyme Microb Technol* **2011**, *49* (5), 441–445. https://doi.org/10.1016/j.enzmictec.2011.07.005.

(43) Naik, R. R.; Tomczak, M. M.; Luckarift, H. R.; Spain, J. C.; Stone, M. O. Entrapment of Enzymes and Nanoparticles Using Biomimetically Synthesized Silica. *Chemical Communications* **2004**, No. 15, 1684. https://doi.org/10.1039/b404586f.

(44) Forsyth, C.; Yip, T. W. S.; Patwardhan, S. V. CO2 Sequestration by Enzyme Immobilized onto Bioinspired Silica. *Chem. Commun.* **2013**, *49* (31), 3191–3193. https://doi.org/10.1039/C2CC38225C.

(45) Jo, B. H.; Seo, J. H.; Yang, Y. J.; Baek, K.; Choi, Y. S.; Pack, S. P.; Oh, S. H.; Cha, H. J. Bioinspired Silica Nanocomposite with Autoencapsulated Carbonic Anhydrase as a Robust Biocatalyst for CO $_2$ Sequestration. *ACS Catal* **2014**, *4* (12), 4332–4340. https://doi.org/10.1021/cs5008409.

(46) Deng, Z.; Shi, Q.; Tan, J.; Hu, J.; Liu, S. Sequence-Defined Synthetic Polymers for New-Generation Functional Biomaterials. *ACS Mater Lett* **2021**, *3* (9), 1339–1356. https://doi.org/10.1021/acsmaterialslett.1c00358.

(47) Ganesh, S. D.; Saha, N.; Zandraa, O.; Zuckermann, R. N.; Sáha, P. Peptoids and Polypeptoids: Biomimetic and Bioinspired Materials for Biomedical Applications. *Polymer Bulletin* **2017**, *74* (8), 3455–3466. https://doi.org/10.1007/s00289-016-1902-1.

(48) Udugamasooriya, D. G.; Dineen, S. P.; Brekken, R. A.; Kodadek, T. A Peptoid "Antibody Surrogate" That Antagonizes VEGF Receptor 2 Activity. *J Am Chem Soc* **2008**, *130* (17), 5744–5752. https://doi.org/10.1021/ja711193x.

(49) Chongsiriwatana, N. P.; Patch, J. A.; Czyzewski, A. M.; Dohm, M. T.; Ivankin, A.; Gidalevitz, D.; Zuckermann, R. N.; Barron, A. E. Peptoids That Mimic the Structure, Function, and Mechanism of Helical Antimicrobial Peptides. *Proceedings of the National Academy of Sciences* **2008**, *105* (8), 2794–2799. https://doi.org/10.1073/pnas.0708254105.

(50) Statz, A. R.; Park, J. P.; Chongsiriwatana, N. P.; Barron, A. E.; Messersmith, P. B. Surface-Immobilised Antimicrobial Peptoids. *Biofouling* **2008**, *24* (6), 439–448. https://doi.org/10.1080/08927010802331829.

(51) Statz, A. R.; Meagher, R. J.; Barron, A. E.; Messersmith, P. B. New Peptidomimetic Polymers for Antifouling Surfaces. *J Am Chem Soc* **2005**, *127* (22), 7972–7973. https://doi.org/10.1021/ja0522534.

(52) Gangloff, N.; Ulbricht, J.; Lorson, T.; Schlaad, H.; Luxenhofer, R. Peptoids and Polypeptoids at the Frontier of Supra- and Macromolecular Engineering. *Chem Rev* **2016**, *116* (4), 1753–1802. https://doi.org/10.1021/acs.chemrev.5b00201.

(53) Kirshenbaum, K.; Barron, A. E.; Goldsmith, R. A.; Armand, P.; Bradley, E. K.; Truong, K. T. V.; Dill, K. A.; Cohen, F. E.; Zuckermann, R. N. Sequence-Specific Polypeptoids: A Diverse Family of Heteropolymers with Stable Secondary Structure. *Proceedings of the National Academy of Sciences* **1998**, *95* (8), 4303–4308. https://doi.org/10.1073/pnas.95.8.4303.

(54) Lee, B.-C.; Zuckermann, R. N.; Dill, K. A. Folding a Nonbiological Polymer into a Compact Multihelical Structure. *J Am Chem Soc* **2005**, *127* (31), 10999–11009. https://doi.org/10.1021/ja0514904.

