



Cite this: *Phys. Chem. Chem. Phys.*,  
2021, **23**, 21840

# Physisorption of bio oil nitrogen compounds onto montmorillonite†

Maisa Vuorte,<sup>a</sup> Susanna Kuitunen<sup>b</sup> and Maria Sammalkorpi<sup>a,c</sup>

We assess computationally the adsorption of a series of nitrogen containing heterocycles and fatty acid amides from bio-oil on a model clay surface, Na-montmorillonite. The adsorption energies and conformations predicted by atomistic detail molecular dynamics (MD) simulations are compared against density functional theory (DFT) based molecular electrostatic potentials (MEP) and Hirshfeld, AIM, Merz-Singh-Kollman, and ChelpG charges. MD predicts systematically adsorption *via* cation bridging with adsorption strength of the heterocycles following purine > pyridine > imidazole > pyrrole > indole > quinoline. The fatty acid amides adsorption strength follows the steric availability and bulkiness of the head group. A comparison against the DFT calculations shows that MEP predicts adsorption geometries and the MD simulations reproduce the conformations for single adsorption site species. However, the DFT derived charge distributions show that MD force-fields with non-polarizable fixed partial charge representations parametrized for aqueous environments cannot be used in apolar solvent environments without careful accuracy considerations. The overall trends in adsorption energies are reproduced by the Charmm GenFF employed in the MD simulations but the adsorption energies are systematically overestimated in this apolar solvent environment. The work has significance both for revealing nitrogen compound adsorption trends in technologically relevant bio oil environments but also as a methodological assessment revealing the limits of state of the art biomolecular force-fields and simulation protocols in apolar bioenvironments.

Received 28th April 2021,  
Accepted 10th September 2021

DOI: 10.1039/d1cp01880a

rsc.li/pccp

## 1 Introduction

Biofuels, that is, fuels derived from *e.g.* vegetable oils, animal fats, starch, waste biomass, or microbial biomass, are an environmentally more sustainable alternative to traditional fossil fuels. In comparison to fossil fuels, bio fuel alternatives result in a reduction of CO<sub>2</sub>, hydrocarbons, and particulate matter emissions. Additionally, bio fuels rely on energy from relatively recent carbon fixation, which further decreases net carbon emissions to the environment. Diversification of biofuel sources from traditional food agriculture commodities and efforts to reduce involved land usage has led to rising interest in biomass of marine origins.<sup>1,2</sup> Especially algae

biomass appears promising because of its high growth rate, efficient carbon dioxide fixation, high biomass production rate, and elevated oil yield.<sup>3–5</sup>

The composition of biofuels depends strongly on the specific biomass used as feedstock and the chemical processes employed in conversion to biofuel.<sup>5–9</sup> A particular feature of algae derived bio-oils is their elevated nitrogen compound content, 5–10 wt% of algae bio-oils procured through pyrolysis.<sup>1,4,10</sup> These nitrogen compounds need to be removed from the oil for its use as a clean fuel, as well as to ensure optimal processing and shelf-life of the bio-oil. High concentrations of nitrogen containing compounds are problematic due to contamination of the catalyst, and the formation of NO<sub>x</sub> and N<sub>2</sub>O compounds during combustion of the bio-oil in fuel usage. These function as greenhouse gases, contribute to the depletion of the ozone layer, and have a role in the formation of acid rains and photochemical smog. Hence, denitrogenation is important in bio-oils refining. The nitrogen removal is commonly achieved *via* selective adsorption of nitrogen compounds from the rest of the oil onto the active sites of an adsorbent substrate, such as activated carbon,<sup>11,12</sup> silica–alumina<sup>11</sup> or metal-exchanged zeolites.<sup>11,13,14</sup> The adsorbent may later be regenerated for reuse through solvent washing or thermal treatment.<sup>14</sup>

<sup>a</sup> Department of Chemistry and Materials Science, School of Chemical Engineering, Aalto University, P.O. Box 16100, FI-00076 Aalto, Finland.  
E-mail: maria.sammalkorpi@aalto.fi

<sup>b</sup> Neste Engineering Solutions Oy, P.O. Box 310, FI-06101 Porvoo, Finland

<sup>c</sup> Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, P.O. Box 16100, FI-00076 Aalto, Finland

† Electronic supplementary information (ESI) available: Bootstrapped PMF profiles and decomposition of PMF into Coulombic and Lennard-Jones non-bonded contributions for imidazole adsorption on Na-montmorillonite, as well as partial charge decompositions of the N-heterocycles and fatty acid amides: Charmm GenFF, Hirshfeld, AIM, Merz-Singh-Kollman, and ChelpG charges. See DOI: 10.1039/d1cp01880a



As selective adsorption of nitrogen compounds is in key role in their removal, it is enticing to pursue chemical detail level guidelines to understand and control the adsorption response. Presently, there are only a few targeted studies of nitrogen compound adsorption from algae oils, the most recent of these using phyllosilicate clays, *e.g.* montmorillonite,<sup>15,16</sup> activated carbon,<sup>11,12</sup> Cu-zeolites,<sup>11</sup> silica gels,<sup>11</sup> and aminated metal-organic frameworks.<sup>17</sup> Computational approaches to surfactants in apolar environments have mostly focused on bulk assemblies, see *e.g.* ref. 18–25, or been limited to modelling adsorption of surfactants and other organic compounds at the apolar oil–solid interfaces, see *e.g.* ref. 26–33. Our previous work examines adsorption of vegetable oil minority species (fatty acids, monoglycerides, phospholipids) on silica *via* molecular dynamics simulation and links head group hydrogen bonding capability and charge to adsorption strength.<sup>34</sup>

Here, we pursue the adsorption response of a systematic series of nitrogen containing heterocycles, fatty acid amides, and water using atomistic detail molecular dynamics (MD) simulations. Specifically, physisorption of the compounds on a model clay (Na-montmorillonite) surface is probed in a triglyceride (triolein) solvent that describes a model bio-oil. Montmorillonite is chosen for the work as a common model adsorbent clay for drywashing bio-oils: montmorillonite-based adsorbents have formerly been used for both decolorization and deodorization of bio-oils as well as catalysts for bio fuel synthesis.<sup>15,16,35,36</sup> Experimental studies of nitrogen compound adsorption on montmorillonite are mainly limited to chlorophyll, phospholipids, and sterols.<sup>15,16</sup> We employ MD simulations with the umbrella sampling method to determine the adsorption conformations and energies of individual molecules on the Na-montmorillonite surface. As prior studies<sup>22,37–41</sup> have raised concern on the performance accuracy of the aqueous environment force-field parametrizations used typically in atomistic detail MD simulations when taken to apolar solvent environments, we assess the modelling accuracy *via* density functional theory (DFT) electronic structure calculations with an implicit solvent model. The results enable an objective performance evaluation of the Charmm force field, in particular Charmm GenFF<sup>42</sup> parameters, for examining simple organic molecules in low dielectric, apolar solvent environments. Besides the methodological assessment, the work has significance for revealing systematic nitrogen compound adsorption trends in technologically relevant bio oil environments. In particular, the work provides a mapping of adsorption preference order for the nitrogen compounds, which allows making predictions on the adsorption response based on the accessibility of the nitrogen but also bring out methodological pitfalls for modelling biobased surfactants in organic solvents.

## 2 Computational methods

### 2.1 Molecular dynamics simulations

The adsorption of six nitrogen containing, heterocyclic compounds, five fatty acid amides, and water from triglyceride

(triolein) solvent onto Na-montmorillonite was studied in terms of adsorption conformation and free energy changes accompanying adsorption events *via* MD simulations. All MD simulations were carried out using the Gromacs 2016.5 simulation package.<sup>43</sup> The montmorillonite surface was built based on the montmorillonite crystal structure (super cell) provided as part of the INTERFACE force field database,<sup>44</sup> with Na<sup>+</sup> counterions and optimized Al to Mg substitutions. This leads to chemical formula Na<sub>0.321</sub>[Si<sub>4</sub>O<sub>8</sub>][Al<sub>1.679</sub>Mg<sub>0.321</sub>O<sub>2</sub>(OH)<sub>2</sub>]. The chemical structures of the examined heterocyclic nitrogen compounds and fatty acid amides are presented in Fig. 1.

