Soft Matter

Adsorption of Amino Acids on Graphene: Assessment of Current Force Fields

Journal:	*Soft Matter*
Manuscript ID	SM-ART-12-2018-002621.R1
Article Type:	Paper
Date Submitted by the Author:	28-Jan-2019
Complete List of Authors:	Dasetty, Siva; Clemson University, Chemical and Biomolecular Engineering Barrows, John; Medical University of South Carolina - College of Medicine, Biochemistry and Molecular Biology Sarupria, Sapna; Clemson University, Chemical and Biomolecular Engineering

Adsorption of Amino Acids on Graphene: Assessment of Current Force Fields

Siva Dasetty[a], John K. Barrows[b], and Sapna Sarupria[a*]

[a]*Department of Chemical & Biomolecular Engineering, Clemson University, Clemson, SC 29634 and*

[b]*Department of Biochemistry & Molecular Biology, Medical University of South Carolina, Charleston, SC 29425*

Abstract

We compare the free energies of adsorption (ΔA_{ads}) and the structural preferences of amino acids on graphene obtained using the non-polarizable force fields — Amberff99SB-ILDN/TIP3P, CHARMM36/modified-TIP3P, OPLS-AA/M/TIP3P, and Amber03w/TIP4P/2005. The amino acid–graphene interactions are favorable irrespective of the force field. While the magnitudes of ΔA_{ads} differ between the force fields, the relative free energy of adsorption across amino acids is similar for the studied force fields. ΔA_{ads} positively correlates with amino acid–graphene and negatively correlates with graphene–water interaction energies. Using a combination of principal component analysis and density-based clustering technique, we grouped the structures observed in the graphene adsorbed state. The resulting population of clusters, and the conformation in each cluster indicate that the structures of the amino acid in the graphene adsorbed state vary across force fields. The differences in the conformations of amino acids are more severe in the graphene adsorbed state compared to the bulk state for all the force fields. Our findings suggest that the force fields studied will give qualitatively consistent relative strength of adsorption across proteins but different structural preferences in the graphene adsorbed state.

* ssarupr@g.clemson.edu

I Introduction

Graphene is an atom thick two-dimensional carbon nanomaterial with physical and chemical properties suitable for a diverse set of applications, such as defense equipment, cosmetics, textiles, drug delivery and energy storage.[1–6] As the applications of graphene have grown, so have the concerns regarding its biocompatibility.[7–9] While unmodified graphene may be toxic, graphene coated with proteins was found to be less harmful.[10,11] Accordingly, understanding the association of proteins and graphene is important. Detailed knowledge about the formation of protein–graphene complexes can also advance bio-based applications of graphene, such as in the development of effective biosensors, drug-delivery agents, and protein separation processes.[12–14] To enhance our ability to engineer such complexes, an understanding and preferably an ability to predict *a priori* the strength of adsorption and structure of the protein upon adsorption to graphene is essential. To this end, several studies have examined the conformational changes of proteins adsorbed on graphene using a wide array of experimental and computational techniques.[15–20] The emerging picture from these studies is that the adsorption behavior depends on the interplay of protein–water–graphene interactions. This fine balance of interactions has made it challenging to predict the adsorbed state and binding energy of a protein on graphene.

Developing quantitative structure-function relationships for protein–graphene interactions requires extensive information on the structure and the activity of the proteins adsorbed on graphene. While the ability of current wet-laboratory tools to provide high-resolution structural information in interfacial regions is limited[21], molecular simulations are well-suited for such investigations[22]. The success of molecular simulations strongly depends on their ability to describe the interaction details of the system appropriately. Ideally, quantum mechanics (QM) would be the most accurate approach, however, such calculations are currently infeasible for systems with more than a few hundred atoms.[23,24] Alternatively, atomic simulations with accurate description of the potential energy of the system can capture the protein–graphene interactions in aqueous media. A common choice for describing proteins in protein–graphene systems is using the non-polarizable protein force fields, such as Amber[25], CHARMM[26], OPLS[27], and GROMOS[28]. Graphene is described by the Lennard-Jones (LJ) parameters of the aromatic carbon atoms from the respective force fields.[29–35] However, majority of these force fields have been developed using data sets focused on folded proteins

in aqueous media. As a result, the reliability of these force fields to study the interfacial behavior of proteins has been questioned.[36–38] Some attempts to address this include modifying the graphene-water interactions. For example, Werder *et al.*,[39] suggested modified LJ parameters to reproduce water–graphene contact angle. Hughes *et al.*, argued in favor of incorporating the polarizability of graphite/graphene to correctly capture the interfacial water density and orientations.[37] However, there has been no study so far that has systematically investigated the differences in the protein–graphene behavior across the different force fields. It is well-appreciated that the structural properties of proteins in bulk differ with force fields.[40–43] It is therefore possible that the thermodynamic and structural properties of the protein–graphene interactions will differ with force fields, and the conclusions drawn from simulations will be force field dependent. Thus, there is a pressing need to assess the effects of force field choices on protein–graphene interactions in a systematic fashion.

In this manuscript, we study adsorption of all twenty natural amino acids on graphene calculated with four force fields and their compatible water models. Our choice of amino acid–graphene systems is motivated by two key reasons: (i) It is computationally feasible to calculate precise free energies of adsorption of amino acid–graphene systems. (ii) Amino acids are the primary building blocks of proteins, and therefore, a natural place to start. Furthermore, information on amino acid–graphene systems will aid the development of coarse-grained models for protein–graphene interactions using bottom-up approaches.[19,44] Also, amino acid–graphene interactions are themselves of interest for various applications.[45,46] We investigate the differences in the binding preferences of amino acids with the latest versions of the commonly employed force fields, viz., Amber[25,47], CHARMM[26,48], and OPLS[49,50] with their compatible water models. We calculate the free energy of adsorption (ΔA_{ads}) using extensive molecular dynamics (MD) and umbrella sampling simulations for the twenty natural amino acids. We analyze conformational preferences using dihedral distributions and cluster analysis of the configurations of adsorbed amino acid. Our results indicate that the relative binding strength of amino acids on graphene is qualitatively similar across the different force fields. We find that the structures of amino acids in the graphene adsorbed state have notable force field dependence.

II Methods

We investigate the effect of force field on adsorption behavior of amino acids on graphene sheet. Four force fields with their compatible water models are evaluated — Amberff99SB-

ILDN[47] with TIP3P[51], CHARMM36[48] with modified-TIP3P (mTIP3P)[52], OPLS-AA/M[50] with TIP3P[51], and Amber03w[53] with TIP4P/2005[54]. See SI for discussion on the differences in the force fields. We characterize the adsorption behavior using the potential of mean force (PMF) between all twenty natural amino acids and graphene calculated using MD simulations and umbrella sampling method. Umbrella sampling[55] of the resulting 80 systems performed amount to an aggregate of 84 μs of simulation. We evaluate ΔA_{ads} and the enthalpic components contributing to the free energy. Detailed analysis of the backbone dihedral distributions and conformations sampled by the amino acids in the graphene adsorbed state are performed and compared across force fields.

A. System setup and simulation details

Amino acids were capped by acetyl (ACE) and n-methyl amide (NMA) groups at the N- and C- termini, respectively (Fig. 1). The starting structures of all the twenty natural amino acids were constructed with LEaP tool of Amber14.[56] Histidine was modeled in its neutral form. Potential energy of each amino acid was minimized in vacuum by steepest-descent algorithm until the maximum force on any atom in the system was less than 10 kJ/mol/nm or the potential energy of the system converged.

A periodic graphene sheet with a surface area of 3.2×3.4 nm^2 was created using an open source Python script[57] and x_2top tool of GROMACS-4.5.5. No bonded interactions within the sheet were considered since the positions of all graphene atoms were fixed in our simulations. This mimics a defect free graphene surface[58] or a nanotube with a diameter much greater than an extended conformation of an amino acid. LJ parameters of the aromatic carbon atom of the force field representing the amino acids were used to describe the graphene carbon atoms. No partial charges were assigned to the graphene carbon atoms and the intra-graphene LJ interactions were set to zero. Geometric means were used for cross-pair LJ parameters with OPLS force field, and Lorentz-Berthelot combination rules were used for Amber and CHARMM force fields.

The CHARMM36 force field parameters in GROMACS format were downloaded from the MacKerell lab's website.[59] The OPLS-AA/M parameters in CHARMM format were downloaded from the Jorgensen lab website.[60] An in-house Perl script was used to convert OPLS-AA/M force field parameters from CHARMM to GROMACS format. Amber03w force field parameters in GROMACS format were provided by Dr. Jeetain Mittal.[53]

The minimized amino acid configuration and generated graphene sheet were combined

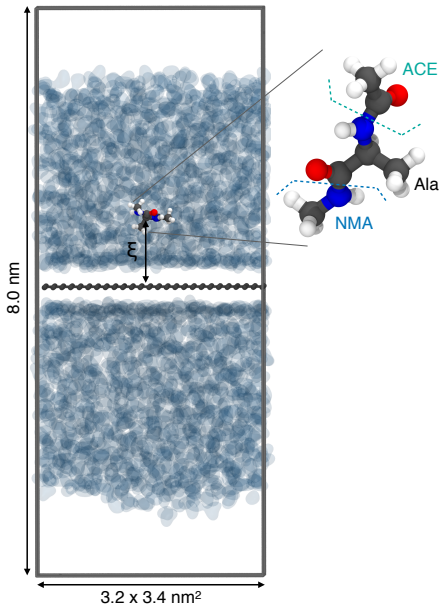

FIG. 1. Illustration of the system set-up in umbrella sampling simulations. The amino acid with the capping groups and the graphene sheet (gray) are represented by ball and stick model. Water is shown in the background (blue).

