Electronic Supplementary Information

Conformationally Directed Assembly of Peptides on 2D Surfaces Mediated by Thermal Stimuli

Tyler D. Jorgenson[1,2], Madelyn Milligan[2,3], Mehmet Sarikaya[1,2,3,4], René M. Overney[*1,4]

[1]Molecular Engineering and Sciences Institute, University of Washington, Seattle, WA
[2]GEMSEC, Genetically Engineered Materials Science and Engineering Center, University of Washington, Seattle, WA
[3]Department of Material Science and Engineering, University of Washington, Seattle, WA
[4]Department of Chemical Engineering, University of Washington, Seattle, WA

S1: Peptide Synthesis

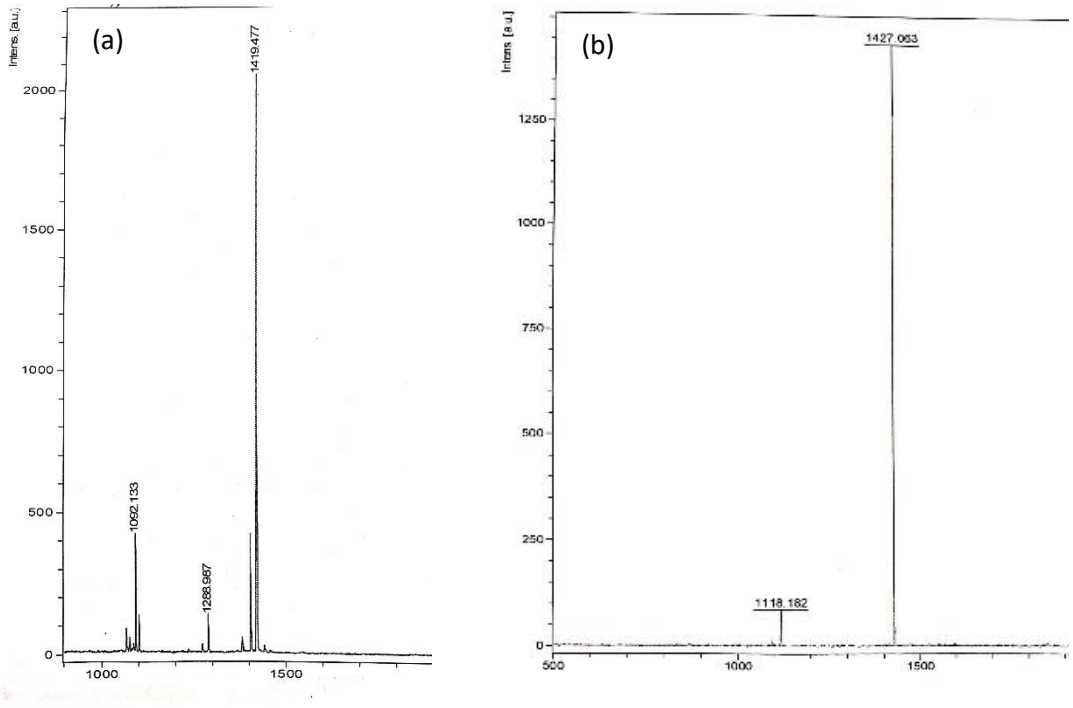

Figure S1. Mass Spectra of Synthesized Peptides. (a) WT spectra. The theoretical molecular weight of WT is 1383.4 Daltons. The observed mass spectra peak is at 1419.5 Daltons meaning there is potentially an attached K^+ ion. The purity is approximately 80% based on relative intensities. (b) Mass Spectra for M2. The theoretical molecular weight of M2 is 1427.5 Daltons. The observed mass spectra peak is at 1427.05 Daltons. The purity based on intensities is approximately 90%.

