Langmuir
Article
pubs.acs.org/Langmuir

Probing the Molecular Mechanisms of Quartz-Binding Peptides

Ersin Emre Oren,[†,‡,▽] Rebecca Notman,[§,▽] Il Won Kim,[⊥] John Spencer Evans,[⊥] Tiffany R. Walsh,[§,¶] Ram Samudrala,[‡] Candan Tamerler,[†,#] and Mehmet Sarikaya*[,†]

[†]*Department of Materials Science and Engineering, University of Washington, Seattle, Washington,* [‡]*Department of Microbiology, University of Washington, Seattle, Washington,* [§]*Department of Chemistry, University of Warwick, Coventry, CV4 7AL, U.K.,* [⊥]*Laboratory for Chemical Physics, New York University, New York, New York, 10010,* [¶]*Centre for Scientific Computing, University of Warwick, Coventry, CV4 7AL, U.K., and* [#]*Molecular Biology and Genetics, Istanbul Technical University, Maslak, Istanbul, 80626, Turkey.* [▽] *Denotes joint first authors.*

Received January 5, 2010. Revised Manuscript Received May 1, 2010

Understanding the mechanisms of biomineralization and the realization of biology-inspired inorganic materials formation largely depends on our ability to manipulate peptide/solid interfacial interactions. Material interfaces and biointerfaces are critical sites for bioinorganic synthesis, surface diffusion, and molecular recognition. Recently adapted biocombinatorial techniques permit the isolation of peptides recognizing inorganic solids that are used as molecular building blocks, for example, as synthesizers, linkers, and assemblers. Despite their ubiquitous utility in nanotechnology, biotechnology, and medicine, the fundamental mechanisms of molecular recognition of engineered peptides binding to inorganic surfaces remain largely unknown. To explore propensity rules connecting sequence, structure, and function that play key roles in peptide/solid interactions, we combine two different approaches: a statistical analysis that searches for highly enriched motifs among *de novo* designed peptides, and, atomistic simulations of three experimentally validated peptides. The two strong and one weak quartz-binding peptides were chosen for the simulations at the quartz (100) surface under aqueous conditions. Solution-based peptide structures were analyzed by circular dichroism measurements. Small and hydrophobic residues, such as Pro, play a key role at the interface by making close contact with the solid and hindering formation of intrapeptide hydrogen bonds. The high binding affinity of a peptide may be driven by a combination of favorable enthalpic and entropic effects, that is, a strong binder may possess a large number of possible binding configurations, many of which having relatively high binding energies. The results signify the role of the local molecular environment among the critical residues that participate in solid binding. The work herein describes molecular conformations inherent in material-specific peptides and provides fundamental insight into the atomistic understanding of peptide/solid interfaces.

1. Introduction

Natural hierarchical architectures of biological hard tissues are highly functional and form through molecular recognition of solids by proteins that mediate biomineralization.[1−3] Similarly, the development of advanced bioinorganic materials, therapeutic devices, and nanoparticles depends mostly on the biocompatibility of solid surfaces, again, at the biomolecule/material interface.[4−6] As a consequence, there has been a growing interest recently in the utilization of biocombinatorial peptide libraries, such as phage, cell surface, and yeast display,[7−9] to identify peptide sequences with desired material binding specificity[10−12] and the use of these peptides for synthesis, assembly, and formation of controlled materials structures for practical implementation in technology and medicine.[13−16] Although the solid-binding peptide sequence database has grown enormously, the field is still in a taxonomic phase. In particular, despite their versatility, due mostly to structural and chemical diversity of the amino acids, the ubiquitous utility of engineered peptides is based largely only on empirical understanding of solid-binding characteristics.[10−12,14−16] There is a growing need, however, for an atomistic understanding of the mechanism(s) by which explicit peptide sequences interface with specific solid surfaces for building robust foundations for peptide-based hybrid molecular technologies of the future.[17−19]

To understand molecular recognition of a solid by a peptide, we employ a number of complementary approaches combining bioinformatics, molecular simulations, and biophysical experimental analyses. Our goal is to explore a set of propensity rules that will connect amino acid content, sequence, and molecular structure with peptide/material interaction and affinity. We build

*To whom correspondence should be addressed. E-mail: sarikaya@u.washington.edu. Tel.: +1 206 543 0724. Fax: +1 206 543 6381. Address: Materials Science and Engineering, University of Washington, 327 Roberts Hall, Box 352120, Seattle, WA, 98195.

