

# Large-scale atomistic model construction of subbituminous and bituminous coals for solvent extraction simulations with reactive molecular dynamics<sup>☆</sup>

Pilsun Yoo<sup>a,\*</sup>, Gang Seob Jung<sup>a</sup>, Matthew R. Ryder<sup>b</sup>, Frederic Vautard<sup>c</sup>, Ercan Cakmak<sup>b</sup>, Sungsool Wi<sup>d</sup>, Matthew C. Weisenberger<sup>e</sup>, Edgar Lara-Curzio<sup>f</sup>, Jonathan P. Mathews<sup>g,\*\*</sup>, Stephan Irle<sup>a,\*\*\*</sup>

<sup>a</sup> Computational Sciences and Engineering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37830, USA

<sup>b</sup> Materials Science and Technology Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37830, USA

<sup>c</sup> Chemical Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

<sup>d</sup> National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL, 32310, USA

<sup>e</sup> Center for Applied Energy Research, University of Kentucky, Lexington, KY, 40511, USA

<sup>f</sup> Energy Science and Technology Directorate, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, 37830, USA

<sup>g</sup> The Leone Family Department of Energy and Minerals Engineering, and the EMS Energy Institute, The Pennsylvania State University, 126 Hosler Building, University Park, PA, 16802, USA

## ARTICLE INFO

### Keywords:

Bottom-up carbonaceous model construction  
Atomistic coal model  
Solvent extraction  
Reactive molecular dynamics simulations of coal  
Ensemble of macromolecules

## ABSTRACT

Large-scale atomistic models for complex polycyclic aromatic hydrocarbon systems help understand the chemical properties and behaviors of complex feedstocks such as coal or petroleum. However, the development and utilization of large-scale models remain limited due to the difficulty in achieving the varied structural characteristics necessary to capture stochastic nature of these feedstocks. We demonstrate a systematic workflow to construct stochastic molecular systems from a broad analytical suite: high-resolution transmission electron microscopy (HRTEM), carbon-13 nuclear magnetic resonance spectroscopy (<sup>13</sup>C NMR), laser desorption/ionization mass spectroscopy (LDI-MS), and elemental analysis. We present a model construction and analysis utility of a new Python-based module. We selected one subbituminous and three high-volatile bituminous coals to construct large-scale models (~40,000 atoms). The constructed models were utilized to examine the affinity for solvent extraction (naphthalene or tetralin) and the effect of structural properties (e.g., aromatic cluster size, functional groups, and cross-linking) in reactive molecular dynamics simulations. Complex chemical reactions were monitored with bond order transitions, intermediates formation, and mass distributions. Reactive molecular dynamics simulations suggest a plausible chemical extraction process and products for the complex fossil feedstocks. The results indicated that radical formations with bond breaking of bridging oxygens and carbons were required at high temperatures to facilitate hydrogenation and extraction of gas molecules from radical-free molecules. We observed that aliphatic chains of tetralin were easily decomposed and combined with radicals to form small size of molecules with aryl bonding, mainly increasing molecules in the 500–1000 Da, while naphthalene had little impact on chemical extraction process.

<sup>☆</sup> NOTICE OF COPYRIGHT: This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

\* Corresponding author.

\*\* Corresponding author.

\*\*\* Corresponding author.

E-mail addresses: [yoop@ornl.gov](mailto:yoop@ornl.gov) (P. Yoo), [jpm10@psu.edu](mailto:jpm10@psu.edu) (J.P. Mathews), [irles@ornl.gov](mailto:irles@ornl.gov) (S. Irle).

<https://doi.org/10.1016/j.carbon.2024.118939>

Received 16 November 2023; Received in revised form 6 February 2024; Accepted 14 February 2024

Available online 14 February 2024

0008-6223/© 2024 Elsevier Ltd. All rights reserved.

