Electronic Supplementary Information For:

Material Composition and Peptide Sequence Affects Biomolecule Affinity to and Selectivity for *h*-Boron Nitride and Graphene

Nermina Brljak,[a] Atul D. Parab,[a] Rahul Rao,[b] Joseph M. Slocik,[b] Rajesh R. Naik,[b] Marc R. Knecht,[a,c,†,] Tiffany R. Walsh [d,†,*]*

[a]Department of Chemistry, University of Miami, 1301 Memorial Drive, Coral Gables, Florida 33146, United States,

[b]Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio, 45433, United States

[c]Dr. J.T. Macdonald Foundation Biomedical Nanotechnology Institute, University of Miami, UM Life Science Technology Building, 1951 NW 7th Ave, Suite 475, Miami, Florida, 33136, United States

[d]Institute for Frontier Materials, Deakin University, Geelong, Victoria 3216, Australia

(†: These authors contributed equally)

*To whom correspondence should be addressed: MRK: Phone: (305) 284-9351, email: knecht@miami.edu; TRW: Phone: +61 (35) 227-3116, email: tiffany.walsh@deakin.edu.au

Computational Methodology

Simulation Details

System Setup:
For the surface-adsorbed systems, each system simulated consisted of a single peptide chain, two single-layer graphene or *h*-BN substrates, ~3200 water molecules and a sufficient number of Cl^- counter-ions to ensure the overall charge neutrality of the system. The overall cell dimensions were ~49×43×80 $Å^3$, and ~50×43×80 $Å^3$ for graphene and *h*-BN respectively. The two 2D sheets were placed ~48 Å apart along the direction perpendicular to the surface plane, and the intervening gap was filled with liquid water. The number of waters was adjusted to ensure the density of water in the centre of this sandwich configuration was equivalent to the density of liquid water at 300 K and 1 atm.

For the peptide "in-solution" simulations (in the absence of the 2D material), the simulation comprised a single peptide chain, ~7000 water molecules and sufficient counter-ions to ensure overall charge neutrality of the cell. The simulation cell had dimensions of approximately 60 × 60 × 60 $Å^3$.

All three peptide chains were modeled in the zwitterionic form, i.e. the N- and C- termini of the peptide were uncapped (*i.e.* modeled as NH_3^+ and COO^-, respectively), consistent with the peptides used experimentally.

REST-MD Simulation Details:
All simulations were performed using Gromacs 2018.[1] The GRAPPA force-field and BoNi-CHARMM force-field[2,3] were used to model the interactions with the graphene and *h*-BN surfaces respectively. The peptides were modeled using the CHARMM22* force-field,[4,5] and the modified version of TIP3P water[6,7] was employed. All simulations were performed in the canonical (*NVT*) ensemble, at 300 K with temperature regulated by a Nose-Hoover[8,9] thermostat. Note that all replicas were run at the same thermal temperature of 300 K. The LJ non-bonded interactions were tapered to zero between 10 and 11 Å, whereas the electrostatic interactions were treated using a particle mesh Ewald (PME)[10] summation with a real-space cutoff of 11 Å. A time step of 1 fs was used, each system simulated for 20×10^6 MD steps and exchanges between adjacent replicas attempted every 1000 steps (≡ 1ps). Coordinates were saved every 1000 steps.

The REST approach is a version of Hamiltonian replica exchange molecular dynamics (H_REMD) that provides advanced sampling of the conformational space of a system.[11-13] For fuller details of the technique, its implementation in GROMACS and investigation of the improvements sampling efficiency, we refer readers to Terakawa *et al.*, Wright *et al.* and Tang *et al.*[11-13]

In the present study the 'effective temperature' window spanned 300-430 K across sixteen replicas. The λ (scaling) values used for the replicas were 0.000, 0.057, 0.114, 0.177, 0.240, 0.310, 0.382, 0.458, 0.528, 0.597, 0.692, 0.750, 0.803, 0.855, 0.930, and 1.000. The initial configuration

of the peptide in each replica differed, covering a range of different secondary structure motifs, e.g. α–helix, β-turn, PPII helix and random coil.