The studied nitrogen compound set is selected because the chemical composition of algae derived bio oils can vary significantly depending on factors such as algae species and chemical processing methods used to process the biomass, such as lipid extraction or pyrolysis. During chemical processing, organic macromolecules, such as lipids and proteins, may depolymerize and recombine to form organic chemical species not present in the original biomass feedstock. Such species include for example fatty acid amides. Particularly, the combination of residual water in the feedstock and elevated temperature will rapidly hydrolyze proteins into amino acids. These then further decompose into ammonia and amines. Ammonia or amines that react with free fatty acids yield primary and secondary fatty acid amides, respectively.<sup>7</sup> Additionally, diverse nitrogen containing heterocycles are abundant in algae derived bio-oils.<sup>4,8,45</sup> By covering the adsorption of heterocyclic compounds and fatty acid amides as a series that varies systematically in structure, see Fig. 1, we cover the species that can be expected to be present relatively well and also obtain a generalizable trend in the adsorption responses.

The montmorillonite surface and the nitrogen adsorbates (N-heterocycles and fatty acid amides) are described in fully atomistic detail level using the INTERFACE force field<sup>44</sup> and Charmm GenFF 4.1,<sup>42</sup> respectively. Notably, the INTERFACE force field enables fully dynamic montmorillonite slab and surface atoms. This allows for possible morphing of the surface structure. The triolein solvent is described by the Charmm c27 complementary united-atom description by Henin *et al.*<sup>46</sup> Our previous work points toward Charmm c27 interactions being better balanced for apolar solvent environment simulations than the more recent Charmm parametrizations.<sup>37</sup> The usage of a united atom description for the solvent leads to a significant speedup of the calculations. For water, the tip3p<sup>47,48</sup> explicit water model is used. Notably, both Charmm and INTERFACE force fields were originally parametrized in aqueous environments and for fully hydrated biospecies. As the systems examined here contain charged species (montmorillonite Na<sup>+</sup> counter ions) in a low dielectric medium (triolein solvent), all results obtained for these systems require a careful and critical examination.

To prepare the simulation configurations, 100 triolein molecules were first inserted into a rectangular 3.6 nm × 3.6 nm × 25 nm simulation box. This leads to an excessively sparse, random mixture of triolein solvent. A montmorillonite surface 1.57 nm (2 repeat layers) in thickness was set to match the



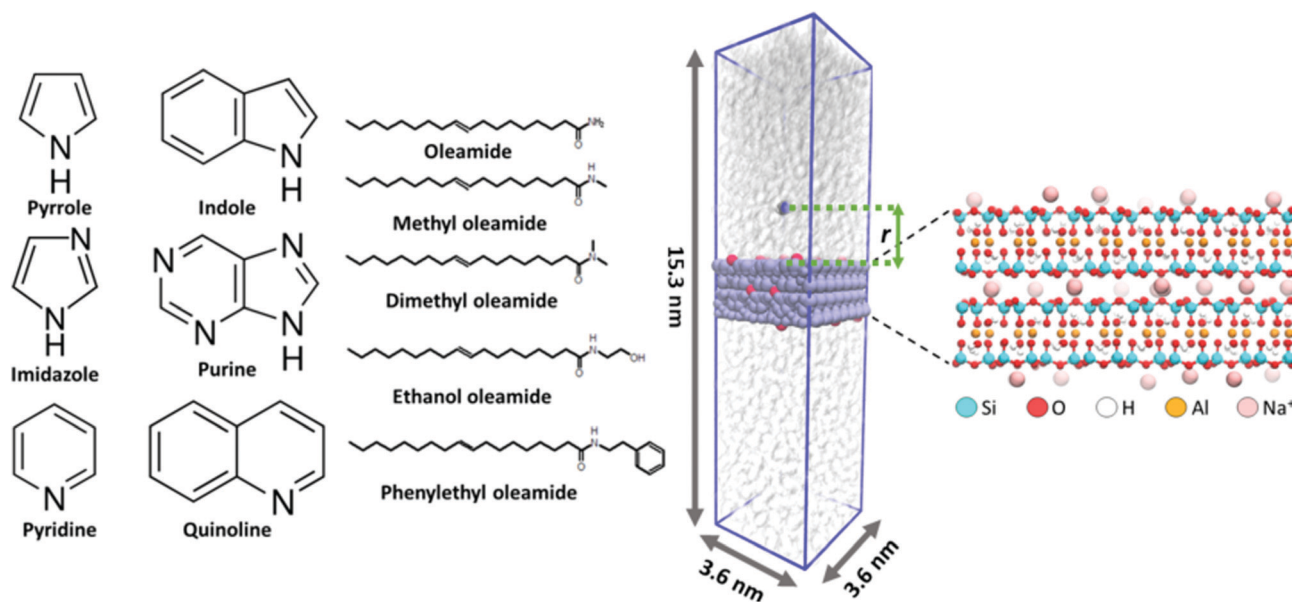


Fig. 1 Left: Chemical structures of the nitrogen containing heterocyclic compounds and fatty acid amides examined in this work. Middle: Visualization of an example initial configuration and the simulation box are presented at right. The montmorillonite surface is shown in light purple, Na<sup>+</sup> ions in red, and triolein solvent in translucent white. The PMF reaction coordinate *r* denotes the perpendicular distance between adsorbate center of mass and the plane defined by the topmost montmorillonite Si atoms. Right: Structure of the simulated two layer montmorillonite surface.

simulation box size in *x*- and *y*-dimensions as an infinite surface plane slab that spans the box in the *x,y*-plane (periodic boundary conditions). The simulation box was then gradually scaled smaller in the direction perpendicular to the surface (*z*-direction). Each scaling step was followed by a short 200 step steepest descent energy minimization. This was continued until appropriate density of the triolein solvent<sup>49,50</sup> was reached. Next, a 10 ns *NPT* ensemble initial equilibration using a semi-isotropic Berendsen barostat<sup>51</sup> with a reference pressure of 1 bar, compressibility of  $4.5 \times 10^{-5} \text{ bar}^{-1}$ , and a time constant of 2.0 ps was performed. In this, the average temperature was maintained at 343.15 K (70 °C) using the stochastic velocity rescale thermostat developed by Bussi *et al.*<sup>52</sup> with a time constant of 0.5 ps. After this initial equilibration of the triglyceride solvent, a single adsorbate (N-heterocycle or fatty acid amide) molecule was placed 5.5 nm above the montmorillonite surface. During a 50 ps  $\lambda$ -coupling simulation, the interactions between the adsorbate and the rest of the system (triglyceride solvent and montmorillonite surface) were gradually turned on to avoid overlap of atoms. The system was then *NPT* equilibrated for 5 ns using the barostat and thermostat parameters specified above. Visualization of an example initial simulation configuration, the simulation box, and the montmorillonite slab is provided in Fig. 1. During simulation, the surface Na<sup>+</sup> atom positions can fluctuate significantly, resulting in changes in local electrostatics. The effect of this is pronounced due to the low dielectric screening of the triglyceride solvent. To avoid the PMF convergence challenges rising from Na<sup>+</sup> displacements at the surface, the Na<sup>+</sup> ion positions were constrained using  $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  harmonic constraints after the *NPT* equilibration for all later simulation steps.

The binding energy of the adsorbate on the montmorillonite surface was assessed based on the free energy difference estimate of the adsorbate molecule being fully solvated by the triolein *vs.* adsorbed on the surface. This was estimated based on the potential of mean force (PMF) determined *via* umbrella sampling methodology. For the N-heterocycles, the perpendicular component (*z*-direction) of distance between the center of mass of adsorbed molecule and the plane defined by the topmost montmorillonite surface Si atoms was used as the PMF reaction coordinate, see Fig. 1. As the fatty amides are relatively large, the center of mass of the amide group connected to the head group (not the fatty acid tail) was used for the reaction coordinate.

Technical details for obtaining the PMF curves are as follows. To obtain the PMF curves, umbrella sampling windows were generated at 0.1 nm intervals along the reaction coordinate. Each of the simulated windows was run for a total of 200 ns using a 2 fs time step. The adsorbate was constrained in each window by a  $1000\text{--}8000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$  harmonic constraint. Similar to the simulation system preparation, the temperature of the system was controlled by the stochastic velocity rescale thermostat<sup>52</sup> with a time constant of 0.5 ps and an average system temperature of 343.15 K. The average pressure was set to 1.0 bar using the Parrinello–Rahman semi-isotropic barostat with compressibility of  $4.5 \times 10^{-5} \text{ bar}^{-1}$  and a time constant of 2.0 ps.<sup>53,54</sup> The PMF curves were extracted from the trajectories of the simulated windows using the weighed histogram analysis method (WHAM) developed by Kumar *et al.*<sup>55</sup>. The presented PMF profiles represent the average profile of 300 bootstrapped profiles, with error estimate based on the standard deviation of the bootstrapped profiles.



