

Three-dimensional molecular modeling of Dachengzi oil shale kerogen

Xiaoye Wang^(a), Xiangxin Han^(a), Jianhui Tong^(b), Yulong You^(a),
Xiumin Jiang^{(a)*}

^(a) School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

^(b) School of Materials Science and Engineering, Jingdezhen Ceramic Institute, Jingdezhen, 333001, China

Received 27 April 2021, accepted 22 April 2022, available online 10 June 2022

Abstract. *A three-dimensional (3D) molecular model of Dachengzi oil shale kerogen was constructed using quantitative ¹³C direct polarization/magic angle spinning solid-state nuclear magnetic resonance (DP/MAS SSNMR) data and accurate quantum chemistry methods. The heteroatom-containing functional groups were carefully identified using X-ray photoelectron spectroscopy (XPS) and pyrolysis experimental data. A large portion of C₄₈₇H₇₇₈O₄₃N₈S₅ was selected for the model to introduce more types of functional groups and make the model representative. The carbon unit fractions, structural parameters and atomic ratios of the model well matched with the experimental data. The equilibrium structure was obtained by geometry optimization using the density functional theory (DFT) method at the B3LYP/STO-3G level of theory and validated by frequency calculations at the same level. The final geometry is an incompact structure containing a large number of branches, which well reflects the cross-linked molecular structure of the kerogen. The simulated ¹³C NMR spectrum was generated using quantum chemical calculations at the B3LYP/6-31G(d) level. The simulated spectrum is in good agreement with the experimental spectrum, indicating the validity and reliability of the model.*

Keywords: *Dachengzi oil shale, kerogen, molecular modeling, quantum chemistry, simulated ¹³C NMR.*

* Corresponding author: e-mail xiuminjiang@sjtu.edu.cn

1. Introduction

Kerogen is an insoluble organic matter found in sedimentary rocks [1]. When exposed to sufficient thermal energy, kerogen in source rocks is converted into petroleum. In geologic basins, this conversion occurs at low temperatures (~90 to 120 °C) over millions of years. Oil shales are immature rocks exceptionally rich in oil-prone kerogen that have not yet generated much petroleum. Occurring throughout the world [2], oil shales have been exploited as a combustible fuel, but a more desirable product would be shale oil produced artificially in reactors at high temperatures (~500 °C) [1]. Therefore, the pyrolysis characteristics of kerogen are the core issue of oil shale utilization. Since the chemical structure of kerogen fundamentally determines its pyrolysis behavior, exploring the molecular characteristics of kerogen is essential for understanding its reaction mechanism and controlling the yield and composition of the products. However, kerogen is formed during diagenesis from the alteration and condensation of biologic precursors resulting in an extremely complex, high molecular weight macromolecule without a defined chemical composition [3], making it difficult to fully evaluate its chemical structure. Therefore, the characterization of kerogen structure often requires a combination of techniques. For this purpose, various nondestructive techniques such as solid-state nuclear magnetic resonance (SSNMR), X-ray photoelectron spectroscopy (XPS), Fourier transform infrared (FTIR), etc., have been widely used and have provided a comprehensive view of the molecular structure of kerogen [4–8]. Most notably, specific molecular structure data are frequently derived from the ^{13}C SSNMR spectra using the cross polarization/magic angle spinning (CP/MAS) technique. However, the quantification accuracy of CP has been controversial because the latter can only detect about 50% of the ^{13}C nuclei [9, 10]. Recently, quantitative direct polarization/magic angle spinning (DP/MAS) techniques have been used in the characterization of kerogen for more reliable structural information [11, 12].

In order to explain the experimental results and to visualize the structural characteristics of kerogen, scientists have proposed detailed two-dimensional (2D) molecular models of different kerogens since the 1970s [13–19]. These models cannot be considered as a representative of the actual chemical structure of kerogen, because they are based on analytical data that are an average of a large number of kerogen molecules with different structures. Nevertheless, the 2D kerogen model is still a powerful tool for synthesizing experimental results and validating the compatibility of the data from different analytical methods. In the light of 2D models, the three-dimensional (3D) molecular modeling of kerogen has progressed significantly in the last decade [20–23]. 3D models can present much more structural information compared to 2D models, such as spatial shape [24, 25], functional groups that protrude from the main body of the model as the connecting positions with other molecular structures [21, 26],

groups involved in non-bonded interactions due to spatial proximity [21, 27] and microporosity [23, 28]. These structural properties determine the interaction forms between functionalities within a kerogen molecular unit, between different kerogen molecular units, and between kerogen and its external environment such as minerals and catalysts, and can substantially affect the simulations of reactions involving kerogen, such as pyrolysis and combustion.

Simulation techniques have been applied to the 3D modeling of kerogen for geometry optimization and property prediction. Initially, geometry optimizations were mainly performed using molecular mechanics calculations due to the limitations of computational power [24, 29]. However, molecular mechanics is based on classical mechanics which is theoretically inaccurate at the atomic and subatomic scales. In contrast, quantum chemistry methods use the laws of quantum mechanics and the values of fundamental constants such as the speed of light and the masses and charges of electrons and nuclei. Therefore, they can accurately predict the structures and energies for a wide range of molecular systems. Recently, a few kerogen models optimized using quantum chemistry methods have been proposed. Based on the 2D Siskin model [15], Orendt et al. [20] generated five conformers of Green River kerogen using *ab initio* quantum chemistry methods at the HF/STO-3G level of theory, and the lowest energy conformer was validated by comparing the simulated ^{13}C NMR spectrum at the PBE1PBE/4-31G level with the experimental spectrum. Guan et al. [25] constructed 3D models of a series of Huadian kerogen isomers using the density functional theory (DFT) method at the B3LYP/STO-3G level and studied the relationship between model stability and structural parameters. It is worth mentioning that DFT methods require a similar amount of computational resources as HF methods, but can provide more accurate results because they include the effects of electron correlation which makes DFT calculations efficient and widely used in molecular simulation studies [30–33].

