



## Full Length Article

# Construction of multidimensional structure model of the Mesoproterozoic Xiamaling shale kerogen, northern North China

Chengbo Guo<sup>a</sup>, Meijun Li<sup>a,b,\*</sup>, Xiaoqiang Liu<sup>b</sup>, Hong Xiao<sup>a</sup>, Qingyong Luo<sup>a</sup>, Qiuya Han<sup>a</sup>, Wenke Li<sup>c</sup>, Junhao Ren<sup>a</sup>

<sup>a</sup> National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing 102249, China

<sup>b</sup> College of Petroleum, China University of Petroleum (Beijing) at Karamay, Xinjiang 834000, China

<sup>c</sup> College of Chemistry and Environmental Engineering, Sichuan University of Science and Engineering, Zigong 643000, China

## ARTICLE INFO

## Keywords:

Kerogen

Structural characteristics

Molecular model

<sup>13</sup>C NMR

Molecular dynamics simulation

## ABSTRACT

In the southeastern Xuanlong sag of Northern China, the organic-rich black shale of the Mesoproterozoic Xiamaling Formation holds significant potential as a source rock. Through advanced analytical techniques including elemental analysis, <sup>13</sup>C NMR, XPS, FTIR, and XRD, the elemental composition and chemical structure were thoroughly characterized, and the molecular structure of the Xiamaling shale kerogen was established. The <sup>13</sup>C NMR analysis revealed the carbon composition of kerogen, indicating 75.19 % aliphatic carbon, 22.06 % aromatic carbon, and 2.75 % carbonyl carbon, with quantitative support from FTIR and XRD spectra. The dominant methylene carbons exhibited a low average methylene chain length (Cn) of 3.87, suggesting that they are primarily short-branched or alicyclic compounds rather than long straight chains. This conclusion was verified by the absence of a sharp peak at a wavenumber of 719.95 cm<sup>-1</sup> in the FTIR spectrum. Oxygen-containing functional groups observed in <sup>13</sup>C NMR spectra were also identified in FTIR spectra at a wavenumber range of 3500–1000 cm<sup>-1</sup>. XPS analysis indicated that nitrogen-containing groups were amino, quaternary, pyridine, and primary pyrrole groups, and sulfur existed as sulfoxide, aromatic sulfur, aliphatic sulfur, and sulfone. Finally, a relatively rational 2D molecular structure with the chemical formula of C<sub>278</sub>H<sub>346</sub>N<sub>6</sub>O<sub>26</sub>S<sub>2</sub> was established by comparing the predicted and experimental <sup>13</sup>C NMR spectra. Using MD simulation, five optimized 2D molecules were combined to construct a robust 3D kerogen model with a physical density of 1.067 g/cm<sup>3</sup>. Comparatively, the aliphatic structure of the Mesoproterozoic Xiamaling kerogen model consists of some monocyclic alkanes and polycyclic saturated structures, resembling Paleozoic kerogen models but differing from Mesozoic counterparts. The constructed kerogen model addresses the gap in the Precambrian kerogen modeling and sets the stage for in-depth exploration of pyrolysis mechanisms in the future.

## 1. Introduction

Black shale plays a pivotal role as a source rock, encompassing a wealth of mineral matrix and organic content. Kerogen, as the primary product of organic matter, drives the production of shale oil and gas through thermal and chemical processes [1–3]. Comprising predominantly C, H, and O, with trace amounts of S and N elements, kerogen is categorized into Type I, II, and III based on the H/C and O/C ratios [2]. The potential for shale oil and gas resources hinges chiefly on the chemical structure within kerogen [4–7].

In recent decades, a range of physical and chemical techniques, including Pyrolysis-gas chromatography-mass spectrometry (PY-GC-

MS), solid state <sup>13</sup>C nuclear magnetic resonance (<sup>13</sup>C NMR), X-ray photoelectron spectroscopy (XPS), Fourier transform infrared spectroscopy (FTIR), and X-ray diffraction (XRD), have been employed to investigate the structural attributes of kerogens from different regions. Using acquired elemental data, functional group information, and structural parameters, distinct two-dimensional (2D) and three-dimensional (3D) kerogen structure models have been developed.

One notable kerogen structural model is the Tertiary Green River oil shale kerogen model, denoted by the chemical formula of C<sub>645</sub>H<sub>1017</sub>N<sub>19</sub>O<sub>17</sub>S<sub>4</sub> [8]. This 2D model was developed based on insights into structural attributes, hydrocarbon content, and functional groups found in both the solid oil shale and its shale oil. These characteristics

\* Corresponding author at: National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing 102249, China.

E-mail address: [meijunli@cup.edu.cn](mailto:meijunli@cup.edu.cn) (M. Li).

<https://doi.org/10.1016/j.fuel.2024.131572>

Received 12 August 2023; Received in revised form 13 March 2024; Accepted 20 March 2024

0016-2361/© 2024 Elsevier Ltd. All rights reserved.

were assessed through techniques such as  $^{13}\text{C}$  NMR,  $^{29}\text{Si}$  NMR, and MS analyses [8]. Furthermore, the corresponding 3D model of the Tertiary Green River kerogen is established via ab initio and molecular mechanics calculations [9]. The molecular structure of kerogen exhibits variation corresponding to its type and maturity. As a result, prior research has established kerogen molecular models across diverse types and levels of maturity using a range of experimental techniques. For instance, Behar and Vandembroucke developed three types of kerogen models representing three distinct evolution stages (early diagenesis, early catagenesis, and late catagenesis) [10]. Ungerer et al. [11] utilized insights from ultimate analysis,  $^{13}\text{C}$  NMR, and XPS analyses [12] to develop six distinct kerogen models, each characterized by different kerogen types and thermal maturity. These models, encompassing atomic ratios, functional group composition, density, and heat capacity, closely align with empirical data [11]. Furthermore, these models have been instrumental in probing the adsorption mechanisms of shale gas within kerogen [13], as well as investigating the influences of maturity [14], water content [14,15], and salinity [15] on the adsorption process.

As oil shale research advances in China, researchers have established kerogen models of oil shale in different regions. This endeavor aims to delve deeper into the distinctive structural disparities among kerogens and their ensuing hydrocarbon generation traits. Ru et al. [16] proposed plausible 2D and 3D structures for Huadian kerogen, validated through  $^{13}\text{C}$  NMR spectrum simulation and physical density analysis. They further scrutinized reactive sites, deducing that the thermal maturation process of the molecular structure of kerogen could be categorized into three distinct stages [16]. Different from the work of Ru et al. [16], Guan et al. [17] and Wang et al. [18] respectively developed 12 isomeric 2D models for Huadian and Yaojie kerogens, varying the grid count. Their investigation extended to generating 3D models using PY-GC-MS,  $^{13}\text{C}$  NMR, XPS, FTIR, and XRD techniques. Finally, rigorous testing involved comparing predicted and experimental  $^{13}\text{C}$  NMR spectra, along with assessing total energy minima for the Huadian and Yaojie kerogen models. As a result, they concluded that optimal kerogen model design entails minimizing total energy through the thoughtful selection of grid counts [17,18]. Utilizing the constructed 3D model of Chang 7 oil shale kerogen, Wu et al. [19] respectively developed two expansive systems featuring 3 and 5 kerogen models to investigate the pyrolysis mechanism across a temperature range of 2000–3000 K via reactive force field molecular dynamics. In contrast, Zhao et al. [20,21] directed their efforts towards analyzing in-situ pyrolysates within the extensive Fushun

kerogen system at six temperatures from 1600 to 2400 K.

In this study, kerogen was isolated from the Xiamaling shale in the Xiahuayuan area. To achieve the elemental composition and structural features of kerogen, elemental analysis,  $^{13}\text{C}$  NMR, XPS, FTIR, and XRD techniques are performed on the Xiamaling kerogen samples. Based on the visualization data provided by these experiments, a 2D molecular structure model of the Xiamaling kerogen is reconstructed by comparing the predicted and experimental  $^{13}\text{C}$  NMR spectra. Subsequently, a 3D kerogen model is obtained through MD simulation. In addition, to better understand the differences in molecular structure characteristics of the Precambrian, Paleozoic, and Mesozoic kerogens, four kerogen models from Paleozoic and Mesozoic are selected for comparison.

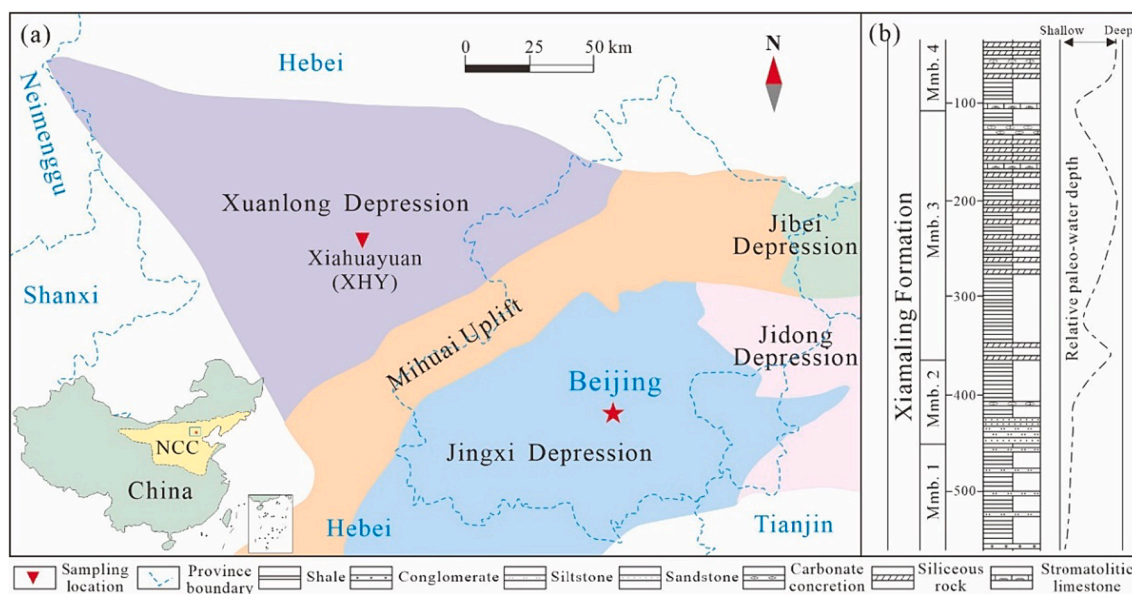
## 2. Samples and experiments

### 2.1. Sample

Three marine shale outcrop samples were collected from the Member 3 of Xiamaling Formation (XML3) located in the Xiahuayuan area of the southeastern Xuanlong Depression, situated in the northern part of North China (Fig. 1). All samples were obtained following the removal of overlying vegetation and oxidized surface rocks (Fig. 2a), ensuring their freshness and absence of ferric oxide stains in their surface (Fig. 2c–d–) [22]. It is remarkable that samples exhibited high ignitability due to their exceptionally high organic matter content (Fig. 2b) [22]. Before kerogen preparation, the basic geochemical properties of the shale samples were obtained through total organic carbon analysis and Rock-Eval pyrolysis (Table 1). The shale samples underwent fine grinding and subsequent demineralization to yield kerogen samples following the standard procedure outlined by Durand and Nicaise [23].

### 2.2. Elemental analysis

The element composition of the kerogen samples was assessed utilizing both elemental analyzer and X-ray fluorescence (XRF) spectrometer. The element analyzer determined the percentage of carbon (C), hydrogen (H), oxygen (O), nitrogen (N), and total sulfur (S) within the samples. The estimation of pyritic sulfur in the kerogens relied on the iron content, assuming that all iron in the isolated kerogens existed in the form of pyrite, as indicated by XRF analysis. The organic sulfur content ( $S_{\text{org}}$ ) within the kerogens was determined by subtracting the



**Fig. 1.** (a) Geographic map showing the sampling location in the Xiahuayuan area of the southeastern Xuanlong Depression, northern North China; (b) Detailed schematic stratigraphic column of the Xiamaling Formation in the Xiahuayuan area (modified from [22]).



