Side-chain effects on the co-existence of emergent nanopatterns in amino acid adlayers on graphene

Journal:	*Nanoscale*
Manuscript ID	NR-ART-02-2020-001333.R1
Article Type:	Paper
Date Submitted by the Author:	18-May-2020
Complete List of Authors:	Awuah, Joel; Institute for Frontier Materials, Walsh, Tiffany; Deakin University, Institute for Frontier Materials

SCHOLARONE™
Manuscripts

Journal Name

Side-chain effects on the co-existence of emergent nanopatterns in amino acid adlayers on graphene[†]

Joel B. Awuah,[a] and Tiffany R. Walsh[*a]

The spontaneous tendency of amino acid adlayers to self-assemble into ordered patterns on non-reactive surfaces is thought to be chiefly influenced by amino acid termination state. Experiments have shown that different side chains can produce different patterns, with a distinction drawn between side chains that can support hydrogen bonds or electrostatic interactions, and those that are hydrophobic. However, as is demonstrated in this work, this distinction is not clear cut, implying that there is currently no way to predict in advance what type of pattern will be formed. Here, we use molecular dynamics simulations of amino acid adlayers in neutral, zwitterion, and neutral-zwitterion states for two types of amino acids, either histidine or alanine, adsorbed at the *in-vacuo* graphene interface. The findings from these simulations yield a framework of predictive rules for anticipating the types of patterns that a given amino acid can produce. In contrast to earlier studies on adlayers of tryptophan and methionine on graphene that reveal the presence of only a single type of pattern motif, the canonical dimer row, here we find that emergent patterns of histidine and alanine adlayers supported the co-existence of several different types of motifs, influenced by the different side-chain characteristics. For alanine, the compact side-chain does not support hydrogen bonding and engages weakly with the surface, leading to the emergence of a new dimer row configuration in addition to the canonical dimer row motif. On the contrary, for histidine, the side-chain supports hydrogen bonding, leading to the emergence of a dimer row motif different from the canonical dimer row, co-existing with several different monomer row motifs. On this basis, we propose that emergent canonical dimer row patterns are more likely for amino acids with side-chains that are non-compact and that also lack extensive hydrogen bonding capacity, and that engage strongly with the underlying substrate. These findings provide a fundamental basis to rationally guide the design of desired self-assembled nanostructures on planar surfaces.

1 Introduction

The self-organisation of biomolecules adsorbed on solid surfaces into ordered nanopatterns is a promising route for the development of a wide-range of graphene-based nano-devices, including field-effect transistors (FET) and optical biosensors, and for diagnostic and therapeutic strategies.[1–8] The ability to control this nanostructuring and non-covalent functionalisation of graphene surfaces is a fundamental requirement for achieving this. A key step required to facilitate this control is to elucidate the elementary mechanisms influencing the self-assembly of biomolecules at the graphene interface.[1–3,5,8–10] Small biomolecules, such as amino acids, representing the basic unit of peptides and proteins, are ideal structures to obtain fundamental insights into this self-ordering process.[9,11–24] Despite these considerable prospects, a deeper understanding of the fundamental mechanisms controlling biomolecule self-assembly at the molecular-level is still needed to enable these advances.[1–6,10]

Amongst the elementary mechanisms influencing pattern formation in amino acid adlayers deposited on non-reactive solid surfaces, several experimental and simulation studies have identified inter-molecular interactions as the primary driving force accounting for the formation of ordered supramolecular structures, albeit with additional influences from the molecule-surface interactions.[11,13–26] These inter-molecular interactions could be mediated *via* e.g. hydrogen bonding networks, charge-charge interactions, or other more specialised contacts between the molecules. Notably, inter-molecular interactions within amino acid adlayers have been reported to be strongly influenced by the termination

a Institute for Frontier Materials, Deakin University, Geelong, VIC 3216, Australia. Tel: +61 (0)3 5227 3116; E-mail: tiffany.walsh@deakin.edu.au.

state, which is based on the charges present at the N- and C-termini.[11,13,19,27] Generally, zwitterionic assembly *i.e.* mediated by charge-charge interactions, is experimentally considered as the construction motif for ordered patterns in amino acids adsorbed on unreactive substrates.[11,13,17–20,24] However, a recent simulation study has demonstrated that ordered patterns may not be driven solely by charge-charge interactions, but also any type of interaction involving a charge.[27]

Despite the identification of the amino acid's termination state as the primary driving force for ordering[11,13–28], the influence of other fundamental factors affecting pattern formation, such as the amino acid side-chains themselves, remain unclear.[17,19,29] Therefore, there is a need to further explore the spontaneous emergence of these ordered structures, in terms of the connection between the patterns formed and the different amino acid side-chain characteristics.

Amino acids can self-assemble into different types of 1D and 2D ordered nanostructures, including monomer rows, dimer rows and complex tetramer structures, upon adsorption on solid surfaces.[9,12] A combination of several experimental techniques, e.g. scanning tunnelling microscopy (STM) and near edge x-ray absorption fine structure spectroscopy (NEXAFS), have proven valuable in characterising the ordered patterns and structures formed on low-reactivity substrates.[11,13–20,24] The overall lateral arrangement of the amino acid adlayers can be probed *via* scanning tunnelling microscopy (STM),[11,13–20,24] and the orientation of the adsorbed individual molecules can in principle be determined using near edge x-ray absorption fine structure spectroscopy (NEXAFS).[19,21,30] Despite these experimental achievements, gaining molecular-level information on the functionalities of amino acid side-chains and their influence on the emergent structures that are spontaneously formed as a function of time can be challenging to accomplish *via* experiment alone.

Amino acids can be categorised on the basis of the different physico-chemical features of their side-chains, which may account for different binding modes and orientations of amino acids adsorbed on planar surfaces.[9,17,19,31–33] Therefore, amino acids with different side-chain characteristics might self-organise into different types of patterns on the same substrate. However, this might not always be the case, given that amino acids with distinct side-chain characteristics, for example tryptophan (Trp) and methionine (Met), have been identified to form similar canonical dimer row structures upon adsorption on unreactive surfaces.[11,17,27] By combining the outcomes of molecular simulations with existing experimental data, a more complete insight into the role of the side-chains on pattern formation can be determined, by revealing how different side-chains can influence both inter-molecule and molecule-surface interactions.

Regarding experimental efforts, pattern formation in amino acid adlayers on low-reactivity substrates, mainly noble metal surfaces and graphite, has been extensively studied.[11,13–20,24] More specifically, using STM, extended nanowires composed of canonical dimer rows of Met were observed in ultrahigh vacuum (UHV) conditions by Humblot *et al.* on Au(111)[13,18] and by Schiffrin *et al.* on Ag(111)[11], and by Riemann and Nelson on graphite under liquid octanol.[20] In all cases, the authors reported that the

Met adsorbed with the side-chains aligned parallel to the surface. Similarly, using STM and NEXAFS, Riechert *et al.* found that tyrosine (Tyr) formed highly-ordered 2D rows induced by molecular dimerisation on the Ag(111) surface under UHV conditions.[19] However, their NEXAFS data showed that the aromatic side-chains were tilted with reference to the surface, resulting in π-π stacking interactions between adjacent molecules. This accounted for interdigitation between adjacent Tyr side-chains within the long-ranged rows, contrary to that of Met on Ag(111) under UHV conditions.[11,19]