S2: AFM Analysis Technique

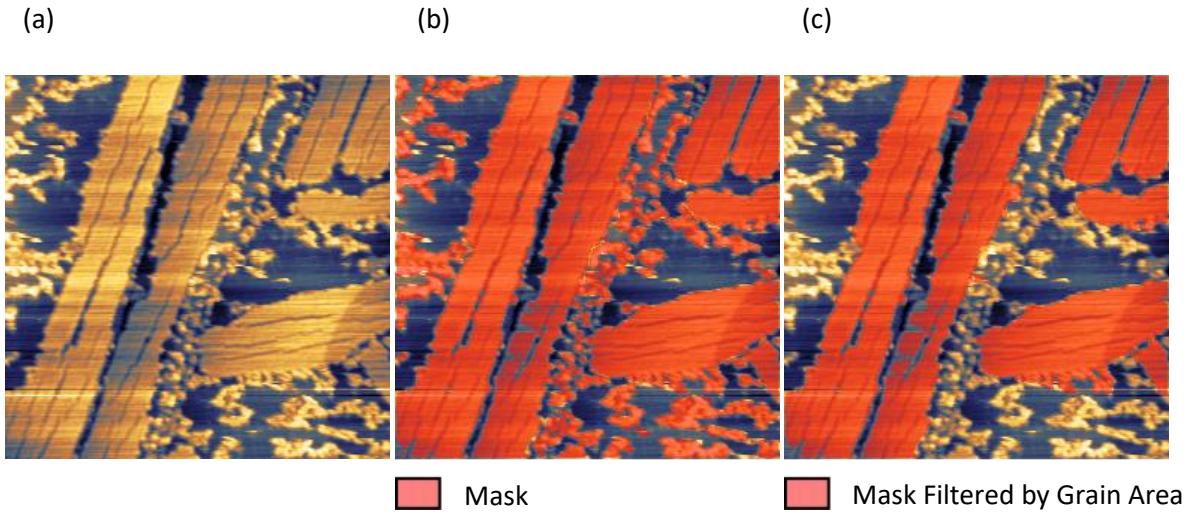

Figure S2. Example Masking Technique for Structural Analyses. (a) AFM images were processed using the Gwyddion software package. The images are initially flattened and cropped to remove any graphite step edges that might cause faulty masking and further processing. If too many step edges or image artifacts are present the phase or amplitude image can typically also be used in place of the height image. (b) A mask is applied to the flattened image by determining a minimum height, phase, or amplitude depending on the image of choice that separates the peptide from the underlying graphite surface. Some additional manual editing of the mask was performed to remove erroneously masked graphite when the computer was unable to automatically exclude it. Moreover, manual editing of the mask was performed when assembled and amorphous peptide structures were erroneously connected as part of the same grain. The mask is then filtered by a variety of grain parameters depending on the quality of the image. In part (c) the mask is filtered by grain size. This was often the most effective at separating the ordered and amorphous regions.