(1) Mann, S. *Nature* **1988**, *332*, 119–124.
(2) Weiner, S.; Addadi, L. *J. Mater. Chem.* **1997**, *7*, 689–702.
(3) Sarikaya, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14183–14185.
(4) Dorozhkin, S. V.; Epple, M. *Angew. Chem., Int. Ed.* **2002**, *41*, 3130–3146.
(5) Horbett, T. A., Brash, J. L., Eds. In *Proteins at Interfaces II: Fundamentals and Applications*, American Chemical Society: Washington, DC, 1995; Vol *602*.
(6) Loo, C.; Lin, A.; Hirsch, L.; Lee, M. H.; Barton, J.; Halas, N.; West, J.; Drezek, R. *Technol. Cancer Res. Treat.* **2004**, *3*, 33–40.
(7) Smith, G. P. *Science* **1985**, *228*, 1315–1317.
(8) Hoess, R. H. *Chem. Rev.* **2001**, *101*, 3205–3218.
(9) Chao, G.; Lau, W. L.; Hackel, B. J.; Sazinsky, S. L.; Lippow, S. M.; Wittrup, K. D. *Nat. Protoc.* **2006**, *1*, 755–768.
(10) Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. *Nature* **2000**, *405*, 665–668.
(11) Sarikaya, M.; Tamerler, C.; Schulten, K.; Jen, A.; Baneyx, F. *Nat. Mater.* **2003**, *2*, 577–585.
(12) Naik, R. R.; Brott, L. L.; Carlson, S. J.; Stone, M. O. *J. Nanosci. Nanotechnol.* **2002**, *2*, 95–100.
(13) Niemeyer, C. M. *Angew. Chem., Int. Ed.* **2001**, *40*, 4128–4158.
(14) Mao, C. B.; Flynn, C. E.; Hayhurst, A.; Sweeney, R.; Qi, J. F.; Georgiou, G.; Iverson, B.; Belcher, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6946–6951.
(15) Sarikaya, M.; Tamerler, C.; Schwartz, D. T.; Baneyx, F. *Annu. Rev. Mater. Res.* **2004**, *34*, 373–408.
(16) Sano, K. I.; Sasaki, H.; Shiba, K. *Langmuir* **2005**, *21*, 3090–3095.
(17) Seeman, N.; Belcher, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6451–6455.
(18) Berry, C. C.; Curtis, A. S. G. *J. Phys. D: Appl. Phys.* **2003**, *36*, R198–R206.
(19) Tamerler, C.; Sarikaya, M. *Nano Focus* **2009**, *3*, 1606–1615.

our study on several known recent findings, which includes the following: (a) In general, for a given set of solid-binding peptides, different sequences have different binding properties;[10−12,14−16] (b) in a given peptide, the order of the amino acids, that is, the sequence, is more significant than the amino acid content;[20] and (c) molecular conformations and architectures have a major effect on peptide binding to solid surfaces.[21] Despite recent advances, how peptides atomistically interact with inorganic surfaces and the rules governing these processes are still not understood. Here we provide a comprehensive interrogation of sequence, structure, and surface binding properties of solid-binding peptides.

In a previous study,[20] using the sequence information present within a pool of biocombinatorially selected quartz binding dodecapeptides, we first *de novo* designed peptides with predictable binding affinities to quartz. The binding affinities of the peptides were verified experimentally.[20] In this report, we investigate the molecular mechanisms that underpin the correlation between peptide sequence and its binding affinity. With the confidence we gained in the knowledge-based design approach,[20] here we generated a pool of 10^9 peptides and identified the top 100000 sequences having the highest similarity scores. We searched for residue motifs that are enriched within these designed high-affinity quartz-binding sequences as potentially responsible domains for binding. Independently, we also performed atomistic simulations to reveal any basis for molecular interactions in solution and at the aqueous peptide−quartz interface. For the atomistic simulation study, we chose three of the knowledge-base designed quartz-binding peptides, two with the highest (S1 and S2, i.e., strong-1 and strong-2) and one with the lowest (W1, i.e., weak-1) similarity scores. We found, interestingly, that the two different computational approaches, that is, atomistic simulations based on molecular dynamics and motif enrichment factor based on similarity analysis, provided molecular features consistent with each other that are likely to facilitate surface interactions. Furthermore, we also complemented our simulated solution structure data with circular dichroism (CD) spectroscopy observations, with which the modeling predictions concurred. Our findings indicate the importance of such an approach in eliciting general rules that govern peptide-inorganic affinity. The comprehensive approach established here could open up new vistas for probing, quantifying, and ultimately exploiting newly designed peptides that possess controllable properties at inorganic interfaces, leading to practical synthesis, assembly, and fabrication routes of functional nanostructured materials for use in technology and medicine.[4−6,13,18]

2. Materials and Methods

2.1. Calculation of the Enrichment of Amino Acid Motifs. To find the enriched amino acid motifs in the de novo designed peptides, we first calculated the observed frequency of each motif (ρ^{O}_{motif}) by counting them and then normalizing the counts by the total number of motifs present in the entire set of peptides. Once the observed motif frequencies are obtained, the probability of having an n amino acid long motif within a given N amino acid long peptide (P^{O}_{motif}) can be calculated using the following equation:

$$P^{O}_{motif} = 1 - (1 - \rho^{O}_{motif})^{(N-n+1)} \quad (1)$$

Later to see whether these domains are enriched over the combinatorial selection or not, we compared them with the expected frequency P^{E}_{motif} of such domains within a given N amino acid long peptide. The observed amino acid distribution of the Ph.D.-12 library, used for the combinatorial selection of the quartz binding peptides, is given in the Ph.D.-12 phage display library kit.[22] The likelihood of having an n amino acid long motif by chance can be calculated using the preceding empirical data as follows:

$$P^{E}_{motif} = 1 - (1 - \prod_{i=1}^{n} P_i^{aa})^{(N-n+1)} \quad (2)$$

where, P_i^{aa} is the probability of having the amino acid aa, which is in the ith position of the motif in the Ph.D.-12 library.[22] Finally, the *motif enrichment factors*, EF_{motif}, can be calculated by comparing the already calculated observed and expected frequencies as follows.

$$EF_{motif} = \frac{P^{O}_{motif}}{P^{E}_{motif}} \quad (3)$$

The *motif enrichment factors* indicate how much the motif is enriched within the designed peptides, and the higher valued motifs are assumed to be the one responsible for binding.