Fig. S1 of ESI,<sup>†</sup> presents an example set of these 300 bootstrapped profiles, calculated for imidazole, to enable assessment of PMF profile and adsorption energy convergence. The reported adsorption energies correspond to the difference between the PMF minimum, *i.e.* the adsorption site at the adsorbate surface, and the plateau value at large distance.

To assess the contributions of Lennard-Jones (LJ) and Coulombic non-bonded interactions to the overall PMF profiles, we performed a decomposition of the PMF data curve to LJ and Coulombic contributions for imidazole as the test case. The decomposed PMF is presented as Fig. S2 in the ESI.<sup>†</sup> The data supports the conclusion that the contribution of LJ-interactions to adsorption is much smaller than the Coulomb interaction contributions. The difference is so large that even taking into account the differences in molecule sizes, for species such as in this work the overall adsorption energy can be expected to scale with molecule point charges, particularly the adsorbing imidazole ring nitrogen. Imidazole is employed as the test case here due to its two ring nitrogens.

## 2.2 Electronic structure calculations

Electronic structure calculations were carried out for the same set of N-heterocycles and fatty acid amides as the MD simulations, see Fig. 1. For the fatty acid amides, the hydrocarbon tail was truncated from 18 to 6 carbons and the double bond was omitted. Initial molecular geometries were constructed using Avogadro: an open-source molecular builder and visualization tool, version 1.1.1.<sup>56</sup> The configurations were energy minimized *in vacuo* using the general Amber force field (GAFF)<sup>57</sup> included in Avogadro. Gaussian g16<sup>58</sup> was used for subsequent geometry optimization of the prepared molecular structures and single point calculations for population analysis to derive Hirshfeld,<sup>59</sup> Merz-Singh-Kollman (MK),<sup>60,61</sup> and ChelpG<sup>62</sup> partial charges, as well as to calculate molecular electrostatic potential (MEP). The PBE0 hybrid functional<sup>63,64</sup> and def2-TZVPP basis set<sup>65,66</sup> were used. The solvent was modelled by an implicit CPCM solvent model<sup>67,68</sup> with relative permittivity of 3.109 and refractive index of 1.477 based on literature values for triolein.<sup>69,70</sup> Atoms in molecules (AIM) analysis<sup>71–74</sup> was carried out to derive AIM<sup>75</sup> partial charges based on the calculated electron densities. VMD was used for MEP visualizations.<sup>76</sup>

## 3 Results and discussion

To assess the adsorption energies of the nitrogen containing compounds, first the PMF curves corresponding to the free energy landscape observed by the nitrogen compound near the montmorillonite surface were calculated. For the heterocycles, the curves are presented in Fig. 2 and the corresponding adsorption energies  $\Delta G_{\text{ads}}$  collected in Table 1. The data points toward adsorption strength of the examined N-heterocycles being purine > pyridine > imidazole > pyrrole > indole > quinoline.

Let us next consider the binding conformations adopted by the heterocycles. Fig. 3 plots the radial distribution functions

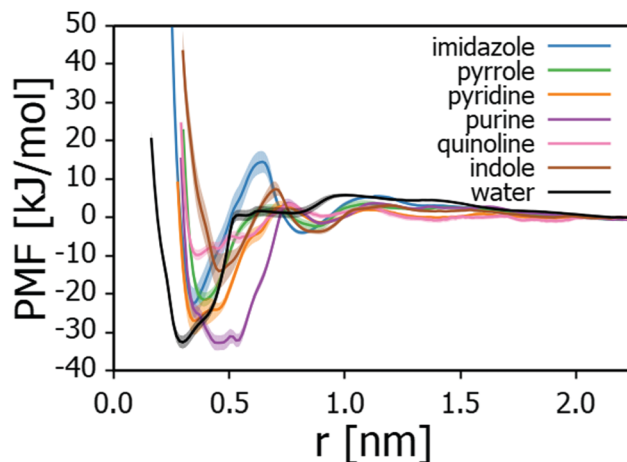


Fig. 2 Potential of mean force (PMF) curves for N-heterocycles and water adsorbing onto Na-montmorillonite in triolein solvent. Distance  $r$  is the perpendicular distance between heterocycle center of mass and Na-montmorillonite surface topmost layer Si atoms. Shading indicates the PMF error estimate.

Table 1 Adsorption energies  $\Delta G_{\text{ads}}$  [kJ mol<sup>-1</sup>] for heterocyclic nitrogen containing compounds and water on montmorillonite

Adsorbate	$\Delta G_{\text{ads}}$ [kJ mol <sup>-1</sup> ]
Quinoline	$-9.8 \pm 1.4$
Indole	$-14.0 \pm 4.7$
Pyrrole	$-21.4 \pm 1.7$
Imidazole	$-22.4 \pm 4.0$
Pyridine	$-26.9 \pm 3.6$
Purine	$-32.7 \pm 2.0$
Water	$-32.5 \pm 1.7$

(RDFs) of N-heterocycle nitrogens with respect to montmorillonite surface Na<sup>+</sup> ions. The RDFs were calculated based on the umbrella sampling windows corresponding to PMF minima in Fig. 2 over the last 30 ns of simulation time. The data shows that adsorption involves strong interaction of the adsorbate with the montmorillonite surface Na<sup>+</sup> cations for all examined heterocycle species. Most examined heterocycles bind directly with the surface Na<sup>+</sup> cations, as shown by the sharp localized short distance peak. For example, purine adsorbs on Na-montmorillonite in a parallel conformation with the surface, through interaction of a basic nitrogen with the surface Na<sup>+</sup> ions. This adsorption conformation is supported by existing literature.<sup>77</sup> However for indole and pyrrole, a more spread-out binding at larger separation from the surface cation is predicted. This presumably results from these heterocycles lacking a similar specific binding site for a cation. Both pyrrole<sup>78</sup> and indole<sup>79,80</sup> have been shown to bind cations through cation- $\pi$  interactions. The use of static charges in a non-polarizable force field, such as Charmm GenFF, means this effect is not captured. Calculations using polarizable versions of both Charmm<sup>81–83</sup> and OPLS<sup>84</sup> force fields have shown that cation-benzene interaction in gas phase is greatly underestimated by non-polarizable force fields. Interestingly for this work,



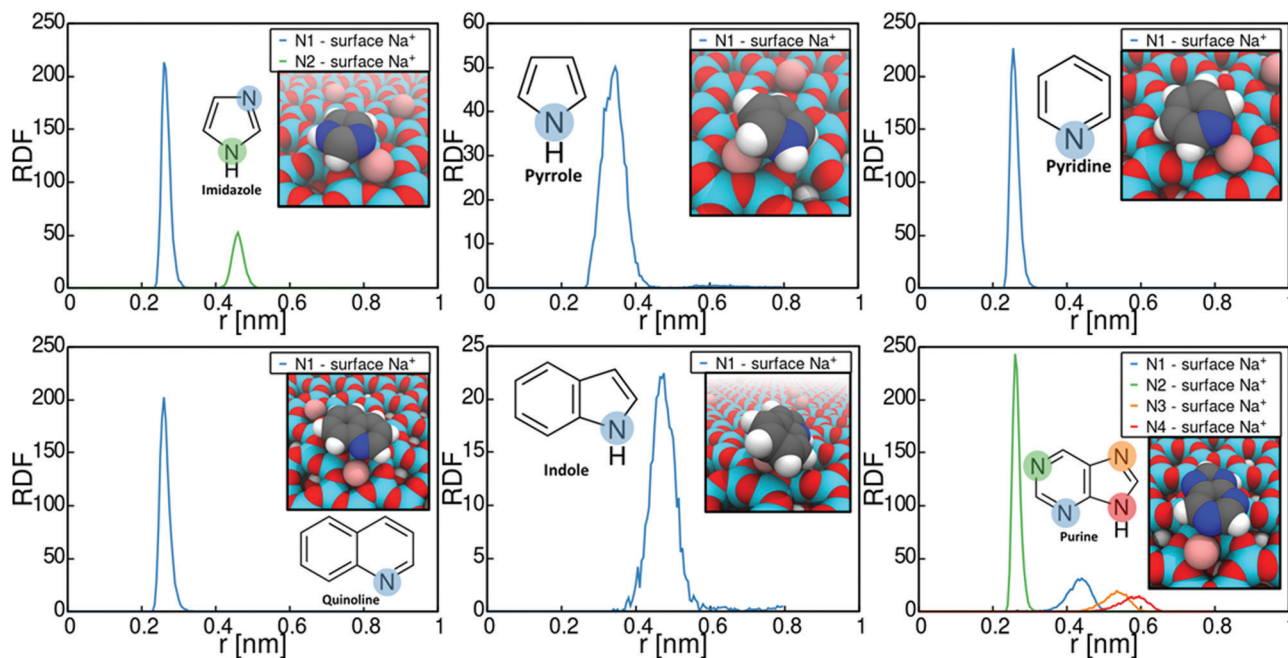


Fig. 3 Radial distribution function (RDF) of N-heterocycle nitrogen with respect to montmorillonite surface  $\text{Na}^+$  ions. Insets are visualizations of molecule adsorption conformations at the PMF minimum.

