

# Adsorption of non-ionic surfactants on organoclays in drilling fluid investigated by molecular descriptors and Monte Carlo random walk simulations

Dina Kania<sup>a,b,\*</sup>, Robiah Yunus<sup>a,b</sup>, Rozita Omar<sup>a</sup>, Suraya Abdul Rashid<sup>a,c</sup>, Badrul Mohamed Jan<sup>d</sup>, Akmal Aulia<sup>e</sup>

<sup>a</sup> Department of Chemical and Environmental Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Malaysia

<sup>b</sup> Institute of Plantation Studies, Universiti Putra Malaysia, 43400 UPM, Serdang, Malaysia

<sup>c</sup> Institute of Advanced Technology, Universiti Putra Malaysia, 43400 UPM, Serdang, Malaysia

<sup>d</sup> Department of Chemical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

<sup>e</sup> Integration and Analytics, Carigali Hess Operating Company Sdn. Bhd., 50450 Kuala Lumpur, Malaysia

## ARTICLE INFO

### Keywords:

Hydrophobicity  
Organoclay  
Drilling fluid  
Non-ionic surfactant  
Molecular descriptors  
Monte Carlo simulations

## ABSTRACT

Non-ionic surfactants have been used as a rheology control additive in drilling fluids to prevent flocculation of solids, such as organoclays, and maintain mud dispersion. Here, the fundamental phenomena involved in non-ionic surfactant adsorption on organoclays and how it affects the rheology of synthetic-based drilling fluids were elucidated by the analysis of molecular descriptors and Monte Carlo simulations. Using the Random Forests machine learning algorithm software, the non-ionic surfactant adsorption on organoclays was found to be affected mainly by the hydrophobicity and molecular shape of hydrophobic chains of non-ionic surfactants. Molecular descriptor calculations indicated that the hydrophobic interaction and van der Waals forces are dominant factors. Then, the Monte Carlo simulations provided the empirical effect of the non-ionic surfactant hydrophobic chains on their self-assembly to organoclays. Molecules with two chains were found to be easily adsorbed to the surface due to molecular size, while molecules with three and four chains occupy more sites and interact with each other more frequently, forming larger clusters. Consequently, non-ionic surfactants with more hydrophobic chains were predicted to improve rheology of drilling fluids and form stable mud emulsions. This approach is useful in predicting the effects of new additive in drilling fluid formulation.

## 1. Introduction

Organoclays are hydrophobic materials produced by modifying clays and clay minerals with various organic compounds through an intercalation process [1,2]. In drilling fluid application, organoclays are widely used as rheological additives to build the viscosity of synthetic-based mud (SBM) [3]. The viscosity is a result of interfacial interactions and interparticle interaction of organoclay particles [4–6]. However, due to the contamination of certain cations such as calcium or low pH condition during drilling a wellbore, clay layers interact and form a flocculated structure, resulting in high rheological properties and a high amount of fluid loss [4]. In such cases, chemical thinners are used to neutralise the positive charge on the edge of clay layers and prevent particle flocculation.

For non-ionic thinners, steric stabilization is the term that takes

place for the stabilization of colloidal particles by uncharged macromolecules to prevent them from flocculating [7,8]. The adsorbed non-ionic thinners alter the rheological behaviour of clay suspensions accordingly, mostly reducing the yield point and gel structure of the mud. The yield point defines the ability of the drilling fluid to remove the drilling cuttings from below the drill bit to the surface through the annulus [9,10] and the value is affected by the electrostatic attraction between clay particles [4,5,11].

In our previous study, we investigated three types of non-ionic polyol ester surfactants as a thinner and a lubricant for SBM [12]. These long-tailed non-ionic surfactants had a strong thinning ability, reducing mud yield point to 50% at a concentration of 1% (v/v). We found that higher molecular weight polyol ester with more hydrophobic chains caused a higher rheology. This may be due to the ability of branched molecule to bind the adjacent lattice of clay together, but its high

\* Corresponding author at: Department of Chemical and Environmental Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Malaysia.

E-mail address: [dinakania@gmail.com](mailto:dinakania@gmail.com) (D. Kania).

<https://doi.org/10.1016/j.apsusc.2020.148154>

Received 4 August 2020; Received in revised form 17 September 2020; Accepted 21 September 2020

Available online 14 October 2020

0169-4332/ © 2020 Elsevier B.V. All rights reserved.

molecular weight causes a more expanded clay lattice. In a recent study, Madejová et al. [13] investigated the arrangement of cationic surfactants in organoclay by IR and NMR spectroscopy. It was found that with the increasing of alkyl chain length, surfactants adopted more ordered conformation. However, in drilling fluid, it is difficult to measure the self-assembly of surfactants in the mud in the laboratory, as the drilling fluids are very thick and many other chemicals and solid additives are also presented in the system.

Molecular simulation is usually used to provide the underlying mechanism for adsorption of liquid additives to solids from the molecular perspective [14,15]. Currently, molecular descriptors are widely used to predict quantitative structure relationship of chemical substances [16–18]. Molecular descriptors represent molecules by generating different combinations of structure representation, atom/bond weighting schemes, and mathematical functions [19]. For example, molecular descriptors can identify parameters that are significant to the adsorption of non-ionic polymers. Song et al. [20] found that the hydrophobic number and hydrophobic block length are strongly correlated with the affinity between non-ionic polymers and a thermoplastic polymer resin. Zhao et al. [21] found that hydrophobicity, polar surface area and molecular weight were the main factors affecting the adsorption affinity model and the capacity of cationic chemicals on an activated charcoal. It also had 64% accuracy with the measured adsorption affinity. Recently, Brusseau [22] found molar volume descriptor effective for predicting the quantitative structure–property relationship (QSPR) of non-ionic surfactants adsorption. The predictive molar volume indicated a reasonable representation of the influence of molecular size on the driving force of hydrophobic adsorption.

Monte Carlo simulations use repeated random generation to estimate distributions and obtain statistical properties for various situations [23]. Monte Carlo simulations have been used to understand fundamental aspects on crystal growth on a surface [24], adsorption of oxygen and hydrogen on metal surfaces [25], adsorption of organic micropollutants on hydrophobic zeolites [26], diffusion process using one dimensional Monte Carlo random walk method [27,28], adsorption of CO<sub>2</sub> and water in the interlayer of Na-montmorillonite [29,30], and many more. The number of uniform  $n$  random points generated is an essential part of each Monte Carlo system. Ferreira et al. [31] developed an accurate representation of the pore size distribution gamma-alumina support catalyst, porosity, and internal surface area with the standard error on the average surface area being 0.6% for a population of 30 samples, which is considered to be sufficiently adequate. By using Monte Carlo simulations, Göttl et al. [24] have created an intuitive picture for graphene growth on germanium and given insights into the driving forces of the crystals. In other study, Makaremi et al. [29] used multiphase Gibbs ensemble Monte Carlo simulations to study the structural and transport properties of equilibrium water – CO<sub>2</sub> mixtures in the interlayer of clays. The algorithm employed three basic trial moves of particles: particle displacement, particle exchange, and volume change.

Random walk method has been used to model the adsorption of non-ionic surfactants and polymers. Cory and Rodgers [32] suggested that the adsorption of non-ionic surfactants onto a surface is due to the van der Waal's interaction, which forms a random monolayer. The model was simple but realistic in describing the molecule deposition. Budinski-Petković and Kozmidis-Luburić [33] also studied the irreversible adsorption of directed self-avoiding random walks of various lengths on a square lattice with equal deposition probability using Monte Carlo simulations. The deposition rate increased with the number of components in the mixture and the deposition rate slowed down nearing the saturation state. Brown and Chakrabarti [34] described that in aqueous media, the hydrophobic chains of polymers tend to aggregate via a short-range attractive interaction in order to shield each other from the solvent. Accordingly, the formation of structures modified the rheological properties of the solution. Balazs et al. [35] reviewed that hydrophobic interactions of non-ionic surfactants drove

the chains to self-assemble spontaneously. The review summarized the effect of branched chains in polymers. At the same concentration, the polymer having four chains formed larger aggregates, approximately two to four times larger than the structures that are created from the two-chain polymers. As a result, the solution containing the four-sticker polymer had higher viscosity than the two-sticker polymer solution containing the smaller aggregates.