S3: Evolution of Peptide Solution Structure

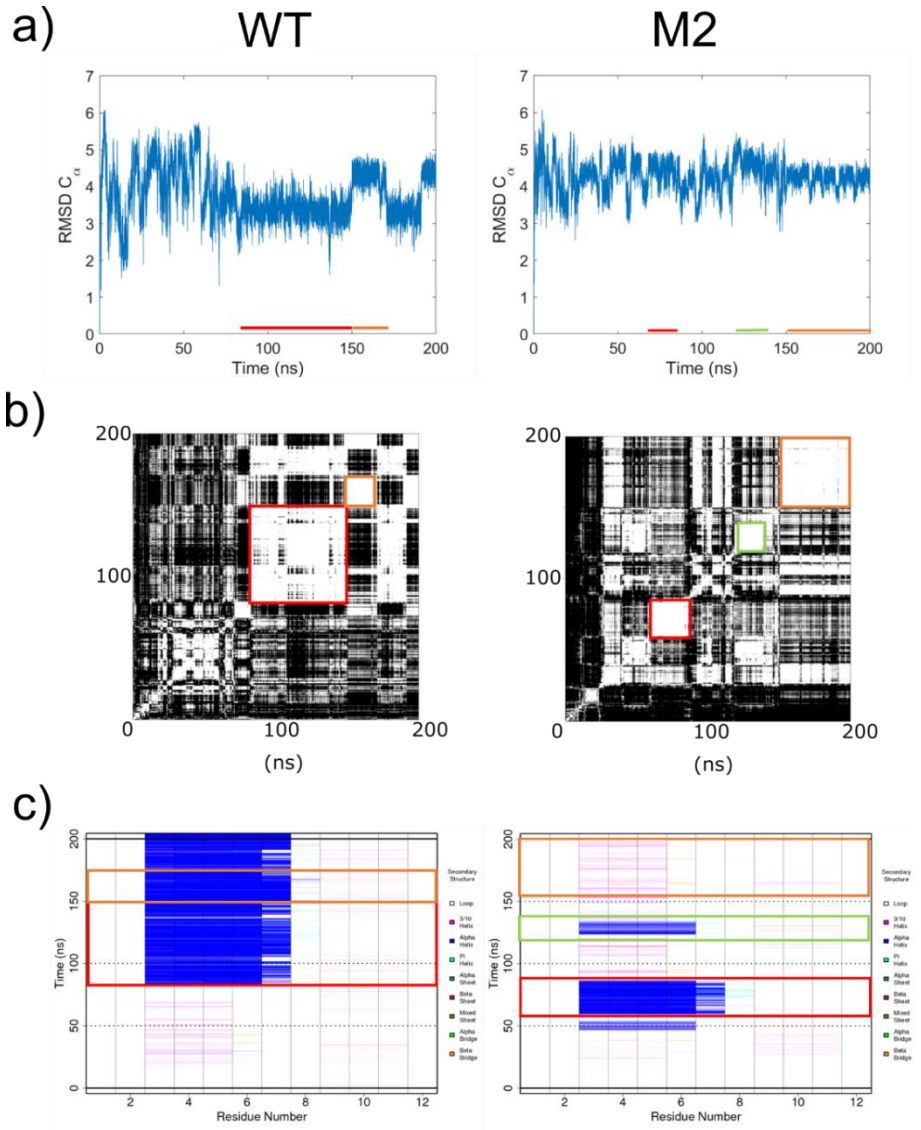

Figure S3. The Evolution of GrBP5-WT and M2 Solution Structure. (a) Plots of the Cα RMSD as the simulations progressed. Showing a potential conformational switching behavior for WT. (b) RMSD maps show the structural relationship between all structures within the ensemble of structures generated during the simulation. Black indicates a RMSD of greater than 4 Å while white indicates an RMSD of less than 2 Å. The white boxes highlighted correspond to the times of switching Cα RMSD for WT. (c) Color coded plot showing the evolution of secondary structure during the simulation. Blue represents an α-helix, pink represents a 3/10 helix, cyan represents a π-helix. Color coded, red, orange and green superimposed boxes correspond to the different conformational ensembles observed in the RMSD maps in (b) showing that the different stable states correspond to different secondary structures. For WT the primary difference was in ASP8. M2 folds in and out of the three observed conformations. Differences in observed conformations are further discussed in Supplementary Comment 3. A more fine grained ensemble clustering presented in the main text was performed to further validate these analyses and determine all structural states observed during the simulation.

S4: Backbone Conformational Propensities in Solution

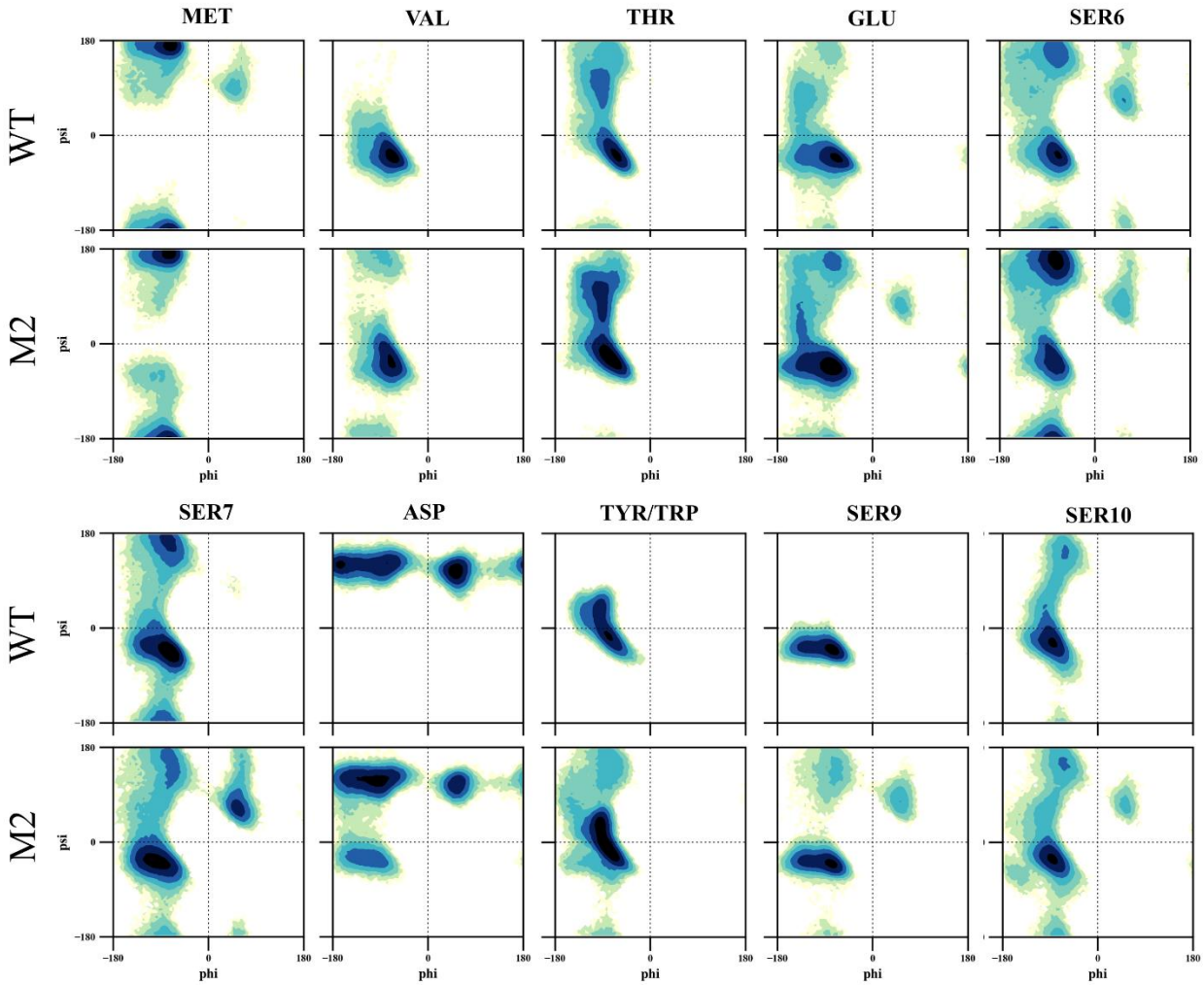

Figure S4. (Φ, Ψ) Conformational Propensities of WT and M2. The upper left-hand quadrant is indicative of a β-sheet, while the lower left-hand quadrant is indicative of right-handed α-helical conformational space. The upper right-hand side is representative of left-handed α-helical conformational space. The darkest blue indicates the highest population while lighter blues and greens represent lower population. The population is scaled by the total number of angle pairs to get a probability density. Residues Glu, Ser6, Ser7, Asp8, Trp9, and Ser 10 and 11 sampled more conformational space in M2 than WT. These differences were related to the unstable alpha helix and inconsistency in a hydrogen bonding network between Ser 10 and 11 and Asp8. These distributions were used to estimate energetics of the observed peptide solution structures. See Supplementary Comment 2 for specifics on the calculations of free energies and entropies.

