Developing an Amorphous Kerogen Molecular Model Based on Gas Adsorption Isotherms

Hyeonseok Lee1*, Mehdi Ostadhassan1, Kouqi Liu1, Bailey Bubach1

1Department of Petroleum Engineering, University of North Dakota, Grand Forks, ND 58202, United States
*Corresponding author e-mail: hyeonseok.lee@und.edu, mehdi.ostadhassan@und.edu

Abstract

Computational modeling has been increasingly utilized to study kerogen or amorphous organic material. Additionally, simulation techniques have been acquired to analyze phenomenon that involves organic matter such as gas adsorption/desorption in recent years. This is happening due to the attention that organic matter is receiving in regards to its role in exploiting shale plays and carbon sequestration in rock layers with an abundance of organic matter. Kerogen models are continuously being developed but still require further research due to its complexity and variability. In this study, three different kerogen models, with varying chemical compositions from the Bakken, were constructed through molecular simulation and quantum mechanics calculations. To build the initial models, first, data from 13-C nuclear magnetic resonance (13C-NMR), X-ray photoelectron spectroscopy (XPS), and X-ray absorption near-edge structure (S-XANES) were input into the molecular dynamics simulation software (LAMMPS). Then, to verify our model configuration and chemical composition accuracy, N2 and CO2 gas adsorption behavior of the models were predicted and compared with the isotherm experimental results on physical samples that had a similar geochemical composition and thermal maturity. We realized the adsorption of CO2 is more significant compared to N2 which was also observed similarly in the experiments. Finally, although there was some discrepancy between the adsorption patterns in experiments and simulation, which was inferred to the pore structure dissimilarities, there was a good overall agreement in the isotherm and adsorption outcome.

Keywords: organic matter; Bakken; macro-molecular model; gas adsorption.

1. Introduction
The decline in the production from conventional resources and the increase in energy consumption has encouraged growth in shale oil and gas according to the U.S. Department of Energy (DOE) and U.S. Energy Information Administration (EIA). This shift in focus towards unconventional resources, specifically shale oil and gas, now dominates the energy industry, leading to continuous developments on how to extract from these reservoirs [4, 5]. Shale oil and gas, however, are expensive to be extracted compared to their conventional counterparts and require more advanced technologies [1, 8, 30]. Since shale plays are highly heterogeneous and complex, their exploitation requires a better insight and deeper understanding of their constituent components. In other words, shale play characterization is a task that requires a larger variety of experimental and computational methods. Among these components, organic matter or kerogen, which is the source of hydrocarbons [6, 9], is poorly understood compared to inorganic minerals in regards to its molecular structure. Kerogen is the amorphous and insoluble portion of the organic matter and due to its molecular origin, the composition is very complex [2, 6].

The kerogen compound is organic and mainly consists of carbon, hydrogen, oxygen, nitrogen, and sulfur. Major structural and compositional changes can occur to the molecule as it undergoes maturation, due to the high temperatures and pressures it endures, as a function of depth and burial [22]. This chemical transformation is a complex procedure which makes predicting the stereochemistry properties of kerogen highly exacting. Therefore, building the macro-molecular model for kerogen is challenging and 3D models are continuously being developed and corrected with the advancements in computational methods [9, 10].

The first kerogen model was published by Burlingame et al. (1969) by studying the kerogen extracted from the Green River Shale [13]. However, the suggested structure did not represent a comprehensive chemical structure of the sample since it did not contain molecular topology. Next, Siskin et al. (1995) input additional chemical structures and proposed a new updated model for kerogen, particularly by adding the functional groups with oxygen and nitrogen [14]. Both of these studies developed their models and provided a general chemical formula for the kerogen.

With recent advancements in computational 3D modeling, investigating kerogen’s molecular structure has increased significantly and has been expanded to study various phenomenon involving the organic matter. This becomes more interesting when the computational simulation results agree with the experimental outcome [15, 16, 19]. For example, recent studies regarding computational methods have expanded beyond building the organic matter molecular structure into simulating the gas adsorption and desorption processes [13].