Based on previous studies, the use of chemical databases to reflect molecular properties of chemicals combined with Monte Carlo random-walk simulation may be appropriate for predicting, visualizing and analyzing data in drilling fluids. In this study, we predicted the self-assembly of three non-ionic surfactants on organoclays using molecular descriptors and a Monte Carlo simulation. Structural and physico-chemical descriptors were used to classify and provide insights into the associated structural properties of different non-ionic surfactants that influences their adsorption processes on a hydrophobic surface. Then, the effect of branched structure of non-ionic surfactants on their adsorption to organoclays was investigated via Monte Carlo random walk method. This is the first time that the calculation of the molecular structure of chemicals combined with the Monte Carlo simulations are used to explain the underlying phenomenon in the drilling fluid. The present work is an extension of our previous research in the area of adsorption of polyol esters on hydrophobic surfaces. The understanding of the non-ionic surfactant arrangement on the organoclay surface is important in selecting and optimizing the use of surfactant for drilling fluid applications.

## 2. Methods

Two stages of the method were performed: (1) molecular descriptors and data processing; (2) Monte Carlo random walk simulations. In the first stage, the molecular descriptors of each non-ionic surfactants were computed by ChemoPy software. Random Forests was then used to rank the importance of a series of descriptors and eliminate variables with small importance. The objective of the first stage was to obtain a set of descriptors that are most related to the yield point of SBM. In the second stage, Monte Carlo random walk simulations used random number generation to move a molecule of on the lattice and to deposit it based on probability. This stage was aimed to investigate the deposition of molecules having different number of chains/branching and to correlate with the surface phenomena in drilling fluids.

### 2.1. Non-ionic surfactants

Three types of non-ionic surfactants were investigated, namely neopentyl glycol ester (NPGE; C<sub>41</sub>H<sub>76</sub>O<sub>4</sub>), trimethylolpropane ester (TMPE; C<sub>60</sub>H<sub>110</sub>O<sub>6</sub>), and pentaerythritol ester (PEE; C<sub>77</sub>H<sub>140</sub>O<sub>8</sub>). As shown in Fig. 1, NPGE has two hydrophobic chains, TMPE has three hydrophobic chains, and PEE has four hydrophobic chains. The isomeric SMILES (simplified molecular input line entry system) codes of each non-ionic surfactant were then generated from ChemSpider for molecular descriptor calculations. SMILES was developed by Weininger [36] and is widely used to express the molecular structure of a molecule [37].

### 2.2. Molecular descriptors and data processing

The SMILES of each non-ionic surfactants were submitted to the ChemoPy software to describe the structure of each surfactant molecule, a method used by Cao et al. [38]. ChemoPy software calculated a multitude of descriptors based on 1D, 2D, and 3D-input of each non-ionic surfactant molecule, using the SMILES as the input. The ChemoPy software can calculate a total of 1135 descriptors. These descriptors represent a variety of physicochemical (such as hydrophobicity, molecular weight, and molar refractivity), topological (such as E-state indices, topological indices, connectivity and shape indices), and

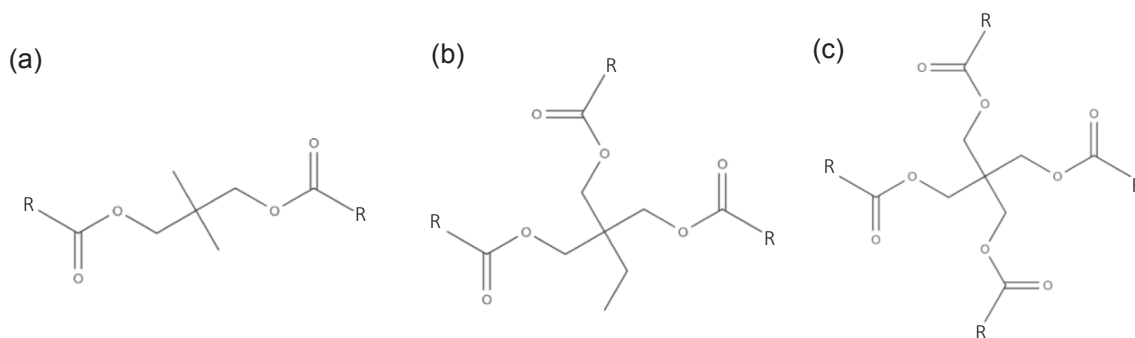


Fig. 1. Molecular structure of different non-ionic surfactants (a) NPGE, (b) TMPE, and (c) PEE.

electrostatic (such as topological polar surface area, partial charges, and van der Waals surface area) properties.

The descriptors with missing values (or null) were excluded from the dataset during the screening process. Then, the experimental yield point of synthetic-based mud (SBM) measured from our previous study [12] were then correlated with molecular descriptors to find the important descriptors relative to the yield point using the Random Forests' out-of-bag (OOB) samples [39]. The analysis was performed using yield point data of SBM containing 2% non-ionic surfactant, which is 5 lb/100ft<sup>2</sup>, 8 lb/100ft<sup>2</sup> and 14 lb/100ft<sup>2</sup> for NPGE, TMPE, and PEE, respectively. Random Forest is a regression tree-based group learning technique consisting of a collection of unpruned regression trees used collectively to determine the output value. The Random Forests screening was carried out in the Weka software, which is a collection of samples that were not used in the construction of each tree that makes up the model. The measure of importance method can yield a tornado chart that ranks all variables from the most important to the least important. Such method employs a permutation approach that shuffles the values of each descriptor separately. If a permuted descriptor vector impacts the prediction significantly relative to another descriptor vector, this permuted descriptor is more important than the other descriptor.

### 2.3. Monte Carlo random walk simulations

The Monte Carlo simulations were performed using a random-walk approach based on a previous study by Budinski-Petković and Kozmidis-Luburić [33] to correlate the physical conditions of non-ionic surfactants with surface adsorption phenomena. The concept was to deposit the non-ionic surfactant molecules as monolayers on a two-dimensional square lattice, while other conditions and restrictions were adapted from Budinski-Petković and Kozmidis-Luburić [33] and Cory and Rodgers [32] as described below.

The methodology flowchart for Monte Carlo simulations are presented in Fig. 2. The simulation domain was a 100 × 100 square lattice used as a system boundary, as illustrated in Fig. 3. The adsorption site boundary with the size of 60 × 60 square lattice was located in the middle of the domain. The system boundary was referred as the “boundary”, and the adsorption sites boundary as the “organoclay”. The boundary was represented by a set  $\mathcal{B}$  such that an arbitrary point

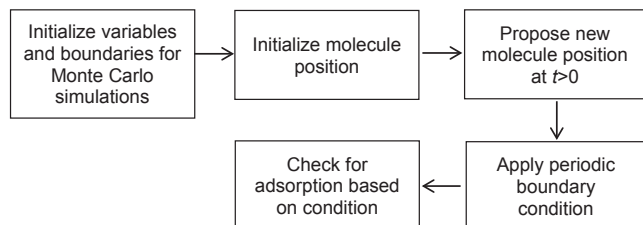


Fig. 2. Methodology flowchart for Monte Carlo random walk simulations.

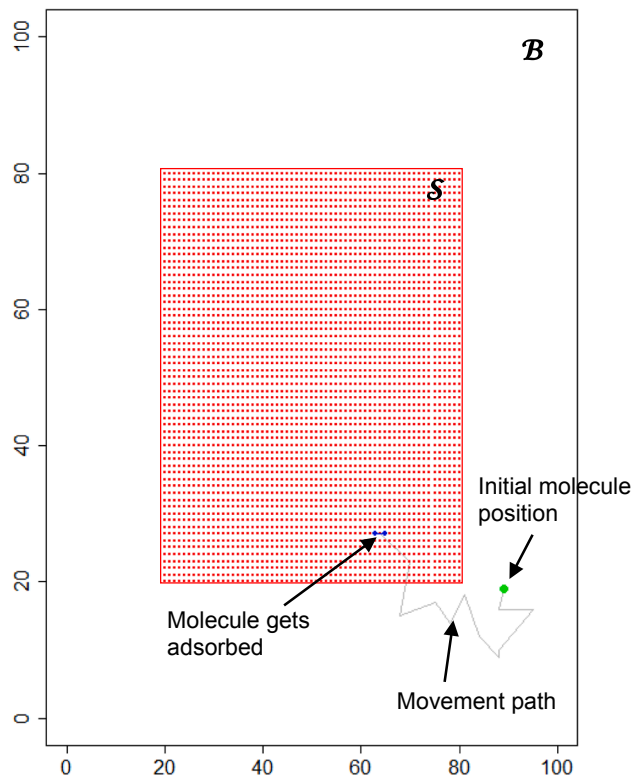


Fig. 3. The simulation domain with the hypothetical of molecule movement path.