In previous studies, the combination of accurate quantum chemistry methods and nonquantitative ^{13}C CP/MAS SSNMR data has been used for kerogen modeling [22, 25]. However, the CP results can be very inconsistent with the elemental analysis results [24], and theoretically, the error caused by CP can significantly influence the quantum chemical calculation results because of the sensitivity of simulation to structural parameters [24, 25, 34]. Thus, accurate structural data from the quantitative DP technique are necessary for any quantum chemistry study of kerogen. Now, no molecular model of kerogen based on DP SSNMR and optimized by quantum chemistry methods has been proposed. In this study, a 3D molecular model of Dachengzi oil shale kerogen was built using quantitative ^{13}C DP/MAS SSNMR data [35]. In addition, by adding the nitrogen- and sulfur-containing functional groups obtained from XPS and pyrolysis experimental data, the model is improved compared with a previous model, which contains only carbon, hydrogen and

oxygen [4, 35, 36]. The model has a large molecular weight for quantum chemistry study, which allows a comprehensive consideration of functional groups and makes the model more representative. To avoid the inaccuracy caused by molecular mechanics calculations, the geometry optimization and frequency calculations of the 3D model were performed using quantum chemistry methods. In addition, the simulated ^{13}C NMR spectrum was generated from quantum chemical calculations to make a meaningful comparison with the experimental spectrum and to evaluate the model.

2. Modeling and computation

2.1. Model construction

The kerogen studied in this work originates from the oil shale sample from Dachengzi mine in Huadian city, China. This kerogen belongs to Type I according to the results of elemental analysis (see Table 1). The nitrogen and sulfur contents are low, indicating that there will be few nitrogen and sulfur units in the model. However, even at these low concentrations, these functionalities have a significant influence on the properties of kerogen [37, 38]. Thus, we introduced up to eight nitrogen atoms and five sulfur atoms into the model, resulting in an initial molecular weight of ~ 7850 Da and an initial chemical formula of $\text{C}_{487}\text{H}_{762}\text{O}_{60}\text{N}_8\text{S}_5$. This model size is reasonable as the molecular weight of kerogen molecular units is generally in the range of 5000 to 10000 Da [3, 16]. Moreover, this size is affordable for quantum chemical calculations.

Table 1. Elemental composition of Dachengzi oil shale kerogen from experiment and the molecular model, wt%

	Experimental data	Kerogen model
C	72.17	77.03
H	9.48	10.33
O ^a	11.77	9.06
N	1.38	1.48
S ^b	1.98	2.11
H/C ^c	1.576	1.598
O/C ^c	0.122	0.088
N/C ^c	0.016	0.016
S/C ^c	0.010	0.010

^a Organic oxygen, ^b total sulfur, ^c atomic ratio

The conditions of the quantitative ^{13}C DP/MAS SSNMR experiment of this kerogen have been reported previously in detail by Wang et al. [35]. The DP spectrum (see Fig. 1) and analytical results (see Table 2) were used to determine the types of carbon-containing functional groups and obtain accurate fractions of these groups, and seven aliphatic groups, six aromatic groups, carboxylic groups and carbonylic groups were identified. Using the chemical formula mentioned above, the number of each carbon type in the model was determined (see Table 2). Furthermore, the structural parameters of the kerogen were derived from the fractions of carbon units (see Table 2). The structural parameters are highly meaningful for modeling because they give an overall view of the kerogen structure directly. Consequently, the main carbon framework can be conveniently constructed using the structural parameters and carbon unit numbers. It is worth mentioning that the distribution of oxygen-containing functional groups was also determined from SSNMR data, which means that only the oxygen atoms connected to carbon atoms were taken into account.

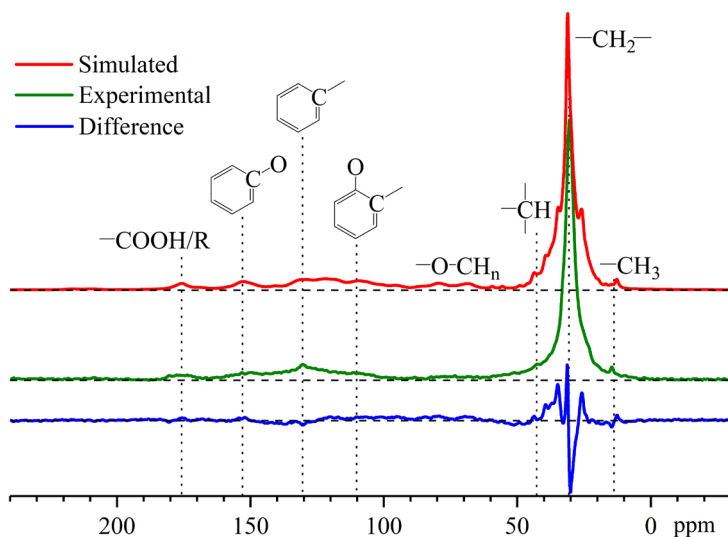


Fig. 1. Comparison between the simulated ^{13}C NMR spectrum and the experimental ^{13}C DP/MAS SSNMR spectrum [35] of Dachengzi oil shale kerogen.

