What makes a good graphene-binding peptide? Adsorption of amino acids and peptides at aqueous graphene interfaces†

Zak E. Hughes,*[a] and Tiffany R. Walsh[a]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

Investigation of the non-covalent interaction of biomolecules with aqueous graphene interfaces is a rapidly expanding area. However, reliable exploitation of these interfaces in many applications requires that the links between the sequence and binding of the adsorbed peptide structures be clearly established. Molecular dynamics (MD) simulations can play a key role in elucidating the conformational ensemble of peptides adsorbed at graphene interfaces, helping to elucidate these rules in partnership with experimental characterisation. We apply our recently-developed polarisable force-field for biomolecule-graphene interfaces, GRAPPA, in partnership with advanced simulation approaches, to probe the adsorption behaviour of peptides at aqueous graphene. First we determine the free energy of adsorption of all twenty naturally occurring amino acids (AAs) via metadynamics simulations, providing a benchmark for interpreting peptide-graphene adsorption studies. From these free energies, we find that strong-binding amino acids have flat and/or compact side chain groups, and we relate this behaviour to the interfacial solvent structuring. Second, we apply replica exchange with solute tempering simulations to efficiently and widely sample the conformational ensemble of two experimentally-characterised peptide sequences, P1 and its alanine mutant P1A3, in solution and adsorbed on graphene. For P1 we find a significant minority of the conformational ensemble possesses a helical structure, both in solution and when adsorbed, while P1A3 features mostly extended, random-coil conformations. In solution this helical P1 configuration is stabilised through favourable intra-peptide interactions, while the adsorbed structure is stabilised via interaction of four strongly-binding residues, identified from our metadynamics simulations, with the aqueous graphene interface. Our findings rationalise the performance of the P1 sequence as a known graphene binder.

Introduction

Since its discovery a decade ago, graphene has been the subject of intense scrutiny,[1] due to its promise in a wide range of applications, including biological imaging,[2] desalination,[3] biochemical sensors[4–6] and biomedicine.[7,8] A number of these applications capitalise on the selective interaction of biomolecules with graphitic interfaces, in aqueous conditions. To better exploit these bio-nanotechology applications it is necessary to gain greater understanding of the the structure/property relationships of biomolecules at graphitic interfaces.[9,10] Developing such an understanding is a non-trivial task that will involve experimental, theoretical and simulation studies.

From screening experiments, peptide sequences that selectively bind to either the basal plane or edges of graphene flakes have been identified.[11,12] However, the quantitative binding affinities of these peptides adsorbed at aqueous graphene interfaces have not yet been reported. In addition, experimental methods have not yet been able to directly determine the structures of the peptides when adsorbed *under aqueous conditions* for a wide range of biomolecule/materials interfaces. Microscopy has been used to observe the structure of adsorbed peptides at graphitic interfaces under dry conditions.[12–14] However, the relationship between the interfacial structure of peptides when dried, *vs.* in-solution conditions is not yet understood; it is highly likely that the drying process strongly affects the resulting packing morphology of the adsorbed chains. Existing structural data obtained for peptides adsorbed at aqueous interfaces are often (but not always) based on indirect techniques such as circular dichroism spectroscopy,[13,15] although advances have been made in for in-solution NMR in the adsorbed state.[16] While such studies can provide valuable information, making connections between

† Electronic Supplementary Information (ESI) available: Details of the REST simulations; the side-chain contact sites; cluster populations of peptides both in solution and adsorbed to the graphene interface; analysis of aromatic-aromatic residue interactions of P1 in solution; analysis of the intra-peptide hydrogen bonding of P1 and P1A3 in solution; free energy of adsorption profiles of amino acids to graphene interface; enthalpies of adsorption of amino acids to graphene interface from previous studies; exemplar replica mobilities; number of clusters identified as a function of MD steps from the REST simulations; composition of the secondary structure of the peptides via analysis of their backbone dihedral angles; probability distribution of distances of residues from graphene interface during the REST simulations. See DOI: 10.1039/b000000x/

[a] *Institute for Frontier Materials, Deakin University, Geelong, Australia. Fax: +61 (0)3 5227 1103; Tel: +61 (0)3 5247 9160; E-mail: zhughes@deakin.edu.au*

the sequence of a peptide to its structure and binding affinity remains highly challenging.

Molecular dynamics (MD) simulations can play a valuable role in elucidating the structure of adsorbed peptides species at aqueous inorganic interfaces, thereby enabling deeper connections between structure and function to be established. However, there are specific challenges that molecular simulation must overcome to reach its full potential in this regard. First, the force-fields (FFs) used to model the interface must describe the interaction of the peptide with the aqueous substrate appropriately. A general lack of comprehensive, quantitative structural and binding data from experiment for these interfaces currently prevents a definitive verification of such FFs. Second, many materials-binding peptides are thought to be intrinsically disordered; the corresponding potential energy landscape describing the peptide conformational ensemble is anticipated to be complex. Therefore careful conformational sampling is an indispensable tool for ensuring meaningful results[17–29], since insufficient sampling may give rise to misleading conclusions. The challenge of extensive conformational sampling is made more acute by the need to use an explicit description of liquid water,[30,31] since spatial structuring of interfacial solvent is thought to be a key determinant in peptide-materials binding.[26,30] However, this requirement to describe liquid water at the molecular level leads to inefficiencies in conventional advanced sampling methods such as temperature-based replica-exchange MD.[25] Therefore, the search for advanced sampling approaches that are effective, relatively economical to use under aqueous conditions, and readily implemented in mainstream molecular simulation software packages is a subject of active development. Here, we have used replica-exchange with solute tempering (REST) MD[32–35], as it meets these criteria, and the outcomes from REST compare favourably with temperature-based REMD benchmarks for peptide-materials simulations[25].

Many of the the previous studies of biomolecules with graphitic substrates have used traditional biomolecular force-fields (FFs) such as CHARMM, Amber and so forth to model the interaction of the biomolecules/water with the graphitic nano-structures.[36–44] In all of these cases, the FF used captured the interaction between the adsorbates and the graphitic interface solely through van der Waals interactions, *i.e.* neglecting polarisation effects. While this approximation can be understood from a practical point of view, a number of recent studies have shown that polarisation can influence the behaviour of the the graphitic-aqueous interfaces[45–48] as well as the interaction of biomolecules with such substrates, especially in the case of charged species.[49–51] One of the major challenges in incorporating polarisation effects into the FF is the increased computational cost. For example the AMOEBAPRO FF is a high quality FF that accounts for the polarisation of atoms, but the computational cost associated with using it is high,[49,51] thus limiting the time- and/or length-scales of the simulation, and/or necessitating the use of an implicit solvent.[10,49,50,52] However, the use of an implicit solvent can bias the sampling of the system, promoting certain forms of secondary structure at the expense of others.[53–59] This increased computational cost arising from the FF directly impacts on the requirement, outlined above, for extensive conformational sampling.