2.2. CD Experiments. Lyophilized S1, S2, and W1 peptides were individually dissolved in distilled deionized water to produce stock solutions. Each stock solution was then diluted to 30, 20, 15, 12, 9, and 6 μM for CD spectrometry measurements in 100 μM Tris·HCl (pH 7.5).[23] On the basis of the concentration variation studies, the optimal concentration was chosen as 15 μM. All CD spectra were obtained at 25 °C with an AVIV 60 CD spectrometer, running 60DS software version 4.1t. The CD spectrometer was previously calibrated with *d*-10-camphorsulfonic acid. Wavelength scans were conducted from 185 to 260 nm with appropriate buffer background subtraction.[23] For each spectra, three scans were averaged using 1 nm bandwidth and a scanning rate of 0.5 nm/sec. Mean residue ellipticity [θM] is expressed in deg·cm^2·dmol^{-1}.[23]

2.3. Molecular Dynamics Simulations. *Peptide Solution Simulations*. For the peptide-solution systems we used replica-exchange molecular dynamics (REMD) simulations[24] as implemented in the Gromacs Package, version 3.3.3,[25] in the NPT ensemble using 22 replicas spanning temperatures 290.0−400.2 K. The REMD simulations were 10 ns in duration, yielding a 220 ns total simulation time for each peptide. Exchanges between replicas were attempted every 1 ps. The peptides were modeled using the CHARMM forcefield[26] and water was modeled using the modified TIP3P potential.[27] See Supporting Information for more details.

Clustering Analysis. A clustering algorithm[28] was used to identify related conformations of the peptides from the trajectory, based on differences in the root mean squared deviation in the backbone atom positions (see Supporting Information for more details).

Flexibility Analysis. Essentially, for each frame in each of the three REMD trajectories, we chose each residue in turn, and calculate the localized "persistence length" of the backbone (both forward and backward from that point in the chain), where the length is given in terms of the peptide repeat unit length. We then average these persistence lengths over all frames in the production trajectory (see Supporting Information for more details).

Peptide-Surface Simulations. Here we modeled the nonbonded interactions between the peptide and the aqueous quartz

(20) Oren, E. E.; Tamerler, C.; Sahin, D.; Hnilova, M.; Seker, U. O. S.; Sarikaya, M.; Samudrala, R. *Bioinformatics* **2007**, *23*, 2816–2822.
(21) Hnilova, M.; Oren, E. E.; Seker, U. O. S.; Wilson, B.; Collino, S.; Evans, J. S.; Tamerler, C.; Sarikaya, M. *Langmuir* **2008**, *24*, 12440–12445.
(22) *Ph.D.-12 Phage Display Library Kit Instruction Manual, version 2.7*; New England Biolabs Inc.: Ipswich, MA, **2006**; p 21.
(23) Collino, S.; Kim, I. W.; Evans, J. S. *Cryst. Growth. Des.* **2006**, *6*, 839–842.
(24) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
(25) Lindahl, E.; Hess, B; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
(26) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(28) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236–240.

Table 1. Motif Enrichment Factors of Various Motifs within the Top-Scored 100 000 Sequences Designed to Bind Quartz

	2 aa Motif	MEF	3aa Motif	MEF	4aa Motif	MEF
1	PW	15	PPW	54	PPPW	142
2	WP	13	PWP	47	PPWP	137
3	WL	10	LPW	39	PPWL	121
4	WM	8	PWL	39	WPPW	113
5	WW	8	WPP	38	LPPW	111
6	MW	8	WLP	34	PWLP	110
7	LW	7	PWM	31	LPWP	109
8	WH	5	PWW	31	PWPP	108
9	QW	5	WPW	28	PPWM	100
10	PP	5	MWP	28	WPPP	94
11	SW	5	WMP	28	WLPP	92
12	WQ	4	MPW	27	PPWW	90
13	LP	4	LWP	26	PWPW	88
14	HW	4	WWP	24	PLPW	84
15	WS	4	PWH	23	PWMP	83
16	MP	4	PMW	22	LPWL	79
17	WY	4	WPL	22	MWPP	79
18	WF	3	PLW	21	PWLW	78
19	NW	3	WLW	20	PWWP	77
20	AW	3	WPH	19	MPWP	76

interface. The fully hydroxylated quartz (100) surface was modeled using the silica force field[29] (designed to be compatible with the CHARMM force field used for the peptides). An equilibrated quartz surface from our previous simulations[30] was used to form a slab containing 6480 SiO_2 atoms with a surface area of 49.5 nm^2 and a thickness of approximately 2.5 nm. For each peptide, between 8 and 21 different starting conformations and orientations with respect to the surface were considered. Each simulation was run for 5−10 ns. All other simulation parameters were the same as for the REMD simulations, except that the system temperature was kept at 300 K. Peptide-surface adsorption energies were calculated using the methods of Yang et al[31] (see Supporting Information for more details).

3. Results and Discussions

In the quest for the amino acid motifs that lead to quartz binding, that is, 2-, 3-, and 4-residue long sequences that are contiguous in the peptides, we first defined a metric, the *motif enrichment factor* (MEF). The MEF indicates the enrichment level of a given motif within the knowledge-based-designed high-affinity peptides compared to their probability of occurrence in the original, experimentally used, biocombinatorial library (see Materials and Methods). In Table 1, we present three sets of top 20 motifs with the highest MEF for dyad, triad, and tetrad motifs present in the top-scoring 100 000 peptides. It is clear that PW, PPW, and PPPW are the most enriched motifs found among the highest scoring quartz-binding peptides. The highest scoring peptides mainly contain Pro along with bulky hydrophobic groups, such as Trp and Leu. One may propose, at this juncture, that these combinations of Pro, Trp, and Leu may contribute to quartz binding either directly *via* interacting with the surface, or indirectly, by dictating a given peptide to have a specific structure so that other residues can interact closely with the solid. To find out which one of these two scenarios, alone or in combination, is effective in peptide binding to quartz, we resorted to the results of the atomistic simulation studies. The atomistic studies were carried out simultaneously, in parallel with the statistical analysis, for an independent exploration of peptide/solid interactions using