(55) Robertson, E. J.; Battigelli, A.; Proulx, C.; Mannige, R. V.; Haxton, T. K.; Yun, L.; Whitelam, S.; Zuckermann, R. N. Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets. *Acc Chem Res* **2016**, *49* (3), 379–389. https://doi.org/10.1021/acs.accounts.5b00439.

(56) Yadav Schmid, S.; Ma, X.; Hammons, J. A.; Mergelsberg, S. T.; Harris, B. S.; Ferron, T.; Yang, W.; Zhou, W.; Zheng, R.; Zhang, S.; Legg, B. A.; Van Buuren, A.; Baer, M. D.; Chen, C.-L.; Tao, J.; De Yoreo, J. J. Influence of Peptoid Sequence on the Mechanisms and Kinetics of 2D Assembly. *ACS Nano* **2024**, *18* (4), 3497–3508. https://doi.org/10.1021/acsnano.3c10810.

(57) Eastwood, J. R. B.; Jiang, L.; Bonneau, R.; Kirshenbaum, K.; Renfrew, P. D. Evaluating the Conformations and Dynamics of Peptoid Macrocycles. *J Phys Chem B* **2022**, *126* (28), 5161–5174. https://doi.org/10.1021/acs.jpcb.2c01669.

(58) Lee, B.-C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N. Biomimetic Nanostructures: Creating a High-Affinity Zinc-Binding Site in a Folded Nonbiological Polymer. *J Am Chem Soc* **2008**, *130* (27), 8847–8855. https://doi.org/10.1021/ja802125x.

(59) Murnen, H. K.; Khokhlov, A. R.; Khalatur, P. G.; Segalman, R. A.; Zuckermann, R. N. Impact of Hydrophobic Sequence Patterning on the Coil-to-Globule Transition of Protein-like Polymers. *Macromolecules* **2012**, *45* (12), 5229–5236. https://doi.org/10.1021/ma300707t.

(60) DeStefano, A. J.; Mengel, S. D.; Bates, M. W.; Jiao, S.; Shell, M. S.; Han, S.; Segalman, R. A. Control over Conformational Landscapes of Polypeptoids by Monomer Sequence Patterning. *Macromolecules* **2024**, *57* (4), 1469–1477. https://doi.org/10.1021/acs.macromol.3c02338.

(61) Mannige, R. V.; Haxton, T. K.; Proulx, C.; Robertson, E. J.; Battigelli, A.; Butterfoss, G. L.; Zuckermann, R. N.; Whitelam, S. Peptoid Nanosheets Exhibit a New Secondary-Structure Motif. *Nature* **2015**, *526* (7573), 415–420. https://doi.org/10.1038/nature15363.

(62) Zhao, M.; Sampath, J.; Alamdari, S.; Shen, G.; Chen, C.-L.; Mundy, C. J.; Pfaendtner, J.; Ferguson, A. L. MARTINI-Compatible Coarse-Grained Model for the Mesoscale Simulation of Peptoids. *J Phys Chem B* **2020**, *124* (36), 7745–7764. https://doi.org/10.1021/acs.jpcb.0c04567.

(63) Swanson, H. W. A.; van Teijlingen, A.; Lau, K. H. A.; Tuttle, T. Martinoid: The Peptoid Martini Force Field. *Physical Chemistry Chemical Physics* **2024**, *26* (6), 4939–4953. https://doi.org/10.1039/D3CP05907C.

(64) Mirijanian, D. T.; Mannige, R. V.; Zuckermann, R. N.; Whitelam, S. Development and Use of an Atomistic CHARMM-Based Forcefield for Peptoid Simulation. *J Comput Chem* **2014**, *35* (5), 360–370. https://doi.org/10.1002/jcc.23478.

(65) Weiser, L. J.; Santiso, E. E. A CGenFF-based Force Field for Simulations of Peptoids with Both Cis and Trans Peptide Bonds. *J Comput Chem* **2019**, *40* (22), 1946–1956. https://doi.org/10.1002/jcc.25850.

(66) Roy Dennington; Todd Keith; John Millam. GaussView, Version 6.1.1. Semichem Inc.: Shawnee Mission, KS 2019.

(67) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

(68) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J Comput Chem* **2009**, NA-NA. https://doi.org/10.1002/jcc.21367.