Simulation Analysis:

The degree of residue-surface contact was determined by calculating the fraction of the total REST-MD reference trajectory (*i.e.* the trajectory that corresponds to the unscaled Hamiltonian, λ=0.000) that a reference site on a residue was less than or equal to a cutoff distance of the 2D surface. The reference sites and cutoff-distances (for graphene) were the same as those used in previous work.[14] For the *h*-BN interface, we used the same procedure, based on residue-surface contact information, as published in earlier studies[15] to establish appropriate cutoffs for all residue types present in the three peptides. The resultant distance-based cut-offs were evaluated to be the same as those already published for the peptide/graphene interface.[14]

To determine the Boltzmann-weighted ensemble of peptide conformations, the Daura clustering algorithm,[16] with a 2 Å cutoff for the atoms in the peptide backbone was employed. The clustering analysis was performed over the full 20,000 frames of the reference trajectory. The percentage population of each distinct structure was determined from the fraction of the total 20,000 frames that were assigned to that distinct structure.

Peptide secondary structure was examined using the STRIDE algorithm.[17] In each case, each frame in the relevant 20,000 frame trajectory was analysed, via STRIDE assignment of secondary structure motif on a residue-by-residue basis. The total number of instances of each secondary structure assignment was counted and the overall percentage fraction of that motif was calculated.

Convergence of the REST-MD simulations can be monitored by different metrics, such as the number of peptide backbone clusters as a function of REST-MD simulation steps, and the degree of swapping between replicas as a function of REST-MD simulation steps (referred to herein and in previous work as replica mobilities). Representative data of each type of analysis are shown in Figures S1 and S2, indicating satisfactory performance.

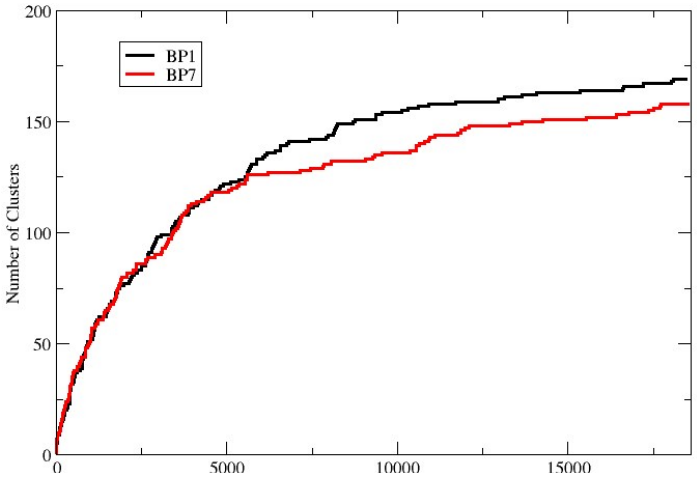

Figure S1. Number of peptide backbone clusters as a function of REST-MD simulation progress, for adsorption of BP1 and BP7 at the aqueous graphene interface.