$b = (b_x, b_y) \in \mathcal{B} \subset \mathbb{N}^2$ . Hence,  $\{b_1, b_2, \dots, b_{n_b}\} = \{(b_{x_1}, b_{y_1}), (b_{x_2}, b_{y_2}), \dots\} \in \mathcal{B}$ . The clay particle was represented by a set  $\mathcal{S}$  such that an arbitrary point  $s = (s_x, s_y) \in \mathcal{S} \subset \mathcal{B}$ . Similarly, we had the set of all points in site, that is,  $\{s_1, s_2, \dots, s_{n_s}\} = \{(s_{x_1}, s_{y_1}), (s_{x_2}, s_{y_2}), \dots\} \in \mathcal{S}$ .

Fig. 3 also illustrates how a molecule moves from a random site to a sorption site. There are two types of system boundary: first, the boundary system, which is the system domain, and second, the boundary of an organoclay ( $\mathcal{S}$ ). Hence,  $\mathcal{S} \subseteq \mathcal{B}$ . The initial molecule position is randomly placed outside of  $\mathcal{S}$  (a green dot). The adsorption site ( $\mathcal{S}_c$ ) of the hydrophobic chain inside the  $\mathcal{S}$  is shown as red dots. The study focused on the interaction of different structure of polyol esters on the surface. Thus, the mixed aggregates formed due to adsorption were not calculated.

Next, the step size,  $n_s$ , was assigned, such that  $n_s \in \mathbb{Z}^+$  and  $n_s \geq 1$ , while the location of non-ionic surfactant,  $l$ , was defined as  $P_l(t) = (x_l(t), y_l(t))$ . The system was time dependent because the molecules moved to different positions as time elapsed during the mixing of drilling fluid. Hence, this location at the next time step was defined as,

$$\begin{bmatrix} x_l(t + \Delta t) \\ y_l(t + \Delta t) \end{bmatrix} = \begin{bmatrix} x_l(t) \\ y_l(t) \end{bmatrix} + \begin{bmatrix} \mathcal{U}(-n_s, n_s) \\ \mathcal{U}(-n_s, n_s) \end{bmatrix} \quad (1)$$

$$\Delta x_l(t + \Delta t) = x_l(t + \Delta t) - x_l(t) = \mathcal{U}(-n_s, n_s)$$

$$\Delta y_l(t + \Delta t) = y_l(t + \Delta t) - y_l(t) = \mathcal{U}(-n_s, n_s)$$

$$P_l(t + \Delta t) = P_l(t) + \langle \mathcal{U}(-n_s, n_s), \mathcal{U}(-n_s, n_s) \rangle \quad (2)$$

Note that  $\mathcal{U}(-n_s, n_s)$  was a uniform distribution bounded by  $-n_s$  and  $n_s$ . The vector  $\Delta P(t) = \langle \Delta x(t), \Delta y(t) \rangle$ . Hence, it was simply written as,

$$P_l(t + \Delta t) = P_l(t) + \Delta P_l(t + \Delta t) \quad (3)$$

Where

$$\Delta P_l(t + \Delta t) = \begin{bmatrix} \mathcal{U}(-n_s, n_s) \\ \mathcal{U}(-n_s, n_s) \end{bmatrix} \quad (4)$$

The total number of non-ionic surfactant molecules initially present in the system was defined as  $n_l$ . Each of the lubricant molecules would travel through different routes due to  $\Delta P(t)$ . The number of simulation steps was defined as  $n_t$ . The total amount of polyol ester that get adsorbed onto the organoclay sites can be different as  $n_t$  was varied.

The maximum number of molecules within the boundary of the system was defined as  $\beta$ . Therefore, for a given adsorption site  $\mathcal{S}_c$  at a given coordinate  $(x, y)$ , the number of molecules can be less than or equal to  $\beta$ . The adsorption probability was set and defined as  $n_p$ . Note that  $n_p$  and  $\beta$  is the same for all organoclay sites  $\mathcal{S}_c \in \mathcal{S}$ .

The polyol ester molecule moved in a periodic boundary condition. Hence, if  $P_l(t + \Delta t) \notin \mathcal{B}$ , then the molecule movement is described as,

$$x_l(t + \Delta t) = \begin{cases} x_l(t + \Delta t) & 0 \leq x_l(t + \Delta t) \leq n_{\mathcal{B}} \\ x_l(t + \Delta t) - n_{\mathcal{B}} & n_{\mathcal{B}} < x_l(t + \Delta t) \\ x_l(t + \Delta t) + n_{\mathcal{B}} & x_l(t + \Delta t) < 0 \end{cases} \quad (5)$$

$$y_l(t + \Delta t) = \begin{cases} y_l(t + \Delta t) & 0 \leq y_l(t + \Delta t) \leq n_{\mathcal{B}} \\ y_l(t + \Delta t) - n_{\mathcal{B}} & n_{\mathcal{B}} < y_l(t + \Delta t) \\ y_l(t + \Delta t) + n_{\mathcal{B}} & y_l(t + \Delta t) < 0 \end{cases} \quad (6)$$

To determine whether a molecule had been adsorbed, the locations of the last vertex of each chain were observed. In the case of a 2-chain non-ionic surfactant molecule, the vertices of each chain were  $(x_l + 1, y_l)$  and  $(x_l - 1, y_l)$  at  $\mathcal{S}_c$ . For a 3-chain non-ionic surfactant molecule, the vertices of the chains would be  $\{(x_l + 1, y_l), (x_l - 1, y_l), (x_l, y_l - 1)\}$ . For a 4-chain molecule, the vertices of the hydrophobic chains were  $\{(x_l + 1, y_l), (x_l - 1, y_l), (x_l, y_l - 1), (x_l, y_l + 1)\}$  on a organoclay site. The illustration of each type of non-ionic surfactant is shown in Fig. 4.

The simulations were run to calculate the percentage of organoclay sites occupied by non-ionic surfactants and the percentage of non-ionic

**Table 1**  
Range of parameters for Monte Carlo simulations.

| Parameters                              | Unit    | Range   |
|-----------------------------------------|---------|---------|
| Number of steps ( $n_t$ )               | –       | 25–200  |
| Non-ionic surfactant molecule ( $n_l$ ) | –       | 25–200  |
| Step size ( $n_s$ )                     | lattice | 1–5     |
| Adsorption probability ( $n_p$ )        | –       | 0.2–0.6 |

surfactant adsorbed on the organoclay, varying the input parameters such number of polyol ester molecule ( $n_l$ ), number of steps ( $n_t$ ), step size ( $n_s$ ), and the probability of non-ionic surfactant to be adsorbed ( $n_p$ ). The range of input parameters was given in Table 1. Finally, the simulation output, namely the percentage of organoclay sites occupied by non-ionic surfactants and the percentage of non-ionic surfactants adsorbed to the organoclay, were calculated using Eqs. (7) and (8). The simulations were repeated three times and the mean value was reported.

$$\%organoclay \text{ sites occupied} = \frac{\mathcal{S}_o}{\mathcal{S}_c} \times 100\% \quad (7)$$

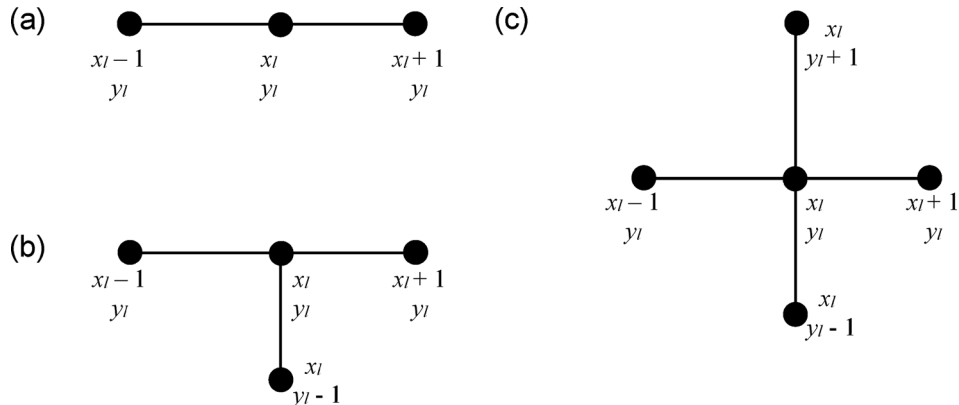
$$\%surfactant \text{ adsorbed} = \frac{l_{\mathcal{A}}}{n_l} \times 100\% \quad (8)$$

### 3. Results and discussion

In this study, computational molecular calculations and the Monte Carlo adsorption simulation were used to gain an understanding of the extent of the adsorption mechanism of non-ionic surfactants on hydrophobic surfaces such as organoclay. The nature of non-ionic surfactant molecules, such as molecular weight, polarity, and hydrophobicity, was considered, while the Monte Carlo random walk simulations focused on the effects of non-ionic surfactant structure on the formation of adsorption layers on organoclay particles.