(69) Huang, J.; MacKerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J Comput Chem* **2013**, *34* (25), 2135–2145. https://doi.org/10.1002/jcc.23354.

(70) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J Chem Inf Model* **2012**, *52* (12), 3155–3168. https://doi.org/10.1021/ci3003649.

(71) Torkelson, K.; Naser, N. Y.; Qi, X.; Li, Z.; Yang, W.; Pushpavanam, K.; Chen, C.-L.; Baneyx, F.; Pfaendtner, J. Rational Design of Novel Biomimetic Sequence-Defined Polymers for Mineralization Applications. *Chemistry of Materials* **2024**, *36* (2), 786–794. https://doi.org/10.1021/acs.chemmater.3c02216.

(72) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(73) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J Chem Theory Comput* **2015**, *11* (4), 1864–1874. https://doi.org/10.1021/ct5010406.

(74) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J Chem Phys* **1984**, *81* (8), 3684–3690. https://doi.org/10.1063/1.448118.

(75) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput Phys Commun* **2014**, *185* (2), 604–613. https://doi.org/10.1016/j.cpc.2013.09.018.

(76) Pietrucci, F.; Laio, A. A Collective Variable for the Efficient Exploration of Protein Beta-Sheet Structures: Application to SH3 and GB1. *J Chem Theory Comput* **2009**, *5* (9), 2197–2201. https://doi.org/10.1021/ct900202f.

(77) The PLUMED Consortium. Promoting Transparency and Reproducibility in Enhanced Molecular Simulations. *Nat Methods* **2019**, *16* (8), 670–673. https://doi.org/10.1038/s41592-019-0506-8.

(78) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angewandte Chemie International Edition* **1999**, *38* (1–2), 236–240. https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M.

(79) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J Mol Graph* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(80) Rodríguez, F.; Glawe, D. D.; Naik, R. R.; Hallinan, K. P.; Stone, M. O. Study of the Chemical and Physical Influences upon in Vitro Peptide-Mediated Silica Formation. *Biomacromolecules* **2004**, *5* (2), 261–265. https://doi.org/10.1021/bm034232c.

(81) Knecht, M. R.; Wright, D. W. Functional Analysis of the Biomimetic Silica Precipitating Activity of the R5 Peptide from Cylindrotheca Fusiformis. *Chemical Communications* **2003**, No. 24, 3038. https://doi.org/10.1039/b309074d.

(82) Senior, L.; Crump, M. P.; Williams, C.; Booth, P. J.; Mann, S.; Perriman, A. W.; Curnow, P. Structure and Function of the Silicifying Peptide R5. *J Mater Chem B* **2015**, *3* (13), 2607–2614. https://doi.org/10.1039/C4TB01679C.

(83) Lutz, H.; Jaeger, V.; Berger, R.; Bonn, M.; Pfaendtner, J.; Weidner, T. Biomimetic Growth of Ultrathin Silica Sheets Using Artificial Amphiphilic Peptides. *Adv Mater Interfaces* **2015**, *2* (17), 1500282. https://doi.org/10.1002/admi.201500282.

(84) Yang, W.; Hellner, B.; Baneyx, F. Self-Immobilization of Car9 Fusion Proteins within High Surface Area Silica Sol–Gels and Dynamic Control of Protein Release. *Bioconjug Chem* **2016**, *27* (10), 2450–2459. https://doi.org/10.1021/acs.bioconjchem.6b00406.

(85) Buckle, E. L.; Sampath, J.; Michael, N.; Whedon, S. D.; Leonen, C. J. A.; Pfaendtner, J.; Drobny, G. P.; Chatterjee, C. Trimethylation of the R5 Silica-Precipitating Peptide Increases Silica Particle Size by Redirecting Orthosilicate Binding. *ChemBioChem* **2020**, *21* (22), 3208–3211. https://doi.org/10.1002/cbic.202000264.

(86) Kröger, N.; Deutzmann, R.; Sumper, M. Silica-Precipitating Peptides from Diatoms. *Journal of Biological Chemistry* **2001**, *276* (28), 26066–26070. https://doi.org/10.1074/jbc.M102093200.

(87) Kröger, N.; Lorenz, S.; Brunner, E.; Sumper, M. Self-Assembly of Highly Phosphorylated Silaffins and Their Function in Biosilica Morphogenesis. *Science (1979)* **2002**, *298* (5593), 584–586. https://doi.org/10.1126/science.1076221.