**Fig. 2.** Overview of the characteristics of black shales from the Xiamaling Formation. (a) The outcrop of the Xiamaling Formation black shales in the Xiahuayuan area (XHY), Hebei province. The vegetation and oxidized surface layers were removed using a pickaxe, exposing fresh rock surfaces; (b) Ignition of the black shales; (c-d) The outcrop samples of black shales from the XHY section (modified from [22]).

**Table 1**

Total organic carbon content and Rock-Eval pyrolysis results of black shales studied in the Xiamaling Formation.

| Sample | Lithology   | TOC (wt %) | S <sub>1</sub> (mg/g) | S <sub>2</sub> (mg/g) | S <sub>1</sub> + S <sub>2</sub> (mg/g) | T <sub>max</sub> (°C) | HI (mg/g) |
|--------|-------------|------------|-----------------------|-----------------------|--|-----------------------|-----------|
| XML3-1 | Black shale | 6.43       | 0.43                  | 25.24                 | 25.67                                  | 434                   | 392.53    |
| XML3-2 | Black shale | 12.10      | 1.35                  | 64.67                 | 66.02                                  | 438                   | 534.46    |
| XML3-4 | Black shale | 7.36       | 0.59                  | 35.61                 | 36.20                                  | 442                   | 483.83    |

pyrite sulfur from the overall sulfur content [2,24].

### 2.3. Structural characterization

The <sup>13</sup>C Nuclear Magnetic Resonance (<sup>13</sup>C NMR) analyses were conducted using a Bruker AVANCE III spectrometer, operating at 600.19 MHz for 1H NMR and 150.92 MHz for <sup>13</sup>C NMR, equipped with a 4-mm double-resonance probe head and a ZrO<sub>2</sub> rotor. Spectra were recorded in CDCl<sub>3</sub> solutions, utilizing tetramethylsilane (TMS) as an internal standard. The spinning speed was set at 5,000 Hz, and the CP-TOSS/MAS (TOSS: total side band suppression; MAS: magic angle spinning; CP: Cross polarization) method was employed to investigate sample structures and restrain sidebands. The spectral width was 100,000 Hz, and a recycle delay time of 1 s was utilized with a total of 4,000 scans. It is essential to acknowledge that the <sup>13</sup>C CP/TOSS NMR technique inherently presents limitations in estimating non-protonated aromatic carbons. Based on the previously reported chemical shift [18,25–28] and the ratio of Gaussian to Lorentz distribution, the <sup>13</sup>C NMR spectrum was fitted by Peakfit v4.12 software. This analysis facilitated the determination of the relative proportions of different carbon types present in the Xiamaling kerogen. The results of the <sup>13</sup>C NMR analysis are outlined in Table 3.

X-ray Photoelectron Spectroscopy (XPS) spectra were acquired employing a K-Alpha X-ray photoelectron spectrometer (Thermo Fisher Scientific, UK), utilizing a monochromatic Al K $\alpha$  X-ray source (excitation energy = 1468.6 eV). The X-ray spot size was set to 400  $\mu$ m, and spectra

were collected within a range of 0–1350 eV, employing a pass energy of 100 eV for broad scans and 30 eV for individual elements. The binding energies were calibrated with respect to the C 1s signal at 284.8 eV. To quantify the proportions of distinct heteroatom functional groups, spectra for C 1s (carbon 1s), N 1s (nitrogen 1s), and S 2p (sulfur 2p) were subjected to fitting using Avantage software. The form of heteroatoms and the respective binding energy ranges were derived from previous literature sources [18,25,27–31], as summarized in Table 5.

Infrared (IR) spectroscopy analysis was conducted using a Bruker Vertex 70 V spectrometer, employing the potassium bromide (KBr) pellet method [32,33]. To prepare the pellets, the kerogen sample (1 mg) was first finely ground to a size of approximately 200 mesh and subsequently mixed with KBr (500 mg). The sample presser was then utilized to compress the resultant blend into discs measuring 10 mm in diameter, applying a pressure of 13 MPa for 5 min. The FTIR spectrum was obtained with a resolution of 4 cm<sup>−1</sup>, spanning the wavenumber range of 4000–400 cm<sup>−1</sup>.

The X-ray diffraction (XRD) spectra of the kerogen samples were gained through an Olympus Innov-x BTX XRD, utilizing Cu K $\alpha$  radiation with a Ni filter (operating at 40 kV and 40 mA). The diffraction patterns were captured within a range spanning from 5° to 55° 2 $\theta$ , employing a scanning rate of 3° 2 $\theta$  min<sup>−1</sup>.

### 2.4. Molecular dynamics calculation method

The five optimized structural models of the Xiamaling kerogen were



placed into a periodic box using the Amorphous Cell module in Materials Studio 2017 software. The resulting system contained a total of 3290 atoms. To avoid overlap between functional groups and aromatic rings, the initial cell was set at a low density of  $0.3 \text{ g/cm}^3$  [34]. Finally, to gain reliable density, a series of molecular dynamics simulations were implemented by way of the Forcite module, adopting the simulation methods used by Liu et al. [34]. The three steps are as follows: (1) The geometry optimization was performed to minimize the energy of the cell and achieve reasonable cell parameters with the detailed parameter setting: Quality, Ultra-fine; Algorithm, Smart; Energy,  $2.0\text{e-}5$ ; Max iterations, 5000,000; Forcefield, COMPASS [34]. (2) An anneal dynamics calculation was executed to explore the optimal conformational space with the following parameters: Quality, Ultra-fine; Initial temperature, 300 K; Mid-cycle temperature, 700 K; Dynamics steps per ramp, 20,000; Heating ramps per cycle, 5; Annealing cycles, 5; Total number of steps, 5000; Forcefield, COMPASS [34]. (3) A series of MD calculations were carried out to ensure the reliability of the model density with the following parameters: Ensemble, NPT; Quality, Ultra-fine; Temperature, 298.15 K; Pressure,  $1.0\text{e-}4 \text{ GPa}$ ; Time step, 1 fs; Total simulation time, 1200 ps; Thermostat, Nose; Barostat, Berendsen; Forcefield, COMPASS [34].

### 3. Results and discussion

#### 3.1. Basic geochemical properties of source rocks

The bulk geochemical composition of the Xiamaling shales is presented in Table 1. The analyzed Xiamaling black shales have very high total organic carbon (TOC) contents in the range of 6.43–12.10 wt% with a mean value of 8.63 wt%. The quantity of thermo-vaporized free hydrocarbons ( $S_1$ , mg/g) during pyrolysis up to  $200^\circ\text{C}$  ranges from 0.43 to 1.35 mg/g, with an average of 0.79 mg/g. Meanwhile, the average content of residual hydrocarbons ( $S_2$ , mg/g) released from thermal cracking of mature organic matter at pyrolysis temperatures up to  $650^\circ\text{C}$ , is 41.84 mg/g, varying from 25.24 to 64.67 mg/g. Genetic potential ( $S_1 + S_2$ , mg/g) is in the range of 25.67–66.02 mg/g with a high average value of 42.63 mg/g. These results show that the three shale samples of the Xiamaling Formation belong to the excellent source rocks [22]. The temperature ( $T_{\text{max}}$ ,  $^\circ\text{C}$ ) at which the  $S_2$  peak reaches its maximum displays low values within  $434$  to  $442^\circ\text{C}$ , with an average of  $438^\circ\text{C}$ . Furthermore, the hydrogen index (HI,  $\text{HI} = (100 \times S_2)/\text{TOC}$ , mg/g) contents span from 392.53 to 534.46 mg/g (average = 470.27 mg/g). The relationship between HI and  $T_{\text{max}}$  illustrating kerogen type

classification [35] for the Xiamaling shales is depicted in Fig. 3a.

#### 3.2. Element composition of kerogen

A comprehensive overview of the results from elemental analysis is presented in Table 2. The H/C and O/C atomic ratios range from 1.080 to 1.240 (average = 1.163) and 0.031–0.062 (average = 0.042), respectively. The N/C atomic ratio varies from 0.016 to 0.018 with a mean of 0.017. The estimation of organic sulfur from total sulfur revealed that the XML3-1 and XML3-4 kerogen samples contain 0.51 wt % and 0.48 wt% organic sulfur, respectively, while the XML3-2 sample lacks organic sulfur. Therefore, according to the sulfur content of the three samples, it can be inferred that there is a small amount of organic sulfur in the kerogen structure of the Member 3 of the Xiamaling Formation. Fig. 3b illustrates the H/C versus O/C relationship for kerogen type classification [36]. In the cross-plots of HI versus  $T_{\text{max}}$  and H/C versus O/C, the three Xiamaling Formation sample points are distributed in the areas of type II<sub>1</sub>–II<sub>2</sub> (Fig. 3a) and type I–II<sub>1</sub> kerogen (Fig. 3b), respectively. Overall, the organic matter (OM) in the Xiamaling shales is categorized as type II kerogen with low maturity, which is consistent with previous reports [22,37].

#### 3.3. Structural features

##### 3.3.1. $^{13}\text{C}$ NMR analysis

This study focused on characterizing the carbon skeleton structure within the Xiamaling kerogen. To achieve this, kerogen samples XML3-1, XML3-2, and XML3-4 were chosen for solid state  $^{13}\text{C}$  NMR spectroscopy analysis. The NMR spectra of kerogen typically encompass three prominent categories: aliphatic carbon band ( $0$ – $90 \text{ ppm}$ ), aromatic carbon band ( $90$ – $165 \text{ ppm}$ ), and carbonyl carbon band ( $165$ – $220 \text{ ppm}$ ) [12,16,25,38]. The  $^{13}\text{C}$  NMR spectra of the Xiamaling kerogen samples are presented in Fig. 4. Notably, in comparison to the aromatic carbon and carbonyl carbon bands, the aliphatic carbon band significantly dominates the  $^{13}\text{C}$  NMR spectra. This observation underscores the prominence of aliphatic carbon as the principal carbon skeletal structure within the Xiamaling kerogen.

To ascertain the relative proportions of various carbon types within the chemical structure, a curve fitting analysis was conducted on the  $^{13}\text{C}$  NMR spectrum of the XML3-1 kerogen sample (depicted in Fig. 5). The relative areas of the fitted peaks are listed in Table 3. The carbon structure of kerogen in the Xiamaling Formation is characterized by a notably high aliphatic carbon ratio, amounting to 75.19 %. Notably, the