S5: Side Chain Conformational States of WT

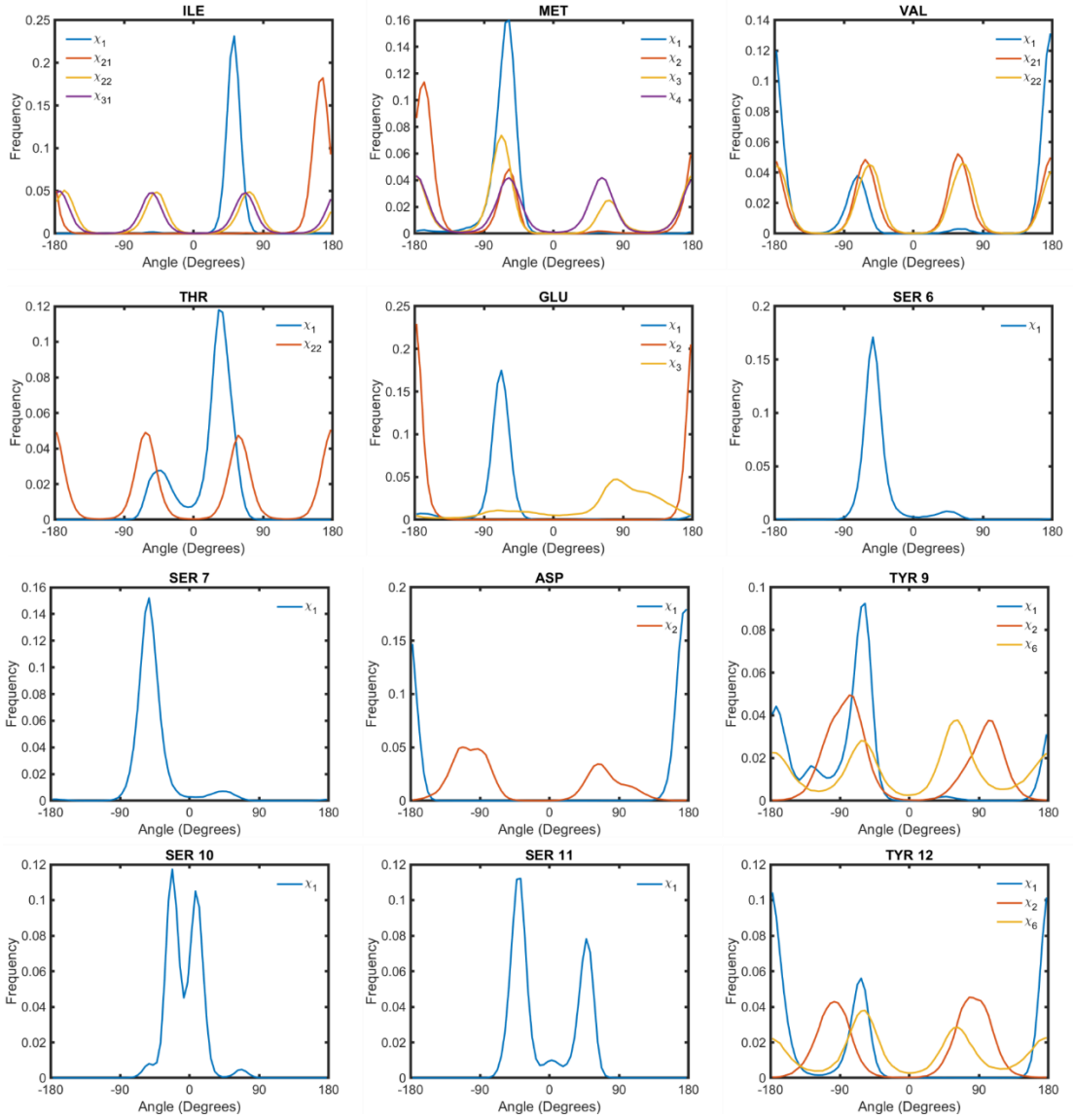

Figure S5. Side Chain Conformational States of WT. The Chi angles for each side chain were binned by every 5 degrees. The population density was determined by dividing the counts in each bin by the total number of counts. Hydrogen bonding networks between Ser 10 and 11 and Asp 8 caused different rotameric states to be sampled that a "free" serine like Ser 6 or 7. These distributions were used to calculate rotamer entropy.

