Grant agreement No. 640979

ShaleXenvironmenT

Maximizing the EU shale gas potential by minimizing its environmental footprint

H2020-LCE-2014-1
Competitive low-carbon energy

D4.1
New models for clay and kerogen pores in shale rocks

WP 4 – Modeling of confined fluids

Due date of deliverable	Month 18 - February 2017
Actual submission date	03/04/2017
Start date of project	September 1st 2015
Duration	36 months
Lead beneficiary	ARMINES
Last editor	Andrey Kalinichev
Contributors	ARMINES, NCSR “D”, ICPF
Dissemination level	Public (PU)

This Project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 640979.

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

History of the changes

Version	Date	Released by	Comments
1.0	17-02-17	M. Vasileiadis	First draft of the report
1.1	20-02-17	K. D. Papavasileiou, L.D. Peristeras	Updates
1.2	27-02-17	A.G.Kalinichev	Updated and combined with reports of other participants (ARMINES, ICPF)
1.3	31-03-17	A.G.Kalinichev	Updated and combined with reports of other participants (ARMINES, ICPF)

Table of contents

List of figures

List of tables

Key word list

Kerogen, Models, Micropores, Porosity characterization, Voronoi tessellation, clay minerals, clay particle edges.

Definitions and acronyms

Acronyms	Definitions
NORM	Naturally occurring radioactive material
GAFF	Generalized Amber Force Field
PSD	Pore size distribution
LPD	Limiting pore distribution
MC	Monte Carlo
MD	Molecular Dynamics
RMC	Reverse Monte Carlo
QMD	Quenched Molecular Dynamics
HRMC	Hybrid Reverse Monte Carlo
MD-HRMC	Molecular Dynamics Hybrid Reverse Monte Carlo

1. Introduction

1.1. General context

Although the existence of shale gas has been known for decades, the technology for the economically viable extraction of shale gas was made available only recently [1]. The experience that obtained during the years of shale gas exploitation, together with the information that has been accumulated, have revealed a number of issues associated with shale gas extraction. The environmental impact of key technologies necessary for shale gas extraction (such as hydraulic fracturing) is poorly understood, with evidence that they may be related to important issues such as pollution of potable water [2] and seismic phenomena [3]. Furthermore, it is currently established that each shale field is unique, therefore requiring adaptation of the available methods for optimal yield. These facts pose a pressing demand for scientific research that would help understand the underlying physical phenomena. In this way it will be possible to progressively tackle the various open issues that will ultimately enable an undoubtedly green exploitation of proven available energy sources in a sustainable and beneficiary for the society way.

The study of the phenomena taking place at the atomistic level in a shale rock is of immense importance. It is related to the transportation of the shale gas, the fracturing fluid, and the naturally occurring radioactive materials (NORM) [4]. The understating of the factors controlling the migration of these substances would have profound implications in the design of processes and methods for gas extraction, and elimination of any potential environmental cost in an economically viable and sustainable manner. Computational modeling methods are expected to play an important role towards the understanding of the complex physical and chemical phenomena, tacking place in the highly confined environment of shale rocks. For these reasons, it is important to have realistic models of the primary materials constituting shale gas reservoirs. It is currently assumed that shales of interest are composed primarily of clays (kaolinite, muscovite, smectites) and other minerals (quartz, calcite) with dispersed nodules of organic matter. The extractable shale gas is predominantly located in the nodules or organic matter. The insoluble (in common organic solvents) part of the organic matter found in a shale rock is usually defined as kerogen [5].

In shale formations, shale gas (typically a mixture of methane, ethane and propane with a very small amount of butane and heavier hydrocarbons as well as carbon dioxide and nitrogen) is stored as free gas, adsorbed gas, and dissolved gas. The free gas gathers in the fractures and pores of a shale rock, the adsorbed gas occurs on the surfaces of both the organic material and clay minerals, and the dissolved gas enriches the organic matter. The adsorbed and dissolved gas is in equilibrium with a homogeneous free gas phase in an interconnected shale pore structure. If some shale pores are not interconnected, there may be a departure from the equilibrium between the adsorbed and dissolved gas, and the free

gas. The amount of free gas is relatively easy to estimate based on the temperature and formation pressure, its porosity, and the fraction of porosity which is filled by the gas. The contribution of adsorbed gas to the total gas in place (GIP), estimated to reach up to 60%, is still poorly understood, mainly due to complexity of the gas adsorption on shale. Generally, most of the adsorption area is located in the organic material and the contribution of the adsorbed gas to the total GIP is less significant in the inorganic matrix [6].

Kerogen is a complex organic material with an amorphous porous carbon skeleton and exhibits significant pore-shape and pore-connectivity variations. Depending on its geographic origin, maturity and sedimentary history, kerogen displays a broad range of density and, chemical composition in terms of atomic contents and chemical functions, porosity and tortuosity. The van Krevelen diagram [7,8] (Figure 1), a plot of hydrogen-to-carbon (H/C) and oxygen-to-carbon (O/C) atomic ratios, provides a tool to distinguish different types of kerogen in terms of depositional origin: type I (lacustrine), type II (marine), type III (terrestrial), and type IV (originating from residues) and maturity. Increasing the maturity of kerogens due to exposure to high temperature and pressure over geological time scales leads to a shift of the kerogen H/C and O/C atomic ratios in the van Krevelen diagram from the top-right to bottom-left corners.

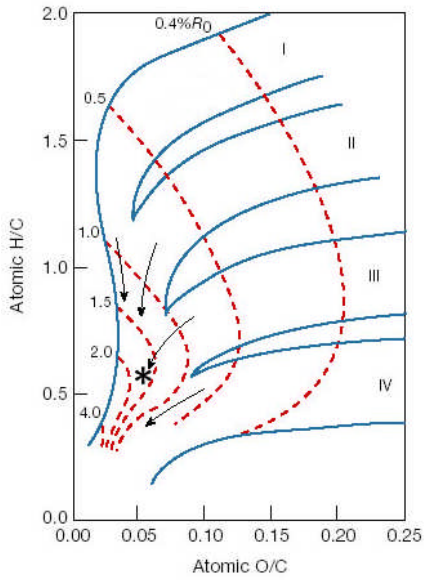


Figure 1: The van Krevelen diagram, a plot of hydrogen-to-carbon (H/C) and oxygen-to-carbon (O/C) atomic ratios.

Here we are interested in generating realistic and representative molecular models of kerogen and shale rock host (clay and other minerals) for later use in the computational molecular modelling of their interactions with fluid mixtures relevant to the shale gas technology, such as the shale gas itself and/or the fracturing fluid. This is a complex task requiring addressing a number of challenges:

1) Kerogen is an amorphous solid composed of a mixture of different not very well defined macromolecules of varying size.

2) Not only kerogen in every shale field is different, but also the characteristics of kerogen in different depths of the same well vary significantly.

3) There are limited tools available for detailed characterization of the porosity of amorphous materials and their performance for the analysis.

4) Large size structures needed to study amorphous material is problematic.

5) Clay mineral structures are very diverse as they incorporate a significant degree of compositional and structural disorder.

6) In most previous molecular simulations, the structure of clay particles are represented by semi-infinite layers, i.e., they do not have any edges. This simplification is acceptable to a certain extent, but real clay particles always have a finite size and should be terminated by lateral surfaces or edge surfaces.

7) Edge surfaces of clay particles can exhibit adsorption sites highly different from the ones on the basal surfaces of clay layers, which is extremely important for the interaction of clay particles in shale rocks with kerogen and with naturally occurring radioactive materials (NORM).

1.2. Deliverable objectives

1.2.1. Kerogen models

In recent years, there have been increasing efforts to use atomistic simulations and provide molecular-level insight into shale gas behaviour in different types of kerogen. The structure of real kerogens is unknown and experimental techniques such as XPS, ^{13}C NMR and S-XANES are utilised to characterise kerogen samples in terms of their elemental composition (H, C, O, S, N), functional groups containing carbon and hydrogen, oxygen, sulphur, and nitrogen, percentage of aromatic carbon with attachments, average number of aromatic carbons per polyaromatic cluster, and average aliphatic carbon chain length [9].

Our first objective is to generate new realistic and representative molecular models of kerogen, with primary focus on the porosity based on a systematic methodology. The unoccupied space in the structure is where most shale gas is hosted and the route of its transportation. Given the diverse experimentally determined characteristics of porosity, which indicate the unique characteristics of each shale field, the method that is sought here should preferably allow easy tuning of the porosity. In this way, it may be easier to map each constructed model with one or more shale reservoirs on the basis of their porosity. Capturing the correct characteristics of the pore, such as the pores size distribution (PSD), the limiting pore diameter distribution (LPD), the specific pore surface, the specific pore volume, will enable as later to have reliable calculation of the diffusion relevant fluids (shale gas and fracturing fluid).

Our second objective is the development of a methodology and a general purpose software for the calculation of the characteristics of the porosity in amorphous materials. The approach should be accurate, fast, scalable with system size and applicable to any generated model for kerogen. Having a wealth of information on the pores is necessary for the categorization of the various models and at the same time, enables the systematic calculation of their barrier properties and the study of the shale gas transport through these materials.

1.2.2. Clay models and other shale mineral models

Muscovite, kaolinite, and quartz are identified as the major components found in the Bowland shale according to its averaged mineralogical composition provided by WP3. In addition to these minerals, smectites can also be found, even if in smaller amounts, in other shale rocks. In order to study the molecular level interactions of the shale with kerogen, fracturing fluids, hydrocarbons, and NORM (radium, strontium, and barium), it is important to consider a first step in which the contribution of each of the minerals to the overall properties of the shale is evaluated separately. The objective of the current stage is to develop new molecular models of muscovite, smectite, kaolinite, and quartz based on the recent experimental and molecular modeling data, including the mineral particles of finite size, i.e., with realistically represented edges. Bulk atomistic models of smectites have already been recently developed using the same approach [10,11].

2. Methodological approach

2.1. Robust methodology and general purpose software for the study of porosity

There are numerous methods are reported in literature for the re-construction of porous carbonaceous materials. In a recent review [12] various methods were categorized into constitutive and mimetic:

- *Constitutive* methods
  - Slit type pores of graphite
  - Coarse grained and other approximate methods
- *Mimetic* methods
  - Quenched Molecular Dynamics (QMD)
  - Reverse Monte Carlo (RMC)
  - Hybrid reverse Monte Carlo (HRMC)
  - MD-HRMC

The latest method has been used specifically for the construction of kerogen models consisting of C, H and O, with the use of OREBO force field [13]. In general, the use of the various RMC base methods would require experimental data (such as the radial distribution

function, $g(r)$) and reactive force fields. These are major limitations since the most accurate force fields require extremely long computational times [12] and furthermore no parameterization exists for all elements of interest such as S. For this reason, we adopted another approach reported in literature.

2.1.1. Construction of bulk kerogen models

Initially, we adopt a representative model of the kerogen molecule and then a systematic method is used for the construction of realistic bulk kerogen configurations based on this molecule. The problem of molecular modelling of kerogen is long lasting, with the first efforts to elucidate the structure of kerogen dating back the 1980s. The first 2D molecules of kerogen were constructed by Behar and Vandenbrouke [14], followed by Siskin [15]. Later Siskins 2D model was converted into a 3D representation [16]. One last approach towards this direction was made by Ungerer et al. [17-19]. For different types of kerogen at different maturation stages they developed small representative molecular structures and used them to construct bulk kerogen models and to study various of their properties. In our work, we adopt the mature type II molecule published by Ungerer et al. since this type of kerogen is of interest in many industrially important shale fields. This molecule has been given the name II-D and is shown in Figure 2.

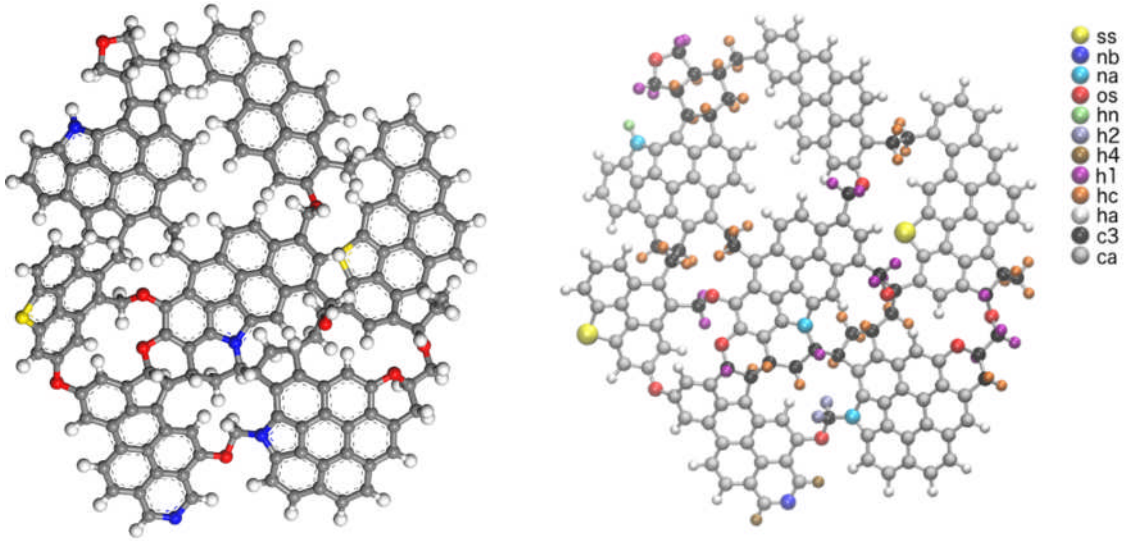


Figure 2: Molecular model of a kerogen unit representing overmature type II kerogen

The red, blue, yellow, grey, and white spheres represent oxygen, nitrogen, sulphur, carbon, and hydrogen atoms, respectively. The chemical formula of the kerogen unit is $C_{175}H_{102}O_9S_2N_4$, its molecular weight M_w=2468.9g/mol, and the atomic ratios H/C=0.58, O/C=0.051, S/C=0.011, and N/C=0.023. (GAFF atom type assignments are shown on the right: black (c3), grey (ca), purple (h1), light blue (h2), ochre (h4), white (ha), light orange (hc), lime (hn), blue (nb), red (os), yellow (ss)).

Two methodological approaches have been then used to generate molecular models of kerogene by two collaborating groups (ICPF and NCSR "D"). These approaches used different force field parameterizations in order to ensure the robustness of the resulting models.

Ungerer et al. [19] have used the PCFF force field. However, this option was not selected for our work, since due to its functional form PCFF is not compatible with other force field parameterizations intended for clays (CVFF, ClayFF) and other small molecules of interest that diffuse in shale gas relevant systems.

In the first approach (ICPF), the construction of large kerogen model structures was performed using the CVFF [20] which is a standard and reliable force field for simulations of organic molecules and is compatible with ClayFF [21]. A rather general method was applied in which model units with 170 to 260 carbon atoms (300 to 650 atoms in total) that match as closely as possible the elemental analysis, the distribution of functional groups, and the structural features of kerogen were first considered. These units then served as building blocks for the generation of dense porous kerogen structures. The initial kerogen unit had a chemical formula $C_{175}H_{102}O_9S_2N_4$, the molecular weight is M_w=2468.9g/mol, and its molecular model is shown in Figure 2, which also provides the composition and structural parameters of the unit.