with water molecules to give the simulation system shown in Fig. 1. The system comprised of ∼1935 water molecules. A vapor region was incorporated in our simulation system to avoid the use of barostat with fixed atoms (graphene) that can lead to artifacts in the simulations.[61] The vapor region enabled us to maintain constant pressure (0 atm) during the simulations. Simulations were performed in the NVT ensemble. Systems with a net non-zero charge were neutralized using sodium or chloride counterions. All systems were subjected to energy minimization with steepest-descent algorithm. The bonds between hydrogen and the heavy atoms were constrained using LINCS.[62] The systems were equilibrated at 300 K for 1 ns. A time step of 2 fs was used for all MD simulations. During equilibration, velocity-rescaling thermostat[63] with separate coupling groups for the amino acid, solvent and graphene sheet were used, with time constants of 3 ps, 1 ps and 1 ps, respectively. Final production simulations were performed at 300 K for 10 ns. We employed Nosé-Hoover thermostat[64,65] during the production runs with the same temperature coupling groups as for the equilibration simulations. For each system, the non-bonded cutoff distances used during parameterization of the respective force field were employed – 1.0 nm in Amber99SB-ILDN and Amber03w force fields, 1.2 nm with force switch scheme starting at 1.0 nm for

CHARMM36 force field, and 1.1 nm with potential switch scheme starting at 0.9 nm for OPLS/AA/M force field. The long range electrostatic interactions were calculated using particle mesh Ewald technique.[66] Neighbor lists were updated with the Verlet scheme.[67]

B. PMF and ΔA_{ads} calculations

Umbrella sampling was used to determine the PMF between the amino acids and graphene. PMF was calculated along the distance normal to the graphene surface and the center of mass (COM) of the amino acids (ξ). We used a force constant (k) of 8000 kJ/mol/nm^2 for 0.4 nm $\leq \xi \leq$ 0.8 nm and k = 4000 kJ/mol/nm^2 for 0.9 nm $\leq \xi \leq$ 2.0 nm. Spacing of 0.05 nm for 0.4 nm $< \xi <$ 0.8 and 0.1 nm for 0.9 nm $< \xi <$ 2.0 nm between windows provided good overlap between the distribution of neighboring windows. Weighted histogram analysis method was used for constructing the PMF from the distributions of ξ in each window.[68]

For each window, amino acid was placed at the reference ξ value and simulations were performed for 10 ns each. ξ values were saved every 1 ps generating 10000 data points for analysis in each umbrella sampling window. Simulations of every system were performed five times with different starting velocities to estimate the error bars. The reported averages and error bars were obtained by Bayesian bootstrapping method discussed in Hub et al.[69] In total, we generated 100 bootstrapped histograms to construct 100 PMFs using the five independent runs as sample data. To check for system size effects, we performed umbrella sampling simulations of Gly with Amber99SB-ILDN/TIP3P in a box of volume 4.26×4.18×12 nm^3 with 4462 water molecules. No system size effects were observed (SI Fig. S1).

ΔA_{ads} was calculated by integrating the PMFs obtained from umbrella sampling. We defined the adsorbed state using the cumulative distribution function (CDF) of ξ obtained in the unbiased simulations. Specifically, the adsorbed state contains all the configurations with ξ up to CDF of <= 0.95. Other definitions depending on the mean force and a hard cutoff of ξ = 0.5 nm were also tested. We found that the estimated ΔA_{ads} to be statistically similar in all the definitions. The simulation set up and parameters in the unbiased simulations are the same as those described in the previous section. Further details of the ΔA_{ads} calculations and unbiased simulations are provided in the SI. Enthalpic contributions to ΔA_{ads} were calculated using unbiased simulations of amino acid–graphene–water systems. The configurations obtained from the unbiased simulations were used to calculate the dihedral distributions and for the cluster analysis of the configurations sampled by the amino

acids in the graphene adsorbed state. Additionally, simulations of amino acids in bulk water (i.e., with no graphene sheet present) were performed to compare bulk and adsorbed states.

III Results and Discussion

A. PMF profiles of twenty amino acids with four force fields

Fig. 2 shows PMF (W(ξ)) profiles obtained with the tested force fields for all twenty amino acids. The amino acid–graphene interactions are favorable irrespective of amino acid and force field. The qualitative nature of PMFs is similar for all amino acids across the force fields studied here, and are consistent with previous studies.[18,70–75] To compare differences in the PMFs, we categorize local features of the PMF profile into i) the largest value of ξ at which amino acid–graphene interactions are favorable (ξ^*), ii) gradient of PMF with respect to ξ (∇W(ξ)), and iii) magnitude and curvature of PMF near the local and global minima. These characteristics are highlighted in Fig. 2(c) for Val.

The three local features of the PMF profiles vary with both amino acids and force fields. A barrier is observed between $\sim$0.8 nm$< \xi <$1.0 nm for several amino acids in case of Amber03w/TIP4P/2005. This contrasts with the other force fields that do not show such a feature. To probe the origins of this barrier we calculated the PMF of Val and Phe by replacing the water model with TIP3P. We used both Amber03w/TIP3P and Amber03*/TIP3P since the latter are a compatible pair (Fig. S2). Interestingly, we found that the barrier-like feature disappeared when the water model was replaced to TIP3P.

The differences in ξ^* are negligible between the force fields for all amino acids. Interestingly, there are noteworthy differences in ξ^* values across amino acids for any given force field. For example, the smallest amino acid Gly has a higher ξ^* value than some of the non-polar amino acids with larger side chains. With Gly, the higher ξ^* is possible only if it maintains an extended conformation. The ξ^* value is highest for Trp, Phe, Tyr, and Arg, suggesting peptides containing these amino acids can have favorable interactions with graphene at larger distances than other amino acids, if we ignore peptide sequence and structural effects.

For 0.6 nm$< \xi <$0.8 nm region, the PMFs for all force fields are similar except that OPLS-AA/M/TIP3P gives slightly lower values of the PMF. Differences in ∇W(ξ) in the 0.6 nm$< \xi <$0.8 nm region across amino acids exist for all force fields. In case of some amino acids such as Val, Ile, Pro, Gln Arg, and Lys, ∇W(ξ) is positive indicating that these amino acids are strongly attracted to graphene. In other cases, ∇W(ξ) is zero or moderately

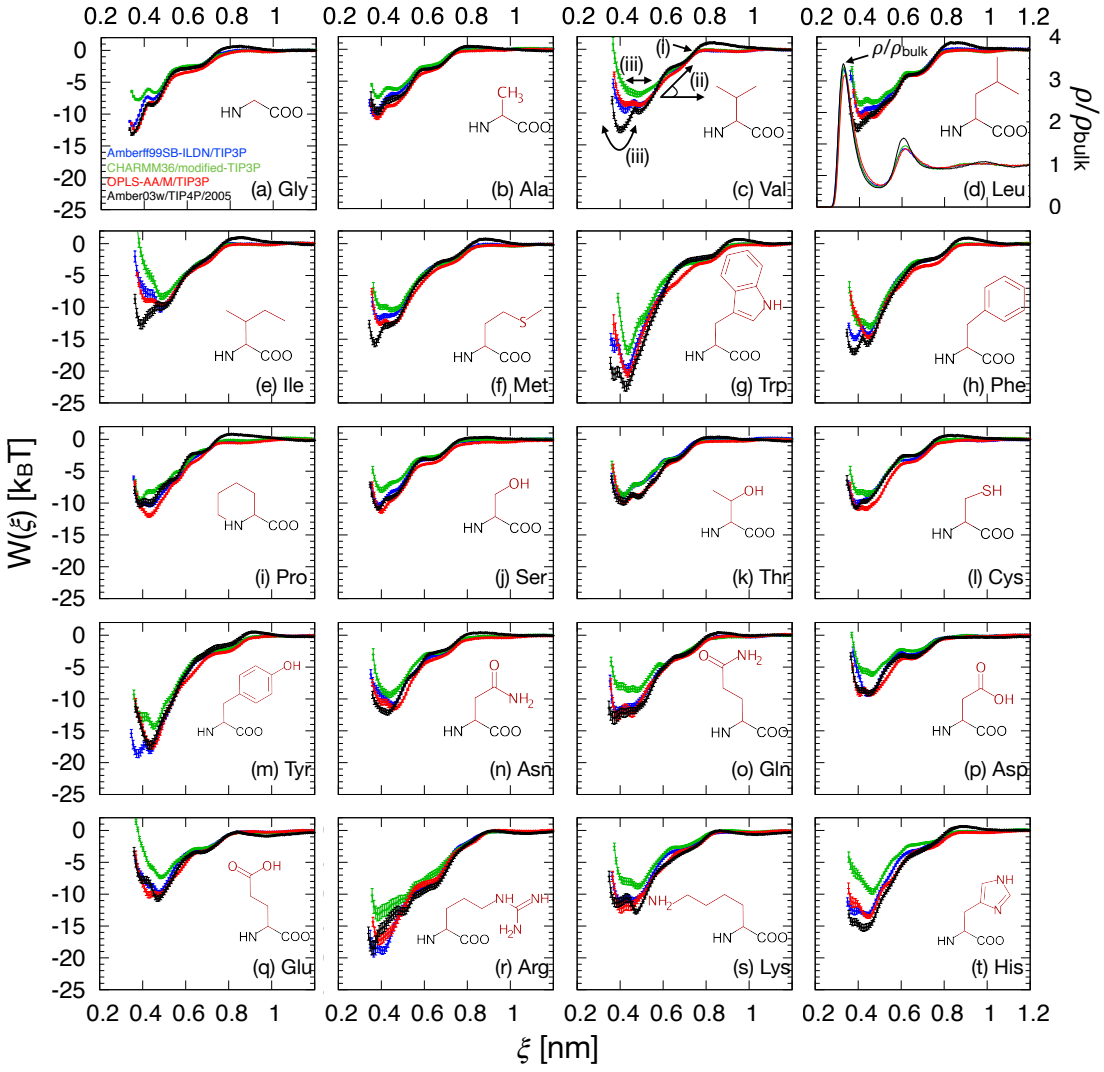

FIG. 2. PMF (W(ξ)) between amino acid and graphene at 300 K. Color code: Blue: Amberff99SB-ILDN/TIP3P, Green: CHARMM36/mTIP3P, Red: OPLS-AA/M/TIP3P, Black: Amber03w/TIP4P/2005. ξ is the normal distance between center of mass of the amino acid and graphene surface. Bootstrapped average profiles are shown. The error bars indicate standard errors on the average of the bootstrapped PMFs. Inset of each panel shows the chemical structure of the amino acid. The terminal capping groups are not shown for clarity. Density of the water oxygen (ρ) along ξ normalized by the average bulk (ρ_{bulk}) density of the respective water model is shown in Panel (d). The arrows in Panel (c) represent the three features described in the main text. All PMFs are shifted to be equal to zero at $\xi = 1.8$ nm. Marvin sketch[76] was used to draw the chemical structures.

negative with respect to ξ until ξ ~0.6 nm. The 0.6 nm distance coincides with the location of the second peak of the water density (second water layer) along ξ (Fig. 2d). Penna *et al.*[77] suggested that the favorable interactions between the second layer of water and amino acids led to longer residence time of the amino acids at distances between 0.45-0.75 nm in studying graphite binding peptide adsorption on graphite. It is possible that a similar phenomenon results in the zero or moderately negative slopes in $\nabla W(\xi)$ around 0.6 nm observed here.