Table 2. Structural data of Dachengzi oil shale kerogen from ^{13}C DP/MAS SSNMR [35] and the molecular model


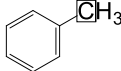



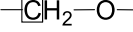
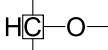
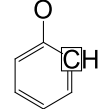
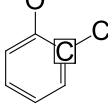
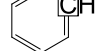
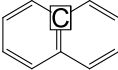
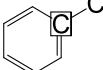
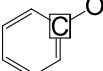


Carbon unit	Character	Location	Experiment mol% (number)	Model mol% (number)
Aliphatic methyl	f_{al}^1		0.74 (4)	0.82 (4)
Aromatic methyl	f_{al}^{a}		0.14 (1)	0.21 (1)
Chain-end methylene	$f_{\text{al}}^{2\text{t}}$		0.32 (2)	0.82 (4)
Methylene	f_{al}^2		66.79 (325)	66.12 (322)
Methine	f_{al}^3		4.34 (21)	4.72 (23)
Oxy-methylene	$f_{\text{al}}^{\text{O}2}$		2.34 (11)	2.46 (12)
Oxy-methine	$f_{\text{al}}^{\text{O}3}$		2.69 (13)	2.26 (11)
Ortho-oxyaromatic protonated	f_{a}^{HO}		3.80 (19)	3.70 (18)
Ortho-oxyaromatic branched	f_{a}^{SO}		2.11 (11)	2.26 (11)
Aromatic protonated	f_{a}^{H}		3.14 (15)	3.08 (15)
Bridging ring junction	f_{a}^{B}		2.98 (14)	3.29 (16)
Aromatic branched	f_{a}^{S}		4.29 (21)	4.11 (20)
Oxy-aromatic	f_{a}^{O}		3.91 (18)	3.70 (18)
Carboxylic	f_{COO}		1.72 (9)	1.85 (9)
Carbonylic	f_{CO}		0.69 (3)	0.62 (3)

Table 2 (continued)

Structural parameter	Definition	Experiment mol% (number)	Model mol% (number)
Aromaticity	$f_a = f_a^{\text{HO}} + f_a^{\text{SO}} + f_a^{\text{H}} + f_a^{\text{B}} + f_a^{\text{S}} + f_a^{\text{O}}$	20.23%	20.12%
Aliphaticity	$f_{\text{al}} = f_{\text{al}}^1 + f_{\text{al}}^{\text{a}} + f_{\text{al}}^{2\text{t}} + f_{\text{al}}^2 + f_{\text{al}}^3 + f_{\text{al}}^{\text{O}2} + f_{\text{al}}^{\text{O}3}$	77.37%	77.41%
Average aliphatic carbon chain length	$f_{\text{al}} / (f_a^{\text{SO}} + f_a^{\text{S}} + f_a^{\text{O}}) \leq Cn \leq f_{\text{al}} / (f_a^{\text{SO}} + f_a^{\text{S}})$	7.5–12.1	7.9
Substitute degree of aromatic rings	$\sigma = (f_a^{\text{SO}} + f_a^{\text{S}} + f_a^{\text{O}}) / f_a$	0.51	0.5
Average carbons per aromatic cluster	$C = 6f_a / (f_a - 2f_a^{\text{B}})$	8.5	8.9

The XPS data published by Tong et al. [4] was used to determine the distribution of nitrogen- and sulfur-containing groups. For nitrogen-containing groups, pyrrolic, pyridinic and amino groups were introduced into the model. For sulfur-containing groups, aromatic sulfur, aliphatic sulfur and sulfone/sulfoxide groups were introduced into the model. Other nitrogen- and sulfur-containing groups were not considered due to their low contents.

Combining the carbon framework and the heteroatom-containing groups, an initial 2D kerogen model was obtained. The molecular structure was adjusted to make the model parameters as close to the experimental results as possible. During the adjustment, the aromaticity was matched with the experimental data, and at the same time, the H/C atomic ratio was made as close to the elemental analysis result as possible, which, however, was difficult because these two parameters negatively correlate.

2.2. Geometry optimization

All the quantum chemical calculations were conducted on the PI supercomputer at Shanghai Jiao Tong University, using the Gaussian 09 program [39] and a compute node configured with 16 CPU cores and 256 GB of memory. The total computing time for the optimization and frequency jobs was ~11 days. Based on the adjusted 2D molecular kerogen model, an initial 3D model was constructed using the GaussView 5.0.9 visualization program. The initial model is not a stable conformation and needs to be optimized to an equilibrium structure corresponding to a potential energy minimum. The geometry optimization was performed using the density functional theory (DFT) method at the B3LYP [40–42]/STO-3G [43, 44] level of theory. The B3LYP method is popular in quantum chemical calculations because it can provide accurate results for a wide range of molecular systems at a modest

computational cost. The small basis set STO-3G was used to keep the cost affordable for such a large model size. In addition, the default convergence criteria of the Gaussian program were used for the geometry optimization. In order to identify whether the optimized model is located at a minimum on the potential energy surface, we performed frequency calculations at the same level of theory as the optimization and checked the frequencies.

2.3. Simulation of ^{13}C NMR spectrum

The NMR calculations were carried out at the B3LYP/6-31G(d) level and the absolute shielding values were subtracted from that of tetramethylsilane (TMS) calculated at the same level to obtain the chemical shifts. The simulated ^{13}C NMR spectrum was generated using 1.5 ppm Lorentzian broadening in the aliphatic region and 5 ppm Lorentzian broadening in the aromatic region for the chemical shifts.