To generate a representative kerogen model structure, 12 kerogen units were placed to a large simulation box of initial kerogen density about 0.05 g/cm^3 with periodic boundary conditions in all three dimensions. This the initial kerogen configuration was then gradually compressed and cooled down from a high temperature to pressure and temperature of typical reservoir conditions. The kerogen units were initially placed into the simulation box randomly without overlap and with the same orientation. We chose the same initial orientation of the kerogen units since they are flat and we anticipated that they should tend to stack parallel in a compressed state. The compression and cooling of the initial configuration was performed in step-wise fashion and Figure 3 displays an example of the time variation of the system temperature and kerogen density during molecular dynamics (MD) generation of the kerogen structures. The final structure typically contained ultramicropores (pore width < 1 nm) represented by the void spaces left between kerogen units in the compressed state. To introduce micropores (pore width of 1-2 nm) into our kerogen models, we used a dummy particle of varying size during the MD relaxation procedure. The dummy particle was then removed from the compressed kerogen structure, leaving a void where gas adsorption is likely to occur.

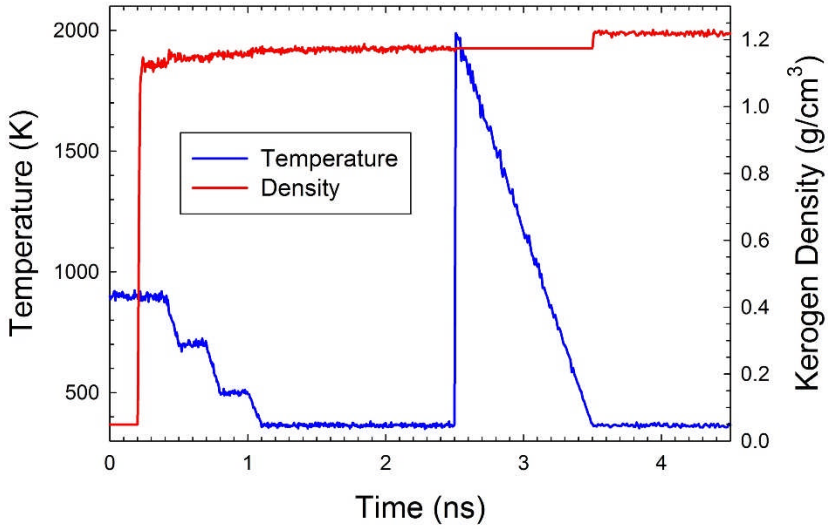


Figure 3: Example of time variation of the system temperature and kerogen density during the molecular dynamics generation of the porous kerogen structure.

The constructed kerogen structures were systematically characterised by calculating the geometric pore size distribution (PSD), pore limiting diameter, D_{min}, maximum pore size, D_{max}, accessible surface area, S_A, and pore volume, V_{pore}, and by analysing the pore network accessibility [22]. The PSD was determined by a Monte Carlo procedure that probed the kerogen pore space to find the largest spheres that contained test points and did not overlap with the atoms of the kerogen structure. The test points were randomly placed in the simulation box with no overlaps with the kerogen atoms. Each sphere then defined the volume of a void space that can be covered by a sphere of radius r or smaller, $V_p(r)$. The derivative $-dV_p/dr$ is the PSD and was obtained by numerical differentiation of $V_p(r)$ [23]. The value of D_{max} is a by-product of the PSD calculation.

The accessible surface area corresponds to the positions of the centre of a probe particle rolling over the kerogen atoms. It is thus defined as a locus of the points that represents the location of the probe particle at a distance of the collision diameter σ from a kerogen atom and at least a distance σ from all other atoms of the kerogen structure. As a probe particle, we used nitrogen atom with σ_N=0.3314nm to directly relate the calculated S_A to that measured in the BET adsorption experiments. An analogous Monte Carlo approach was used for calculation of the pore volume where a He-size probe of σ_{He}=0.258nm was used to provide values of V_{pore} consistent with that determined experimentally by helium pycnometry.

The analysis of the pore space accessibility aimed to detect if the pore network in the kerogen structures was fully accessible to a spherical probe of the size chosen to be equal to σ_N. The kerogen pore network was accessible to the spherical probe if, for this probe, it was possible to construct a continuous trajectory from one face of the simulation box to the opposite face of the simulation box without overlapping with the kerogen atoms. The value of D_{min} is a by-product of the percolation analysis. The pore limiting diameter, D_{min}, corresponds to a maximum probe size for which a pore network still percolates.

For the second approach (NCSR "D"), two other general propose force fields were selected:

a) Dreiding force field [24] with Gasteiger atomic point charges [25] was cused as a simple widely available force field, which is easy to implement and use;
b) The General AMBER force field (GAFF) [26], which is a force field extensively used for modelling molecules usually hosted by kerogen, e.g. small organic molecules. Testing on a wide array of molecules showed that GAFF produces minimized structures closer to their crystallographic counterparts, (i.e. lower root mean square displacement) compared to Dreiding.

Both of these force fields are also compatible with ClayFF [27]. In order to derive GAFF parameters, the selected representative molecular structure of the II-D kerogen was optimized by means of DFT calculations, using the Becke 3-parameter Lee-Yang-Parr (B3LYP) functional [28-30], coupled with the 6-31G* basis set. Then partial atomic charges were obtained from the electrostatic potential (ESP) of the optimized geometry. The ESP was calculated at the HF/6-31G* level of theory, according to the Mertz-Kollman population analysis scheme [31,32]. Then under the assumption of small deviations of the conformation during the dynamical simulations atomic charges were derived according to the Restrained Electrostatic Potential (RESP) protocol [33], utilizing the ANTECHAMBER module [34] of the AMBER12 suite of programs [35].

Similarly to the previous approach, the construction of larger bulk kerogen models consisted of two major stages. Initially a number of molecules were placed in a box at low density, the system was heated to a high temperature and then a cooling protocol was applied to the structure. Initial configurations were constructed at low density of 0.01 g cm^{-3} using the Amorphous Builder software integrated in the Scienomics MAPS platform [36]. Amorphous builder employs a configurational bias scheme for the placements of molecules in a box at a pre-specified density. Repeated attempts to construct initial configurations at higher density, closer to the actual system density, by using this approach have all failed.

In order to investigate ways to impose and control porosity, apart from examining the effects of the force fields tested, two more factors are examined: (i) the effect of the system size and (ii) following Ungerer et al. [19], introduction of LJ dummy particles. Three system sizes are compared, namely a small one comprising 15 II-D molecules, a medium of 50 II-D molecules and a large one having 100 molecules. With respect to the dummy particles, unlike previous efforts, the parameters varied here are both their number and their size. The size of the dummy particles varies from 15 Å to 40 Å and their number from one to three. The mass of the particles is that of Ar, while their interaction parameter is 0.20 kcal mol^{-1}.

Because of the fact that the II-D molecule is a bulky plain molecule, with very limited flexibility (almost rigid), it is anticipated that the ergodic hypothesis will not hold at the conditions examined. For this reason, multiple configurations are constructed for each size and each LJ dummy particle set up as it is shown in Table 1.

Table 1: Details of the number of configurations constructed, categorized based on (a) the size of the system, (b) the size and the number of the LJ particles and (c) the applied force field. The bar (-) represents not attempted calculations.

# II-D molecules	LJ particle		# configurations	
	σ (Å)	# particles	Dreiding	GAFF
15	None		16	16
	15	1	16	-
50	None		5	6
	15	1	-	6
		2	-	6
		3	-	6
	20	1	-	6
		2	-	6
		3	-	6
	30	1	-	6
		2	-	6
	40	1	-	6
100	None		5	6

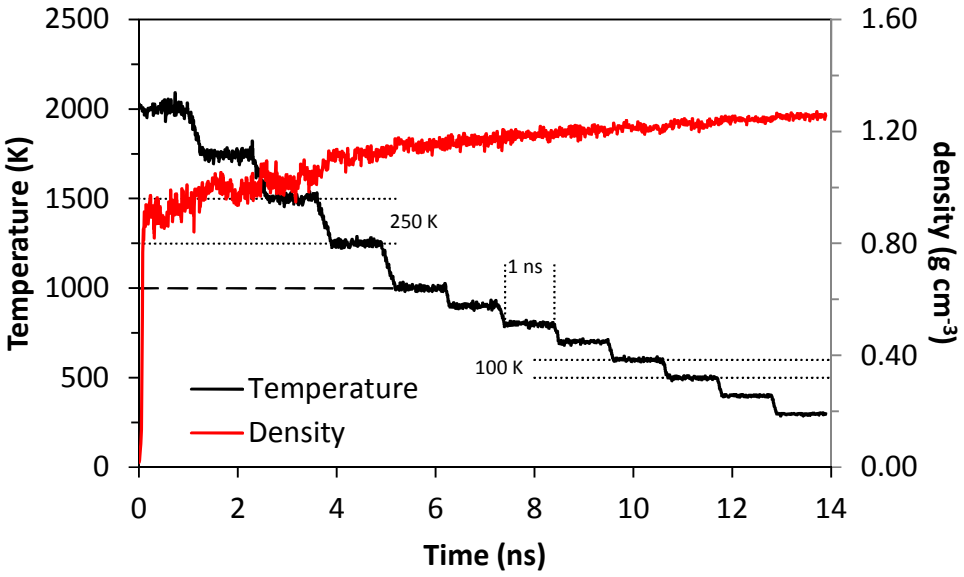


Figure 4: Staged cooling protocol for the construction of 15 II-D bulk kerogen models demonstrated for a run using GAFF.

The system is cooled to 1000 K by four cycles of 0.3 ns cooling at rate of 833K ns^{-1} followed by 1 ns of NPT equilibration at each cycle. Further reduction of the temperature to 300 K is achieved by cooling steps lasting 0.1 ns with rate equal to 1000 K ns^{-1}. The density evolution along the runs is also shown for comparison.

Every configuration that was generated using amorphous builder as detailed in Table 1 was then relaxed using MD simulations [37] (performed using LAMMPS code [38]) at the isothermal isobaric statistical ensemble (NPT), by progressively lowering the temperature

according to a cooling protocol. Throughout the NPT simulations, pressure was kept constant at 1 atm. The cooling protocol followed for the "15 molecules" system is shown in Figure 4. The simulations started at 2000 K, with an equilibration of 1 ns. The system was then cooled using a stepwise procedure. At each step, the temperature was reduced linearly from the initial to final temperature, followed by a small NPT equilibration. The temperature steps are bigger but with a lower cooling rate for high temperatures down to 100K. Then shorter steps with slightly higher cooling rate cool the system down to 300 K.

A similar protocol is followed for the medium and large systems sizes, but in order to mitigate the computational cost, simulations started at 1100 K cooled down to 300 K by steps of 0.05 ns at rate 2000 K ns^{-1} followed by 0.8 ns of NPT equilibration. For systems containing LJ dummy particles, the same cooling protocols were used. At the end of these runs and after the LJ dummy particle is removed, an extra step of 2 ns NPT equilibration at 298.15 K and 398.15 K is applied. During the MD simulations, the Nose-Hoover thermostat and barostat were used to control the temperature and pressure, respectively [39,40]. The integration time step was set to 1 fs for all calculations. Electrostatic interactions summations are performed using the traditional Ewald summation technique in every simulation. A simple cut-off at 12 Å was applied to all non-bonded interactions.

Estimation and characterization of the porosity

Two programs were then used for the elucidation of the characteristics of the porosity: (a) Poreblazer [22] and (b) an in-house developed software based on Voronoi space tessellation. Poreblazer is a grid based methodology that is used only for the calculation of the pore size distribution (PSD) with default settings (UFF atom size [41]) except for the cublet size. For the small and medium sized systems, the cublet edge is set to 0.5 Å, while for the 100 II-D systems the size of the cublet is set to 0.7 Å. This choice was made in order to mitigate the computational cost. It is worth noticing that, for medium system sizes, even the 0. 5 Å cublet edge resulted in increased memory resources demands and significantly large computation time.

Voronoi tessellation methodology

The Voronoi tessellation methodology was developed for the detailed characterisation of amorphous and crystalline materials. Its core is a Voronoi tessellation of the material, followed by an analysis with algorithms and concepts of the graph theory. The software is written in hybrid code written primarily in python with parts of the calculations performed by FORTRAN subroutines for significant acceleration. The calculation of the Voronoi polyhedra is performed by Voro++ [42]. The properties targeted here are:

- the CH_4 accessible specific volume,
- the CH_4 accessible specific surface,
- the maximum and limiting pore diameters in the structure [43],
- the identification of individual pores,

- the determination of pores that are percolated and
- the identification of diffusion paths in percolated pores.

The flowchart of the methodology is shown in Figure 5. The analysis starts with the generation of the Voronoi polyherdra (hereafter referred as Voronoi cells or simply cells) given the radii of the atoms composing the material. In this study, we used the Van der Waals radii of the atoms. The defining vertices, edges and faces of the Voronoi cells that result from Voro++ is the information necessary for the subsequent analysis. A spherical probe is also assumed in order to identify the accessible instead of the free volume. The radii of the atoms are then increased by the radius of the probe. In the following text, whenever atomic radii are referred, we refer to these increased radii. Since we are interested in shale gas, the assumed probe is CH_4, the primary component of shale gas. The intention is to model CH_4 with the TraPPE force field [44] and for this reason the assumed probe has a radius of 3.73 Å, the radius of methane in this force field.

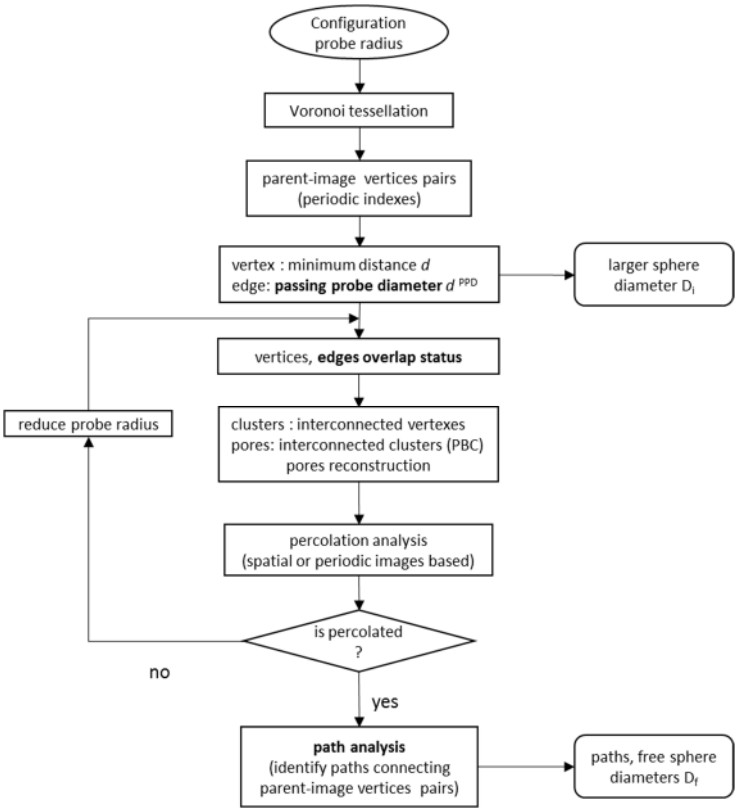


Figure 5: Flowchart of the Voronoi based analysis of the void space in amorphous materials.