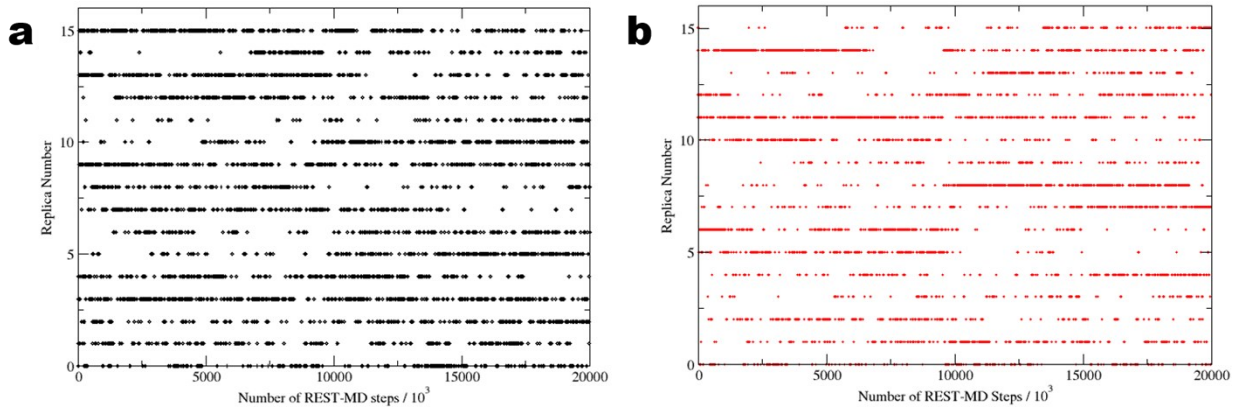

Figure S2. Representative plots of replica mobility (exchange between replicas) for BP1 adsorbed at the aqueous *h*-BN interface for a) replica 0, and b) replica 15.

Experimental Methodology

Materials:

Sodium dodecyl sulfate was purchased from Sigma and ammonium persulfate was acquired from BDH Chemicals. Thermal release tape was sourced from Semiconductor Equipment Corporation. Single layer graphene and *h*-BN on copper foil were obtained from Grolltex. All peptides (> 95% purity) were commercially synthesized from Genscript. The Au QCM sensors used were purchased from Biolin Scientific. Milli Q water (18.2 MΩ·cm) was used for all the experiments. All chemicals were used without additional purification.

Methods:

All QCM measurements were conducted using a Q-Sense E4 instrument from Biolin Scientific. Single layer graphene or *h*-BN grown on a Cu foil were first transferred onto thermal release tape carefully to ensure there were no air bubbles. After an ozone-UV treatment of 5 min, they were etched with $(NH_4)_2S_2O_8$ (100 mg/mL in Water) for 3 h, followed by rinsing with water and drying with N_2 gas. The graphene or *h*-BN thermal tape was adhered to the Au QCM sensors at room temperature and released by heating to 125 °C using a hotplate. XPS spectra of the deposited *h*-BN were collected on an M-Probe Surface Science XPS spectrometer. Spectra were collected using a spot size of 800 mm, 0.01 eV step, and averaged over 50 scans. Prior to each QCM measurement, the sensors were cleaned with a 5 min ozone-UV treatment. This was followed by washing with water for 30 min, 4 wt% sodium dodecyl sulfate (SDS) for 30 min, and water again for 30 min, all at a flow rate of 150 µL/min. Measurements started with a flow of water for 10 min followed by 30 min of the peptide solution at varying concentrations (2.5-20 µg/mL) at 22.5 °C. Using the Sauerbrey equation, the change in frequency is directly related to the mass of peptide adhered to the surface which enables the rate parameters to be determined.

Peptide Conformational Analysis: Free in Solution

For each of the sequences, BP1, BP7, and P1, the peptide secondary structure was investigated in the absence of the 2D material surface (i.e. free in solution). Circular dichroism (CD) spectroscopy (Figure S3) showed that all three peptides were predominantly random coil in structure. Interestingly, the BP7 and BP1 peptides had a similar spectrum with a signal positioned at ~198 nm, while the P1 peptide displayed a signal that was shifted to ~195 nm.