When the goal is building a reliable molecular computational model for kerogen, two separate approaches can be taken. The first approach is the realistic 3D kerogen model which was introduced by Ungerer et al. (2015) [12]. They analyzed different kerogen types (based on their biogenic origin) by using molecular dynamics and quantum mechanics based on a set of experimental data. This 3D model was created based on analyzing the experimental data that provided the authors with both functional groups and various chemical structures of the samples.
The second approach is to build the amorphous model through quenching molecular dynamics. In this method, the disordered carbons transformed to the amorphous solid structure by using the reactive force field [10, 20]. Such models have been referred to and reused numerously in different studies such as the investigation of gas adsorption behavior and kerogen swelling [16, 17]. The previous attempts of building the molecular structure of kerogen, are just the beginning. Their development, excluded necessary elements that are found in the structure, meaning that their accuracy and use is limited. Moreover, the proposed models necessitate further study to meet the expectation that the kerogen models are compatible with naturally occurring kerogen to understand the fundamental phenomenon that involves the organic matter, i.e. gas adsorption/desorption and diffusion [9].

By exposing organic matter to various types of gas such as methane, nitrogen and carbon dioxide at specific pressures and temperatures, adsorption isotherms can be obtained. The isotherms are used to provide information about organic matter’s pore network. Knowing the adsorption capacity of the organic matter becomes important regardless of pore network, when organic rich shale plays are considered for sequestration. In addition, recent studies are suggesting injection of carbon dioxide into organic rich shales can increase their production potential [39], while the amount of injected CO$_2$ should be studied since the gas can adsorb into organic matter as well. Hence, in order to precisely estimate the capacity of organic matter in terms of adsorption for sequestration and/or associated mechanisms for enhanced oil recovery, the structure of organic matter should be known.

The Bakken Formation is one of the largest unconventional shale oil plays in North America and is currently being studied for potential CO$_2$ enhanced oil recovery and sequestration [39]. Based on what is said above, the objective of this study is 1) to generate a new representative molecular model of kerogen at the thermally immature (pre-oil window) stages for the Bakken, with a primary focus on the chemical composition and structure and 2) predict adsorption capacity of the models in terms of CO$_2$ quantity. In order to achieve the goals, three different organic matter models are developed via computational/simulation methods based on experimental chemical compositional data that is collected from an immature Bakken sample from literature. Then, CO$_2$ adsorption isotherms are predicted from the generated models and compared with experimental gas adsorption data that is collected on the physical sample in the lab. While the validity of the model is examined with a comparison between gas adsorption data predicted from the models and experimental ones, we were able to use isotherms as a criterion or constraint to tune the models as well.

2. Methods

This study proposes a new conceptual model for kerogen from the Bakken Shale. Moreover, the models are verified via the gas adsorption isotherm experiments and molecular simulation. The general chemical composition and existing functional group data are input into the simulation software from the literature, while additional gas adsorption experiments were carried out on
the Bakken kerogen samples with similar characteristics. The construction of macromolecule kerogen models consists of four major steps.

a) The molecular structure of the physical sample B1 was analyzed through experimental data, then the details of chemical composition and structures including functional groups were decided.
b) The details from experimental data were utilized to prepare the models regarding the aromatic and functional group molecules using drawing software.
c) A number of atoms and molecules were positioned in a rectangular simulation box with different initial configurations. The simulation ran at high temperatures during the construction process.
d) In order to get a partial charge of elements, the macromolecular models were optimized by density functional theory (DFT) calculations.

The comparison of gas adsorption behavior between physical kerogen samples and the kerogen models is performed with the following three major steps:

a) The experiment of nitrogen and carbon dioxide gas adsorption was conducted on isolated kerogen samples B2 and B3.
b) The molecular models that were built through the above construction stages were initially placed at both ends of the simulation box, and the number of gas molecules was distributed in terms of the experimental data.
c) Gas adsorption simulations run at the same temperature and pressure of the experiment, and then a comparison of the results are made.

2.1. Analysis and model preparation

A variety of methods can be utilized to provide us with the chemical composition of the organic matter. $^{13}$C-NMR can examine the chemical structures and parameters related to carbon in any organic material. The sulfur and nitrogen content are revealed through the X-ray absorption near-edge structure (XANES) and finally XPS is capable of quantifying several functionalities in carbonaceous materials associated with carbon, oxygen, sulfur, and nitrogen [31]. Kelemen (2007) reported chemical composition and structural information in a comprehensive study on a diverse origin of kerogen including the Bakken which were the input parameters for this study and were acquired by using $^{13}$C-NMR, XPS and XSANES [21].