3. Results and discussion

3.1. Molecular model

The 2D molecular model of Dachengzi kerogen is shown in Figure 2 and the corresponding initial 3D model is shown in Figure 3a. The model has a composition of $\text{C}_{487}\text{H}_{778}\text{O}_{43}\text{N}_8\text{S}_5$ and a molecular weight of 7594 Da. It is neutral and has no unpaired electrons. Four grid structures composed of aliphatic chains and aromatic clusters were introduced into the model to present the intramolecular cross-linked structure. Moreover, a large number of open chains were determined based on the number of chain-end methylene groups.

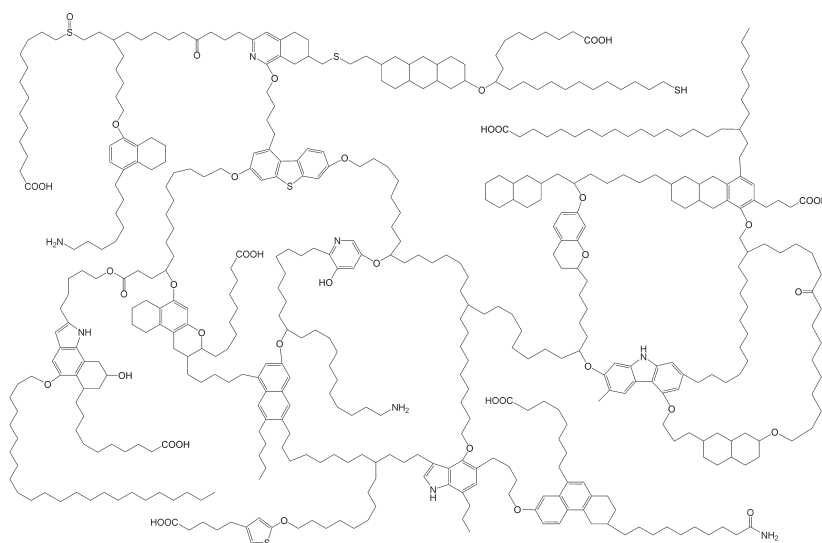


Fig. 2. 2D molecular model of Dachengzi oil shale kerogen. The chemical formula is $\text{C}_{487}\text{H}_{778}\text{O}_{43}\text{N}_8\text{S}_5$.

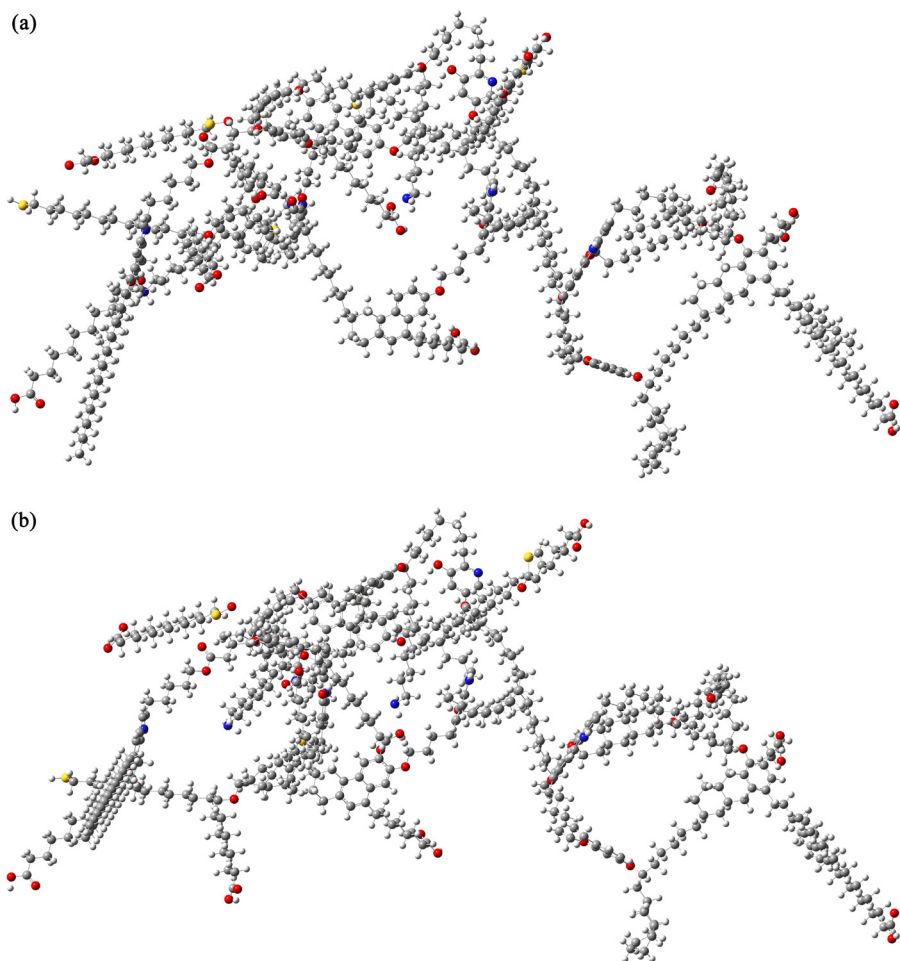


Fig. 3. 3D molecular model of Dachengzi oil shale kerogen before (a) and after (b) geometry optimization.