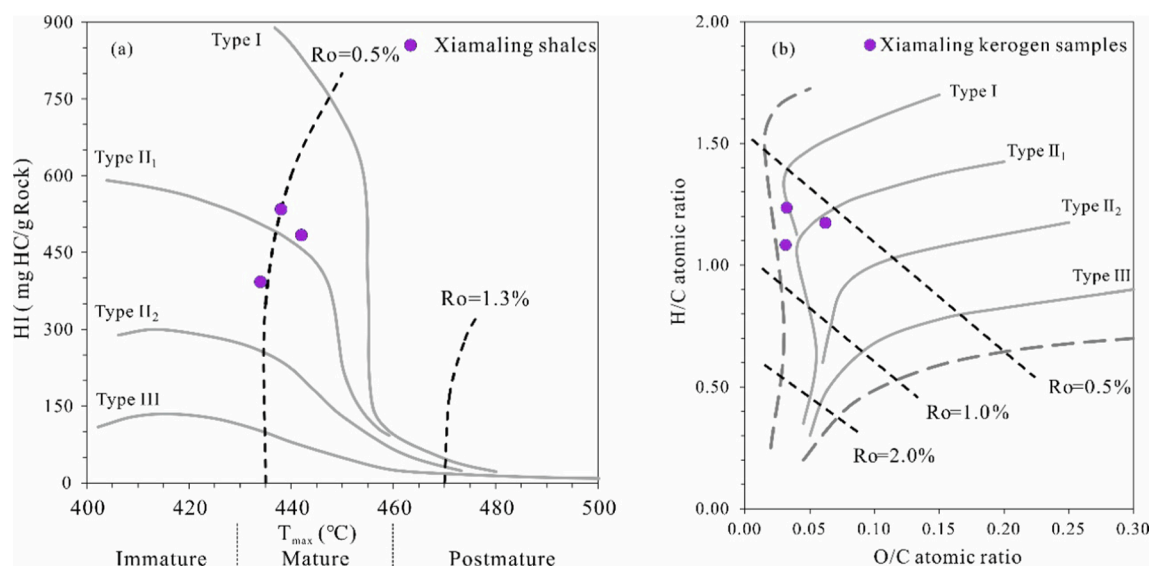


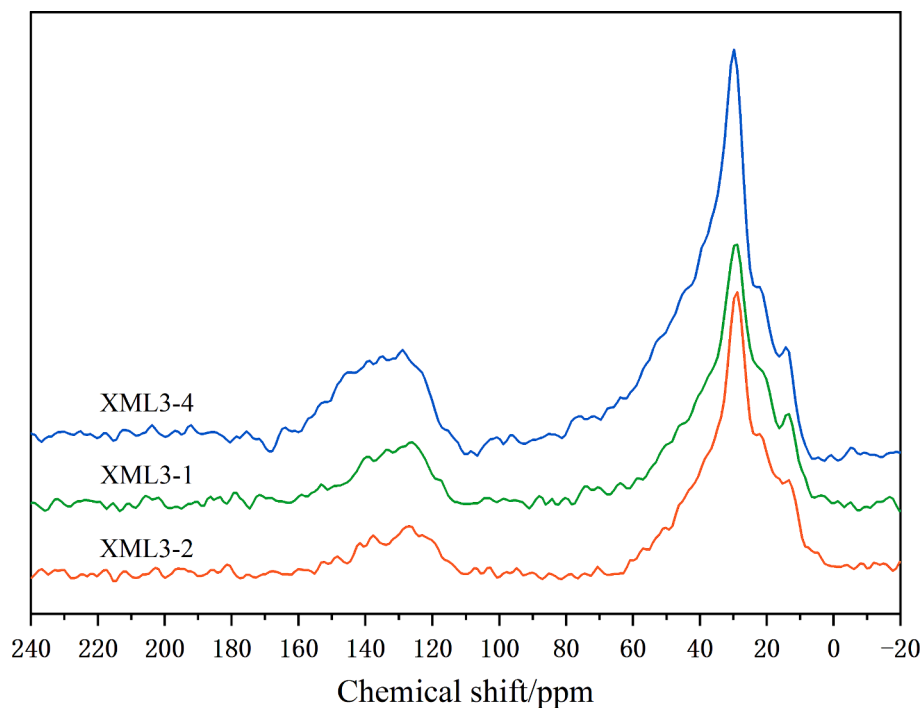
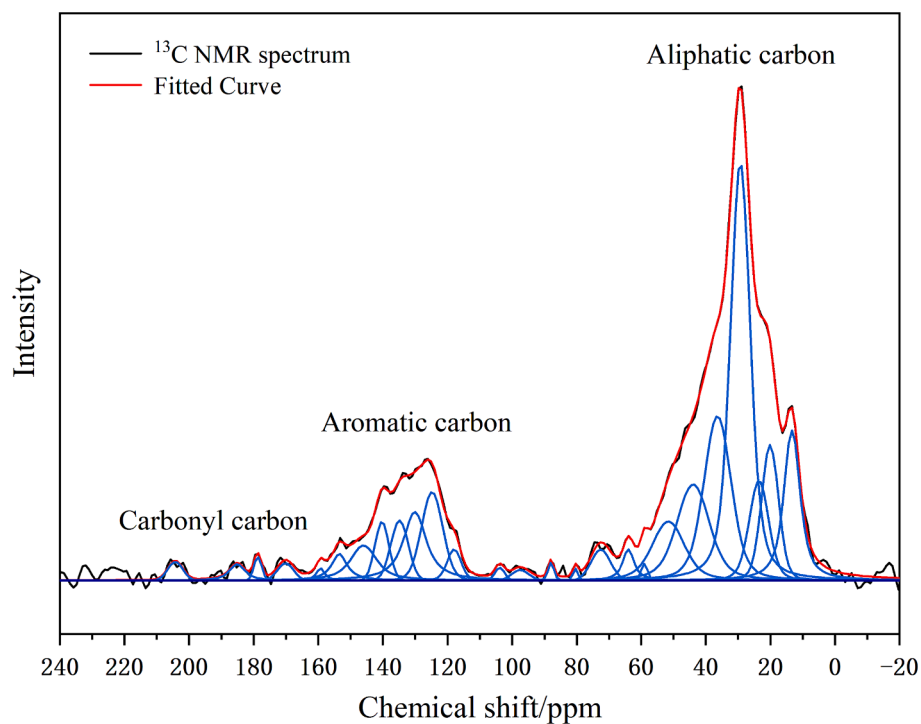
Fig. 3. (a) Relationship between HI and  $T_{\text{max}}$ ; (b) H/C versus O/C, displaying the kerogen types and thermal maturity of the analyzed Xiamaling shale samples.



**Table 2**

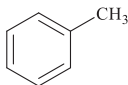
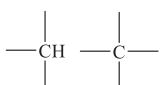
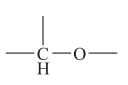
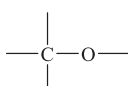
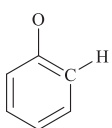
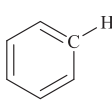
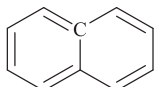
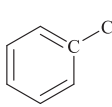
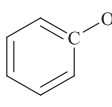
Elemental composition of the Xiamaling kerogen samples.

| Sample | C (wt%) | H (wt%) | O (wt%) | N (wt%) | S <sub>total</sub> (wt%) | Fe (wt%) | S <sub>org</sub> (wt%) | H/C   | O/C   | N/C   | S/C   |
|--------|---------|---------|---------|---------|--------------------------|----------|------------------------|-------|-------|-------|-------|
| XML3-1 | 52.14   | 5.10    | 4.31    | 1.07    | 1.34                     | 1.45     | 0.51                   | 1.170 | 0.062 | 0.018 | 0.007 |
| XML3-2 | 50.38   | 5.19    | 2.15    | 0.94    | 2.93                     | 5.79     | —                      | 1.240 | 0.032 | 0.016 | —     |
| XML3-4 | 67.60   | 6.10    | 2.82    | 1.37    | 1.04                     | 0.97     | 0.48                   | 1.080 | 0.031 | 0.017 | 0.005 |

**Fig. 4.**  $^{13}\text{C}$  NMR spectra of the Xiamaling shale samples, showing the carbon skeleton structure of the embedded kerogen.**Fig. 5.** The experimental  $^{13}\text{C}$  NMR spectrum and fitting curve, demonstrating the aliphatic carbon band (0–90 ppm), aromatic carbon band (90–165 ppm), and carbonyl carbon band (165–220 ppm) of the XML3-1 kerogen.

**Table 3**

Carbon types and their relative area in the XML3-1 kerogen.

| Carbon distribution | Carbon types                 | Symbol        | Structure  | Chemical shift/ppm | Relative area/% |
|---------------------|------------------------------|---------------|--|--------------------|-----------------|
| Aliphatic (75.19 %) | Aliphatic methyl             | $f_{al}^M$    | $—CH_3$  | 13.32              | 8.06            |
|                     | Aromatic methyl              | $f_{al}^A$    |   | 20.11              | 6.21            |
|                     | Amet-methylene               | $f_{al}^{AM}$ | $—\overset{H_2}{\underset{ }{C}}—\overset{H_2}{\underset{ }{C}}—CH_3$  | 23.54              | 5.39            |
|                     | Methylene                    | $f_{al}^H$    | $—\overset{H_2}{\underset{ }{C}}—\overset{H_2}{\underset{ }{C}}—\overset{H_2}{\underset{ }{C}}—$   | 29.30              | 24.58           |
|                     | Methine, quaternary carbon   | $f_{al}^D$    |   | 36–50              | 21.83           |
|                     | Oxy-aliphatic                | $f_{al}^O$    | $H_3C—O—\overset{H_2}{\underset{ }{C}}—O—$<br><br> | 50–90              | 9.12            |
| Aromatic (22.06 %)  | <i>ortho</i> -oxy-protonated | $f_{ar}^{O1}$ |    | 95–120             | 1.95            |
|                     | Protonated                   | $f_{ar}^H$    |   | 124.78             | 4.86            |
|                     | Bridgehead                   | $f_{ar}^B$    |   | 130.11             | 5.59            |
|                     | Branched                     | $f_{ar}^C$    |   | 135–145            | 7.74            |
|                     | Oxy-aromatic                 | $f_{ar}^{O2}$ |   | 150–160            | 1.93            |
| Carbonyl (2.75 %)   | Carboxyl                     | $f_a^C$       | $—COOH/R$  | 165–190            | 2.04            |
|                     | Carbonyl                     | $f_a^O$       | $\begin{array}{c} O \\    \\ —C— \end{array}$  | 190–220            | 0.71            |

prominent peak spanning 23–35 ppm in the aliphatic band (0–90 ppm in Fig. 5) signifies that the aliphatic structure in Xiamaling kerogen is dominated by methylene carbon. The sharp peak at 13.32 ppm is associated with methyl carbon. In the aliphatic region (0–90 ppm in Fig. 5), the second most significant peak, spanning 36–50 ppm, corresponds to methine and quaternary carbon, contributing to 21.83 % (Table 3), which is close to the proportion of methylene carbon. As shown in Table 3, the Xiamaling kerogen exhibits a relatively low aromaticity ( $f_{ar}$ )

value of 22.06 %. This value is nearly identical with the aromaticity observed in the Cretaceous Nengjiang Formation kerogen within the Songliao Basin [39]. The aromaticity of 22.06 % indicates the presence of approximately 22 aromatic carbon atoms per 100 carbon atoms in the Xiamaling kerogen structure, comprising either 4 benzene rings or 2 naphthalene structures. In the aromatic band (90–165 ppm in Fig. 5), branched aromatic carbon is the predominant component, constituting 7.74 %, followed by bridged aromatic carbon at 5.59 % (Table 3).

Furthermore, the result of  $^{13}\text{C}$  NMR analysis (Table 3) reveals that the carboxyl carbon content (2.04 %) exceeds that of carbonyl carbon (0.71 %) in the kerogen structure of the Xiamaling Formation.

Significant structural parameters about the carbon skeleton within the Xiamaling kerogen were computed employing the information from Table 3. These calculated parameters have been presented in Table 4. The average methylene chain length (Cn) is 3.87, indicating a prevalence of short-branched chains or alicyclic structures in the kerogen, as opposed to long, straight chains. Furthermore, the branched degree of the aliphatic chain (BI) is 0.29, indicating a highly branched structure in Xiamaling kerogen. The substitutive degree of the aromatic ring ( $\delta$ ) is 0.44, suggesting approximately 3 carbons in each aromatic ring undergo adjacent group substitution. This substitution pattern aligns with findings in Yaojie and Fushun type II kerogens [18,28]. The aromatic cluster size ( $X_{BP}$ ) is an important structural parameter that reflects the condensation level among aromatic groups within kerogen structures [12,18,20,26,38]. The calculated  $X_{BP}$  value for the Xiamaling kerogen is 0.34, which is close to the  $X_{BP}$  value of indole (0.33) but lower than that of anthracene (0.40). This suggests that the aromatic compounds in the Xiamaling kerogen can be represented by structures containing 2–3 aromatic rings, indicating a relatively moderate level of condensation.