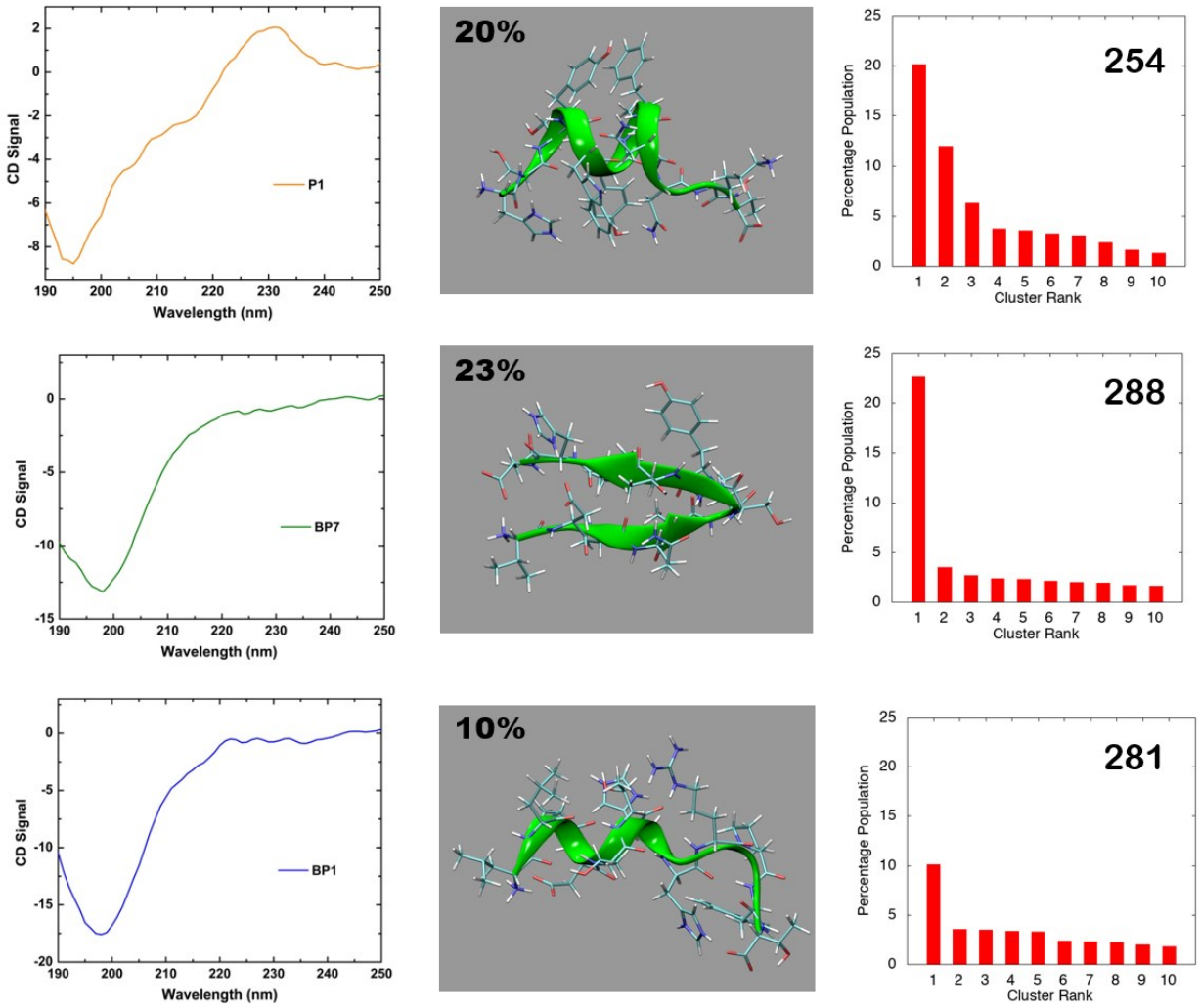

Figure S3. Summary of conformational analyses. Left panel: circular dichroism spectra; middle panel: representative snapshot of the most populated cluster; right panel: population (expressed as a percentage of the total ensemble) of the top-ten most populated clusters, with total number of clusters indicated in the top right corner.

REST-MD simulations of the peptides free in solution were analysed for secondary structure using both clustering analyses (Figure S1) and also the STRIDE algorithm (Table S1). In the clustering analysis the ensemble is classified into groups of similar structures. The outcomes of this analysis provides information in terms of 1) the overall number of similar structures (denoted here as clusters), and 2) their relative population in the ensemble. The most populated clusters for

all peptides in solution are presented in Figure S1, along with the populations of the top-ten most populated clusters. In each case, the population of the most populated cluster is not dominant (*i.e.* under 50%), and each peptide supports a high number of clusters (254-288). The most populated cluster in each case features secondary structure, with P1 supporting 20% as partially-helical structure, BP7 supporting 23% as a beta hairpin, and BP1 supporting 10% as a distorted helix. However, the most populated cluster structures are not representative of the entire ensemble, given that the remaining clusters (250+ clusters) were chiefly unstructured, corresponding with a random coil assignment. Despite the low population fraction of each such random-coil cluster structure, the sheer number of these overwhelm the contribution of the most populated cluster, thus resulting in an overall random-coil character for each peptide conformational ensemble.