To overcome these dual challenges of predicting the conformational ensemble of biomolecules adsorbed at aqueous graphitic nano-structures via the use of a polarisable FF, while at the same time economising on the computational expense, we recently developed a new polarisable FF for aqueous biographitic interfaces, GRAPPA.[51] In GRAPPA, the polarisation of the graphene surface is described via a rigid-rod dipole, while the water/adsorbate species are modelled using the CHARMM family of FFs.[60,61] While the description of polarisation within GRAPPA is not as rigorous compared with AMOEBAPRO, the computational cost is drastically reduced in comparison. For example, the free-energy of adsorption of amino acids (AAs) at the aqueous graphene interface can be determined from metadynamics simulations[62] using GRAPPA,[51] something that would be impracticable using the AMOEBAPRO FF. Initial testing of GRAPPA showed that this FF can reproduce the spatial and orientational ordering of water at graphene- and CNT-aqueous interfaces reported from first-principles MD simulations.[63–65]

As mentioned previously, the free energy of adsorption of graphene-selective peptides to the aqueous graphene interface has not been reported either by experiment or simulation. Using a combination of steered MD and non-equilibrium thermodynamic integration Mijajlovic *et al.* were able to determine the free-energy of adsorption for a pentapeptide with a graphene surface via molecular simulation. However, such simulations have a significant computational cost, are technically challenging to perform, and produce results with very large error estimates. In contrast, the determination of the adsorption free-energy/enthalpy of individual AAs[39–41,44,51] can achieved more readily. While the interplay between the peptide sequence, adsorbed conformation(s), and peptide-interface binding means we cannot simply assume that a residue within a peptide will adsorb in the same manner as the corresponding AA,[18,30,66,67] such data do provide a valuable benchmark that can be used to interpret results from both simulation and experiment. A number of studies have recently reported calculations of the adsorption *enthalpies* of all twenty AAs to graphene under aqueous conditions, using a variety of different FFs,[39–41] none of which are polarisable. However, to the authors knowledge, the only reported free-energies of adsorption for AAs are in our previous study outlining the development of the GRAPPA FF,[51] where free energies of adsorption where calculated for eight AAs (Ala, Arg, Asp, Gly,

HisA, Phe, Tyr and Trp). In this study we have predicted the adsorption free energy for all twenty naturally occurring AAs, including both the protontated and non-protontated form of histidine.

These data presented herein will then be used to help interpret the adsorption behaviour of a known graphene-binding peptide, P1, and its mutant analogue, P1A3.[10,37,68] Simulations of P1 and P1A3 adsorbed at a graphene interface, have been reported previously.[10,37,39,69] However, none of these previous studies used a polarisable FF, or made use of any advanced conformational sampling approaches. Our results have revealed new, previously unseen behaviour of P1 and highlight the importance of considering both the description that the FF provides, and the degree of conformational sampling.

Methods

MD Simulations

Simulations of the AAs and peptides in explicit water at graphene interfaces were performed using GROMACS version 4.5.5.[70] The PLUMED plugin[71] was used to apply the well-tempered metadynamics approach[72] for the AA adsorption simulations. The CHARMM-modified version of the TIP3P[73,74] water model was used for the water molecules and the CHARMM22* FF parameters used for the AAs/peptides.[60,61] For these simulations a time-step of 1 fs was used with the Lennard-Jones (LJ) non-bonded interactions switched off between 10.0 and 11.0 Å and a cut-off of 13.0 Å used for the PME summation.[75]

The free energy of adsorption of all twenty naturally occurring AAs was calculated. The L-chiral forms of the amino acids were modelled with the amino acids capped by acetyl and N-methyl groups at the N- and C-termini, respectively. The AAs were modelled according to their likely protonation state at pH 7, with either a Na^+ or Cl^- used as a counterion to ensure overall charge neutrality where necessary. Histidine was modelled in both the charge neutral (HisA) and protonated (HisH) state. Each system contained 2325 water molecules, between two graphene sheets 44.27×38.34 Å, separated by 44 Å of liquid water and 36 Å of vacuum. The number of water molecules in the system was such to yield a bulk density of water in the centre of the inter-slab space equal to that of a simulation cell of bulk water at the same ambient temperature and pressure. All of these metadynamics simulations were performed in the Canonical (*NVT*) ensemble at a temperature of 300 K. The bias was applied on the position of the centre of mass of each AA along the z-axis (*i.e.* the direction perpendicular to the graphene surface). Gaussians of 0.5 Å width were deposited every 1 ps for a total simulations time of 150 ns per AA, and the initial Gaussian height was set to 0.15 kJ mol^{-1}. A well-tempered metadynamics bias factor of 10 was used.

The P1 (HSSYWYAFNNKT) and P1A3 (HSSAAAAFNNKT) peptides were simulated when adsorbed to a graphene interface as well as when free in solution (*i.e* without the presence of a graphene sheet). Both the Lys and His residue were modelled in their protonated states and two Cl^- were used as a counterions to ensure overall charge neutrality. Each system contained 6846 water molecules. For the surface-adsorbed simulations a graphene sheet 63.9×59.6 Å was separated from its periodic image by 57.5 Å of liquid water. For the simulations of the free peptides in solution, a cubic cell of dimensions $63.9 \times 59.6 \times 54.0$ Å was used. We used the Terakawa implementation of the REST approach[34] for the peptide simulations, with a total of sixteen replicas used in each case. Each REST simulation was run for 20 ns, with a temperature window spanning 300-433 K and different initial structures used for each replica. For full details of the REST simulations refer to the ESI†, section 'REST Simulation Details' and to our previous studies.[17,18,22,25] Snapshots of the trajectory were saved every 1 ps.

Analysis

For the AA adsorption free energy simulations, the zero-point of the free-energy was calculated as the average free-energy at a distance greater than 15.0 Å from the surfaces. The uncertainty was determined from the difference between the final free-energy and the average free-energy over the last 5 ns of simulation. See Fig. S1 in the ESI†for examples of barrier re-crossings and simulation convergence. For the simulations of the peptide, the Boltzmann-weighted ensemble of configurations at 300 K was determined from a cluster analysis of the full trajectory of the baseline (unscaled) potential. The clustering analysis was performed over all the backbone atoms of the peptide via the Daura method,[76] with a 2 Å cutoff. The secondary structure of the peptides was analysed according to two different definitions. In the first scheme, we assigned a secondary structure motif on the basis of the ϕ and ψ backbone dihedral angles, with each secondary structure motif corresponding to a range of angles as characterised in previous works.[17,25] In the second scheme, we assigned the secondary structure based on the hydrogen bonding of the backbone as defined by the dictionary of secondary structure of proteins (DSSP) program.[77] Analysis of the contact between each residue and the graphene surface was evaluated using a procedure previously developed for other materials-binding peptides.[18,22] In brief, each residue type was assigned a reference site in the side chain for determining the residue-surface separation. The residue was then deemed to be in contact with the graphene sheet if the distance between the site and the substrate was less than a cutoff distance. The cutoff distance for each residue was assigned on the basis of previous work and inspection of the distance distribution profiles. The side-chain

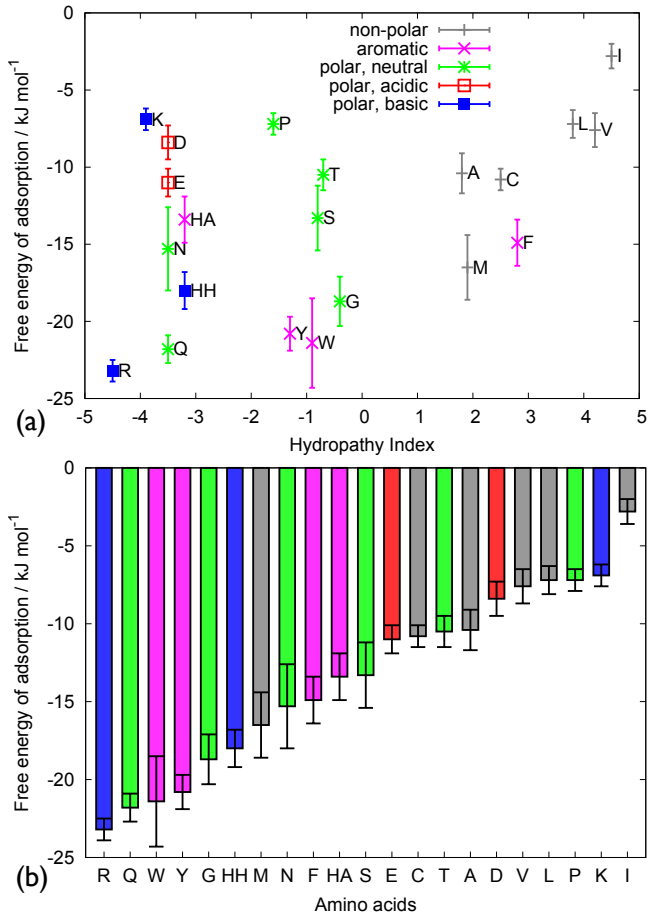

Fig. 1 Free energy of adsorption of all twenty amino acids shown as (a) a function of the hydropathy index and (b) in bar chart format.

sites and residue/surface distance cutoffs for each residue are provided in Table S1 in the ESI†.