### 3.3.2. XPS analysis

In this section, the XPS technique was employed to characterize the various forms of oxygen (O), nitrogen (N), and sulfur (S), along with their respective content within oil shale, kerogen, and coal [17,18,20,25,28,38,40]. The XPS spectra of XML3-1, XML3-2, and XML3-4 kerogen samples are shown in Fig. 6, illustrating the elemental peaks present in the three kerogen samples and the corresponding peak intensity. Although there are slight variations in the O 1s and C 1s peaks, these spectra exhibit similarities in the relative content of each element, and the XML3-1 spectrum was chosen for a comprehensive quantitative analysis of elemental composition in Xiamaling Formation kerogen. Table 5 summarizes the derived element concentrations from the XPS spectrum for XML3-1 kerogen. C dominates the kerogen structure at 89.00 %, followed by O at 8.61 %, N at 1.64 %, and S at 0.37 %. In addition to C, O, N, and S, XML3-1 kerogen contains trace amounts of Cl and Fe. The 0.36 % Cl originates from residual acid treatment, while 0.03 % Fe primarily corresponds to pyrite content [38].

The accurate determination of binding modes between carbon (C) and oxygen (O) through O 1s peak-fitting is hindered by the interference of nitrogen (N) and sulfur (S) [38]. As a result, to precisely discern the atomic state of oxygen, the C 1s signal is employed instead of the O 1s signal [25,38]. The XPS spectra of C 1s, N 1s, and S 2p were meticulously fitted to deduce the quantity of each respective O, N, and S functional group present within the XML3-1 kerogen sample. These XPS spectra, accompanied by the associated peak splitting curves and fitting curves, are presented in Fig. 7. The resultant analytical results are presented in Table 6.

The XPS C 1s spectrum of the XML3-1 kerogen (Fig. 7a) displays four distinct peaks situated at 284.80, 286.48, 287.66, and 289.06 eV. These peaks are attributed to different types of carbon and oxygen groups, including aliphatic and aromatic carbon, C—O, C—OH, O=C—O, and

O=C. The XPS analysis results, presented in Table 6, highlight that C—O and C—OH groups are the most abundant oxygen functionalities, followed closely by O=C—O and C=O. This result further confirms the observations derived from the  $^{13}\text{C}$  NMR analysis, as demonstrated in Table 3. Based on the computed percentages provided in Table 6, it can be deduced that there are approximately 6 C—OH and C—O groups, 5 O=C—O groups, and 3 O=C groups per 100 carbon atoms within the kerogen structure.

Building on earlier XPS investigations in kerogens and coals [18,25,27–31], a comprehensive analysis of the XPS N 1s spectrum for the XML3-1 kerogen unveiled the presence of five distinct nitrogen species: pyridine, amine, pyrrole, quaternary, and chemisorbed nitrogen oxides ( $\text{NO}_x$ ,  $x = 1, 2$ ), as illustrated in Fig. 7b. Delving into the conclusive data presented in Table 6, it becomes evident that pyrrolic nitrogen constitutes the dominant form of organic nitrogen within the Xiamaling kerogen, accounting for a substantial 43.4 % of all organic nitrogen species. This observation aligns with analogous findings reported in earlier kerogen studies [17,18,20,25,30,31]. Additionally, a minor quantity of amino, quaternary, and pyridine compounds contributes to the composition of the organic nitrogen structure within the kerogen. Notably, the Xiamaling kerogen sample also encompasses a discernible quantity of chemisorbed nitrogen oxides. This observation underscores that partial nitrogen oxides remain within the kerogen, potentially due to their strong associations with minerals or limited reactivity during the kerogen isolation procedure [17,18]. Comparable instances of such nitrogen oxides have been previously documented in other kerogen and coal studies [17,18,20,25,28,30].

The XPS S 2p spectrum of the XML3-1 kerogen underwent fitting involving eight distinct peaks (Fig. 7c). This comprehensive analysis revealed that sulfur atoms manifest in six discernible forms, namely pyrite, aliphatic sulfur, aromatic sulfur, sulfoxide, sulfone, and sulfate. This result shares similarities with the sulfur-containing compounds identified in Marcellus shale kerogen [31] and Jincheng No. 15 anthracite [29]. Quantitative scrutiny, as highlighted in Table 6, underscores that sulfur primarily takes on organic forms, with a marginal presence of pyrite and sulfate. From the perspective of the relative content of each organic sulfur functional group, sulfoxide emerges as the dominant functional group, constituting approximately 29.72 % of the composition, followed by aromatic sulfur, aliphatic sulfur, and sulfone.

### 3.3.3. FTIR analysis

FTIR has been widely utilized for qualitative analysis of functional groups within kerogen structures [17,18,25,28,30,38]. Standard FTIR spectra of three kerogen samples are displayed in Fig. 8. As shown in Fig. 8, very weak peaks are observed near  $431\text{ cm}^{-1}$ , which are caused by pyrite minerals. Except for a small trend error in the peak intensity between  $1400$  and  $1610\text{ cm}^{-1}$ , the three kerogens show similar characteristics of the FTIR spectra, and the spectrum of the XML3-1 kerogen is utilized to qualitatively analyze the functional groups in the kerogen structure.

As shown in the FTIR spectrum of XML3-1 in Fig. 8, the pronounced, high-intensity, sharp peaks observed at approximately  $2922.48$  and  $2852.03\text{ cm}^{-1}$  can be attributed to the asymmetric and symmetric stretching of alkyl  $\text{CH}_2$  groups. Notably, this aligns with the potent resonance signal at  $29.30\text{ ppm}$  in the  $^{13}\text{C}$  NMR spectrum (Fig. 5), effectively signifying the preeminence of  $\text{CH}_2$  groups within the kerogen structure. It is noteworthy that there is an absence of a discernible, well-defined peak around  $719.95\text{ cm}^{-1}$ , a wavenumber typically associated with the skeletal vibration of long-chain alkanes ( $(\text{CH}_2)_n$ ,  $n \geq 4$ ). This observation further substantiates the deduction drawn from the  $^{13}\text{C}$  NMR analysis that the average methylene chain length (Cn) is a low value of 3.87. Consequently, the preponderance of methylene carbons primarily exists in the form of short-branched chains or alicyclic structures within the Xiamaling kerogen. Moreover, a prominent peak at approximately  $1454.55\text{ cm}^{-1}$  is observed, which can be attributed to the asymmetrical bending vibration of  $\text{CH}_3$  groups and the shear vibration

**Table 4**

Structural parameters of the XML3-1 kerogen obtained from the curve-fitted  $^{13}\text{C}$  NMR spectrum.

| Structure parameters                 | Symbol   | Definition  | Experiment |
|--------------------------------------|----------|---|------------|
| Average methylene chain length       | Cn       | $\text{Cn} = (f_{al}^B + f_{al}^H)/f_{ar}^C$                          | 3.87       |
| Branched degree of aliphatic chain   | BI       | $\text{BI} = f_{al}^D/f_{al}$   | 0.29       |
| Substitutive degree of aromatic ring | $\delta$ | $\delta = (f_{ar}^C + f_{ar}^{O2})/f_{ar}$                            | 0.44       |
| Aromatic clusters size               | $X_{BP}$ | $X_{BP} = f_{ar}^B/(f_{ar}^{O1} + f_{ar}^H + f_{ar}^C + f_{ar}^{O2})$ | 0.34       |



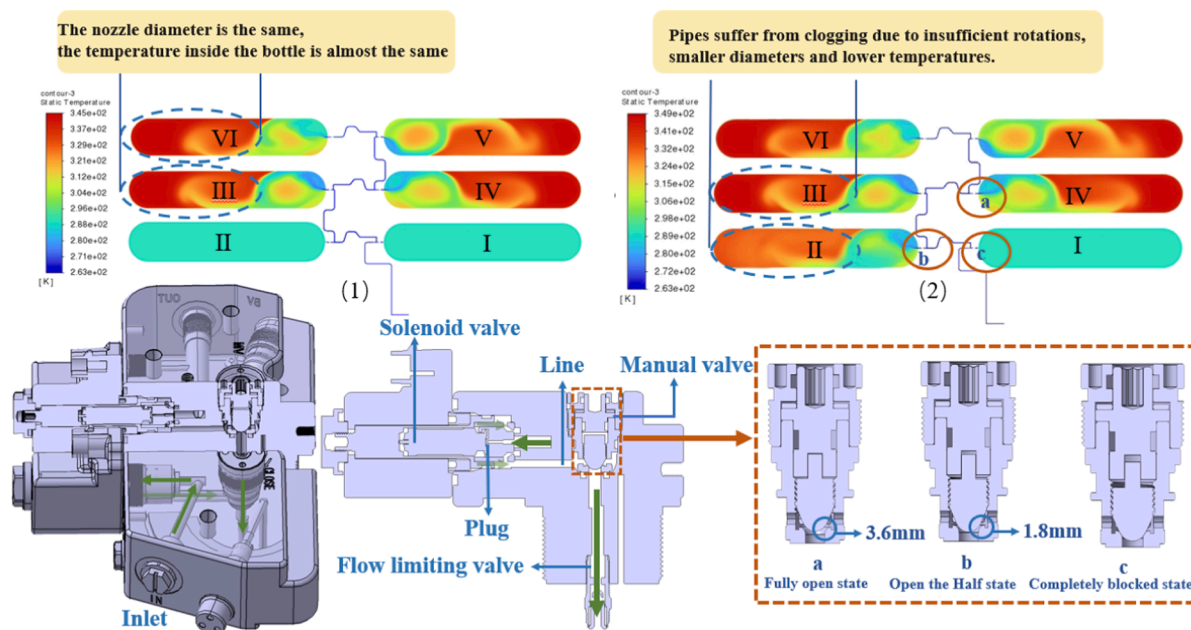


Fig. 6. XPS spectra of the Xiamaling kerogen samples, illustrating the elemental composition within the kerogen.

Table 5

Elemental composition of kerogen derived from XPS spectrum of the XML3-1 sample.

| Element | Binding energy/eV | Atomic/% |
|---------|-------------------|----------|
| C 1s    | 284.96            | 89.00    |
| O 1s    | 532.37            | 8.61     |
| N 1s    | 400.04            | 1.64     |
| S 2p    | 164.59            | 0.37     |
| Cl 2p   | 199.96            | 0.36     |
| Fe 2p   | 726.79            | 0.03     |

of  $\text{CH}_2$  groups. Conversely, a comparably subtle peak is discernible at  $1373.69\text{ cm}^{-1}$ , primarily stemming from the asymmetric bending of  $\text{CH}_3$  groups. This detailed observation effectively emphasizes that the Xiamaling kerogen contains a lower concentration of  $\text{CH}_3$  groups compared to  $\text{CH}_2$  groups, aligning with the results obtained from the  $^{13}\text{C}$  NMR analysis (as detailed in Table 3). A relatively broad peak at  $1607.64\text{ cm}^{-1}$  in the kerogen spectrum is clearly visible, positioned to the left of the  $\text{CH}_3$  and  $\text{CH}_2$  peaks (as depicted in Fig. 8). This particular peak corresponds to the aromatic stretching vibration of the  $\text{C}=\text{C}$  bond. Furthermore, the presence of two modest peaks at  $3053.62\text{ cm}^{-1}$  and  $879.11\text{ cm}^{-1}$  can be attributed to the aromatic  $\text{C}-\text{H}$  stretching vibration and out-of-plane deformation, respectively. Collectively, these spectral features suggest that the Xiamaling kerogen is characterized by a limited presence of aromatic rings.

The FTIR spectrum, in addition, offers some qualitative insights into the oxygen-containing functional groups present. Notably, a minor sharp peak is discernible at  $1702.47\text{ cm}^{-1}$ , arising from the  $\text{C}=\text{O}$  stretching of carboxyl and/or carbonyl groups. Correspondingly, the relatively subdued resonance signal observed at  $165\text{--}220\text{ ppm}$  in the  $^{13}\text{C}$  NMR spectra (Fig. 5) aligns with these carboxyl/carbonyl groups. Within the higher wavenumber range of the FTIR spectrum (Fig. 8), a pronounced, medium-intensity, broad band emerges at around  $3431.43\text{ cm}^{-1}$ . This particular feature predominantly stems from  $\text{O}-\text{H}$  stretching vibrations. Additionally, the appearance of a broadband at approximately  $1246.91\text{ cm}^{-1}$  provides evidence for the presence of ethers ( $\text{C}-\text{O}$ ) and alcohols ( $\text{C}-\text{OH}$ ) within the Xiamaling kerogen.