As shown in Table 2, the carbon-containing functional group fractions and the structural parameters of the model are in good agreement with the experimental data, indicating that the model can accurately describe the carbon framework of the kerogen. In order to balance the high content of methine groups and the long chain characteristics, several saturated cycles were introduced into the model. Most carboxylic groups ($-\text{COOH}/\text{R}$) in the model were set to carboxyl groups ($-\text{COOH}$) to make the fractions of oxyaliphatic and methyl groups in the model consistent with the SSNMR results, while the others were set to ester groups ($-\text{COOR}$). The carboxyl groups were located at chain ends, which represent the connecting positions of the kerogen molecule with the environment [26].

The atomic ratios (see Table 1) of the model match the elemental analysis results fairly well, except that the O/H atomic ratio is slightly lower than the experimental value. The difference in O/H atomic ratio is probably caused by the SSNMR measurements underestimating the content of the oxygen-containing groups, which are frequently connected to mineral components [26] and adjacent to paramagnetic centers [9]. As a consequence, the polarizations of these groups decay extremely fast and their signals are weakened. Since there is usually a negative correlation between the H/C atomic ratio and the aromaticity of kerogen, the simultaneous match of these two parameters indicates the reliability of the DP SSNMR data.

The nitrogen- and sulfur-containing functional groups in the model are shown in Table 3. According to the XPS results proposed by Tong et al. [4], the ratio of pyrrolic, pyridinic and amino groups is about 3:1:2, and the ratio of aromatic sulfur, aliphatic sulfur and sulfone/sulfoxide groups is about 4:3:5. We slightly increased the fraction of pyridinic groups for a better correlation between the model and shale oil products, because the results of electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) reported by Tong et al. [38] show a high concentration of pyridinic compounds in the Dachengzi shale oil. More specifically, the ESI FT-ICR MS results indicate that pyridines and tetrahydroquinolines are dominant among the basic nitrogen compounds, and indoles and carbazoles are the major components of the neutral nitrogen compounds. In addition, an amide group was introduced into the model because a large amount of amides has been detected in shale oil and shale char [38, 45].

Table 3. Heteroatom-containing functional groups in Dachengzi oil shale kerogen molecular model


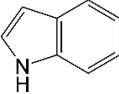
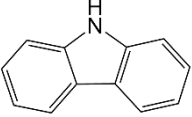
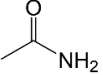

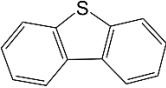


Group	Structure	Number
Pyridine		2
Indole		2
Carbazole		1
Amino	—NH_2	2
Amide		1

Table 3 (continued)

Group	Structure	Number
Thiophene		1
Dibenzothiophene		1
Aliphatic sulfur		2
Sulfoxide		1

It is noteworthy that the aliphatic sulfur and sulfone/sulfoxide groups are the major precursors of H_2S and SO_2 in oil shale pyrolysis experiments [36, 46, 47]. However, the fractions of these groups from XPS are inconsistent with the pyrolysis results of Dachengzi oil shale, which show that H_2S is much more abundant among the gaseous products than SO_2 [36]. Considering that the XPS results may be affected by the oxidation of surface functional groups, we reduced the number of sulfone/sulfoxide groups to one in order to better match the experimental results. Furthermore, pyrolysis experiments [36] show that most of the sulfur in the Dachengzi oil shale transfers to the non-condensable gas and semi-coke, and the two fractions are similar, indicating that the fractions of aliphatic sulfur and aromatic sulfur in the kerogen are similar because the sulfur-containing compounds in non-condensable gas and semi-coke are mainly from the aliphatic and aromatic sulfur in the kerogen, respectively. Thus, the number of both aromatic and aliphatic sulfur functional groups in the model was set to two. The specific forms of aromatic sulfur groups in the model were determined to be thiophenes and dibenzothiophenes according to the types of sulfur-containing compounds in Dachengzi shale oil [36].

3.2. Equilibrium structure from quantum chemistry

The geometry optimization of the kerogen model generated 3513 basis functions and converged in 234 steps. The energies of the model before and after optimization were -24390.4607 and -24390.7136 Hartrees, respectively. The final equilibrium structure is shown in Figure 3b. Unlike the kerogen models optimized using molecular mechanics or semi-empirical methods [22, 24], the present model did not distinctly fold during the optimization process. The final geometry is still a relatively incompact structure and only slightly different from the initial model. This is a natural result of the quantum chemistry methods searching for a local energy minimum near the initial structure. Consequently, the present model contains a large number

of branches, which well reflect the cross-linked molecular structure of the kerogen and represent the connections with other organic units and mineral matrix, indicating the reliability of quantum chemical calculations. In addition, the Type I kerogen model proposed by Bousige et al. [23], which is generated from an initial configuration containing a random distribution of atoms using a molecular dynamics-hybrid reverse Monte Carlo (MD-HRMC) reconstruction method, also has the same structural features.

The frequency calculations yield 3957 frequencies, among which are 21 imaginary frequencies between -43 and 0 cm^{-1} . In fact, it is very difficult to avoid low imaginary frequencies for large molecular systems with many flexible long chains like the present model. However, Ungerer et al. [22] have discussed imaginary frequencies of large flexible molecules and considered low frequencies (-50 cm^{-1} to 0 cm^{-1}) to have a minor influence on the model properties. In addition, low imaginary frequencies are usually associated with the whole vibration of long aliphatic carbon chains, which is exactly the same as the result from checking vibration modes using the GaussView program, and have a negligible influence on the stability of the functional groups and the entire model.