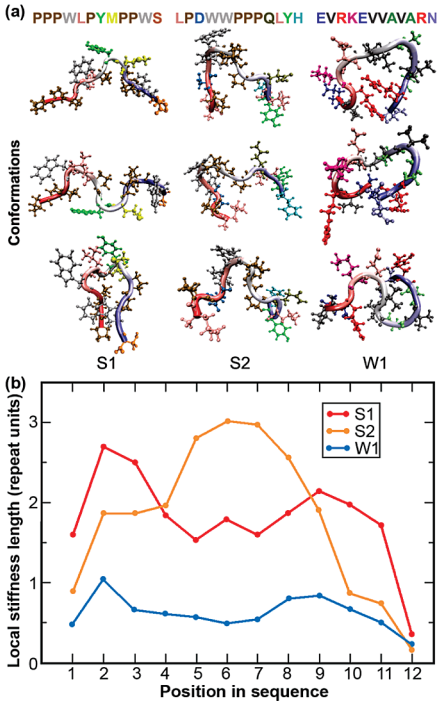

Figure 1. (a) Representative structures of the key conformations of the strong and weak-binding peptides. Water is omitted for clarity. (b) Trajectory-averaged localized stiffness diameter (expressed as multiples of the peptide repeat unit length) as a function of chain position for S1, S2, and W1. Nonpolar: A, L, M, P, V, W; polar: H, N, Q, S, Y; acidic: D, E; and basic: K, R.

the highest (S1 and S2) and lowest (W1) scoring designed peptides whose binding to quartz has been confirmed experimentally in a previous study[20] (Figure SI1 in Supporting Information).

Prior to performing any interfacial simulations, we first studied molecular conformations of each of the peptides in solution using replica exchange molecular dynamics simulations in the absence of a solid surface (see Materials and Methods). A clustering algorithm was used to identify highly occurring conformations in each of the three peptides. While there were only 41 and 47 clusters in S1 and S2, respectively, W1 had as many as 94 clusters. The three most populated cluster conformations for each of the peptides are shown in Figure 1a. The most populated cluster in S1 and S2 constitute 21% and 41%, respectively, of the total conformations, while in W1, the percentage is only 9%. The presence of a large number of clusters in W1, each with a low population, suggests that the weak binder can access a greater number of energetically favorable conformations compared with S1 or S2. To explore the issue of conformational freedom further, we first calculated traditional measures of peptide flexibility such as the radius of gyration or the end-to-end distance, and found these metrics incapable of differentiating between the three peptides. These metrics are more appropriate in the study of peptides with well-defined secondary structures. Since these three peptides do not fall into this category (*vide infra*), we quantify the conformational freedom using a localized "stiffness" length (Figure 1b), which measures an average distance, forward and backward along the chain, for which the backbone conformation persists (see Materials and Methods). The plots in Figure 1b indicate that, while W1 has no substantially persistent backbone structure, S1 and S2 have contiguous regions in their chain backbones that are locally rigid. Thus, S1 can be viewed as a set of two mesogenic, proline-rich regions connected by a semiflexible spacer, while S2 can be described as a central mesogenic region flanked by semiflexible chains.

(29) Lopes, P. E. M.; Murashov, V.; Tazi, M.; Demchuk, E.; MacKerell, A. D. *J. Phys. Chem. B* **2006**, *110*, 2782–2792.
(30) Notman, R.; Walsh, T. R. *Langmuir* **2009**, *25*, 1638–1644.
(31) Yang, M.; Stipp, S. L. S.; Harding, J. H. *Cryst. Growth. Des.* **2008**, *8*, 4066–4074.

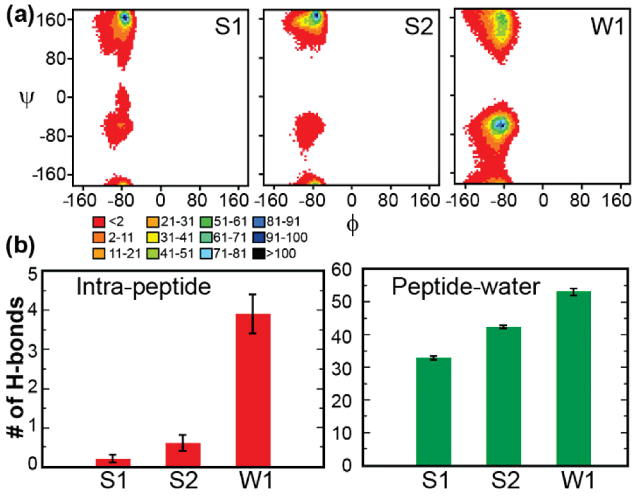

Figure 2. Characterization of the peptides in solution: (a) Ramachandran plots of the strong and weak-binding peptides, colored according to occupancy; (b) average number of intrapeptide and peptide−water hydrogen bonds for each of the peptides.