Results and discussion

Adsorption free energy of amino acids

The adsorption free-energies of the AAs are shown as a function of their the Kyte-Doolittle hydropathy index (HI) in Fig. 1(a), and in a bar chart in Fig. 1(b), numeric values are given in Table 1. All species displayed affinity for adsorption at graphene aqueous interfaces in a range of ~ -3 to ~ -23 kJ mol^{-1}. The species with the strongest affinity for the aqueous graphene interface are Arg, Gln, Trp and Tyr while the most weakly adsorbing species are Ile, Lys, Pro, Leu and Val. Previous studies using non-polarisbale FFs have shown a correlation between the enthalpy of adsorption and either the mass or hydrophilicity of the AAs.[39–41] The free-energies calculated here showed little correlation with either. While Arg, Gln, and

Table 1 Free-energies of adsorption (ΔFE / kJ mol^{-1}) for the amino acids on the graphene surface determined from metadynamics simulations

Amino acid	ΔFE / kJ mol^{-1}
Ala	-10.4 ± 1.3
Arg	-23.2 ± 0.7
Asn	-15.3 ± 2.7
Asp	-8.4 ± 1.1
Cys	-10.8 ± 0.7
Glu	-11.0 ± 0.9
Gln	-21.8 ± 0.9
Gly	-18.7 ± 1.6
HisA	-13.4 ± 1.5
HisH	-18.0 ± 1.2
Ile	-2.8 ± 0.8
Leu	-7.2 ± 0.9
Lys	-6.9 ± 0.7
Met	-16.5 ± 2.1
Phe	-14.9 ± 1.5
Pro	-7.2 ± 0.7
Ser	-13.3 ± 2.1
Thr	-10.5 ± 1.0
Trp	-21.4 ± 2.9
Tyr	-20.8 ± 1.1
Val	-7.6 ± 1.1

Trp all possess large side-chains, Gly has the smallest side-chain of the set and yet also adsorbs strongly. Although the strongly hydrophobic AAs (HI > 3.5), Ile, Leu and Vel, are all weakly adsorbing, so is the highly hydrophilic Lys. Instead, we propose that the nature of the side-chain appears to play a determining role in the adsorption strength of the AA to the interface. Large, planar groups *e.g.* phenol, indole and guanadinium correlate with strong adsorption, as do compact side-chains such as seen for glycine. In our previous study we identified and rationalised the strong affinity between the amide group, which is also planar, and the graphene surface;[51] therefore it follows that both Gln and Asn are strongly adsorbing species. In contrast, their acidic analoges, Glu and Asp, are both weaker binders. The protonation of histidine increases the adsorption strength of His significantly, a result which is consistent with the polarisability of the graphene surface. In contrast to the planar nature of the functional groups of the strong binders, the weakly binding species are either bulky (Ile, Leu, Val), have a constrained geometry (Pro) or possess a long alkyl chain (Lys). The ability of the adsorbate to conform to a planar geometry can be related to the structure of water at the interface,[51] with the greatest density of water predicted to lie ~ 3 Å above surface, and with a second water layer being found at ~ 6 Å from the surface (the density profile of water at the graphene interface is shown in Fig. S2(a)†). Thus, bulky hydrophobic groups adsorbed at the in-

Table 2 The strongest and weakest binders for the different AA free-energy/enthalpy studies

Study	Metric	Strongest	Weakest
This work	ΔG_{ads}	Arg > Gln > Trp > Tyr	Ile < Lys ≈ Pro ≈ Leu ≈ Val
Pandey *et al.*	ΔH_{ads}	Trp > Arg > Tyr > Phe	Gly < Ala < Ser
Camden *et al.*	ΔH_{ads}	Arg > Gln > Asn	Val < Ile ≈ Leu ≈ Phe
Dragneva *et al.*	$\Delta H_{\text{ads}}^{\text{Dra}}$	Arg > Trp > Tyr	Thr ≈ Pro

terface will be exposed to this dense water region in the first interfacial solvent layer, making the adsorption of these AAs less favourable.

The effect of interfacial water structuring on the adsorption of the AAs can also be inferred from the free-energy profiles of the different species, shown in Figs. S2 and S3 in the ESI†. For all 21 adsorbates the global minimum is found at the graphene interface, with most profiles also showing a local minimum further from the surface. This local minimum is related to the 'solvent mediated' adsorption of the AAs, where the AA is separated from the graphene interface by the first layer of interfacial water molecules. This 'solvent mediated' minimum has also been observed for the adsorption of AAs and peptides at metal interfaces,[18,78,79] particularly for silver,[18,80] and mineral[26,30] interfaces.

As mentioned previously, to the authors knowledge no previous studies have reported the *free-energies* of adsorption of AAs at the aqueous graphene interface. However, a number of previous studies have reported calculations of the adsorption enthalpy of AAs on graphene.[39–41,44] In these previous studies, Pandey *et al.* used the, non-polarisable, AMBER ff99SB FF and calculated ΔH_{ads} for capped AAs. While Camden *et al.* calculated ΔH_{ads} for G-X-G tripeptides (in the zwitterionic form) using the non-polarisbale TEAM FF. In contrast, Dragneva *et al.*[41] defined the adsorption enthalpy differently to that of Pandey *et al.* and Camden *et al.*, and predicted ΔH_{ads} for both the capped and zwitterionic forms of the AAs. In addition, recent simulation studies investigating the adsorption of AAs on the Au(111) surface (using a polarisable FF) have shown that the adsorption enthalpies of AAs may differ significantly from their corresponding free-energies of adsorption.[22,80,81] As such, a direct quantitative comparison between these previous studies (based on enthalpies) and our current work (based on free energies) is not appropriate. However, some discussion of the qualitative findings of the different studies can be made, in particular concerning the relative affinity of the different AAs for graphene.

Table 2 summarises the strongest and weakest binders found for each study, including our present results, while Fig. S4 shows the adsorption enthalpy as a function of hydropathy index for the Camden and Pandey studies. While there are some similarities between the results of the previous studies and the present work, a number of key differences exist, in particular for those residues that adsorb moderately or weakly at the aqueous graphene interface. All studies found Arg to be a strong binder, and all but Camden *et al.* identified Trp and Tyr to also be strong binding. Both our present work and Camden *et al.* predict Gln as a strongly adsorbing species. There is more variation in the weakly adsorbing AAs; Ile, Leu and Val were identified as weakly binding species by both Camden *et al.* and in the present study. Pro is found by both ourselves and Dragneva *et al.* to be amongst the weakest binders. However, in stark contrast with our data that identify Gly as a strong binder, Pandey et al found the opposite. Comparing the overall results of the studies we also see different correlations, Camden *et al.* found a general trend between a decrease in hydropathy and stronger adsorption, something that was not observed in the other studies (Fig. S4 in the ESI†). However, the results of Pandey *et al.* do correlate stronger adsorption to the mass of the AA; their findings are consistent with the fact the interaction between the graphene and adsorbates in their study is based on LJ interactions only and not on other contributions such as polarisation. Unfortunately, none of these studies investigated the effect of the protonation of His, thus not allowing any evaluation of the theoretically greater ability of a polarisable FF to model the effects of protonation state to be judged. Overall, there appears to be a general consensus that Arg, Trp, Tyr and Gln are strong binders, but there is less agreement on the status of the other adsorbates.