Additionally, Krebs *et al.* reported chevron-like structures comprising highly-ordered monomer rows for both histidine (His) and Tyr adlayers on graphite in octanol using STM,[17] contrary to the findings obtained for Met adsorbed on graphite in liquid octanol.[20] Krebs *et al.* attributed these adsorption behaviours to the dipole moment of the side-chains of His and Tyr, which can alter the inter-molecular interactions such as the preferred head-to-head (C-terminus/N-terminus) to head-to-tail (terminus/side-chain arrangements).[17] In other studies, Smerieri *et al.* used STM and electron spectroscopies to characterise the self-assembly of S-Glumatic acid (S-Glu) on the Ag(100) surface in UHV conditions.[14] The authors observed that S-Glu self-organised into several distinct patterns at different temperatures, and that the hydrogen bonds formed *via* head-to-head, tail-to-tail or head-to-tail interactions stabilised the adsorbed nanostructures.[14] However, in this instance, the various and systematic permutational possibilities of the different protonation states of both the C-terminus and the side-chain could be extremely complex. In summary, these studies suggest that the amino acid side-chain could be influential in the type of adsorbed structures that spontaneously emerge on solid surfaces. Nonetheless, while experimental efforts have shown that different amino acids can indeed produce different patterns, there is presently no clear link to enable a prediction of which type of pattern is most likely to appear. Molecular dynamics (MD) simulations offer a complementary promising pathway to obtaining these details.[8,34,35]

Focusing on molecular simulations of amino acids at graphene/graphite interfaces, several theoretical studies have reported the influence of amino acid side-chains on their adsorption modes and configurations. For example, amino acids with aromatic and extended side-chains, including Trp, Tyr, Met and Arginine (Arg) are reported to bind to graphene mainly *via* the side-chain, which is typically oriented flat on the surface.[20,31,33,36–41] In contrast, other amino acids, such as Glycine (Gly), Alanine (Ala) and Aspartic acid (Asp) were predicted to adsorb on graphene *via* the amide backbone and were also aligned parallel to the surface.[31–33,36–38] However, in most of these theoretical studies, only a single isolated amino acid and/or dimers adsorbed on graphene/graphite were considered.[20,31,33,36–41] While valuable, such studies cannot bridge the knowledge gaps related to adlayer patterning. Instead, a model of adsorbed adlayers on graphene surface is required to address these questions.[27]

In this regard, reports on modelling the dynamics of amino acid adlayers on non-reactive solid surfaces in vacuum are relatively few in number.[26,27,35,42] A recent MD simulation study by Awuah and Walsh[27] reported the influence of termination state

on the emergence of pattern formation in both Trp and Met ad-layers at the graphene interface in vacuum. These simulations revealed the spontaneous emergence of canonical dimer row structures in zwitterionic-containing adlayers. The predicted feature dimensions of these dimer rows were consistent with STM data of Met on Ag(111) and Au(111),[11,13,18] and Tyr on Ag(111) in UHV conditions.[19] However, this study did not distinguish the influence of the side-chains on the type of patterns formed, because both amino acids featured similar side-chain characteristics; i.e. they contained an extended and hydrophobic side-chain group, with this side-chain possessing no/low capacity for contributing to directional inter-molecular interactions (i.e. hydrogen bonds).

On the basis of previous experiments and simulation data, and our current work described herein, several rules regarding the type of emergent pattern are proposed. First, it is hypothesised that amino acids that have side-chains i) with no capacity to support hydrogen bonding, and ii) are not compact, will support the spontaneous emergence solely of canonical dimer rows. Furthermore it is suggested that for amino acids with compact side-chains that i) do not engage strongly with the surface and ii) have no capacity to hydrogen bond, will likely self-organise into several co-existing motifs. Moreover, if the amino acid contains a side-chain that supports hydrogen bonding interactions, then it is hypothesised that this amino acid will spontaneously self-organise into several co-existing emergent patterns.

Here, we seek further insights into this influence, using MD simulations. To be clear, we do not seek to evaluate if the emergent structures mature into longer-ranged patterns, but seek to focus on the *spontaneous emergence* of these nanopatterns. For comparison, we have modelled amino acid adlayers comprising either neutral, or zwitterionic species (referred to herein as pure systems), or a 1:1 mixture of neutral and zwitterionic states (referred to herein as mixed systems) comprising either histidine (His) adlayers or alanine (Ala) adlayers, adsorbed on a graphene surface *in vacuo*, using simulated annealing molecular dynamics (SAMD) simulations, using the GRAPPA force-field[31] combined with CHARMM22* force-field[43,44]. These amino acids feature contrasting side-chain characteristics, *i.e.* His contains a positively charged imidazole functional group side-chain that can support hydrogen bonding interactions, whereas Ala features a relatively compact methyl group side-chain that does not support hydrogen bonding, and engages weakly with graphene. We also evaluate the effect of the presence of trace water under UHV conditions. Additionally, as a preliminary test of our hypothesis, we modelled zwitterionic glycine (Gly) adlayers adsorbed on a graphene surface *in vacuo*.

2　Computational Methods

All work presented here was carried out using GROMACS v5.1.3.[45] Two sets of simulations of graphene interfaces were considered; the anhydrous graphene/amino acid interface, and the graphene/trace-water/amino acid interface containing trace amounts of water, herein denoted as 'dry' and 'wet' interfaces respectively. Both systems contained a single graphene sheet (~128 Å × ~120 Å) and an amino acid adlayer comprising 200 amino acids. All amino acids were initially positioned in randomised

lateral positions, at an initial vertical distance ($C_{graphene}$–C_{α}) of 5 Å from the graphene surface. In the wet system a monolayer of 200 water molecules (randomly arranged in the plane) separated the amino acid adlayer and graphene sheet, at a $C_{graphene}$–O_{water} distance of 3 Å. These initial spatially-randomised arrangements of the amino acid adlayers and water monolayer, as shown in Fig. S1 and S2 in the ESI†, were generated using PACKMOL.[46,47] The GRAPPA force-field[31] was used to model the bio-graphene interactions, and the amino acids and water molecules were described using the CHARMM22* force-field[43,44] and the CHARMM-modified version of the TIP3P water model respectively.[48,49]

For all systems, we modelled three different types of adlayers, similar to the models used in earlier work.[27] The first and second types of adlayers featured 200 amino acids either purely in the charge-neutral (N) state or the zwitterionic (Z) state, herein referred to as *pure* adlayers, whereas the third type of adlayer was composed of a mixture of 100 neutral molecules and 100 zwitterionic molecules, herein denoted as the *mixed* adlayer. The L-chiral forms of His and Ala were modelled with the neutral state capped by the amine (NH_2) and carboxylic (COOH) groups at the N- and C-termini respectively, while that of the zwitterionic state were capped by charged ammonium (NH_3^+) and carboxylate (COO^-) groups. Additionally, the His side-chain was protonated at the NE2 site, as shown in Fig. S3 in the ESI†. For each system investigated, a 2D periodic boundary condition (PBC) (*i.e.* in the lateral *x* and *y* dimensions only) was applied. Two soft repulsive walls were set below and above the graphene sheet, with a vacuum layer of ~26 Å separating the amino acid adlayers and the second reflective wall located above the graphene. The justification of these simulation setups and the combination of these force-fields have been previously tested and implemented by Awuah and Walsh to study pattern formation in amino acid adlayers at the *in-vacuo* graphene interface.[27]

The simulated annealing molecular dynamics (SAMD) simulation approach has been employed throughout.[50] All simulations were performed in the Canonical (*NVT*) ensemble, with the temperature maintained *via* the use of the Nosé-Hoover thermostat.[51,52] Newton's equations of motion were propagated using the leap-frog algorithm[53]. An integration time-step of 1 fs was used and frames were saved every 1 ps during the SAMD runs. The Particle Mesh Ewald (PME) electrostatic summation[54] was used, along with a cutoff in Lennard-Jones (LJ) nonbonded interactions of 12 Å. Similar to the SAMD protocol implemented in our previous study[27], the amino acid adlayers were subjected to five SA cycles (with four phases in each cycle) of simulated annealing, with a temperature window spanning 300-450 K, while maintaining the temperatures of the graphene slab and water (where applicable) at 300 K throughout. The duration of each cycle was 1ns, amounting to a total SAMD run of 5 ns. Furthermore, for all systems, three independent SAMD runs were performed, with the same initial adlayer arrangement as the starting configuration, but with different randomly-generated initial velocities. This procedure has been previously tested and found appropriate for studying amino acid adlayer-graphene interfaces.[27] Full details of analyses are provided in the ESI†.