In most cases, the PMFs obtained with OPLS-AA/M/TIP3P and the two Amber force fields in the ξ <0.6 nm region are statistically similar. The net amino acid–graphene interactions are comparatively less favorable with CHARMM36/mTIP3P. Despite the same bonded (excluding torsional) and LJ parameters of amino acid and graphene in Amberff99SB-ILDN/TIP3P and Amber03w/TIP4P/2005, the PMF for ξ <0.6 nm is lower in magnitude with the latter for Val, Leu, Ile, Met, Trp, Phe, Asn, Lys, and His. Similarly, graphene in OPLS-AA/M/TIP3P and CHARMM36/mTIP3P has the same LJ parameters. Yet, the PMF obtained with CHARMM36/mTIP3P for ξ <0.6 nm is lower in magnitude than OPLS-AA/M/TIP3P. Another difference observed in the PMFs relate to observation of double wells in certain cases. In general, Amber03w/TIP4P/2005 has higher tendency of resulting in double wells. As expected, for any given force field, PMFs across the various amino acids show differences.

B. Free energy of adsorption of amino acids

PMF was integrated to calculate the free energy of adsorption, ΔA_{ads} (Fig. 3). The adsorbed state was defined using the CDF of ξ obtained in the unbiased simulations. All the configurations with ξ up to a CDF of <= 0.95 are considered to be in adsorbed state. Details of ΔA_{ads} calculations, and the values and standard errors of ΔA_{ads} are provided in the SI. Consistent with the PMF results, ΔA_{ads} is similar between the two Amber force fields and OPLS-AA/M/TIP3P. CHARMM36/mTIP3P results in relatively higher ΔA_{ads} values for all the amino acids. We also compare the ΔA_{ads} values calculated here with those reported by Hughes and Walsh[70]. They used CHARMM22*[78] for amino acid, polarizable force field, GRAPPA for graphene[37,79] and mTIP3P for water. The resulting ΔA_{ads} values suggest less favorable amino acid–graphene interactions compared to CHARMM36/mTIP3P. Poblete et al.,[72,75] recently compared the performance of CHARMM36/TIP3P (not mTIP3P) and GRAPPA in predicting equilibrium binding constants of small molecules with diverse functional groups. They found that the agreement between simulation and experimental

data was stronger for CHARMM36/TIP3P compared to GRAPPA. This indicates that more rigorous parameterization is required for incorporating polarization effects in graphene.

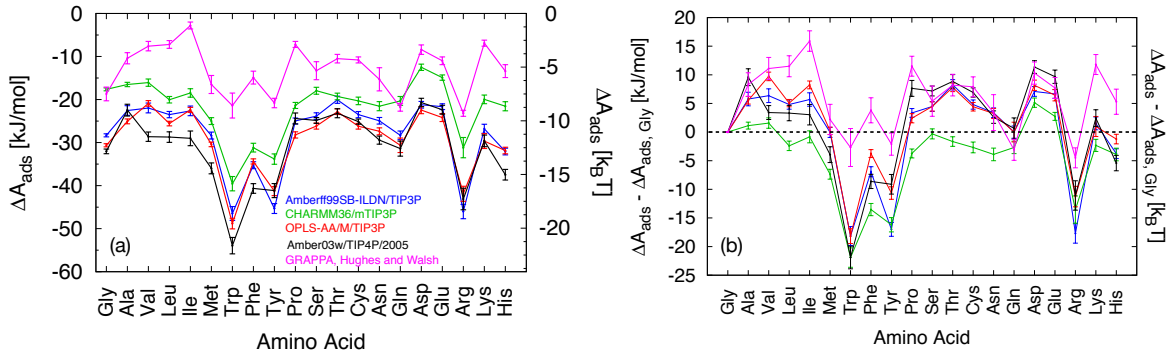

FIG. 3. (a) ΔA_{ads} of amino acids obtained with Amberff99SB-ILDN/TIP3P (blue), OPLS-AA/M/TIP3P (red), CHARMM36/mTIP3P (green), and Amber03w/TIP4P/2005 (black) is shown. In each case, the averages and standard errors on the average obtained using the bootstrapped PMFs are shown. ΔA_{ads} of each amino acid on graphene reported by Hughes and Walsh[70] using GRAPPA and CHARMM22*/mTIP3P is shown.[78] (b) ΔA_{ads} of amino acids relative to that of glycine. The amino acids on the x-axis are arranged according to their side chain chemistry: non-polar (Gly, Ala, Val, Leu, Ile, Met), aromatic (Trp, Phe, Tyr), polar (Pro, Ser, Thr, Cys, Asn, Gln), and charged (Asp, Glu, Arg, Lys, His).

The force fields show quantitative differences in the estimates of ΔA_{ads}. The relative strength of binding between different amino acids is similar for the two Amber force fields and OPLS-AA/M/TIP3P. This becomes evident when we compare ΔA_{ads} across different amino acids relative to Gly (Fig. 3b). Such similarity in the trends indicates an inherent convergence of the force fields in describing the interactions of each amino acid with its surrounding environment. Nevertheless, the possibility of propagation of the minor variances in ΔA_{ads} between force fields into major differences in describing the interactions of proteins with graphene cannot be disregarded. This is true especially because ΔA_{ads} of larger proteins is generally not the simple sum of the ΔA_{ads} of individual amino acids.[70,80]

The free energies of adsorption of amino acids on graphene are not reported by wet-laboratory experiments to date for validating any of the studied force fields. Indirect validations can be performed by comparing the binding preferences of amino acids on graphene with the corresponding observations from *ab initio*, simulations, and wet-laboratory studies.[45,71,81–85]

The binding preferences of the amino acids on graphene estimated in our study and that reported in the literature are summarized in Fig. 4.

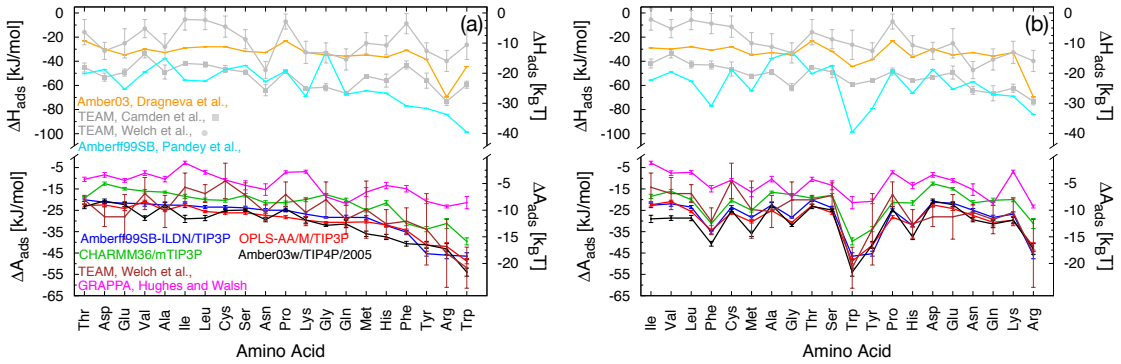

FIG. 4. Interaction energy (ΔE_{ads})/enthalpy (ΔH_{ads})/free energy of adsorption (ΔA_{ads}) reported in the literature and obtained in our study. In panel (a) ΔA_{ads} calculated with Amberff99SB-ILDN is used for sorting the amino acids on the x-axis. In panel (b), the amino acids are arranged according to KD hydrophobic scale values[86]. Color code of ΔA_{ads} is same as in Fig. 3. ΔE_{ads} or ΔH_{ads} calculated using unbiased MD simulations with various force fields reported in Camden *et al.*,[81] (298 K) (gray, square points), Welch *et al.*,[71] (298 K) (gray, circle points), Dragneva *et al.*,[87] (310 K) (orange), and Pandey *et al.*,[44] (300 K) (cyan) are shown. Standard errors on ΔE_{ads} are not reported in Dragneva *et al.*[87] and Pandey et al.[44]. In Camden et al.[81], the amino acids are capped with glycine.

Trp, Arg, Tyr, and Phe are among the most favorable amino acids that interact with graphene according to all the non-polarizable force fields. These amino acids were identified as the strongest binders by Pandey *et al.*[44] using Amber99SB/TIP3P and Dragneva *et al.*[87] (except Phe) using Amber03/TIP3P. Both studies used ΔE_{ads} to assess amino acid–graphene interaction strength. This agreement between the trend of ΔE_{ads} with amino acid reported by Pandey et al.,[44] and our ΔA_{ads} results point to the role of direct interactions between amino acids and graphene in adsorption and is discussed in Section III C.

Berry and coworkers[71,81] reported the ΔH_{ads} and ΔG_{ads} (Gibbs free energy of adsorption) for all twenty amino acids adsorbing on trilayer graphite surface. The amino acids were capped with Gly residues on both N- and C- termini (zwitterionic form), compared to the ACE and NMA groups used in our study. Non-polarizable force field (TEAM) based on automated assignment of force field parameters from fragments of molecules was used. The

partial charges of water reported by the authors are similar to that of SPC/E[88]. Despite this different approach, the ΔH_{ads}[81] indicates Arg and Trp to be among the strongest binders. The other strong binders include Asn, Lys, Gln, and Gly. The ΔG_{ads} values reported by Welch et al.,[71] and ΔA_{ads} obtained in our study are statistically similar.