(88) Gascoigne, L.; Magana, J. R.; Atkins, D. L.; Sproncken, C. C. M.; Gumi-Audenis, B.; Schoenmakers, S. M. C.; Wakeham, D.; Wanless, E. J.; Voets, I. K. Fractal-like R5 Assembly Promote the Condensation of Silicic Acid into Silica Particles. *J Colloid Interface Sci* **2021**, *598*, 206–212. https://doi.org/10.1016/j.jcis.2021.04.030.

(89) Wallace, A. K.; Chanut, N.; Voigt, C. A. Silica Nanostructures Produced Using Diatom Peptides with Designed Post-Translational Modifications. *Adv Funct Mater* **2020**, *30* (30), 2000849. https://doi.org/10.1002/adfm.202000849.

(90) Sprenger, K. G.; Prakash, A.; Drobny, G.; Pfaendtner, J. Investigating the Role of Phosphorylation in the Binding of Silaffin Peptide R5 to Silica with Molecular Dynamics Simulations. *Langmuir* **2018**, *34* (3), 1199–1207. https://doi.org/10.1021/acs.langmuir.7b02868.

(91) Lechner, C. C.; Becker, C. F. W. Exploring the Effect of Native and Artificial Peptide Modifications on Silaffin Induced Silica Precipitation. *Chem Sci* **2012**, *3* (12), 3500. https://doi.org/10.1039/c2sc20687k.

(92) Lutz, H.; Jaeger, V.; Schmüser, L.; Bonn, M.; Pfaendtner, J.; Weidner, T. The Structure of the Diatom Silaffin Peptide R5 within Freestanding Two-Dimensional Biosilica Sheets. *Angewandte Chemie International Edition* **2017**, *56* (28), 8277–8280. https://doi.org/10.1002/anie.201702707.

(93) Montagna, M.; Brückner, S. I.; Dianat, A.; Gutierrez, R.; Daus, F.; Geyer, A.; Brunner, E.; Cuniberti, G. Interactions of Long-Chain Polyamines with Silica Studied by Molecular Dynamics Simulations and Solid-State NMR Spectroscopy. *Langmuir* **2020**, *36* (39), 11600–11609. https://doi.org/10.1021/acs.langmuir.0c02157.

(94) Yang, W.; Zhou, Y.; Jin, B.; Qi, X.; Cai, B.; Yin, Q.; Pfaendtner, J.; De Yoreo, J. J.; Chen, C.-L. Designing Sequence-Defined Peptoids for Fibrillar Self-Assembly and Silicification. *J Colloid Interface Sci* **2023**, *634*, 450–459. https://doi.org/10.1016/j.jcis.2022.11.136.

(95) Hamm, L. M.; Giuffre, A. J.; Han, N.; Tao, J.; Wang, D.; De Yoreo, J. J.; Dove, P. M. Reconciling Disparate Views of Template-Directed Nucleation through Measurement of Calcite Nucleation Kinetics and Binding Energies. *Proceedings of the National Academy of Sciences* **2014**, *111* (4), 1304–1309. https://doi.org/10.1073/pnas.1312369111.

(96) Alamdari, S.; Pfaendtner, J. Impact of Glutamate Carboxylation in the Adsorption of the α-1 Domain of Osteocalcin to Hydroxyapatite and Titania. *Mol Syst Des Eng* **2020**, *5* (3), 620–631. https://doi.org/10.1039/C9ME00158A.

(97) Sampath, J.; Pfaendtner, J. Amphiphilic Peptide Binding on Crystalline vs. Amorphous Silica from Molecular Dynamics Simulations. *Mol Phys* **2019**, *117* (23–24), 3642–3650. https://doi.org/10.1080/00268976.2019.1657192.

(98) Lechner, C. C.; Becker, C. F. W. A Sequence-Function Analysis of the Silica Precipitating Silaffin R5 Peptide. *Journal of Peptide Science* **2014**, *20* (2), 152–158. https://doi.org/10.1002/psc.2577.

(99) Strobl, J.; Kozak, F.; Kamalov, M.; Reichinger, D.; Kurzbach, D.; Becker, C. F. Understanding Self-Assembly of Silica-Precipitating Peptides to Control Silica Particle Morphology. *Advanced Materials* **2023**, *35* (11). https://doi.org/10.1002/adma.202207586.