Voronoi tessellation methodology: calculation of accessible volume and surface

The total accessible volume is calculated as the sum of the unoccupied volume of each Voronoi cell. It is calculated as the total volume of the cell, decreased by the volume of the fraction of the atom that lies within the cell. The volume of the atom that is within the cell is calculated using the algorithm developed by Dodd and Theodorou [45]; this algorithm calculates the remaining volume by cutting planes. In this case, the atom is the sphere and

the planes are the phased of the Voronoi cell. An identical approach results to the estimation of the accessible surface of the structure.

Voronoi tessellation methodology: Identification of accessible vertices and edges

The remaining analysis relays on the probe accessible vertices and edges of the Voronoi cells. These lay in the probe accessible space and hold the necessary information. Each accessible geometric locus is characterized by the biggest probe that can access its points. An accessible vertex is characterized by the biggest probe that it may host, while an accessible edge is characterized by the biggest probe that can traverse it.

Each Vertex is checked for overlaps with each nearest atoms and if found as non-overlapping the minimum edge from the neighboring atoms is calculated. This distance represents the biggest probe it may host. The edges are regarded as accessible if both its vertices are accessible and it does not intersect any of the atoms of the material. The latest condition is verified by solution of the sphere cutting ray problem with the atoms of the cells.

The accessible edges are characterized by the *passing pore diameter*, d_k^{PPD}, the biggest probe that can freely traverse the edge, including its constituting vertices. In order to calculate d_k^{PPD} we consider two cases. The case were the edge intersects the triangle defined by the three cells (i.e. atoms) by which it is shared, and the case in which the edge does not intersect the triangle (Figure 6). In the first case (Figure 6a) the d_k^{PPD} is equal to the diameter of a sphere whose center lays at the middle to edge and its surface is tangent to the three atoms defining the triangle. In the second case (Figure 6b) the distance from the surface of the atoms of equally distributed points (including the vertices) on the edge is calculated. The passing pore diameter d_k^{PPD} is then specified as the minimum of these distances.

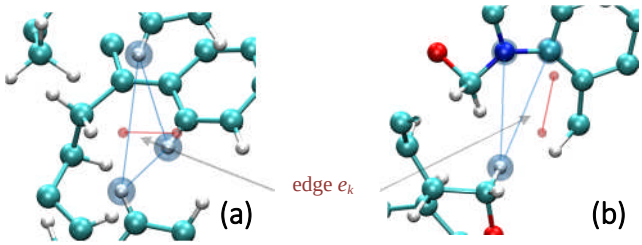


Figure 6: Non-overlapping edges.

Each non-overlapping edge, e_k, (red line), is assigned with the passing probe diameter i.e. the maximum diameter of the spherical probe which is possible to pass from the one constituent vertex to the other depending on the relative position of edge with the atoms by which it is shared: (a) e_k intersects with the triangle formed by the atoms by which e_k is shared and (b) e_k does not intersect with the triangle created by the atoms whose cells share e_k.

The periodic boundary conditions (PBC) imposed in our models as a standard practice may result at pairs or parent-image non-overlapping vertices (v_i, $v_{i'}$). At this point, the periodic indices $\mathbf{i} = (i_x, i_y, i_z)$ of each image $v_{i'}$ are identified for use in the following analysis.

Voronoi tessellation methodology: Identification of pores

The analysis of the resulted network of non-overalapping vertices and edges leads to the identification of the pores of the material. The network of accessible vertices and edges can be regarded as a bidirectional, disconnected (in the general case) graph *G*, while the clusters of interconnected vertices are the constituent connected sub-graphs, G_i. As a result, all the algorithms developed in the context of graph theory can be used here for the purposes of the characterization of the porous space of kerogen.

Topologically, the G_i clusters may be connected due to the periodic boundary conditions imposed in our models (Figure 7).

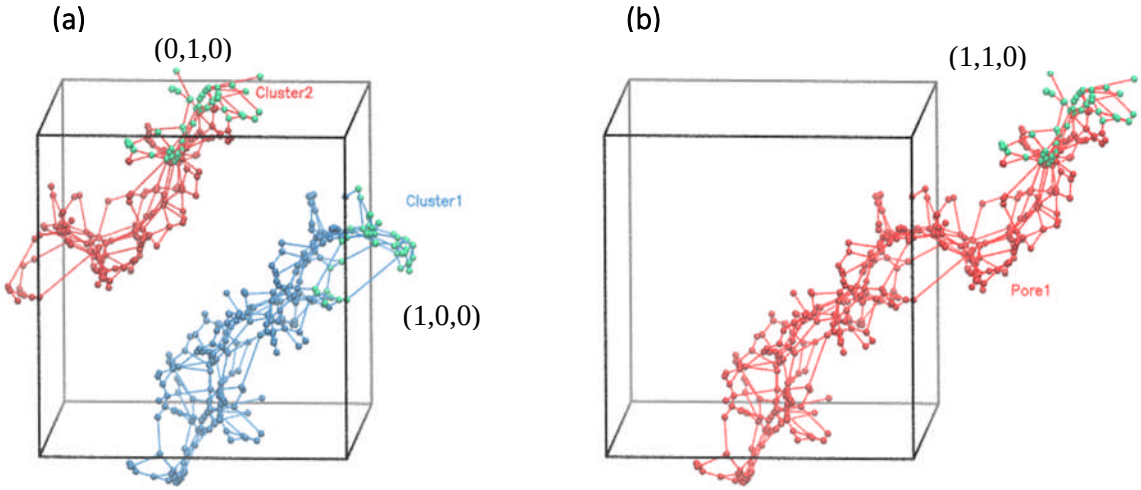


Figure 7: Merging two clusters, Cluster1 (C_1) with Cluster2 (C_2).

*(**a**) This merging may result in a pore where the image vertices still exist. (**b**) In this case, the periodic indexes are properly updated during the merging procedure and indicate a percolation channel characterized with the same direction.*

In this context, a pore is a set of clusters that share a pair of parent-image vertices i.e. if at least one vertex $v_{i'}^{C_2}$ of the second cluster C_2 is the periodic image of a vertex $v_i^{C_1}$ belonging to the first cluster C_1 or *vice versa*. Each pore of the system is then reconstructed by merging together cluster sharing parent-image pairs until no such pair can be found. Two clusters (C_1, C_2) are merged by translating the vertices of C_2 by the minimum vector separating a shared pair ($v_i^{C_1}, v_{i'}^{C_2}$). After the translation, the periodic indices of the image vertices that either themselves or their parents belong to C_2, are updated according to the translation performed. If the image vertices of the resulting cluster ($C_{12} = C_1 \cap C_2$) coincide with their parents (i.e. their periodic indexes after the update are zero) they are removed and the connectivity of their parents is updated properly taking care that no duplicated edges exist.

The free/accessible volume of each pore then is calculated as a sum of the individual shells' accessible volume as described earlier. It is worth noticing that if a Voronoi cell is shared between two or more pores, only the part of the cell belonging to the pore under consideration is taken into account.

Voronoi tessellation methodology: Percolation analysis

Each of the previously identified pores were then examined for percolation. A pore is percolated if at least one pair of parent-image vertices $v_i, v_{i'}$ exist in the pore. The set of unique periodic indices (i_x, i_y, i_z), of the image vertex, $v_{i'}$ indicates the direction(s) of percolation of the pore, i.e. which faces of the simulation box can be freely reached by a traversing probe.

To systematically trace the percolated diffusion paths in percolated pore it is useful to define the notion of an *island.* An island is a set of vertices that have the following properties:

- They are all image vertices of the same pore
- They all have the same periodic index **i**
- The are interconnected by accessible edges

The m^{th} island of periodic index **i** is notated as $I'^{m}_{\mathbf{i}}$,. Equivalently parent islands can be defined and they are indexed as $I^{m}_{\mathbf{i}}$. Then, all possible pairs of islands $I^{m}_{\mathbf{i}}, I'^{m'}_{\mathbf{i}}$ $m, m' = 1,2, \dots N^{I}_{\mathrm{i}}$ (N^{I}_{i} being the number of islands for periodic index **i**) are considered and for every pair a path connecting the two *islands* is found.

A path is defined as a sequence of connected edges. The length of the path is equal to the number of its edges. In a percolated pore a path of length ℓ connecting a pair of parent-image vertices, $v_i, v_{i'}$ (or equivalently a pair of parent-image islands $I^{m}_{\mathbf{i}}, I'^{m'}_{\mathbf{i}}$) as $P^{\ell}_{v_i, v_{i'}}$. The shortest path is the then determined as an optimization problem:

$$P^{min}_{v_i, v_{i'}} = \min_{\ell} P^{\ell}_{v_i, v_{i'}} \quad (1)$$

In the context of amorphous materials, we are interested in identifying the path that has the potential to allow faster and easier diffusion in the material. This is the path that is simultaneously the shortest possible path composed of edges of the larger possible passing pore diameter, d^{PPD}_{k}. These two conditions are not necessarily conflicting or not. This task is performed using the Dijkstra algorithm, implemented in networkx python library. While the unbiased Dijkstra algorithm identifies the path of minimum length connecting two vertices by assigning weights to the edges results in the identification of a path that meets the set criteria. Therefore, in order to find the path that meets our criteria for the pair $I^{m}_{\mathbf{i}}, I'^{m'}_{\mathbf{i}}$, we assign appropriate weights, w_k, to the edges of the pore:

$$w_k = (M - d^{PPD}_{k} + m)\, a \quad (2)$$

where, a is a scale factor used to increase the separation between the w_k that correspond to the smaller, m, and the larger, M, values of d^{PPD}_{k} values found in the pore:

$$M = \max_{k}\{d^{PPD}_{k} : k \in pore\ edges\}, \text{and} \quad (3)$$

$$m = \min_k \{d_k^{PPD} : k \in pore\ edges\}. \quad (4)$$

It should be noted that in the way the weights w_k are defined the larger the d_k^{PPD}, the smaller the corresponding weight. Therefore, the traced paths meet the criteria set and the maximum probe that can traverse the path corresponds to the edge with the largest weight.

Due to the nature of the systems, we are interested in, it is possible that the path traced for the pair $I_{\mathbf{i}}^{m}, I'^{m'}_{\mathbf{i}}$ does not include the edge corresponding to the real maximum passing probe diameter. To overcome this problem, the edge with the maximum weight, w_k, i.e. the minimum passing probe diameter of the path identified is eliminated from the pore's graph and the path is recalculated a procedure repeated until the graph becomes disjoined (i.e. no more paths can be found). The free particle diameter corresponds to the examined islands pair, $d^{PPD}_{I_{\mathbf{i}}^{m}, I'^{m'}_{\mathbf{i}}}$ is the maximum value of the calculated minimum passing pore diameters of all the traced paths. In turn, the maximum $d^{PPD}_{I_{\mathbf{i}}^{m}, I'^{m'}_{\mathbf{i}}}$ for all the island pairs corresponds to the limiting pore diameter $d^{PPD}_{i_{pore}}$ of a percolated pore.

Voronoi tessellation methodology: Maximum and limiting pore diameters

Finally the maximum included sphere, D_I, that can be hosted in the structure and the limiting pore diameter or free sphere diameter, D_f, are calculated. D_I, is obtained easily as the maximum value of the distances d_j that characterize the accessible vertices:

$$\mathrm{D}_I = \max(\{d_j\}) \quad (5)$$

D_f the critical penetrant diameter above which free diffusion in the structure would not be possible. It is by definition the maximum value of the calculated limiting pore diameters:

$$\mathrm{D}_\mathrm{f} = \max\{d^{PPD}_{pore\,i}\} \quad (6)$$

given that percolated pores exist in the system for the probe diameter in use. Thus, when the accessible volume comprises no percolated pores, the assumed probe is reduced gradually, the pores are recalculated and percolation analysis follows. This procedure is repeated until the probe is sufficiently small to have percolated pores. Then diffusion paths are then identified and the limiting pore diameters and free sphere diameter D_f are calculated accurately.

2.1.2. Construction of bulk and edge clay models

Montmorillonite (smectite) and muscovite (illite) are both 2:1 dihoctahedral phyllosilicates. They consist of an octahedral layer of aluminum oxide (AlO_6) sandwiched between two tetrahedral layers of silicon oxide (SiO_4), giving a so-called TOT layer. Both clay minerals always bear a permanent charge and pH-dependent charge. In addition to their charged structures these clays also have large surface areas, and these two features make them very reactive and important in many geological and environmental processes, and also for many industrial applications.

The permanent charge generally appears due to Al/Si substitutions in the tetrahedral layers for muscovite (giving a layer charge of -2*e*) and both Mg/Al substitutions in the octahedral layers and/or Al/Si substitutions in the tetrahedral layer for montmorillonite (giving a layer charge of -0.75*e*). This permanent charge is compensated by cations adsorption (muscovite), or cations and water adsorption (montmorillonite) on the basal surfaces of both clays, that are obtained by cutting these clay minerals along the (001) direction. The molecular structure and properties of such basal surfaces are sufficiently well studied and realistic reliable models already exist (e.g., [10,11,27,46]).

The pH-dependent charge mostly originates from hydrolysis of the broken Al-O, Si-O, or Mg-O bonds at the edges of the TOT layers [47]. Different edge facets can be obtained depending on the cleavage direction considered. The surface energies upon cleaving and hydrolysis of the broken bonds have been determined for the most common dioctahedral phyllosilicate edges facets. These studies reveal that if we consider the chemisorption of the same amount of water molecules after cleavage, the most dominant edge facet appear along the (110) direction [48-50] followed by the (010) edge that is more stable than the remaining (100) and (130) edge facets which need more water molecules to compete with the first two [48,49]. Hence the clay models that will be constructed and studied in this work will be based on the three main cleaving directions: the (001) direction so-called basal surface, and the (110) and (010) edge facets.

To construct the basal and edge structures of muscovite and montmorillonite, we started from the orthogonal pyrophyllite structure from Drits et al. [51] (Figure 8).

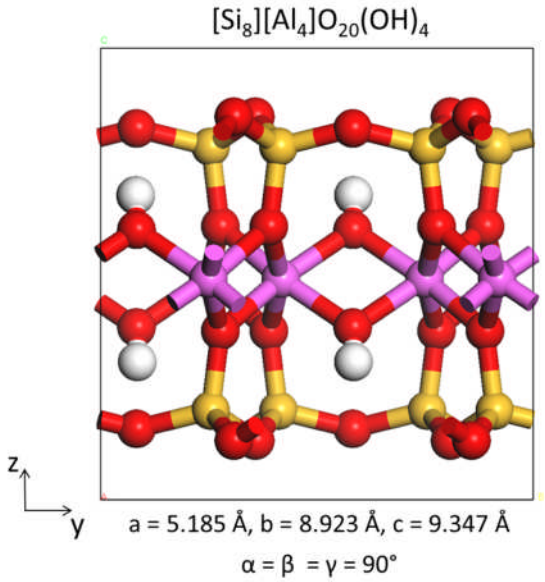


Figure 8: Unit cell structure of pyrophyllite. Yellow – Si, pink – Al, red – O, grey – H.