ΔA_{ads} obtained with GRAPPA indicate Gly, Trp, Gln and Arg as most favorable graphene binders. This contrasts the results from non-polarizable force fields where Gly and Gln are not strong binders. In addition, according to GRAPPA the interaction between Phe–graphene is less favorable than Gly–graphene unlike non-polarizable force fields. Ile, Pro, and Lys are found to form weak amino acid–graphene complexes with GRAPPA consistent with our results. While there is no direct experimental data, the relative affinity of different aromatic amino acids was reported in Mallineni *et al.*[84]. In their cyclic voltammetry and photoluminescence experiments, all the aromatic amino acids were found to strongly interact with graphene. Similarly, several density functional theory based studies found that aromatic amino acids and Arg strongly bind with graphene compared to Gly, Val, and Asp.[82,83,85] These reported trends in relative binding strength of amino acids are consistent with those observed using the non-polarizable force fields.

Does size and/or chemistry of the amino acid dictate the trends observed in ΔA_{ads}? To analyze this, we compared the variation of $\Delta A_{ads}/\Delta E_{ads}/\Delta H_{ads}$ with different hydrophobic scales, and molecular weights of the amino acids. Fig. 4b shows the resulting changes in $\Delta A_{ads}/\Delta H_{ads}$ with the widely used Kyte-Doolittle (KD) hydrophobic scale.[86] KD scale is based on the free energy of transfer of the amino acid between liquid–water and water–vapor phase, and surface exposure of residues in crystal structures.[86] The hydrophobic nature of the amino acids decreases from Ile to Arg in the plot. Except for the ΔH_{ads} calculated with TEAM force field, the correlations between $\Delta A_{ads}/\Delta E_{ads}$ with other force fields and KD scale are negligible. Similarly, no correlations are observed with the atomic hydrophobic scale[89] (see Fig. S4). ΔA_{ads} calculated in this study moderately correlates with the molecular weight of the amino acids. At first glance, these observations indicate that the size of amino acids has more impact on the formation of amino acid–graphene complex than the amino acid chemistry. To further elucidate the origins in the differences in ΔA_{ads}, we probe the enthalpic contributions to ΔA_{ads}.

C. Correlations between ΔA_{ads} and components of potential energy

We breakdown the enthalpic contributions to ΔA_{ads} in terms of the interaction energy between amino acid, graphene and water. Specifically, we calculate interaction energy (E) as the sum of LJ and electrostatic interactions between amino acid and graphene, graphene and water, water and water, amino acid and water, and intramolecular interaction energy of the amino acid. For the remainder of the discussion we focus primarily on amino acid–graphene ($\Delta E_{ads,AG}$) and graphene–water interaction energy ($\Delta E_{ads,GW}$). The Δ refers to the difference in the adsorbed and bulk states and is computed as the difference in E between the systems with amino acid adsorbed on graphene and amino acid in bulk (i.e. $\Delta E = E_{ads} - E_{bulk}$). We used unbiased simulations and the window with $\xi = 2.0$ nm of the umbrella sampling simulations to calculate E_{ads} and E_{bulk}, respectively. The reported averages and error bars are the means and the standard errors obtained by using Bayesian bootstrapping[90,91] technique on the ΔE obtained from five independent runs.

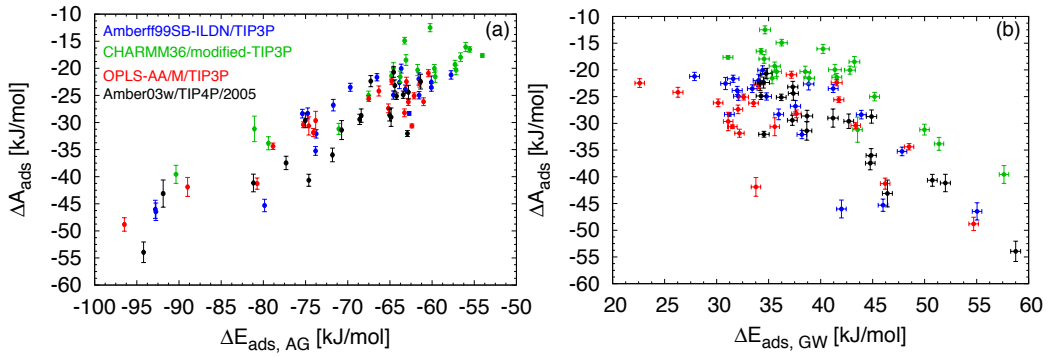

FIG. 5. Correlations between free energy of adsorption and (a) change in amino acid–graphene interaction energies, $\Delta E_{ads,AG}$, and (b) change in graphene–water interaction energies, $\Delta E_{ads,GW}$.

Fig. 5 shows the change in ΔA_{ads} with $\Delta E_{ads,AG}$ and $\Delta E_{ads,GW}$ calculated with the four force fields. ΔA_{ads} has a strong positive correlation with $\Delta E_{ads,AG}$ and a moderate negative correlation with $\Delta E_{ads,GW}$. None of the other potential energy components display a correlation with ΔA_{ads} (SI Fig. S5). The strong correlation between ΔA_{ads} and $\Delta E_{ads,AG}$ observed here is also consistent with the observations of Zerze et al.[92]. They found that the protein–surface interactions predominantly dictate the behavior of the protein (Trp-Cage and GB1 hairpin) on the surface compared to the hydrophobicity of the surface[93]. This correlation between $\Delta E_{ads,AG}$ and ΔA_{ads} explains the correlation observed between ΔA_{ads}

and molecular weight of the amino acid. These observations suggest that the intermolecular interactions of the exposed residues to graphene can be used as a zeroth-order approximation to predict the binding affinities of proteins to graphene.

D. Differences in conformations of amino acids with force fields

It is well known that the conformational preferences of proteins in bulk are different among the widely employed force fields.[40–43,94,95] A well-appreciated problem with the earlier versions of Amber, CHARMM, and OPLS is a bias in the structure towards either α-helix or β-sheets.[95–98] Therefore, these force fields have been revised with updated torsional parameters based on quantum simulations and experimental data (details are provided in SI Section 2). The revision protocols primarily differ in the reference backbone space used for scanning the side chain dihedral space, and the weight factors used to minimize the potential energies determined by molecular and quantum mechanics simulations.[47,50,99] The backbone dihedral angle parameters (Fourier coefficients and angles) of all amino acids other than Gly and Pro are generally determined by using the ϕ-ψ space of alanine dipeptide. For this reason, the ϕ-ψ distributions of all amino acids except Gly and Pro resemble Ala ϕ-ψ distributions, and are referred to as Ala-like amino acids. The Gly and Pro parameters are optimized separately because of their unique side chains – Gly has no C_β and Pro side chain is bonded to the backbone. It is therefore, pertinent to investigate the differences in conformations of adsorbed amino acids and to enquire whether they are simply a manifestation of the differences in bulk water. To this end, we calculated the backbone dihedral angle distributions of each amino acid sampled with different force fields in bulk water and in the graphene adsorbed state.

1. Conformations of amino acids in bulk

Fig. 6a shows the sampled ϕ-ψ space for Val, Thr, Gln, Phe, and Lys in bulk water with the four force fields. We focus on these five amino acids to facilitate the discussion. The sampled ϕ-ψ space for all the amino acids in bulk water with the four force fields are given in SI Fig. S6. For the purpose of the discussion, we follow the nomenclature of ϕ-ψ space proposed by Hollingsworth and Karplus.[100] The distributions can be compared across the amino acids and across the different force fields using two features – (i) the regions of the ϕ-ψ space sampled, and (ii) the population of the conformations in these regions. When comparing across amino acids, the ϕ-ψ space sampled is similar for all Ala-like amino acids however, the population of the conformations sampled in these regions differ with amino

acid since it depends on the side chain (each column in Figs. 6a and S6).

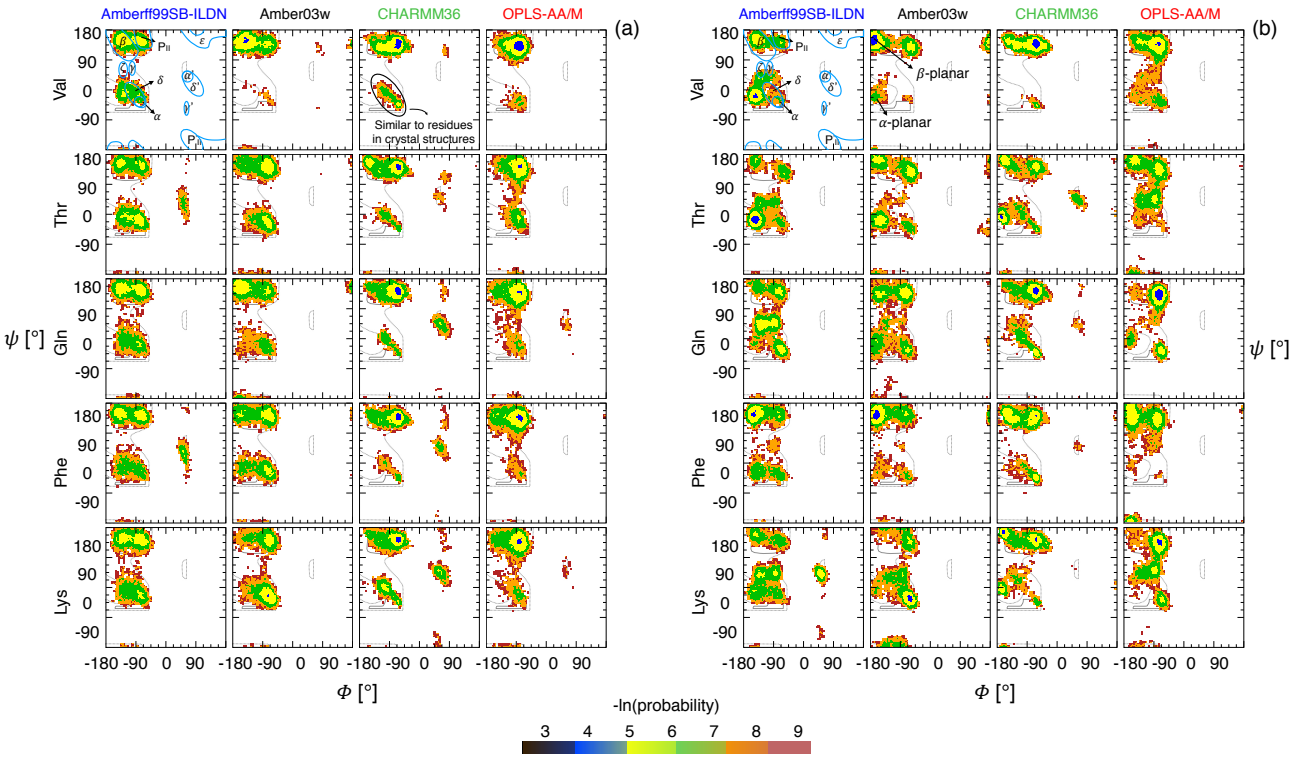

FIG. 6. The distribution of the backbone torsional parameters (ϕ-ψ) obtained for Val, Thr, Gln, Phe, and Lys in (a) bulk water and (b) graphene adsorbed state with different force fields. The background contours (shown in gray) represent the sterically allowed Ramachandran angles.[101] The classification of the ϕ-ψ space proposed by Hollingsworth and Karplus[100] is shown for Val in Amberff99SB-ILDN/TIP3P. Cool to warm colors in the heat map indicate high to low probabilities. We used Poblete *et al.*,[75] nomenclature for referring to the new regions explored in the graphene adsorbed state by amino acids (Val with Amber03w/TIP4P/2005 in (b)).