S6: Side Chain Conformational States of M2

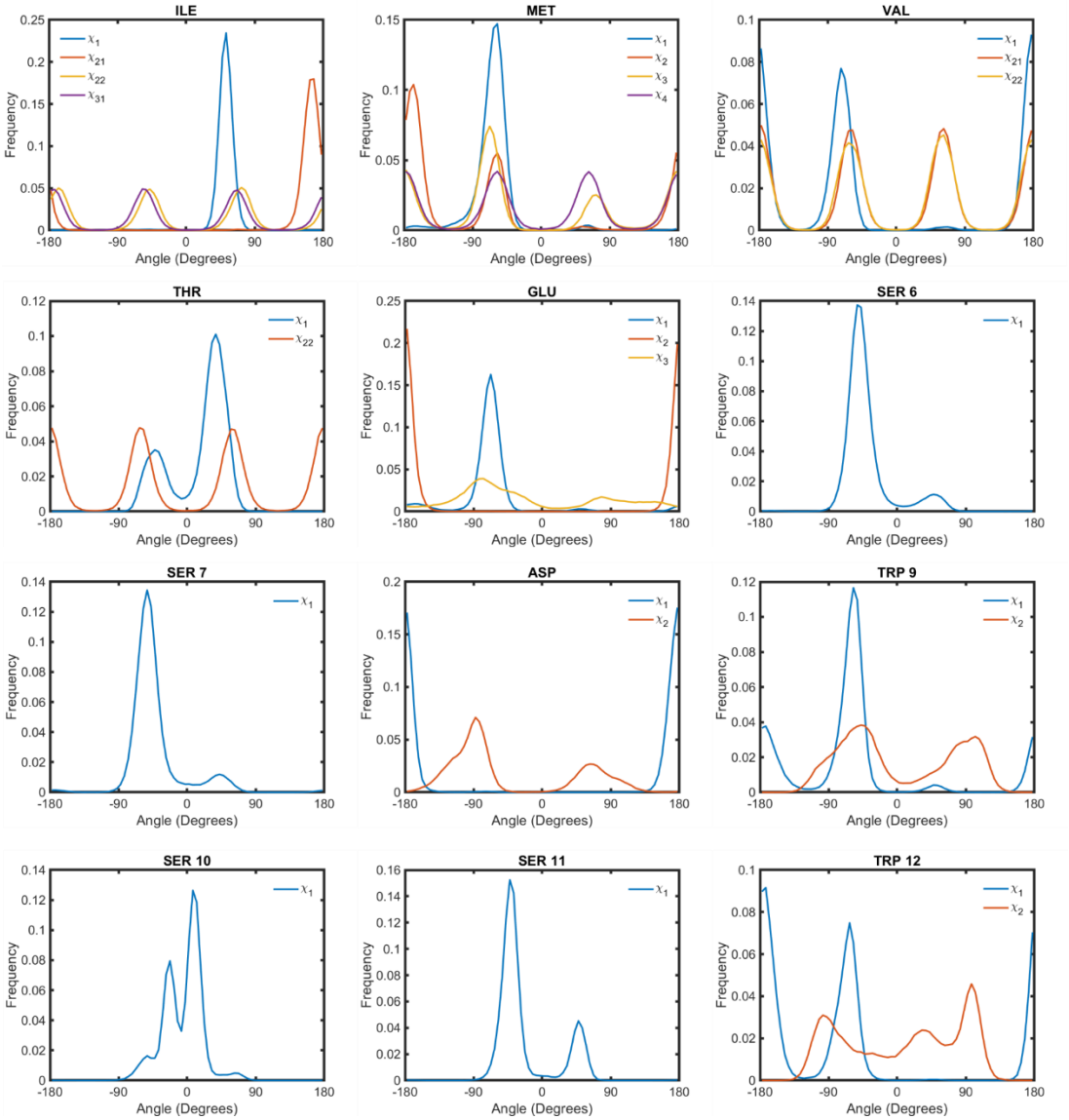

Figure S6. Side Chain Conformational States of M2. The Chi angles for each side chain were binned by every 5 degrees. The population density was determined by dividing the counts in each bin by the total number of counts. As with WT, Ser 10 and 11 have a high density of non-standard rotameric states associated with a hydrogen bonding network with Asp 8.