3 Results and discussion

To briefly summarise, SAMD simulations were performed for pure adlayers (adlayers of His or Ala that are either entirely neutral or entirely zwitterionic) and mixed adlayers (adlayers of His or Ala comprising a 1:1 mixture of neutral and zwitterionic forms) adsorbed at the graphene interface *in vacuo*. As detailed in the Methodology, these simulations were initiated with spatially randomised positions for all amino acids on the surface. Adlayers in the absence of trace water are denoted as 'dry' systems, whereas adlayers that featured trace amounts of water are denoted as 'wet' systems.

Exemplar movies of the pure neutral, pure zwitterionic, and mixed dry interfaces are provided as supplementary data, along with descriptive text provided in the ESI† Section 'Video Description'.

3.1 Adlayer Morphology

Snapshots of the resultant morphologies of the dry pure and mixed adlayers are provided in Fig. 1. Corresponding snapshots for the wet graphene interface are provided Fig. S4 in the ESI†. For both His and Ala, morphologies of the pure zwitterion adlayers, and the mixed adlayers, revealed incipient, spontaneously-ordered patterns at both dry and wet graphene interfaces, highlighted in green in Fig. 1. Adlayers comprising purely neutral species showed little evidence of spontaneous ordering on graphene. This outcome is broadly consistent with simulations reported for tryptophan (Trp) and methionine (Met) adlayers on graphene *in vacuo*.[27] However, regarding the zwitterionic systems, previously we found that Trp and Met supported only one type of ordered motif, *i.e.* the canonical dimer row. In contrast, both His and Ala supported several motifs that simultaneously co-existed in the one (zwitterion-containing) adlayer.

The pure neutral adlayers exclusively formed a single-layered morphology, resulting in a relatively small percentage of exposed graphene ($\sim$ 36% and $\sim$ 56% for pure His-N and Ala-N adlayers respectively). The zwitterion-containing adlayers formed a multi-layered morphology, resulting in a greater average percentage of exposed graphene ($\sim$ 54% for pure His-Z and $\sim$ 74% for Ala-Z adlayers), as indicated in Fig. S5 and S6, and Table S1 in the ESI†. Further characterisation of the adlayer morphologies is provided in the ESI† Section 'Adlayer Morphology'. In summary, the overall trend in the amount of exposed graphene, in addition to the single- and multi-layered organisation of the pure neutral and zwitterion-containing adlayers respectively, indicated that the termination state had a substantial influence on the lateral and vertical arrangements of the amino acids in the adlayers. These findings are consistent with the previously-proposed zwitterion-driven mechanism of multi-layered organisation of amino acid adlayers[27], summarised in Fig. 1b.

3.2 Ordered Motifs within Adlayer Morphologies

Snapshots of representative ordered nanostructures found within His and Ala adlayers at the dry graphene interface are provided in Fig. 2 and 3 respectively. Figures S9 and S10 in the ESI† show the corresponding ordered motifs at the wet graphene interface. The four motifs identified within the His adlayers are denoted as the staggered dimer row, and the terminus-terminus, ring-terminus, and alternating monomer rows. The two motifs spontaneously produced by Ala are referred as canonical and staggered-parallel dimer rows. A comparison of these motifs and the adlayers in which they emerged are summarised in Table 1. Overall, different types of spontaneously ordered motifs, co-existent in the zwitterion-containing His and Ala adlayers on graphene, were identified. Specifically, for the His adlayers, four different types of spontaneously emergent motifs were identified (Figure 2), with at least three motifs found in the one system for each independent SAMD simulation, for both the zwitterion and mixed adlayers. In contrast, the Ala adlayers featured spontaneous emergence of two different types of dimer row structure (Fig. 3) in the one system for each independent SAMD simulation, for both zwitterion and mixed adlayers. These findings differed from those reported for Trp and Met adlayers on graphene, where a single type of motif, *i.e.* the canonical dimer row, was observed.[27] This suggests that amino acids with different side-chain characteristics can initially self-organise into different types of ordered structures upon deposition on non-reactive substrates.

Previous simulations reported observations of spontaneous emergence of only canonical dimer rows for both Trp and Met adlayers at graphene interface *in vacuo*.[27] The formation of these traditional dimer rows was found to be mediated mainly by charge-driven interactions between the N- and C-termini, with the side-chains mostly positioned flat on the graphene and assumed to be spectators in the motif. Consequently, the molecules were arranged in a terminus-to-terminus configuration, with the side-chains protruded outward to form the exterior surface of the dimer row.[27] This configuration might be linked to the fact that the adsorbed geometries of the individual amino acids may be chiefly influenced by the strong engagement of the side-chain with graphene.

3.2.1 His Adlayer Motifs

Dimers in the His staggered dimer row were adsorbed with the imidazole rings mostly aligned parallel to the surface, as indicated in Fig. 2, similar to that of the indole side-chain in Trp.[27] However, within the staggered dimer row motif, the lateral adjacent pairs (Fig. 2b) were arranged in a terminus-to-terminus configuration, with the ammonium and carboxylate groups facing each other. The longitudinal adjacent molecules in the row (Fig. 2b) interacted in a carboxylate-to-ring configuration. The lateral adjacent molecules were mostly aligned anti-parallel to each other (rings pointing in opposite directions, illustrated with magenta arrows in Fig. 2a). Due to this staggered feature, we suggest that this motif might ultimately mature into chevron-like structures as reported for His on graphite in octanol, by Krebs *et al.*[17] Additionally, the staggered dimer row motif emerged only in the adlayers composed purely of zwitterions (Table 1).

In the case of the monomer row motifs present in the His adlayers, three different types of molecular arrangements were noted. The ring-terminus monomer row motif (Fig. 2a) was featured in both zwitterion and mixed adlayers (Fig. S9a and S9c in the
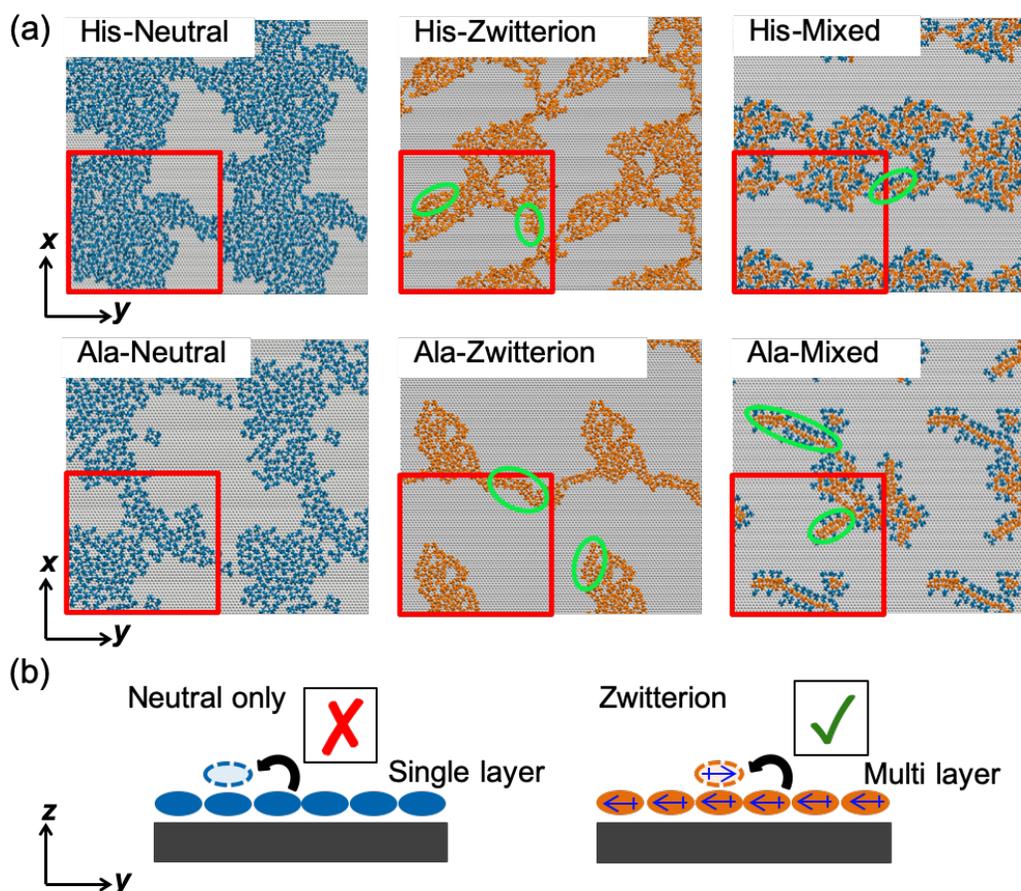
Fig. 1 (a) Snapshots of representative morphologies of neutral, zwitterion and mixed adlayers of His and Ala on graphene (2×2 supercell). The graphene sheet is shown in grey, neutral and zwitterion amino acids are shown in blue and orange respectively. Red squares represent the simulation periodic cell boundaries, green indicates spontaneously ordered motifs. (b) Proposed mechanism of single- and multi-layered organisation of amino acid adlayers.