(001) Surfaces

To get the basal surface of this pyrophyllite crystal, the structure was cut through the (001) direction and the resulting unit cell was repeated along a and b vectors with factors of 4 and

2, respectively. The resulting cell has a surface area of 20.74 Å x 17.85 Å with the chemical formula $[Si_{64}][Al_{32}]O_{160}(OH)_{32}$. Based on the respective layer charges of muscovite and montmorillonite the following chemical compositions should be obtained in the cells.

- Muscovite $[Si_{48}Al_{16}][Al_{32}]O_{160}(OH)_{32}Ka_{16}$ *(Musc)*
- Tetr.-charge montmorillonite $[Si_{58}Al_6][Al_{32}]O_{160}(OH)_{32}Na_6$ (*T-Mont*)
- Oct.-charge montmorillonite $[Si_{64}][Al_{26}Mg_6]O_{160}(OH)_{32}Na_6$ (*O-Mont*)
- Mixed tetr.-oct.-charge montmorillonite $[Si_{62}Al_2][Al_{28}Mg_4]O_{160}(OH)_{32}Na_6$ (*TO-Mont*)

The recently developed *Supercell* program [52] was used to randomly assign substitutions in the TOT layers of pyrophyllite in order to generate several substitution arrangements corresponding to the chemical compositions mentioned above. For each of the 4 cases, the substituted TOT layer structures were sorted based on their electrostatic energy (calculated directly by Supercell program) and the three lowest energy structures were selected. Then these structures are repeated again along a and b vectors (surface area now is 41.48 Å x 35.70 Å) and combined together to provide a three TOT-layered structures. In the case of muscovite, two of the three interlayers were filled with potassium ions, and the third one is expanded to create a vacuum space (60 Å thick) in contact with the external surfaces on that will be filled by aqueous solution or a more complex fluid. For montmorillonite, two of the three interlayers were filled with sodium ions and a water content equivalent to one water monolayer. This is the most sable hydration state of sodium in montmorillonite according to previous experimental and simulation studies. The third interlayer is expanded as for muscovite to create a 60 Å thick vacuum space that will be filled by aqueous solution (Figure 9). Note that these four clay models will be used to study how the basal surface properties vary with the charge localization in the TOT layers, and also with the amount of this charge in the case of muscovite and tetrahedral-charge montmorillonite.

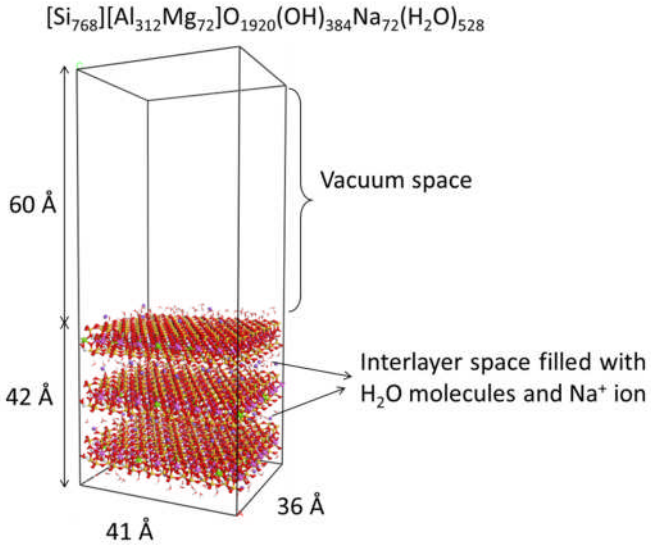


Figure 9: Octahedral-charge montmorillonite (001) structure.

(110) Surfaces

The unit cell structure of pyrophyllite from Drits et al. [51] was cut along the (110) plane and repeated by factors of 4 and 2 along the *b* and *c* crystallographic vectors to generate a simulation supercell with surface area of 20.64 Å x 35.83 Å. Note that the crystal's stoichiometry is kept the same throughout the procedure. The edge structures obtained correspond to the so-called AC-type periodic bond chains (PBC) [53] and bear dangling Si-O bonds and 5-fold Al atoms in the octahedral layer. The chemisorption of H_2O molecules on these edges yields Si-OH groups. Further physisorption of H_2O molecules on these 5-fold Al atoms will then give Al-OH_2 groups at the edges in addition to the Si-OH groups.

The surface topology of the (110) pyrophyllite edge at neutral pH is provided in Figure 10 as suggested by crystal growth theory [53] that also highlights the changes occurring when a tetrahedral/octahedral substitution is inserted in the structure. Ab-initio MD simulation studies [54] have observed a similar non-substituted (110) edge topology. However, the most stable surface terminations in the presence of a tetrahedral or octahedral substitution differ from what is suggested by crystal growth theory. Moreover, it also reveals that the Mg substituting atom at the edge is likely to occur either in 5-fold or 6-fold coordination. In a recent simulation study, the effect of octahedral substitution and layer charge on the edge structure of montmorillonite was studied [55] using classical ClayFF force field. It appears from that study that the Mg substituting atom at the edge tends to adopt the 5-fold coordination with increasing layer charge. In these calculations the less charged structures were probed than the ones we are considering in our work (-0.25*e* and -0.375*e* vs -0.75*e*).

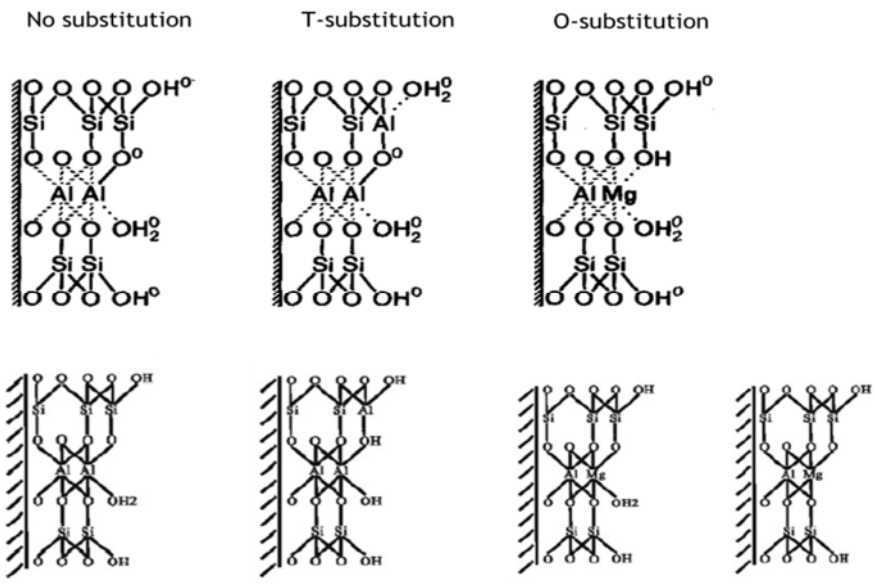


Figure 10: (110) surface topologies from crystal growth theory ([53], upper row) and from ab-initio MD simulation ([54], lower row).

As for the (001) surfaces, the *Supercell* program [52] was used here to randomly assign the substitutions in the TOT layers to generate three statistically independent TOT layers for each of the *Musc*, *T-Mont*, *O-Mont*, and *TO-Mont* model structures. For *O-Mont* and *TO-Mont* the Mg/Al substitution positions were modified in order to have one of the two parallel (110) edge surface bearing an Mg substitution in a solvent accessible position, and the other edge bearing an Mg substitution in the linking PBC (see Figure 11 below). The other substitutions were randomly distributed among the inner PBC chains. Based on the previous simulation evidence the following coordination numbers were defined for the metal atoms in the octahedral layers of the two edge surfaces:

- Edge with Mg/Al substitution at a solvent accessible position: out of 4 metal atoms (1 Mg and 3 Al), one Mg and one Al are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Mg/Al substitution at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.

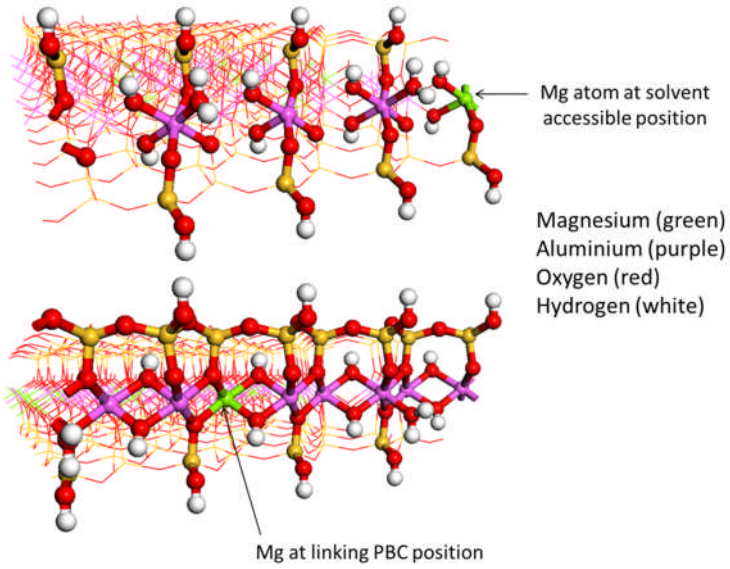


Figure 11: (110) edge facets surface showing the solvent accessible and linking PBC positions, as well as the coordination of the different metal atoms at the edges.

For *TO-Mont* in addition to the octahedral substitutions assigned as mentioned above, the Al/Si substitutions positions were modified so that in one TOT layer there is one Al substitution in a solvent accessible position in one of the two T layers, and one Al substitution in the linking PBC position in the other T layer (see Figure 10, surface topology from ref [54]). Hence the surface topology in this case is similar to that of *O-Mont* (see above) with Al(tet)-OH-Al(oct)-OH groups instead of Si(tet)-O-Al(oct)-OH_2 groups. The following coordination numbers are defined for the metal atoms in the octahedral layers of the two edge surfaces:

- Edge with Mg/Al and Al/Si substitutions at a solvent accessible position: out of 4 metal atoms (1 Mg and 3 Al), one Mg and one Al are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Mg/Al and Al/Si substitutions at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.

As previously, for *T-Mont* the Al/Si substitutions positions were modified so that in one TOT layer there is one Al substitution in a solvent accessible position in one of the two T layers, and one Al substitution in the linking PBC position in the other T layer. The remaining Al/Si substitutions are randomly distributed among the inner PBC chains.

The same also applies for muscovite, but this case in one TOT layer there are 2 Al substitutions in solvent accessible positions in one of the two T layers, and 2 Al substitutions in the linking PBC positions in the other T layer (see Figure 3) and the remaining Al/Si substitutions randomly distributed among the inner PBC chains. For *T-Mont* and muscovite we assume that the tetrahedral layer charge also affects the edge metal atoms coordination. The following coordination numbers are defined for the metal atoms in the octahedral layers of the two edge surfaces for the two structures:

- Edge with Al/Si substitutions at a solvent accessible position: out of 4 metal atoms (4 Al), two Al are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Mg/Al and Al/Si substitutions at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.

The obtained structures are further enlarged by a factor of 2 along the b vector to get 3 TOT layers of 41.28 Å x 35.83 Å surface areas that are combined similarly to the (001) surface to generate our simulation cells.

(010) Surfaces

The unit cell structure of pyrophyllite from Drits et al. [51] was cut along the (010) plane, keeping the crystal's stoichiometry and the resulting structure is repeated by factors of 4 and 2 along the *a* and *b* crystallographic vectors to generate a supercell cell with surface area of (20.74 Å x 35.69 Å). The edges structures obtained correspond to so-called B-type periodic bond chains (PBC) [53] and bear dangling Si-O bonds and 5-fold Al atoms. The chemisorption of water molecules on these edges surfaces yields Si-OH and Al-OH groups. Further physisorption of water molecules on these 5-fold Al atoms give additional Al-OH_2 groups on the edges. According to the crystal growth theory [53], the (010) edge surface topology is defined as shown in Figure 12 below for different cases with no substitution, tetrahedral substitution, or octahedral substitution. The corresponding surface topologies from simulation studies [54] are also provided in Figure 12. In both cases, we can see that the octahedral substituting Mg atoms bear two OH_2 molecules in a 6-fold coordination unlike the Al atoms that bear 1 OH and 1 OH_2 groups. Moreover, it has been shown from ab-initio calculation that when the edge is exposed to water the Mg-(OH)(OH_2) configuration at the

edge readily transforms into Mg-$(OH_2)_2$ through a proton capture from the Si-OH that form dangling Si-O bond that is later saturated by capturing a proton from the Al-(OH)(OH_2) [54,56]. When a tetrahedral substitution is present in the edge octahedral Al atoms can occur, with the same probability either 6-folded or 5-folded (Figure 12).

The substitution distribution in the layers for muscovite and montmorillonite models (*T-Mont*, *O-Mont*, and *TO-Mont)* follows all the rules mentioned in the previous section for each of these cases. Unlike the (110) edges, the substituting Mg atoms at the edge are always 6-fold coordinated. Hence for the muscovite and montmorillonite structures we have the following coordination environment for the metal atoms in the octahedral layers of the two parallel edge surfaces:

- Edge with Mg/Al substitution at a solvent accessible position: out of 4 metal atoms (1 Mg and 3 Al), one Mg and one Al are 6-fold coordinated and the other two are 5-fold coordinated.
- Edge with Mg/Al substitution at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Mg/Al and Al/Si substitutions at a solvent accessible position: out of 4 metal atoms (1 Mg and 3 Al), one Mg and one Al are 6-fold coordinated and the other two are 5-fold coordinated.
- Edge with Mg/Al and Al/Si substitutions at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Al/Si substitutions at a solvent accessible position: out of 4 metal atoms (4 Al), two Al are 5-fold coordinated and the other two are 6-fold coordinated.
- Edge with Mg/Al and Al/Si substitutions at the linking PBC position: out of 4 metal atoms (4 Al) two are 5-fold coordinated and the other two are 6-fold coordinated.

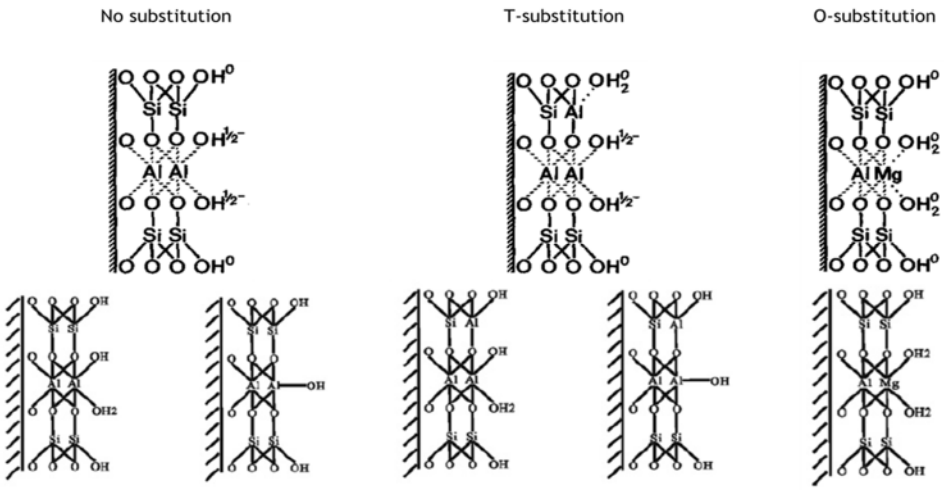


Figure 12: (010) surface topologies from crystal growth theory theory ([53], upper row) and from ab-initio MD simulation ([54], lower row).