Turupcu *et al.* determined the interaction energy between  $\text{K}^+$  ion and benzene in water and THF using the cation- $\pi$  interactions enabled OPLS/2020 force field as  $-7.5 \text{ kcal mol}^{-1}$  ( $-31.4 \text{ kJ mol}^{-1}$ ) and  $-6.7 \text{ kcal mol}^{-1}$  ( $-28.0 \text{ kJ mol}^{-1}$ ) respectively.<sup>84</sup> In the approach, manual tuning of interactions is required to match *ab initio* and experimental target data.<sup>81,83,84</sup> Furthermore, as hydrated  $\text{Na}^+$  ions rarely directly participate in cation- $\pi$  interactions with benzene, this approach has not been pursued here. Nevertheless, the lack of competing cation- $\pi$  interaction in the simulations steers adsorption to occur *via* an electronegative atom such as the ring nitrogen, instead of the  $\pi$ -conjugated rings.

The significance of the strong surface cation correlation in the binding is that the type of cation present on the montmorillonite surface is likely to have a very strong influence on adsorption. Adsorption is driven mainly by Coulombic interactions while the Lennard-Jones contribution to the overall adsorption energy is minimal, see Fig. S2 in the ESI,<sup>†</sup> for decomposition of the imidazole adsorption PMF. This means adsorption energy can be expected to scale with the magnitude of the interacting point charges. Particularly, cation valency can be expected to be very significant in guiding adsorption, with higher valency leading to stronger adsorption. However, size of the surface cation will also affect steric availability of especially the amide carbonyl oxygen and divergence from the established adsorption preference can be expected.<sup>85-87</sup> An indirect conclusion of the strong cation correlation of the adsorption is that also the presence of *e.g.* water on the montmorillonite surface can be expected to significantly affect adsorption strength. With water present on surface, adsorption may occur either *via* displacement of water by the adsorbate or *via* water

bridging. For example, according to existing literature, the adsorption of *e.g.* nitroaromatic compounds in aqueous environments onto smectite clays involves the displacement on water at the clay surface by the adsorbate, followed by complexation with the clay surface cations.<sup>88</sup>

It is worth noting that periodic plane wave calculations derived adsorption energies of purine bases adenine ( $-84.5 \text{ kJ mol}^{-1}$ ) and guanine ( $-115.4 \text{ kJ mol}^{-1}$ )<sup>77</sup> on Na-montmorillonite *in vacuo* are higher compared to the adsorption energy of purine calculated here ( $-32.7 \text{ kJ mol}^{-1}$ ), likely explained by the lack of charge transfer in the model. Similarly elevated in comparison to the here presented MD modelling derived values are the theoretical literature values for binding energies of pyrrole -  $\text{Na}^+$  ( $-107.9 \text{ kJ mol}^{-1}$ )<sup>78</sup> and indole -  $\text{Na}^+$  ( $-131.8$  to  $-136.5 \text{ kJ mol}^{-1}$ )<sup>79,80,89</sup> complexes in gas phase. The elevated adsorption energy values are, however, calculated in vacuum or gas phase. Furthermore, evidence of chemisorption of, *e.g.* pyridine onto montmorillonite in aqueous environments exists.<sup>90</sup> It should be noted that the MD model used here only accounts for physisorption of the adsorbate on the adsorbent. Finally, the adsorption energies reported in Table 1 correspond to single molecule adsorption on a surface and therefore cooperative and competitive adsorption effects (effect of other adsorbate molecules present on the surface) have not been assessed here.

The adsorption geometries inferred from Fig. 3 and the adsorption energies of Table 1 show that the accessibility of the non-protonated basic nitrogen (in pyridine and imidazole) leads to higher adsorption energies than those resulting from  $\text{Na}^+$  cation interaction with the protonated nitrogens (in pyrrole and indole). The nonprotonated nitrogens in pyridine and



Table 2 Adsorption energies  $\Delta G_{\text{ads}}$  [kJ mol<sup>-1</sup>] for fatty acid amides and water on montmorillonite

Adsorbate	$\Delta G_{\text{ads}}$ [kJ mol <sup>-1</sup> ]
Ethanol oleamide	-20.2 ± 1.3
Phenylethyl oleamide	-20.3 ± 1.3
Dimethyl oleamide	-23.0 ± 1.3
Oleamide	-32.6 ± 0.9
Methyl oleamide	-46.7 ± 1.8
Water	-32.5 ± 1.7

imidazole and their lone pairs are not involved in the formation of a delocalized  $\pi$ -orbital and thus do not contribute to the formation of resonance structures.<sup>91</sup> This nitrogen lone pair explains the basicity of pyridine comparable to regular amines and imines.<sup>91</sup>

Table 2 summarises the adsorption energies of the examined fatty acid amides on Na-montmorillonite and the corresponding PMF profiles are presented in Fig. 4. The adsorption strengths follow the order: methyl oleamide > oleamide > dimethyl oleamide > phenylethyl oleamide > ethanol oleamide. In general, adsorption strength can be expected to be proportional to the steric availability of the headgroup for this type of linear, analogous molecules. Indeed, except for methyl oleamide, the results point toward introducing a bulkier headgroup to the amide significantly lowering the adsorption strength. This influence persists even if the substitute headgroup has some hydrogen bonding capability, such as ethanol oleamide. Methyl oleamide appears to be an outlier in the series by its significantly stronger adsorption than what predicted by its headgroup accessibility. Its adsorption strength is likely an artefact arising from the Charmm GenFF parametrization (imbalance of charge distribution vs. other interactions) and discussed below.

Fig. 5 plots the radial distribution function of the carbonyl oxygen and amide nitrogen in relation to the montmorillonite

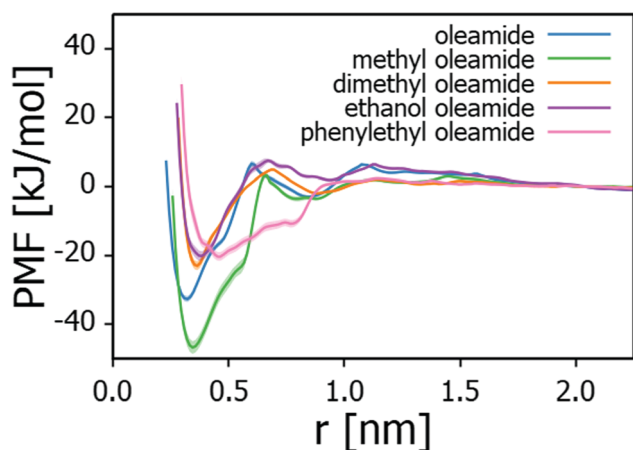


Fig. 4 Potential of mean force (PMF) curves for fatty acid amides adsorbing onto Na-montmorillonite in triolein solvent. Distance  $r$  is the perpendicular distance between fatty acid amide head group center of mass and Na-montmorillonite surface topmost layer Si atoms. Shading indicates the PMF error estimate.

surface Na<sup>+</sup> cations, as well as visualizations of the binding. The data reveals the adsorption sites and conformations. Adsorption of all the examined amides occurs systematically *via* interaction of the carbonyl oxygen with the montmorillonite surface Na<sup>+</sup> ions. This means that the adsorption strength can indeed be expected to be directly related to the steric availability of the carbonyl oxygen. Consequently, the adsorption response is directly subject to the partial negative charge assigned to the carbonyl oxygen atom in the force field parametrization.