(100) Hare, S. R.; Pfaendtner, J. Elucidating the Role of Catalytic Amino Acid Residues in the Peptide-Mediated Silica Oligomerization Reaction Mechanism. *Physical Chemistry Chemical Physics* **2022**, *24* (6), 3664–3674. https://doi.org/10.1039/D1CP04542C.

(101) Zexer, N.; Kumar, S.; Elbaum, R. Silica Deposition in Plants: Scaffolding the Mineralization. *Ann Bot* **2023**. https://doi.org/10.1093/aob/mcad056.

(102) Mao, C. M.; Sampath, J.; Pfaendtner, J. Molecular Driving Forces in the Self-Association of Silaffin Peptide R5 from MD Simulations. *ChemBioChem* **2024**. https://doi.org/10.1002/cbic.202300788.

(103) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J Mol Graph Model* **2006**, *25* (2), 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005.

(104) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J Comput Chem* **2004**, *25* (9), 1157–1174. https://doi.org/10.1002/jcc.20035.

(105) Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J Chem Theory Comput* **2012**, *8* (11), 4405–4412. https://doi.org/10.1021/ct300613v.

(106) MacKerell, A. D.; Banavali, N.; Foloppe, N. Development and Current Status of the CHARMM Force Field for Nucleic Acids. *Biopolymers* **2000**, *56* (4), 257–265. https://doi.org/10.1002/1097-0282(2000)56:4<257::AID-BIP10029>3.0.CO;2-W.

(107) Heinz, H.; Lin, T.-J.; Kishore Mishra, R.; Emami, F. S. Thermodynamically Consistent Force Fields for the Assembly of Inorganic, Organic, and Biological Nanostructures: The INTERFACE Force Field. *Langmuir* **2013**, *29* (6), 1754–1765. https://doi.org/10.1021/la3038846.

(108) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J Comput Chem* **1997**, *18* (12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

(109) Pfaendtner, J.; Bonomi, M. Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics. *J Chem Theory Comput* **2015**, *11* (11), 5062–5067. https://doi.org/10.1021/acs.jctc.5b00846.

(110) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics. *J Phys Chem B* **2006**, *110* (8), 3533–3539. https://doi.org/10.1021/jp054359r.

(111) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with Adaptive Gaussians. *J Chem Theory Comput* **2012**, *8* (7), 2247–2254. https://doi.org/10.1021/ct3002464.

(112) Ozaki, M.; Sakashita, S.; Hamada, Y.; Usui, K. Peptides for Silica Precipitation: Amino Acid Sequences for Directing Mineralization. *Protein Pept Lett* **2018**, *25* (1), 15–24. https://doi.org/10.2174/0929866525666171214111007.

(113) Alshammari, B. H.; Lashin, M. M. A.; Mahmood, M. A.; Al-Mubaddel, F. S.; Ilyas, N.; Rahman, N.; Sohail, M.; Khan, A.; Abdullaev, S. S.; Khan, R. Organic and Inorganic Nanomaterials: Fabrication, Properties and Applications. *RSC Adv* **2023**, *13* (20), 13735–13785. https://doi.org/10.1039/D3RA01421E.

(114) Brett, M. W.; Gordon, C. K.; Hardy, J.; Davis, N. J. L. K. The Rise and Future of Discrete Organic–Inorganic Hybrid Nanomaterials. *ACS Physical Chemistry Au* **2022**, *2* (5), 364–387. https://doi.org/10.1021/acsphyschemau.2c00018.

(115) Khalid, K.; Tan, X.; Mohd Zaid, H. F.; Tao, Y.; Lye Chew, C.; Chu, D.-T.; Lam, M. K.; Ho, Y.-C.; Lim, J. W.; Chin Wei, L. Advanced in Developmental Organic and Inorganic Nanomaterial: A Review. *Bioengineered* **2020**, *11* (1), 328–355. https://doi.org/10.1080/21655979.2020.1736240.

(116) Lachowski, K. J.; Vaddi, K.; Naser, N. Y.; Baneyx, F.; Pozzo, L. D. Multivariate Analysis of Peptide-Driven Nucleation and Growth of Au Nanoparticles. *Digital Discovery* **2022**, *1* (4), 427–439. https://doi.org/10.1039/D2DD00017B.