The STRIDE analyses complemented these clustering data, and again indicated that random coil characteristics dominated the conformational ensemble in each case, followed by β-turn contributions (Table S1). Nonetheless, each sequence exhibited small differences in the proportions of the non-random-coil contributions. Specifically, P1 featured the highest fraction of helical content (12%) of the three sequences, whereas BP7 was remarkable for the relatively high content of β-sheet (as opposed to β-turn) motif. This finding is consistent with the reasonably similar appearance of the experimentally-determined CD spectra for each of the three peptides.

Table S1. Percentage contribution of secondary structure motifs determined from STRIDE analyses of the "in-solution" molecular dynamics simulation trajectories. RC corresponds to random coil.

Motif	BP1 / %	BP7 / %	P1 / %
RC	54	57	49
Alpha helix	6	1	13
3-10 helix	2	1	0
Beta sheet	1	12	0
Beta turn	37	29	38

Table S2. Percentage contribution of secondary structure motifs determined from STRIDE analyses of the surface-adsorbed molecular dynamics simulation trajectories on the two substrates, for the three peptides. RC corresponds to random coil.

Graphene

Motif	BP1 / %	BP7 / %	P1 / %
RC	66	68	55
Alpha helix	4	1	11
3-10 helix	1	0	0
Beta sheet	0	0	0
Beta turn	29	31	34

h-BN

Motif	BP1 / %	BP7 / %	P1 / %
RC	54	59	48
Alpha helix	7	2	17
3-10 helix	2	2	0
Beta sheet	0	1	0
Beta turn	37	36	35