ESI† show this motif in the zwitterion adlayer). This motif was constructed *via* interactions between the C-terminus and imidazole ring, consistent with experimental data reported for His on graphite in liquid octanol.[17] The lateral molecular arrangement within this structure was comparable to that of the longitudinal adjacent molecules in the staggered dimer row. As a result, the inter-molecular distance (C_α-C_α) between the longitudinal adjacent pairs along the row for these two structures was ~ 9.2 Å, broadly comparable with the corresponding experimental value of His on graphite (~ 10.0 Å)[17]. However, any discrepancy here may be linked to several factors, including the environment (liquid octanol), the substrate (graphite) and sample preparation, as previously discussed by Awuah and Walsh.[27]

For the terminus-terminus monomer row motif, the molecules within the row were aligned parallel to each other, *i.e.* with their side-chains pointing in the same direction (Fig. 2a), and with the ring in a (mostly) tilted orientation with respect to the graphene plane. The amino acids interacted in a terminus-to-terminus configuration, with no contribution from the imidazole rings, but without supporting a second row as per the canonical dimer row. The separation between the molecules along the row *via* calculation of the average C_α-C_α distance was estimated as $\sim$

5.6 Å, which was found comparable to the inter-dimer distance along the Met and Trp canonical dimer rows.[27] Similar to the ring-terminus monomer row, this motif was present in both the zwitterion and mixed adlayers. Fig. S9b in the ESI† shows this motif in the mixed adlayer.

The alternating monomer row, shown in Fig. 2a and Fig. S9f in the ESI†, was present only in the mixed adlayers, where row formation was driven by several types of inter-molecular interactions, including C-terminus/ring, C-terminus/C-terminus, N-terminus/ring and N-terminus/C-terminus interactions. Due to these different interactions pairings, the His molecules were aligned either parallel or anti-parallel to each other (Fig. 2a), with the imidazole rings positioned mostly in a tilted orientation. The inter-molecular spacing per C_α-C_α distance was estimated as ~ 5.8 Å.

Overall, several types of emergent ordered structures were found to co-exist in the His adlayers. The emergence of these structures was mostly affected by the character of the adlayers, *i.e.* either pure zwitterionic or a mixture of neutral and zwitterionic. Our simulations suggest that while the spontaneous emergence of ordered patterns may depend on the protonation state of the amino acid, the types of patterns formed appears to be in-
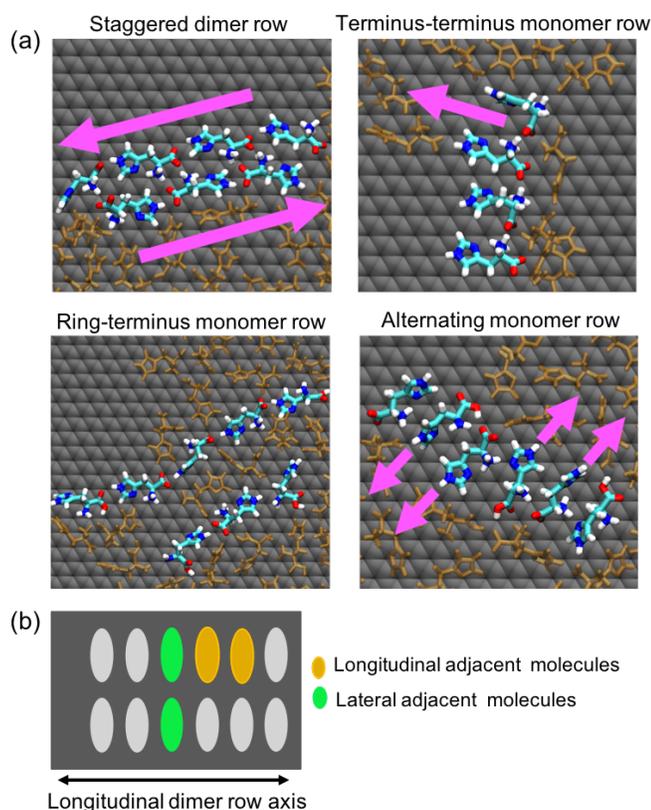
Fig. 2 (a) Zoomed in snapshots in plan view of representative spontaneously-emergent ordered motifs within pure zwitterion (top) and mixed (bottom) His adlayers on the dry graphene surface. Surrounding amino acids in the layer are shown in translucent orange. Magenta arrows indicate directions of C-terminus to imidazole ring. (b) Schematic showing the longitudinal and lateral adjacent molecules in a dimer row.

chain orientations might be linked to the compact nature of the methyl side-chain and its weak engagement with graphene. This outcome in side-chain orientation variability was not observed for amino acids with larger hydrophobic side-chains, such as Trp and Met adlayers on graphene. [27]
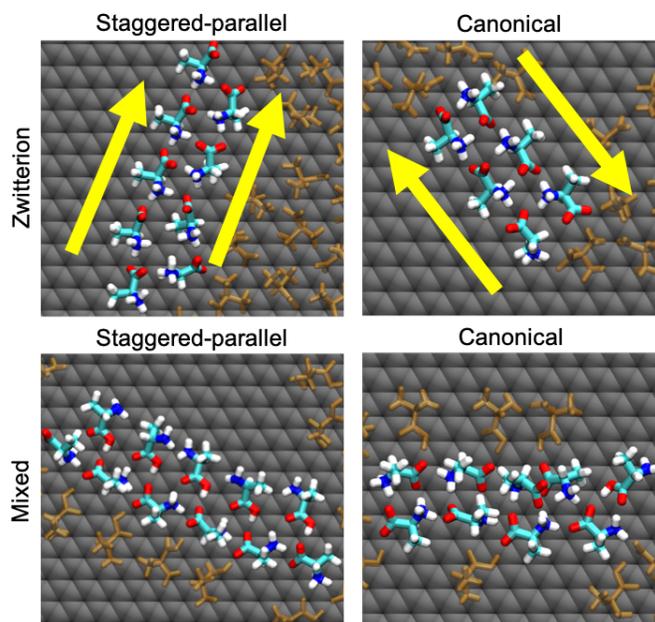
Fig. 3 Zoomed in snapshots in plan view of representative spontaneously-emergent ordered motifs within pure zwitterion (top) and mixed (bottom) Ala adlayers on the dry graphene surface. Surrounding amino acids in the layer are shown in translucent orange. Yellow arrows indicate the directions of the N-terminus to C-terminus along the row edge.

fluenced by the side-chain-supported interactions, supporting our hypothesis.