The assessment of adsorption of the nitrogen compounds in this work is on unhydrated Na-montmorillonite with no water present in the system. However, montmorillonite and other smectite clays are prone to surface hydration and significant swelling due to water in the interlayer space of the clay.<sup>92,93</sup> Hence, we determined also the adsorption response and energy of water. Water adsorption energy is -32.5 kJ mol<sup>-1</sup>, see Fig. 2. This means that water at the montmorillonite surface, if present, will be bound at the surface by the Na<sup>+</sup> ions. It may also influence the adsorption mechanism and adsorption strength of adsorbates. Notably, prior works have shown that adsorption related cation bridging primarily occurs at oil-wetted clay interfaces, while on hydrated clay surfaces cation bridging does not occur due to hydration of the ions.<sup>94</sup> This is in agreement with deductions of the current work.

The here employed empirical force-fields, similar to most biomolecular force fields, such as Charmm, Amber, Gromos, and OPLS, are unpolarizable and rely on static partial charges of atoms.<sup>95,96</sup> Furthermore, the parametrizations are optimized for aqueous environments and fully hydrated systems. Despite their shortcomings, empirical biomolecular force fields are successfully used to describe complex biochemical environments. However, the lack of polarization is known to become especially problematic for low dielectric environments, such as in apolar solvents,<sup>97,98</sup> inside lipid membranes,<sup>99,100</sup> and in hydrophobic protein core regions<sup>101</sup>. The parametrization procedure of *e.g.* the Charmm GenFF force field<sup>42</sup> that is used in the MD simulations here involves matching of Hartree-Fock (HF) level interaction energies and distances for the compounds and water, along with additional matching to experimental properties. Because of the parametrization approach, the effect of electronic polarization and dielectric charge screening in aqueous environments are somewhat baked into the empirically derived interaction parameters of the force fields; however, the extent of this is unclear.

To assess the effect of this, we next compare the Charmm GenFF parameters, especially electrostatics in terms of partial charges, against DFT electronic structure calculations. The examination focuses on electrostatics because the adsorption to montmorillonite surface occurs here *via* cation bridging.

The molecular electrostatic potentials (MEPs) corresponding to the N-heterocycles and truncated fatty acid amides are presented in Fig. 6 and 7 respectively. The MEP describes the electrostatic potential felt by a positive test charge, *i.e.* proton, at a certain coordinate around a molecule, disregarding polarization of the molecule in response to the test charge. Typically,



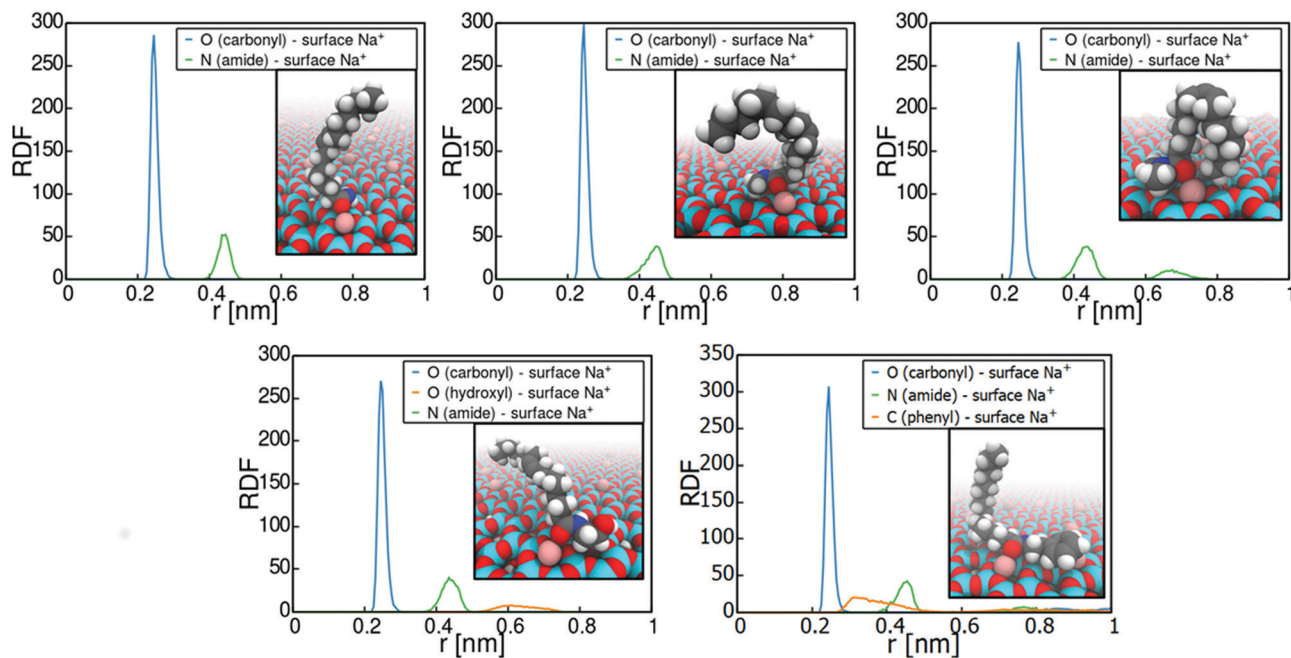


Fig. 5 Radial distribution function (RDF) of fatty acid amide functional oxygens, amide nitrogens, and aromatic carbons (phenylethyl oleamide) with respect to montmorillonite surface  $\text{Na}^+$  ions. Insets are visualizations of molecule conformations at the PMF minimum. The hydrocarbon tail remains freely solvated in the triolein solvent and adopts a multitude of conformations.

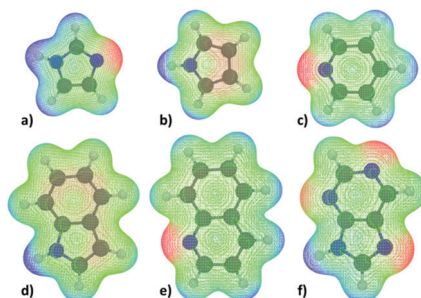


Fig. 6 Visualizations of the molecular electrostatic potential (MEP) of (a) imidazole, (b) pyrrole, (c) pyridine, (d) indole, (e) quinoline, and (f) purine superimposed onto the molecular structure. Red corresponds to regions of negative potential and blue to regions of positive potential as experienced by a positive test charge (*i.e.* proton).

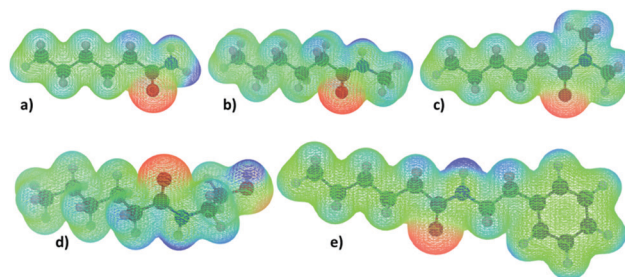


Fig. 7 Visualizations of the molecular electrostatic potential (MEP) of (a) hexamide, (b) *N*-methylhexamide, (c) *N*-dimethylhexamide, (d) *N*-ethanolhexamide, and (e) *N*-phenylethylhexamide superimposed onto the molecular structure. Red corresponds to regions of negative potential and blue to regions of positive potential as experienced by a positive test charge (*i.e.* proton).

MEP is visualized as a molecular surface that is based on either the van der Waals radii of the atoms, solvent accessible surface, or a constant value of electron density. The latter, with an isovalue of  $0.01e/a_0^3$ , where  $e$  is unit charge and  $a_0$  the Bohr radius, was used here.

The data of Fig. 6 and 7 shows that regions of negative electrostatic potential are concentrated at the nonprotonated basic nitrogens for *N*-heterocycles and at the carbonyl oxygens for the fatty acid amides. For *N*-ethanolhexamide, the alcohol hydroxyl oxygen introduces a second region of negative electrostatic potential. Comparison of the negative electrostatic potential regions with the adsorption geometries inferred from Fig. 3 and 5 reveals that MEP can be considered as a strong predictor for the site of adsorption *via* cation bridging for most

of the examined adsorbent species. However for molecules with multiple possible adsorption sites, such as purine, discrepancy between the MEP prediction of most negative potential sites and the MD-simulation predicted adsorption sites can be observed.