A Ramachandran plot gives the distribution of the dihedral angles φ against ψ in a peptide. It has characteristic signatures which indicate peptide secondary structures such as α-helix, β-sheet, or random coil. On the basis of the Ramachandran plot for each peptide (Figure 2a), we find that none of the peptides have any well-defined α-helix or β-sheet secondary structure. The φ/ψ angles within the strong binders are mainly clustered at (−75°,160°), characteristic of polyproline type II (PPII) structure. In contrast, the weak binder can be classified having random-coil (RC) structure. Our simulations indicate that W1, which is proline poor, is able to explore a greater region of φ/ψ space than the strong binders; supporting the well-known fact that proline can reduce the conformational freedom of peptides. For each peptide, we also analyzed the average number of intrapeptide and peptide−water hydrogen bonds, as shown in Figure 2b. We find that there are fewer internal hydrogen bonds in S1 and S2 compared with W1. This large number of intrapeptide hydrogen bonds in the weak-binding peptide is attributed to the presence of a greater number of polar/charged residues, coupled with its relatively high backbone flexibility. Although polar/charged groups are present in the strong binders, intrapeptide interactions are not maximized due to spatial constraints, that is, the localized rigidity imposed between these moieties. We propose, therefore, that the comparatively greater internal stability of W1 may not be conducive to promoting peptide−surface interactions, while the strong binders may be driven to the quartz surface to achieve stability via interfacial interactions.

These theoretical findings were also confirmed by experimental analysis (Figure 3). The strong binding sequences exhibit a (−) circular dichroism (CD) ellipticity band centered near 200 nm ($\pi-\pi^*$ transition), and a (+) ellipticity band centered near 230 nm (n−π^* transition), consistent with the presence of PPII secondary structure.[32,33] Interestingly, the n−π^* transition band intensities for S1 and S2, vary; we attribute this to the differences in Pro content in each sequence, that is, six and four Pro in S1 and S2, respectively. In comparison, W1 features a (−) ellipticity band at 198 nm, consistent with the presence of RC conformation in equilibrium with other secondary structures.[23,34] Aside from being the flexibility of W1, we also attribute the presence of RC

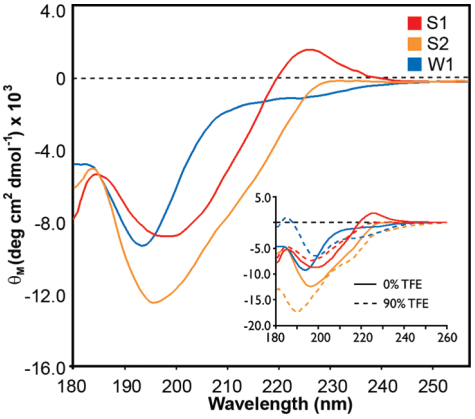

Figure 3. Circular-dichroism spectra of 15 μM S1, S2, and W1 peptides at pH 7.5. Inset spectra shows the effect of 2,2,2-trifluoroethanol (TFE) (90% v/v, dashed lines) on peptide conformation.

conformation to the spatial proximity of charges in the sequence stretch −EVRKE− which may destabilize the conformation.

The addition of the structure-stabilizing solvent, 2,2,2-trifluoroethanol (TFE) to each peptide solution leads to interesting effects in their conformational behavior (Figure 3), yielding a reduction in the intensities of the PPII-specific bands at 200 and 230 nm for S1 and S2, respectively. No significant wavelength shifts in the ellipticity bands are shown for S1, indicating that PPII structure still persists in the presence of TFE and that no other detectable secondary structures are forming for S1. In contrast, for S2 we note a 6 and 2 nm blue shift in the $\pi-\pi^*$ and n−π^* transition band wavelengths, respectively, accompanied by the appearance of a weak (−) band at approximately 220 nm. This suggests that S2 is shifting away from PPII structure toward other secondary structures such as RC (190 nm) or α-helix (220 nm).[23,32,33] This type of transition is not seen for S1 which has a higher Pro content (Figure 3). The conformational response of Pro-deficient W1 to TFE leads to a 5 nm red shift in the $\pi-\pi^*$ ellipticity band and the appearance of a weak (−) band at approximately 220 nm, indicating a folding transition away from RC structures toward α-helix.[23,32,33] Thus, each peptide exhibits different degrees of folding propensity in the presence of TFE which we attribute to different Pro content of each of the peptides.

Following the solution-based studies, we performed molecular dynamics (MD) simulations of each of the three peptides interacting with the quartz (100) surface to evaluate possible binding mechanisms at the molecule/solid interface. First, we carried out several simulations of each peptide in a wide range of starting molecular configurations and orientations on the surface to widen our sampling of conformational space. The interaction energies for the best binding trajectories for S1, S2, and W1 are summarized in Figure SI2 (in Supporting Information).[20] We found that all sequences were capable of binding to the solid substrate in many different conformations, a result that has been seen in similar types of studies in other systems.[35] As shown in Figure SI2 (in Supporting Information), the binding conformations for W1 generally displayed weaker binding on average compared with S1 and S2. Furthermore, while the binding strength of W1 fell sharply after the fourth lowest trajectory, S1 and S2 displayed favorable binding energies for all the trajectories studied. We therefore suggest that the observed binding affinity of S1 and S2 may be due to a combination of favorable energetic and entropic effects, that is, they support a larger number of possible strong-binding configurations that all have relatively

(32) Cubelis, M. V.; Caillez, F.; Blundell, T. L.; Lovell, S. C. *Proteins: Struct. Funct. Bioinform.* **2005**, *58*, 880–892.
(33) Chellgren, B. W.; Creamer, T. P. *Biochemistry* **2004**, *43*, 5864–5869.
(34) Evans, J. S. *Curr. Opin. Colloid Interface Sci.* **2003**, *8*, 48–54.
(35) Tomasio, S. D.; Walsh, T. R. *Mol. Phys.* **2007**, *105*, 221–229.