Peptide In-solution REST Simulations

There has been little previous simulation work investigating the structure of P1 in solution. Pandey *et al.* used Monte Carlo simulations using coarse-grained force-fields to investigate the in-solution aggregation of multiple P1 peptides.[82] However, to the authors' knowledge there have been no reports of molecular simulation studies using advanced sampling to investigate the structure of single molecules of P1 or P1A3 in solution. Recent work on materials-binding peptides has shown their intrinsically disordered nature demands extensive conformational sampling.[18,22,25] At the same time it has been shown that replica exchange with solute tempering (REST) is a particularly effective approach for predicting the conformational ensemble of such peptides.[17,18,22,25]

In our cluster analysis, each distinct cluster is related to a set of thermally-accessible structures at 300 K, which we propose is related to a basin on the potential energy landscape. The population of each cluster (basin) is defined as the fraction of the 20,000 trajectory frames that are assigned to each cluster. In Fig. S6(a) we show the growth in the number of clusters as a function of REST MD steps. By 20 ns of REST simula-

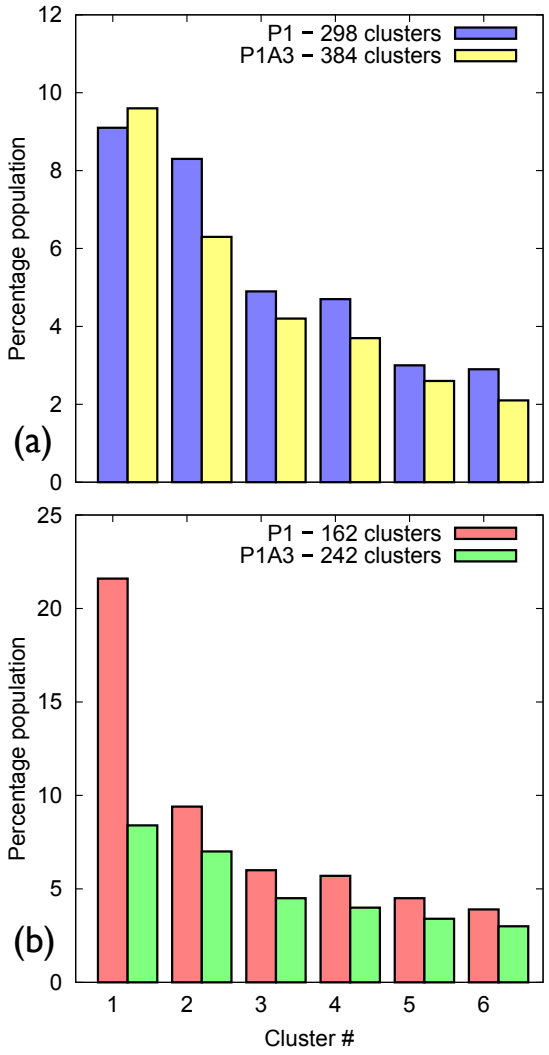

Fig. 2 Percentage population of the top six most highly-populated clusters of P1 and P1A3 (a) in solution and (b) adsorbed at the graphene interface.

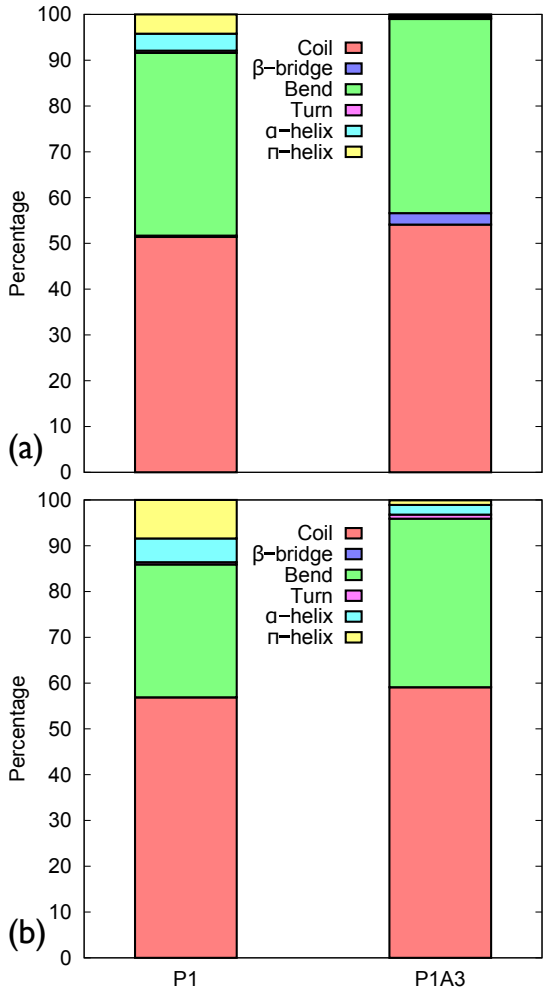

Fig. 3 The percentage of different secondary structure categories of the overall conformational ensemble of the peptides in solution determined via (a) analysis of the backbone dihedral angles and (b) using the DSSP definitions.

tion, the number of clusters has started to plateau indicating the approach of equilibration. Fig. 2(a) and Table S2† give the relative population of the most likely thermally-accessible structures (clusters) in the overall ensemble (referred to herein as the cluster population). The high total number of clusters (298 and 384 for P1 and P1A3 respectively), together with the fact that no single cluster accounts for over 10% of the population for either peptide, indicates that the neither peptide has one dominant conformation (or even two, or three conformations) that it will adopt in aqueous solution. Also, mutating the three aromatic residues to alanine has dramatically increased the number of thermally-accessible conformations of the peptide.

The apparently intrinsically disordered character of the peptides was confirmed by analysis of their secondary structure, as shown in Figs. 3(a) and S7(a)†. However, it is important to note that helical character (*i.e.* from the Ramachandran analysis) does not necessarily translate into helical *structure*. Using the secondary structure definitions defined based on the ϕ and ψ angles, Fig. S7(a)†, there is no single secondary structure category that dominates the conformational ensemble of either peptide. If the more rigidly-defined DSSP terminology is used, Fig. 3(a), then the coil and bend motifs dominate both P1 and P1A3. In general, the secondary structure motifs are similar for the both the original peptide and the mutant when free in solution. However, P1 features a relatively greater amount of helical (α and π) character. Indeed, the top cluster of P1 (Fig. 4(a)) featured a helical structure located in the centre of the peptide. In contrast, none of the top ten

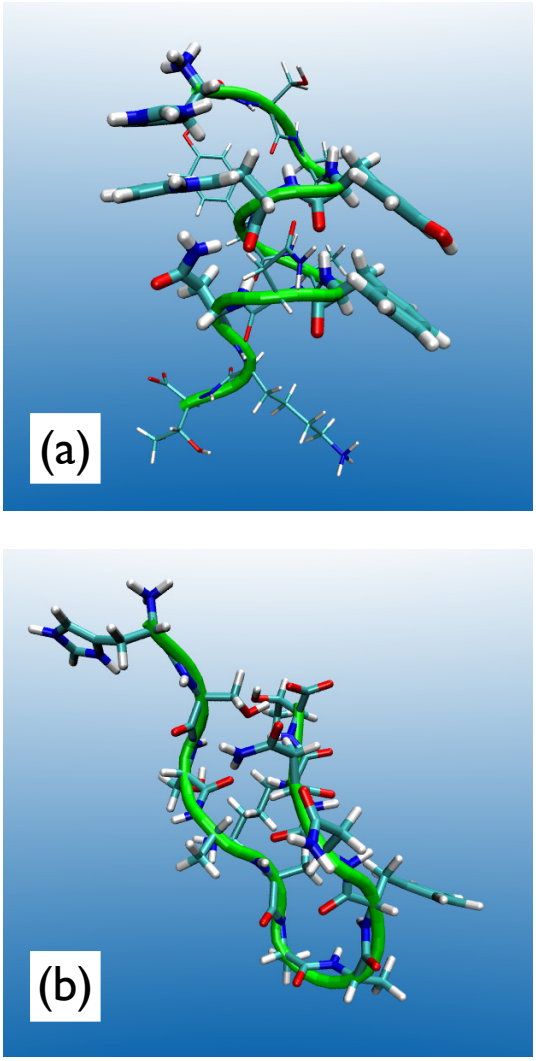

Fig. 4 Snapshots of the most populated clusters of (a) P1 (9.1%) and (b) P1A3 (9.6%) in solution.

