Supplementary Information: Impact of Glutamate Carboxylation in the Adsorption of the α-1 Domain of Osteocalcin to Hydroxyapatite and Titania

Sarah Alamdari[1], *Jim Pfaendtner[1]

[1]Dept. of Chemical Engineering, University of Washington, Seattle 98195-1750

*Email: jpfaendt@uw.edu

Gamma Carboxyl Glutamic Acid Forcefield Parameters

The following list includes the atom names, and the corresponding CHARMM36 atom types that were used to model γ-carboxyglutamic acid (Gla). These were adjusted from existing CHARMM36 atom types. A general CH atom was used to model the C_γ atom, and partial charges for that atom were adjusted as shown by column 3 so that the entire Gla amino acid had an overall charge of -2, and each carboxylate group retained the same overall charge. This approach has also been used in previous studies.[1,2] Other studies have shown that this Gla-carboxylate/Ca^{2+} interaction has both been over and underestimated by previous iterations of the CHARMM forcefield parameters, but have not yet been investigated yet with the CHARMM36 release, or the $Ca^{1.5+}$ model used by INTERFACE.[3,4] In this study, due to the different ions used within the surface model constraints, a generalized approach parameterized from CHARMM36 was used.

Partial charges for Gla

```
  N    NH1   -0.470  0
 HN      H    0.310  1
 CA    CT1    0.070  2
 HA    HB1    0.090  3
 CB   CT2A   -0.180  4
HB1    HA2    0.090  5
HB2    HA2    0.090  6
 CG    CT1   -0.290  7
HG1    HA1    0.090  8
JA1     CC    0.620  9
OE1     OC   -0.760 10
OE2     OC   -0.760 11
JA2     CC    0.620 12
OE3     OC   -0.760 13
OE4     OC   -0.760 14
  C      C    0.510 15
  O      O   -0.510 16
```

Table S1. Specific simulation setup details for PT-WTE-MTD

Simulation	Peptide	Surface	Charge			Counterions		
			Surface	Peptide	Total	$Ca^{1.5+}$	Ca^{2+}	Cl^-
1	dOC	HAP	0	-2	-2	2	-	1
2	OC	HAP	0	-5	-5	4	-	1
3	dOC	Neutral Hydroxylated TiO_2	0	-2	-2	-	1	0
4	OC	Neutral Hydroxylated TiO_2	0	-5	-5	-	3	1

Table S2. Specific simulation setup details for wtMTD

Simulation	Peptide	Surface	Charge			Counterions		
			Surface	Peptide	Total	$Ca^{1.5+}$	Na^+	Cl^-
5	GLU	HAP	0	-1	-1	2	-	2
6	GLA	HAP	0	-2	-2	2	-	1
7	GLU	Neutral Hydroxylated TiO_2	0	-1	-1	-	1	0
8	GLA	Neutral Hydroxylated TiO_2	0	-2	-2	-	2	0

Table S3. Temperatures used for each replica of PT-WTE

HA		TiO_2	
Replica	**Temperature (K)**	**Replica**	**Temperature (K)**
1	300	**1**	300
2	305	**2**	305
3	310	**3**	311
4	316	**4**	316
5	321	**5**	322
6	326	**6**	328
7	332	**7**	333
8	337	**8**	339
9	343	**9**	345
10	349	**10**	352
11	355	**11**	358
12	361	**12**	364
13	367	**13**	371
14	373	**14**	377
15	380	**15**	384
16	387	**16**	391
17	393	**17**	398
18	399	**18**	405
19	406	**19**	412
20	414	**20**	419
21	421	**21**	426
22	428	**22**	434
23	435	**23**	442
24	442	**24**	450
25	450	**-**	-

Convergence of Enhanced Sampling Simulations and Reweighting

Binding free energies were calculated using a Boltzmann averaged difference of the Helmholtz free energy from the surface adsorbed and solution states, as shown in Equations S1 and S2. Here r represents the peptide center of mass (COM) distance to the surface. The region from r_0 to r_1 is the region where the peptide COM was considered bound (distance <= 3.0), and

r_2 is the distance at which the wall was placed as described in the Methods section. The time series monitoring binding energies are shown in Figure S1.

$$A_{surface} = \frac{1}{r_1 - r_0} \int_{r_0}^{r_1} e^{-\frac{A(r)}{k_B T}} dr \quad , \qquad A_{solution} = \frac{1}{r_2 - r_1} \int_{r_1}^{r_2} e^{-\frac{A(r)}{k_B T}} dr \tag{S1}$$

$$A_{binding} = -k_B T \ln\left(\frac{A_{surface}}{A_{solution}}\right) \tag{S2}$$

The binding energy reported in Figure S1 was calculated by taking the mean of the Boltzmann weighted binding energies calculated during the last 30% of simulation time. The binding energies for capped Glu and Gla were calculated to be -17.55 kJ/mol ± 0.27 kJ/mol and -22.7 kJ/mol ± 1.05 kJ/mol on HA, and -15.4 kJ/mol ± 1.13 kJ/mol and -26.0 kJ/mol ± 0.97 kJ/mol on TiO_2, respectively. The binding energies of dOC are -16.0 ±1.2 kJ/mol and -40.5±0.9 kJ/mol for OC on HA, and -67.9 ±2.5 kJ/mol for dOC and -6.1 ±1.7 kJ/mol for OC on TiO_2 respectively. Errors reported were calculated by taking the standard deviation of binding energy from the last 30% of the simulation.