Next, we determined based on the electronic structure calculations the Hirshfeld,<sup>59</sup> AIM,<sup>75</sup> Merz-Singh-Kollman (MK),<sup>60,61</sup> and ChelpG<sup>62</sup> partial charges corresponding to each atom site in the adsorbents. The partial atomic charge is not a quantum mechanical observable and therefore is not uniquely defined. Partial charge is highly method dependent, as demonstrated here by our comparison of the different partial charge determining methods. Fig. 8 presents a comparison of these



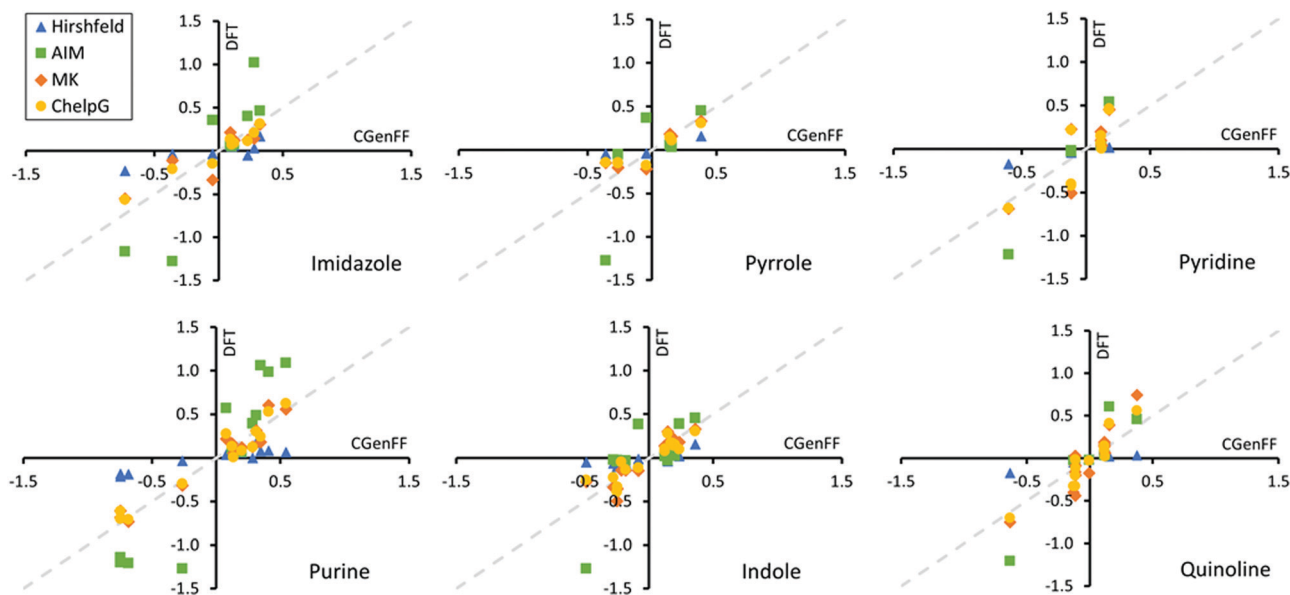


Fig. 8 Comparison of the DFT calculations based Hirshfeld, AIM, Merz-Singh-Kollman (MK), and ChelpG charges with the Charmm GenFF partial charges for the studied N-heterocycles. Deviation from diagonal (dashed grey line) corresponds with a mismatch to the partial charge.

DFT derived partial charges for the heterocycles to those of the Charmm GenFF. Exact values of the partial charges for the entire set of model compounds including the fatty acid amides are presented in Fig. S3–S5 of ESI.† Fig. 8 shows that both Hirshfeld and AIM charges differ significantly from Charmm GenFF, particularly for atoms at either extreme of the molecule charge distribution. The AIM partial charges are systematically larger and Hirshfeld charges smaller in magnitude in comparison to Charmm GenFF partial charges. Meanwhile, both MK and ChelpG result in values that match well with the Charmm GenFF partial charges. The data for the fatty acid amides is analogous, see ESI.†

To better understand the deviations, let us consider the charge determination schemes a bit more in detail. Hirshfeld population analysis involves partition of electron density in real space and assigned partial charge is dependent on the relative electron-withdrawing power of an atom relative to its environment.<sup>59,102</sup> AIM charges are determined by dividing the molecular charge density into atomic contributions based on overall molecular topology through construction of zero-flux surfaces of the electronic density. MK and ChelpG partial charge schemes aim to replicate the MEP at various grid points around the molecule.<sup>61,62</sup> The two methods differ in definition of these grid points. ChelpG uses a regularly spaced cubic lattice, while the MK scheme uses grid points on nested Connolly surfaces. Since the MK and ChelpG charges are fitted to reproduce MEP, atoms that lie deeply buried inside the molecule can vary in their charge which leads to uncertainty. On the other hand, the Charmm GenFF partial charges have been determined based on HF/6-31G(d) level theory interaction energies in aqueous environment, *e.g.*, matching interactions with single water molecules placed at configurations that would arise *via* hydrogen bonding and/or lone pair interaction sites.<sup>42</sup>

The resulting partial charges of atoms are tuned until sufficient agreement with quantum mechanical calculations based water interaction distance and interaction energies. Commonly MK charges at the MP2/6-31G(d) theory level are used as initial estimates for partial charges.<sup>42</sup> The Charmm GenFF parametrization compound set includes a wide variety of heterocycles, which often provide a scaffold for the majority of pharmaceuticals. This explains the good match between the MK, ChelpG and Charmm GenFF partial charges for the examined heterocycles. Mismatch between Charmm GenFF and the MK and ChelpG charges is greater for the fatty acid amides, see Fig. S5 of ESI,† for which parameter assignment was done mainly through analogy, accounting for neighbouring bonded atoms.<sup>103</sup> Notably, atomic charges can depend strongly on the conformation of the molecule. This means the partial charges of a fixed point charge model, as in the case of most biomolecular force-fields, should be averaged over several molecular conformations for flexible molecules, such as the amides here. Here, partial charges were determined based only on a single molecular conformation. The provided DFT based charges are to indicate systematic deviations due to the apolar solvent environment. However, based on the good match with Charmm GenFF, the effect of this dielectric screening is minimal. Here the continuum solvent model with a relative permittivity of 3.109 is closely comparable to calculations in gas phase.

The good match between Charmm GenFF and MEP derived partial charges (MK and ChelpG) indicates that the Charmm GenFF based molecular dynamics simulations employ charge distributions that can be considered to accurately represent the equivalent of DFT simulations MEP even in this apolar solvent environment. Thus, the predictions of adsorption geometries and trends in adsorption energies for molecules with a single likely adsorption site can be considered reliable in comparison





to quantum chemical calculations. This is supported by the good agreement with adsorption conformations from literature. However, subtle differences in electron density for molecules with multiple possible adsorption sites are clearly not captured by Charmm GenFF charge partition. For example, Charmm GenFF uses identical charges for the amide nitrogens and carbonyl oxygens of *N*-ethanolhexamide and *N*-phenylethylhexamide, despite the apparent differing chemical environments. Both MK and ChelpG predict more negative charges for both amide N and carbonyl O of *N*-phenylethylhexamide than those of Charmm GenFF. The persistence of this trend for other molecular configurations is unclear. Additionally for purine, Charmm GenFF assigns a charge difference of  $0.062e$  between the two nitrogens of the six-membered ring. This charge difference is not replicated by any of the charge calculation schemes. Together with the lack of polarization and proper charge screening in a low dielectric medium, such as the triolein solvent of this work, such charge distribution is likely to produce inaccurate adsorption geometries in the force-field based MD simulations. However, as the overall charge trends are in line with the various charge determination schemes, the qualitative response can be expected to persist.

## 4 Conclusions

The adsorption of six nitrogen containing heterocyclic compounds and five fatty acid amides from model vegetable oil solvent to model phyllosilicate adsorbent (Na-montmorillonite) was examined *via* MD simulations and DFT electronic structure calculations. Adsorption energies and configurations corresponding to single molecule physisorption on dehydrated Na-montmorillonite were determined. We find that the adsorption of the examined heterocyclic compounds occurs through interaction of the ring nitrogen, particularly the basic, non-protonated nitrogen, with the montmorillonite surface  $\text{Na}^+$  cations. The examined fatty amides adsorb *via* interaction of the carbonyl oxygen with the montmorillonite  $\text{Na}^+$  ions. Adsorption of the heterocycles is dictated by the basicity of the ring nitrogen and, when applicable, the strength of cation- $\pi$  interaction, while for the fatty acid amide series, the accessibility of the carbonyl group and the bulkiness of the head group govern the adsorption strength. Throughout, the adsorption takes place *via* a cation bridging mechanism, *i.e.* the montmorillonite surface  $\text{Na}^+$  ions are the adsorption site foci.