#### 3.1. Molecular structure descriptors

The molecular structure descriptors were calculated using the ChemoPy software, providing a total of 1135 descriptors. The screening of important descriptors that affect drilling fluid's yield point was conducted using Random Forests regression tree-based group learning technique, leading to 18 important descriptors of non-ionic surfactants. The screening was conducted since not all descriptors have a relation to the measured yield point, which is the resistance of initial flow of fluid. Their importance rank is summarized in Table 2. The analysis was performed using our previous experimental yield point data, which is 5 lb/100ft<sup>2</sup>, 8 lb/100ft<sup>2</sup> and 14 lb/100ft<sup>2</sup> for SBM containing 2% of NPGE, TMPE, and PEE, respectively, while the yield point of a blank

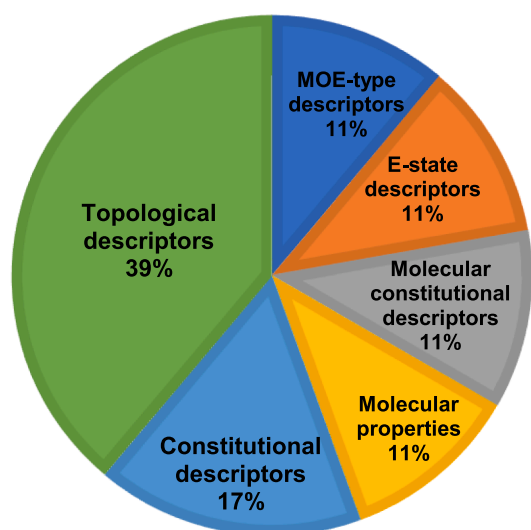


**Fig. 4.** The vertices of a molecule having (a) two hydrophobic chains (b) three hydrophobic chains, and (c) four hydrophobic chains.



**Table 2**  
Selected ChemoPy descriptors ranked by Random Forests.

| Rank | Descriptors                                           | Description                                        | Non-ionic surfactant type |         |         |
|------|-------------------------------------------------------|----------------------------------------------------|---------------------------|---------|---------|
|      |                                                       |                                                    | NPGE                      | TMPE    | PEE     |
| 1    | <i>EstateVSA1</i>                                     | MOE E-state indices and surface area contributions | 30.5                      | 37.7    | 75.9    |
| 2    | <i>Shet</i>                                           | Sum of the E-State indices of hetero atoms         | 35.3                      | 56.1    | 77.3    |
| 3    | <i>Nhev</i>                                           | Number of heavy atoms                              | 45                        | 66      | 85      |
| 4    | <i>MW</i>                                             | Molecular weight                                   | 556.4                     | 816.6   | 1052.8  |
| 5    | <i>MR</i>                                             | Molar refractivity                                 | 194.6                     | 283.9   | 364.1   |
| 6    | <i>Nhyd</i>                                           | Count of hydrogen atoms                            | 76                        | 110     | 140     |
| 7    | <i>logP<sup>2</sup></i>                               | Square of logP value                               | 173.5                     | 365.7   | 590.2   |
| 8    | <i>Tsch</i>                                           | Schiultz index                                     | 51,692                    | 130,384 | 236,440 |
| 9    | <i>W</i>                                              | Wiener index                                       | 13,372                    | 33,601  | 60,808  |
| 10   | <i>IDET</i>                                           | Total information index on distance equality       | 4705.1                    | 10584.9 | 17,874  |
| 11   | <i>nta</i>                                            | Number of all atoms                                | 121                       | 176     | 225     |
| 12   | <i>LabuteASA</i>                                      | Labute's Approximate Surface Area                  | 280.5                     | 410.0   | 526.8   |
| 13   | <i>ISIZ</i>                                           | Molecular size index                               | 837.1                     | 1312.8  | 1758.1  |
| 14   | <i>ZM1</i>                                            | Zagreb index with order 1                          | 184                       | 270     | 348     |
| 15   | <i>Thara</i>                                          | Harary number                                      | 161.9                     | 281.9   | 407.4   |
| 16   | <i>TIAC</i>                                           | Total information index on atomic composition      | 134.6                     | 196.9   | 253.4   |
| 17   | <i>Shev</i>                                           | Sum of E-State indices of non-hydrogen atoms       | 86.5                      | 126.7   | 163.9   |
| 18   | <i>ncarb</i>                                          | Count of C atoms                                   | 41                        | 60      | 77      |
|      | Experimental yield point, lb/100 ft <sup>2</sup> [12] |                                                    | 5                         | 8       | 14      |



**Fig. 5.** Distribution of various descriptor types correlated with mud yield point.

sample (without non-ionic surfactant) was 19 lb/100ft<sup>2</sup> [12]. Each descriptor has been colour-coordinated by the type of descriptor, i.e. constitutional descriptors, topological descriptors, connectivity indices, electropological state (E-state) indices, autocorrelation descriptors, charge descriptors, and molecular properties, as shown in Fig. 5. The Random Forests regression has a high correlation ( $R^2 = 0.89$ ) with the yield point of drilling fluid. However, since the calculation dealt with too many descriptors and most of the descriptors are not related to the yield point, the correlation has quite high RMSE, which is 7.14.

As shown in Table 2, the most important descriptors are *EstateVSA1*, which is a molecular operating environment (MOE) type descriptor, that calculates the sum of van der Waals surface area (VSA) contributions to E-state. The *EstateVSA1* value of PEE, which is 75.9, is much higher than the values of TMPE and NPGE (37.7 and 30.5, respectively), indicating that PEE has the highest van der Waals surface area that contributes to the interaction with clay surfaces. Meanwhile,

the second most important descriptor, *Shet*, represents the sum of the E-state indices of heteroatoms. Heteroatom is any atom that is not carbon or hydrogen. For non-ionic surfactants used in this study, heteroatom refers to oxygen. The E-state index combines the electronic state of the bonded atom within the molecule with its topological nature in the context of the whole molecular skeleton [40]. PEE has the highest sum E-state indices (*Shet*) of 77.3 because PEE has the most lone pairs of valence electrons of its oxygen atoms in the polyol groups.

It was previously described that the higher yield point for SBM containing non-ionic polyol ester was due to larger clay aggregates formed which caused by higher molecular weight polyol ester [12]. Based on Table 2, *nhev* (the number of heavy atoms) and *MW* (molecular weight) are important descriptors. *MW* is strongly linked to molecular size. PEE has the highest *MW* of 1052.8 compared to TMPE and NPGE (816.6 and 556.4, respectively), which can provide a larger van de Waals attraction to the clay surface. The *nhev* of PEE is highest compared to TMPE and NPGE, which indicates that the molecular mass had a significant effect on the hydrophobic adsorption of non-ionic compounds, as was the case in the previous work by Song et al. [20]. Similarly, yield point of SBM increases with the increase in descriptors' values related to the molecular size, including *MR* (molar refractivity), *nta* (number of all atoms), *ISIZ* (molecular size index), and *ncarb* (count of C atoms).

The *logP<sup>2</sup>* descriptor measures the hydrophobic nature of the molecules [41] and therefore has a strong correlation with non-ionic surfactant adsorption. *LogP<sup>2</sup>* of NPGE, TMPE, and PEE is 173.5, 365.7, and 590.2, respectively, showing that PEE has higher hydrophobic contributions of atoms to its molecule. Czajka et al. [42], Song et al. [20] and Rosen [43] explained that the affinity of non-ionic polymer or surfactant with the surface strongly depends on the length and the number of hydrophobic chains. An increase in hydrophobic chain units will increase the Gibbs free energy ( $\Delta G$ ) of adsorption, thereby increasing the efficiency of adsorption. Thus, a greater number in hydrophobic chains of non-ionic surfactant theoretically increase the adsorption. Crucial changes of yield point of drilling fluid occurred when the non-ionic surfactant present, which are possibly caused by the adsorption of non-ionic surfactants on organoclay. The results in Table 2 show that the yield point of drilling fluid increases with an increase of hydrophilic/hydrophobic index of non-ionic surfactant.