(117) Qi, X.; Zhao, Y.; Lachowski, K.; Boese, J.; Cai, Y.; Dollar, O.; Hellner, B.; Pozzo, L.; Pfaendtner, J.; Chun, J.; Baneyx, F.; Mundy, C. J. Predictive Theoretical Framework for Dynamic Control of Bioinspired Hybrid Nanoparticle Self-Assembly. *ACS Nano* **2022**, *16* (2), 1919–1928. https://doi.org/10.1021/acsnano.1c04923.

(118) Yang, W.; Cai, B.; Lachowski, K. J.; Yin, Q.; De Yoreo, J. J.; Pozzo, L. D.; Chen, C.-L. Insights into the Biomimetic Synthesis of 2D ZnO Nanomaterials through Peptoid Engineering. *J Phys Chem Lett* **2023**, *14* (43), 9732–9739. https://doi.org/10.1021/acs.jpclett.3c01882.

(119) Lachowski, K. J.; Chiang, H. T.; Torkelson, K.; Zhou, W.; Zhang, S.; Pfaendtner, J.; Pozzo, L. D. Anisotropic Gold Nanomaterial Synthesis Using Peptide Facet Specificity and Timed Intervention. *Langmuir* **2023**, *39* (45), 15878–15888. https://doi.org/10.1021/acs.langmuir.3c01577.

(120) Larson, H.; Cossairt, B. M. Indium–Poly(Carboxylic Acid) Ligand Interactions Modify InP Quantum Dot Nucleation and Growth. *Chemistry of Materials* **2023**, *35* (15), 6152–6160. https://doi.org/10.1021/acs.chemmater.3c01309.

(121) Gary, D. C.; Terban, M. W.; Billinge, S. J. L.; Cossairt, B. M. Two-Step Nucleation and Growth of InP Quantum Dots via Magic-Sized Cluster Intermediates. *Chemistry of Materials* **2015**, *27* (4), 1432–1441. https://doi.org/10.1021/acs.chemmater.5b00286.

(122) Ripberger, H. H.; Harvey, S. M.; Cossairt, B. M. Editorial: The Molecular Underpinnings of Nanoscale Semiconductor Synthesis. *Frontiers in Nanotechnology* **2023**, *5*. https://doi.org/10.3389/fnano.2023.1229232.

(123) Dou, F. Y.; Harvey, S. M.; Mason, K. G.; Homer, M. K.; Gamelin, D. R.; Cossairt, B. M. Effect of a Redox-Mediating Ligand Shell on Photocatalysis by CdS Quantum Dots. *J Chem Phys* **2023**, *158* (18). https://doi.org/10.1063/5.0144896.

(124) Park, N.; Beck, R. A.; Hoang, K. K.; Ladd, D. M.; Abramson, J. E.; Rivera-Maldonado, R. A.; Nguyen, H. A.; Monahan, M.; Seidler, G. T.; Toney, M. F.; Li, X.; Cossairt, B. M. Colloidal, Room-Temperature Growth of Metal Oxide Shells on InP Quantum Dots. *Inorg Chem* **2023**, *62* (17), 6674–6687. https://doi.org/10.1021/acs.inorgchem.3c00161.

(125) Simancas, R.; Dari, D.; Velamazán, N.; Navarro, M. T.; Cantín, A.; Jordá, J. L.; Sastre, G.; Corma, A.; Rey, F. Modular Organic Structure-Directing Agents for the Synthesis of Zeolites. *Science (1979)* **2010**, *330* (6008), 1219–1222. https://doi.org/10.1126/science.1196240.

(126) Wang, Z.; Chu, W.; Zhao, Z.; Liu, Z.; Chen, H.; Xiao, D.; Gong, K.; Li, F.; Li, X.; Hou, G. The Role of Organic and Inorganic Structure-Directing Agents in Selective Al Substitution of Zeolite. *J Phys Chem Lett* **2021**, *12* (38), 9398–9406. https://doi.org/10.1021/acs.jpclett.1c01448.

(127) Bernini, M. C.; Snejko, N.; Gutierrez-Puebla, E.; Brusau, E. V.; Narda, G. E.; Monge, M. Á. Structure-Directing and Template Roles of Aromatic Molecules in the Self-Assembly

Formation Process of 3D Holmium–Succinate MOFs. *Inorg Chem* **2011**, *50* (13), 5958–5968. https://doi.org/10.1021/ic102472u.