3.2.2 Ala Adlayer Motifs

For the Ala adlayers, two types of dimer row motifs were noted with distinct molecular ordering in both the zwitterionic and mixed adlayers, namely a staggered-parallel dimer row and the canonical dimer row, at both dry and wet graphene interfaces (Fig. 3 and Fig. S10 in the ESI†). In both cases, the side-chain could assume several orientations within the ordered motifs, for example, oriented away from, pointing to or arranged sideways on the graphene (Fig. S11 in the ESI†), with no discernible impact on the resultant nanopattern. The broad range of Ala side-

The difference between the molecular arrangement of the two Ala dimer row motifs appeared to be due to the positions of the carboxylate and ammonium groups of the lateral adjacent molecules. For example, along the staggered-parallel dimer row, the C- and N-termini of the lateral adjacent molecules were pointed towards the same direction, respectively (direction of C-terminus indicated with yellow arrows in Fig. 3). However, in the canonical dimer row, the respective C- and N-termini of the lateral adjacent dimers were oriented in opposite directions, as indicated in Fig. 3. A consequence of these two different structural arrangements was a slight increase ($\sim$ 0.4 Å) in the separation distance between the lateral adjacent molecules in the staggered-parallel dimer row, as shown in Fig. S12 in the ESI†.

Table 1 Comparison of spontaneously-emergent dimer row (DR) and monomer row (MR) motifs for the various amino acids and adlayers in which the motifs emerge.

Amino acid	Motifs	Adlayer	Ref.
Tryptophan	Canonical DR	Zwitterion, Mixed	27
Methionine	Canonical DR	Zwitterion, Mixed	27
Histidine	Staggered DR	Zwitterion	This work
	Terminus-terminus MR	Zwitterion, Mixed	
	Ring-terminus MR	Zwitterion, Mixed	
	Alternating MR	Mixed	
Alanine	Canonical DR	Zwitterion, Mixed	This work
	Staggered-parallel DR	Zwitterion, Mixed	

Additionally, for the staggered-parallel dimer row that emerged in the mixed adlayers, the lateral adjacent molecules comprised one neutral state paired with two zwitterionic species (Fig. 3 and Fig. S10c in the ESI†). In contrast, the lateral adjacent molecules within the canonical dimer row comprised one neutral state paired with one zwitterionic state, as shown in Fig. 3 and Fig. S10d in the ESI†. Nonetheless, the inter-molecular spacing along both dimer rows was estimated as ~ 5.7 Å (Fig. S12 in the ESI†). In the absence of experimental data for Ala adlayers on graphene in UHV, this value was found comparable to that of Met and Trp adlayers on graphene [27]. This comparable feature size for these amino acid adlayers may be due to the fact that their spontaneously-emergent ordered motifs were driven primarily by termini-to-termini interactions. That said, our data indicated that the size of an amino acid side-chain can have significant impact on the geometry of the amino acid on the substrate, which ultimately can influence the type of ordered patterns formed in the adsorbed adlayers.

Contrary to the ordered motifs found in the zwitterion-containing adlayers, no sign of emergent ordered structures was apparent in the adsorbed pure neutral adlayers for either His or Ala. However, similar to the findings of neutral Trp and Met adlayers on graphene [27], discrete dimer and trimer units, formed *via* carboxylic-carboxylic interactions, were found, as indicated in Fig. 4. These dimer and trimer units did not further associate to form rows or other patterns, which clearly suggested that the formation of ordered motifs in amino acid adlayers is mainly driven by charge-induced interactions. However, the absence of charge-induced interactions seemed to influence the population of the discrete dimer and trimer units in the neutral adlayers, as further detailed in the ESI† section 'Dimers and Trimers'.

Further characterisation of the adlayers is provided in the ESI† (Section 'Additional Characterisation'), including inter-molecular 2D radial distribution functions, the (modest) effect of the presence of trace water, and detailed adsorption geometry analyses.

3.3 Interactions within Adlayers

To further explore the influence of the side-chain characteristics on the interplay between the inter-molecular interactions and the variety and types of nanopatterns, the inter-molecular hydrogen bond interactions within the amino acid adlayers was quantified. These hydrogen bond analyses captured both conventional hydrogen bonds (*i.e.* no charges involved) and electrostatic interactions. The breakdown of the total number of inter-amino acid hydrogen bonds within His and Ala adlayers into various categories (compared to that of Trp and Met) [27] is provided in Fig. 5. Exemplar snapshots of typical hydrogen bond motifs within His and Ala adlayers are indicated in Fig. 6. Full details of all the inter-amino acid hydrogen bond interactions and their average numbers are provided in Fig. S16-S23 in the ESI†. Fig. 5 indicates a relatively higher number of hydrogen bonds within the zwitterion-containing adlayers compared with the neutral adlayers, consistent with our previous findings of Trp and Met adlayers on graphene. [27]

Within the pure zwitterionic adlayers of both His and Ala, the
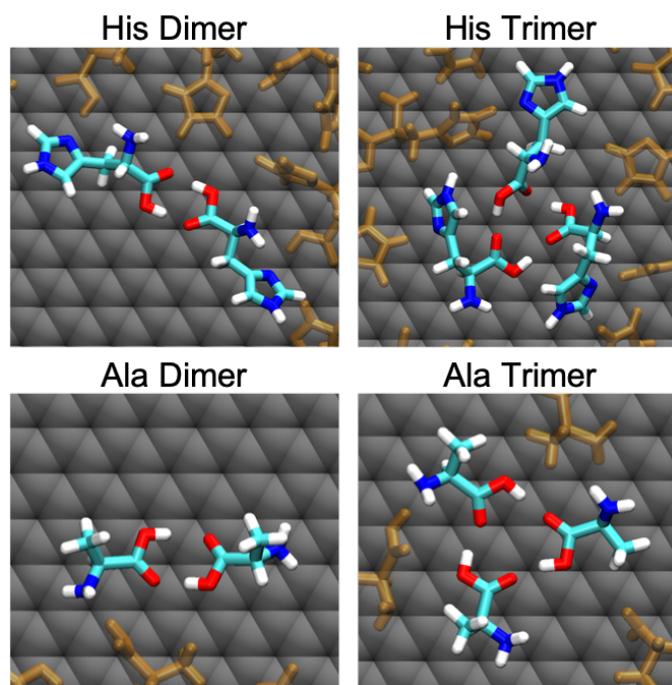
Fig. 4 Plan view zoomed in snapshots of the discrete dimer and trimer units identified in pure neutral adlayers for His (top) and Ala (bottom) at the dry graphene surface. Surrounding amino acids in the layer are shown in translucent orange.

largest contributor to the inter-molecular hydrogen bonds was the carboxylate-ammonium (C-term/N-term) pairing, shown in Figures 6d and 6f. This hydrogen bond motif is consistent with both experimental and theoretical studies. [11,13,18–20,27] Nevertheless, the carboxylate-ring (C-term/Ring) pairing showed a considerable contribution to the inter-molecular interactions within the His adlayer. Moreover, within the pure neutral adlayers, the carboxylic-ring (C-term/Ring) pairing (Figures 6a and b) was the largest hydrogen bond contributor for His, while the carboxylic-carboxylic (C-term/C-term) pairing (Fig. S16 in the ESI†) was the second largest, with a significant contribution from the Ring/Ring pairing (Fig. 6c). In summary, the character of hydrogen-bonding in the His adlayer was substantially influenced by the imidazole ring side-chain. On the contrary, the C-term/C-term pairing (Fig. 6e) was the largest contributor for Ala within the neutral adlayer, which explains the relatively higher percentage of discrete dimer and trimer units predicted for Ala.