The conformations of each amino acid including Gly and Pro varies with force fields (each row in Figs. 6 and S6). The regions of ϕ-ψ space sampled with Amberff99SB-ILDN/TIP3P and Amber03w/TIP4P/2005 are similar for each amino acid. However, differences in the probabilities of the conformations can be observed — while the former does not show a preference for any particular region, the latter shows enhanced probability in either α-helical or β-sheet region in case of most amino acids. A major deviation in the regions of ϕ-ψ space sampled is observed for Val, where Amber03w/TIP4P/2005 has much lower sampling near the α-helix and δ-region compared to the other amino acids. This is consistent with

previous observations[98]. Such a contrast between Val and other amino acids is not observed for the other force fields. The differences in the ϕ-ψ distributions are higher between the Amber-based force fields and the other two force fields. The sampling near α-helical region with CHARMM36/mTIP3P is narrower compared to other three force fields. This narrower distribution agrees well with the ϕ-ψ distributions of residues in protein crystal structures[100]. In addition, CHARMM36/mTIP3P shows sampling in $\alpha^{'}$ and $\delta^{'}$ regions which are sparsely, if at all, sampled in other force fields for most amino acids. The OPLS-AA/M/TIP3P shares several similarities in its ϕ-ψ distributions with Amber-based and CHARMM36/mTIP3P force fields. A notable difference is that the bridge regions (ζ, γ) between the α-helical and β-sheets are sampled more in OPLS-AA/M/TIP3P compared to the other force fields. The overall differences across force fields are qualitatively consistent with those reported by Vitalini et al.[98] using μs long simulations. This suggests that our simulations have successfully sampled the key conformational features.

2. Conformations of amino acids in graphene adsorbed state

How do the differences in the dihedral distributions across the force fields observed in bulk water manifest in the distribution of amino acid conformations in the graphene adsorbed state? The ϕ-ψ distributions of the Ala-like amino acids in graphene adsorbed state are shown in Fig. 6b for selected amino acids and Fig. S7 for all amino acids. The ϕ-ψ distributions of all amino acids in graphene adsorbed state differ from their distributions in bulk water.

In each scenario, the graphene sheet altered ϕ-ψ distributions by affecting the ratio of the low to high energy conformations observed in bulk water. Broadly one could classify two types of changes: (i) change in the populations of already favored (sampled) conformations (ii) sampling of new regions that were not sampled in bulk water. Across all force fields for most amino acids, there is a decrease in β-sheet, regions vicinal to α-helix (i.e., δ-region), and the P_{II} region. Consistent increase in the sampling of β-planar, and α-planar regions is observed, which are regions not sampled in bulk water. However, there are differences across amino acids for a given force field. That is, the specific regions with increased or decreased sampling are dependent on the nature of the amino acid. For example, while the α-planar region is sampled more for Thr when adsorbed to graphene, the bridge regions show increased sampling in case of Gln for Amberff99SB-ILDN/TIP3P.

Consistent with previous studies[37,75], we observe conformational preference towards α-

planar and β-planar structure upon adsorption to graphene. Ala samples α-planar configurations in both Amber-based force fields and CHARMM36/mTIP3P. This region is relatively less sampled in case of OPLS-AA/M/TIP3P. Poblete *et al.*[75] attributed these configurations to the stabilization of graphene–amino acid interactions through amide-π bonds. Poblete *et al.*[75] and our results indicate that the non-polarizable description of graphene in Amber-based and CHARMM36/mTIP3P force fields appear to capture the amide-π bond features.

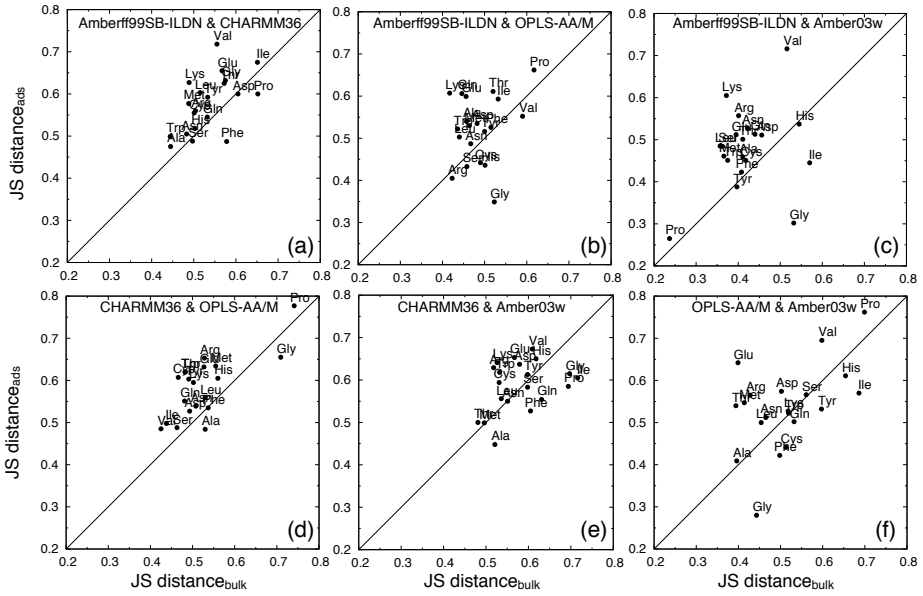

FIG. 7. Variation of the JS distance between the backbone dihedral angles estimated with different pairs of force fields. (a) Amberff99SB-ILDN/TIP3P & CHARMM36/mTIP3P, (b) Amberff99SB-ILDN/TIP3P & OPLS-AA/M, (c) Amberff99SB-ILDN/TIP3P & Amber03w/TIP4P/2005, (d) CHARMM36/mTIP3P & OPLS-AA/M, (e) CHARMM36/mTIP3P & Amber03w/TIP4P/2005, and (f) OPLS-AA/M/TIP3P & Amber03w/TIP4P/2005. The x- and y-axis refer to the JS distance between the ϕ-ψ distributions of the indicated force fields in bulk and adsorbed states, respectively. If JS distance is 0, the ϕ-ψ distributions obtained with the indicated force fields are identical. The solid line in each plot is the y=x line.

Another noteworthy observation is that in several cases the changes in the ϕ-ψ distribution upon adsorption to graphene appear to be subtly influenced by the distributions in bulk state. For instance, consider the ϕ-ψ distributions of Gln. In case of CHARMM36/mTIP3P, Gln samples a narrow region (diagonal) in α-helix and surrounding regions. In the graphene adsorbed state, Gln maintains this narrow diagonal feature in this region while sampling

more towards the bridge region. Similarly, with Amber03w/TIP4P/2005, Gln samples α-helix and δ-region in bulk state. In the adsorbed state Gln continues to sample these regions while also sampling the bridge regions. This suggests that the bulk behavior has some influence on the ϕ-ψ distributions of the adsorbed state. We hypothesize that the force fields have different barriers (free energy landscapes) in the ϕ-ψ space in bulk state, and when adsorbed to graphene some of these barriers are overcome in case of few amino acids in different force fields. Again considering the example of Gln, this would indicate that Gln is unable to overcome the barriers for Gln to escape the α-surrounding region in CHARMM36/mTIP3P in the adsorbed state, while barriers within the α-surrounding region are low for Amber03w/TIP4P/2005, thus enabling Gln to sample those configurations in the adsorbed state. We further quantify the differences in bulk and adsorbed states using Jensen-Shannon (JS) distance between the ϕ-ψ distributions (see Fig. S8)[102–105]. JS distance is an information theory based metric that is used to measure the distance between distributions of a random variable. Its value is bounded between 0 and 1, where 0 implies that the two distributions are identical. The values for most amino acids lie between 0.3 and 0.8 indicating that the distributions in the adsorbed state are different from those in bulk for all force fields.

Are the differences observed across various force fields in the graphene adsorbed state primarily an effect of the differences in bulk state? To answer this, we calculated the JS distance between the ϕ-ψ distributions between force fields in bulk water and graphene adsorbed state. In Fig. 7 (also see SI Fig. S9) we plot the correlations between these for various pairs of force fields. There is moderate linear trend in the JS distance observed in bulk and adsorbed state between few force field pairs. This suggests that the observed conformational differences in graphene adsorbed state are partly a result of the differences in the bulk state. Amino acids deviating from y=x line in Fig. 7 indicate the degree of differences in the ϕ-ψ distributions across the indicated force fields between the bulk and adsorbed states. The points above the y=x line indicate that the differences in ϕ-ψ distributions between the two force fields for the adsorbed state are larger than those in bulk state for those amino acids. In most cases, majority of the amino acids are above the y=x line implying that we might observe greater force field dependence when studying protein behavior on graphene surfaces than in bulk.