(128) Sinnwell, M. A.; Miller, Q. R. S.; Palys, L.; Barpaga, D.; Liu, L.; Bowden, M. E.; Han, Y.; Ghose, S.; Sushko, M. L.; Schaef, H. T.; Xu, W.; Nyman, M.; Thallapally, P. K. Molecular Intermediate in the Directed Formation of a Zeolitic Metal–Organic Framework. *J Am Chem Soc* **2020**, *142* (41), 17598–17606. https://doi.org/10.1021/jacs.0c07862.

(129) Huang, Q.; Yang, Y.; Qian, J. Structure-Directed Growth and Morphology of Multifunctional Metal-Organic Frameworks. *Coord Chem Rev* **2023**, *484*, 215101. https://doi.org/10.1016/j.ccr.2023.215101.

(130) Nudelman, F.; Sommerdijk, N. A. J. M. Biomineralization as an Inspiration for Materials Chemistry. *Angewandte Chemie International Edition* **2012**, *51* (27), 6582–6596. https://doi.org/10.1002/anie.201106715.

(131) Cha, J. N.; Shimizu, K.; Zhou, Y.; Christiansen, S. C.; Chmelka, B. F.; Stucky, G. D.; Morse, D. E. Silicatein Filaments and Subunits from a Marine Sponge Direct the Polymerization of Silica and Silicones *in Vitro*. *Proceedings of the National Academy of Sciences* **1999**, *96* (2), 361–365. https://doi.org/10.1073/pnas.96.2.361.

(132) 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* **1998**, *95* (11), 6234–6238. https://doi.org/10.1073/pnas.95.11.6234.

(133) Crookes-Goodson, W. J.; Slocik, J. M.; Naik, R. R. Bio-Directed Synthesis and Assembly of Nanomaterials. *Chem Soc Rev* **2008**, *37* (11), 2403. https://doi.org/10.1039/b702825n.

(134) Aggeli, A.; Bell, M.; Carrick, L. M.; Fishwick, C. W. G.; Harding, R.; Mawer, P. J.; Radford, S. E.; Strong, A. E.; Boden, N. PH as a Trigger of Peptide β-Sheet Self-Assembly and Reversible Switching between Nematic and Isotropic Phases. *J Am Chem Soc* **2003**, *125* (32), 9619–9628. https://doi.org/10.1021/ja021047i.

(135) Wang, S.; Ge, X.; Xue, J.; Fan, H.; Mu, L.; Li, Y.; Xu, H.; Lu, J. R. Mechanistic Processes Underlying Biomimetic Synthesis of Silica Nanotubes from Self-Assembled Ultrashort Peptide Templates. *Chemistry of Materials* **2011**, *23* (9), 2466–2474. https://doi.org/10.1021/cm2003885.

(136) Ramanathan, R.; Campbell, J. L.; Soni, S. K.; Bhargava, S. K.; Bansal, V. Cationic Amino Acids Specific Biomimetic Silicification in Ionic Liquid: A Quest to Understand the

Formation of 3-D Structures in Diatoms. *PLoS One* **2011**, *6* (3), e17707. https://doi.org/10.1371/journal.pone.0017707.

(137) Tomizaki, K.; Ahn, S.-A.; Imai, T. Synthesis of Silica Nanofibers Templated by Self-Assembled Peptide Nanostructures. *Transactions of the Materials Research Society of Japan* **2012**, *37* (4), 541–546. https://doi.org/10.14723/tmrsj.37.541.

(138) Meegan, J. E.; Aggeli, A.; Boden, N.; Brydson, R.; Brown, A. P.; Carrick, L.; Brough, A. R.; Hussain, A.; Ansell, R. J. Designed Self-Assembled Β-Sheet Peptide Fibrils as Templates for Silica Nanotubes. *Adv Funct Mater* **2004**, *14* (1), 31–37. https://doi.org/10.1002/adfm.200304477.

(139) Acar, H.; Garifullin, R.; Guler, M. O. Self-Assembled Template-Directed Synthesis of One-Dimensional Silica and Titania Nanostructures. *Langmuir* **2011**, *27* (3), 1079–1084. https://doi.org/10.1021/la104518g.