### 3.3.4. XRD analysis

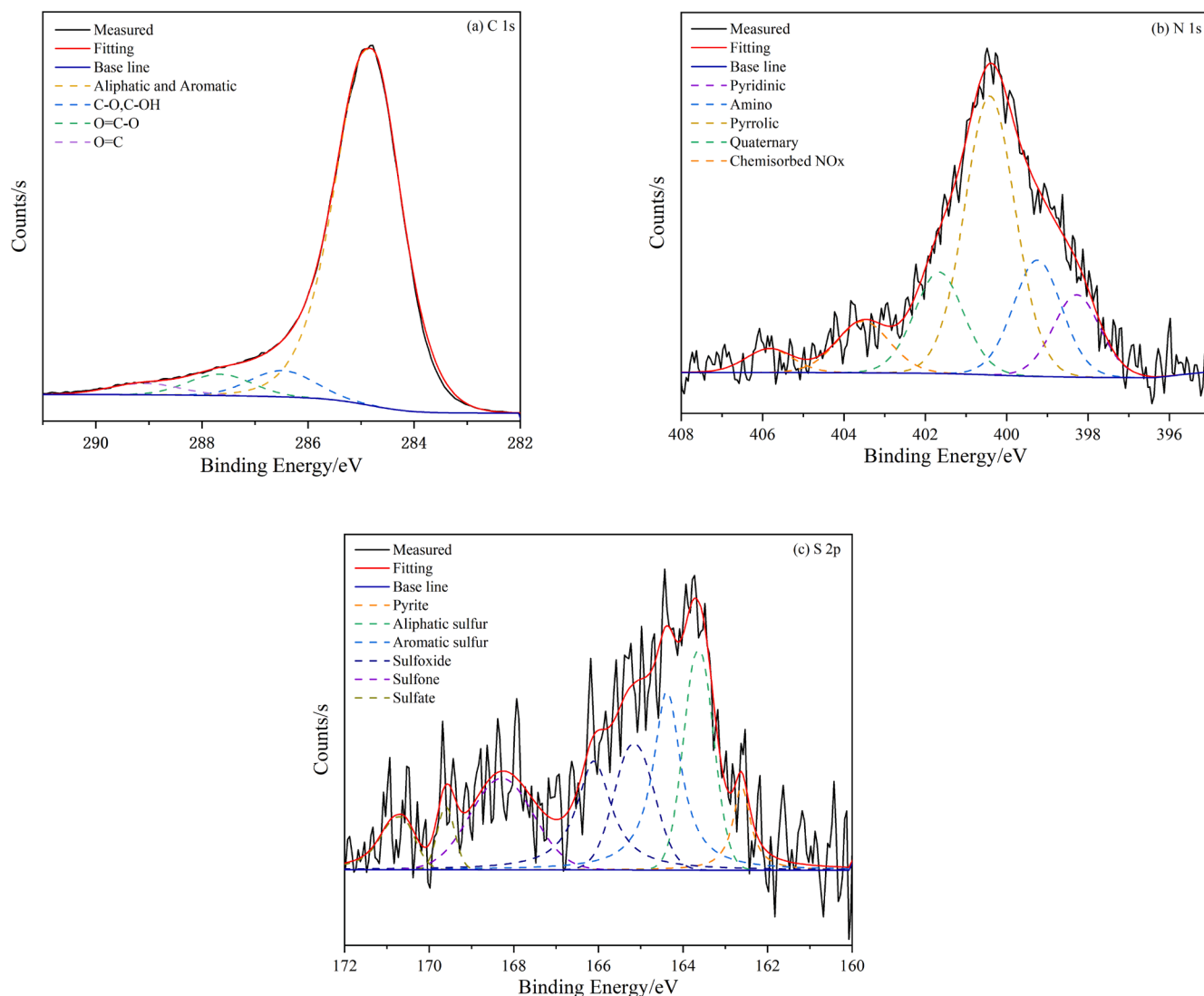
The XRD spectra of three kerogen samples from the Xiamaling Formation are shown in Fig. 9. Notably, The high-intensity peaks, along with minor sharp peaks observed in Fig. 9, can be attributed to the presence of quartz (Q) and pyrite (Py). It is worth acknowledging that complete removal of these inorganic minerals via chemical methods is challenging [2,18,28]. The XRD spectra of the three kerogens are fairly similar to each other, except for some slight differences in the intensity of the observed pyrite peaks. Compared with the XRD spectra of the XML3-2 and XML3-4 kerogen samples, the XRD spectrum of the XML3-1 kerogen exhibits notable peaks in  $2\theta = 26^\circ$  and  $42.7^\circ$ , along with low-density pyrite peaks in the range of  $2\theta = 38^\circ\text{--}44^\circ$ . As a result, the XRD spectrum of the XML3-1 kerogen was selected for detailed analysis of the kerogen structure.

A medium-intensity, broad hump at around  $2\theta = 20^\circ$ , designated as the  $\gamma$ -peak, emerges in the spectrum of XML3-1 kerogen. This hump is attributed to the diffraction of naphthenes and n-alkanes [17,18,25,28]. A weak peak at  $2\theta = 26^\circ$ , referred to as the 002 peak, serves as an indicator of the presence of aromatic structures. Additionally, the (100) peak at  $42.7^\circ$ , representing the in-plane structure of aromatic hydrocarbons, is discernible but with low intensity. This suggests that only a minor portion of aromatic carbons in the Xiamaling kerogen contribute to organized clustering, while most form disordered clusters interconnected by diverse bridge bonds [25,30]. These XRD analytical findings align with the  $^{13}\text{C}$  NMR and FTIR analyses, underscoring the dominance of aliphatic methylene groups within the Xiamaling kerogen.

### 3.4. Establishment of 2D kerogen molecular structure

The chemical compositions and structural attributes of the Xiamaling kerogen were comprehensively assessed through elemental analysis,  $^{13}\text{C}$  NMR, XPS, FTIR, and XRD methodologies. These results reveal that the kerogen predominantly comprises short-branched chains, cycloalkanes, and aromatic compounds. The subsequent steps delineate the process of constructing the molecular model for the Xiamaling kerogen.

In the first step, the initial chemical formula of Xiamaling kerogen was established based on elemental analysis data. Considering previous studies [16,17,19,41,42], the molecular mass of kerogen models usually falls within the range of 3000 to 6000 Da. An initial bulk molecular formula of  $\text{C}_{284}\text{H}_{334}\text{O}_{18}\text{N}_5\text{S}_2$  was derived from the elemental analysis of



**Fig. 7.** XPS (a) C 1s spectrum, (b) N 1s spectrum, (c) S 2p spectrum, and their respective fitting curves for qualitative and quantitative characterization of O, N, and S heteroatom functional groups in the XML3-1 kerogen.

**Table 6**

XPS analysis results for heteroatom functional groups of the XML3-1 kerogen.

| Elemental peak | Functionality                 | Binding energy/eV | Atomic/% |
|----------------|-------------------------------|-------------------|----------|
| C 1s           | Aliphatic and aromatic carbon | 284.80            | 86.46    |
|                | C—O and C—OH                  | 286.48            | 5.94     |
|                | O=C—O                         | 287.66            | 4.78     |
|                | O=C                           | 289.06            | 2.82     |
|                |                               |                   |          |
| N 1s           | Pyridine                      | 398.28            | 11.93    |
|                | Amino                         | 399.24            | 17.35    |
|                | Pyrrole                       | 400.41            | 43.40    |
|                | Quaternary                    | 401.66            | 15.37    |
|                | Chemisorbed nitrogen oxides   | 402–406           | 11.96    |
| S 2p           | Pyrite                        | 162.61            | 6.12     |
|                | Aliphatic sulfur              | 163.61            | 18.61    |
|                | Aromatic sulfur               | 164.38            | 20.49    |
|                | Sulfoxide                     | 165–166           | 29.72    |
|                | Sulfone                       | 168.28            | 16.30    |
|                | Sulphate                      | 169–170.7         | 8.76     |

the Xiamaling kerogen, resulting in a molecular mass of 4169.83 Da. This size is akin to the Chang 7 type II kerogen (formula:  $C_{300}H_{297}O_{18}N_{13}S_{10}$ ) [19], demonstrating similarity in scale. In the

second step, the carbon skeleton of the kerogen model was established by drawing on insights from the  $^{13}C$  NMR experiment. Notably, the products of kerogen pyrolysis, known as asphaltene molecules, are recognized for being primarily comprised of 4–8 aromatic ring structures [43]. Huang et al. [44] proposed that the size of aromatic clusters in the initial type II kerogen ( $\%Ro = 0.76$ ) ranged from 1 to 4 in the Bohai Bay Basin. In contrast, Zhang et al. [39] identified sizes of 1–2 in type II kerogen ( $\%Ro = 0.67$ ) from the Nengjiang Formation in the Songliao Basin. Hence, based on the  $X_{BP}$  value and the proportion of aromatic carbon in Xiamaling kerogen, a fundamental carbon skeleton with 1–4 aromatic rings was established for aromatic cluster structures. To meet the requirement for highly branched aliphatic carbons, several polycyclic saturated structures were incorporated. Short methylene chains were introduced as connectors between aromatic rings and cyclanes, forming the molecular structure model for the Xiamaling kerogen. In the third step, heteroatom functional groups were identified and incorporated into the kerogen molecule. Oxygenated groups such as ether, hydroxyl, carboxyl, and carbonyl were considered, consistent with results from  $^{13}C$  NMR, XPS, and FTIR analyses. Organic nitrogen- and sulfur-containing functional groups were determined based on XPS experimental data (Table 5). Specifically, the model included 1 pyridine, 1 amino, 3 pyrroles, and 1 quaternary nitrogen for organic nitrogen species. For organic sulfur species, 2 sulfur atoms in the initial kerogen

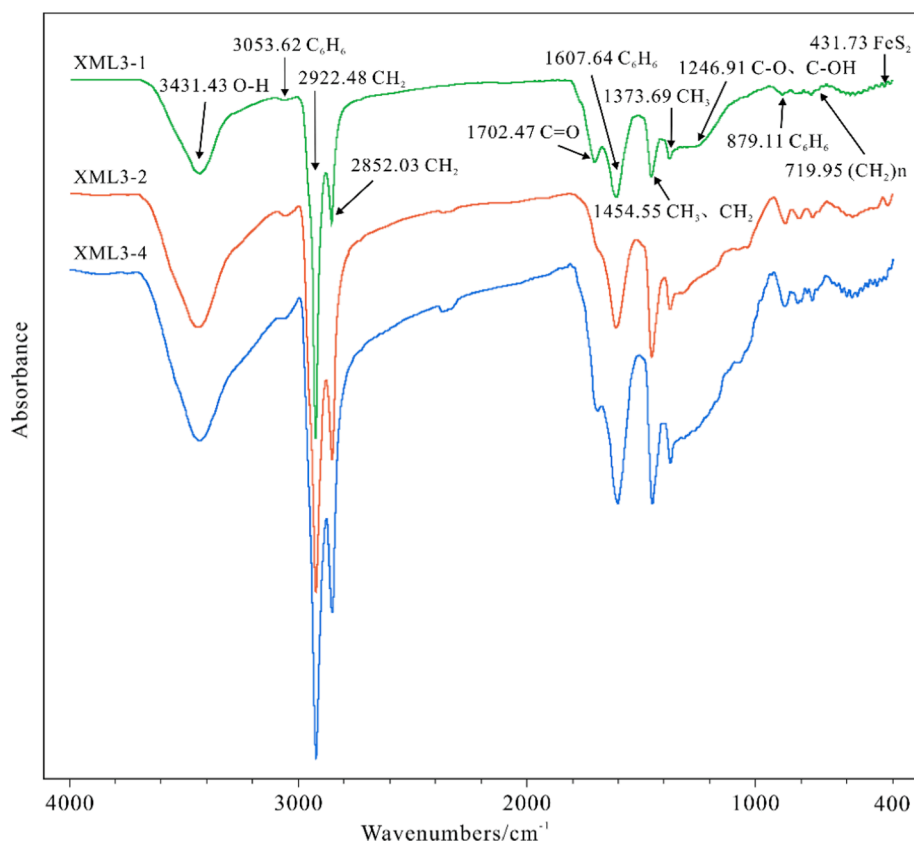


Fig. 8. FTIR spectra of the Xiamaling kerogen samples for qualitative functional group characterization.