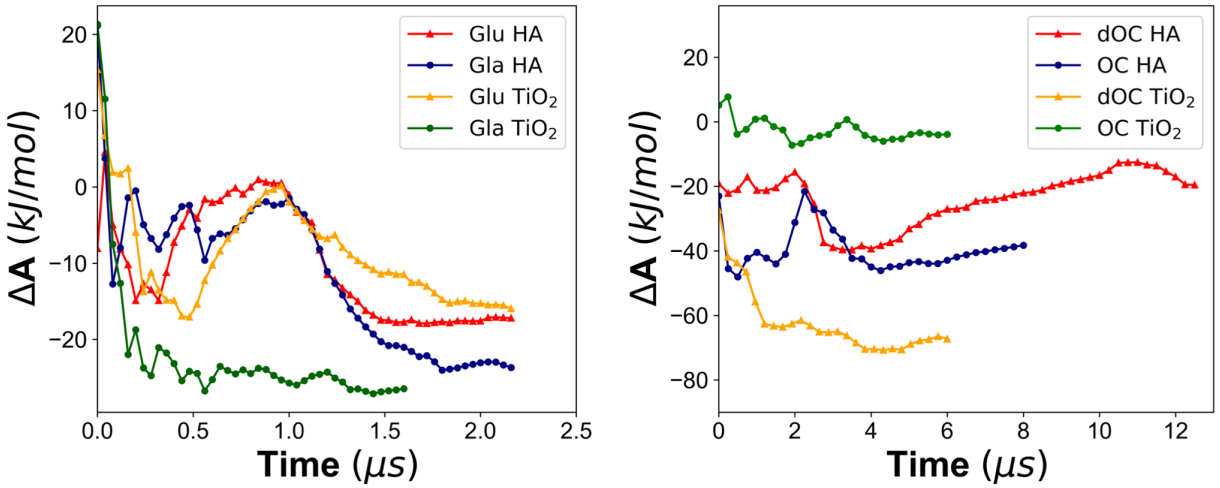

Figure S1. Binding free energy calculated from Equation S2 as a function of time for all wtMTD simulations of capped amino acids (left) plotted every 2.5 ns, for clarity. Binding energy calculated for the 300K replica in PTMetaD-WTE simulations (right) plotted every 20 ns, for clarity.

Free-energies and other equilibrium averages were obtained from the biased simulations via reweighting. For all reweighting analyses we used static bias approximation. That is, the trajectory of interest is re-processed using the final MetaD bias potential from each simulation and each frame is assigned a weight based amount of bias given by the use of the final static bias; weights are obtained with the Torrie-Valleau[5] scheme (n.b., this method was inspired by Branduardi et al.[6] and was shown to provide exact agreement with reference potentials in multiple prior studies[7,8])

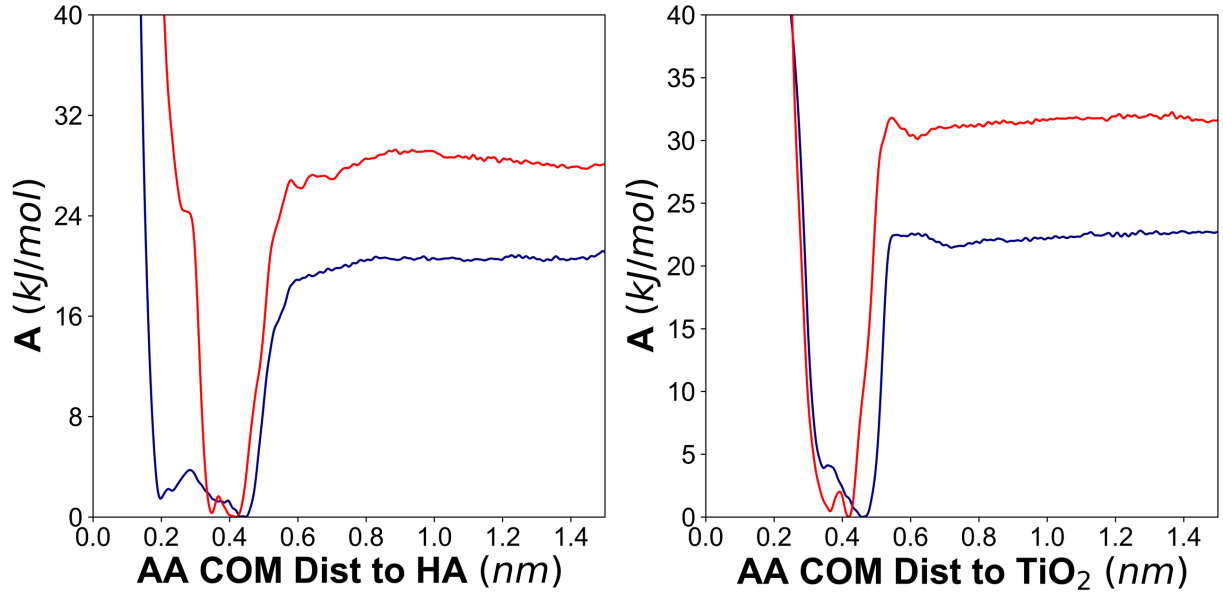

Figure S2. Free energy profile for Glu (blue) and Gla (red) binding onto HA (left) and TiO_2 (right). Whereas the binding profiles in Figure 3 and 4 show the reweighted free-energy as a function of the side-chain COM, these profiles represent the free-energy of the COM of the entire amino acid, derived from the sum of the MetaD bias potential