3. Orientational preferences of graphene adsorbed amino acids

To elaborate further on the conformations of amino acids when adsorbed to graphene beyond ϕ-ψ distributions, we use cluster analysis[106–108] that is typically used in unsupervised machine learning. We use the heavy atom of amino acid–graphene distances, backbone dihedral angles, and intramolecular interaction energy of amino acid as the raw data to perform cluster analysis. We eliminate linear correlations in the data by projecting the data along the top principal component vectors, which are identified using principal component analysis. For identifying clusters in the projected data, we employ accelerated hierarchical density-based spatial clustering of applications with noise (HDBSCAN*) algorithm[106,107]. We preserve the relative orientation of the atoms within the amino acid with respect to graphene by maintaining the net rotational and translational variance around the vector normal to graphene. This approach enables us to identify the most prominent conformations sampled in the simulations. Further methodological details are provided in SI. The clusters identified are projected onto ϕ-ψ-d subspace, where d is the distance between center of mass of side chain and graphene. We do so to facilitate the comparison of conformations sampled across different force fields. However, we caution the readers that clusters identified as similar in this subspace could have differences in other dimensions.

Results from cluster analysis are shown in Fig. 8 and 9 for Val and Trp, respectively. Results for the remaining amino acids are shown in SI. In the case of Val, the distance between the side chain center of mass and graphene (d) sampled in the prominent clusters is similar for the two Amber-based force fields and OPLS-AA/M/TIP3P. Few conformations that contribute to $\sim$2% (cluster 7 and 9) of the total structures that have the side chain at $d >$ $\sim$0.7-0.8 nm are observed in CHARMM36/mTIP3P. Such structures are not found with other force fields.

In the largest cluster (cluster 0) of adsorbed Val for Amber03w/TIP4P/2005 (Fig. 8), the backbone is parallel to the graphene surface and the side chain is oriented towards water. This conformation displays β-planar nature. In the next largest cluster (cluster 1), ϕ-ψ angles are close to the P_{II} region and the sidechain is oriented parallel to the surface. In cluster 1, the C-terminus fluctuates to and fro from the surface. Cluster 2 shares the ϕ-ψ space with cluster 0 and 1, and has more variance in its conformation relative to graphene surface. In contrast to cluster 1, the N-terminus position fluctuates from the surface in

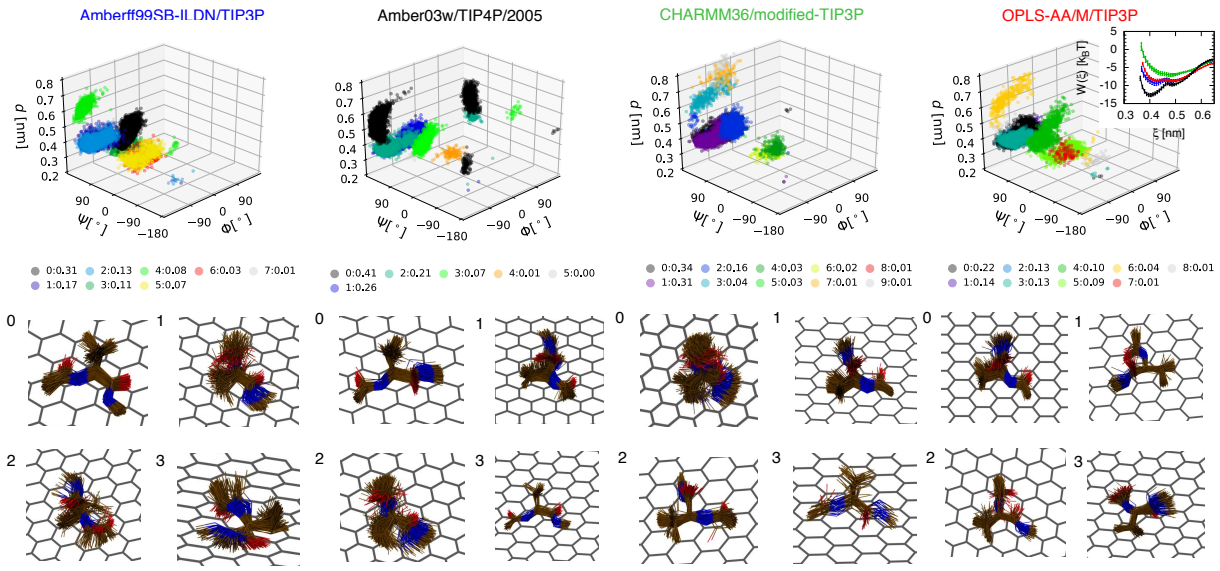

FIG. 8. The clusters of Val in the ϕ-ψ-d subspace obtained with the four force fields. d is the z-distance between the center of mass of side chain and the graphene surface. Each color in each plot represents a different cluster. The cluster labels are shown in the legend below each plot. The numbers next to each color in the legend indicate the fraction of total structures in that cluster. Note that the sum of the fractions is not one because the points identified as noise are not shown. The superimposed structures (every 10 frames) obtained by performing a 2D rotational and translational fit along the plane of the graphene surface (represented by lines) from the top four clusters are shown for each force field. A 3D fit would make the structures look more similar but we would loose information on the orientation of the amino acid with respect to graphene (Fig. S10). Cluster numbers are included alongside each snapshot. Carbon, nitrogen, oxygen are colored as brown, blue, and red, respectively in the structures. Hydrogen atoms and water are not shown for visual clarity. The PMF profiles of Val for ξ <0.65 obtained with the four force fields is also shown.

cluster 2. Clusters 0, 1 and 2 contain ∼88% of the total structures and collectively indicate that Val is stabilized on graphene by having at least one amide bond parallel to the graphene surface in Amber03w/TIP4P/2005.

In the largest cluster (cluster 0) in Amberff99SB-ILDN/TIP3P which comprises ∼31% of the total structures, the sidechain of Val orients towards water and the backbone aligns with the surface, similar to that in case of cluster 0 of Amber03w/TIP4P/2005. How-

ever, the difference in clusters 0 of the two force fields lies in the ϕ-ψ region sampled. In Amber03w/TIP4P/2005, the region corresponds to β-planar while in Amberff99SB-ILDN/TIP3P the region sampled is P_{II}. Clusters 1 and 2 which have ∼30% of the structures correspond to clusters 2 and 1 (respectively) in Amber03w/TIP4P/2005 which comprise ∼47% of the structures. Cluster 3 of Amberff99SB-ILDN/TIP3P has ∼11% contribution and belongs to the bridge region in the ϕ-ψ space. This configuration is rarely observed in Amber03w/TIP4P/2005.

The most populated clusters (0, 1, and 2) in CHARMM36/mTIP3P are located within β-region in the ϕ-ψ space. Cluster 0 in CHARMM36/mTIP3P resembles cluster 2 of Amber03w/TIP4P/2005 and cluster 1 of Amberff99SB-ILDN/TIP3P. Cluster 1 of CHARMM36/mTIP3P and cluster 1 of Amber03w/TIP4P/2005 are similar. Cluster 2 of CHARMM36/mTIP3P, ∼16% of the structures, comprises of extended conformations. These are rarely observed in other force fields. Cluster 0 of OPLS-AA/M/TIP3P is similar to cluster 1 of Amber03w/TIP4P/2005. The other top clusters (cluster 1 and 2) of OPLS-AA/M/TIP3P are around the β-region, and the sidechain aligns with the surface in these clusters.

The collective differences indicate that Amber03w/TIP4P/2005 restricts the conformations of Val to few structures, and its dominant cluster 0 (∼41%) is rarely observed with other force fields. This coincides with one of the dominant well in the PMF of Val obtained with Amber03w/TIP4P/2005. Though OPLS-AA/M/TIP3P and Amberff99SB-ILDN/TIP3P have similar PMF profiles, the distribution of structures within the clusters are different. CHARMM36/mTIP3P shares some of the clusters with Amberff99SB-ILDN/TIP3P and Amber03w/TIP4P/2005 but has a lower value of PMF for ξ <0.6 nm. Thus, for similar (different) PMFs we observe a diversity (similarity) of structures across different force fields.

In contrast to Val, Trp shows lesser differences across the different force fields. In the top four clusters that contribute to >80% of the total structures sampled, all the force fields indicate stacking between the aromatic group of Trp and graphene. The dominant cluster (cluster 0) in all force fields is similar. This cluster is relatively less sampled (∼35%) with Amber03w/TIP4P/2005 compared to other force fields (>40%). The backbone configuration of cluster 0 is around β-region with C-terminus closer to the surface than the N-terminus. Other clusters vary in the relative position of C- and N- terminus with respect to graphene and sampling in the ϕ-ψ space. However, they share the feature of Trp sidechain and graphene aromatic groups stacking. The configuration of cluster 1 in CHARMM36/mTIP3P

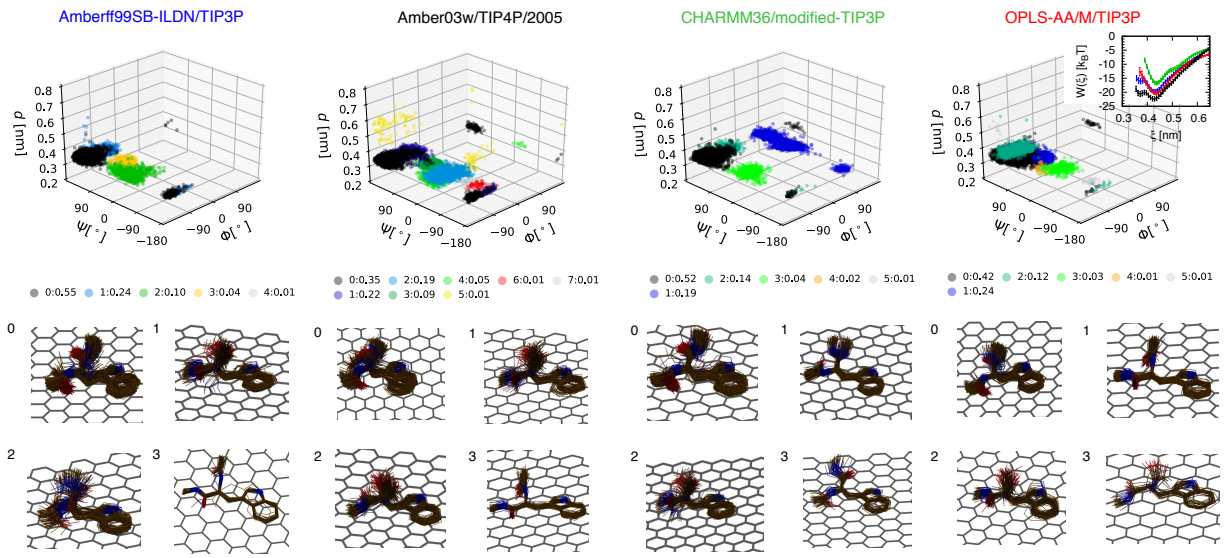

FIG. 9. The clusters of Trp in the ϕ-ψ-d subspace obtained with the four force fields. Color code is same as that in Figure 8.

is around the left-handed α-helix region (α' and δ') and is rarely sampled in other force fields. The remaining clusters in CHARMM36/mTIP3P and OPLS-AA/M/TIP3P have similar sampling in ϕ-ψ space. The prominence of the stacked configurations across all force fields indicates that the π-π interactions between Trp and graphene dominate the Trp-graphene interactions.