3. Summary of activities and research findings

3.1. Kerogen structure characterization and analysis

For the kerogen models generated by the first approach (ICPF), the following analysis has been already performed. Starting from different initial configurations of kerogen units, 20 compressed configurations for each size of the LJ dummy particle were typically generated. The density equilibration was only slightly affected by the initial configurations. For adsorption simulations, we then selected a high density structure out of the 20 generated configurations that had the lowest configurational energy. In addition, we ran a 1ns MD and collected ten structures separated by 0.1ns for this structure. The dummy particle was then removed from the collected structures. The structures were statistically equivalent and represented the overmature type II kerogen. The size of corresponding simulation boxes was (4-5)nm x (3-3.5)nm x (2.5-3)nm.

Table 2 summarises the various properties used to characterise the kerogen structures. Figure 13 displays the PSD of the kerogen structures with different microporosity corresponding to the dummy particle with σ_{LJ} ={0, 0.9, 1.1, 1.3, 1.5}nm. Results in Table 1 and Figure 13 were averaged over the collected structures. Figure 14 provides examples of the kerogen structures with low and high microporosity. First, we see that density of the kerogen structures was typically between 1.2 and 1.3g/cm^3 which compared quite well with a range of the experimental density 1.2 to 1.4g/cm^3 found for mature and overmature kerogens [57]. Second, the PSD corresponding to no dummy particle suggested the presence of ultramicropores of size from 0.3 to 0.5nm, see also Figure 14a. The dummy particle created a host micropore (where gas adsorption is likely to occur) proportional to σ_{LJ} as indicated by the pronounced peak in PSDs, see also Figure 5b. Besides the host micropore the kerogen structures corresponding to σ_{LJ}={0, 0.9, 1.1, 1.3, 1.5}nm also contained smaller micropores and ultramicropores. Experimental PSDs for isolated kerogens cannot be reliably measured. However, experimental PSDs for shale (i.e., including organic and inorganic matter) suggest the existence of micropores larger than 1nm [58, 59], cf. PSDs corresponding to σ_{LJ}={1.3, 1.5}nm. Finally, the values of S_A, V_{pore}, and D_{max} gradually increased with increasing σ_{LJ} while values of D_{min} did not vary significantly with σ_{LJ}.

Table 2: The kerogen density, ρ, accessible surface area, S_A, pore volume, V_{pore}, pore limiting diameter, D_{min}, and maximum pore size, D_{max} of the kerogen structures with different microporosity corresponding to the dummy particle of the size σ_{LJ}. The simulation uncertainties are given in the last digits as subscripts.

σ_{LJ} (nm)	ρ (g/cm^3)	S_A (m^2/g)	V_{pore} (cm^3/g)	D_{min} (nm)	D_{max} (nm)
0	1.298_4	4.7_{16}	0.056_2	0.170_{11}	0.503_{22}
0.9	1.274_5	36.9_{34}	0.065_3	0.197_{24}	0.783_{13}
1.1	1.247_6	76.7_{57}	0.084_4	0.232_{20}	0.901_{45}
1.3	1.231_5	131.2_{75}	0.094_3	0.231_{23}	1.153_{18}
1.5	1.208_4	150.6_{49}	0.118_3	0.188_{11}	1.350_{14}

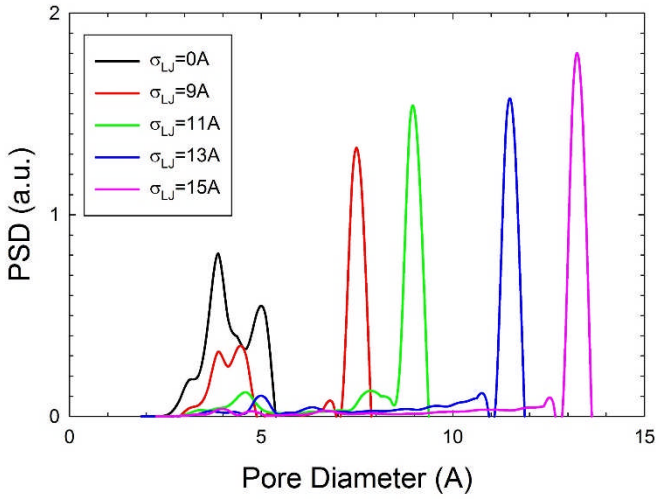


Figure 13: Geometric pore size distribution, PSD, as a function of the pore diameter for the porous kerogen structures with different microporosity. The microporosity was introduced via the dummy particle of size σ_{LJ}.

Grand Canonical Monte Carlo (GCMC) simulations were employed to study the adsorption of two proxies of the shale gas (pure methane and mixture of 82% of methane, 12% of ethane and 6% of propane) in the kerogen structures. The shale gas adsorptions are compared with GCMC simulation of CO2 adsorption in the kerogen structures. Hydrocarbons are modelled using the all-atom OPLS force field [60] while carbon dioxide is represented by the EPM2 model [61]. In these trial GCMC simulations, the kerogen structures were considered rigid for the sake of computational efficiency, neglecting possible swelling of adsorbent by the adsorbed molecules.

Table 3: The adsorbed amount of methane, n_{CH4}, in the microporous structures of overmature type II kerogen at a temperature of 365K and pressure of 275bar.

σ_{LJ} (nm)	n_{CH4} (mmol/g)
0	0.83_8
0.9	1.03_4
1.1	1.39_9
1.3	1.61_6
1.5	2.00_2

Table 4: The adsorbed amount of the shale gas (i.e., mixture of 82% of methane, 12% of ethane and 6% of propane), n_{CH4}, n_{C2H6}, and n_{C3H8}, in the microporous structures of overmature type II kerogen at a temperature of 365K and pressure of 275bar.

σ_{LJ} (nm)	n_{CH4} (mmol/g)	n_{C2H6} (mmol/g)	n_{C3H8} (mmol/g)
0	0.56_7	0.15_4	0.05_3
0.9	0.56_4	0.19_3	0.14_4
1.1	0.79_9	0.21_3	0.20_4
1.3	0.81_4	0.31_3	0.28_4
1.5	1.01_2	0.30_1	0.39_1

Table 5: The adsorbed amount of carbon dioxide, n_{CO2}, in the microporous structures of overmature type II kerogen at a temperature of 365K and pressure of 275bar.

σ_{LJ} (nm)	n_{CO2} (mmol/g)
0	1.16_7
0.9	1.46_5
1.1	1.88_9
1.3	2.08_7
1.5	2.57_1

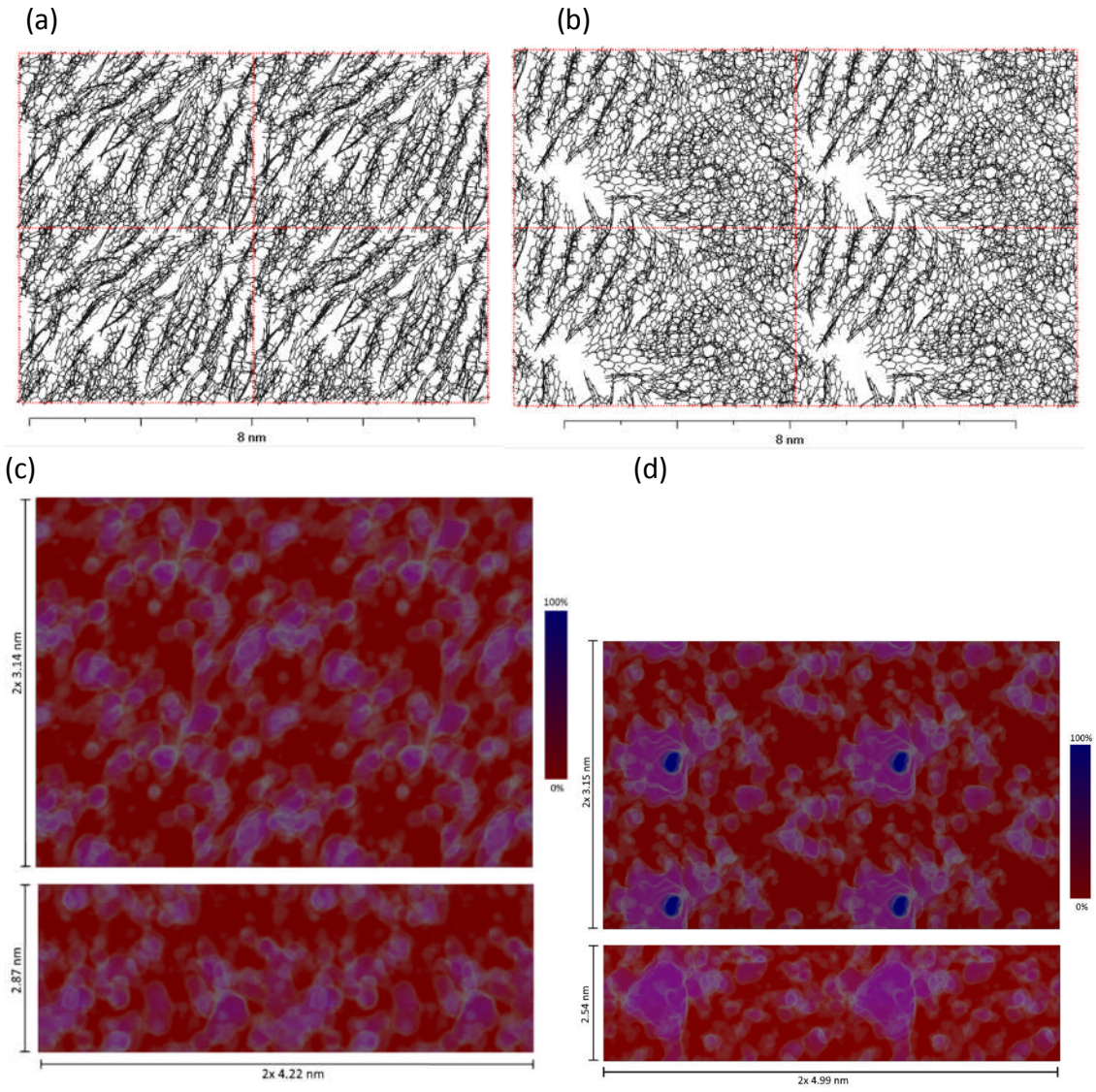


Figure 14: Examples of the kerogen structures with low and high microporosity.

The left column corresponds to the structure generated without the dummy particle while the right column is for the structure generated with the dummy particle of size σ_{LJ}=1.3nm. We displayed the periodic images for the sake of clarity. In portions (a) and (b) of the figure, the red, blue, yellow, grey, and white spheres represent oxygen, nitrogen, sulphur, carbon, and hydrogen atoms, respectively. The portions (c) and (d) of the figure show the front and top view of the kerogen structures with the cutting plane in the middle of simulation box to demonstrate percolation of the pore network probed by nitrogen sphere.

Tables 3-5 list the simulated adsorbed amount of methane, mixture of methane, ethane and propane as a proxy of the shale gas, and carbon dioxide, respectively. We first see that adsorbed amount, n, increased with rising σ_{LJ} although that increase was small for the kerogen structure corresponding to σ_{LJ}=0.9nm. Second, gas adsorption onto the kerogen structures was stronger for CO_2 than for CH_4 and the ratio n_{CO2}/n_{CH4} was about 1.3. The same finding about CO_2 and CH_4 adsorption onto similar type II kerogen structures was reported by Sui and Yao [62]. A stronger CO_2 adsorption in comparison with CH_4 adsorption on Barnett's shale samples was also experimentally measured by Heller and Zoback [63]. This can be attributed to stronger van der Waals and electrostatic gas-gas and gas-adsorbent interactions in the case of CO_2 adsorption. To better understand the stronger CO_2 adsorption, we performed an additional GCMC simulation for the kerogen structure corresponding to σ_{LJ}=1.5nm and switched-off the partial charges in the CO_2 model. We found that the adsorbed amount decreased by about 10% with respect to the original CO_2 model, demonstrating an enhancement of CO_2 adsorption due to electrostatic interactions. However, the adsorbed amount of CO_2 of the modified model was still higher than that for CH_4, demonstrating stronger van der Waals interactions for CO_2-kerogen systems compared to CH_4-kerogen systems. Third, our results for the proxy of Barnett's shale gas, i.e., a mixture of 82% of methane, 12% of ethane and 6% of propane indicated preferential adsorption of propane and to a lesser extent ethane. An example is that for the kerogen structure without the host micropore σ_{LJ}=0nm), the adsorbed mixture contained about 74% of CH_4, 20% of C_2H_6 and 6% of C_3H_8. With increasing microporosity (i.e., with increasing σ_{LJ}), the molar percentage of methane decreased, that of ethane was roughly constant (i.e., around 20%) and that of propane gradually increased. For instance, for the kerogen structure corresponding to σ_{LJ}=1.5nm, the adsorbed mixture consisted of about 59% of CH_4, 18% of C_2H_6 and 23% of C_3H_8. The strong preferential adsorption of propane molecules can be associated with the condensation of the propane molecules in larger micropores since, in contrast to methane and ethane (which are in supercritical state), propane is in subcritical state above the vapour pressure. A similar observation was reported by Falk et al. for subcritical dodecane in model kerogens [64].

For the kerogen models constructed using the second approach (NCSR "D"), a more detailed structural analysis has been performed.

Density

The density as function of temperature for the different system sizes for the kerogen model structure created without the aid of LJ dummy particles for Dreiding and GAFF force fields is shown in Figures 15 and 16, respectively. All values are obtained as averages over time and for different runs. In both cases, the average value of the density decreases with increasing system size reaching a limiting value already for the 50 molecules system. As expected, the statistical uncertainty becomes narrower with increasing system size. The GAFF force field produces denser structures for every system size and temperature by a factor of 6.9% on average and a maximum of 16%. The system size effect is evident for both force fields. These values cannot be compared directly with any experimental results but are within the experimental range.

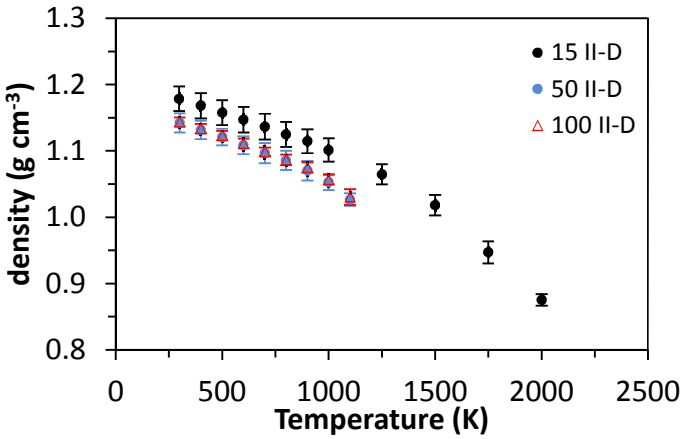


Figure 15: Density of kerogen structure as a function of temperature along the cooling procedure for Dreiding force field.