most populated clusters of P1A3 showed any significant helical structure. Considering that alanine is often proposed as a helix-promoting residue, the fact that the original peptide sequence possesses a greater degree of helical character (and helical structure) than the mutant may seem counter-intuitive. However, the snapshot of P1 shown in Fig. 4(a) indicates a plausible reason why the helical form of P1 is more stable than that of P1A3. The spacing of the aromatic groups in the P1 sequence facilitates favourable interactions between these groups when this peptide assumes a helical structure. Thus, the helix configuration of P1 appears to be stabilised through π-π interactions, in addition to the usual backbone/backbone hydrogen bonds. These interactions cannot, by definition, be supported in the mutant sequence.

To investigate the intra-peptide interactions present in the helical configuration of P1, a 10 ns regular MD simulation was performed using the helical configuration as the starting structure. Previous simulations of another intrinsically disordered peptide have found that a distance of < 6 Å is indicative of non-covalent interaction.[17] Table S3 in the ESI†gives the fraction of the trajectory that the distance between the aromatic groups is within this cutoff for P1 in solution, for both the REST simulation and the 10 ns standard MD simulation. From these data it is apparent that there are π-π stacking interactions between His1-Trp5 and Tyr4-Phe8 in the helical configuration of P1. Table S4 in the ESI† gives the intra-peptide hydrogen bonding, obtained for the reference replica trajectory of both P1 and P1A3 as well as for a 10 ns regular MD simulation of the top cluster of P1. On average there is a greater number of hydrogen-bonds present in the helical configuration of P1 than in the full conformational ensemble of either P1 and P1A3. This further indicates that the helical configuration of P1 is stabilised by favourable intra-peptide interactions. It should also be noted that on average P1A3 forms as many H-bonds as P1, in contrast to P1, however, the H-bonding in P1A3 is via side-chain/side-chain and backbone/side-chain hydrogen bonds rather than backbone-backbone hydrogen bonds. This is consistent with the hairpin-like structure identified as the top cluster of P1A3 (Fig. 4(b)).

The most populated cluster of P1 (the helical configuration) still only accounts for $< 10\%$ of the entire conformational ensemble. The majority of P1 configurations, like those of P1A3, are in general extended configurations. However, the presence of this helical structure highlights the need to use advanced sampling techniques such as REST to study instrinsically disordered peptides. Even a very long 'brute force' MD simulation might not have revealed this minor, but significant, configuration. Results of previous REST simulations of materials-binding peptides, both free in solution, and when adsorbed at aqueous gold, silver and quartz interfaces, have not featured the strong helical content as observed in the case of P1. In these previous studies, as in the present work, we carried out REST simulations where the initial configurations of the replicas spanned a range of different secondary structural motifs, including the initial configuration of a single replica possessing an α-helix structure. On the basis of this evidence, we believe that the relative populations of the different secondary structural motifs of the initial structures is not likely to be responsible for the degree of helical content in the configurational ensemble of P1. However, further systematic studies on the effect of the initial configurations on the final ensemble populations of adsorbed peptides may be warranted to rigourously confirm this.

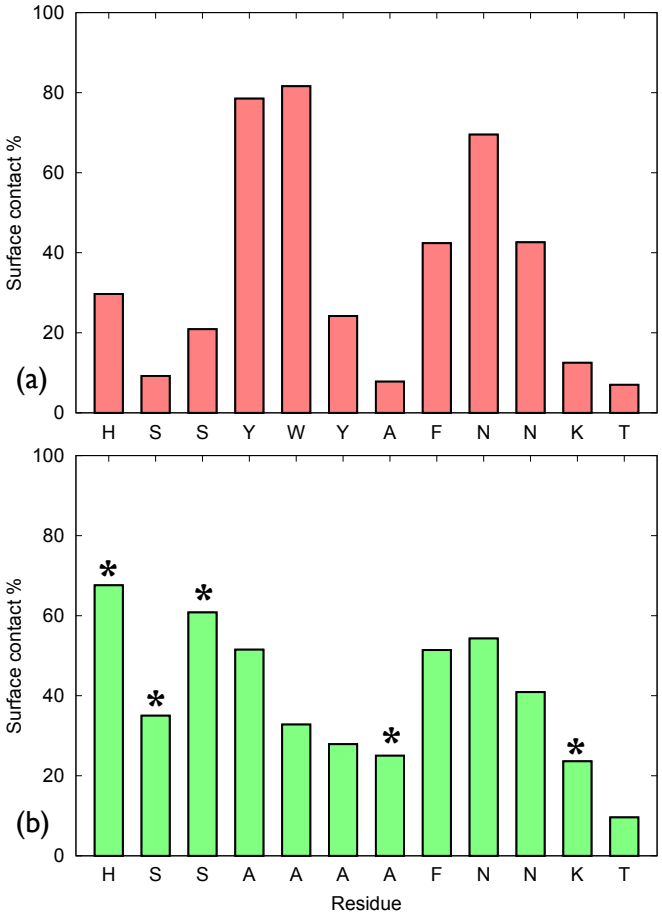

Fig. 5 The degree of contact for each peptide residue with the graphene surface for (a) P1 and (b) P1A3. Those residues marked by stars show an increase in contact of $\geq$10% when mutated from P1 to P1A3.

Adsorbed Peptide REST Simulations

Fig. 5 shows the contact fraction of each residue with the aqueous graphene interface for both P1 and P1A3. In Figs. S8 and S9 in the ESI† we provide the probability distribution of distance from the surface for each side-chain site. In the case of P1 the Y4, W5 and N9 residues are strongest contact points, with F8 and N10 also showing significant propensity to make contact with the graphene sheet. Our corresponding calculations of the AA adsorption free energy indicated Tyr, Trp, Phe and Asn as being strong binders (see Fig. 1 and Table 1). However, Y6 is not contact with the surface for large proportion of the trajectory, despite its substantial binding free energy in *amino acid* form, highlighting the fact that the behaviour of residues within a peptide depends upon the sequence and structure of the peptide as well as the individual constituent residues. As expected, the mutation of the three aromatic residues to alanine in P1A3 changes the distribution of contact points in the peptide. The fraction of the trajectory that A6 and A7 were in contact with the surface was significantly reduced compared to Y6 and W7 in P1. In contrast, H1 in P1A3 increased its binding propensity for the graphene interface, becoming the strongest contact point for the mutant peptide. Residues S2 and S3 also showed an increase in their surface contact fraction. In going from P1 to P1A3, the changes in contact distribution of the C-terminal half of the peptide are less pronounced than those seen for the first six residues, suggesting that the mutation imparts a localised effect on the binding. The distribution of distances of the side-chain sites from the graphene interface (Figs. S8 and S9†) is quite similar for residues 7-12. In contrast, the distributions of residues 1-6 undergo greater change, even for H1, S2 and S3, where these residues are common to both peptide sequences. Overall, the degree of residue-surface contact in P1A3 shows less variation difference between the strongest and weakest binding residues, compared with P1.