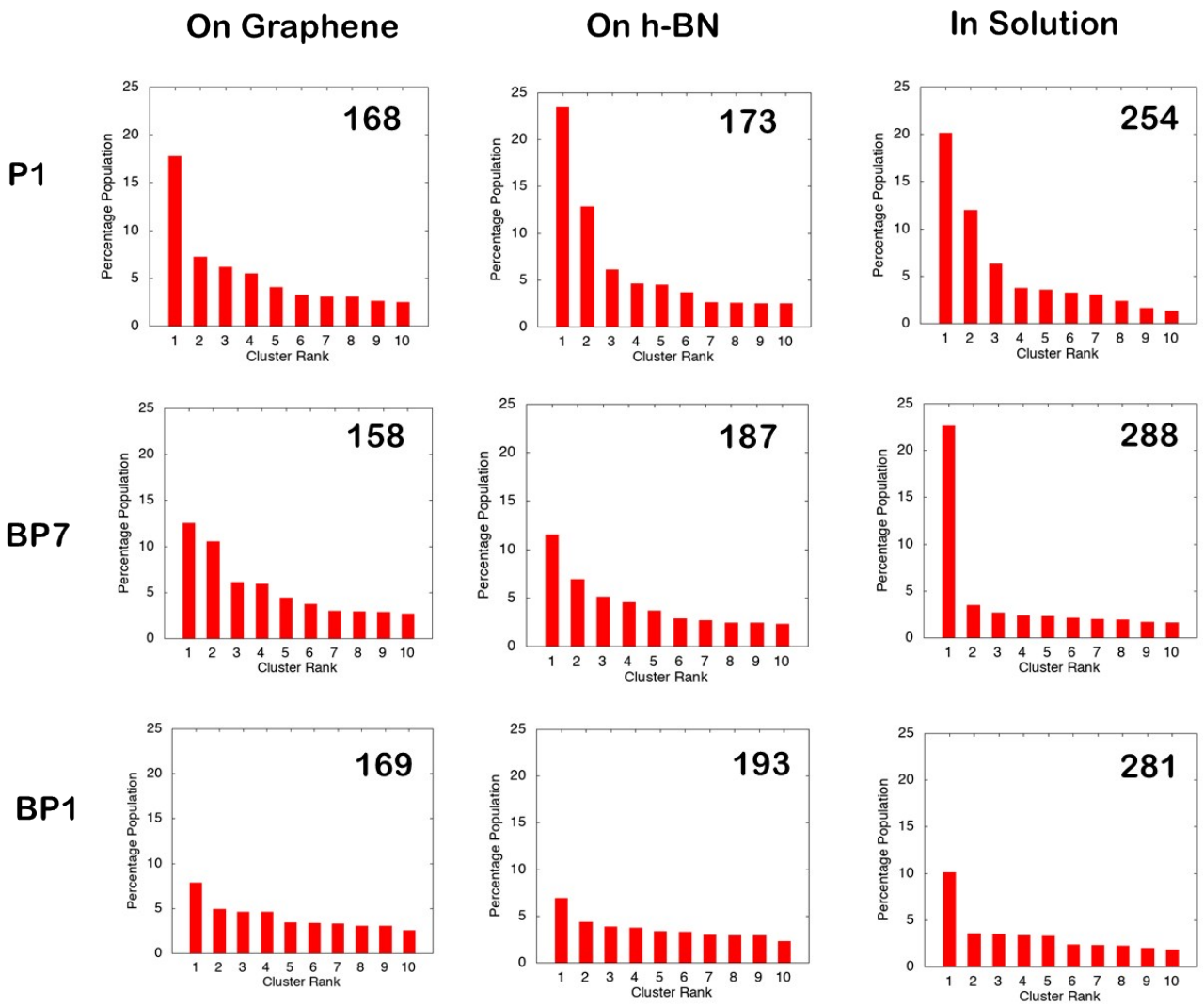

Figure S4. Cluster population distributions for the top-ten most populated clusters, in both the surface-adsorbed and un-adsorbed states, as determined from the molecular dynamics simulations. The total number of clusters is indicated in the inset.

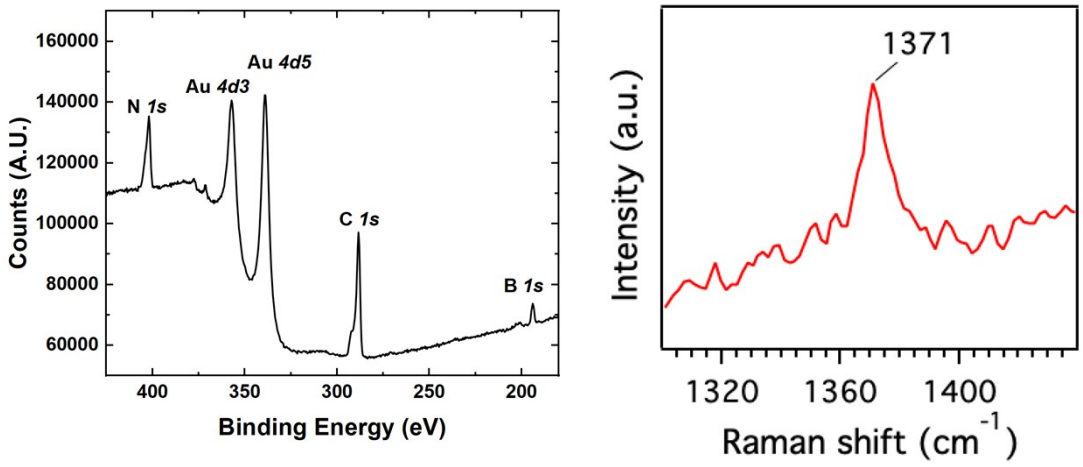

Figure S5. Confirmation of *h*-BN deposition onto the QCM sensor surface. (left panel) XPS spectra of the *h*-BN coated Au QCM sensor. Peaks associated with the Au surface and *h*-BN are evident. The presence of carbon likely arises from adsorbed material that is desorbed and removed prior to QCM analysis. (right panel) Raman spectrum collected (using 633 nm excitation) from *h*-BN on the Au QCM sensor. The sharp peak at 1370 cm^{-1} corresponds to the E_{2g} mode of *h*-BN.