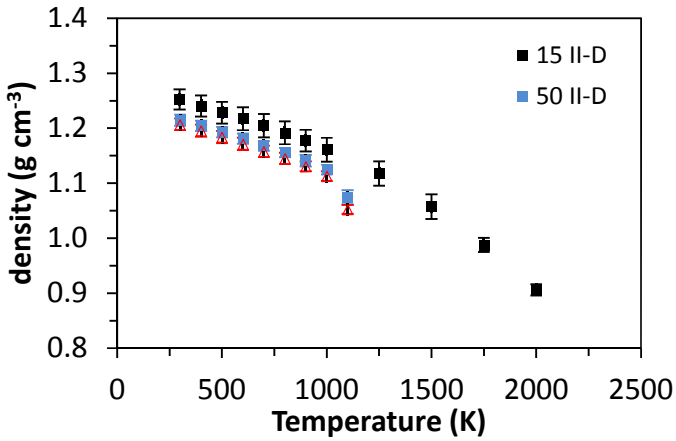


Figure 16: Density of kerogen structure as a function of temperature along the cooling procedure for GAFF force fields.

As it has already been mentioned, a number of structures was created with the aid of LJ dummy particles. Despite the fact that this approach was also tested for Dreiding and small systems is it is mainly focused to medium size (50 II-D molecules) structures using GAFF. It was regarded unnecessary to study large systems since the bulk density reached a limiting value for the medium sized systems. In addition to that, GAFF is more accurate than Dreiding at least in the structural reproduction of small organic molecules. Figure 17 shows the effect of the LJ dummy particles and pressure on the calculated density for these configurations at 298 K. The structures generated with the aid of LJ dummy particles are less dense with decreasing density as the size of the particle(s) increase, a finding that is in agreement with intuition. This behaviour is not observed for the largest 40 Å LJ particle, where the structure is denser than the structure constructed with a 30 Å particle. The same is the effect of the number of LJ particles on the density, with decreasing density as the number of particles increases. The majority of the models are denser than 1.1 g cm^{-3} with the smallest observed density of 1.04 g cm^{-3} obtained when two 30 Å LJ particles are used. With the exception of the case of the one LJ particle with diameter 20 Å, the increase of the pressure from 1 to 250 atm results in an increase of the density around 0.1%, indicating the small effect of the pressure on the density.

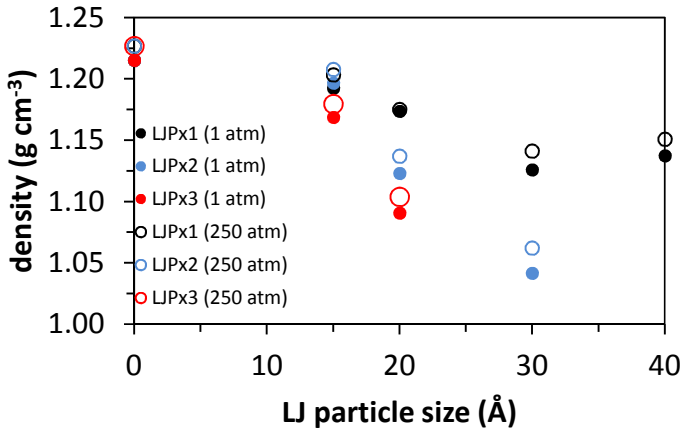


Figure 17: Density as function of number and size of LJ dummy particles (LJP) for runs with 50 II-D molecules using GAFF force field. Temperature is 298.15 K.

Porosity without the aid of LJ particles

The pore size distributions calculated for the systems constructed without the presence of LJ dummy particle(s) at 298 K are shown in Figures 18 and 19 for the Dreiding and GAFF force fields, respectively. For the small system sizes of 15 II-D molecules, the PSD is calculated also at 600 K. In all cases, the smallest pore diameter is approximately 2.5 Å. The maximum pore diameter observed using Dreiding force field is roughly 12 Å for all sizes. The biggest pores constructed with GAFF and 15 or 50 II-D molecules approach the value of 10 Å wile biggest pore size jumps to 14 Å for the large 100 II-D molecules systems. For both force fields, temperature is found to have negligible effect on the pore sizes of the constructed kerogen models, a finding suggesting limited mobility of the bulky II-D molecules in a congested environment, even at temperatures as high as 600 K. As the system size increases, larger pores appear in the configurations, (decrease of the height peak of the PSD with increased value for the bigger pores) while only for the GAFF 100 II-D molecule system this is combined with an increased maximum pore diameter from 12 to 14 Å. This indicates that the PSD is system size dependent and that it is necessary to move to relatively large system sizes consisting of at least several tens of molecules when simulating kerogen with this type of models.

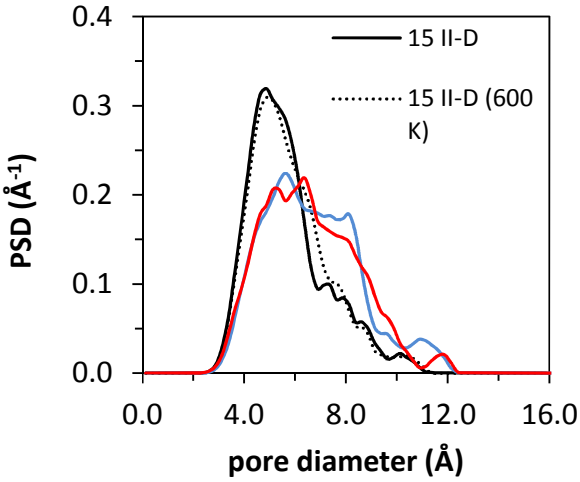


Figure 18: Pore size distribution (PSD) for different structure sizes constructed with Dreiding force field at 298.15K. The small system size PSD at 600K is also shown.

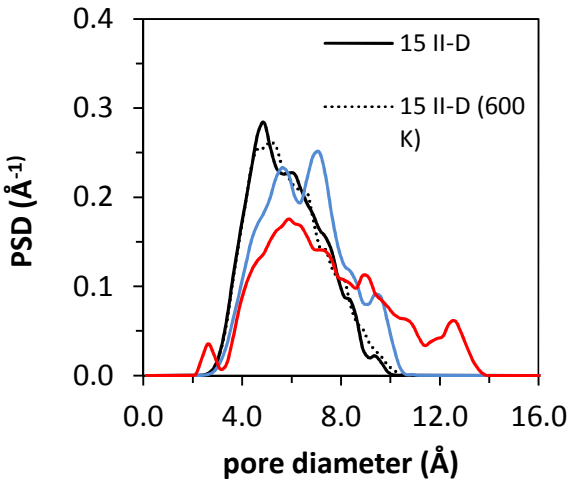


Figure 19: Pore size distribution (PSD) for different structure sizes constructed with GAFF at 298.15K. The small system size PSD at 600K is also shown.

Even though PSD describes the type of pores of the material, it is not sufficient to provide a complete picture of the pathways available for the diffusion of fluids in kerogens. In this respect, the limiting pore diameter or free sphere radius is necessary as a rough estimate for the size of the molecules possible to diffuse in the material. The calculated distribution of LPD for the two force fields as function of the system size is shown in Figure 20. Temperature is 298.15K for all cases. In addition to that, for the small systems size the temperature effect is shown by calculations also at 600K. In all cases the LPD value does not exceed the diameter of TraPPE CH_4 i.e. 3.73 Å which indicates that even the diffusion of small molecules such as CH_4 is difficult to occur especially in a static structure. The effect of system size and temperature in the LPD is in general considered minor.

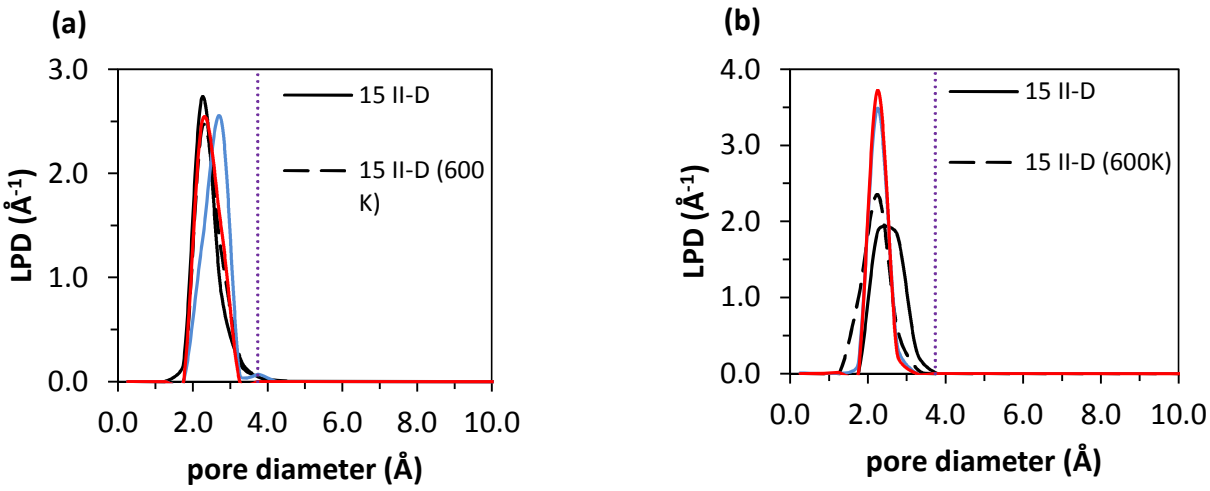


Figure 20: Limiting pore diameter distribution (PSD) for different structure sizes constructed with the (a) Dreiding and (b) GAFF force fields at 298.15K, respectively. The small system size LPD at 600K is also shown.

Porosity with the aid of LJ particles

The effect of the number of LJ particles can be seen in Figure 21. The PSD, LPD, maximum pore diameter and % porosity (calculated as the fraction of CH_4 accessible volume with respect to the total volume of the system) distributions are shown for the 50 molecules GAFF structures, constructed with the aid of different number of 15 Å LJ dummy particles. The use of LJ dummy particle(s) increases the accessible volume in the structure as it is evident from probability density of the porosity. This extra space is distributed in larger pores, as it is indicated by the broadened PSD (Figure 21b) with maximum pore diameters approaching 20 Å. The presence of bigger pores is also reflected in the MPD distribution (Figure 21d). In general, inclusion of more LJ particles allows the construction of bigger pores in addition to the small pores (~2-5Å) that remain in the structure. On the other hand, LPD distribution –a key characteristic for the study of diffusion (Figure 21c) – is not significantly affected by the increase in the number of LJ particles.

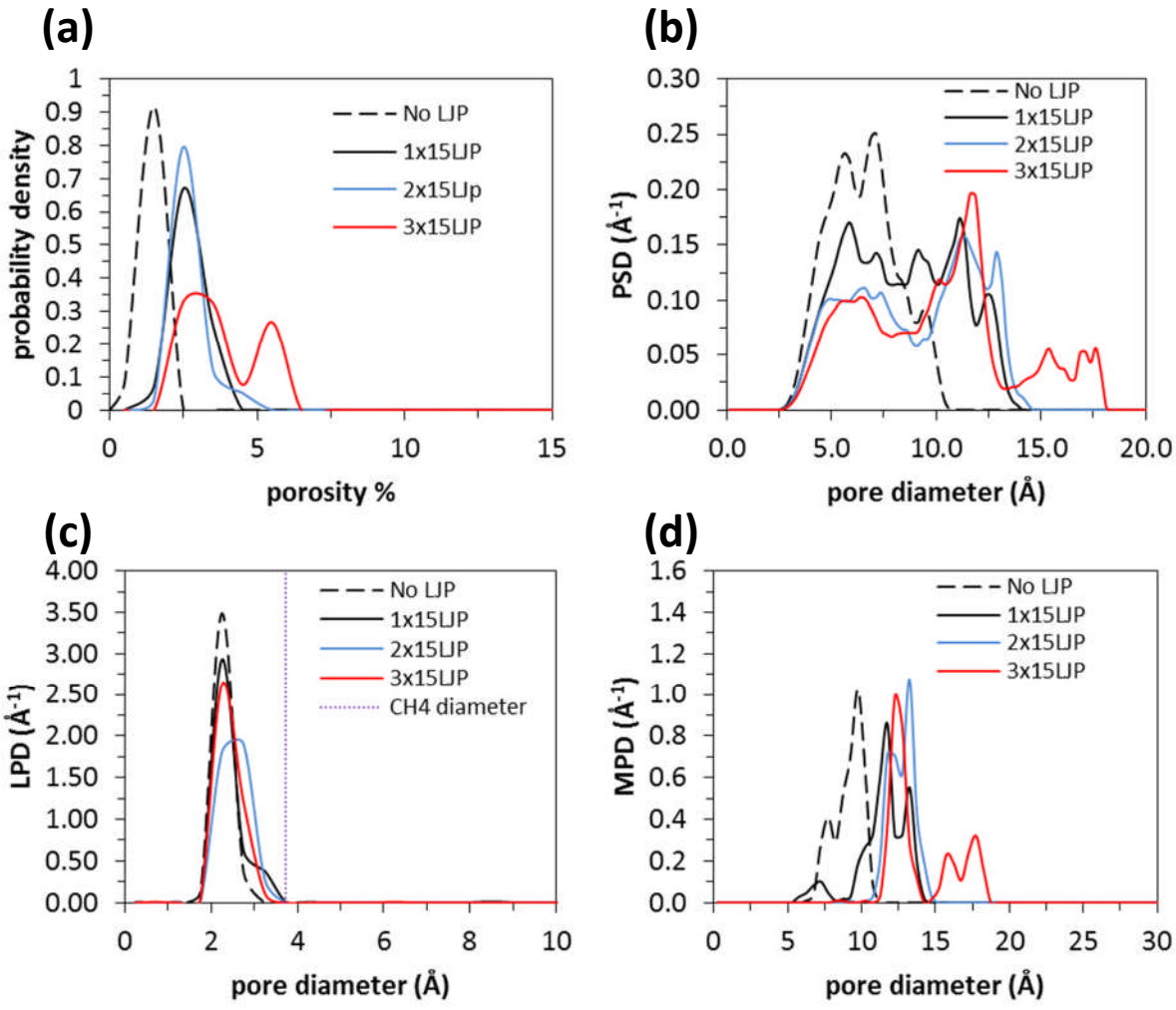


Figure 21: Accessible volume characteristics for 50 II-D bulk kerogen systems modeled using GAFF and constructed with the aid of different number of 15 Å LJ dummy particles. (a) probability density of porosity, (b) PSD (c) LPD distribution and (b) MPD.