Fig. S6(b)† shows the number of distinct clusters found over the course of the simulation for the adsorbed peptides. In comparison with the solution simulations, the number of clusters found for the adsorbed peptides is reduced, by 35-45 % (in the adsorbed state there are 162 and 242 clusters for P1 and P1A3 respectively). The population distribution of the most likely structures is shown in Fig. 2(b) (see also Table S2 in the ESI†). The behaviour of P1A3 is similar to that when in solution, with the population of no single cluster exceeding 10%, and showing a gradual decay in population of the clusters. In the case of P1 a different behaviour is observed, with the top ranked conformation accounting for 21.6% of the entire conformational ensemble. Apart from this top cluster, the population distribution of P1 is similar to that of P1A3.

Snapshots showing the top three clusters of P1 when adsorbed on graphene are shown in Fig. 6. Remarkably, the top ranked cluster is a helical structure, while clusters two and three (indeed the majority of clusters) are in extended configurations. Thus, while the majority of P1 peptides adsorbed on graphene have extended configurations, there will be a noticeable fraction of the ensemble that will be present in a helical type configuration. We propose that the 'anomalous' population of cluster 1 could be explained by the fact that the helical configuration in the top cluster places Y4, W5, F8 and N9 all on one facet of the helix, thus locating all four residues (which are also identified by our AA calculations as strong binders) in contact with the surface collectively, yielding four strong anchoring points. As a consequence of this spatial alignment, H1 and Y6 are also arranged together on the upper facet of the helix (facing away from the graphene surface), facilitating a π-π stacking interaction between the two. Thus, the helical configuration found as the top cluster is likely to posses both favourable peptide-graphene binding interactions and favourable intra-peptide interactions. This configuration

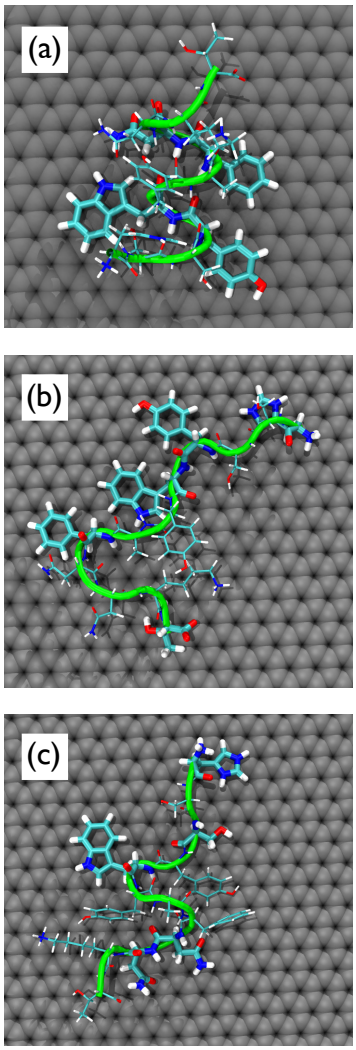

Fig. 6 The three most populated clusters of P1 adsorbed on graphene, (a) cluster 1 (21.6%), (b) cluster 2 (9.4%) and (c) cluster 3 (6.0%).

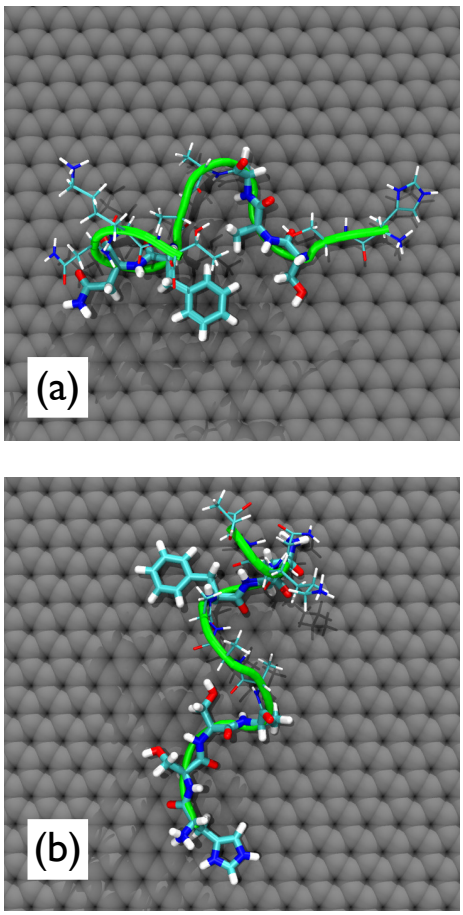

Fig. 7 The two most populated clusters of P1A3 adsorbed on graphene, (a) cluster 1 (8.4%) and (b) cluster 2 (8.3%).

also partly explains the very different contact behaviour seen for the two tyrosine residues (Fig. 5), since the helical configuration, by construction, assures strong surface contact for Y4 and impedes contact for Y6. In the case of P1A3 (Fig. 7) the top ten clusters all have extended configurations, with no significant helical structure observed. The increase in binding propensity of H1 in P1A3 can be explained by the fact that the P1A3 peptide does not support helical structures, thus, in constrast with P1, enabling H1 to make substantial surface contact.

Analysis of the secondary structure of the two peptides at the graphene interface is shown in Figs. 3(b) and S7(b)†. The secondary structure defined by the dihedral backbone angles, Fig. S7(b)†, is similar to that found for the peptides free in solution. The α, β, PPII and miscellaneous motifs are all found to be present with no one motif dominating for either P1 or P1A3. For the secondary structure defined using DSSP, changes between the solution and adsorbed peptides are more apparent, Fig. 3(b), particularly in the case of P1. As for the peptides in solution, the most common motif was coil, followed by bend; however, the amount of helical content (α-helix and π-helix) increased for both peptides, particularly for P1. This reinforces the results of the cluster analysis, in that while most of the conformational ensemble of P1 and P1A3 is associated with extended configurations, P1 has a small but significant contribution associated with a helical structure.

The REST simulations of P1 on graphene reveal that the four-residue Y4, W5, F8, N9 motif is able to stabilise a helical configuration at the aqueous graphene interface. Referring back to our adsorption free energy results, the in-plane alignment of Y4, W5, F8, N9 assures that all four side-chains are in close contact with the surface and avoid the highly-structured first interfacial solvent layer. On the other hand, in this con-

figuration H1 and Y6 reside on the opposite face of the helix, and thus also avoid this first water layer, being placed above it instead. While this configuration only accounts for $\sim 20\%$ of the population of adsorbed peptide configurations, and the majority of configurations were extended, it is not a negligible contribution to the ensemble. The mutation of Y4 and W5 in P1A3 makes an adsorbed helical configuration less favourable and therefore far less helical content is observed for P1A3 (Fig. 7). Instead H1 becomes a major anchor point for the peptide, as does S3 to a lesser extent.

As mentioned previously, although P1 is thought to be a stronger binder than P1A3, the experimental free-energy of adsorption of either peptide sequence has not yet been reported. While these REST simulations cannot in themselves provide an estimate of the free energy of adsorption for the two peptides, it is possible to estimate the entropic and enthalpic contributions to the free-energy of adsorption.[18,22] As reported previously,[18] the conformational entropy of a peptide can be calculated from

$$S_{\mathrm{conf}} = -\sum_{i=1}^{n} p_i \ln(p_i) \quad (1)$$

where n is the total number of clusters for a given system and p_i is the population (fraction) of configurations that belong to the i^{th} cluster. Thus, the larger S_{conf} is, the greater the number of thermally-accessible structures supported by that peptide. The enthalpic binding score, Ω, is a measure of the enthalpic contribution to the binding, as has been introduced previously.[18,22] Here, we define Ω as:

$$\Omega = \sum_{i=1}^{12} c_i \Delta F E_{\mathrm{ads}}^{\mathrm{AA}} \quad (2)$$

where c_i is the fraction of the trajectory that each residue spends in contact with the surface, and $\Delta F E_{\mathrm{ads}}^{\mathrm{AA}}$ is the free energy of adsorption of the associated amino acid. Thus, the more negative the value of Ω, the greater the enthalpic contribution to the binding at the interface. The theoretical limit of Ω (in the extremely unlikely situation of every residue being in perfect contact with the surface over the whole trajectory) for P1 and P1A3 is -181 and -149 kJ mol^{-1} respectively. Table 3 gives the values of these properties for both peptides. In the adsorbed state, as expected, P1A3 has a much larger conformational entropy contribution compared with P1. However, Ω is more negative for P1 than for P1A3, implying that the difference in the conformational entropy contribution may be offset by the more favourable enthalpic contribution to the binding.