The observed cation bridging adsorption mechanism means that the type and valency of the montmorillonite surface cations, but also the presence of water on the surface can be expected to greatly affect adsorption strength. Notably, when excess water is present in the system, hydration of the montmorillonite surface and the polar headgroups of the adsorbates can be expected. We assessed also water adsorption to the montmorillonite surface. If water is present in the system to wet the surface to hydrated montmorillonite, the results leave under debate whether the N-compounds on a hydrated

montmorillonite surface adsorb *via* displacement of water from the surface to cationic sites or *via* water bridging.

The electronic structure calculations enabled determining the reliability and quality of the MD-simulation results. We conclude that the molecular electrostatic potential provides a good accuracy predictor for the site of adsorption for the here observed cation bridging type adsorption response. This good match with the MEP and the MD simulations results of the adsorption arises because for most of the examined molecules, the Charmm GenFF partial charges mirror the trends in molecular electron density. This was further analyzed *via* determining the Hirshfeld, AIM, MK and ChelpG partial charges. We noted that discrepancies between the DFT calculated electron density and Charmm GenFF arise for molecules with multiple possible adsorption sites and for molecules with poorer analogy in the force field parametrization set, such as the fatty acid amide head groups examined here. For such species, adsorption geometry predictions based on MD simulations remain unreliable. An additional source of error is the lack of polarization in the Charmm force field. This is especially apparent for interaction of ionic groups or ions in low dielectric mediums, such as the triolein solvent of this work. As a result, electrostatic interactions can be considered overestimated by a factor of 2.<sup>104</sup> This consideration leads to a direct conclusion that the adsorption energies presented in this work for physisorption are significant overestimates in the apolar solvent environment considered here.

We presented here a systematic examination of nitrogen compound adsorption from bio oils *via* computational means. The obtained work demonstrates MD simulations can be used to extract reliable information about selective nitrogen adsorption but the findings also raise warranted attention to limitations of the molecular modelling approaches in quantitatively accurate adsorption modelling for apolar solvent systems. The significance of the work is that the adsorption findings provide a design guideline for adsorption response of related systems but the findings also provide means to assess accuracy of molecular level modelling approaches for similar apolar solvent systems. Overall, we conclude that molecular dynamics based modelling serves as a valuable tool in providing molecular level description and guidelines of physisorption responses in bio-oils.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The authors thank Rasmus Kronberg and Prof. Kari Laasonen for assistance with the electronic structure calculations and for useful discussions. This work was partly financed by the Forum and Neste Foundation project 20210130 (M. V.). The authors acknowledge the financial support of the Academy of Finland Grant No. 309324 (M. S.). We are grateful for the support by the FinnCERES Materials Bioeconomy Ecosystem. Computational resources by CSC IT Centre for Science,



Finland, RAMI – RawMatTERS Finland Infrastructure, and Aalto Science-IT project are also gratefully acknowledged.

## Notes and references

- E. S. Shuba and D. Kifle, *Renewable Sustainable Energy Rev.*, 2018, **81**, 743–755.
- K. Satyanarayana, A. Mariano and J. Vargas, *Int. J. Energy Res.*, 2011, **35**, 291–311.
- X. Miao, Q. Wu and C. Yang, *J. Anal. Appl. Pyrolysis*, 2004, **71**, 855–863.
- B. Maddi, S. Viamajala and S. Varanasi, *Bioresour. Technol.*, 2011, **102**, 11018–11026.
- L. Sheng, X. Wang and X. Yang, *Bioresour. Technol.*, 2018, **247**, 14–20.
- O. Palardy, C. Behnke and L. M. Laurens, *Energy Fuels*, 2017, **31**, 8275–8282.
- S. Chiaberge, I. Leonardis, T. Fiorani, G. Bianchi, P. Cesti, A. Bosetti, M. Crucianelli, S. Reale and F. De Angelis, *Energy Fuels*, 2013, **27**, 5287–5297.
- S. Wang, Q. Wang, X. Jiang, X. Han and H. Ji, *Energy Convers. Manage.*, 2013, **68**, 273–280.
- A. Ross, J. Jones, M. Kubacki and T. Bridgeman, *Bioresour. Technol.*, 2008, **99**, 6494–6504.
- N. H. Zainan, S. C. Srivatsa and S. Bhattacharya, *Fuel*, 2015, **161**, 345–354.
- F. Li, L. Katz and S. Qiu, *Ind. Eng. Chem. Res.*, 2019, **58**, 3959–3968.
- F. Li, L. Katz and Z. Hu, *ACS Sustainable Chem. Eng.*, 2019, **7**, 16529–16538.
- M. Hazrat, M. Rasul and M. M. K. Khan, *Proc. Eng.*, 2015, **105**, 865–876.
- A. J. Hernández-Maldonado and R. T. Yang, *Angew. Chem.*, 2004, **116**, 1022–1024.
- P. Falaras, I. Kovanis, F. Lezou and G. Seiragakis, *Clay Miner.*, 1999, **34**, 221–232.
- E. Santillan-Jimenez, R. Pace, S. Marques, T. Morgan, C. McKelphin, J. Mobley and M. Crocker, *Fuel*, 2016, **180**, 668–678.
- B. N. Bhadra and S. H. Jhung, *Chem. Eng. J.*, 2020, **388**, 124195.
- P. A. Penttila, S. Vierros, K. Utriainen, N. Carl, L. Rautkari, M. Sammalkorpi and M. Österberg, *Langmuir*, 2019, **35**, 8373–8382.
- O.-P. Lehtinen, R. W. N. Nugroho, T. Lehtimaa, S. Vierros, P. Hiekkataipale, J. Ruokolainen, M. Sammalkorpi and M. Österberg, *Colloids Surf., B*, 2017, **160**, 355–363.
- S. Abel and M. Marchi, *J. Phys. Chem. B*, 2020, **124**, 11802–11818.
- S. Abel, N. Galamba, E. Karakas, M. Marchi, W. H. Thompson and D. Laage, *Langmuir*, 2016, **32**, 10610–10620.
- R. de Souza, R. Ratochinski, M. Karttunen and L. Dias, *J. Chem. Inf. Model.*, 2019, **60**, 522–536.
- A. Khoshnood and A. Firoozabadi, *Langmuir*, 2015, **31**, 5982–5991.
- S. Vierros and M. Sammalkorpi, *Phys. Chem. Chem. Phys.*, 2018, **20**, 6287–6298.
- S. Vierros and M. Sammalkorpi, *ACS Omega*, 2019, **4**, 15581–15592.
- M. Doig and P. J. Camp, *Phys. Chem. Chem. Phys.*, 2015, **17**, 5248–5255.
- M. Doig, C. P. Warrens and P. J. Camp, *Langmuir*, 2013, **30**, 186–195.
- J. L. Bradley-Shaw, P. J. Camp, P. J. Dowding and K. Lewtas, *Langmuir*, 2016, **32**, 7707–7718.
- J. L. Bradley-Shaw, P. J. Camp, P. J. Dowding and K. Lewtas, *Phys. Chem. Chem. Phys.*, 2018, **20**, 17648–17657.
- S. J. Eder, A. Vernes and G. Betz, *Langmuir*, 2013, **29**, 13760–13772.
- Y. Xiong, T. Cao, Q. Chen, Z. Li, Y. Yang, S. Xu, S. Yuan, J. Sjöblom and Z. Xu, *J. Phys. Chem. C*, 2017, **121**, 5020–5028.
- J. Zhong, P. Wang, Y. Zhang, Y. Yan, S. Hu and J. Zhang, *Energy*, 2013, **59**, 295–300.
- Q. Liu, S. Yuan, H. Yan and X. Zhao, *J. Phys. Chem. B*, 2012, **116**, 2867–2875.
- M. Vuorte, S. Vierros, S. Kuitunen and M. Sammalkorpi, *J. Colloid Interface Sci.*, 2020, **571**, 55–65.
- M. Munir, M. Ahmad, M. Rehan, M. Saeed, S. S. Lam, A. Nizami, A. Waseem, S. Sultana and M. Zafar, *Environ. Res.*, 2021, **193**, 110398.
- M. Munir, M. Ahmad, M. Mubashir, S. Asif, A. Waseem, A. Mukhtar, S. Saqib, H. S. H. Munawaroh, M. K. Lam and K. S. Khoo, *et al.*, *Bioresour. Technol.*, 2021, **328**, 124859.
- S. Vierros and M. Sammalkorpi, *J. Chem. Phys.*, 2015, **142**, 094902.
- S. Vierros and M. Sammalkorpi, *Phys. Chem. Chem. Phys.*, 2015, **17**, 14951–14960.
- I. V. Kopanichuk, V. A. Novikov, A. A. Vanin and E. N. Brodskaya, *J. Mol. Liq.*, 2019, **296**, 111960.
- D. P. Geerke and W. F. van Gunsteren, *ChemPhysChem*, 2006, **7**, 671–678.
- D. P. Geerke and W. F. van Gunsteren, *J. Phys. Chem. B*, 2007, **111**, 6425–6436.
- K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes and I. Vorobyov, *et al.*, *J. Comput. Chem.*, 2010, **31**, 671–690.
- M. J. Abraham, T. Murtola, R. Schulz, S. Pall, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19–25.
- H. Heinz, T. J. Lin, R. K. Mishra and F. S. Emami, *Langmuir*, 2013, **29**, 1754–1765.
- Y. J. Bae, C. Ryu, J.-K. Jeon, J. Park, D. J. Suh, Y.-W. Suh, D. Chang and Y.-K. Park, *Bioresour. Technol.*, 2011, **102**, 3512–3520.
- J. Henin, W. Shinoda and M. L. Klein, *J. Phys. Chem. B*, 2008, **112**, 7008–7015.
- W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- A. D. MacKerell Jr, D. Bashford, M. Bellott, R. L. Dunbrack Jr, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo and S. Ha, *et al.*, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
- M. J. E. O'Neil, *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 14th Revised edn, Elsevier Health Sciences, London, United Kingdom, 2006.