*Tsch*, *W*, *IDET*, *ISIZ*, *ZM1*, *Thara*, and *TIAC* are topological descriptors that characterize the constitution and configuration of a molecule by a single index number. The indices derive from the distance matrix in both topological and topographic forms [44] and the values are used to establish the relationship between the structure of the compounds and the physicochemical properties. The results show that the increase in these topological indices has resulted in higher mud yield point. However, these indices cannot be intuitively linked to the current understanding of mud yield point. It is common that some descriptors, which serve as numerical descriptions, cannot be directly interpreted [45]. These types of descriptors are mostly useful in developing quantitative structure–activity relationship (QSAR) models and diversity analysis.

It was found that molecular descriptors commonly describing the adsorption between the organic molecule and the hydrophilic surface were unimportant to the yield point. For example, descriptors such as *TPSA* (topological polar surface area), *pol* (polarity) and *naccr* (H-bond acceptor) had no significant attribute toward yield point. *TPSA* and *PSA* (polar surface area) descriptors reflect to the capacity of the molecule to form H-bonds due to oxygen or nitrogen [46,47]. The polarity of non-ionic surfactants represents its polar head which has high interaction with metal surfaces. Polar sites cannot be easily adsorbed onto hydrophobic surfaces, while the non-polar and weak polar ones are reverse [48,49]. In SBM, the hydrophilic group of non-ionic surfactants may participate in hydrogen bonds when the oxygen atom of alcohol and the brine form H-bonds. Since the H-bond is unlikely to happen between its hydrophobic tails and organoclay, the H-bond acceptor descriptor,

*naccr*, was found to be insignificant to yield point. This further proved that the yield point of the mud is not affected by H-bond interaction.

As shown in Fig. 5, topological descriptors are dominating (39%). Topological descriptors describe numerical values based on a hydrogen-suppressed graph representation of the molecule. It shows that yield point, which is normally affected by the attractive force among colloidal particles in drilling fluid, depends on the topology of non-ionic surfactant such as size, shape, symmetry, branching and cyclicity. Seven topological descriptors were considered to be significant descriptors. The next important category is the constitutional descriptors (17%) representing the number of atomic contributions. This includes molecular weight, which is proportional to the ester chain number of non-ionic surfactants. The number of ester chains relates to the hydrophobic effect that causes the adsorption of non-ionic surfactants. Other descriptor types suggest that the contribution of molecular mass, hydrophobicity, and molar refractivity are important, which is consistent with the fact that the non-ionic surfactant tail site is hydrophobic in nature. As a result, the hydrophobic tails can bind to the hydrophobic surface of organoclay. Hence, the architecture of non-ionic surfactant, i.e. two-chain, three-chain, or four-chain, matters on the adsorption. Meanwhile, the contribution of partial charge descriptors is not shown, indicating that electronic effects of non-ionic surfactant are not significant to yield point.

Overall, molecular descriptors have classified and provided insights into the associated structural properties of non-ionic surfactants that influences their adsorption processes on a hydrophobic surface, such as hydrophobicity, molecular size, molecular weight, and molecular shape. The findings also showed that the polarity and hydrophilicity of non-ionic surfactants in the interaction between non-ionic surfactant and organoclay are negligible. Hence, hydrophobic interactions and van der Waals forces are dominant in describing the interaction between non-polar tails of non-ionic surfactant and organoclay.

### 3.2. Self-assembly behaviour of polyol ester chains

Monte Carlo simulation was conducted in order to develop better understanding of the molecular arrangement of non-ionic surfactant molecules with different number of branching on a sorbent surface (organoclay). The model is simplistic in nature. Non-ionic surfactant molecules consisting of an aliphatic (non-polar) tail and a polar head group were employed. Different molecule structures have a constant chain length with a variation in the number of chains (i.e. 2-chain for NPGE, 3-chain for TMPE, and 4-chain for PEE). Through molecular descriptor analysis, it was suggested that the mechanisms of non-ionic surfactant adsorption on the organoclay are mainly physical adsorption, due to hydrophobic self-assembly interaction and van der Waals forces. The changes in the molecular architecture of various surfactants are believed to impact their ability to attach to the surface. The effects of parameters such as number of steps ( $n_t$ ), number of non-ionic surfactant molecule ( $n_l$ ), step size ( $n_s$ ), and the probability of non-ionic surfactant to be adsorbed ( $n_p$ ), were defined in the R programming language. Fig. 6 shows the example of simulation output for different types of non-ionic surfactant molecules.

First, a step size of  $n_s = 5$ , adsorption probability  $n_p = 0.4$ , and non-ionic surfactant molecule  $n_l = 100$  conditions were used to study the number of steps ( $n_t$ ) dependence. The simulations were run at  $n_t$  between 25 and 200. The  $n_t$  was counted by the number of steps for depositing the molecules and scaled by the total number of lattice sites. Fig. 7 shows the mean values of non-ionic surfactant adsorbed (%) and organoclay site occupied by the non-ionic surfactant (%) at different number of steps ( $n_t$ ). Error bars represent the standard deviation of the data. It is shown that branching the molecules to 3- and 4-chain decreases the number of adsorbed molecules. NPGE is shown to have more organoclay sites adsorbed compared to TMPE and PEE. Schenk et al. [50] have explained that the linear molecules are packed more efficiently within the adsorbate structure, expressed in molecules/unit

cell, compared to the branched molecules due to size entropy effects. This is because the smaller molecule finds it easier to fill the “gaps” at high loads within the adsorbate.

However, since one molecule of PEE can occupy more sites due to a larger number of chains, the percentage of sites occupied is higher than NPGE, i.e. for 200 number of steps, PEE occupied 8.22% of site, while NPGE occupied 5% of site. The adsorption efficiency and dispersion of surfactants are greatly affected by the chain of surfactants [35]. Since all molecule chains obeyed the excluded volume criterion, which means that one adsorption site in the lattice can be occupied by one monomer only at the same time, the chance of PEE and TMPE molecules to become adsorbed is less than NPGE, approximately 17%.

Next, the effect of the number of non-ionic surfactant molecules was simulated by varying the  $n_l$ , as shown in Fig. 8. Fixed values of number of steps  $n_t = 100$ , step size  $n_s = 5$ , and adsorption probability  $n_p = 0.4$  were used. As  $n_l$  increased, the relationship between  $n_l$  and the percentage of non-ionic surfactant adsorbed on the clay shows no clear trend. This means that the number of  $n_l$  not adsorbed in  $\mathcal{S}$  is comparable to the number of  $n_l$  in  $\mathcal{S}_c$ . This behaviour can be understood by considering the one-site-one-monomer restriction. Increased number of molecules moving in a random-walk direction may cause molecule jamming at a late stage of deposition due to inaccessible sorbent site [33]. If the site is inaccessible, the simulator would stop the attempt to deposit the molecule, but would increase the time by one unit. Nevertheless, the percentage of organoclay occupied increased in a fixed lattice boundary ( $100 \times 100$ ) when  $n_l$  increased.

Simulations were also performed for various length of random walks (step size) defined as  $n_s$ , and the results are shown in Fig. 9. Fixed values of number of steps  $n_t = 100$ , non-ionic surfactant molecule  $n_l = 100$ , and adsorption probability  $n_p = 0.4$  were used. The plot of  $n_s$  vs non-ionic surfactant adsorbed indicates that it is still in line with the previous results. The 2-chain NPGE molecules are easier to get adsorbed due to its molecular shape, hence  $n_l$  of NPGE adsorbed is greater by more than 16% compared to molecules having more chains.