In the case of the mixed adlayers, the overall trends in the inter-molecular hydrogen bonding broadly tally with that of Trp and Met. Specifically, there were relatively higher proportion of charge-charge interactions between zwitterions (zwitterion-zwitterion, as labelled Z-Z M in Fig. 5) compared with neutral-neutral interactions (N-N M in Fig. 5) for both amino acids. A moderate degree of interaction was also present between the neutral and zwitterions (N-Z M in Fig. 5), particularly in the case of the His adlayers, indicating an absence of phase segregation (*i.e* solely zwitterion-to-zwitterion or neutral-to-neutral interactions) within the mixed adlayers. In terms of the neutral-zwitterion hydrogen bond interactions in the mixed His adlayer
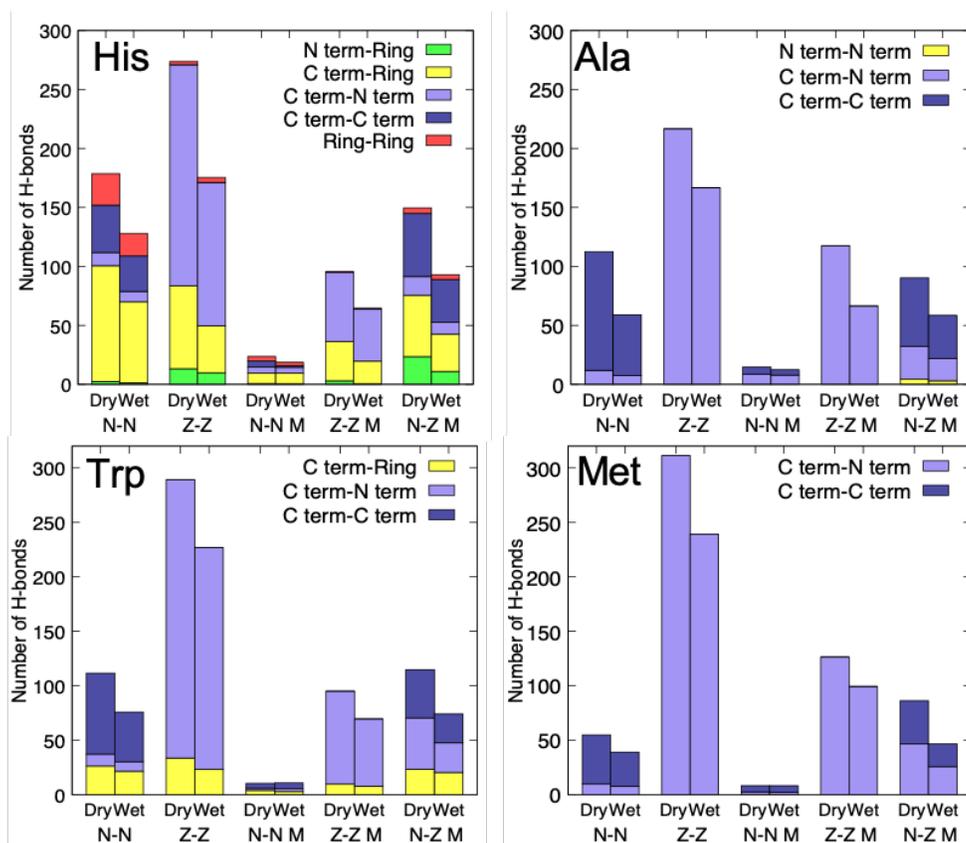
Fig. 5 The average number and types of inter-amino acid hydrogen bonds for His, Ala, Trp and Met in the first layer of the adsorbed adlayers. N-N, Z-Z and N-Z represent neutral-neutral, zwitterion-zwitterion and neutral-zwitterion interactions respectively. M denotes mixed adlayers. Data for Trp and Met were taken from Reference 27.
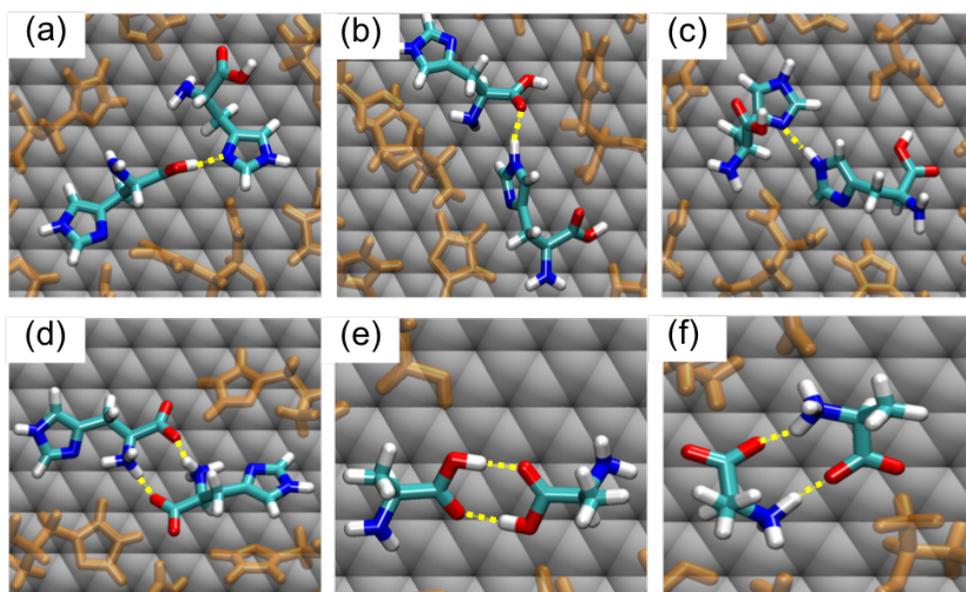
Fig. 6 Exemplar snapshots of representative (a-c) ring-supported and (d-f) termini-supported hydrogen bond motifs in (a-d) histidine and (e and f) alanine adlayers. Hydrogen bonds are indicated in yellow dotted lines. Surrounding amino acids in the layer are shown in translucent orange.

(Fig. 5, N-Z M), the amount of C-term/C-term contacts and C-term/Ring contacts were almost comparable. However for the mixed Ala adlayer, the neutral-zwitterion interactions (Fig. 5b, N-Z M) were dominated by the C-term/C-term pairings (Fig. S21 in the ESI†). Overall, within the His adlayers, the interactions are not dominated by inter-termini contacts, and instead the imida-

zole side-chain supports several types of hydrogen bonding inter-action, which appears to be a significant factor in the emergence of several ordered motifs. In contrast the side-chains of Ala, Met, and Trp have no capacity to hydrogen bond, and their interactions were dominated by hydrogen bonds between the N- and C-termini. However, despite these commonalities, Ala appeared able to support a range of co-existent motifs whilst Met and Trp do not. This highlights the complexity of the factors that underpin pattern formation, which extend beyond a mere question of 'hydrophobic *vs.* hydrogen-bonding. This clearly suggests that the spontaneously-emergent canonical dimer row motifs were mainly mediated by termini-termini contacts. In contrast, the His data indicate that the capacity of the amino acid side-chain to contribute to directional inter-molecular interactions can enable the simultaneous co-existence of several ordered patterns. The influence of the presence of trace water was also investigated for His and Ala adlayers (Fig. 5, and Figures S24-S26), revealing a drop in the number of amino acid/amino acid hydrogen bonds relative to the dry case. However, this did not disrupt the formation of ordered motifs. These hydrogen-bonding trends at the wet interface were consistent with those of Trp and Met.[27]

Results obtained from a previous study[27] together with the present data are therefore consistent with the hypothesis that amino acids comprising side-chains that have no capacity to support hydrogen bonding, and are not compact, such as Trp and Met, will self-organise into a single motif, namely a canonical dimer row. Also, for a given substrate, amino acids with compact side-chains that do not engage strongly with the surface, *e.g.* Ala, can self-organise into several co-existing motifs, *e.g.* canonical and staggered-parallel dimer rows. Furthermore, if the amino acid contains a side-chain that can support hydrogen bonding interactions, *e.g.* His, then this amino acid can also spontaneously self-organise into several co-existing ordered patterns, including monomer and dimer rows. As a prelude to a further in-depth investigation of this hypothesis, we performed an initial test by modelling an adlayer comprising purely of 200 zwitterionic Gly molecules at the dry graphene surface. Like Ala, this amino acid also has a compact side-chain (a single hydrogen atom), which makes this a relatively simple test case. On this basis, we expected that Gly would support the same co-existent motifs as observed for Ala.