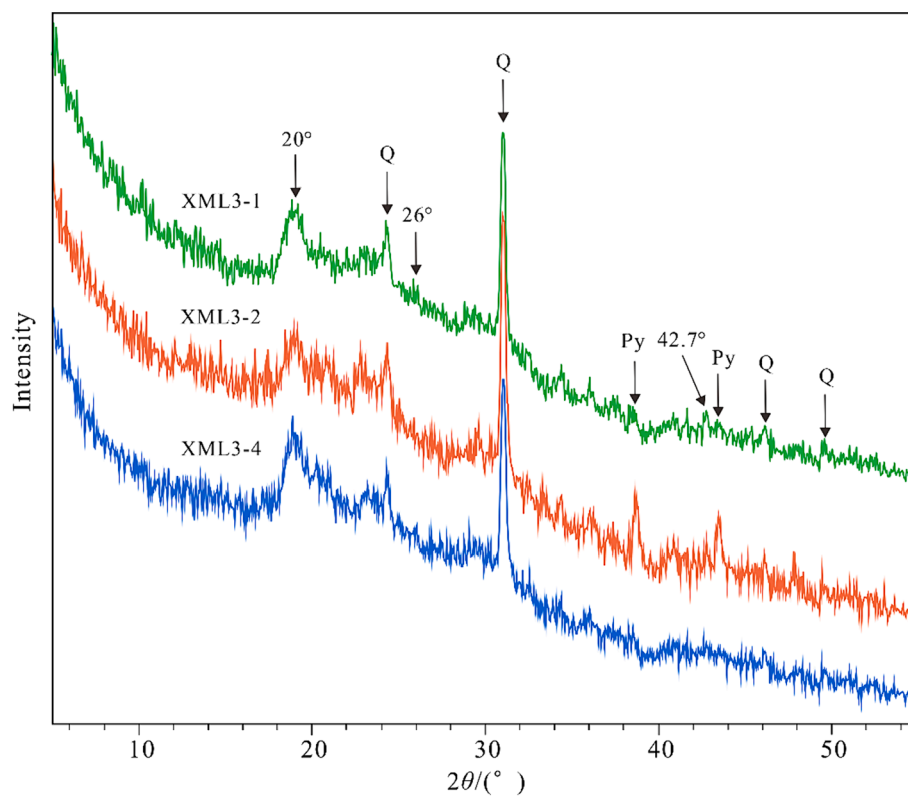


Fig. 9. XRD spectra of the Xiamaling kerogen samples for the qualitative analysis of mineral composition and carbon structure within the kerogen structure (Q = quartz, Py = pyrite).



molecule were associated with 1 thiophene and 1 sulfoxide. In the fourth step, the initial 2D kerogen model was constructed by interlinking these fragments. In the fifth step, a  $^{13}\text{C}$  NMR spectrum for the initial 2D kerogen molecule was predicted by the MestReNova software. The NMR predictors within the software utilized the fastest algorithms to compute the  $^{13}\text{C}$  NMR spectrum of the required predicted molecular structure [20]. The validity of this method has been substantiated through reports and validation studies by Wang et al. [45] and Zhao et al. [20]. In the sixth step, a comparison was made between the predicted and experimental  $^{13}\text{C}$  NMR spectra. It is well-established that the spatial arrangement of carbon atoms and their neighboring atoms exerts a significant influence on the chemical shift of individual carbon atoms and, consequently, the overall  $^{13}\text{C}$  NMR spectrum of the 2D kerogen model [31]. Consequently, the congruence between predicted and experimental  $^{13}\text{C}$  NMR spectra has been recognized as a robust criterion for evaluating the accuracy of generated 2D kerogen or coal models [16,20,31,38,44–49]. In the concluding phase, an iterative process encompassing steps 4 to 6 was executed to iteratively refine and substitute aliphatic carbons, aromatic carbons, and heteroatom functional groups within the kerogen molecule, ultimately achieving a satisfactory alignment between the predicted and experimental  $^{13}\text{C}$  NMR spectra.

The comparison between the  $^{13}\text{C}$  NMR spectra and the final kerogen molecule is exhibited in Fig. 10, while Table 7 presents the structural parameters derived from both experimental and modeling approaches. Notably, the simulated spectrum closely aligns with the experimental one, exhibiting a general agreement despite minor discrepancies. Additionally, the carbon structural parameters and the composition of oxygen and nitrogen heteroatom functional groups in the established kerogen model are very close to the experimental  $^{13}\text{C}$  NMR and XPS results. For instance, a key parameter  $X_{\text{BP}}$ , representing the size of aromatic clusters, was calculated from the  $^{13}\text{C}$  NMR data and found to be 0.34 (as shown in Table 4), matching the expected  $X_{\text{BP}}$  of the simulated structure (0.34 from Table 7). However, because the current computing power limits the size of the established kerogen model and there is a limited amount of sulfur in Xiamaling Formation kerogen, only aromatic sulfur and sulfoxide are considered in the sulfur functional group of Xiamaling Formation kerogen model. Generally speaking, this congruence underscores the reliability of the molecular model, characterized by a chemical formula of  $\text{C}_{278}\text{H}_{346}\text{N}_6\text{O}_{26}\text{S}_2$  and a total molecular weight of 4251.86 Da. The model effectively captures the essential attributes of the Xiamaling kerogen sample with a high degree of accuracy.

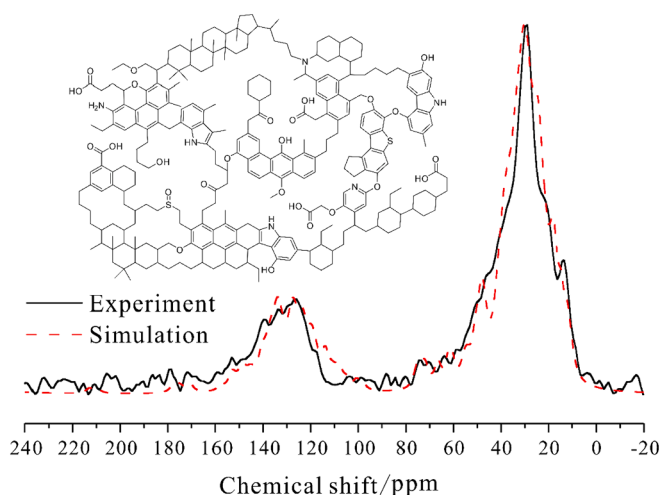


Fig. 10. 2D structural model of the Xiamaling kerogen ( $\text{C}_{278}\text{H}_{346}\text{N}_6\text{O}_{26}\text{S}_2$ ) and correlation between the experimental and simulated  $^{13}\text{C}$  NMR spectra.

Table 7

Structural parameters of the Xiamaling kerogen.

| Property | Structural parameter | Experimental | Modeling |
|----------|----------------------|--------------|----------|
| C groups | Cn                   | 3.87         | 2.35     |
|          | BI                   | 0.29         | 0.25     |
|          | $\delta$             | 0.44         | 0.50     |
|          | $X_{\text{BP}}$      | 0.34         | 0.34     |
| O groups | C—O and C—OH         | 43.87 %      | 52.00 %  |
|          | O=C—O                | 35.30 %      | 40.00 %  |
|          | O=C                  | 20.83 %      | 8.00 %   |
| N groups | Pyridine             | 13.55 %      | 16.67 %  |
|          | Amino                | 19.71 %      | 16.67 %  |
|          | Pyrrole              | 49.30 %      | 50.00 %  |
|          | Quaternary           | 17.46 %      | 16.67 %  |
| S groups | Aliphatic            | 21.86 %      | 0 %      |
|          | Aromatic             | 24.07 %      | 50 %     |
|          | Sulfoxide            | 34.92 %      | 50 %     |
|          | Sulfone              | 19.15 %      | 0 %      |

### 3.5. Construction of 3D kerogen model

To create a reliable 3D kerogen structure model, the five optimized Xiamaling kerogen molecules were put into a periodic box. Subsequently, a sequence of essential calculations, including geometry optimization, anneal dynamic, and MD simulations, were sequentially executed. Fig. 11a graphically depicts that the equilibrium temperature of the Xiamaling kerogen cell stabilized at 298.147 K, with the initial temperature closely approximating the average temperature. Significantly, density emerges as a fundamental physical parameter in characterizing kerogen. Consequently, numerous studies have adopted simulated density as a benchmark for assessing the validity of 3D kerogen structure models [16,28,50,51]. The equilibrium state of kerogen density was achieved during the latter 200 ps of dynamic simulation, as demonstrated in Fig. 11b. The resulting average density of the kerogen cell settled at  $1.067 \text{ g/cm}^3$ , in agreement with the density observed in both simulated [11,51,52] and experimental [53] investigations of type II kerogen. The final configuration of the Xiamaling kerogen cell is visually represented in Fig. 12. Specifically, the Xiamaling kerogen model possesses the following cell parameters:  $a = 40.85 \text{ \AA}$ ,  $b = 28.17 \text{ \AA}$ ,  $c = 32.54 \text{ \AA}$ ,  $\alpha = 69.46^\circ$ ,  $\beta = 101.27^\circ$ , and  $\gamma = 79.64^\circ$ .

### 3.6. Comparison with other kerogen models

The Mesoproterozoic Xiamaling kerogen model (Fig. 10) presented in this study marks the first-ever Precambrian kerogen model developed. In order to discern the structural attributes of Precambrian and Phanerozoic kerogen molecular models, four representative kerogen models were chosen: Lower Cambrian Niutitang Formation [38], Upper Devonian Duvernay shale [11], Triassic Chang 7 Member [19], and Upper Cretaceous Nenjiang Formation [39], as depicted in Fig. 13. These models exemplify type II kerogens characterized by low immaturity, akin to that of the Xiamaling kerogen.

The five chosen type II kerogens, among them the Xiamaling kerogen (Fig. 10), display an elevated aromatic carbon ratio ( $f_{\text{ar}}$ ) exceeding 25 %, and their aliphatic chain branching index (BI) falls within the range of 0.11 to 0.53. This significantly contrasts with the less immature type I kerogens [16,17,25,54], which typically exhibit a lower  $f_{\text{ar}}$  of 10 %–17 % and a BI within the range of 0.045–0.085. Furthermore, their average methylene chain length ( $\text{Cn} = 0.18\text{--}3.35$ ) is considerably lower than one-third of the Cn parameter (10–28) observed in type I kerogens [16,17,25,54]. These findings highlight that type II kerogens possess not only shorter aliphatic chains but also a greater prevalence of aliphatic rings and aromatic structures, in line with previous reports [2].