3.3. Model validation

Since the chemical shifts of ^{13}C nuclei are mainly related to hybridization and substituent electronegativity [48], it is predictable that the kerogen model constructed from the carbon unit data in Table 2 can reproduce the general characteristics of the experimental ^{13}C DP/MAS SSNMR spectrum via theoretical calculations. However, other effects such as long-range interaction, γ -gauche steric compression and heteroatoms, which depend on the uncertainties of the positional arrangement of functional groups and the spatial configuration of the model, also move the chemical shifts, especially for large-sized molecular systems [48]. Moreover, quantum chemistry methods can better introduce these effects into NMR spectrum simulations than empirical methods [20, 24]. As a result, the simulated NMR spectrum from quantum chemical calculations can differ from the experimental spectrum, and comparing the simulated spectrum with the experimental spectrum is still important for evaluating the consistency between the model and the true chemical structure of the kerogen.

The simulated ^{13}C NMR spectrum was generated by overlapping the broadened chemical shifts from quantum chemical calculations, as shown in Figure 1. In order to validate the kerogen model, we compared the simulated ^{13}C NMR spectrum with the experimental ^{13}C DP/MAS SSNMR spectrum (see Fig. 1). Overall, the band intensities and line shapes of the carbon-containing functional groups in the simulated spectrum are in good agreement with those in the experimental spectrum, indicating that the structural features of the kerogen are well reproduced by the model. The aromaticity obtained

by integrating the simulated band intensities in the range of 90–170 ppm is 19.5%, which is very close to the experimental result (20.23%).

For a better comparison, the total band area of the simulated spectrum was adjusted to be equal to that of the experimental spectrum, and the difference spectrum was obtained, as shown in Figure 1. The difference in the aromatic region is close to zero, indicating that the distribution of aromatic groups is well matched. In the aliphatic region, the intensities of the oxy-aliphatic carbon bands and the edge of the main aliphatic band in the simulated spectrum are slightly stronger than those in the experimental spectrum. After checking the simulated chemical shift distribution in the GaussView program, it was found that heteroatoms play an important role in this difference. For example, the chemical shifts of ortho-oxyaromatic protonated carbon nuclei occur at lower frequencies, i.e. to the left, because the adjacent nitrogen atoms and the chemical shifts of methylene carbon nuclei occur at higher frequencies due to the adjacent sulfur atoms, which increases the intensity of the oxy-aliphatic carbon band. The two peaks in the range of about 26–33 ppm in the difference spectrum are mainly derived from a small difference (~0.5 ppm) in the positions of the narrow methylene peaks in the experimental and simulated spectra, which is caused by the systematic error in the computational method. Therefore, these two peaks do not indicate any significant difference between the signal distributions of the simulated and experimental spectra, and thus the peaks may be ignored.

4. Conclusions

In this study, a 3D Dachengzi kerogen model was constructed based on quantitative ^{13}C DP/MAS SSNMR data. The carbon unit fractions, structural parameters and atomic ratios of the model all match well with the experimental data. According to the XPS and pyrolysis experiment results from previous works, heteroatom-containing functional groups were identified in detail, including nitrogen-containing functional groups (pyridine, indole, carbazole, amino and amide groups) and sulfur-containing functional groups (thiophene, dibenzothiophene, aliphatic sulfur and sulfoxide groups). The model has a composition of $\text{C}_{487}\text{H}_{778}\text{O}_{43}\text{N}_8\text{S}_5$ and a large molecular weight of 7594 Da, which allows more types of functional groups to be introduced into the model and makes it more representative.

Geometry optimization was performed on the 3D kerogen model to obtain an equilibrium structure, using the DFT method at the B3LYP/STO-3G level of theory. The final geometry is an incompact structure containing a large number of branches, which is distinctly different from the compact models generated by molecular mechanics calculations but is similar to the true molecular structure of the kerogen, indicating the reliability of the quantum chemistry method. The optimized model was verified as a conformation

corresponding to a potential energy minimum by frequency calculations at the same level of theory.

In order to evaluate the model, the simulated ^{13}C NMR spectrum was generated at the B3LYP/6-31G(d) level and compared with the experimental spectrum. The intensities and line shapes of the carbon bands in the simulated spectrum are in good agreement with the experimental results, indicating the validity and reliability of the model.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 51876122, No. 51766004).

REFERENCES

1. Tissot, B. P., Welte, D. H. *Petroleum Formation and Occurrence*. Springer-Verlag, New York, 1984.
2. Knaus, E., Killen, J., Biglarbigi, K., Crawford, P. An overview of oil shale resources. In: *Oil Shale: A Solution to the Liquid Fuel Dilemma* (Ogunsola, O. I., Hartstein, A. M., Ogunsola, O., eds.). ACS Symposium Series, **1032**, American Chemical Society, Washington, DC, 2010, 3–20.
3. Vandenbroucke, M., Largeau, C. Kerogen origin, evolution and structure. *Org. Geochem.*, 2007, **38**(5), 719–833.
4. Tong, J. H., Han, X. X., Wang, S., Jiang, X. M. Evaluation of structural characteristics of Huadian oil shale kerogen using direct techniques (solid-state ^{13}C NMR, XPS, FT-IR, and XRD). *Energy Fuels*, 2011, **25**(9), 4006–4013.
5. Fletcher, T. H., Gillis, R., Adams, J., Hall, T., Mayne, C. L., Solum, M. S., Pugmire, R. J. Characterization of macromolecular structure elements from a Green River oil shale, II. Characterization of pyrolysis products by ^{13}C NMR, GC/MS, and FTIR. *Energy Fuels*, 2014, **28**(5), 2959–2970.
6. Kelemen, S. R., Freund, H., Gorbaty, M. L., Kwiatek, P. J. Thermal chemistry of nitrogen in kerogen and low-rank coal. *Energy Fuels*, 1999, **13**(2), 529–538.
7. Kelemen, S. R., Afeworki, M., Gorbaty, M. L., Sansone, M., Kwiatek, P. J., Walters, C. C., Freund, H., Siskin, M., Bence, A. E., Curry, D. J., Solum, M., Pugmire, R. J., Vandenbroucke, M., Leblond, M., Behar, F. Direct characterization of kerogen by X-ray and solid-state ^{13}C nuclear magnetic resonance methods. *Energy Fuels*, 2007, **21**(3), 1548–1561.
8. Bansal, V. R., Kumar, R., Sastry, M. I. S., Badhe, R. M., Kapur, G. S., Saxena, D. Direct estimation of shale oil potential by the structural insight of Indian origin kerogen. *Fuel*, 2019, **241**, 410–416.
9. Wind, R. A., Maciel, G. E., Botto, R. E. Quantitation in ^{13}C NMR spectroscopy of carbonaceous solids. In: *Magnetic Resonance of Carbonaceous Solids*