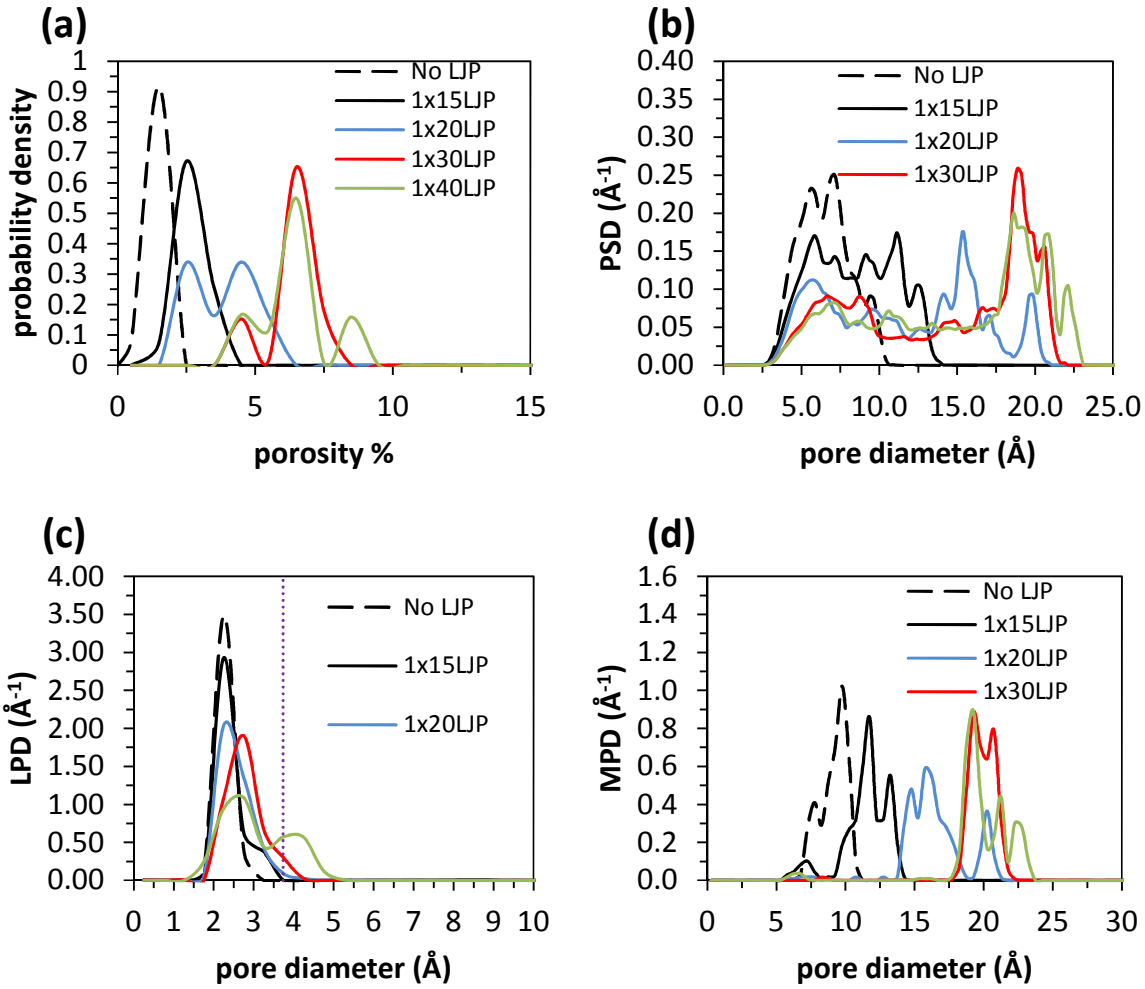


Figure 22: Accessible volume characteristics for 50 II-D bulk kerogen systems modeled using GAFF and constructed with the aid of varying size (15 to 40 Å) of a single LJ dummy particle. (a) probability distribution of the porosity, (b) PSD, (c) LPD and (d) MPD.

The effect of the variation of the LJ dummy particle diameter on the characteristics of the accessible volume is shown in Figure 22. Porosity can also be increased by increasing the LJ dummy particle size, as it is intuitively expected. While this is true in principle, there seems to be an upper limit in the achieved porosity. Increasing the size from 30Å to 40Å, results in no significant increase in the induced porosity (Figure 22a). In every case, no more than 10% porosity can be attained. Comparing to the number of LJ particles, variation of their size is much more effective in inducing porosity in the sense that equal or bigger porosity can be achieved with a small increase of the size of a single particle.

Furthermore, increasing the LJ dummy particle's diameter results in larger pores, as it is indicated by the PSDs of Figure 22b. The broadening trend of the PSD is not observed when the diameter of the LJ particle is increased from 30 to 40 Å, which is consistent with what is observed in the distribution of the % porosity. This suggests an upper limit in the pore sizes that we may aim for using one LJ particles with an upper limit approximately equal to 23 Å. There is also a clear effect of increasing the LJ particle size the MPD, which is translated towards bigger pores a trend, which also does not when LJ particle size increases from 30 to 40 Å (Figure 22d). The cease of the LJ particle size effect on pore diameter is a consequence of the size of the II-D molecule, whose diameter is approximately 28 Å in its gas phase

conformation and can be even smaller in the bulk. With respect to the LPD distribution, the size of the LJ particles has a much more significant effect compared to changing the number of LJ particles (Figure 22c). Limiting pore distributions become broader and less sharp, reaching values greater than the diameter of CH_4 (something which couldn't be achieved with variation of then number of LJ dummy particles) up to 5 Å. It is evident that the increase of LJ dummy particle size is a much more effective in controlling the major characteristics of porosity, being the only way that allows the creation of diffusion channels in the bulk (i.e. limiting pore diameters greater than the diameter of CH_4).

The influence the particles of size 20 and 30 Å have on the LPD is shown in Figure 23. It is clear that an increased number of larger particles can result in well-shaped, wide percolated pores that allow diffusion of large probes. A considerable number of pores with limiting diameters higher than the diameter of CH_4 and up to 10 Å are created.

Overall, no single model shares all types of micropores that are known to exist in kerogen. Despite this, the various combinations of LJ dummy particle numbers and sizes have made it possible to create structures of different porosity characteristics that are all present in kerogen. These characteristics are closely linked to phenomena that are of interest in modelling systems relevant to shale gas industry.

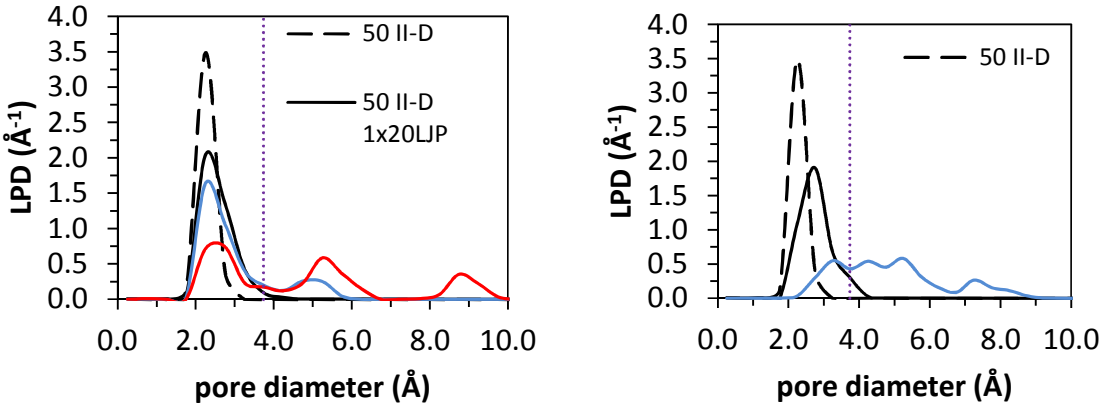
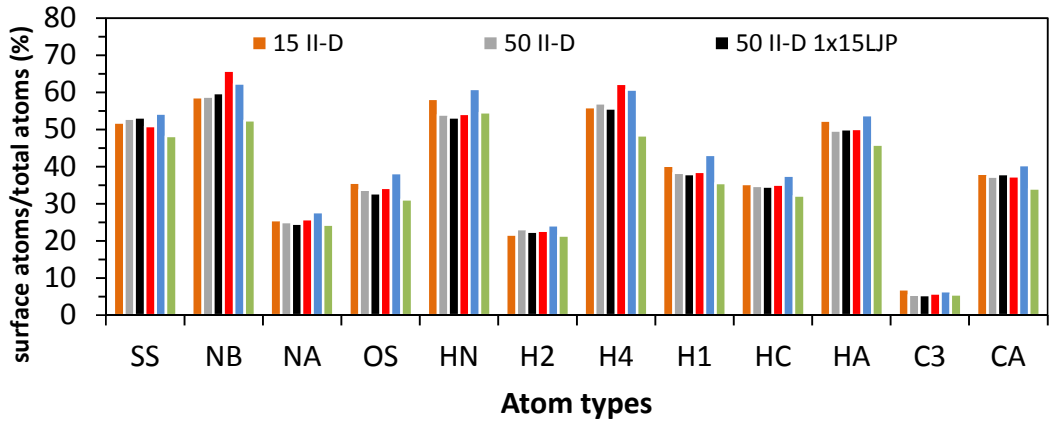


Figure 23: LPD for different number of LJ particles of size 20 Å (left) and 30 Å (right).

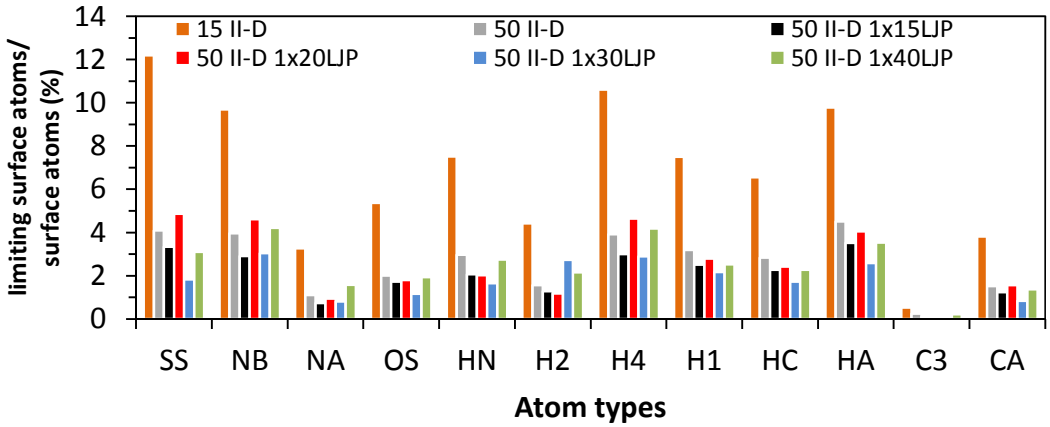
Comparison with experimental data

Although direct quantitative comparisons between the calculated quantities presented here and experimental results is not possible, all the models that have been generated are in general within or close to experimental range. The majority of the models give densities between 1.1 and 1.3 g cm^{-3}, values very close the commonly observed densities for type II kerogen [57]. PSDs are in accordance with the experimental pores sizes of micropores, having diameters of a few Å [13,65]. Ultimately, to the best of our knowledge, the only information on limiting pores is related to Barnett shale, where experiments have proven the existence of pore throats of less than 5 Å [12,13], a finding that corroborates our calculations.

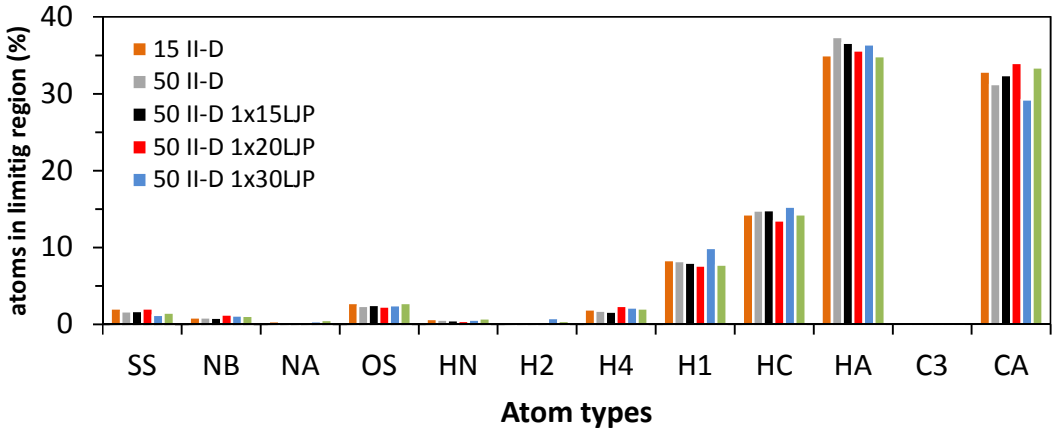
Pore surface analysis



(a)



(b)



(c)

Figure 24: GAFF atom type analysis of the surface of the pores. GAFF atom type analysis of the surface of the pores.

(a) Percentage of atoms of each type that form the surface with respect to the total number of atoms of this type in the whole structure, (b) Percentage of the surface atoms that are found at the limiting regions of the pore (limiting regions are defined as a radius of 3.73·1.2=4.476 Å around the pore). (c) Percentage of the atoms of the limiting regions that are of specific type.

All the previous analysis revealed that CH_4 accessible pores can be percolated only when a LJ dummy particle is used. Other simulation parameters such as the system size or the force field seem to have minor effect on the more important characteristics of the porosity. A considerable amount of percolated configuration has only be achieved only for the case of 3x20 Å, 2x30 Å and 1x40 Å LJ dummy particle sets. On the other hand, when probing the accessible volume with a 2 Å particle, all approaches with both force fields result in a large number of percolated configurations. For this reason, analysis of the diffusion paths is performed with probe diameter equal to 2 Å, so that pore characteristics of all procedures are acquired.

Figure 24a shows the % of each GAFF atom type (Figure 2) located at the percolated pore surface, as a fraction of the total number of this atom type in the configuration. A tendency is of specific atom types to lay at the pore surface instead of the bulk is observed. An interesting example is the amino hydrogens (atom type hn). In each configuration, approximately half are found at the surface of pores a very interesting finding given that this functional group may be involved to strong interactions such as hydrogen bonding. This finding may have important implications when studying the diffusion of polar molecules. The fraction of the atoms of the pore that are found at limiting diameter regions is shown in Figure 24b. We define as limiting regions the pore segments, which are accessible by probes 20 % larger than CH_4 i.e. of 4.476 Å radius. The reduction of bars in Figures 24(a,b) does not follow a specific trend. Finally, Figure 24c shows the fraction of each atom type in the limiting region, with respect to the total number of atoms in the limiting region. There is a measurable fraction of amino groups, indicated by the presence of hn atoms, while the limiting region environment is dominated by aromatic carbon and hydrogen ca and ha respectively.

4. Conclusions and future steps

A model molecule representative of mature type II kerogen was used for the construction of several bulk kerogen models comprising realistic porosity. Three force fields were employed and their effects were compared. Specifically, CVFF, the Dreiding force field with Gasteiger atomic charges, and GAFF with *ab-initio* derived point charges were utilized. Staged cooling of low density structures using *NPT* (1 atm) MD atomistic molecular simulations created dense structures at room temperature. For imposing and controlling porosity in a systematic manner, a varying number of LJ dummy particles of different sizes was used. The structures were characterized on the basis of their porosity, one of their most crucial characteristics for the study of diffusion of fluid relevant to shale gas industry. For this reason, a robust methodology based on Voronoi tessellation of the material was developed for amorphous materials and implemented in a generic code.