Snapshots of the resultant representative ordered motifs found in the adsorbed zwitterionic Gly adlayer are shown in Fig. 7. As expected, both staggered-parallel and canonical dimer rows were found to co-exist, comparable to that of Ala. Similar to the methyl side-chain in Ala, the side-chain of Gly can assume several orientations with respect to the surface plane without disrupting the emergent ordered motifs. The motifs formed in the zwitterionic Gly adlayer broadly support our hypothesis. However, given the significant effect of other factors that influence the types of ordered structures formed on a given substrate, such as temperature[14–16,19,19,24] and coverage[11], we can not firmly conclude that the emergence of different types of motifs can be extrapolated from the side-chain characteristics alone.

In summary, our simulations demonstrate that different types of spontaneously-emergent motifs can co-exist in graphene-
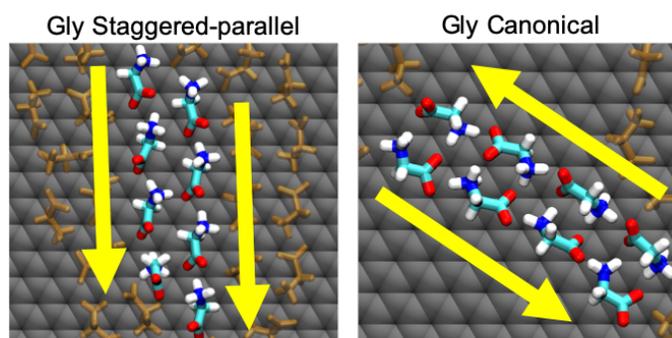
Fig. 7 Zoomed in snapshots in plan view of representative staggered-parallel and canonical dimer row patterns for Gly in the pure zwitterion adlayer on the dry graphene. Surrounding amino acids in the layer are shown in translucent orange. Yellow arrows indicate the directions of the N-terminus to C-terminus along the row edge.

adsorbed His and Ala adlayers that contain some zwitterionic content. From the different co-existent motifs observed here, it is clear that molecular arrangements in the ordered structures are influenced by the different characteristics of the amino acid side-chains. These side-chain characteristics extend beyond a 'hydrogen-bonding or not' scenario' and include the bulkiness, level of engagement with surface, as well as the capacity to support hydrogen bonding. Further qualification on what is meant by the level of engagement with the surface is also warranted. Here, a distinction should be made between directional engagement with the surface, as is facilitated via π-π stacking of aromatic rings with the graphene surface, and strong engagement conferred by general, non-directional van der Waals interactions which can collectively contribute to strong surface engagement via a large lateral footprint on the surface. We suggest Trp belongs to the former category and Met to the latter, and yet both produce the same (single) nanopattern. This suggests that strong engagement with the surface, and not the mechanism by which this strong engagement is conferred, is the key metric here.

Although environment effects (e.g. in-vacuo vs. in-liquid environments, temperature, pH and so forth) are not a focus of the current work, the question of possible disruption of the zwitterionic charge-charge interactions by the presence of salt is relevant. Previous studies have indicated interesting cation-π interactions on the graphene surface,[55] which prompted our preliminary investigation into the influence of ionic contaminants in the presence of trace water. Specifically we performed SAMD simulations of the zwitterionic Ala adlayer in the presence of 200 water molecules and 200 salt ions, specifically equal numbers of Na^+ and Cl^{-1}. Overall, we found our conclusions were insensitive to the presence of these ionic contaminants, with the same types of co-existent patterns identified as before (summarised in Figure S33 in the ESI†). Further investigations are not within the scope of this work, but should be addressed in future studies.

Some limitations of our current study warrant more detailed comments. Although co-existence of incipient ordered motifs in adsorbed adlayers on graphene was identified here, investigation of the subsequent mechanism of maturation of these motifs into

longer-ranged rows (as reported by experiments) was not an objective of this work. This task is outside the scope of the current study and would require the use of advanced techniques specialised for modelling long time-scale events. In addition, the current work did not probe the effect of molecular chirality on the types of emergent motifs, since only the L-chiral forms of His and Ala were considered here. However, it is not unreasonable to expect that adlayers entirely comprising the R-chiral forms would also yield similar results, given that adlayers comprising either entirely of either enantiomer of a single type of amino acid have been experimentally reported to form similar patterns.[13] Furthermore, we reiterate here that all molecular simulations hinge on the quality of the representation of the interatomic potentials, *i.e.* force-fields. As such, the force-field combination used here has been subjected to rigorous testing in terms of both aqueous solution and *in vacuo* environments.[27,31] Finally, although the current simulations have identified that a 50% fraction of zwitterionic content can produce nanostructures, this work has not sought to identify the minimum fraction of zwitterionic content required to support this.

4   Conclusions

Using molecular dynamics simulations, the influence of amino acid side-chain characteristics on the types of incipient ordered nanostructures spontaneously formed in amino acid adlayers adsorbed at the *in vacuo* graphene interface was investigated. The findings suggested a predictive framework for anticipating what types of pattern(s) can emerge for a given amino acid. This framework goes beyond a binary argument of 'hydrogen-bonding vs. hydrophobic' type of distinction. Specifically, the findings suggest an hypothesis that amino acids with side-chains that have no capacity to support hydrogen bonding and are not compact should self-organise into a single type of motif, the canonical dimer row. However, amino acids with compact side-chains (that do not engage strongly with the surface) are suggested to self-organise into several co-existing motifs. Furthermore, amino acids with a side-chain that can support hydrogen bonding can also spontaneously self-organise into several emergent co-existent nanopatterns. These hypotheses were explored via modelling of adlayers of two types of amino acids with contrasting side-chain characteristics, histidine and alanine. Spontaneous emergence of nanopatterning was found only in zwitterion-containing adlayers, with no evidence of ordered motifs in neutral adlayers, in agreement with previous work. Both histidine and alanine supported the co-existence of several ordered nano-motifs, in contrast with previous studies of Trp and Met adlayers. Alanine, with a compact side-chain that engaged weakly with graphene, and with no capacity to support hydrogen bonding, produced the emergence of a staggered dimer row configuration as well as the canonical dimer row. On the contrary, histidine adlayers supported a non-canonical dimer row motif co-existent with several different monomer row motifs, ascribed to the hydrogen bonding capacity of the side-chain. Preliminary simulations on the glycine adlayer supported the current hypotheses. Also, the presence of trace water did not adversely affect the emergent nanopatterns. Overall, in partnership with experiments, these findings provide a ra-

tional knowledge-based guide towards producing programmable, surface-induced nanostructures.

Conflicts of interest

There are no conflicts to declare.