- (Botto, R. E., Sanada, Y., eds.). ACS Advances in Chemistry, No. 229, American Chemical Society, Washington, DC, 1992, 3–26.
10. Smernik, R. J., Schwark, L., Schmidt, M. W. I. Assessing the quantitative reliability of solid-state ^{13}C NMR spectra of kerogens across a gradient of thermal maturity. *Solid State Nucl. Magn. Reson.*, 2006, **29**(4), 312–321.
 11. Mao, J. D., Fang, X. W., Lan, Y. Q., Schimmelmann, A., Mastalerz, M., Xu, L., Schmidt-Rohr, K. Chemical and nanometer-scale structure of kerogen and its change during thermal maturation investigated by advanced solid-state ^{13}C NMR spectroscopy. *Geochim. Cosmochim. Acta*, 2010, **74**(7), 2110–2127.
 12. Gao, Y., Zou, Y. R., Liang, T., Peng, P. A. Jump in the structure of Type I kerogen revealed from pyrolysis and ^{13}C DP MAS NMR. *Org. Geochem.*, 2017, **112**, 105–118.
 13. Yen, T. F. Structural aspects of organic components in oil shales. In: *Developments in Petroleum Science* (Yen, T. F., Chilingarian, G. V., eds.). Elsevier, Oxford, 1976, **5**, 129–148.
 14. Behar, F., Vandembroucke, M. Chemical modelling of kerogens. *Org. Geochem.*, 1987, **11**(1), 15–24.
 15. Siskin, M., Scouten, C. G., Rose, K. D., Aczel, T., Colgrove, S. G., Pabst Jr., R. E. Detailed structural characterization of the organic material in Rundle Ramsay Crossing and Green River oil shales. In: *Composition, Geochemistry and Conversion of Oil Shales* (Snape, C. E., ed.). Springer Netherlands, Dordrecht, 1995, 143–158.
 16. Lille, Ü., Heinmaa, I., Pehk, T. Molecular model of Estonian kukersite kerogen evaluated by ^{13}C MAS NMR spectra. *Fuel*, 2003, **82**(7), 799–804.
 17. Wang, Q., Hou, Y. C., Wu, W. Z., Yu, Z., Ren, S. H., Liu, Q. Y., Liu, Z. Y. A study on the structure of Yilan oil shale kerogen based on its alkali-oxygen oxidation yields of benzene carboxylic acids, ^{13}C NMR and XPS. *Fuel Process. Technol.*, 2017, **166**, 30–40.
 18. Huang, Z. K., Liang, T., Zhan, Z. W., Zou, Y. R., Li, M. W., Peng, P. A. Chemical structure evolution of kerogen during oil generation. *Mar. Petrol. Geol.*, 2018, **98**, 422–436.
 19. Chu, W. Y., Cao, X. Y., Schmidt-Rohr, K., Birdwell, J. E., Mao, J. D. Investigation into the effect of heteroatom content on kerogen structure using advanced ^{13}C solid-state nuclear magnetic resonance spectroscopy. *Energy Fuels*, 2019, **33**(2), 645–653.
 20. Orendt, A. M., Pimienta, I. S. O., Badu, S. R., Solum, M. S., Pugmire, R. J., Facelli, J. C., Locke, D. R., Chapman, K. W., Chupas, P. J., Winans, R. E. Three-dimensional structure of the Siskin Green River oil shale kerogen model: A comparison between calculated and observed properties. *Energy Fuels*, 2013, **27**(2), 702–710.
 21. Katti, D. R., Thapa, K. B., Katti, K. S. Modeling molecular interactions of sodium montmorillonite clay with 3D kerogen models. *Fuel*, 2017, **199**, 641–652.
 22. Ungerer, P., Collell, J., Yiannourakou, M. Molecular modeling of the volumetric and thermodynamic properties of kerogen: Influence of organic type and maturity. *Energy Fuels*, 2015, **29**(1), 91–105.