Discussion of Possible Origins of Residue Specific Binding on *h*-BN vs. Graphene

The physical origins of the indicative residue binding specificity on h-BN vs. graphene demonstrated by this work are currently not definitively resolved and represents a substantial challenge to resolve in future. The answer to the question, e.g. "why do bulky aliphatics prefer to bind to h-BN over graphene?", has not yet been established and requires further, integrated investigation from both experimental and modelling perspectives.

One possible (currently speculative) answer is that given that the anticipated leucine/valine–surface interaction (in the absence of water) can only be driven by van der Waals contributions, which would be similar for both *h*-BN and graphene, this points to differences in the interfacial solvent structuring at the graphene vs. *h*-BN interfaces, and suggests that entropic contributions due to displacement (and therefore release) of interface-bound waters might give rise to this specificity effect. For example, earlier work[18] comparing solvent structuring at graphene and *h*-BN interfaces showed an excess of hydrogen-bond acceptors at the *h*-BN interface, which is the reverse of what has been observed for the water/graphene interface. Clearly resolution of this interesting question is outside the scope of the current work and will require significant future effort to resolve, but it is a question that could be addressed in future.

REFERENCES

(S1) Abraham, M.; Murtola, T.; Schulz, R.; Pálla, S.; Smith, J.C.; Hess, B; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. Software X **2015**, *1-2*, 19-25.

(S2) Wright, L. B.; Rodger, P. M.; Corni, S.; Walsh, T. R. GolP-CHARMM: First-Principles Based Force Fields for the Interaction of Proteins with Au(111) and Au(100). *J. Chem. Theory Comput.* **2013**, *9*, 1616-1630.

(S3) Wright, L. B.; Rodger, P. M.; Walsh, T. R.; Corni, S. First-Principles-Based Force Field for the Interaction of Proteins with Au(100)(5 × 1): An Extension of GolP-CHARMM. *J. Phys. Chem. C* **2013**, *117*, 24292-24306.

(S4) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

(S5) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization? *Biophys. J.* **2011**, *100*, L47-L49.

(S6) 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*, 926-935.

(S7) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996**, *105*, 1902-1921.

(S8) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255-268.

(S9) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695-1697.

(S10) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

(S11) Terakawa, T.; Kameda, T.; Takada, S. On easy implementation of a variant of the replica exchange with solute tempering in GROMACS. *J. Comput. Chem.* **2011**, *32*, 1228-1234.

(S12) Wright, L. B.; Walsh, T. R. Efficient conformational sampling of peptides adsorbed onto inorganic surfaces: insights from a quartz binding peptide. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4715-4726.

(S13) Tang, Z.; Palafox-Hernandez, J. P.; Law, W.-C.; Hughes, Z. E.; Swihart, M. T.; Prasad, P. N.; Knecht, M. R.; Walsh, T. R. Biomolecular Recognition Principles for Bionanocombinatorics: An Integrated Approach To Elucidate Enthalpic and Entropic Factors. *ACS Nano* **2013**, *7*, 9632-9646.

(S14) Hughes, Z.E.; Walsh, T. R. What Makes a Good Graphene-Binding Peptide?, *J. Mater. Chem. B*, **2015**, *3*, 3211-3221.

(S15) Z. Tang, J. P. Palafox-Hernandez, W.-C. Law, Z. E. Hughes, M. T. Swihart, P. N. Prasad, M. R. Knecht and T. R. Walsh, *ACS Nano*, 2013, **7**, 9632-9646.

(S16) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem. Int. Ed.* **1999**, *38*, 236-240.

(S17) Andersen, C.A.; Palmer, A.G.; Brunak S.; Rost B. *Structure* **2002** *10*, 175-84.

(S18) Budi, A.; Walsh, T.R. A Bespoke Force Field To Describe Biomolecule Adsorption at the Aqueous Boron Nitride Interface. *Langmuir* **2019**, *35*, 16234-16243.