Clustering

Clustering was done using the gromos[9] method developed by Daura et al. implemented in GROMACS. Gromos is a geometric clustering method that uses a root mean square deviation (RMSD) distance metric to group structures given a cutoff value. Peptides were clustered based on RMSD of the alpha carbon atoms in the peptide backbone using a cutoff value of 0.2 nm.[10]

Structures were processed every two to four frames using the ensemble of biased structures at 300K, classified based on their COM distance to the surface (distance <= 3.0 nm away from the surface). The cluster weights from this analysis were then reweighted as described above to correct for the effect of the MetaD bias. The top 3 clusters for dOC on HA and TiO2 are included in Figure S3, and OC on HA and TiO2 are included in S4. The top 3 structures in Figure S3 and S4 represent at least 50% of the peptide conformations observed at the surface. The top 3 solution structures are shown in Figure S5 and S6, to demonstrate conformational preferences in solution. Convergence of clustering shown in Figure S5, was performed by calculating the total number of clusters as a function of trajectory frames clustered.

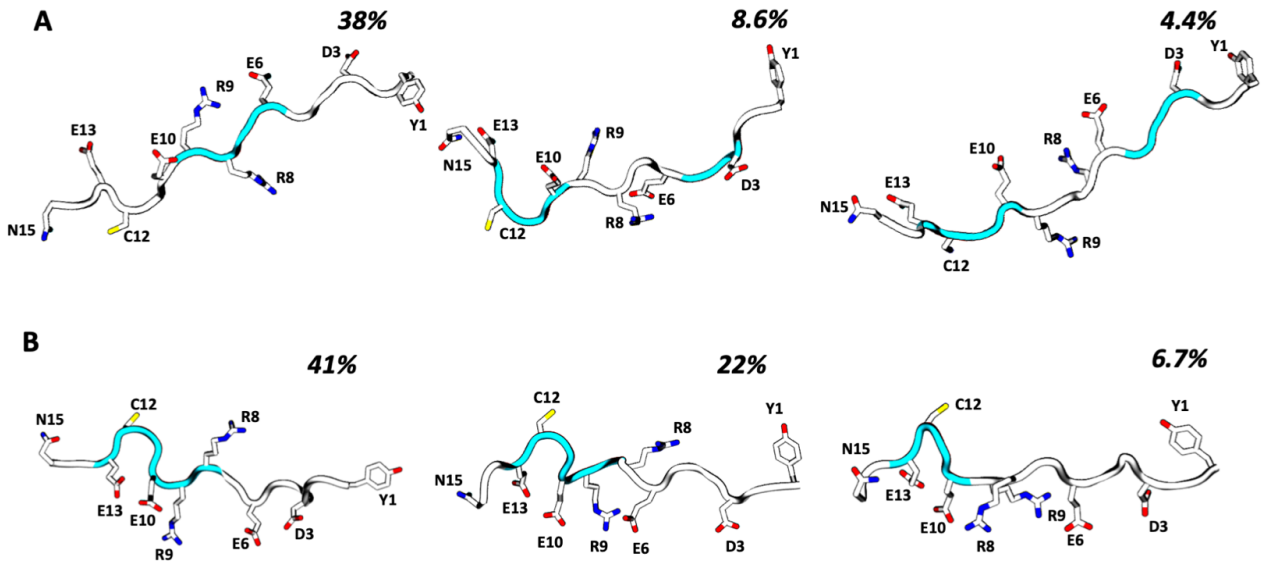

Figure S3. Top 3 structures from clustering analysis for adsorbed dOC in row A) HA and row B) TiO_2 surfaces. Percentages listed over the structures correspond to the reweighted % of simulation that structure represents.