In this study, it was decided to mainly refer to $^{13}$C-NMR data which provides us with more detailed chemical information to construct the bulk kerogen model. In addition, XPS and XSANES data were used for characterization of organic nitrogen and sulfur species. It is stated by Kelemen (2007) [21], that the Bakken sample belongs to the immature (pre-oil window) type II kerogen with $T_{\text{max}}$ (419 °C) and Hydrogen index (HI) (580 mg/g), and the sample (B1 was used in this study to construct the 3D computational model.
The ratios between carbon, hydrogen, sulfur, and nitrogen atoms were decided by considering the $^{13}$C–NMR, XANES, and XPS analysis results, and they were added to the molecular simulation stage. We determined that around 35% of total carbon comprises aromatic structures which contain nitrogen and sulfur such as pyridine, pyrrole, and thiophene. The functional groups related to sulfur were set as sulfate and sulfoxide structures, moreover the carbonyl and ether which are oxygen related functional groups were assigned.

In our study, Samples B2 and B3 are tested for gas adsorption for further model verifications while their geochemical characteristics are reported in table 1, obtained from Rock-Eval 6 pyrolysis. It can be seen that these two samples have the same Tmax of 429 (°C), and HI index of 555 and 513 (mg/g) correspondingly. Based on table 1 and figure 1, all samples belong to type II kerogen and in the pre-oil window (immature) [38].

2.2. Molecular Dynamics Simulation (MDS)

The molecular dynamics simulation that was used to build the kerogen model utilized a closed system (no mass and energy transfer). Several types of potential force fields have been developed to realistically represent the overall forces that are exerted on atoms in a molecule in a wide range of substances with different applications in molecular dynamics research. Assisted Model Building with Energy Refinement (AMBER) force field, which is used extensively for simulating biomolecules, is one of the force fields that utilize empirical energy approaches to calculate interactions between molecules in small molecules [29]. General AMBER Force Field (GAFF) [27,28] is an extension of AMBER and is suitable for all organics that contain C, N, O, H and S in their structure. In this study, the all-atom GAFF potential model was used for molecular dynamics simulation so that single atoms of the system have their GAFF parameters interacting with other neighboring atoms. All simulations were conducted through LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) [36] with force field parameters such as the bond, angle, dihedral, and non-bonded energy terms calculated based on the GAFF parameter library as follows:

\[
E_{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \\
\sum_{i<j} 4\varepsilon \left[ \frac{\sigma_{ij}}{r_{ij}}^{12} - \frac{\sigma_{ij}}{r_{ij}}^6 \right] + \sum_{i<j} \frac{q_i q_j}{\varepsilon r_{ij}} , r_{ij} < r_c
\]

The GAFF force field ($E_{total}$) that signifies the potential model is represented by Eq. (1) where $r_{ij}$ is the distance between sites i and j, $r_c$ is the cutoff radius beyond which the short-range interactions are neglected, $\sigma_{ij}$ and $\varepsilon$ are the L-J parameters, $q_i$ and $q_j$ represent the charges on sites i and j, and C is a unit conversion constant. The potentials are the added overall terms to calculate the total intermolecular interactions. The first three terms express the intramolecular energy representing forces associated with bonds, bond angles, and torsions respectively while the last two terms are the functions of Lennard-Jones potential and pairwise electrostatic
interactions between the atoms. The simulation employed the Lennard-Jones (LJ) potential model for short-range site-site interactions, and a particle-particle particle-mesh solver (PPPM) [33, 37] to compute the long-range Columbic interaction. In addition to the determination of the total number of atoms and the ratios between them at the preparation stage as explained earlier, molecules and atoms are set by molecular drawing software, Avogadro [24] and Packmol [35]. All of the ingredient molecules were assigned by Gasteiger-Marsili sigma charge [25] at the initial stage of the molecular mechanics.

Also, a Nose-Hoover thermostat was applied to the number of atoms (N), simulation volume (V) and temperature (T) canonical ensemble so that the system could maintain a constant temperature throughout the simulation. For constructing the kerogen model process, all simulations were carried out at 400K with the time step of 1 fs for runs up to 30 ns in length (30 million steps). The molecular dynamics simulation allowed the ingredient molecules to interact with neighboring atoms and to reposition themselves in the system. Then, the structural model for disordered molecules was generated by the given force field parameters while the simulations run.

2.3. Quantum mechanics

In the potential energy of the system that is given by Eq (1), the pairwise electrostatic interactions between neighboring atoms with a partial charge \( q_i \) and \( q_j \) are quantified by Coulomb’s law. However, in previous MDS stage, the electrons are not considered in molecular mechanics force fields including the general AMBER force field. Due to the fact that the electrostatic interactions are inherited into empirical parameters which are effective for the ground state electron configuration of given covalent bonds, therefore, it is impossible for the molecular mechanics to simulate electronic rearrangements. On the other hand, quantum mechanics can acquire a description of the electronic degrees of freedom and compute the electronic wave function.