GAFF is found to produce denser structures than Dreiding and CVFF. Increasing the system size decreases the density, with a limiting value obtained for medium sizes structures. Including dummy particle(s) results in less dense structure. The density in most cases decreases both as function of the number and size of the LJ particles.

A specific force field selection has no significant effect to PSD and LPD. In all cases, the maximum pore sizes obtained are around 1nm (without the aid of LJ particles). The limiting pores obtained are around 2 Å well below the diameter of CH_4 whose diffusion we intend to study. Temperature has also negligible effect on PSD and LPD of 15 II-D structures suggesting limited mobility of the bulky II-D molecules even at temperatures as high as 600 K. Including

LJ dummy particles increases the porosity of the structures, and affects the pore sizes. Variation of size of LJ dummy particles is much more effective in inducing porosity. Variation of the size of LJ particle(s) results also in bigger pores with a limit around 23 Å for the biggest pore that can be achieved by this methodology as a consequence of the size of II-D molecule. Variation of size of the dummy particles can also lead to bigger limiting pores that fore certain big enough values can exceed the diameter of CH_4 and reach values that approach 10 Å. Analysis of the pore surface reveals a preference of some specific atom types to concentrate at the surface of pores instead of the bulk. Furthermore, a measurable amount of the only functional group that can be involved in strong directional interactions such as hydrogen bonding at the limiting regions.

Bulk and edge surface models for montmorillonite and muscovite clays are developed and using the *Supercell* code [52] and the newly developed ClayFF parameterization for clay edges [66]. A series of unconstrained MD simulations is started for the (001) interfaces of montmorillonite and muscovite with aqueous solutions of $SrCl_2$, $BaCl_2$, and $RaCl_2$ as proxy models of NORM. The parameters of Ra^{2+} cation are not available in the literature and are being developed based on the available calorimetric and structural data.

Future steps in the work to be conducted are:

- Generation of adsorption isotherms for various temperatures of pure substances such as CH_4, C_2H_8, CO_2, H20 and their mixtures, that are relevant to shale gas extraction;
- Study of diffusion of pure substances and their mixtures in kerogen ;
- Develop structural models for other shale rock minerals: quartz, kaolinite, calcite;
- Clay adsorption sites identified and characterized in the present simulations will be further probed by umbrella-sampling potential of mean force calculations to quantify the free energy of adsorption and cation exchange capacity for NORM cations (Sr^{2+}, Ba^{2+}, Ra^{2+}) with other common cations (Na^{+}, Ca^{2+});
- Construct combined clay/kerogen models of generalized shale rock for the molecular simulation study of preferential density distributions and mobility of relevant fluids;
- Identify correlations between the observed transport behavior of the various fluids and one or more porosity or other characteristics of the structures so that the observed behavior can be deeply understood and explained;
- Study different types of kerogen.

5. Publications resulting from the work described

1) L. Michalec and M. Lísal, Molecular Simulation of Shale Gas Adsorption onto Overmature Type II Model Kerogen with Control Microporosity, Molecular Physics 2017, DOI: 10.1080/00268976.2016.1243739.

2) Manolis Vasileiadis, Loukas D. Peristeras, Konstantinos D. Papavasileiou, & Ioannis G. Economou, “Modeling of bulk kerogen porosity: Methods for control and characterization”, in preparation.

3) B.F. Ngouana-Wakou, A.G.Kalinichev, Molecular dynamics simulation of divalent metal cations in montmorillonites: The effect of structural arrangements of isomorphic substitutions in smectites on the clay swelling, adsorption and mobility of interlayer and surface species, in preparation.

6. Bibliographical references

[1] T. Alexander, J. Baihly, C. Boyer, B. Clark, G. Waters, V. Jochen, J. Le Calvez, R. Lewis, C.K. Miller, J. Thaeler, B.E. Toelle, Shale Gas Revolution, Oilfield Rev. 23 (2011) 40–55.

[2] R.B. Jackson, A. Vengosh, T.H. Darrah, N.R. Warner, A. Down, R.J. Poreda, S.G. Osborn, K. Zhao, J.D. Karr, Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction, Proc. Natl. Acad. Sci. 110 (2013) 11250–11255. doi:10.1073/pnas.1221635110.

[3] C. Frohlich, M. Brunt, Two-year survey of earthquakes and injection/production wells in the Eagle Ford Shale, Texas, prior to the MW4.8 20 October 2011 earthquake, Earth Planet. Sci. Lett. 379 (2013) 56–63.

[4] A.W. Nelson, A.W. Knight, D. May, E.S. Eitrheim, M.K. Schultz, Naturally-Occurring Radioactive Materials (NORM) Associated with Unconventional Drilling for Shale Gas, in: D.L. Drogos (Ed.), ACS Symp. Ser., American Chemical Society, Washington, DC, 2015: pp. 89–128.

[5] M. Vandenbroucke, Kerogen: from Types to Models of Chemical Structure, Oil Gas Sci. Technol. 58 (2003) 243–269.

[6] B. P. Tissot and D. H. Welte, Petroleum Formation and Occurrence, Springer: Berlin, Germany, 1984.

[7] D. W. van Krevelen, Coal: Typology, Chemistry, Physics, Constitution, Elsevier: Amsterdam, The Netherlands, 1961.

[8] J. S. Seewald, Nature 426 (2003) 327-333.

[9] S. R. Kelemen, M. Afeworki, M. L. Gorbaty, M. Sansone, P. J. Kwiatek, C. C. Walters, H. Freund, M. Siskin, A. E. Bence, D. J. Curry, M. Solum, R. J. Pugmire, M. Vandenbroucke, M. Leblond, and F. Behar, Energy Fuels 21 (2007) 1548-1561.

[10] B.F. Ngouana-Wakou, AG Kalinichev: Structural arrangements of isomorphic substitutions in smectites: Molecular simulation of the swelling properties, interlayer structure, and dynamics of hydrated Cs-montmorillonite revisited with new clay models. Journal of Physical Chemistry C 118 (2014) 12758-73.

[11] N. Loganathan, A.O. Yazaydin, G.M. Bowers, A.G. Kalinichev, R.J. Kirkpatrick: Cation and water structure, dynamics, and energetics in smectite clays: A molecular dynamics study of Ca-hectorite. Journal of Physical Chemistry C 120 (2016) 12429-39.

[12] J.C. Palmer, K.E. Gubbins, Atomistic models for disordered nanoporous carbons using reactive force fields, Microporous Mesoporous Mater. 154 (2012) 24–37.

[13] C. Bousige, C.M. Ghimbeu, C. Vix-Guterl, A.E. Pomerantz, A. Suleimenova, G. Vaughan, G. Garbarino, M. Feygenson, C. Wildgruber, F.-J. Ulm, R.J.-M. Pellenq, B. Coasne, Realistic molecular model of kerogen's nanostructure, Nat. Mater. 15 (2016) 576–582.

[14] F. Behar, M. Vandenbroucke, Chemical modelling of kerogens, Org. Geochem. 11 (1987) 15–24.

[15] M. Siskin, C.G. Scouten, K.D. Rose, T. Aczel, S.G. Colgrove, R.E. Pabst, Detailed Structural Characterization of the Organic Material in Rundle Ramsay Crossing and Green River Oil Shales, in: C. Snape (Ed.), Compos. Geochem. Convers. Oil Shales, Springer Netherlands, Dordrecht, 1995: pp. 143–158.

[16] A.M. Orendt, I.S.O. Pimienta, S.R. Badu, M.S. Solum, R.J. Pugmire, J.C. Facelli, D.R. Locke, K.W. Chapman, P.J. Chupas, R.E. Winans, Three-Dimensional Structure of the Siskin Green River Oil Shale Kerogen Model: A Comparison between Calculated and Observed Properties, Energy Fuels. 27 (2013) 702–710.

[17] J. Collell, G. Galliero, F. Gouth, F. Montel, M. Pujol, P. Ungerer, and M. Yiannourakou, Micropor. Mesopor. Mat. 197 (2014) 194-203.

[18] J. Collell, P. Ungerer, G. Galliero, M. Yiannourakou, F. Montel, and M. Pujol, Energy Fuels 28 (2014) 7457-7466.

[19] P. Ungerer, J. Collell, M. Yiannourakou, Molecular Modeling of the Volumetric and Thermodynamic Properties of Kerogen: Influence of Organic Type and Maturity, Energy Fuels. 29 (2015) 91–105.

[20] P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, Proteins 4 (1988) 31-47.

[21] A.G. Kalinichev, P.P. Kumar, R.J. Kirkpatrick: Molecular dynamics computer simulations of the effects of hydrogen bonding on the properties of layered double hydroxides intercalated with organic acids. Philosophical Magazine 90 (2010) 2475 - 88.

[22] L. Sarkisov, A. Harrison, Computational structure characterisation tools in application to ordered and disordered porous materials, Mol. Simul. 37 (2011) 1248–1257.

[23] L. D. Gelb and K. E. Gubbins, Langmuir 15 (1999) 305-308.

[24] S.L. Mayo, B.D. Olafson, W.A. Goddard, DREIDING: a generic force field for molecular simulations, J. Phys. Chem. 94 (1990) 8897–8909.

[25] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, Tetrahedron. 36 (1980) 3219–3228.

[26] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174.

[27] M.Szczerba, A.G.Kalinichev, Intercalation of ethylene glycol in smectites: Several molecular simulation models verified by X-ray diffraction data. Clays and Clay Minerals, 64 (2016) 488-502.

[28] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.

[29] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, J. Phys. Chem. 98 (1994) 11623–11627.

[30] W.J. Hehre, R. Ditchfield, J.A. Pople, Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules, J. Chem. Phys. 56 (1972) 2257–2261.

[31] B.H. Besler, K.M. Merz, P.A. Kollman, Atomic charges derived from semiempirical methods, J. Comput. Chem. 11 (1990) 431–439.

[32] U.C. Singh, P.A. Kollman, An approach to computing electrostatic charges for molecules, J. Comput. Chem. 5 (1984) 129–145.

[33] C.I. Bayly, P. Cieplak, W. Cornell, P.A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, J. Phys. Chem. 97 (1993) 10269–10280.

[34] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, J. Mol. Graph. Model. 25 (2006) 247-260.

[35] D.A. Case, R.M. Betz, W. Botello-Smith, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, J.. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York, P.A. Kollman, AMBER 2016, University of California, San Francisco, 2016.

[36] MAPS platform. Current version is 4.0 | Scienomics, (n.d.). http://www.scienomics.com/products/molecular-modeling-platform (accessed September 22, 2016).

[37] D. Frenkel, B. Smit, Understanding molecular simulation: from algorithms to applications, 2nd ed, Academic Press, San Diego, 2002.

[38] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, J. Comput. Phys. 117 (1995) 1–19.

[39] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A. 31 (1985) 1695–1697.

[40] W. Shinoda, M. Shiga, M. Mikami, Rapid estimation of elastic constants by molecular dynamics simulation under constant stress, Phys. Rev. B. 69 (2004).

[41] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, J. Am. Chem. Soc. 114 (1992) 10024–10035.

[42] C.H. Rycroft, VORO++: A three-dimensional Voronoi cell library in C++, Chaos Interdiscip. J. Nonlinear Sci. 19 (2009) 041111.

[43] M.D. Foster, I. Rivin, M.M.J. Treacy, O. Delgado Friedrichs, A geometric solution to the largest-free-sphere problem in zeolite frameworks, Microporous Mesoporous Mater. 90 (2006) 32–38.

[44] M.G. Martin, J.I. Siepmann, Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes, J. Phys. Chem. B. 102 (1998) 2569–2577.

[45] L.R. Dodd, D.N. Theodorou, Analytical treatment of the volume and surface area of molecules formed by an arbitrary collection of unequal spheres intersected by planes, Mol. Phys. 72 (1991) 1313–1345.

[46] N. Loganathan, A.G. Kalinichev: Quantifying the Mechanisms of Site-Specific Ion Exchange at an Inhomogeneously Charged Surface: Case of Cs+/K+ on Hydrated Muscovite Mica. J. Phys. Chem. C (2017), DOI: 10.1021/acs.jpcc.6b13108.

[47] H. van Olphen, An introduction to Clay Colloid Chemistry: John Wiley and Sons, New York, 318pp, 1977.

[48] W. F. Bleam, G. J. Welhouse, M. A. Janowiak, Clays and Clay Minerals 41 (1993) 305-316.

[49] S. V. Churakov, J. Phys. Chem. B 110 (2006) 4135-4146.

[50] L. P. Lavikainen, J. T. Hirvi, S.Kasa, T. Schatz, T. A. Pakkanen, Theor Chem Acc 134 (2015) 112.

[51] V. A. Drits, S. Guggenheim, B.B. Zviagina, T. Kogure Clays and Clay Minerals 60 (2012) 574-587.

[52] K. Okhotnikov, T. Charpentier, S.Cadars, J. Cheminform. 8 (2016) 141-146.

[53] G. N. White, L. W. Zelany, Clays and Clay Minerals 36 (1988) 141-146.

[54] X.Liu, X.Lu, E.J.Meijer, R.Wang, H.Zhou, Geochim. Cosmochim. Acta 81 (2012) 56-68, 2012.

[55] A.G.Newton, K.D.Kwon, D-K.Cheong, Minerals 6 (2016), 25.

[56] S.V.Churakov, Geochim. Cosmochim. Acta 71 (2007) 1130-1144.

[57] K.S. Okiongbo, A.C. Aplin, S.R. Larter, Changes in Type II Kerogen Density as a Function of Maturity: Evidence from the Kimmeridge Clay Formation, Energy Fuels. 19 (2005) 2495–2499.

[58] M. E. Curtis, R. J. Ambrose, D. Energy, C. H. Sondergeld, and C. S. Rai, Soc. Petr. Eng. 2010, 137693.

[59] D. J. K. Ross and R. M. Bustin, Mar. Petr. Geol. 26 (2009) 916-927.

[60] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. 118 (1996) 11225-11236.

[61] J. G. Harris and K. H. Yung, J. Phys. Chem. 99 (1995) 12021-12024.

[62] H. Sui and J. Yao, J. Nat. Gas Sci. Eng. 31 (2016) 738-746.

[63] R. Heller and M. Zoback, J. Unconventional Oil Gas Resources 8 (2014) 14-24.

[64] K. Falk, R. Pellenq, F.-J. Ulm, and B. Coasne, Energy Fuels 29 (2015) 7889-7896.

[65] C.R. Clarkson, N. Solano, R.M. Bustin, A.M.M. Bustin, G.R.L. Chalmers, L. He, Y.B. Melnichenko, A.P. Radliński, T.P. Blach, Pore structure characterization of North American shale gas reservoirs using USANS/SANS, gas adsorption, and mercury intrusion, Fuel. 103 (2013) 606–616.

[66] M.Pouvreau, J.A.Greathouse, R.T.Cygan, A.G.Kalinichev, Structure and dynamics of hydrated gibbsite and brucite edge surfaces: DFT results and further development of the ClayFF classical force field with metal O-H angle bending terms. J. Phys.Chem.C., (2017) submitted.