## 1. Introduction

Coal is a heterogeneous and complex material comprising various organic molecules and inorganic minerals. The organic matter of coal is made of carbonaceous material that originates from conversions of deposited ancient plants under high pressure and heat [1]. Since the Industrial Revolution, coal has primarily been used for electricity generation through combustion. Although coal remains a vital electricity resource due to its availability and abundance, its usage will gradually change because of environmental concerns about greenhouse gas emissions associated with coal combustion [2] and the opportunities for using coal as a feedstock for manufacturing value-added carbon products such as graphite and carbon fibers. Understanding and utilizing coal compounds for non-power generation and efficient production of value-added products are becoming more critical [3].

However, the complexity and heterogeneity of coal have hindered a thorough characterization of its nature and optimal chemical conversion routes to value-added products. Traditionally, coal has been classified into four ranks depending on the amount of volatile matter or calorific value: lignite, subbituminous, bituminous, and anthracite [4]. Also, the maceral contributions (biomass input and depositional environment-related) influence the different conversion properties due to their chemical structural variations [5]. It is reasonable to assume that the varying ranks and maceral types of coals require different conversion routes, depending on their distinct chemical and physical properties [6–8].

Efforts over the past decades led to an increased understanding of coal characteristics, such as rank, maceral composition, morphology, molecular structure, and compounds of raw materials at certain maturation, owing mostly to advancements in analytical techniques. It is now widely accepted that coal is a collection of three-dimensional networks of macromolecules condensed by combinations of covalent and non-covalent bonds and associated molecules [1,9–13]. Highly informative characterization approaches are proton nuclear magnetic resonance ( $^1\text{H}$  NMR) and solid state carbon-13 nuclear magnetic resonance ( $^{13}\text{C}$  NMR) spectroscopy, which provide molecular structures with aromatic, aliphatic contents, skeletal structure, and hydrogen attachments data [14,15]. Information obtained from mass spectroscopies [16–19] and lattice fringe analysis [20,21] with high-resolution transmission electron microscopy (HRTEM) has also given rise to greater insights into molecular size distributions and subunits constituting macromolecules. Furthermore, X-ray absorption spectroscopy and small-angle neutron scattering (SANS) analysis have provided detailed information about compositions and pore distributions [22–24].

The utilization of advanced experimental techniques and improved computational power has accelerated the examination of coal with atomistic computational models to obtain insight into chemical structures and conversion mechanisms [25–36]. By integrating findings from experimental characterization techniques, models have advanced to encompass a broad spectrum of molecular types and sizes, tailored to precise chemical compositions and mass distributions. Recent research has underscored the significance of comprehensive coal specimen characterizations in facilitating the development of atomistic models at a broader scale, encompassing ranks, maceral types, and compositions [27,37,38]. The complexity and scales of coal models have been significantly increased by combining 5000–100,000 atoms using 2–400 molecules to study various properties further originating from their structural and chemical characteristics for both collections of molecules and individual molecule using quantum mechanics (QM), molecular mechanics (MM) and molecular dynamics (MD) simulations [31,34,39–41].

In recent years, density functional theory (DFT) calculations of small molecular models with functional groups or substituted aromatic molecules have provided insight into experimental results for chemical processes of coal molecules. For example, Hou et al. reported energetic barriers of hydrogen transfer and conversion of radicals in coal with

simplified models by bridging small hydrocarbon molecules to find optimal hydrogen-donor solvents in direct coal liquefaction using B3LYP DFT method [42]. Similarly, Cheng and his coworkers studied the chemisorption property of oxygen on coal models for energetic preferences and possible chemisorption configurations [43].