Over the last few years, there has been substantial consolidation in the knowledge base regarding studies of the molecular simulation of peptides adsorbed to different solid substrates. Despite this, clear guidelines that would allow the identification of the precise sequence of a peptide, guaranteed to exhibit preferrential adsorption to one material over another, remain unresolved. Despite this, some of the factors that affect selective binding across different substrates are emerging. The free energies of adsorption of AAs to aqueous gold, silver and graphene interfaces have been reported, allowing some cross-substrate comparisons to be made, at least on the level of AAs. On this basis, the strongest comparisons between graphene and other substrates can be made for gold, since the gold/peptide interface has been the most extensively studied of this set. Arg, Trp and Tyr adsorb strongly to both gold and graphene interfaces; therefore, we suggest that none of these residues would be useful in constructing a peptide sequence that is selective for binding to gold over graphene or *vice versa*. In contrast, Gly, Asn and Gln are predicted to adsorb more strongly to graphene compared with gold. Thus, we suggest that the inclusion of these residues in a peptide sequence may confer peptide adsorption selectivity of graphene over gold interfaces. However, we reiterate here that the adsorption behaviour of a peptide cannot simply be extrapolated from its constituent residues alone. Clear identification of the sequence traits that can deliver peptide-materials binding selectivity requires further systematic studies, from both experimental and molecular simulation perspectives.

Table 3 The enthalpy binding score and conformational entropies of the two graphene binding peptides

Peptide	Ω/kJ mol^{-1}	S_{ads}
P1	-74	3.51
P1A3	-64	4.14

Conclusions

We have combined GRAPPA, a polarisable FF specifically developed for modelling biomolecule-graphitic interactions, with metadynamics and REST simulations to elucidate clear links between the sequence, conformations and graphene-binding propensity of two experimentally-characterised peptide sequences. Using metadynamics, we predicted adsorption free energies for all twenty naturally-occuring amino acids, providing a benchmark against which the adsorption of residues within a peptide can be compared. All amino acids showed affinity for the graphene surface in a direct, rather than solvent mediated, mode. The strongest binders were found to be Arg, Tyr, Trp and Gln and the weakest Ile. A key finding is that amino acids with flat (or compact) side-chain groups that can avoid extensive contact with the first, highly-structured water layer when adsorbed at the interface have favourable binding, while amino acids with bulky side-chains shown concomittantly weaker binding. REST simulations of the P1 and P1A3 peptides both adsorbed at the graphene interface and while free in solution revealed that both peptides are intrin-

sically disordered, with extended configurations dominating their conformational ensembles. However, a minor, but significant, part of the conformational ensemble of P1 was assigned to a helical configuration, both in-solution and when adsorbed. The stability of this helical configuration in solution was explained by favourable intra-peptide interactions. When adsorbed on the graphene surface, the helical P1 structure presented four strongly-binding residues (Y4, W5, F8, N9), on one facet of the helix, placing these in contact with the substrate, providing additonal stablisation. P1A3 showed diminished contact with the interface in the immediate region of the mutation sites, while other residues in the sequence showed enhanced surface contact. The mutations in positions 4 and 5 appeared to abolish the propensity for the peptide to support a helical structure, with extended structures dominating in the adsorbed state for P1A3. Our findings highlight how the structure of adsorbed peptides depends not only on the residues present but also their local environment. These connections between the sequence and surface-binding of peptides to their structure(s) open a viable route to the development of *de novo* designed peptides for specific applications.

Acknowledgements

The authors thank the Victoria Life Sciences Computational Initiative (VLSCI) and the National Computing Infrastructure (NCI) for provision of computational resources. ZEH and TRW thank **veski** for research funding, and TRW thanks **veski** for an Innovation Fellowship.

References

1 K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva and A. A. Firsov, *Science*, 2004, **306**, 666–669.
2 Z. Liu, S. M. Tabakman, Z. Chen and H. Dai, *Nat Protoc.*, 2009, **4**, 1372–1381.
3 A. Kalra, S. Garde and G. Hummer, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 10175–10180.
4 Y. Liu, X. Dong and P. Chen, *Chem. Soc. Rev.*, 2012, **41**, 2283–2307.
5 M. S. Mannoor, H. Tao, J. D. Clayton, A. Sengupta, D. L. Kaplan, R. R. Naik, N. Verma, F. G. Omenetto and M. C. McAlpine, *Nature Comm.*, 2012, **3**, 763.
6 Y. X. Fang and E. K. Wang, *Chem. Commun.*, 2013, **49**, 9526–9539.
7 Y. Zhang, T. R. Nayak, H. Hong and W. B. Cai, *Nanoscale*, 2012, **4**, 3833–3842.
8 H. C. Zhang, G. Grüner and Y. L. Zhao, *J. Mater. Chem. B*, 2013, **1**, 2542–2567.
9 D. R. Samarajeewa, G. R. Dieckmann, S. O. Nielsen and I. H. Musselman, *Nanoscale*, 2012, **4**, 4544–4554.
10 B. Akdim, R. Pachter, S. S. Kim, R. R. Naik, T. R. Walsh, S. Trohalaki, G. Hong, Z. Kuang and B. L. Farmer, *ACS Appl. Mater. Inter.*, 2013, **5**, 7470–7477.
11 Y. Cui, S. N. Kim, S. E. Jones, L. L. Wissler, R. R. Naik and M. C. McAlpine, *Nano Lett.*, 2010, **10**, 4559–4565.
12 C. R. So, Y. Hayamizu, H. Yazici, C. Gresswell, D. Khatayevich, C. Tamerler and M. Sarikaya, *ACS Nano*, 2012, **6**, 1648–1656.
13 X. Mao, Y. Wang, L. Liu, L. Niu, Y. Yang and C. Wang, *Langmuir*, 2009, **25**, 8849–8853.
14 X. Mao, Y. Guo, Y. Luo, L. Niu, L. Liu, X. Ma, H. Wang, Y. Yang, G. Wei and C. Wang, *J. Am. Chem. Soc.*, 2013, **135**, 2181–2187.
15 Z. Li, T. Uzawa, T. Tanaka, A. Hida, K. Ishibashi, H. Katakura, E. Kobatake and Y. Ito, *Biotech. Lett.*, 2013, **35**, 39–45.
16 P. A. Mirau, R. R. Naik and P. Gehring, *J. Am. Chem. Soc.*, 2011, **133**, 18243–18248.
17 A. H. Brown, P. M. Rodger, J. S. Evans and T. R. Walsh, *Biomacromol.*, 2014, **15**, 4467–4479.
18 J. P. Palafox-Hernandez, Z. Tang, Z. E. Hughes, Y. Li, M. T. Swihart, P. N. Prasad, T. R. Walsh and M. R. Knecht, *Chem. Mater.*, 2014, **26**, 4960–4969.
19 R. H. Meissner, J. Schneider, P. Schiffels and L. Colombi Ciacchi, *Langmuir*, 2014, **30**, 3487–3494.
20 K. A. Ball, A. H. Phillips, D. E. Wemmer and T. Head-Gordon, *Biophs. J.*, 2013, **104**, 2714–2724.
21 K. Ostermeir and M. Zacharias, *Biochim. Biophys. Acta*, 2013, **1834**, 847–853.
22 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.
23 A. Sethi, J. Tian, D. M. Vu and S. Gnanakaran, *Biophys. J.*, 2012, **103**, 748–757.
24 C. Narayanan, D. S. Weinstock, K. P. Wu and R. M. Levy, *J. Chem. Theory Comput.*, 2012, **8**, 3929–3942.
25 L. B. Wright and T. R. Walsh, *Phys. Chem. Chem. Phys.*, 2013, **15**, 4715–4726.
26 J. Schneider and L. Colombi Ciacchi, *J. Am. Chem. Soc.*, 2012, **134**, 2407–2413.
27 K. Moritsugu, T. Terada and A. Kidera, *J. Amer. Chem. Soc.*, 2012, **134**, 7094–7101.
28 J. Mittal, T. H. Yoo, G. Georgiou and T. M. Truskett, *J. Phys. Chem.. B*, 2013, **117**, 118–124.
29 M. Knott and R. B. Best, *PLoS Comp. Biol.*, 2012, **8**, e1002605–1–10.
30 A. A. Skelton, T. Liang and T. R. Walsh, *ACS Appl. Mater. Interfaces*, 2009, **1**, 1482–1491.
31 R. A. Latour, *Biointerphases*, 2008, **3**, FC2–FC12.
32 P. Liu, B. Kim, R. A. Friesner and B. J. Berne, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 13749–13754.
33 L. Wang, R. A. Friesner and B. J. Berne, *J. Phys. Chem.. B*, 2011, **115**, 9431–9438.
34 T. Terakawa, T. Kameda and S. Takada, *J. Comput. Chem.*, 2010, **32**, 1228–1234.
35 L. Wang, Y. Deng, J. L. Knight, Y. Wu, B. Kim, W. Sherman, J. C. Shelley, T. Lin and R. Abel, *J. Chem. Theory Comput.*, 2013, **9**, 1282–1293.
36 S. R. Friling, R. Notman and T. R. Walsh, *Nanoscale*, 2010, **2**, 98–106.
37 S. N. Kim, Z. Kuang, J. M. Slocik, S. E. Jones, Y. Cui, B. L. Farmer, M. C. McAlpine and R. R. Naik, *J. Am. Chem. Soc.*, 2011, **133**, 14480–14483.
38 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.
39 R. B. Pandey, Z. Kuang, B. L. Farmer, S. S. Kim and R. R. Naik, *Soft Matter*, 2012, **8**, 9101–9109.
40 A. N. Camden, S. A. Barr and R. J. Berry, *J. Phys. Chem. B*, 2013, **117**, 10691–10697.
41 N. Dragneva, W. B. Floriano, D. Stauffer, R. C. Mawhinney, G. Fanchini and O. Rubel, *J. Chem. Phys.*, 2013, **139**, 174711.
42 M. Mijajlovic, M. J. Penna and M. J. Biggs, *Langmuir*, 2013, **29**, 2919–2926.
43 N. Todorova, A. J. Makarucha, N. D. M. Hine, A. A. Mostofi and I. Yarovsky, *PLoS Comp. Biol.*, 2013, **9**, e1003360.
44 S. Sheikholeslami, R. B. Pandey, N. Dragneva, W. Floriano, O. Rubel,