IV Conclusions

Given the multiple options of force fields available to study protein–graphene interactions and the absence of experimental data for validation of these force fields, a comparison between their performances is necessary. In this work, we investigate the differences in the description of amino acid–graphene interactions by four empirical non-polarizable force fields — Amberff99SB-ILDN/TIP3P, CHARMM36/mTIP3P, OPLS-AA/M/TIP3P, and Amber03w/TIP4P/2005. We find that the formation of amino acid–graphene complex is favorable for all amino acids and force fields. The PMFs between the amino acids and graphene are qualitatively similar between the force fields. There are quantitative differences in the PMFs especially near the global minimum for each amino acid. In general, the two Amber force fields are in good agreement with each other. CHARMM36/mTIP3P predicts the less favorable adsorption energy for most amino acids amongst the studied force fields.

While there are quantitative differences, the relative binding strength of amino acids are similar between the force fields in most cases. This implies an inherent convergence in the delicate description of amino acid interactions with their surrounding environment across the tested force fields. Furthermore, the relative binding preferences of the amino acids on graphene agree with the available experimental, and *ab initio* data. Given the limited available experimental data, it is possible that this agreement will reduce as experimental data for more amino acids is reported. This might occur either because of inaccurate description of the potential energy of the system or inherent limitations of the non-polarizable force fields. In the latter scenario, polarizable force fields such as Drude-particle[109] or fluctuating charge models[110] can be explored. Previous studies report that polarizability can influence water structure near graphene to varying degrees.[37,111,112]. For all the studied non-polarizable force fields, we observed that ΔA_{ads} strongly correlates with amino acid–graphene, and water–graphene intermolecular interactions.

We observe variations in the structures sampled by the amino acids near the graphene surface across different force fields. We characterize the conformations based on ϕ-ψ distributions and cluster analysis. The results indicate that the differences in conformations in the bulk state do not fully account for those seen in the adsorbed state between the force fields. The differences in the conformations of adsorbed amino acids is a manifestation of differences in bulk state, the configurational energy landscape of the amino acids and graphene-amino acid interactions. Our clustering protocol enables us to identify the most prominent structures sampled for each amino acid with every force field and facilitates the comparison between them. The prominent conformations of the amino acids in the graphene adsorbed state involve amide-π and π-π stacking interactions. For most cases, the adsorbed state has lower preference for the α-helix region. New dihedral space corresponding to α-planar, β-planar, bridge, and P_{II} regions are sampled in the graphene adsorbed state. The specific region that is favored is strongly dependent on both the force field and the amino acid. Larger differences in amino acid structures sampled upon adsorption across different force fields seem to be observed where there are no clear dominant graphene-amino acid interactions. These differences are subdued when a dominant interaction takes over – such as in case of aromatic residues and Arg. It is however, clear that even with similar PMFs differences in conformations are observed.

What can the adsorption of amino acids on graphene tell us about the force field de-

pendence of adsorption behavior of larger molecules, such as peptides and proteins? It could be expected that the different non-polarizable force fields will give similar qualitative behavior when comparing the relative strength of adsorption across different peptides or proteins. Of course, the ΔA_{ads} of larger molecules need not be the sum of ΔA_{ads} of individual amino acids. We are currently investigating this aspect in context of peptide–graphene interactions. It could be hypothesized that the differences in the structures of adsorbed amino acids across force fields would build up to manifest in large discrepancies for protein structure on graphene. It however, can be counter argued that since these force fields are developed based on folded protein structure the differences will eventually converge. Therefore, it is quite possible that there is a non-linear trend in the differences across force fields with respect to size of the protein. Nevertheless, with limited sampling it can be expected that the conformations of the peptides/proteins in the graphene adsorbed state could vary with the force field. While the enhanced sampling methods can help in addressing the sampling challenges in simulations, more experimental data characterizing protein structure at surfaces is necessary to validate the force fields, and improve them as needed.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgments

S. S acknowledges the financial support, in part, by the Defense Threat Reduction Agency (HDTRA-1-16-1-0023), and Clemson University start-up funds. J. K. B acknowledges funding from Clemson University Creative Inquiry program. We thank the Clemson Cyberinfrastructure Technology Integration group for the allotment of computing time on Palmetto cluster.

References

[1] D. Li, W. Zhang, X. Yu, Z. Wang, Z. Su and G. Wei, *Nanoscale*, 2016, **8**, 19491–19509.

[2] Z. E. Hughes and T. R. Walsh, *Nanoscale*, 2015, **7**, 6883–6908.

[3] D. Khatayevich, T. Page, C. Gresswell, Y. Hayamizu, W. Grady and M. Sarikaya, *Small*, 2014, **10**, 1505–1513.

[4] Y. Yu and L. Wu, *Anal. Bioanal. Chem.*, 2013, **405**, 4913–4919.

[5] K. Yang, L. Feng, X. Shi and Z. Liu, *Chem. Soc. Rev.*, 2013, **42**, 530–547.

[6] D. A. C. Brownson, D. K. Kampouris and C. E. Banks, *J. Power Sources*, 2011, **196**, 4873–4885.

[7] B. Luan, T. Huynh, L. Zhao and R. Zhou, *ACS Nano*, 2014, **9**, 663–669.

[8] A. Bianco, *Angew Chem Int. Edit.*, 2013, **52**, 4986–4997.

[9] V. C. Sanchez, A. Jachak, R. H. Hurt and A. B. Kane, *Chem. Res. Toxicol*, 2011, **25**, 15–34.

[10] C. Ge, J. Du, L. Zhao, L. Wang, Y. Liu, D. Li, Y. Yang, R. Zhou, Y. Zhao, Z. Chai and C. Chen, *Proc. Natl. Acad. Sci. U.S.A*, 2011, **108**, 16968–16973.

[11] Y. Chong, C. Ge, Z. Yang, J. A. Garate, Z. Gu, J. K. Weber, J. Liu and R. Zhou, *ACS Nano*, 2015, **9**, 5713–5724.

[12] R. Yue, Q. Lu and Y. Zhou, *Biosens. Bioelectron.*, 2011, **26**, 4436–4441.

[13] Y. Huang, X. Dong, Y. Liu, L. Li and P. Chen, *J. Mater. Chem.*, 2011, **21**, 12358–12362.

[14] O. Karunwi, C. Baldwin, G. Griesheimer, S. Sarupria and A. Guiseppi-Elie, *Nano LIFE*, 2013, **3**, 1343007.

[15] G. Zuo, X. Zhou, Q. Huang, H. Fang and R. Zhou, *J. Phys. Chem. C*, 2011, **115**, 23323–23328.

[16] J. Katoch, S. N. Kim, Z. Kuang, B. L. Farmer, R. R. Naik, S. A. Tatulian and M. Ishigami, *Nano Lett.*, 2012, **12**, 2342–2346.

[17] B. Sengupta, W. E. Gregory, J. Zhu, S. Dasetty, M. Karakaya, J. M. Brown, A. M. Rao, J. K. Barrows, S. Sarupria and R. Podila, *RSC Adv.*, 2015, **5**, 82395–82402.

[18] M. J. Penna, M. Mijajlovic, C. Tamerler and M. J. Biggs, *Soft Matter*, 2015, **11**, 5192–5203.

[19] X. Zou, S. Wei, J. Jasensky, M. Xiao, Q. Wang, C. L. Brooks III and Z. Chen, *J. Am. Chem. Soc.*, 2017, **139**, 1928–1936.

[20] Z. E. Hughes and T. R. Walsh, *Nanoscale*, 2018, **10**, 302–311.

[21] S. Morsbach, G. Gonella, V. Mailänder, S. Wegner, S. Wu, T. Weidner, R. Berger, K. Koynov, D. Vollmer, N. Encinas, S. L. Kuan, T. Bereau, K. Kremer, T. Weil, M. Bonn, H. Butt and K. Landfester, *Angew Chem. Int. Edit.*, 2018, 12626–12648.

[22] M. Deighan and J. Pfaendtner, *Langmuir*, 2013, **29**, 7999–8009.

[23] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García and J. M. Soler, *Phys. Status Solidi (b)*, 1999, **215**, 809–817.

[24] L. E. Ratcliff, S. Mohr, G. Huhs, T. Deutsch, M. Masella and L. Genovese, *WIRES Comput. Mol. Sci.*, 2017, **7**, e1290.

[25] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.

[26] A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. R. Stote, J. Straub, J. Watanabe, M.and Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.

[27] W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

[28] X. Daura, A. E. Mark and W. F. Van Gunsteren, *J. Comput. Chem.*, 1998, **19**, 535–547.

[29] J. Guo, X. Yao, L. Ning, Q. Wang and H. Liu, *RSC Adv.*, 2014, **4**, 9953–9962.

[30] Z. Gu, Z. Yang, L. Wang, H. Zhou, C. A. Jimenez-Cruz and R. Zhou, *Sci. Rep-UK*, 2015, **5**, 10873.

[31] W. Lv, G. Xu, H. Zhang, X. Li, S. Liu, H. Niu, D. Xu and R. Wu, *Sci. Rep-UK*, 2015, **5**, 7572.