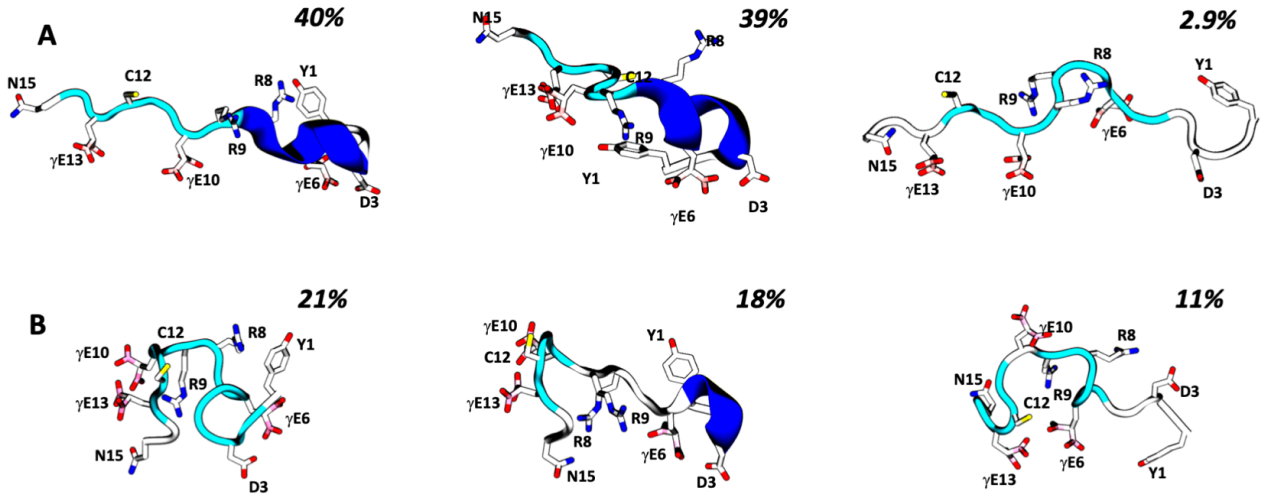

Figure S4. Top 3 structures from clustering analysis for adsorbed OC in row A) HA and row B) TiO_2 surfaces. Percentages listed over the structures correspond to the reweighted % of simulation that structure represents.

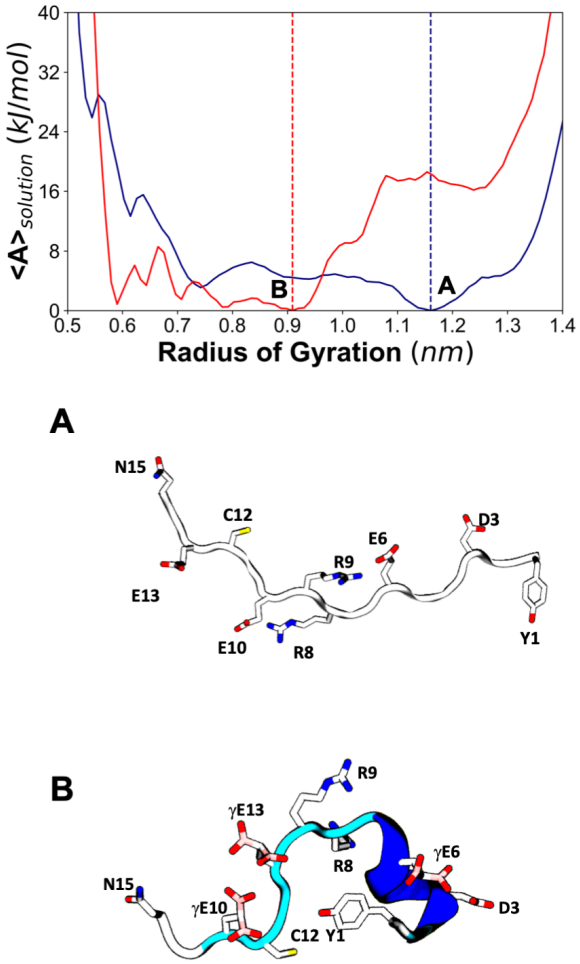

Figure S5. Free energy profiles for dOC and OC in HA systems along the radius of gyration for structures in solution (COM distance > 3.0 nm)