S. A. Barr, Z. Kuang, R. Berry, R. Naik and B. Farmer, *J. Chem. Phys.*, 2014, **140**, 204901.
45 T. A. Ho and A. Striolo, *J. Chem. Phys.*, 2013, **138**, 054117.
46 J. Sala, E. Guardia and J. Martí, *Phys. Chem. Chem. Phys.*, 2012, **14**, 10799–10808.
47 X. Zhao and J. K. Johnson, *Mol. Simul.*, 2005, **31**, 1–10.
48 D. Y. Lu, Y. Li, U. Ravaioli and K. Schulten, *J. Phys. Chem. B*, 2005, **109**, 11461–11467.
49 S. D. Tomásio and T. R. Walsh, *Mol. Phys.*, 2007, **105**, 221–229.
50 S. M. Tomásio and T. R. Walsh, *J. Phys. Chem. C*, 2009, **113**, 8778–8785.
51 Z. E. Hughes, S. M. Tomásio and T. R. Walsh, *Nanoscale*, 2014, **6**, 5438–5448.
52 T. R. Walsh and S. M. Tomasio, *Mol. BioSyst.*, 2010, **6**, 1707–1718.
53 R. H. Zhou and B. J. Berne, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 12777–12782.
54 R. H. Zhou, G. Krilov and B. J. Berne, *J. Phys. Chem. B*, 2004, **108**, 7528–7530.
55 C. Tan, L. Yang and R. Luo, *J. Phys. Chem. B*, 2006, **110**, 18680–18687.
56 A. Huang and C. M. Stultz, *Biophys. J.*, 2007, **92**, 34–45.
57 M. S. Shell, R. Ritterson and K. A. Dill, *J. Phys. Chem. B*, 2008, **112**, 6878–6886.
58 F. Godschalk, S. Genheden, P. Soderhjelm and U. Ryde, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7731–7739.
59 M. Kolar, J. Fanfrlik, M. Lepsik, F. Forti, F. J. Luque and P. Hobza, *J. Phys. Chem. B*, 2013, **117**, 5950–5962.
60 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, 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. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
61 S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Biophys. J.*, 2011, **100**, L47–L49.
62 A. Laio and M. Parrinello, *Prot. Natl. Acad. Sci. USA*, 2002, **99**, 12562–12566.
63 G. Cicero, J. C. Grossman, E. Schwegler, F. Gygi and G. Galli, *J. Am. Chem. Soc.*, 2008, **130**, 1871–1878.
64 M. K. Rana and A. Chandra, *J. Chem. Phys.*, 2013, **138**, 204702.
65 C. Calero, J. Marti, E. Guardia and M. Masia, *J. Chem. Theory Comput.*, 2013, **9**, 5070–5075.
66 S. Corni, M. Hnilova, C. Tamerler and M. Sarikaya, *J. Phys. Chem. C*, 2013, **117**, 16990–17003.
67 G. Nawrocki and M. Cieplak, *J. Phys. Chem. C*, 2014, **118**, 12929–12943.
68 M. J. Pender, L. A. Sowards, J. D. Hartgerink, M. O. Stone and R. R. Naik, *Nano Letters*, 2006, **6**, 40–44.
69 Y. Cui, S. N. Kim, R. R. Naik and M. C. McAlpine, *Acc. Chem. Res.*, 2012, **45**, 696–704.
70 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
71 M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci and R. A. Broglia, *Comput. Phys. Comm.*, 2009, **180**, 1961–1972.
72 A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2008, **100**, 020603.
73 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
74 E. Neria, S. Fischer and M. Karplus, *J. Phys. Chem.*, 1996, **105**, 1902–1921.
75 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
76 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem. Int. Ed.*, 1999, **38**, 236–240.
77 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
78 H. Heinz, B. L. Farmer, R. B. Pandey, J. M. Slocik, S. S. Patnaik, R. Pachter and R. R. Naik, *J. Am. Chem. Soc.*, 2009, **131**, 9704–9714.
79 M. J. Penna, M. Mijajlovic and M. J. Biggs, *J. Am. Chem. Soc.*, 2014, **136**, 5323–5331.
80 Z. E. Hughes, L. B. Wright and T. R. Walsh, *Langmuir*, 2013, **29**, 13217–13229.
81 M. Rosa, S. Corni and R. Di Felice, *J. Chem. Theory Comput.*, 2014, **10**, 1707–1716.
82 R. B. Pandey, Z. Kuang and B. L. Farmer, *PLoS ONE*, 2013, **8**, e70847.