S7: Asymmetric Contact Map

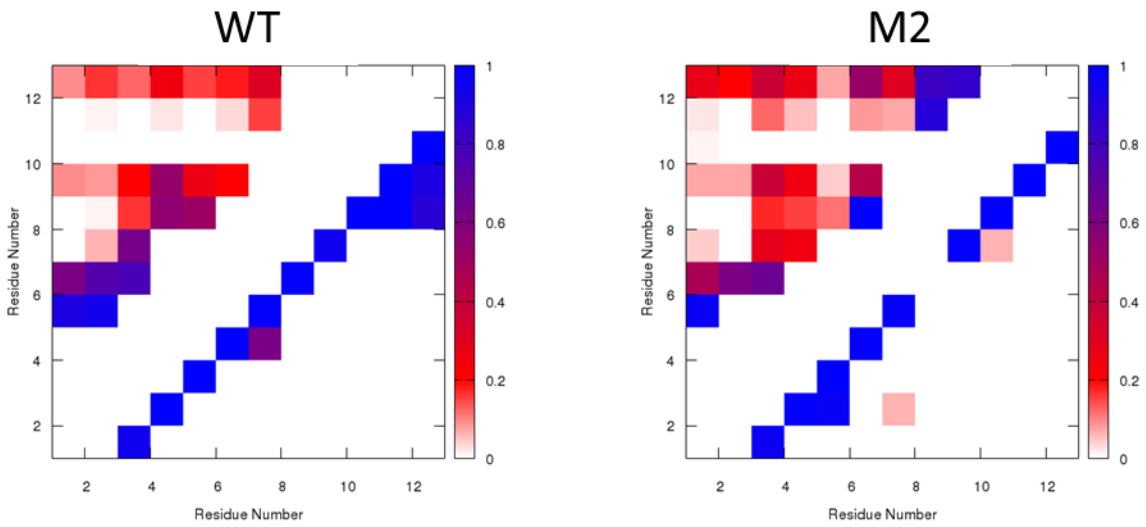

Figure S7. Asymmetric Contact Map. The asymmetric contact map compares contacts present in the starting structure (native contacts) to contacts that developed during the simulation (non-native contacts) These contacts can be side-chain/side-chain, main-chain/main-chain, or side-chain/main-chain van der Waals or hydrogen bonding interactions. The frequency of native (lower-right) and non-native (upper-left) contacts are displayed via a color bar in which blue indicates high frequency while white represents low frequency. The x- and y-axes are the residue position from N to C-terminus. For both WT and M2 most of the contacts present in the starting structure were retained through-out the simulation. The high frequency blue and purple non-native contacts observed in WT in the N-terminus is the development of the alpha-helix. As discussed in the main text, this alpha helix is less stable in M2 thus the contact frequency is lower. Given the arbitrary nature of the starting structure for WT and M2, all native and non-native contacts must be compared. For example, contacts in M2 between Trp12 and Asp8 and Trp9 are considered non-native however, the analogous contacts in WT are considered native. These contacts are a well-defined, stable hydrogen bonding network between the C-terminal serines, main-chain nitrogens of Trp12 and Asp8 that is present in both peptides discussed in Comment S3. The non-native contacts between Trp12 and the N-terminus are more frequent than the analogous interactions in WT suggesting a reason why the N-terminal alpha helix may be unstable.

S8: Peptide Adsorption to Graphene Simulations

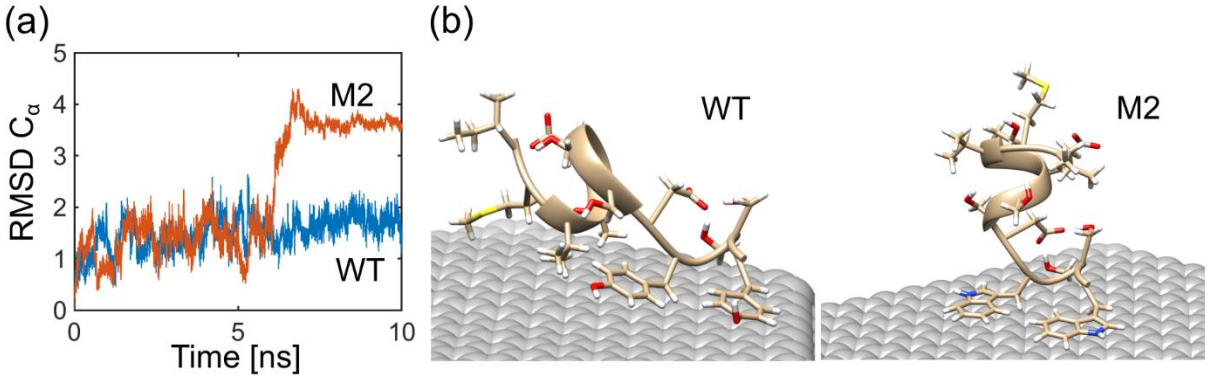

Figure S8. MD adsorption of WT and M2 to a graphene sheet. (a) RMSD of WT and M2 post adsorption reveals the RMSD for WT increased to about 1.75 Å while for M2, the RMSD values oscillated around 1.5 Å before a sudden shift to 3.75 Å. The small change in RMSD for WT suggests that the solution structure is maintained upon adsorption, *i.e.* within the 2 Å structural deviation used to determine the conformational states in solution. (b) Representative structures of the equilibrium WT and M2 peptide on a graphene surface. A more detailed discussion of the graphene simulation set up and analysis can be found in Supplementary Comment 4.