To obtain the quantitative electrostatic properties and optimized geometries for our kerogen model, after molecular dynamics simulation was performed, quantum mechanical (QM) calculations using the ORCA software package [40] was employed to clarify the binding mechanisms between atoms and within molecules. The partial charge of each atom is calculated based on the density functional theory (DFT). DFT-B3LYP/6-31G(d) Method/Basis set combination, which is considered to be suitable for the organic compound, was used for geometric optimization of the Bakken kerogen models [45]. Hirshfeld atomic population [23] was acquired to obtain atomic partial charges since it is less basis-set dependent and can be derived for optimal partitioning of electron density. The partial charge obtained from the quantum mechanics calculation was substituted for the initial partial charge which had set without the polarization of atoms.

2.4. Gas adsorption simulation detail
This study carried out molecular dynamics of carbon dioxide and nitrogen gas adsorption processes by using osmosis and reverse osmosis technique with LAMMPS simulation package [3, 7, 34, 36]. The molecules in the constructed models were initially tethered to their optimized position and synchronized with experimental temperature and gas density (see Figure 2). The amount of CO$_2$ and N$_2$ molecules were decided by the densities of CO$_2$ and N$_2$ at desired temperature and pressure from experiment data [18, 41]. All partial charges of the kerogen models were employed through the geometric optimization and electron density calculations. For this simulation, the harmonic potential was employed between the model and gas molecules, and the site to site non-bonding interaction potential that was used is shown with the following formula:

$$E_{ij} = 4\varepsilon \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon R_{ij}}$$  \hspace{1cm} (2)

Similar parameters as Eq. (1) were used, and these interactions then influenced all of the active sites to get the intermolecular relation. In our simulation, the parameters for this interaction potential in gas adsorption simulation are shown in Table 2. The simulation box with a cubic length of 120 Å, is subjected to periodic boundary conditions. This boundary condition was set in order to maintain a constant temperature of 77 K and 273 K which is the true experimental gas adsorption temperatures and applied with the Nose-Hoover thermostats. This molecular dynamics simulation ran for up to one million timesteps to make sure the density of adsorbate molecules has reached equilibration.

2.5. Gas adsorption on the Bakken kerogen

Gas adsorption experiments are done on isolated kerogen from the bulk shale with the procedures found in Liu et al. (2017) [26]. Briefly, we collected the samples and then removed the bitumen by using the mixture of the methanol and toluene. Then we added the HCl into the solid residue to remove the carbonates. After that, HF was added to remove the silicate minerals and finally pyrite was removed by using CrCl$_2$.

After solid kerogen is isolated from the rock matrix, the kerogen was degassed for at least 8 hours at 110°C to remove moisture and volatiles. Then the kerogen sample was crushed (less than 250 um) and it was loaded into the instruments. Low-pressure nitrogen was measured on a Micromeritics® Tristar II apparatus at 77K while carbon dioxide adsorption was performed on a Micromeritics® Tristar II plus apparatus at 273K. Then the gas adsorption quantity was measured over the relative equilibrium adsorption pressure (P/P0) range of 0.01-0.99, where P is the gas vapor pressure in the system and P0 is the saturation pressure of nitrogen.

3. Results: model verification

3.1. Bakken models
The Bakken Shale models were constructed by analyzing experimental data, molecular mechanics simulation, and quantum mechanics calculations. The whole structures of the models are a complicated mixture of chain and mesh. Figure 3 shows the three macro-molecules which do not have the same chemical composition, geometry, and density. The final chemical compositions of Model (A), (B), and (C) are C\(_{140}\)H\(_{176}\)N\(_{6}\)O\(_{15}\)S\(_{4}\), C\(_{151}\)H\(_{179}\)N\(_{6}\)O\(_{15}\)S\(_{4}\), and C\(_{158}\)H\(_{188}\)N\(_{6}\)O\(_{16}\)S\(_{4}\) respectively.

In order to verify which one of these models are more accurate, the simulation results of these models were compared with the experimental data. Amorphous kerogen models (Model (A), (B), and (C)) are verified based on the experimental and theoretical prediction of gas adsorption data. This means we predicted adsorption isotherms from each generated model and compared that with the isotherm curves that is obtained from the gas adsorption experiments on the physical sample. The estimated density of Model (A), (B), and (C) was 1.1, 1.2, and 1.3 (g/cm\(^3\)) respectively.