Lastly, the adsorption probability ( $n_p$ ) would affect the number of molecules adsorbed, for example that the walk may pass through any lattice site multiple times before it gets adsorbed. Fixed values of number of steps  $n_t = 100$ , non-ionic surfactant molecule  $n_l = 100$ , and step size of  $n_s = 5$ , were applied. The results in Fig. 10 (a) indicate that the percentage of molecule adsorbed increases with the adsorption probability. For molecules with a branched chain structure (3- and 4-chain), the percentage of molecule adsorbed is quite similar. The similarity between TMPE and PEE is that these molecules are more difficult to “stack” than 2-chain molecules such as NPGE. This was supported by Jiang et al. [51] and Vlught et al. [52] which described that linear molecules and molecule with less side chain adsorb to a greater extent on a surface and have a higher packing efficiency and higher loadings per unit cell. A more branched molecule had to seek residence in the channel interiors which is energetically more demanding and therefore required disproportionately higher pressures.

Based on the Monte Carlo random-walk simulations, non-ionic surfactants with a different number of chains have resulted in a specific deposition trend. As noted above, the two hydrophobic chains could easily be adsorbed to the sorbent due to the simpler shape and smaller molecular size of NPGE (556.4 g/mol), compared to the 3- and 4-chain molecules (816.6 and 1052.8 g/mol for TMPE and PEE, respectively). However, as the 4-chain molecule has a more branched-chain and usually heavier molecular weight, this molecule may occupy more sorbent sites. This finding would explain the relationship between the molecular structure of the non-ionic surfactant and the viscosity of SBM. The ability of PEE to occupy more organoclay sites might suggest its ability to form tails and loops that interact with each other, leading to larger aggregates. Thus, SBM containing PEE has higher yield point (14 lb/100ft<sup>2</sup>) compared to the yield point of drilling fluid with TMPE (8 lb/100ft<sup>2</sup>) or NPGE (5 lb/100ft<sup>2</sup>). The non-ionic surfactant with three or four hydrophobic chains and higher molecular weight

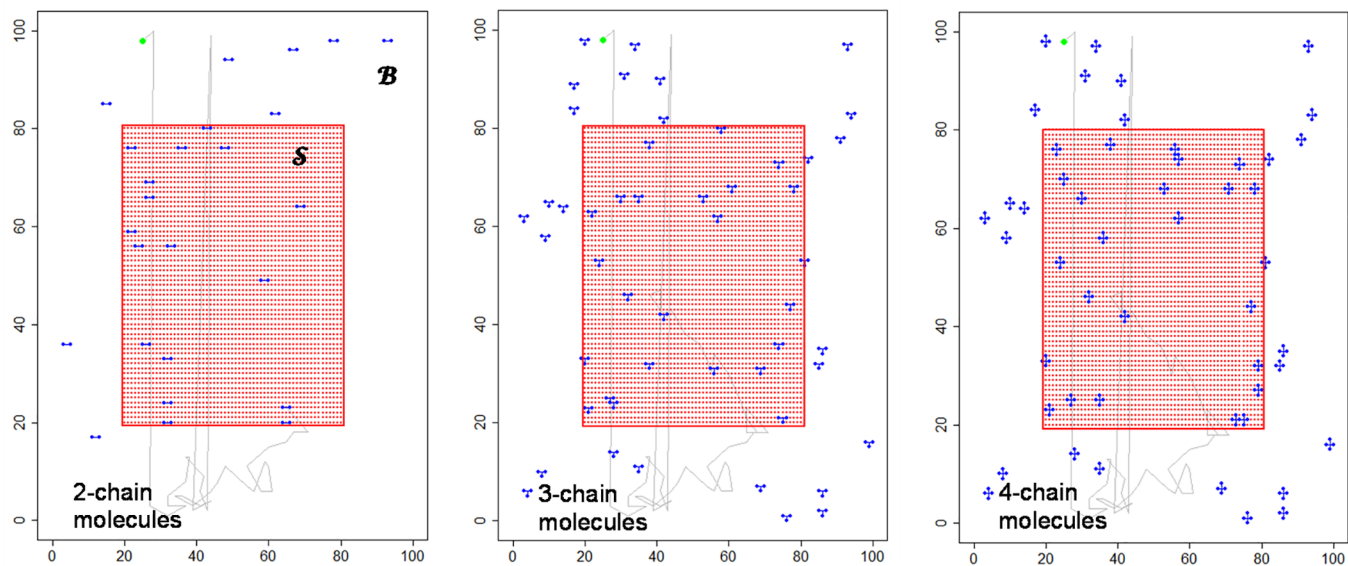
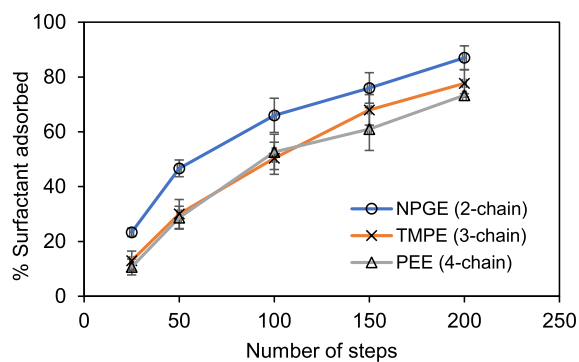
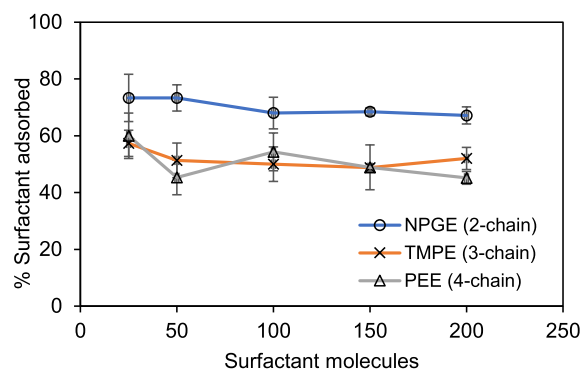


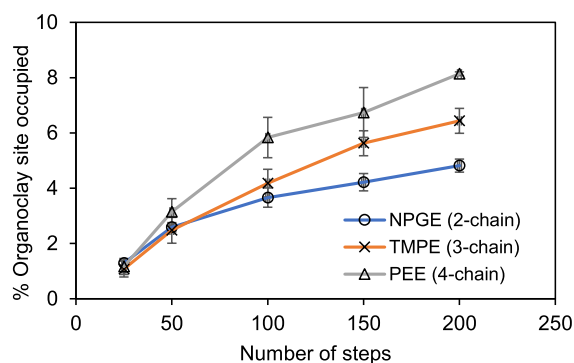
Fig. 6. Predicted simulation output of different molecule chains.



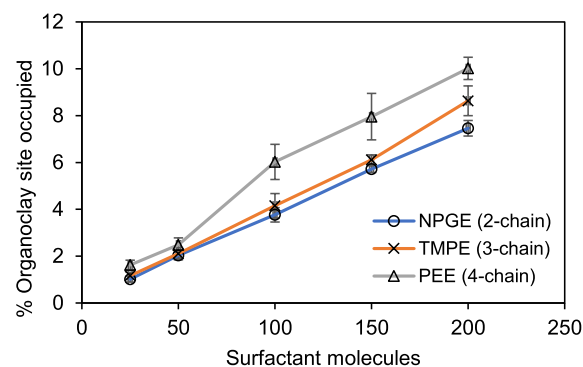
(a)



(a)



(b)



(b)

Fig. 7. Percentage of (a) molecule adsorbed and (b) organoclay occupied at different number of steps.

(PEE = 1052.8 g/mol; TMPE = 816.6 g/mol) would result in higher hydrophobic driving force and bridging flocculation when adsorbed on the organoclay surface.

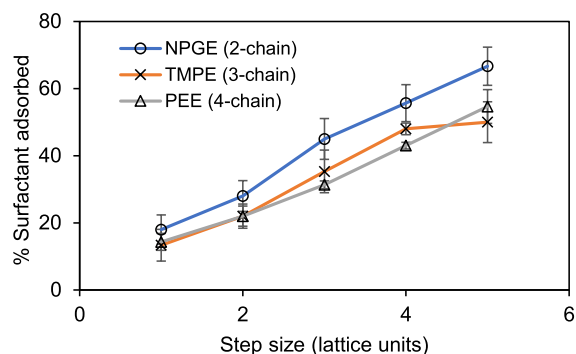
The findings from the Monte Carlo simulations are also able to provide the general picture of the possible deposition process of non-ionic surfactant onto organoclay surfaces and how it implies to the overall mud viscosity. However, to externally validate the simulation works in this study and to measure the non-ionic surfactant adsorption capacity on organoclay in an SBM system is currently not feasible.