S9: Backbone and Side-chain Conformational Similarity

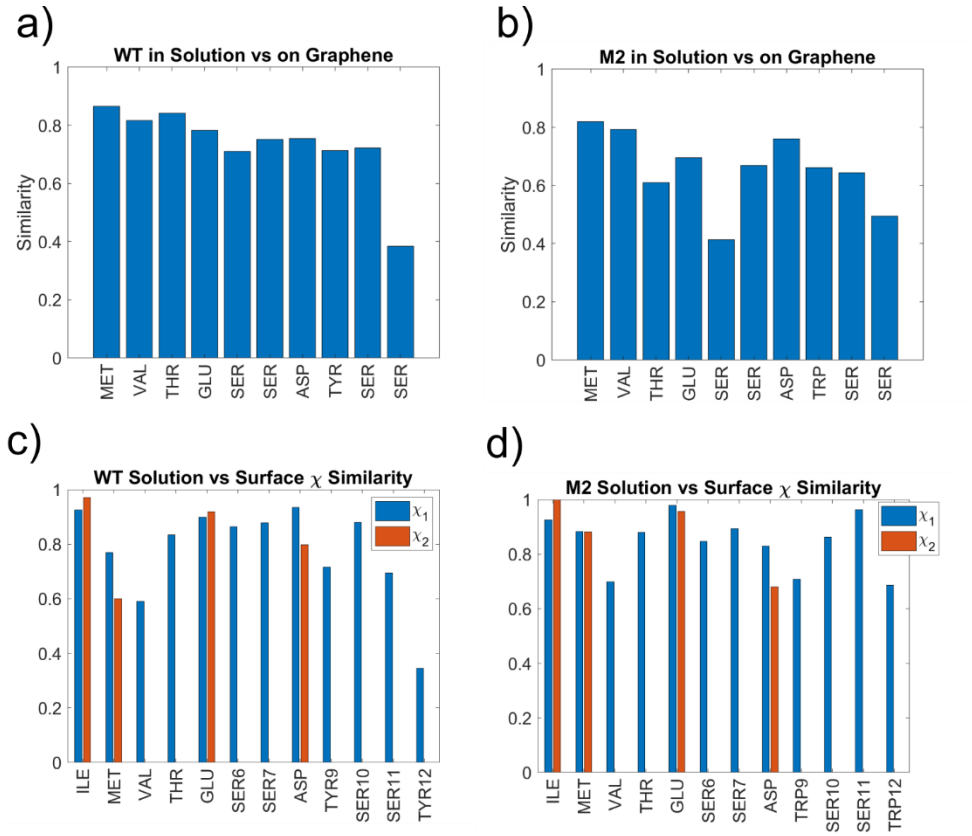

Figure S9. Phi-Psi and Chi Angle Similarity: Solvated vs. Adsorbed. The similarity is a number between 0 and 1 that describes the overlap between two different population distributions., for more information see Supplementary Comment 2. The similarity of WT (a) and M2 (b) backbone phi-psi angles show that the overall conformation remained the same after adsorption to the graphene surface. Side chain conformations for solvated and adsorbed peptides for WT (c) and M2 (d) also shows high similarity for all side-chains except those specifically in contact with the graphene, Tyr, Trp, etc.

Comment S10: *Generation of Ramachandran Plots, Structural Similarity, and Entropic Analysis*

(Φ, Ψ) pairs were binned into a 2D histogram with 72 × 72 bins comprising 5° × 5° increments. The bins were scaled by the total number of (Φ, Ψ) pairs generating a population distribution. (Φ, Ψ) population distributions were compared between the two peptides, GrBP5-WT and GrBP5-M2, as well as between the free solution and interfacial-constrained structures using a similarity metric defined as:

$$M = \sum \min(A(i,j), B(i,j)),$$

in which A and B are two different population distributions and A(i,j) and B(i,j) are the population density in one 5° × 5° bin. Rotamer and χ population distributions were compared using the same metric. For further structural analysis of our system, the rotamer and χ population distributions were compared to the conformational propensities of the individual amino acids.[1] The (Φ, Ψ) and χ1 population distributions were used to estimate the backbone and side-chain conformational entropy as:

$$S = -R \sum \mathrm{A(i,j)}\ \ln[\mathrm{A(i,j)}]$$

Comment S11: Automated Ensemble Clustering

Simulations were clustered using built in ensemble analysis in UCSF Chimera.[2] A step size of 5-6 frames was used for the analysis. This step sizes corresponded to approximately .5 ns. This yielded the ensembles with clear structural differences. Shorter step sizes yielded more ensembles, but many were very similar in terms of structure, i.e., within 1 angstrom RMSD. Even for a step size of 5, some ensembles were very structurally similar to each other. If the ensembles were within 2 angstroms RMSD we consolidated the ensembles together. This led to 56 distinct ensembles for WT and 80 distinct ensembles for M2. The probability for each ensemble was taken as the number of sampled frames containing said ensemble divided by the total number of frames sampled. The probabilities were then used to estimate free energies using:

$$G = -RT \times \ln\left(\frac{p_i}{1 - p_i}\right)$$

In which p_i is the probability of a given ensemble. From the estimated energies the distribution of structures was estimated using:

$$p_i = \frac{e^{E_i/RT}}{\sum e^{E_j/RT}}$$

In which the p_i is the probability of state i and the sum is over all observed states, j. E_i is the estimated free energy of state i.

Comment S12: Structural Differences between Conformational Ensembles

Inspection of the secondary structure of WT showed that the primary structural differences between the 2nd and 3rd ensemble was that Asp formed a π-helix for the first stable structure and

was part of the α-helix in the second stable structure. Furthermore, in the second stable conformation the C-terminal residues formed a 3/10 helix. The conformational switching between the 2nd and 3rd stable WT ensemble appears to be mediated by the (Φ, Ψ) angle pair of Asp, as previously discussed. For M2, the two stable structures that have α-helical secondary structure were very similar to each other; the primary difference being the relative location of the C-terminus to the α-helix. This change also appears to be result of the (Φ, Ψ) angle pair of the Asp. For M2, the residues that form the α-helix sampled more conformational space than in WT which we attributed to the fact that M2 was α-helical for less simulation time. Ser7 in M2 had significant population density in left-handed α-helical conformational space throughout the simulation that was not as significant in WT. This conformational state is associated with the backbone loop in M2 after the helix.