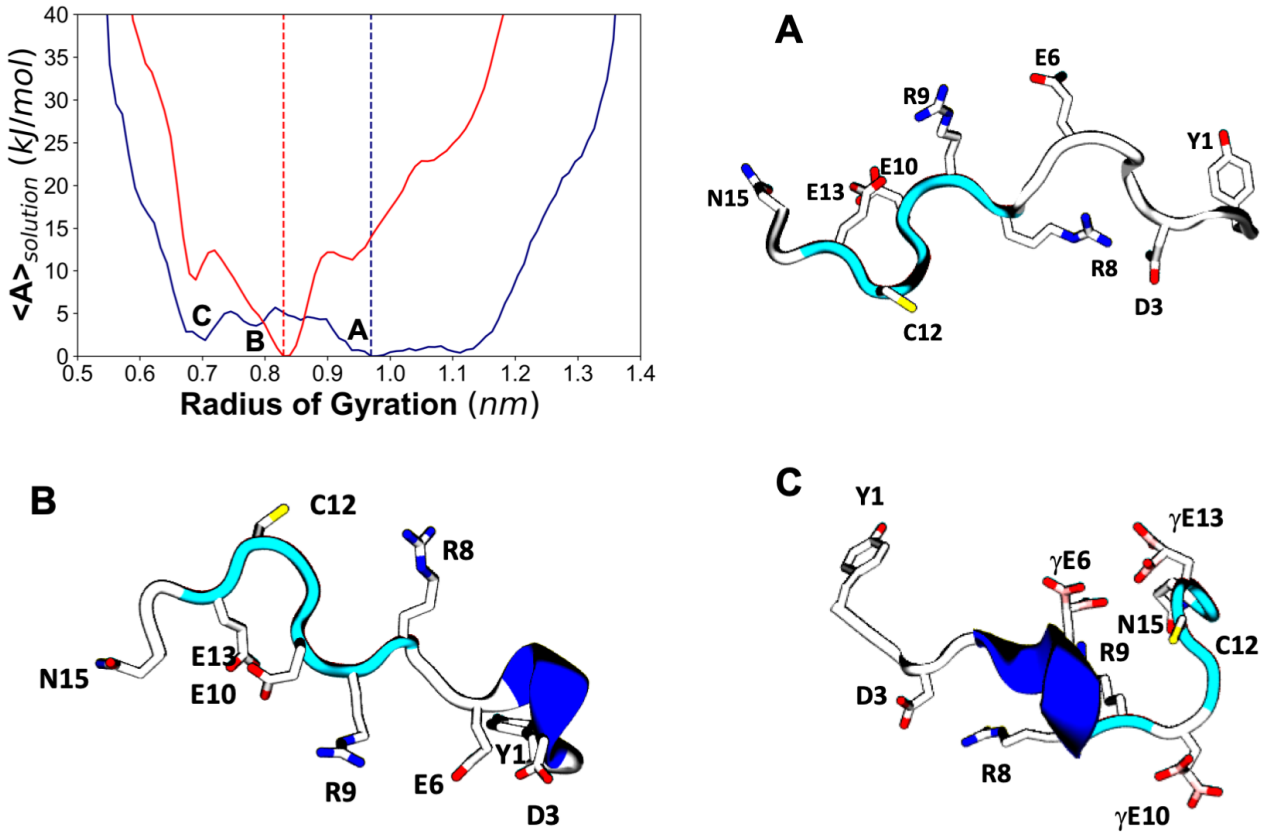

Figure S6. Free energy profiles for dOC (blue) and OC (red) in TiO_2 systems along the radius of gyration for structures in solution (COM distance to the surface > 3.0 nm).

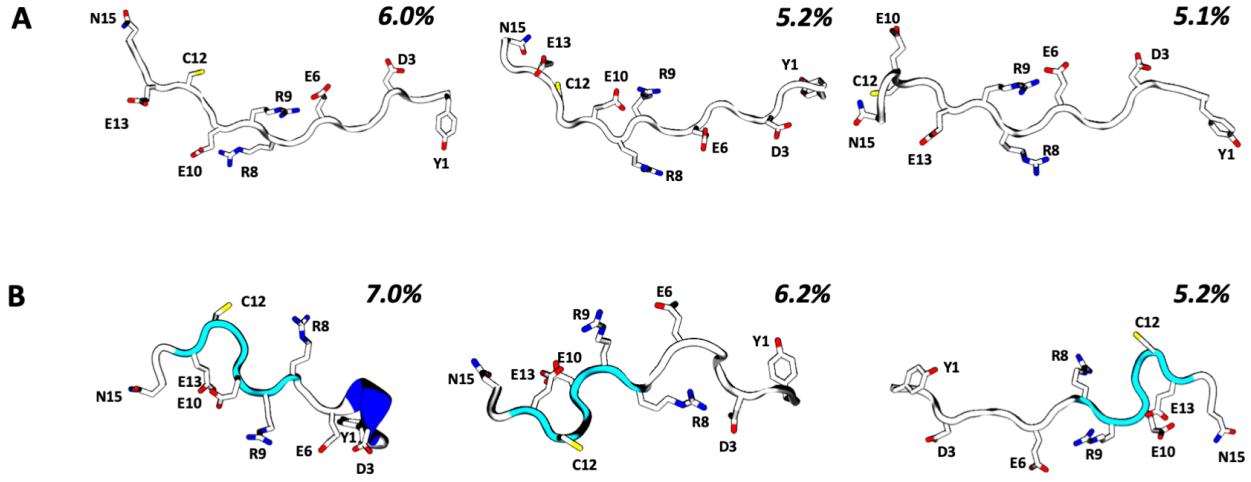

Figure S7. Top 3 structures from clustering analysis for solution dOC in row A) HA systems and row B) TiO_2 systems. Percentages listed over the structures correspond to the reweighted % of simulation that structure represents