Even though the models are originated from the experimental sample B1 from the data that is collected by Kelemen (2007), three models did not have the same chemical structural and chemical composition. The initial position and different simulation box size of each model could vary the simulation system and final results. Since the kerogen is an amorphous organic material, its molecular structure cannot be fixed despite having a similar chemical composition. This work aimed to create the models which have not only diverse structure but also similar chemical composition ratios.

Table 3 summarizes the aromatic carbons in the constructed models that were found compatible with sample B1. Since aromatic carbons were first set up at the initial stages of running the simulation, carbons in aromatic structures are very close to sample B1 in regards to type, bonding, geometries, etc. than other atoms. However, several discrepancies were detected with regards to the methyl group and ether group despite the total number of carbons that was initially fixed. We believe these inconsistencies are because some methyl group was eliminated or added during the optimization process. When the models are comprised of ingredient molecules and atoms, some carbons were subtracted or added so that the number of methyl carbons varies in each model.

The details of all parameters which attributes to chemical composition and geometry are described in Table 3. During the optimization process, aliphatic carbon and hydrogen ingredients experienced addition and subtraction. The free radicals on aliphatic carbons were detected through a local optimization procedure. Initially a limited number of ingredient molecules and atoms allowed the carbons to form free radicals in the system. Figure 4 demonstrates that the defects in the structure were introduced while ingredients initially were wandering in the non-periodic system. In a recent study by Larsen [11], free radicals were measured to be around 0.48 (spins/g x 10E-18) on the physical Bakken sample. As the organic matter goes through maturation, free radicals naturally form and can exist. However, this amount is negligible in our system size so that the adjustment was conducted through addition and subtraction.

The PDF (pair distribution function) profile shows the probability of carbon existence at the distance r(Å) from the carbon, and it is highly related to carbon structure and density. The left
peak position is around 1.5Å which represents aromatic carbons (see Figure 5 a). Since almost
35% of the carbons in the models have an aromatic structure, the first peak from the left is much
larger than others. The three models have similar peak positions with a similar width of PDF, but
the differences of G(r) amplitudes of the models reveals that Model (C) has the largest density
value than the other two. Figure 5 (b) shows the comparison of the three models with sample B1
based on the total oxygen, carbonyl and ether or alcohol groups per 100 carbons. From the image,
the constructed models (A,B,C) contain a higher number of total oxygen per carbon and the
sample B1 has a higher number of ether and alcohol groups compared to the models. The
carbonyl functional on the other hand, has a similar ratio comparing the models with the physical
sample data (B1).

Figure [5] (c) and (d) also indicates that all three models have a major significant percentage of
nitrogen and sulfur than the original input data from 13C-NMR of the physical sample B1. The
constructed models consisted of around 150 carbon atoms in their structure so that existence of
a small number of functional groups would make a big discrepancy in their ratios among the
models. Despite the fact that there are deviations, Figure 5 (b, c, and d) reports that every
functional group measured by Kelemen (2007) are comprised of the models.

3.2. Gas diffusion/adsorption on the surface of kerogen

In this study, we performed the gas adsorption molecular simulation on kerogen to determine
both the adsorption of the gas as well as a tool for verification of our suggested molecular models
(Model (A), (B), and (C)). Pore size is an important parameter to be used to estimate gas
adsorption because the fluid adsorption process occurs into shale kerogen pores [26]. Studies
show that naturally matured kerogen has mainly nano to micro size pore channels so that it is
heavily dependent on Darcy’s law for flow mechanisms [26]. However, this law is not compatible
with the particle transport phenomena under the small pore system [42]. As a result of the
kerogen model building process, pores were randomly generated and the size of them ranges
30^{-1}nm to 70^{-1}nm in the structure of our models. Therefore, the simulation system is mainly
employed to study the intermolecular interactions between particles while gas adsorption occurs,
but not necessarily the adsorption of gas in the pores.

Figure [6] explains the nitrogen gas adsorption results including the experimental data on two
physical samples and the three theoretical models. Samples, B2 and B3, captures around 40.78
and 43.18 (cm³/g STP) of the N₂ gas respectively at 100 kPa, 77K. The nitrogen molecules
adsorption behavior with the model (A) shows that the amount of adsorbed molecules (46.02
cm³/g STP) are close to the two experimental samples. First, the molecules diffused to the model
surface from their initial position in the bulk region. Then, the adsorption/desorption occurred
through the interactions between molecules and the model in the surface region.

Model (B) adsorbed the largest amount of the gas molecules (57.29 cm³/g STP), and the amount
of adsorbed molecules into Model (C) (53.59 cm³/g STP) is also higher than the experimental
samples B2 and B3. After the N₂ molecules diffused and reached the model surface, the mass
transfer to the bulk region is observed. The interaction in the surface region is not strong enough to capture a large number of nitrogen molecules.