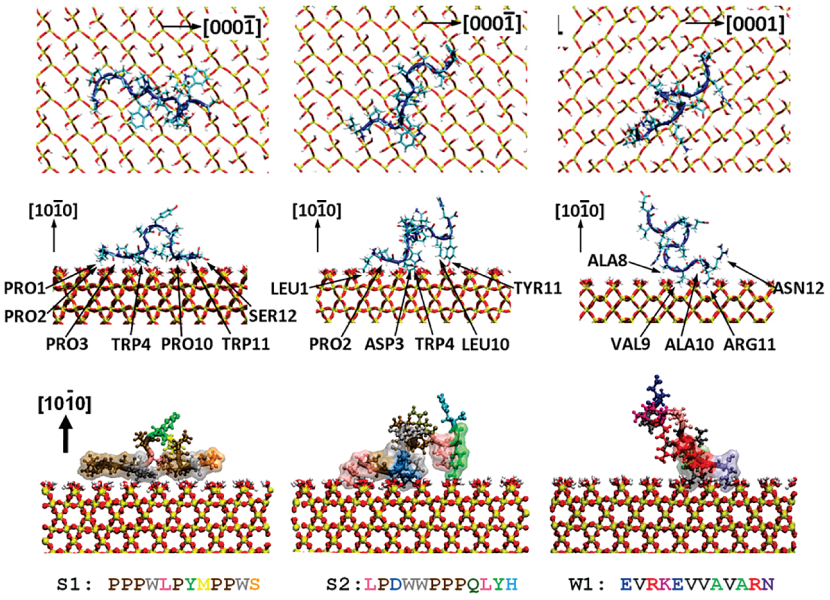

Figure 4. Snapshots of the peptides interacting with the quartz (100) surface for the best binding trajectory in each case. The highlighted residues indicate the surface-bound residues. Water is omitted for clarity.

Table 2. Summary of Residue−Surface Contact Points (Red Text) Obtained from MD Simulations for the Best Five Binding Trajectories in Each Case. The Simulations Had Converged with Respect to the Distance of Each Residue to the Surface

Peptide	S1	S2	W1
Contact Point Residues	PPPWLPYMPPWS	LPDWWPPPQLYH	EVRKEVVAVARN
	PPPWLPYMPPWS	LPDWWPPPQLYH	EVRKEVVAVARN
	PPPWLPYMPPWS	LPDWWPPPQLYH	EVRKEVVAVARN
	PPPWLPYMPPWS	LPDWWPPPQLYH	EVRKEVVAVARN
	PPPWLPYMPPWS	LPDWWPPPQLYH	EVRKEVVAVARN

higher binding strengths than those for the weak-binding sequence. A broad picture that emerges from the analysis of the conformations and the binding strengths of the quartz binding sequences is in a qualitative agreement with experiments (Figure SI1 in Supporting Information).[20] Below, we further interrogate the three peptides for their five best binding trajectories to reveal specific residues interacting with the solid surface.

The residues making contact with the quartz surface in five most stable trajectories are highlighted in Table 2 for S1, S2, and W1. For clarity, three snapshots on the quartz (100) crystallographic surface, each corresponding to the top trajectory for each of the peptides, are shown (Figure 4). Table SI1 summarizes the contact points including the averaged surface-residue distances in the top five trajectories of the peptides. Analysis of these data for S1 and S2 reveals that Pro, Trp, and Leu happen to be the dominant residues in general. In both S1 and S2, the C-termini residues, that is, Ser and His, respectively, also display strong interaction with the surface, including the C-terminal carboxylate group, which consistently made a hydrogen bond with the surface. Thus, our data indicate that the C-terminus carboxylate may co-operatively facilitate surface contact for the 12th, the last, residue. No such influence was noted for the N-terminus. Interestingly, with the exception of Tyr11 in S2, the polar groups in S1 and S2 appear to have not maximized hydrogen-bonding with the surface. Rather, as revealed in Table SI1, the contact is instead dominated by the close proximity of Pro, Trp, and Leu to the surface. Furthermore, these three residues are the only ones in S1 and S2 that could be found *buried* at surface interstitial sites. Previous simulations using the same force field identified local free energy minima at similar molecule−surface separations for hydrophobic residues.[30] This behavior is rationalized by the molecules avoiding unfavorable hydrophobic−solvent interactions, analogous to what has been proposed for antifreeze−protein interactions with ice.[36] In addition to side chain contacts, our data also reveal a substantial contribution from backbone hydrogen bonding. It is well-known that the backbone hydrogen is the only means for some of the residues, such as Pro and Leu, to participate in H-bonding interactions. Trp, on the other hand, can support H-bonding interactions via both the backbone and the indole ring. As a result, Trp confers a limited range of ring orientations relative to the surface, suggesting a conformational adaptability. The behavior of W1 on the surface is distinct from that of the strong binders. Although the initial conformations of W1 have extended structures on the surface, similar to those found for S1 and S2, these do not persist at later stages. For example, in the lowest energy trajectory of W1, eventually only five residues are bound, while the rest of the peptide moves away from the surface. This is unlike the trajectories of the strong binders which, in all cases, persist in extended conformations. As summarized in Table 2 and Table SI1 (in Supporting Information), we conclude that the weak quartz-binding peptide interacts with the solid surface mainly via Val, Ala, and Arg. The nonpolar Val and Ala residues could be attracted to the hydrophobic interstices on the quartz surface as previously observed in simulations of an alanine analogue on quartz.[30] The Arg−surface binding is attributed to a mix of electrostatics and hydrogen bonding.

To assess the probable effects of hydrogen bonding to the stability of peptides on the surface, we also examined the average number of hydrogen bonds that each peptide has both internally

(36) Sicheri, F.; Yang, D. S. C. *Nature* **1995**, *375*, 427–431.