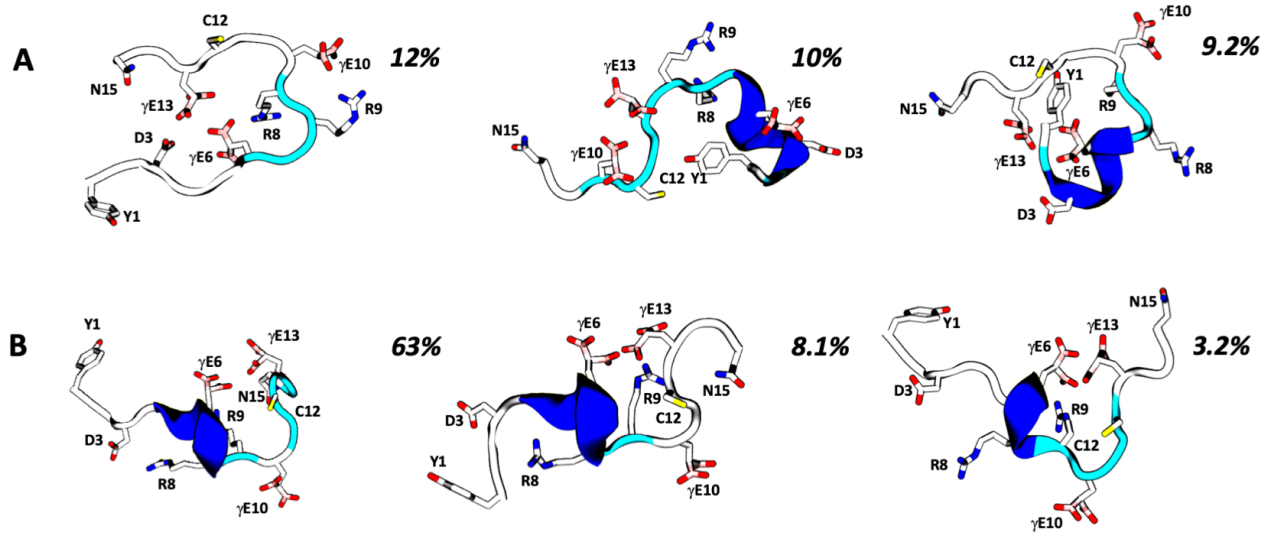

Figure S8. Top 3 structures from clustering analysis for solution OC in row A) HA system and row B) TiO_2 systems. Percentages listed over the structures correspond to the reweighted % of simulation that structure represents.

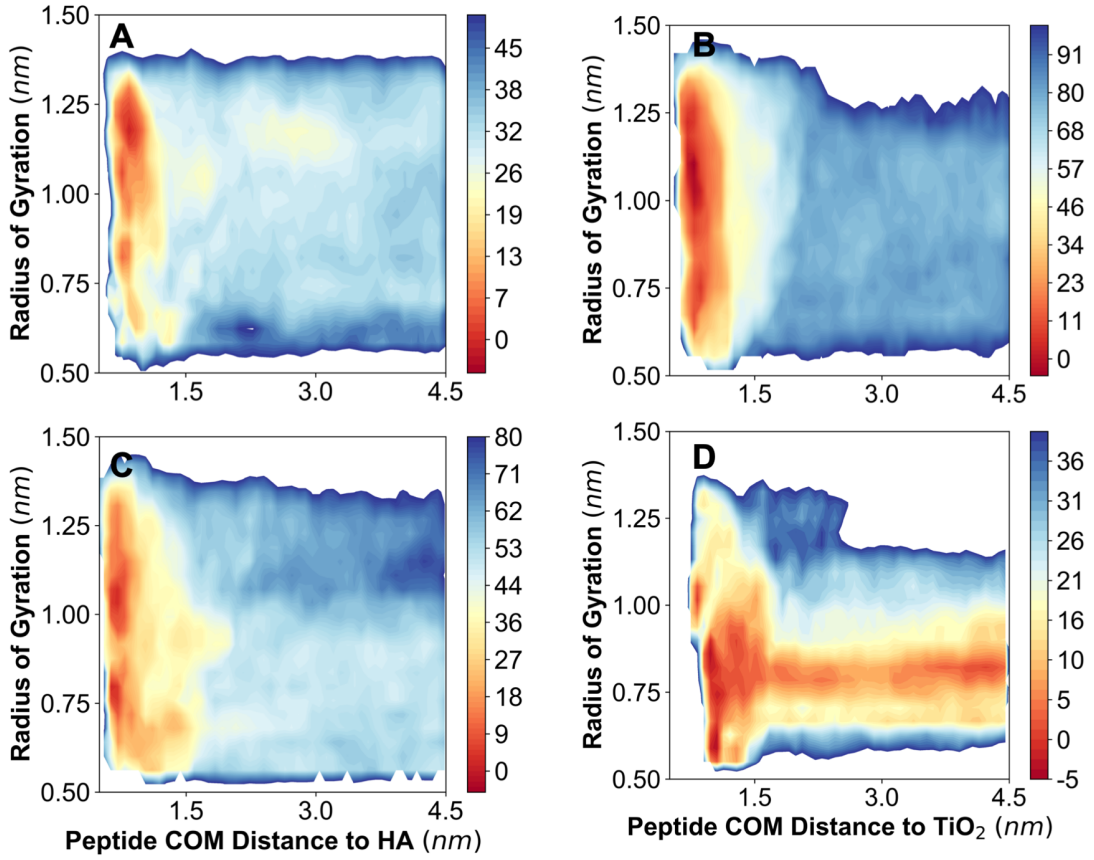

Figure S9. FES for A) dOC on HA, B) dOC on TiO_2, C) OC on HA, and d) OC on TiO2. Note the difference in energy scales, given in kJ/mol for clarity in seeing the minimum energy bound states.

Characterization of surface bound structure

To characterize the minimum energy adsorbed clusters, simulations were reweighted using Torrie-Valleau[5] reweighting scheme using only the surface adsorbed structures (peptide COM distance <= 3.0 nm). The structures were then reweighted along 2 coordination number (CN) CVs, to produce a free energy surface (FES) describing the environment between the peptide, and ions (Ca) or hydroxyls (H) at the surface. The CN is represented by a continuous switching function between atoms i and j following the form given in Equation 3 where r_0, n, and m are adjustable parameters to control the shift, and slope of the switching function.

$$CN = \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^n}{1 - \left(\frac{r_{ij}}{r_0}\right)^m} \tag{3}$$

The reweighted CVs shown are: (1) the CN between each oxygen on every carboxylate in Glu or Gla and all calcium ions (r_0 = 0.35 nm), and (2) the CN between each oxygen and hydrogen on surface hydroxyls (r_0 = 0.2 nm). This was done to capture the amount of surface contacts through the unique interaction with each O on the Glu and Gla residues with the surface. Values of 6, and 12 were used for n and m in both CVs.