[32] D. Zhao, L. Li, D. He and J. Zhou, *Appl. Surf. Sci.*, 2016, **377**, 324–334.

[33] Z. Xu, X. Yang, Q. Wei, W. Zhao, B. Cui, X. Yang and N. Sahai, *Langmuir*, 2018, 7932–7941.

[34] E. Wu, M. Coppens and S. Garde, *Langmuir*, 2015, **31**, 1683–1692.

[35] G. Duan, S. Kang, X. Tian, J. A. Garate, L. Zhao, C. Ge and R. Zhou, *Nanoscale*, 2015, **7**, 15214–15224.

[36] R. A. Latour, *Colloid Surface B*, 2014, **124**, 25–37.

[37] Z. E. Hughes, S. M. Tomásio and T. R. Walsh, *Nanoscale*, 2014, **6**, 5438–5448.

[38] L. Martin, M. M. Bilek, A. S. Weiss and S. Kuyucak, *Interface focus*, 2016, **6**, 20150045.

[39] T. Werder, J. Walther, R. Jaffe, T. Halicioglu and P. Koumoutsakos, *J. Phys. Chem. B*, 2003, **107**, 1345–1352.

[40] O. F. Lange, D. Van der Spoel and B. L. De Groot, *Biophys. J.*, 2010, **99**, 647–655.

[41] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror and D. E. Shaw, *PLOS ONE*, 2012, **7**, e32131.

[42] F. Martín-García, E. Papaleo, P. Gomez-Puertas, W. Boomsma and K. Lindorff-Larsen, *PLOS ONE*, 2015, **10**, e0121114.

[43] P. Robustelli, S. Piana and D. E. Shaw, *Proc. Natl. Acad. Sci. U.S.A*, 2018, 201800690.

[44] R. B. Pandey, Z. Kuang, B. L. Farmer, S. S. Kim and R. R. Naik, *Soft Matter*, 2012, **8**, 9101–9109.

[45] S. J. Rodríguez, L. Makinistian and E. A. Albanesi, *Appl. Surf. Sci.*, 2017, **419**, 540–545.

[46] C. R. So, Y. Hayamizu, H. Yazici, C. Gresswell, D. Khatayevich, C. Tamerler and M. Sarikaya, *ACS Nano*, 2012, **6**, 1648–1656.

[47] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Structure, Function, and Bioinformatics*, 2010, **78**, 1950–1958.

[48] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell Jr, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.

[49] W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657–1666.

[50] M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.*, 2015, **11**, 3499–3509.

[51] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

[52] E. Neria, S. Fischer and M. Karplus, *J. Chem. Phys.*, 1996, **105**, 1902–1921.

[53] R. B. Best and J. Mittal, *J. Phys. Chem. B*, 2010, **114**, 14916–14923.

[54] J. L. F. Abascal and C. Vega, *J. Phys. Chem.*, 2005, **123**, 234505.

[55] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.

[56] D. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *Amber 14*, 2014.

[57] A. Minoia, *Buildcstruct homepage (Version 1.1)*, 2014, `http://chembytes.wikidot.com/buildcstruct`, (accessed January 2016).

[58] C. Lee, X. Wei, J. W. Kysar and J. Hone, *Science*, 2008, **321**, 385–388.

[59] E. P. Raman, J. A. Lemkul, R. Best and A. D. Mackerell Jr, *CHARMM force field homepage*, 2016, `http://mackerell.umaryland.edu/charmm_ff.shtml`, (accessed May 2016).

[60] M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, *OPLS-AA/M force field homepage*, 2016, `http://zarbi.chem.yale.edu/oplsaam.html`, (accessed June 2016).

[61] J. A. Yancey, N. A. Vellore, G. Collier, S. J. Stuart and R. A. Latour, *Biointerphases*, 2010, **5**, 85–95.

[62] H. B, *J. Chem. Theory Comput.*, 2008, **4**, 116–122.

[63] G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

[64] S. Nosé, *Mol. Phys.*, 1984, **52**, 255–268.

[65] W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695.

[66] T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

[67] S. Páll and B. Hess, *Comput. Phys. Commun.*, 2013, **184**, 2641–2650.

[68] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011–1021.

[69] J. S. Hub, B. L. De Groot and D. Van Der Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713–3720.

[70] Z. E. Hughes and T. R. Walsh, *J. Mater. Chem. B*, 2015, **3**, 3211–3221.

[71] C. M. Welch, A. N. Camden, S. A. Barr, G. M. Leuty, G. S. Kedziora and R. J. Berry, *J. Chem. Phys.*, 2015, **143**, 045104.

[72] J. Comer, R. Chen, H. Poblete, A. Vergara-Jaque and J. E. Riviere, *ACS Nano*, 2015, **9**, 11761–11774.

[73] G. Nawrocki and M. Cieplak, *J. Phys. Chem. C*, 2014, **118**, 12929–12943.

[74] N. A. Vellore, J. A. Yancey, G. Collier, R. A. Latour and S. J. Stuart, *Langmuir*, 2010, **26**, 7396–7404.

[75] H. Poblete, I. Miranda-Carvajal and J. Comer, *J. Phys. Chem. B*, 2017, **121**, 3895–3907.

[76] MarvinSketch, *Marvin 17.16.0 (17.16.0) was used for drawing, displaying and characterizing chemical structures, substructures and reactions*, 2017, `http://www.chemaxon.com`, (accessed Jan 2017).

[77] M. J. Penna, M. Mijajlovic and M. J. Biggs, *J. Am. Chem. Soc.*, 2014, **136**, 5323–5331.

[78] S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Biophys. J.*, 2011, **100**, L47–L49.

[79] F. Iori and S. Corni, *J. Comput. Chem.*, 2008, **29**, 1656–1666.

[80] S. Dasetty and S. Sarupria, unpublished work.

[81] A. N. Camden, S. A. Barr and R. J. Berry, *J. Phys. Chem. B*, 2013, **117**, 10691–10697.

[82] H. T. Larijani, M. D. Ganji and M. Jahanshahi, *RSC Adv.*, 2015, **5**, 92843–92857.

[83] P. Singla, M. Riyaz, S. Singhal and N. Goel, *Phys. Chem. Chem. Phys.*, 2016, **18**, 5597–5604.

[84] S. S. K. Mallineni, J. Shannahan, A. J. Raghavendra, A. M. Rao, J. M. Brown and R. Podila, *ACS Appl. Mater. Inter.*, 2016, **8**, 16604–16611.

[85] R. Zhiani, *Appl. Surf. Sci.*, 2017, **409**, 35–44.

[86] J. Kyte and R. F. Doolittle, *J. Mol. Biol.*, 1982, **157**, 105–132.

[87] N. Dragneva, W. B. Floriano, D. Stauffer, R. C. Mawhinney, G. Fanchini and O. Rubel, *J. Chem. Phys.*, 2013, **139**, 174711.

[88] H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.

[89] L. H. Kapcha and P. J. Rossky, *J. Mol. Biol.*, 2014, **426**, 484–498.

[90] D. B. Rubin, *Ann. Stat.*, 1981, 130–134.

[91] R. Bååth, *bayesboot: An Implementation of Rubin's (1981) Bayesian Bootstrap*, 2018, `https://cran.r-project.org/package=bayesboot`, (accessed July 2018).

[92] G. H. Zerze, R. G. Mullen, Z. A. Levine, J. Shea and J. Mittal, *Langmuir*, 2015, **31**, 12223–12230.

[93] R. B. Best and G. Hummer, *J. Phys. Chem. B*, 2009, **113**, 9004–9015.

[94] K. K. Patapati and N. M. Glykos, *Biophys. J.*, 2011, **101**, 1766–1771.

[95] M. D. Smith, J. S. Rao, E. Segelken and L. Cruz, *J. Chem. Inf. Model.*, 2015, **55**, 2587–2595.

[96] R. B. Best, N.-V. Buchete and G. Hummer, *Biophys. J.*, 2008, **95**, L07–L09.

[97] D. Matthes and B. L. De Groot, *Biophys. J.*, 2009, **97**, 599–608.

[98] F. Vitalini, F. Noé and B. G. Keller, *Data in Brief*, 2016, **7**, 582–590.

[99] A. D. MacKerell Jr, *J. Comput. Chem.*, 2004, **25**, 1584–1604.

[100] S. A. Hollingsworth and P. A. Karplus, *Biomolecular concepts*, 2010, **1**, 271–283.

[101] O. Carugo and K. Djinović-Carugo, *Acta. Crystallogr. D.*, 2013, **69**, 1333–1341.

[102] J. Lin, *IEEE T. Inform. Theory*, 1991, **37**, 145–151.

[103] K. Lindorff-Larsen and J. Ferkinghoff-Borg, *PLOS ONE*, 2009, **4**, e4203.

[104] K. C. Wolfe and G. S. Chirikjian, *Entropy*, 2012, **14**, 213–232.

[105] C. L. McClendon, L. Hua, G. Barreiro and M. P. Jacobson, *J. Chem. Theory Comput.*, 2012, **8**, 2115–2126.

[106] L. McInnes, J. Healy and S. Astels, *The Journal of Open Source Software*, 2017, **2**, 205.

[107] L. McInnes and J. Healy, Data Mining Workshops (ICDMW), 2017 IEEE International Conference on, 2017, pp. 33–42.

[108] R. J. Campello, D. Moulavi, A. Zimek and J. Sander, *ACM T Knowl. Discov. D.*, 2015, **10**, 5.

[109] J. A. Lemkul, J. Huang, B. Roux and A. D. MacKerell Jr, *Chem. Rev.*, 2016, **116**, 4983–5013.

[110] S. Patel, A. D. Mackerell Jr and C. L. Brooks III, *J. Comput. Chem.*, 2004, **25**, 1504–1514.

[111] T. A. Ho and A. Striolo, *J. Chem. Phys.*, 2013, **138**, 054117.

[112] A. Striolo, A. Michaelides and L. Joly, *Annu. Rev. Chem. Biol.*, 2016, **7**, 533–556.

We thoroughly investigate the differences in free energy of adsorption and the structures of the amino acids adsorbed on graphene with force fields.