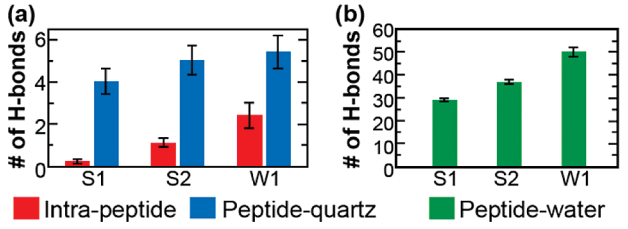

Figure 5. The number of (a) intrapeptide, peptide−quartz, and (b) peptide−water hydrogen bonds, averaged over the best five trajectories of each of the S1, S2, and W1 cases.

and externally, that is, with water and the quartz surface (Figure 5). In both the intrapeptide and peptide−water cases, the trends in the number of hydrogen bonds for all the peptides on the solid surface are similar with what is seen in solution (Figure 2b). On the surface, W1 still has more internal hydrogen bonds, and more peptide−water hydrogen bonds (suggestive of greater internal stability), compared with S1 and S2. The surface-bound W1 loses a certain, although small, number of its internal hydrogen bonds compared to its structure in solution alone. This suggests that W1 has compromised some intrapeptide stability to bind to the surface, although still weakly. In contrast, S1 and S2 do not significantly change the number of intrapeptide hydrogen bonds upon binding to the surface. In fact, within error, S1 and S2 have a comparable number of hydrogen bonds with the surface compared with W1. The analysis here suggests that peptides with a significant internal stability, conferred by either intrapeptide and/or water−peptide interactions, may not be able to form sufficiently strong contacts with the solid surface.

In summary, we generated the motifs from a pool of 100 000 designed peptides and subjected three of these designed sequences (S1, S2, and W1) to an in depth study using MD and CD. To carry out comparable simulations on a greater number of sequences represents a significant and major undertaking; however, the set of simulation data presented here gives us a promising first glimpse into the understanding of peptide/solid interactions. Based on this work, we found a number of factors that might suggest how to distinguish high- and low-affinity quartz-binding peptide sequences, such as PPII secondary structure and the ability to form intrapeptide hydrogen bonds. From the simulations we identified and counted contact points (denoted contact motifs herein) that were contiguous in the sequences (Table SI2 in Supporting Information) and compared these with the most enriched motifs calculated independently (Table 1). In general, there is good agreement between the two sets of results, that is, atomistic simulations and motif enrichment factor. Starting with single residue contacts, the strong binders always had at least one each of Pro, Leu, and Trp in direct surface contact. Further, almost all of the listed contact dyads, and all of the contact triads and tetrads feature at least one Leu, Trp, or Pro, with the exception of two Tyr-containing dyads (*vide infra*). For all possible enriched dyad motifs (Table 1) that S1 and S2 display, such as PW, WP, WL, WW, PP, LP, WS, and MP, our simulations also show these as contact dyads. To be clear, the MEFs that were calculated for the 100000 peptides reveal general trends; the S1 and S2 sequences, for which the simulations were carried out, are specific examples from this large set of sequences. That is, not all the possible motifs found from the MEF analysis that are significant for quartz binding are present in S1 and S2. Similar to dyads, the contact triad data account for all possible enriched motifs present in S1 and S2. Finally, the contact tetrads recover all but one of the possible S1 and S2 tetrads. We suggest that the omission of this tetrad, WPPP, is possibly due its central position in the S2 sequence. Regardless of the initial configuration of S2, one notices that there is a strong *down−up−down−up−down* influence from the backbone on the positioning of the side chains in the central −WPPPQ− stretch such that only the Trp, central Pro, and Gln residues were able to make close surface contact. This effect was not seen at the chain end for PPPW− in S1. We also suggest C-terminus effects (via carboxylate−surface hydrogen bonding) may be biasing the appearance of the −PPWS and −QLYH tetrads. Aside from the agreement between enriched motifs and simulated contact motifs, our data also suggest that Tyr is a promiscuous binder for quartz; it appears to bind regardless of both its position in the sequence and residues that flank it.

Our results not only indicate the importance of individual residues, that is, Pro, Trp, and Leu, but also suggest that the local environment in which each residue is located may play a role. The simulation data suggest that key motifs identified by the MEF calculations are indeed involved in direct interaction with the surface, although this does not rule out additional structural role(s) played by these motifs as well. In this light, we propose that, as a residue, Pro plays a dual role in peptide−quartz interactions. One role may be the extension of backbone adjacent to the Pro-rich domains, thus facilitating the presentation of residues which flank these domains, close to the surface. A corollary is that these Pro-rich regions may hinder the formation of intrapeptide hydrogen bonds. One further role is the capability of the Pro-rich regions themselves in making close surface contacts. The dual role of Pro as a "spacer" as well as a "binder" is well-known for proline-rich regions in proteins.[37] We propose that Pro-rich S1 and S2 may be preferred both enthalpically and entropically: first, via formation of hydrogen bonds with the surface (rather than with itself), second in terms of hydrophobic shielding via burial in the surface interstices, and finally since the Pro-rich regions will lose relatively less entropy upon binding. While in this work Pro is seen to play a significant role in interfacial binding; it would be naive to assume that domains with contiguous Pro content will *always* lead to better binding, regardless of the target surface. In the specific case of quartz, Pro works well because it has a small hydrophobic group; such moieties are thought to be driven to the quartz under aqueous conditions.[30]