Unlike nitrogen gas adsorption experimental conditions, carbon dioxide adsorption simulation performed under different pressure conditions at 273K as shown in Figure 6 [41]. The result of CO₂ gas adsorption isotherms of the samples shows a nearly linear relationship of absorption of gas with respect to the pressure. Also, the models are generally similar in adsorption results with the samples B2 and B3. At lower pressures, the quantity of CO₂ absorbed in all three models is very close to the quantity that is absorbed on B2 and B3 samples. Considering the initial comparison pressure (10 kPa), the model (A), (B), and (C) shows that the total amount of adsorbed molecules are 1.45, 1.35, and 1.30 cm³/g STP, respectively. These are a similar amount with the samples B2 being 1.59 cm³/g STP, and sample B3 measured 1.45 cm³/g STP.

The CO₂ adsorption simulations that have taken place over 50 kPa shows that the main disagreement in the results initiates between the samples and the models at this specific pressure. At 50 kPa, sample B2 with 4.52 cm³/g STP, and sample B3 with 4.35 cm³/g STP amount of adsorbed CO₂ have higher adsorption quantities than the models (model (A): 3.62, (B): 2.70, (C): 3.24 cm³/g STP respectively). As the pressure increases, the deviation between the models and the samples also becomes more significant. Finally, when the simulation and experimental pressures reach 100 kPa, the largest deviation in the amount of adsorbed CO₂ between the models and physical samples is found (sample B2, B3: 6.96, 7.03 cm³/g STP respectively, and the model (A), (B), (C): 10.87, 10.81, 9.01 cm³/g STP respectively). Collectively, it is observed from the results as relative pressure get closer to 1 a notable discrepancy in the quantity of adsorbed CO₂ between the models and physical samples are recorded.

4. Discussion

This study attempted to develop the molecular model of amorphous kerogen through computational methods by inputting experimental data and using gas adsorption to verify the model. First, atoms and molecules were made as the ingredient and they were agglomerated within the designated simulation box which varied based on the initial position of elements. The experimental parameters initially were applied to the system for potential energy minimization with the local optimization process. However, molecular dynamics cannot make a structurally perfect configuration so that defects occur during the building process. Spatial constraint influences the structure and function of atoms network because the simulation space is non-periodic. Thus, it was decided to build three different Bakken kerogen models with different densities. The pores were irregularly spread within the bulk structure. It is important to note that unlike the natural kerogen generation and maturation which is considered a top-down process, the models were generated in a bottom-up process. As the maturity of the organic matter in the shale sample increases, the hydrocarbons detach, and chemical breaking/bonding occur to become stabilized under temperature and pressure. The natural process of maturation which happens under higher pressure and temperatures leads to the generation of hydrocarbons while
the N, S, O ingredients detach from the source of structure. This makes that the kerogen structure becomes more aromatic, better ordered and smaller [46]. Whereas, our models were initiated with small molecules and clusters containing functional groups. These ingredients randomly were relocated through the employed force field, temperature, and pressure. The models randomly developed during the building process so that each model would generate different surface heterogeneity which consists of the geometry of the pore and the charge distribution. Although the computational models and the natural kerogen have different generation path as discussed above, the models not only have amorphous structure like the naturally occurring kerogen, but contains corresponding chemical structures as well.

Our theoretical models consist of small pores (1<nm) which makes them challenging to investigate bulk flow molecular diffusion processes in nano-pores, but suitable to estimate intermolecular behaviors on the surface instead. This simulation focused mainly on the interaction between gas molecules and the model surface and the results of N₂ gas adsorption isotherm prediction at 100kPa and 77K supports the experimental data. (See Figure 6). The model (A) has a similar N₂ molecular adsorption pattern with the samples B2 and B3 while the computational model (B) and (C) adsorbed slightly a larger number of N₂ molecules than an expected amount based on experimental data. We deduce that the difference in geometry of the pore and the charge distribution of each model affect the configuration of N₂ molecules on the kerogen surface and would cause the discrepancy in the amount of N₂ adsorbed by each model surface.