Fig. 10 illustrates that the aliphatic structure within the Mesoproterozoic Xiamaling Formation kerogen model comprises monocyclic alkanes and polycyclic saturated structures. This structural resemblance aligns with the Paleozoic kerogen models (Fig. 13 a and b), in contrast to

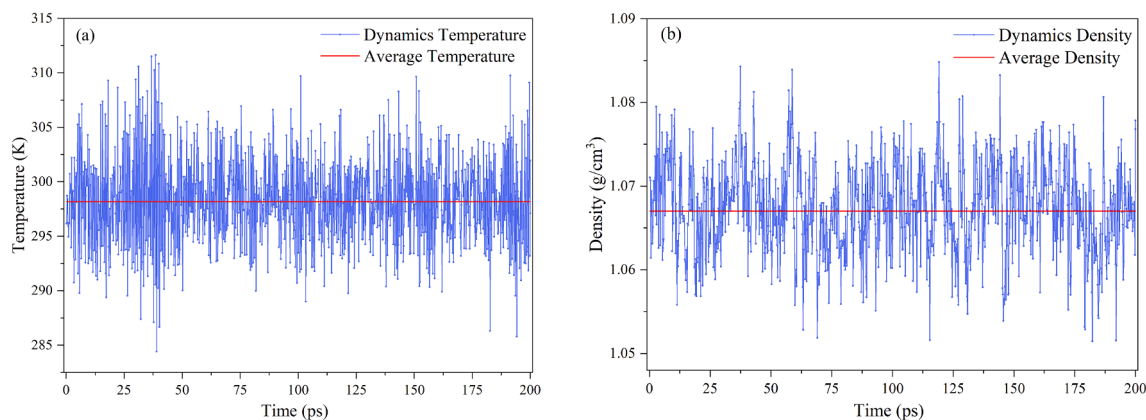


Fig. 11. Temporal evolution of the Xiamaling kerogen (a) Temperature and (b) Density.

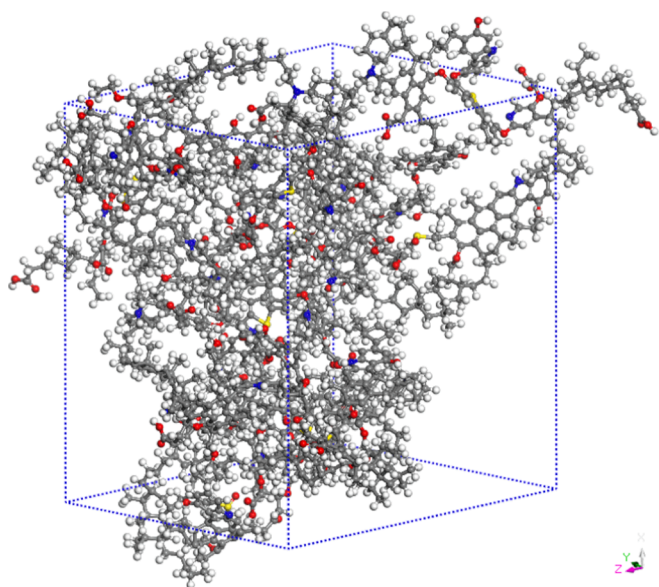


Fig. 12. 3D structure of the Xiamaling kerogen's molecular arrangement at a density of  $1.067 \text{ g/cm}^3$  within a simulation box measuring  $40.85 \times 28.17 \times 32.54 \text{ Å}^3$ .

the observations made in the case of the Mesozoic kerogen models (Fig. 13 c and d). Moreover, both the Xiamaling Formation (Fig. 10) and Duvernay Shale kerogen (Fig. 13b) models exhibit aliphatic structures that include chains with an average methylene chain length parameter  $C_n < 4$ , where the longest chains consist of 5 and 10 carbon atoms, respectively. In contrast, the aliphatic structure of the Niutitang Formation shale (Fig. 13a), except for naphthenes, lacks aliphatic chains due to its notably low  $C_n$  of 0.18 and high BI of 0.53. In contrast to the Niutitang Formation kerogen model, the aliphatic structures within the Mesozoic kerogen molecular models of the Triassic Chang 7 Member (Fig. 13c) and Upper Cretaceous Nenjiang Formation (Fig. 13d) predominantly consist of short, highly branched aliphatic chains. The longest aliphatic chains in these models contain 4 and 7 carbon atoms, respectively.

The kerogen models derived from the Xiamaling Formation (Fig. 10), Niutitang Formation (Fig. 13a), Chang 7 Member (Fig. 13c), and Nenjiang Formation (Fig. 13d) share notable similarities in terms of their aromatic structure characteristics. All these models incorporate varying numbers of aromatic clusters, ranging from 1 to 5, with an average count of aromatic carbons approximately in the range of 8–10 per polyaromatic cluster. In contrast, the kerogen model of the Duvernay shale (Fig. 13b) demonstrates larger polyaromatic structures with aromatic

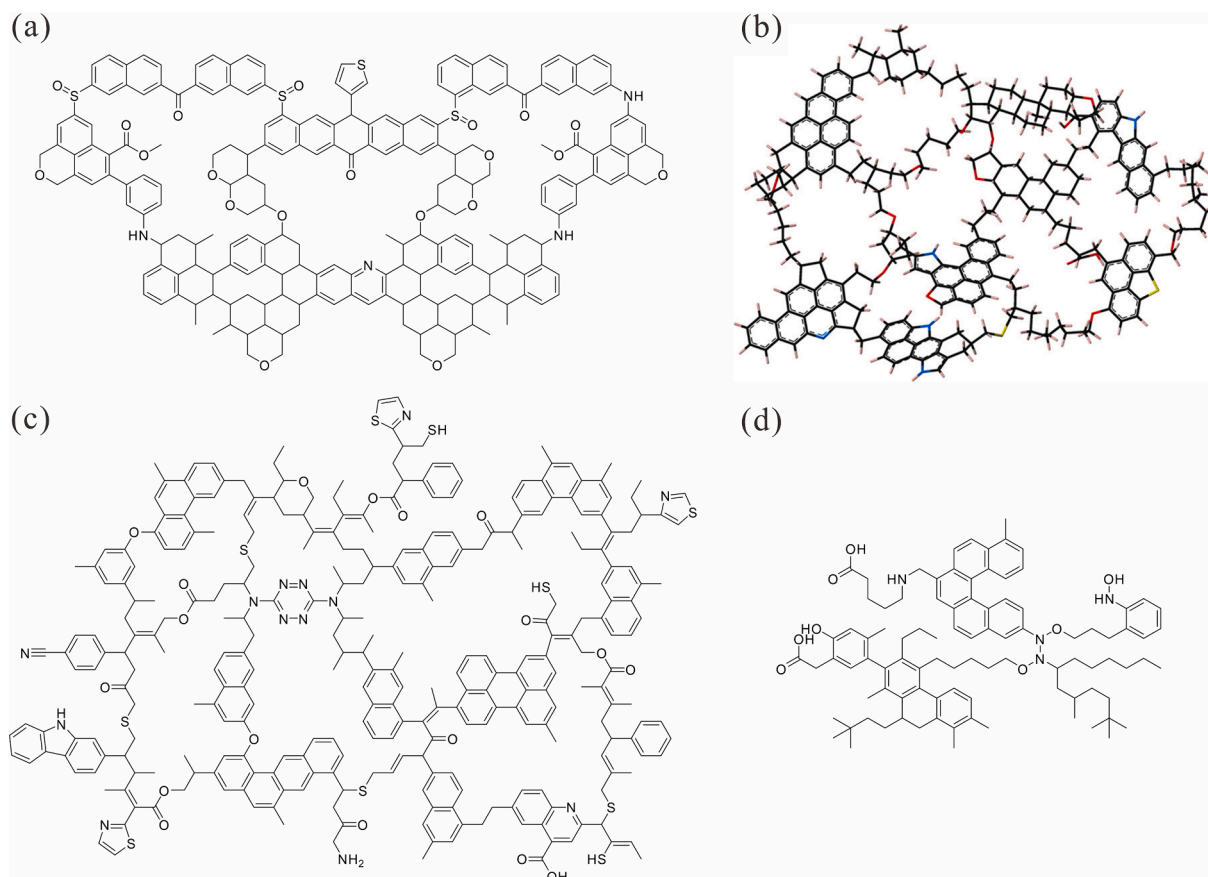
ring numbers spanning from 4 to 6. This model exhibits a maximal average count of carbon atoms per aromatic cluster of 17.5 and a maximum aromatic cluster size parameter ( $X_{BP}$ ) of approximately 0.78.

Regarding heteroatom functional groups, the Chang 7 Member kerogen model (Fig. 13c) is distinctive. It includes thiazole, tetrazine, chinolin, and nitrate as representatives of nitrogen and sulfur functional groups, setting it apart from the other four kerogen models. These other models exclusively feature pyridine, amino, pyrrole, quaternary, sulfide, thiophene, and sulfide bridge as the primary forms of nitrogen and sulfur atoms.

#### 4. Conclusions

This study delved into the molecular structure of kerogen within the Xiamaling Formation, leveraging elemental compositions and structural properties derived from a combination of analytical techniques, including elemental analysis,  $^{13}\text{C}$  NMR, XPS, FTIR, and XRD experiments. Subsequently, a 3D model for the Xiamaling kerogen was developed with a relatively high degree of accuracy. In the end, the following conclusions are reached:

- (1) The molecular structure of Xiamaling kerogen is composed of 75.19 % aliphatic carbon and 22.06 % aromatic carbon. Notably, the relatively low average methylene chain length (3.87) indicates the prevalence of short-branched chains or alicyclics rather than long straight chains in the aliphatic structure. Within the organic nitrogen species, pyrrole is dominant, followed by amino, quaternary, and pyridine. Regarding organic sulfur-containing groups, sulfoxide, aromatic sulfur, aliphatic sulfur, and sulfone are present.
- (2) The establishment of a reasonable molecular structure was achieved with a molecular formula of  $\text{C}_{278}\text{H}_{346}\text{N}_6\text{O}_{26}\text{S}_2$  and a total molecular weight of 4251.86 Da. This structure was validated through the alignment of predicted and experimental  $^{13}\text{C}$  NMR spectra. The central carbon skeleton of the Xiamaling kerogen model predominantly consists of 1–4 aromatic clusters and cycloalkanes. Furthermore, a robust 3D kerogen structure model was obtained through geometry optimization, anneal dynamics calculations and MD simulations. This model exhibits a density of  $1.067 \text{ g/cm}^3$  and comprises a total of 3290 atoms.
- (3) The aliphatic structure of the Mesoproterozoic Xiamaling kerogen model incorporates monocyclic alkanes and polycyclic saturated structures. This pattern is consistent with the aliphatic features observed in Paleozoic kerogen models, but it contrasts with the aliphatic characteristics present in Mesozoic kerogen models.



**Fig. 13.** Type II kerogen models for the Paleozoic and Mesozoic organic-rich shales. (a) Kerogen model of the Lower Cambrian Niutitang Formation ( $C_{210}H_{184}O_{20}N_4S_4$ ) [38]; (b) Type II-B kerogen model (Top of Oil Window) of Upper Devonian Duvernay shale ( $C_{234}H_{263}O_{14}N_5S_2$ ) [11]; (c) Kerogen model of Triassic Chang 7 member ( $C_{300}H_{297}O_{18}N_{13}S_{10}$ ) [19]; (d) Kerogen model (early oil window) of Upper Cretaceous Nenjiang Formation ( $C_{94}H_{124}N_4O_8$ ) [39].

#### CRedit authorship contribution statement

**Chengbo Guo:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Meijun Li:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition. **Xiaoqiang Liu:** Writing – review & editing, Validation, Software, Resources, Methodology, Investigation, Formal analysis. **Hong Xiao:** Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Qingyong Luo:** Resources. **Qiuya Han:** Software, Methodology, Formal analysis. **Wenke Li:** Resources, Investigation, Formal analysis. **Junhao Ren:** Investigation, Formal analysis, Data curation.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgments

This work was funded by the National Natural Science Foundation of China (Grant No.42173054), Natural Science Foundation of Sichuan Province (No.2022NSFSC0182), and partly supported by the 18<sup>th</sup> “Challenge Cup” Competition for National College Students Extracurricular Academic and Technical Works. The authors express sincere

gratitude to the guide by researchers from the Research Institute of Petroleum Exploration & Development. We would like to thank Dr. Berna Hascakir, the reviewer Dr. Rahmouni and the two anonymous reviewers for their careful check and constructive comments.