Our arguments presented above will not extend to inorganic surfaces where Pro adsorption is not favorable. As we discussed here, other peptide characteristics may play significant roles in quartz binding. While the internal stability of W1 may not favor peptide−surface interactions, the strong binders may be driven to the surface to achieve stability via interfacial interactions. Since they are hydrophobic, Pro, Trp, and Leu residues in strong binding peptides could be found *buried* at surface interstitial sites. Further work by using a combined approach, such as ours, is required to identify similar propensities for other residues and solid surfaces. In addition to considering motifs rather than residues in isolation, our results suggest that the positioning of these motifs may also be critical. These findings open up new possibilities in the systematic decoding of what really controls molecular binding at the peptide−inorganic interfaces. The statistical analysis, valuable in identifying these motifs, can be exploited further *via* simulations and experiments where these motifs are placed in different spatial locations along a given sequence; for example, at a chain end, or a chain center; flanked by an enriched motif, or flanked with an under-enriched motif; positioned between two flexible regions, and so forth. Characterization of sequences comprising mixtures and tandem repeats of these motifs would also form an interesting test of our hypotheses and is planned for future studies. In principle it might be interesting to study the binding

(37) Williamson, M. P. *Biochem. J.* **1994**, *297*, 249–260.

characteristics of a dodecapeptide consisting entirely of the same amino acids such as proline. As discussed above, the amino acids present in the enriched motifs interact with the solid surface via a co-operative effect. Although, proline is one of the dominant residues in these motifs, the same effect, however, cannot be induced solely with its tandem repeats. This is due to the induced structural changes and inherent instability[38] of the poly(proline) compared to the peptide units containing the enriched motifs. Also, in practice, the synthesis of poly(proline) is problematic for two reasons: first, because the sequence Pro−Pro is prone to give diketopiperazines,[39] and, second, during the purification of repetitive peptide sequences, fragments lacking one or more amino acids (deletion peptides) are structurally related to the target peptide.[40] For these reasons, we have not attempted to study the binding characteristics of poly(proline) on silica.

The behavior of genetically selected and *de novo* designed peptides share aspects with those found in natural biological−inorganic systems, namely those associated with intrinsically disordered proteins (IDP).[41,42] The presence of IDP-like characteristics, such as structural disorder, PPII structure, and conformational transformability, appear to exist within the quartz-binding peptides studied here. Presently we do not fully understand why IDP-like features such as PPII, but not RC conformation, are associated with higher affinity for quartz. This suggests that specific molecular "rules" may be at work; certain IDP-like features may be more adept at facilitating binding than others. Clearly, resolution of such issues require further detailed examination. If intrinsically disordered sequences are indeed selected during biopanning in the biocombinatorial selection against different materials, such sequences may offer functionally important characteristics as they could be relevant for engineered solid-binding peptides used as molecular linkers, couplers, and assemblers for bionanotechnology and biomimetic materials applications.[10,13,19,43−45]

4. Conclusions

Combining two different approaches, statistical analysis and atomistic simulations, we reveal molecular mechanisms responsible for peptide binding at the aqueous−quartz (100) interface. The work carried out herein identifies intrapeptide interactions and ranks the stability of the sequences in solution thereby providing invaluable data in understanding the behavior of the surface-bound peptides. The peptide−surface studies signify a link between the predicted enriched motifs and the principal points of contact on the surface. The simulations suggest that regional peptide flexibility may also play a role. Our results propose new research directions where the positional dependence of these enriched motifs could be systematically varied to tune potential binding behavior. Knowledge regarding propensity rules connecting sequence, structure, and function is crucial if one aspires to move beyond *taxonomy* and toward *prediction*, to fully exploit these interfacial properties of the peptides at solid surfaces in the controlled synthesis, self-assembly, and formation of functional nanostructured materials under ambient conditions. Our efforts summarized here represent the first steps in this direction. While our results are immediately applicable to quartz, the combined comprehensive approach discussed here is generalizable and could be applied to other engineered peptide−inorganic solid systems in the future.

Acknowledgment. This work was mainly supported (E.E.O., J.S.E., R.S., C.T., M.S.) by the NSF-MRSEC program at the University of Washington and by EPSRC (EP/E02095X/1), U.K (R.N., T.W.). We are grateful to the Centre for Scientific Computing, Warwick, and the National Grid Service (NGS), U.K., for generous access to computing resources. Portions of this work represent Contribution 39 from the Laboratory for Chemical Physics (I.W.K., J.S.E.), New York University and the Turkish State Planning Organization (C.T.). E.E.O., C.T., and M.S. also acknowledge the partial support by NSF BioMat program.

Supporting Information Available: Methodological details on the REMD and peptide−surface simulations and cluster and flexibility analyses; figures showing SPR and Q-dot immobilization assays and peptide/surface interaction energies; tables containing contact residues and their time averaged distances to the surface. This material is available free of charge via the Internet at http://pubs.acs.org.

(38) Guruprasad, K.; Reddy, B. V. B.; Pandit, M. W. *Protein Eng*. **1990**, *4*, 155–161.
(39) Rothe, M.; Mazánek, J. *Angew. Chem., Int. Ed. Engl*. **1972**, *11*, 293.
(40) Dalcol, I.; Rabanal, F.; Ludevid, M. D.; Albericio, F.; Giralt, E. *J. Org. Chem*. **1995**, *60*, 7575–7581.
(41) Uversky, V. N. *Protein Sci*. **2002**, *11*, 739–756.
(42) Tompa, P. *Trends Biochem. Sci*. **2002**, *27*, 527–533.
(43) Mirkin, C. A.; Letsinger, R. L.; Mucic, R. C.; Storhoff, J. J. *Nature* **1996**, *382*, 607–609.
(44) Hartgerink, J. D.; Beniash, E.; Stupp, S. I. *Science* **2001**, *294*, 1684–1688.
(45) Langer, R.; Peppas, N. A. *AIChE J*. **2003**, *49*, 2990–3006.