(140) Wang, Q.; Yu, J.; Zhang, X.; Liu, D.; Zheng, J.; Pan, Y.; Lin, Y. Controlled Biosilification Using Self-Assembled Short Peptides A6K and V6K. *RSC Adv* **2013**, *3* (8), 2784. https://doi.org/10.1039/c2ra22099g.

(141) Han, W.; MacEwan, S. R.; Chilkoti, A.; López, G. P. Bio-Inspired Synthesis of Hybrid Silica Nanoparticles Templated from Elastin-like Polypeptide Micelles. *Nanoscale* **2015**, *7* (28), 12038–12044. https://doi.org/10.1039/C5NR01407G.

(142) Wong Po Foo, C.; Patwardhan, S. V.; Belton, D. J.; Kitchel, B.; Anastasiades, D.; Huang, J.; Naik, R. R.; Perry, C. C.; Kaplan, D. L. Novel Nanocomposites from Spider Silk–Silica Fusion (Chimeric) Proteins. *Proceedings of the National Academy of Sciences* **2006**, *103* (25), 9428–9433. https://doi.org/10.1073/pnas.0601096103.

(143) 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* **2010**, *22* (20), 5780–5785. https://doi.org/10.1021/cm101940u.

(144) Canabady-Rochelle, L. L. S.; Belton, D. J.; Deschaume, O.; Currie, H. A.; Kaplan, D. L.; Perry, C. C. Bioinspired Silicification of Silica-Binding Peptide-Silk Protein Chimeras: Comparison of Chemically and Genetically Produced Proteins. *Biomacromolecules* **2012**, *13* (3), 683–690. https://doi.org/10.1021/bm201555c.

(145) Zane, A. C.; Michelet, C.; Roehrich, A.; Emani, P. S.; Drobny, G. P. Silica Morphogenesis by Lysine-Leucine Peptides with Hydrophobic Periodicity. *Langmuir* **2014**, *30* (24), 7152–7161. https://doi.org/10.1021/la501444t.

(146) Patwardhan, S. V.; Emami, F. S.; Berry, R. J.; Jones, S. E.; Naik, Rajesh. R.; Deschaume, O.; Heinz, H.; Perry, C. C. Chemistry of Aqueous Silica Nanoparticle Surfaces and the Mechanism of Selective Peptide Adsorption. *J Am Chem Soc* **2012**, *134* (14), 6244–6256. https://doi.org/10.1021/ja211307u.

(147) Naik, R. R.; Brott, L. L.; Clarson, S. J.; Stone, M. O. Silica-Precipitating Peptides Isolated from a Combinatorial Phage Display Peptide Library. *J Nanosci Nanotechnol* **2002**, *2* (1), 95–100. https://doi.org/10.1166/jnn.2002.074.

(148) Baio, J. E.; Zane, A.; Jaeger, V.; Roehrich, A. M.; Lutz, H.; Pfaendtner, J.; Drobny, G. P.; Weidner, T. Diatom Mimics: Directing the Formation of Biosilica Nanoparticles by Controlled Folding of Lysine-Leucine Peptides. *J Am Chem Soc* **2014**, *136* (43), 15134–15137. https://doi.org/10.1021/ja5078238.

(149) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; Izadi, S.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kurtzman, T.; Lee, T.-S.; Li, P.; Liu, J.; Luchko, T.; Luo, R.; Manathunga, M.; Machado, M. R.; Nguyen, H. M.; O'Hearn, K. A.; Onufriev, A. V.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Risheh, A.; Schott-Verdugo, S.; Shajan, A.; Swails, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Zhang, S.; Zhao, S.; Zhu, Q.; Cheatham, T. E.; Roe, D. R.; Roitberg, A.; Simmerling, C.; York, D. M.; Nagan, M. C.; Merz, K. M. AmberTools. *J Chem Inf Model* **2023**, *63* (20), 6183–6191. https://doi.org/10.1021/acs.jcim.3c01153.

(150) Qi, X.; Pfaendtner, J. High-Throughput Computational Screening of Solid-Binding Peptides. *J Chem Theory Comput* **2024**, *20* (7), 2959–2968. https://doi.org/10.1021/acs.jctc.3c01286.

(151) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J Comput Chem* **2009**, *30* (13), 2157–2164. https://doi.org/10.1002/jcc.21224.