#### References

- [1] Miknis FP, Smith JW. An NMR survey of United States oil shales. *Org Geochem* 1984;5(4):193–201.
- [2] Vandenbroucke M, Largeau C. Kerogen origin, evolution and structure. *Org Geochem* 2007;38(5):719–833.
- [3] Al-Harabsheh A, Al-Harabsheh M, Al-Otoom A, Allawzi M. Effect of demineralization of El-lajjun Jordanian oil shale on oil yield. *Fuel Process Technol* 2009;90(6):818–24.
- [4] Miknis FP, Netzel DA, Smith JW, Mast MA, Maciel GE.  $^{13}C$  NMR measurements of the genetic potentials of oil shales. *Geochim Cosmochim Acta* 1982;46(6):977–84.
- [5] Miknis FP, Conn PJ. A common relation for correlating pyrolysis yields of coals and oil shales. *Fuel* 1986;65(2):248–50.
- [6] Qin K, Chen D, Li Z. A new method to estimate the oil and gas potentials of coals and kerogens by solid state  $^{13}C$  NMR spectroscopy. *Org Geochem* 1991;17(6): 865–72.
- [7] Freund H, Walters CC, Kelemen SR, Siskin M, Gorbaty ML, Curry DJ, et al. Predicting oil and gas compositional yields via chemical structure–chemical yield modeling (CS-CYM): Part 1 – concepts and implementation. *Org Geochem* 2007;38 (2):288–305.
- [8] Siskin M, Scouten C, Rose K, Aczel T, Colgrove S, Pabst R. Detailed structural characterization of the organic material in Rundle Ramsay Crossing and Green River oil shales. *Fuel* 1995;37(1):143–58.
- [9] Orendt AM, Pimienta ISO, Badu SR, Solum MS, Pugmire RJ, Facelli JC, et al. Three-dimensional structure of the siskin Green River oil shale kerogen model: a comparison between calculated and observed properties. *Energy Fuel* 2013;27(2): 702–10.
- [10] Behar F, Vandenbroucke M. Chemical modelling of kerogens. *Org Geochem* 1987; 11(1):15–24.
- [11] Ungerer P, Collet J, Yiannourakou M. Molecular modeling of the volumetric and thermodynamic properties of kerogen: influence of organic type and maturity. *Energy Fuel* 2015;29(1):91–105.



- [12] Kelemen SR, Afeworki M, Gorbaty ML, Sansone M, Kwiatek PJ, Walters CC, et al. Direct characterization of kerogen by X-ray and solid-state  $^{13}\text{C}$  nuclear magnetic resonance methods. *Energy Fuel* 2007;21(3):1548–61.
- [13] Sui H, Yao J. Effect of surface chemistry for  $\text{CH}_4/\text{CO}_2$  adsorption in kerogen: a molecular simulation study. *J Nat Gas Sci Eng* 2016;31:738–46.
- [14] Zhao T, Li X, Zhao H, Li M. Molecular simulation of adsorption and thermodynamic properties on type II kerogen: influence of maturity and moisture content. *Fuel* 2017;190:198–207.
- [15] Zhou J, Mao Q, Luo KH. Effects of moisture and salinity on methane adsorption in kerogen: a molecular simulation study. *Energy Fuel* 2019;33(6):5368–76.
- [16] Ru X, Cheng Z, Song L, Wang H, Li J. Experimental and computational studies on the average molecular structure of Chinese Huadian oil shale kerogen. *J Mol Struct* 2012;1030:10–8.
- [17] Guan X-H, Liu Y, Wang D, Wang Q, Chi M-S, Liu S, et al. Three-dimensional structure of a Huadian oil shale kerogen model: an experimental and theoretical study. *Energy Fuel* 2015;29(7):4122–36.
- [18] Wang Q, Pan S, Bai J, Chi M, Cui D, Wang Z, et al. Experimental and dynamics simulation studies of the molecular modeling and reactivity of the Yaojie oil shale kerogen. *Fuel* 2018;230:319–30.
- [19] Wu Z, Xu Z. Experimental and molecular dynamics investigation on the pyrolysis mechanism of Chang 7 type-II oil shale kerogen. *J Pet Sci Eng* 2022;209:109878.
- [20] Zhao F, Li B, Che D, Liu S. Chemical structure characteristics and multidimensional model construction of Fushun oil shale kerogen: an experimental and simulation study. *J Mol Struct* 2022;1261:132878.
- [21] Zhao F, Li B, Zhang L, Che D, Liu S. The mechanism of superheated steam affecting the quality of in-situ pyrolysates of oil shale kerogen: part A-saturation of pyrolytic organics. *Fuel* 2022;323:124331.
- [22] Xiao H, Li M, Wang T, You B, Lu X, Wang X. Organic molecular evidence in the ~1.40 Ga Xiamaling formation black shales in North China Craton for biological diversity and paleoenvironment of mid-Proterozoic ocean. *Precamb Res* 2022; 381:106848.
- [23] Durand B, Nicaise G. Procedures for kerogen isolation. In: Durand B, editor. *Kerogen insoluble organic matter from sedimentary rocks*. Paris: Editions Technip; 1980. p. 35–53.
- [24] Orr WL. Kerogen/asphaltene/sulfur relationships in sulfur-rich Monterey oils. *Org Geochem* 1986;10(1):499–516.
- [25] Tong J, Han X, Wang S, Jiang X. Evaluation of structural characteristics of Huadian oil shale kerogen using direct techniques (solid-state  $^{13}\text{C}$  NMR, XPS, FT-IR, and XRD). *Energy Fuel* 2011;25(9):4006–13.
- [26] Hou L, Ma W, Luo X, Tao S, Guan P, Liu J. Chemical structure changes of lacustrine type-II kerogen under semi-open pyrolysis as investigated by solid-state  $^{13}\text{C}$  NMR and FT-IR spectroscopy. *Mar Pet Geol* 2020;116:104348.
- [27] Lee H, Oncel N, Liu B, Kukay A, Altincicek F, Varma RS, et al. Structural evolution of organic matter in deep shales by spectroscopy (1H and  $^{13}\text{C}$  nuclear magnetic resonance, X-ray photoelectron spectroscopy, and Fourier transform infrared) analysis. *Energy Fuel* 2020;34(3):2807–15.
- [28] Wang Q, Wang X, Pan S. The three-dimensional molecular structure model of Fushun oil shale kerogen. *China J Mol Struct* 2022;1255:132380.
- [29] Wang Y-G, Wei X-Y, Xie R-L, Liu F-J, Li P, Zong Z-M. Structural characterization of typical organic species in Jincheng No. 15 anthracite. *Energy Fuel* 2015;29(2): 595–601.
- [30] Wang Q, Hou Y, Wu W, Yu Z, Ren S, Liu Q, et al. A study on the structure of Yilan oil shale kerogen based on its alkali-oxygen oxidation yields of benzene carboxylic acids,  $^{13}\text{C}$  NMR and XPS. *Fuel Process Technol* 2017;166:30–40.
- [31] Liu Y, Liu S, Zhang R, Zhang Y. The molecular model of Marcellus shale kerogen: experimental characterization and structure reconstruction. *Int J Coal Geol* 2021; 246:103833.
- [32] Painter PC, Snyder RW, Starsinic M, Coleman MM, Kuehn DW, Davis A. Concerning the application of FT-IR to the study of coal: a critical assessment of band assignments and the application of spectral analysis programs. *Appl Spectrosc* 1981;35(5):475–85.
- [33] Craddock PR, Le Doan TV, Bake K, Polyakov M, Charsky AM, Pomerantz AE. Evolution of kerogen and bitumen during thermal maturation via semi-open pyrolysis investigated by infrared spectroscopy. *Energy Fuel* 2015;29(4): 2197–210.
- [34] Liu X-Q, Li M, Zhang C, Fang R, Zhong N, Xue Y, et al. Mechanistic insight into the optimal recovery efficiency of CBM in sub-bituminous coal through molecular simulation. *Fuel* 2020;266:117137.
- [35] Mukhopadhyay PK, Wade JA, Kruger MA. Organic facies and maturation of Jurassic/Cretaceous rocks, and possible oil-source rock correlation based on pyrolysis of asphaltene, Scotian Basin, Canada. *Org Geochem* 1995;22(1):85–104.
- [36] Welte DH, Tissot PB. *Petroleum formation and occurrence*. Berlin: Springer-Verlag; 1984.
- [37] Wu J, Li H, Goodarzi F, Min X, Cao W, Huang L, et al. Geochemistry and depositional environment of the Mesoproterozoic Xiamaling shales, northern North China. *J Pet Sci Eng* 2022;215:110730.
- [38] Huang L, Ning Z, Wang Q, Ye H, Wang Z, Sun Z, et al. Microstructure and adsorption properties of organic matter in Chinese Cambrian gas shale: experimental characterization, molecular modeling and molecular simulation. *Int J Coal Geol* 2018;198:14–28.
- [39] Zhang D, Cao H, Lei Y, Wu H, Wang X, Guo X, et al. A study on molecular structural evolution of type II kerogen in a gold tube thermal system: insights from solid-state  $^{13}\text{C}$  NMR. *Fuel* 2023;331:125898.
- [40] Xu F, Pan S, Liu C, Zhao D, Liu H, Wang Q, et al. Construction and evaluation of chemical structure model of Huolinhe lignite using molecular modeling. *RSC Adv* 2017;7(66):41512–9.
- [41] Lille Ü, Heinmaa I, Pehk T. Molecular model of Estonian kukersite kerogen evaluated by  $^{13}\text{C}$  MAS NMR spectra. *Fuel* 2003;82(7):799–804.
- [42] Tong J, Jiang X, Han X, Wang X. Evaluation of the macromolecular structure of Huadian oil shale kerogen using molecular modeling. *Fuel* 2016;181:330–9.
- [43] Mullins OC. The modified Yen model. *Energy Fuel* 2010;24(4):2179–207.
- [44] Huang Z, Liang T, Zhan Z-W, Zou Y-R, Li M, Pa P. Chemical structure evolution of kerogen during oil generation. *Mar Pet Geol* 2018;98:422–36.
- [45] Wang X, Huang X, Lin K, Zhao Y-P. The constructions and pyrolysis of 3D kerogen macromolecular models: experiments and simulations. *Global Chall* 2019;3(5): 1900006.
- [46] Gao Y, Zou Y-R, Liang T, Pa P. Jump in the structure of type I kerogen revealed from pyrolysis and  $^{13}\text{C}$  DP MAS NMR. *Org Geochem* 2017;112:105–18.
- [47] Xiang J-H, Zeng F-G, Liang H-Z, Sun B-L, Zhang L, Li M-F, et al. Model construction of the macromolecular structure of Yanzhou Coal and its molecular simulation. *J Fuel Chem Technol* 2011;39(7):481–8.
- [48] Xiang J-H, Zeng F-G, Li B, Zhang L, Li M-F, Liang H-Z. Construction of macromolecular structural model of anthracite from Chengzhuang coal mine and its molecular simulation. *J Fuel Chem Technol* 2013;41(4):391–400.
- [49] Zhang S, Wang Z, Zhang X, Wei J, Chen F. Construction of molecular structure model of Tunlan coal and its microscopic physicochemical mechanism. *Fuel* 2022; 308:121936.
- [50] Qian Y, Zhan J-H, Lai D, Li M, Liu X, Xu G. Primary understanding of non-isothermal pyrolysis behavior for oil shale kerogen using reactive molecular dynamics simulation. *Int J Hydrogen Energ* 2016;41(28):12093–100.
- [51] Han Q, Li M, Liu X, Xiao H, Ren J, Guo C. A maturation scale for molecular simulation of kerogen thermal degradation. *Org Geochem* 2022;175:104507.
- [52] Huang L, Ning Z, Wang Q, Qi R, Zeng Y, Qin H, et al. Molecular simulation of adsorption behaviors of methane, carbon dioxide and their mixtures on kerogen: effect of kerogen maturity and moisture content. *Fuel* 2018;211:159–72.
- [53] Okiongbo KS, Aplin AC, Larter SR. Changes in type II kerogen density as a function of maturity: evidence from the Kimmeridge clay formation. *Energy Fuel* 2005;19 (6):2495–9.
- [54] Pan S, Zhou H, Wang Q, Bai J, Cui D, Wang X. Experimental and molecular simulation studies of Huadian oil shale kerogen. *ACS Omega* 2022;7:17253–69.