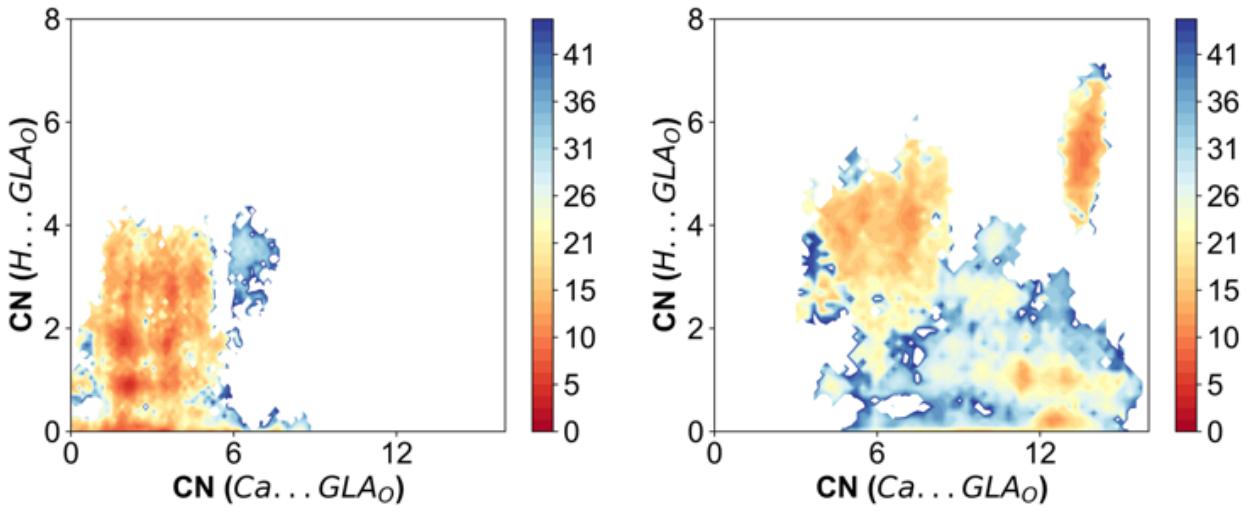

Figure S10. Characterization of the surface bound structures of dOC (left) and OC (right) on HA

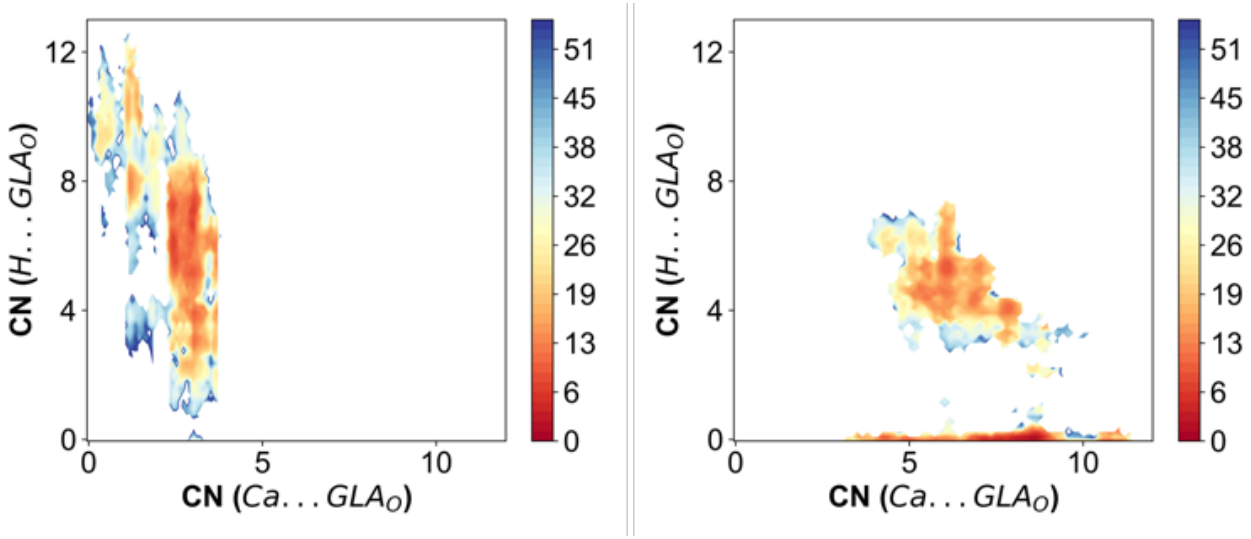

Figure S11. Characterization of the surface bound structures of dOC (left) and OC (right) on TiO_2