23. Bousige, C., Ghimbeu, C. M., Vix-Guterl, C., Pomerantz, A. E., Suleimenova, A., Vaughan, G., Garbarino, G., Feygenson, M., Wildgruber, C., Ulm, F.-J., Pellenq, R. J. M., Coasne, B. Realistic molecular model of kerogen's nanostructure. *Nat. Mater.*, 2016, **15**(5), 576–582.
24. Tong, J. H., Jiang, X. M., Han, X. X., Wang, X. Y. Evaluation of the macromolecular structure of Huadian oil shale kerogen using molecular modeling. *Fuel*, 2016, **181**, 330–339.
25. Guan, X. H., Liu, Y., Wang, D., Wang, Q., Chi, M. S., Liu, S., Liu, C. G. Three-dimensional structure of a Huadian oil shale kerogen model: An experimental and theoretical study. *Energy Fuels*, 2015, **29**(7), 4122–4136.
26. Razvigorova, M., Budinova, T., Tsyntsarski, B., Petrova, B., Ekinci, E., Atakul, H. The composition of acids in bitumen and in products from saponification of kerogen: Investigation of their role as connecting kerogen and mineral matrix. *Int. J. Coal Geol.*, 2008, **76**(3), 243–249.
27. Marzec, A. Intermolecular interactions of aromatic hydrocarbons in carbonaceous materials: A molecular and quantum mechanics. *Carbon*, 2000, **38**(13), 1863–1871.
28. Tesson, S., Firoozabadi, A. Methane adsorption and self-diffusion in shale kerogen and slit nanopores by molecular simulations. *J. Phys. Chem. C*, 2018, **122**(41), 23528–23542.
29. Xu, F., Liu, H., Wang, Q., Pan, S., Zhao, D., Liu, Q., Liu, Y. ReaxFF-based molecular dynamics simulation of the initial pyrolysis mechanism of lignite. *Fuel Process. Technol.*, 2019, **195**, 106147.
30. Pan, S., Wang, Q., Bai, J. R., Chi, M. S., Cui, D., Wang, Z. C., Liu, Q., Xu, F. Molecular structure and electronic properties of oil shale kerogen: An experimental and molecular modeling study. *Energy Fuels*, 2018, **32**(12), 12394–12404.
31. Zou, C. Y., Raman, S., van Duin, A. C. T. Large-scale reactive molecular dynamics simulation and kinetic modeling of high-temperature pyrolysis of the *Gloeocapsomorphaprisca* microfossils. *J. Phys. Chem. B*, 2014, **118**(23), 6302–6315.
32. Zhu, Y. Q., Su, H., Jing, Y., Guo, J. C., Tang, J. L. Methane adsorption on the surface of a model of shale: A density functional theory study. *Appl. Surf. Sci.*, 2016, **387**, 379–384.
33. Zhang, H., Liu, J. X., Wang, X. Y., Jiang, X. M. Density functional theory study on two different oxygen enhancement mechanisms during NO–char interaction. *Combust. Flame*, 2016, **169**, 11–18.
34. Rogel, E. Simulation of interactions in asphaltene aggregates. *Energy Fuels*, 2000, **14**(3), 566–574.
35. Wang, X. Y., You, Y. L., Mu, M., Han, X. X., Shu, J., Jiang, X. M. Structural characterization of Huadian oil shale kerogen by using ¹³C DP/MAS NMR. *Oil Shale*, 2021, **38**(3), 181–198.
36. Tong, J. H. *Behavior and Mechanism of Transformation of Nitrogen and Sulfur Elements in Oil Shale in the Course of Comprehensive Utilization*. PhD dissertation. Shanghai Jiao Tong University, 2013 (in Chinese).

37. Jiang, X. M., Yu, L. J., Yan, C., Han, X. X., Yan, H. L. Experimental investigation of SO₂ and NO_x emissions from Huadian oil shale during circulating fluidized-bed combustion. *Oil Shale*, 2004, **21**(3), 249–257.
38. Tong, J. H., Liu, J. G., Han, X. X., Wang, S., Jiang, X. M. Characterization of nitrogen-containing species in Huadian shale oil by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Fuel*, 2013, **104**, 365–371.
39. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J. *Gaussian 09, Revision B.01*, Gaussian, Inc., Wallingford CT, 2009.
40. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 1988, **38**(6), 3098–3100.
41. Lee, C., Yang, W., Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter*, 1988, **37**(2), 785–789.
42. Miehlich, B., Savin, A., Stoll, H., Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.*, 1989, **157**(3), 200–206.
43. Hehre, W. J., Stewart, R. F., Pople, J. A. Self-consistent molecular-orbital methods. I. Use of gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.*, 1969, **51**(6), 2657–2664.
44. Collins, J. B., von R. Schleyer, P., Binkley, J. S., Pople, J. A. Self-consistent molecular orbital methods. XVII. Geometries and binding energies of second-row molecules. A comparison of three basis sets. *J. Chem. Phys.*, 1976, **64**(12), 5142–5151.
45. Han, X. X., Jiang, X. M., Cui, Z. G., Liu, J. G., Yan, J. W. Effects of retorting factors on combustion properties of shale char. 3. Distribution of residual organic matters. *J. Hazard. Mater.*, 2010, **175**(1–3), 445–451.
46. Burnham, A. K., Bey, N. K., Koskinas, G. J. Hydrogen sulfide evolution from Colorado oil shale. In: *Oil Shale, Tar Sands, and Related Materials* (Stauffer, H. C., ed.). ACS Symposium Series, **163**, American Chemical Society, Washington, DC, 1981, 61–77.

47. Davis, F. A., Panunto, T. W., Awad, S. B., Billmers, R. L., Squires, T. G. Pyrolysis of organic compounds. 1. Flash vacuum pyrolysis (FVP) of coal-model organic sulfides and their S-oxides. *J. Org. Chem.*, 1984, **49**(7), 1228–1230.
48. Silverstein, R. M., Webster, F. X., Kiemle, D. J. *Spectrometric Identification of Organic Compounds*. 7th ed.; John Wiley & Sons, Inc., New York, 2005.