Considering the CO₂ gas adsorption simulation and experimental data, the result shows the CO₂ gas adsorption behavior of the models is close to the experimental data and physical samples at lower relative pressures. It was observed when pressure increases, the models capture more quantity of CO₂ molecules on their surface. We believe this discrepancy is mainly due to the fact that the models have a higher potential surfaces than the samples. It should be noted that since the potentials are overlapped on the surface with the very small pore, the CO₂ and wall interactions, will cause more amount of CO₂ molecules to be adsorbed on the surface with small pores (<1nm) presence compared to the micro (larger) pores [32]. As mentioned above, since the models have smaller pores (30⁻¹nm to 70⁻¹nm ) and a greater number of CO₂ molecules are placed in a fixed system at higher pressures, there has been a higher chance for the CO₂ adsorption on the surface due to gas and kerogen surface interactions. Mass transfer of CO₂ molecules occurred at the beginning of the simulations, then there was not major mass transportation after diffusion of CO₂ molecules into the surface region. Moreover, although the pore size distribution in the organic matter is an important factor for correct estimation of adsorption quantities, non-porous Bakken need to be considered. The study about pore network quantification via a combination of small angle neutron scattering (SANS) and nitrogen adsorption from the Bakken revealed that the Bakken might not have developed any porosity.
Recent study performed on the isolated organic matter of Bakken kerogen in the pre-oil generation window or thermally immature stages which has not generated any hydrocarbons yet.

Surprisingly, the simulation of gas adsorption of CO$_2$ shows that carbon dioxide molecules have migrated to the kerogen model during the process, which can be found inside the model and near the surface as shown in Figure [8]. After the initial mass transfer of CO$_2$ molecules in the bulk region and the adsorption that is occurred on the surface through gas-surface molecular interaction, it seems this the interaction is responsible and strong enough to capture CO$_2$ molecules so that the mass transfer into the bulk region does not take place. This CO$_2$-philic phenomenon occurs in every CO$_2$ adsorption simulation regardless of the kerogen type and pressure. While N$_2$ molecules are randomly diffused in the bulk region except for the molecules that are adsorbed on the surface of the kerogen. It is proved that adsorption of CO$_2$ on the kerogen has a higher magnitude than adsorption of other particles such as CH$_4$ through the adsorption isotherm experiments [43,44]. In the current stage of our simulation of CO$_2$ adsorption/desorption, the combination of the simulation result and the density profile reveals that kerogen models prefer to adsorb CO$_2$ more than N$_2$ while a mixture of these two molecules was not added to the system to compare selectivity. In the next study, the selectivity/adsorption of CH$_4$, H$_2$O, and CO$_2$ mixtures need to be investigated at a variety of conditions such as pore size distributions, temperature, and pressure.

5. Conclusion

This work proposed a new approach to construct amorphous material by considering experimental constraints. Bakken Shale kerogen was selected to be developed with the experimental chemistry data available in the literature. The numerical analysis based on kerogen analytical input was used to determine chemical structures. Computational techniques including integration of molecular dynamics and quantum mechanics facilitated a realistic structure of Bakken kerogen based on the input parameters. We used gas adsorption isotherms obtained from physical kerogen as a constraint to update and verify our model. Based on the results the following conclusions can be made:

- Both physical Bakken kerogen sample and the models from computational methods showed to have similar geochemical composition and structure.
- Adsorption of CO$_2$ molecules was less in lower pressures and the models capture more CO$_2$ molecules on their surface when pressure is higher. This was related to the additional available surface in the models compared to the physical sample. Based on this observation it was concluded that the pore structure plays an important role in adsorption mechanisms as a function of pressure.
- N$_2$ and CO$_2$ gas adsorption behavior of three models were predicted and compared with the isotherm experimental results on physical Bakken kerogen sample. Although there
was some discrepancy between the adsorption behavior of experiments and simulations, there was a good overall agreement in the isotherm and adsorption patterns.

- Based on the simulation results we found, all three created models tend to attract CO\textsubscript{2} molecules more than N\textsubscript{2} molecules which were the same case as gas adsorption experimental results. We concluded the diffusion of N\textsubscript{2} molecules was not strongly affected by the model variations.

- The comparison of Bakken sample and the models corroborated that our building process is reliable to represent amorphous organic material. Furthermore, this process will assist to construct more realistic amorphous material models and to investigate the thermodynamic properties of kerogen.

**References**


Table 1 Properties of the shale samples that were modeled in this study.

<table>
<thead>
<tr>
<th>Property</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{max}$°C</td>
<td>419</td>
<td>429</td>
<td>429</td>
</tr>
<tr>
<td>$HI$(mg/g)</td>
<td>580</td>
<td>555</td>
<td>513</td>
</tr>
<tr>
<td>$VR_o$(%)</td>
<td>0.5</td>
<td>0.54</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Note: The Bakken sample (B1) has detailed chemical composition and structural information [21]. B2 and B3 samples are used for N$_2$, CO$_2$ gas adsorption experiments [26].