Fig. 8. Percentage of (a) molecule adsorbed and (b) organoclay occupied at different molecule concentration.

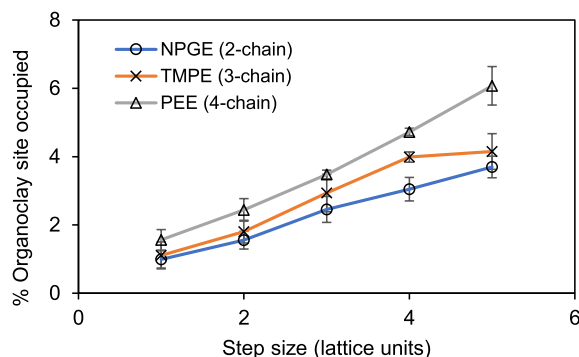
Nonetheless, by comparing our previous experimental findings with our current simulations, the important factors leading to the aggregation process and the macroscopic properties of the mud should be established.

#### 4. Conclusion

The molecular descriptors screening has shown that hydrophobicity, molecular size, molecular weight, and molecular shape are the main

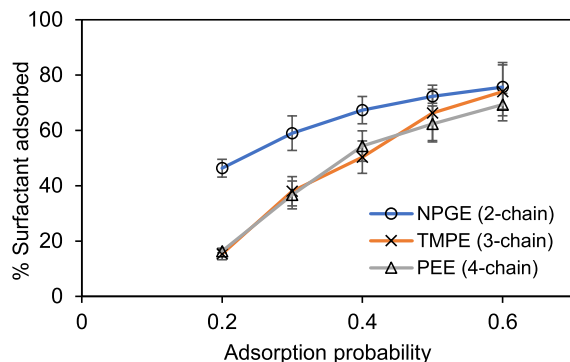


(a)

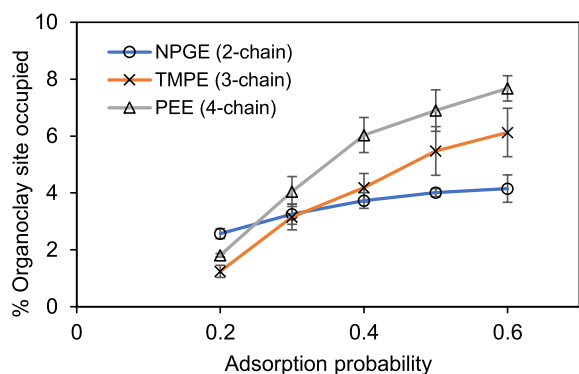


(b)

**Fig. 9.** Percentage of (a) molecule adsorbed and (b) organoclay occupied at different step size.



(a)



(b)

**Fig. 10.** Percentage of (a) molecule adsorbed and (b) organoclay occupied at different adsorption probability.

properties of non-ionic surfactants that affect the rheology, particularly yield point, of synthetic-based mud. The hydrophobic surface of organoclay facilitates strong interactions with non-ionic surfactants NPGE, TMPE, and PEE. Meanwhile, hydrophobic segments in non-ionic surfactants provide the affinity to anchor molecules on the organoclay and their high molecular mass can provide maximum adsorption density. *EstateVSA1* and *Shet* are the most important descriptors to mud yield point, representing the van der Waals surface area contributions and the E-state indices of oxygen, respectively. Overall, 39% of important descriptors are classified as topological descriptors, showing that the topology of different non-ionic surfactants such as size, shape, symmetry, branching and cyclicity affect the yield point of SBM the most. The next 17% important descriptors relate to the number of atomic contributions such as molecular weight, which is proportional to the ester chain number of non-ionic surfactants. These findings are in good agreement with the results of our previous experimental study where PEE has the highest molecular weight and number of heavy atoms, allowing PEE can form larger organoclay aggregates.

The Monte Carlo simulations were conducted to study the effect of branched structure of non-ionic surfactant molecules, i.e. two, three and four hydrophobic chains for NPGE, TMPE, and PEE, respectively, on their self-assembly to organoclays. Fewer branches of NPGE molecules have been found to be more adsorbed. This is because NPGE molecules with two hydrophobic chains pack more efficiently within the sorbent structure due to size entropy effects compared to the three and four-chain molecules. The smaller molecule also tends to fill the gaps between organoclay surface easier at high loads. Yet, molecules with more branches such as TMPE and PEE can occupy more sites due to a larger number of hydrophobic chains. To relate with the experimental yield point of the mud, the ability of PEE to occupy more organoclay sites might suggest its ability to form tails and loops that interact with each other, leading to larger aggregates. Thus, SBM containing PEE has better rheological properties. The use of higher molecular weight non-ionic surfactants with four hydrophobic chains such as PEE may be suitable for mud with a high solid content and where a more stable emulsion is desired. This study shows that the computational molecular calculations is useful to correlate the structure–property relationship of chemical additives with drilling fluid properties. Knowing the relationship and importance of chemical structure to drilling fluid properties by machine learning models can benefit researchers in developing new chemical additives.

#### CRedit authorship contribution statement

**Dina Kania:** Conceptualization, Methodology, Investigation, Writing - original draft. **Robiah Yunus:** Supervision, Writing - review & editing. **Rozita Omar:** Supervision. **Suraya Abdul Rashid:** Supervision. **Badrul Mohamed Jan:** Supervision. **Akmal Aulia:** Software, Visualization, Data curation.

#### Declaration of Competing Interest

None.

#### Acknowledgement

We gratefully appreciate the financial support from Ministry of Higher Education Malaysia for FRGS grant with reference number TK04/UPM/01/1.

#### References

- [1] M. Borisover, et al., Thermal treatment of organoclays: Effect on the aqueous sorption of nitrobenzene on n-hexadecyltrimethyl ammonium montmorillonite, *Appl. Surf. Sci.* 256 (17) (2010) 5539–5544.
- [2] H.P. He, J. Zhu, Chapter 10 - Analysis of Organoclays and Organic Adsorption by