References

1   A. F. Girão, M. C. Serrano, A. Completo and P. A. Marques, *Biomater. Sci.*, 2019, **7**, 1228–1239.

2   B. Zhan, C. Li, J. Yang, G. Jenkins, W. Huang and X. Dong, *Small*, 2014, **10**, 4042–4065.

3   V. Georgakilas, J. N. Tiwari, K. C. Kemp, J. A. Perman, A. B. Bourlinos, K. S. Kim and R. Zboril, *Chem. Rev.*, 2016, **116**, 5464–5519.

4   A. Ambrosi, C. K. Chua, A. Bonanni and M. Pumera, *Chem. Rev.*, 2014, **114**, 7150–7188.

5   J. M. Macleod and F. Rosei, *Small*, 2014, **10**, 1038–1049.

6   F. De Leo, A. Magistrato and D. Bonifazi, *Chem. Soc. Rev.*, 2015, **44**, 6916–6953.

7   Y. Liu, X. Dong and P. Chen, *Chem. Soc. Rev.*, 2012, **41**, 2283–2307.

8   T. R. Walsh and M. R. Knecht, *Chem. Rev.*, 2017, **117**, 12641–12704.

9   L. P. Xu, Y. Liu and X. Zhang, *Nanoscale*, 2011, **3**, 4901–4915.

10   V. Georgakilas, M. Otyepka, A. B. Bourlinos, V. Chandra, N. Kim, K. C. Kemp, P. Hobza, R. Zboril and K. S. Kim, *Chem. Rev.*, 2012, **112**, 6156–6214.

11   A. Schiffrin, A. Riemann, W. Auwärter, Y. Pennec, A. Weber-Bargioni, D. Cvetko, A. Cossaro, A. Morgante and J. V. Barth, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 5279–5284.

12   D. Costa, C. M. Pradier, F. Tielens and L. Savio, *Surf. Sci. Rep.*, 2015, **70**, 449–553.

13   V. Humblot, F. Tielens, N. B. Luque, H. Hampartsoumian, C. Méthivier and C. M. Pradier, *Langmuir*, 2014, **30**, 203–212.

14   M. Smerieri, L. Vattuone, T. Kravchuk, D. Costa and L. Savio, *Langmuir*, 2011, **27**, 2393–2404.

15   I. Tranca, M. Smerieri, L. Savio, L. Vattuone, D. Costa and F. Tielens, *Langmuir*, 2013, **29**, 7876–7884.

16   M. Smerieri, L. Vattuone, D. Costa, F. Tielens and L. Savio, *Langmuir*, 2010, **26**, 7208–7215.

17   E. Krebs, L. Grabill and A. Riemann, *Surf. Sci.*, 2018, **678**, 143–148.

18   A. Naitabdi and V. Humblot, *Appl. Phys. Lett.*, 2010, **97**, 1–4.

19   J. Reichert, A. Schiffrin, W. Auwärter, A. Weber-Bargioni,

M. Marschall, M. Dell'Angela, D. Cvetko, G. Bavdek, A. Cossaro, A. Morgante and J. V. Barth, *ACS Nano*, 2010, **4**, 1218–1226.

20 A. Riemann and B. Nelson, *Langmuir*, 2009, **25**, 4522–4525.

21 S. Fischer, A. C. Papageorgiou, M. Marschall, J. Reichert, K. Diller, F. Klappenberger, F. Allegretti, A. Nefedov, C. Wöll and J. V. Barth, *J. Phys. Chem. C*, 2012, **116**, 20356–20362.

22 A. Kühnle, T. R. Linderoth, M. Schunack and F. Besenbacher, *Langmuir*, 2006, **22**, 2156–2160.

23 M. Mahapatra, L. Burkholder, Y. Bai, M. Garvey, J. A. Boscoboinik, C. Hirschmugl and W. T. Tysoe, *J. Phys. Chem. C*, 2014, **118**, 6856–6865.

24 R. T. Seljamäe-Green, G. J. Simpson, F. Grillo, J. Greenwood, S. M. Francis, R. Schaub, P. Lacovig and C. J. Baddeley, *Langmuir*, 2014, **30**, 3495–3501.

25 T. Popa, E. C. Ting and I. Paci, *Surf. Sci.*, 2014, **629**, 20–27.

26 T. Popa and I. Paci, *J. Phys. Chem. C*, 2015, **119**, 9829–9838.

27 J. B. Awuah and T. R. Walsh, *Small*, 2019, **10.1002/smll.201903403**, 1903403.

28 J. Mou and W. S. Yang, *Ultramicroscopy*, 1992, **42-44**, 1025–1030.

29 W. Xu, M. Dong, H. Gersen, E. Rauls, S. Vázquez-Campos, M. Crego-Calama, D. N. Reinhoudt, E. Lægsgaard, I. Stensgaard, T. R. Linderoth and F. Besenbacher, *Small*, 2008, **4**, 1620–1623.

30 B. A. Collins and H. Ade, *J. Electron Spectrosc. Relat. Phenom.*, 2012, **185**, 119–128.

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

32 R. Zhiani, *Appl. Surf. Sci.*, 2017, **409**, 35–44.

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

34 T. R. Walsh, *Acc. Chem. Res.*, 2017, **50**, 1617–1624.

35 L. B. Wright, N. A. Merrill, M. R. Knecht and T. R. Walsh, *ACS Appl. Mater. Interfaces*, 2014, **6**, 10524–10533.

36 H. Vovusha and B. Sanyal, *RSC Adv.*, 2015, **5**, 67427–67434.

37 Z. E. Hughes and T. R. Walsh, *J. Mater. Chem. B*, 2015, **3**, 3211.

38 R. B. Pandey, Z. Kuang, B. L. Farmer, S. S. Kim and R. R. Naik, *Soft Matter*, 2012, **8**, 9101.

39 H. Vovusha, S. Sanyal and B. Sanyal, *J. Phys. Chem. Lett.*, 2013, **4**, 3710–3718.

40 B. E. Owens and A. Riemann, *Surf. Sci.*, 2014, **624**, 118–129.

41 A. Riemann and B. E. Owens, *Surf. Sci.*, 2010, **604**, 2084–2090.

42 S. Monti, C. Li, H. Ågren and V. Carravetta, *J. Phys. Chem. C*, 2015, **119**, 6703–6712.

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

44 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. K, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586.

45 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

46 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.

47 J. M. Martínez and L. Martínez, *J. Comput. Chem.*, 2003, **24**, 819–825.

48 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.

49 B. Roux, H. A. Yu and M. Karplus, *J. Phys. Chem.*, 1990, **94**, 4683–4688.

50 S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, *Science*, 1983, **220**, 671.

51 W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695.

52 S. Nosé, *J. Chem. Phys.*, 1984, **81**, 511–519.

53 W. F. Van Gunsteren and H. Berendsen, *Mol. Simulat.*, 1988, **1**, 173.

54 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089.

55 G. Shi, Y. Shen, J. Liu, C. L. Wang, Y. Wang, B. Song, J. Hu and H. P. Fang, *Sci. Rep.*, 2015, **4**, 6793.
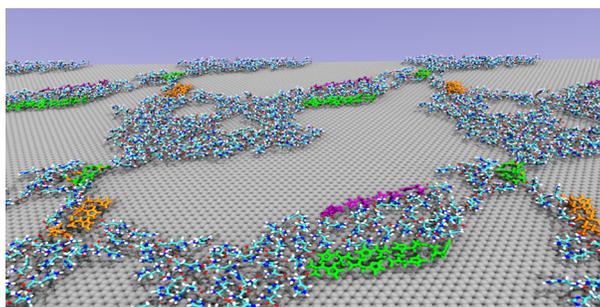
Table of Contents Image
The spontaneous emergence of several co-existing ordered nano-motifs in spatially randomised amino acid adlayers adsorbed on graphene *in vacuo* is influenced by side-chain characteristics.