Characterization of OC helical structure

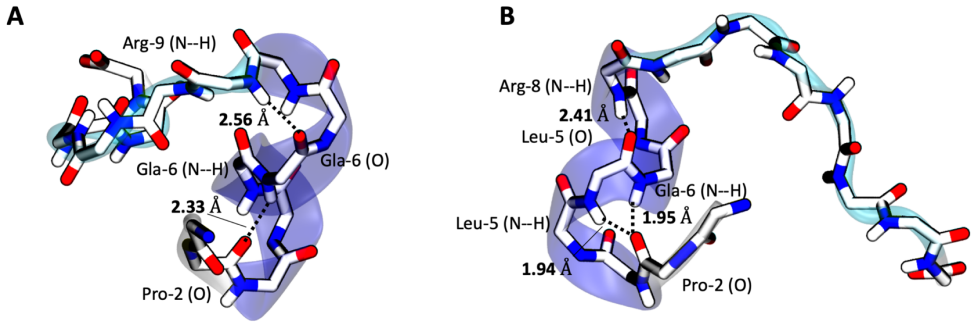

Figure S12. Characterization of the kinked helical structure in OC, showing only the backbone atoms in licorice, and transparent overlay of the helical structures obtained from the proteins adsorbed on HA. **A)** A tight helix formed by i+4 h-bonding with a proline kink in the first helix, followed by i+3 h-bonding in the second helix. **B)** Another tight helix formed by combined i+3, i+4 bonding in the first helix with a proline kink with residues Leu-5 and Gla-6, followed by the same i+3 hydrogen bonding mode in the second helix. Protein structure is shown transparently, and backbone residues are shown in licorice following the same color scheme as the Figure 6.

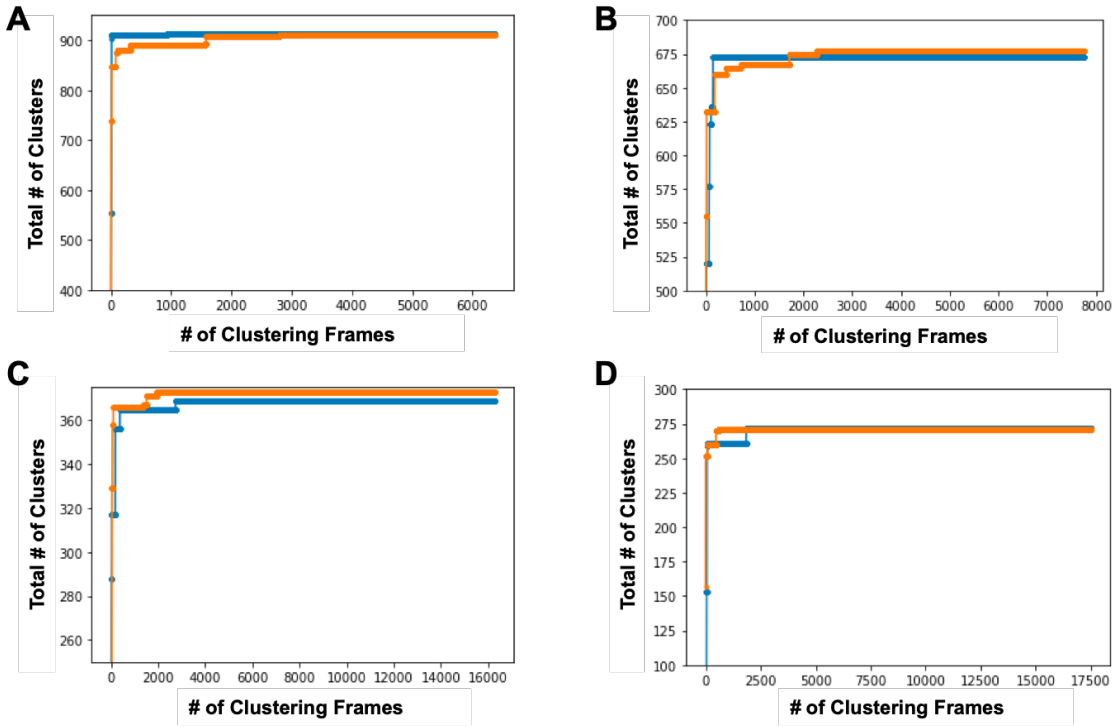

Figure S13. Convergence of clustering calculations for (A) dOC and C) OC on HA; and B) dOC and D) OC on TiO2. Convergence of surface clusters (peptide COM distance <= 3.0 nm from the surface) are given by the blue line, and solution clusters (peptide COM distances > 3.0 nm from the surface) are shown by the orange line.

References

1 S. Banerjee, R. Prasad and P. Sen, *ACS Omega*, 2019, **4**, 496–508.
2 A. B. Sorensen, J. J. Madsen, L. A. Svensson, A. A. Pedersen, H. Østergaard, M. T. Overgaard, O. H. Olsen and P. S. Gandhi, *J. Biol. Chem.*, 2016, **291**, 4671–4683.
3 A. T. Church, Z. E. Hughes and T. R. Walsh, *RSC Adv.*, 2015, **5**, 67820–67828.
4 E. Project, E. Nachliel and M. Gutman, *J. Comput. Chem.*, 2008, **29**, 1163–1169.
5 G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.
6 D. Branduardi, G. Bussi and M. Parrinello, *J. Chem. Theory Comput.*, 2012, **8**, 2247–2254.
7 J. Pfaendtner and M. Bonomi, *J. Chem. Theory Comput.*, 2015, **11**, 5062–5067.
8 C. D. Fu and J. Pfaendtner, *J. Chem. Theory Comput.*, 2018, **14**, 2516–2525.
9 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chemie Int. Ed.*, 1999, **38**, 236–240.
10 Y. Yan, D. Zhang and S.-Y. Huang, *J. Cheminform.*, 2017, **9**, 59.