Visual inspection of Asp in both WT and M2 exposed that it hydrogen bonds with Ser10 and Ser11 while consistently switching between the serines being bound to the same or different hydrogen bond acceptors on Asp. Inspection of side-chain contacts revealed that Trp12 in M2 made frequent non-native contacts with the N-terminal amino acids (Supplementary Figure 5). These contacts were not present, or much less frequent, with the Tyr12 in WT and destabilized the N-terminal α-helix, causing M2 to be more disperse in structure.

(Φ, Ψ) angles with lower similarity resulted from M2 having a less stable solution structure, i.e. greater conformational entropy. The lower similarity value for Ser10 (M~0.45) is for reasons previously discussed. Rotameric states were highly similar (M > .8) between WT and M2. To judge how typical these rotameric population distributions were for the given residue we compared them to the population density determined for the individual residue. This analysis yielded values of M < 0.5 for all the rotamers suggesting the amino acids in the peptide had preferred conformations.

Comment S13: Set-up and Discussion of Graphene Simulations

The aromatic rings of the Tyr or Trp were placed facing the graphene surface as these are believed to be the primary residue associated with binding. WT state 1 and M2 state 2 were chosen because they have clear secondary structure which was assumed to be easier to detect conformational changes, e.g. unfolding, etc. Peptide/graphene simulations were simulated for 10 ns. For WT, the two stable states observed in solution were highly similar in secondary structure with the primary differences being positioning of tyrosines relative to the N-terminal α-helix. Since the graphene surface pinned these tyrosines, we expect these two structures to become conformationally indistinct after adsorption. The large RMSD change for M2 suggests that it switched to a different structural state. In fact, the observed RMSD shift corresponds to the α-helix twisting and changing the conformation from the second to the first structural state shown in Figure 1c. WT and M2 adsorption, as expected, was mediated by the aromatic C-terminus resulting in the development of a 3/10 helix between the pinned aromatic residues.

The secondary structure of WT and M2 remained unchanged and appears to be stabilized by adsorption to graphene. WT and M2 backbone angle distributions narrowed after adsorption. A more in-depth discussion on the effects of adsorption on peptide conformation can be found in the supplementary. Entropic analysis of the adsorption process showed a reduction in backbone conformational entropy of 23.9 $J\ mol^{-1}\ K^{-1}$ and 32.7 $J\ mol^{-1}\ K^{-1}$ for WT and M2 respectively.

Moreover, side-chain entropy decreased by 5.1 J mol^{-1} K^{-1} and 16.8 J mol^{-1} K^{-1} for WT and M2 respectively. The larger entropic loss for M2 over WT adsorption suggests there is a less favorable chemical potential difference, $\Delta\mu$, between the solvated and adsorbed states, thus, giving an energetic explanation for previous experimental results showing slower deposition rates for M2.[3]

Interestingly, for WT, Ser7 started as part of the helix and then intermittently formed either a π-helix or loop structure, while the opposite occurred for M2. Ser7 started as a π -helix and became part of the α-helix. Despite the additional α-helical turn causing the whole peptide to twist, resulting in a large RMSD from the starting structure, the overall secondary structure is largely unchanged. Tyr12 often desorbs from the graphene surface resulting in the loss of the 3/10 helix. Trp12, however, stays in contact with the graphene surface for most of the simulation. This is likely a result of the two-membered indole ring having greater adhesive interactions over the phenol of Tyr.

The same similarity metric, M, was utilized to quantify conformational changes after adsorption to graphene. The (Φ, Ψ) and χ population distributions of the graphene constrained peptides were compared to the population distributions of the corresponding stable ensemble of structures to not bias the similarity. The WT (Φ, Ψ) population distributions showed that most residues were between 60 % and 80 % similar with decreasing similarity going from N to C-terminus. Ser11 was the most different with a similarity value of just under 0.4 resulting from constraints of hydrogen bonding and neighboring pinned residues. For M2, the similarity was generally lower with multiple residues having values close to or below 0.6. Inspection of the population distributions showed that both solution and surface distributions for M2 sampled the same regions of (Φ, Ψ) space but the constrained peptide was more concentrated at a certain location. Similarity analysis of $\chi1$ and $\chi2$ population distributions showed that the side chains sampled much of the same conformational space with only the residues in contact with the graphene having M values below 0.6 (Supplementary Figure 5).

References:

1. Scouras, A. D.; Daggett, V., The Dynameomics rotamer library: amino acid side chain conformations and dynamics from comprehensive molecular dynamics simulations in water. *Protein Science* **2011,** *20* (2), 341-352.

2. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004,** *25* (13), 1605-1612.

3. So, C. R.; Hayamizu, Y.; Yazici, H.; Gresswell, C.; Khatayevich, D.; Tamerler, C.; Sarikaya, M., Controlling Self-Assembly of Engineered Peptides on Graphite by Rational Mutation. *ACS Nano* **2012,** *6* (2), 1648-1656.