- Clay Minerals, in: *Developments in Clay Science*, W.P. Gates, et al., Editors. 2017, Elsevier. p. 310-342.
- [3] G. Zhuang, Z. Zhang, M. Jaber, Organoclays used as colloidal and rheological additives in oil-based drilling fluids: An overview, *Appl. Clay Sci.* 177 (2019) 63–81.
  - [4] K.S. Hafshejani, A. Moslemizadeh, K. Shahbazi, A novel bio-based deflocculant for bentonite drilling mud, *Appl. Clay Sci.* 127–128 (2016) 23–34.
  - [5] A.A. Sulaimon, B.J. Adeyemi, M. Rahimi, Performance enhancement of selected vegetable oil as base fluid for drilling HPHT formation, *J. Petrol. Sci. Eng.* 152 (2017) 49–59.
  - [6] D. Zhou, et al., Influence of different exchangeable cations (Li<sup>+</sup>, Na<sup>+</sup> and Ca<sup>2+</sup>) on the modification effects and properties of organomontmorillonites used in oil-based drilling fluids/muds, *RSC Adv.* 5 (110) (2015) 90281–90287.
  - [7] D.H. Napper, Steric stabilization, *J. Colloid Interface Sci.* 58 (2) (1977) 390–407.
  - [8] P.F. Luckham, S. Rossi, The colloidal and rheological properties of bentonite suspensions, *Adv. Colloid Interface Sci.* 82 (1) (1999) 43–92.
  - [9] W.C. Lyons, G.J. Plisga, *Standard Handbook of Petroleum and Natural Gas Engineering*, Elsevier Science, Houston, 2011.
  - [10] H.C.H. Darley, G.R. Gray, R. Caenn, *Composition and Properties of Drilling and Completion Fluids*, Gulf Publishing Company, Book Division, 2011.
  - [11] E. Kusirini, et al., Synthesis, characterization, and performance of graphene oxide and phosphorylated graphene oxide as additive in water-based drilling fluids, *Appl. Surf. Sci.* 506 (2020) 145005.
  - [12] D. Kania, et al., Nonionic polyol esters as thinner and lubricity enhancer for synthetic-based drilling fluids, *J. Mol. Liq.* 266 (2018) 846–855.
  - [13] J. Madejová, et al., Conformation heterogeneity of alkylammonium surfactants self-assembled on montmorillonite: Effect of head-group structure and temperature, *Appl. Surf. Sci.* 503 (2020) 144125.
  - [14] S.M. Lahalih, I.S. Dairanieh, Development of novel polymeric drilling mud dispersants, *Eur. Polym. J.* 25 (2) (1989) 187–192.
  - [15] D. Ji, et al., Adsorption of C<sub>5</sub>Pe molecules on silica surfaces with different hydrophobicity studied by molecular dynamics simulation, *Appl. Surf. Sci.* 495 (2019) 143624.
  - [16] A.K. Ghose, G.M. Crippen, Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions, *J. Chem. Inf. Comput. Sci.* 27 (1) (1987) 21–35.
  - [17] B. Chandrasekaran, et al., Chapter 21 - Computer-Aided Prediction of Pharmacokinetic (ADMET) Properties, in: R.K. Tekade, Editor, *Dosage Form Design Parameters*, 2018, Academic Press. p. 731-755.
  - [18] T.N.V. de Souza, et al., Adsorption of basic dyes onto activated carbon: Experimental and theoretical investigation of chemical reactivity of basic dyes using DFT-based descriptors, *Appl. Surf. Sci.* 448 (2018) 662–670.
  - [19] A. Mauri, V. Consonni, R. Todeschini, *Molecular Descriptors, Handbook of Computational Chemistry*, Springer, Jackson, 2017, pp. 2065–2093.
  - [20] J. Song, W.E. Krause, O.J. Rojas, Adsorption of polyalkyl glycol ethers and triblock nonionic polymers on PET, *J. Colloid Interface Sci.* 420 (2014) 174–181.
  - [21] Y. Zhao, et al., Adsorptive interaction of cationic pharmaceuticals on activated charcoal: Experimental determination and QSAR modelling, *J. Hazard. Mater.* 360 (2018) 529–535.
  - [22] M.L. Bruseau, The Influence of Molecular Structure on the Adsorption of PFAS to Fluid-Fluid Interfaces: Using QSPR to Predict Interfacial Adsorption Coefficients, *Water Res.* (2019).
  - [23] H.Z. Al Garni, A. Awasthi, Chapter 20 - A Monte Carlo approach applied to sensitivity analysis of criteria impacts on solar PV site selection, in: P. Samui, et al., (Eds.) *Handbook of Probabilistic Models*, 2020, Butterworth-Heinemann. p. 489-504.
  - [24] F. Götl, et al., Exploring driving forces for length growth in graphene nanoribbons during chemical vapor deposition of hydrocarbons on Ge(0 0 1) via kinetic Monte Carlo simulations, *Appl. Surf. Sci.* 527 (2020) 146784.
  - [25] F.O. Sanchez-Varretti, et al., Monte Carlo simulations and cluster-exact approximation applied to H/Cu(100), H/Ag(100) and O/Cu(100) systems, *Appl. Surf. Sci.* 500 (2020) 144034.
  - [26] N. Jiang, et al., The adsorption mechanisms of organic micropollutants on high-silica zeolites causing S-shaped adsorption isotherms: An experimental and Monte Carlo simulation study, *Chem. Eng. J.* 389 (2020) 123968.
  - [27] A. Novikov, D. Kuzmin, O. Ahmadi, Random walk methods for Monte Carlo simulations of Brownian diffusion on a sphere, *Appl. Math. Comput.* 364 (2020) 124670.
  - [28] L. Farnell, W.G. Gibson, Monte Carlo simulation of diffusion in a spatially non-homogeneous medium: A biased random walk on an asymmetrical lattice, *J. Comput. Phys.* 208 (1) (2005) 253–265.
  - [29] M. Makaremi, et al., Multiphase Monte Carlo and Molecular Dynamics Simulations of Water and CO<sub>2</sub> Intercalation in Montmorillonite and Beidellite, *J. Phys. Chem. C* 119 (27) (2015) 15112–15124.
  - [30] E.M. Myshakin, et al., Molecular Dynamics Simulations of Turbostratic Dry and Hydrated Montmorillonite with Intercalated Carbon Dioxide, *J. Phys. Chem. A* 118 (35) (2014) 7454–7468.
  - [31] S. Ferreira, et al., Random porous network generation and 1D mass transfer simulation for gamma-alumina supports, in: A. Espuña, M. Graells, L. Puigjaner (Eds.) *Computer Aided Chemical Engineering*, 2017, Elsevier. p. 91-96.
  - [32] H. Cory, G.J. Rodgers, A three stage model for adsorption of nonionic surfactants, *J. Chem. Phys.* 99 (11) (1993) 8908–8913.
  - [33] L. Budinski-Petković, U. Kozmidis-Luburić, Irreversible deposition of directed self-avoiding random walks on a square lattice, *Physica A* 262 (3) (1999) 388–395.
  - [34] G. Brown, A. Chakrabarti, Structure formation in self-associating polymer and surfactant systems, *J. Chem. Phys.* 96 (4) (1992) 3251–3254.
  - [35] A.C. Balazs, K. Huang, T. Pan, Modeling of amphiphilic polymers and their interactions with nonionic surfactants, *Colloids Surf.*, A 75 (1993) 1–20.
  - [36] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, *J. Chem. Inform. Comput. Sci.*, 1988. 28(1): p. 31-36.
  - [37] F.M. Green, I.S. Gilmore, M.P. Seah, G-SIMS and SMILES: Simulated fragmentation pathways for identification of complex molecules, amino acids and peptides, *Appl. Surf. Sci.* 255 (4) (2008) 852–855.
  - [38] D.-S. Cao, et al., QSAR analysis of the effects of OATP1B1 transporter by structurally diverse natural products using a particle swarm optimization-combined multiple linear regression approach, *Chemometr. Intell. Lab. Syst.* 130 (2014) 84–90.
  - [39] A.L. Teixeira, J.P. Leal, A.O. Falcao, Random forests for feature selection in QSPR Models - an application for predicting standard enthalpy of formation of hydrocarbons, *J. Cheminf.* 5 (1) (2013) 9-9.
  - [40] L.H. Hall, L.B. Kier, Electrotopological State Indices for Atom Types: A Novel Combination of Electronic, Topological, and Valence State Information, *J. Chem. Inf. Comput. Sci.* 35 (6) (1995) 1039–1045.
  - [41] A. Weinebeck, et al., A new QSPR-based prediction model for biofuel lubricity, *Tribol. Int.* 115 (2017) 274–284.
  - [42] A. Czajka, G. Hazell, J. Eastoe, Surfactants at the Design Limit. 2015(1520-5827 (Electronic)).
  - [43] M.J. Rosen, Relationship of structure to properties in surfactants. III. Adsorption at the solid-liquid interface from aqueous solution, *J. Am. Oil Chem. Soc.* 52 (11) (1975) 431–435.
  - [44] J. Galvez, et al., Charge Indexes. New Topological Descriptors, *J. Chem. Inf. Comput. Sci.* 34 (3) (1994) 520–525.
  - [45] O. Mekenyan, M. Cronin, J. Madden, In Silico Toxicology: Principles and Applications, Royal Society of Chemistry, Cambridge, 2010.
  - [46] P. Ertl, B. Rohde, P. Selzer, Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties, *J. Med. Chem.* 43 (20) (2000) 3714–3717.
  - [47] G. Caron, G. Ermondi, Molecular descriptors for polarity: the need for going beyond polar surface area, *Future Med. Chem.* 8 (17) (2016) 2013–2016.
  - [48] X. Zhang, et al., Adsorption of VOCs onto engineered carbon materials: A review, *J. Hazard. Mater.* 338 (2017) 102–123.
  - [49] L.K. Koopal, G.T. Wilkinson, J. Ralston, Adsorption of interacting long-chain surfactant molecules: Isotherm equations, *J. Colloid Interface Sci.* 126 (2) (1988) 493–507.
  - [50] M. Schenk, et al., Separation of Alkane Isomers by Exploiting Entropy Effects during Adsorption on Silicalite-1: A Configurational-Bias Monte Carlo Simulation Study, *Langmuir* 17 (5) (2001) 1558–1570.
  - [51] J. Jiang, et al., Adsorption and separation of linear and branched alkanes on carbon nanotube bundles from configurational-bias Monte Carlo simulation, *Phys. Rev. B* 72 (2005).
  - [52] T.J.H. Vlugt, R. Krishna, B. Smit, Molecular Simulations of Adsorption Isotherms for Linear and Branched Alkanes and Their Mixtures in Silicalite, *J. Phys. Chem. B* 103 (7) (1999) 1102–1118.