Table 2. Parameters related to the adsorbates (CO$_2$ and N$_2$).

<table>
<thead>
<tr>
<th>Molecules</th>
<th>Atoms</th>
<th>Charge</th>
<th>$\sigma$(Å)</th>
<th>$\epsilon/k_B$(K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_2$</td>
<td>N</td>
<td>0.00</td>
<td>3.75</td>
<td>95.20</td>
</tr>
<tr>
<td>$CO_2$</td>
<td>C</td>
<td>0.70</td>
<td>2.80</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>-0.35</td>
<td>3.05</td>
<td>79</td>
</tr>
</tbody>
</table>

Table 3. Structural parameters relevant to carbons in the Bakken kerogen (Sample B1 by Kelemen (2007)) and the constructed models, Model (A), Model (B) and Model (C).

<table>
<thead>
<tr>
<th>Structure</th>
<th>Sample B1</th>
<th>Model (A)</th>
<th>Model (B)</th>
<th>Model (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aromatic</td>
<td>0.35</td>
<td>0.371</td>
<td>0.344</td>
<td>0.33</td>
</tr>
<tr>
<td>Carboxyl/Amide/Carbonyl</td>
<td>0.02</td>
<td>0.028</td>
<td>0.026</td>
<td>0.025</td>
</tr>
<tr>
<td>Protonated aromatic</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Phenoxy/Phenolic</td>
<td>0.02</td>
<td>0.021</td>
<td>0.021</td>
<td>0.021</td>
</tr>
<tr>
<td>Alkyl-substituted aromatic</td>
<td>0.08</td>
<td>0.064</td>
<td>0.064</td>
<td>0.07</td>
</tr>
<tr>
<td>Bridged aromatic</td>
<td>0.09</td>
<td>0.092</td>
<td>0.092</td>
<td>0.092</td>
</tr>
<tr>
<td>Aliphatic</td>
<td>0.63</td>
<td>0.6</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>Methylene/Methine</td>
<td>0.46</td>
<td>0.49</td>
<td>0.35</td>
<td>0.12</td>
</tr>
<tr>
<td>Methyl/Methoxy</td>
<td>0.15</td>
<td>0.107</td>
<td>0.12</td>
<td>0.127</td>
</tr>
<tr>
<td>Alcohol/ether</td>
<td>0.06</td>
<td>0.043</td>
<td>0.043</td>
<td>0.038</td>
</tr>
</tbody>
</table>

Note: the data presented here are ratio per 1 number of carbon.
Figure 1. HI vs. \( T_{\text{max}} \) which shows that each sample belongs type II thermally immature kerogen at the pre-oil window [21,26].
Figure 2. (a) Initial gas adsorption simulation set up with the kerogen models and gas molecules. (b) The system becomes equilibrium status and the gas molecules are diffused.
Figure 3. (a) Constructed Bakken kerogen model (A), (b) Constructed Bakken kerogen model (B) and (c) Constructed Bakken kerogen model (C). Different geometric configuration and chemical compositions with the color code: Carbon: Black; Hydrogen: white; Oxygen: red; Sulfur: yellow and Nitrogen: blue.

Figure 4. Schematic view of defects obtained from the local optimization procedure. The red circle points out carbon atoms presented in black referring to defection and remaining free radical.
Figure 5.  a) Pair distribution functions (PDF) or G(r) of the Bakken kerogen model with different densities. Gray, yellow and blue curves represent, models A, B and C respectively, b) the comparison between the ratio of oxygen-related functional groups to the total number of carbons, c) the comparison between the ratios of nitrogen related functional groups to the total number of carbons in the models and physical sample and d) the comparison of the ratios of Sulfur related functional groups to the total number of carbons.
Figure 6. The comparison of the simulated excess nitrogen (N$_2$) adsorption isotherms between the models (Model (A), Model (B), and Model (C)) with experimental loadings (B2 and B3) at 100 kPa, 77K.

Figure 7. Comparison of simulated excess CO$_2$ adsorption isotherms between the models (Model (A), Model (B), and Model (C)) with experimental loadings (B2 and B3) at 273K. Yellow and orange curves represent experimental data from physical samples while blue, gray and yellow dots are the results from simulations and computational models.
Figure 8. Mass density profile of CO$_2$ from three Bakken kerogen models at 100 kPa and 273K. As a result of the molecular dynamics, this density profile shows that CO$_2$ molecules are crowded near Bakken models at one million timesteps.