- 50 R. E. Weast, *Handbook of Chemistry and Physics*, 69th edn, CRC Press, Boca Raton, FL, 1988.
- 51 H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
- 52 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
- 53 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
- 54 M. Parrinello, A. Rahman and P. Vashishta, *Phys. Rev. Lett.*, 1983, **50**, 1073–1076.
- 55 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011–1021.
- 56 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 17.
- 57 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
- 58 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian'16 Revision B.01*, 2016, Gaussian Inc., Wallingford CT.
- 59 F. L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129–138.
- 60 U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1984, **5**, 129–145.
- 61 B. H. Besler, K. M. Merz Jr and P. A. Kollman, *J. Comput. Chem.*, 1990, **11**, 431–439.
- 62 C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, 1990, **11**, 361–373.
- 63 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
- 64 J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982–9985.
- 65 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 66 F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057–1065.
- 67 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
- 68 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.
- 69 I. Gouw and I. Vlugter, *Fette, Seifen, Anstrichm.*, 1967, **69**, 223–226.
- 70 L. A. Anderson, *Chemical Stimulation of Lipid Production in Microalgae and Analysis by NMR Spectroscopy for Biofuel Applications*, University of California, Davis, 2015.
- 71 W. Tang, E. Sanville and G. Henkelman, *J. Phys.: Condens. Matter*, 2009, **21**, 084204.
- 72 E. Sanville, S. D. Kenny, R. Smith and G. Henkelman, *J. Comput. Chem.*, 2007, **28**, 899–908.
- 73 G. Henkelman, A. Arnaldsson and H. Jónsson, *Comput. Mater. Sci.*, 2006, **36**, 354–360.
- 74 M. Yu and D. R. Trinkle, *J. Chem. Phys.*, 2011, **134**, 064111.
- 75 R. F. Bader, *Acc. Chem. Res.*, 1985, **18**, 9–15.
- 76 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics Modell.*, 1996, **14**, 33–38.
- 77 P. Mignon, P. Ugliengo and M. Sodupe, *J. Phys. Chem. C*, 2009, **113**, 13741–13749.
- 78 A. Gapeev, C.-N. Yang, S. J. Klippenstein and R. C. Dunbar, *J. Phys. Chem. A*, 2000, **104**, 3246–3256.
- 79 R. C. Dunbar, *J. Phys. Chem. A*, 1998, **102**, 8946–8952.
- 80 C. Ruan, Z. Yang, N. Hallowita and M. Rodgers, *J. Phys. Chem. A*, 2005, **109**, 11539–11550.
- 81 E. A. Orabi and G. Lamoureux, *J. Chem. Theory Comput.*, 2012, **8**, 182–193.
- 82 F.-Y. Lin and A. D. MacKerell Jr, *J. Comput. Chem.*, 2020, **41**, 439–448.
- 83 H. M. Khan, C. Grauffel, R. Broer, A. D. MacKerell Jr, R. W. Havenith and N. Reuter, *J. Chem. Theory Comput.*, 2016, **12**, 5585–5595.
- 84 A. Turupcu, J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.*, 2020, **16**, 7184–7194.
- 85 R. Lin, A. Ladshaw, Y. Nan, J. Liu, S. Yiacoumi, C. Tsouris, D. W. DePaoli and L. L. Tavlarides, *Ind. Eng. Chem. Res.*, 2015, **54**, 10442–10448.
- 86 K. S. Walton, M. B. Abney and M. D. LeVan, *Microporous Mesoporous Mater.*, 2006, **91**, 78–84.
- 87 H. Yi, F. Jia, Y. Zhao, W. Wang, S. Song, H. Li and C. Liu, *Appl. Surf. Sci.*, 2018, **459**, 148–154.
- 88 H. Li, B. J. Teppen, C. T. Johnston and S. A. Boyd, *Environ. Sci. Technol.*, 2004, **38**, 5433–5442.
- 89 S. Mecozzi, A. P. West and D. A. Dougherty, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 10566–10571.
- 90 A. R. Swoboda and G. W. Kunze, *Clays Clay Miner.*, 1964, **13**, 277–288.
- 91 V. Ram, A. Sethi, M. Nath and R. Pratap, *The Chemistry of Heterocycles: Chemistry of Six to Eight Membered N,O, S, P and Se Heterocycles*, Elsevier Science, 2019.
- 92 S. Karaborni, B. Smit, W. Heidug, J. Urai and E. van Oort, *Science*, 1996, **271**, 1102–1104.
- 93 M. Segad, B. Jonsson, T. Åkesson and B. Cabane, *Langmuir*, 2010, **26**, 5782–5790.
- 94 T. Underwood, V. Erastova, P. Cubillas and H. C. Greenwell, *J. Phys. Chem. C*, 2015, **119**, 7282–7294.
- 95 A. D. MacKerell Jr, *J. Comput. Chem.*, 2004, **25**, 1584–1604.
- 96 T. Fröhlking, M. Bernetti, N. Calonaci and G. Bussi, *J. Chem. Phys.*, 2020, **152**, 230902.
- 97 I. Vorobyov, V. M. Anisimov, S. Greene, R. M. Venable, A. Moser, R. W. Pastor and A. D. MacKerell, *J. Chem. Theory Comput.*, 2007, **3**, 1120–1133.
- 98 I. V. Vorobyov, V. M. Anisimov and A. D. MacKerell, *J. Phys. Chem. B*, 2005, **109**, 18988–18999.



- 99 G. Klesse, S. Rao, S. J. Tucker and M. S. Sansom, *J. Am. Chem. Soc.*, 2020, **142**, 9415–9427.
- 100 H. Li, J. Chowdhary, L. Huang, X. He, A. D. MacKerell Jr and B. Roux, *J. Chem. Theory Comput.*, 2017, **13**, 4535–4552.
- 101 T. Simonson and C. L. Brooks, *J. Am. Chem. Soc.*, 1996, **118**, 8452–8458.
- 102 S. Saha, R. K. Roy and P. W. Ayers, *Int. J. Quantum Chem.*, 2009, **109**, 1790–1806.
- 103 K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell Jr, *J. Chem. Inf. Model.*, 2012, **52**, 3155–3168.
- 104 I. Leontyev and A. Stuchebrukhov, *Phys. Chem. Chem. Phys.*, 2011, **13**, 2613–2626.