However, coal molecules are, in principle, complex and variable in size and complexity. Hence, investigations of single-molecule properties using DFT are certainly oversimplified, as they exclude steric constraints, entropic configurational and conformational effects, and the overall collective effects of structural components such as van der Waals, hydrogen bonding, and electrostatic interactions. In an attempt to incorporate such collective effects, Li et al. investigated the sulfur transformation mechanism of lignite during the oxidation process with high-temperature reactive MD simulations using a macromolecule ( $\text{C}_{225}\text{H}_{182}\text{N}_4\text{O}_{36}\text{S}_3$ ) derived from Wolfrum's model followed by DFT calculations to confirm reaction paths for the formation of sulfur oxide radicals and sulfur-containing gas [44]. Such studies lead to a wide range of chemical products and reaction mechanism pathways that should be considered when theoretically investigating the chemical conversion properties of coal to inform experimental studies [29,31,32,45,46].

The application of large-scale atomistic models has also been expedited through MM with intermolecular interactions and computations of chemical properties of molecules, including stacking [21], pore size distribution [41], solubility parameters [47], and solvent swelling [48]. Mathews and his team have demonstrated that large-scale atomistic models, constructed with a suite of experimental data, elucidate the specific fractions of coals that can be extracted and enable an exploration of the associated solvent swelling properties in their series of reports [29,48–51]. Moreover, large-scale atomistic models have been actively utilized for investigations of chemical reaction processes, including pyrolysis/combustion and oxidation, mainly using reactive MD simulations based on the reactive force field (ReaxFF) parameters [31,32,44,45,52]. These studies have clearly demonstrated their advantages, as they provide more realistic predictions of intermediates and products that cannot be directly captured in experimental processes and offer insight into chemical reaction paths from structural motifs and functional groups for ensembles of molecular compounds [53–55]. Hence, the construction and application of large-scale atomistic models consistent with experimental characterizations should help design an improved process for engineering coal-derived compounds.

One of the most fundamental challenges in the chemical conversion of coals, other than electricity generation, is the effective extraction and breakdown of desired carbonaceous compounds from which value-added products can be manufactured. Solvent extraction is one common approach to retrieving useful subfractions. Since the first coal extraction report by Pott and Broche in 1934, tetralin (hydrogen donor solvent) has been known to be a suitable solvent for coal digestion and extraction, dissolving most coal fractions [56]. It is important to note that the role of solvents for coal extraction is not limited to liquefying solids by separating associate interactions between macromolecules. Solvents also chemically modify molecules through hydrogen donation, bond breaking, and decomposition [8,57,58]. This fact imposes significant cost factors, hampering the scalability of the solvent extraction process for industrial purposes.

Therefore, it is important to understand the evolution of coal and solvent molecular structures during solvent extraction, but such information is not easily accessible to experimental investigations [8]. For example, it was empirically found that using solvents as hydrogen donors or catalytic hydrogenation of coal was a critical step in preparing mesophase pitch for high-quality carbon fiber productions [59,60]. Hydrogenation is a common practice nowadays, but we are far from a clear picture of the role and impact of hydrogen donation during solvent extraction. In fact, to the best of our knowledge, a systematic comparison of different solvents to understand the role of hydrogenation for the solvent extraction process of coal in large-scale atomistic model

simulations has not yet been attempted, largely perhaps due to the lack of realistic atomistic coal models with sufficient unit cell size, derived from experimental characterizations. The currently available processes for large-scale atomistic model construction require human intervention in several steps using commercial software tools [29,32,34]. Accessible software tools will allow one to quickly build and utilize reasonable models to provide the basis for further simulations [37] to aid experimental studies.

In this work, we have therefore developed an open-source software tool for efficiently constructing large-scale atomistic models for coal materials with relevant experimental characterizations based on a randomized bottom-up construction approach. Our tool automatizes the atomistic model construction process based on structural data from HRTEM,  $^{13}\text{C}$  NMR, LDI-MS and elemental composition analysis. The user supplies a desirable total number of molecules for the coal model to be constructed, and the tool will then create a “library” of molecules matching the experimental mass distribution and incorporate appropriate numbers and types of chemical functional groups, heteroatoms, polycyclic aromatic hydrocarbon (PAH) units, and crosslinking bridges, which are helpful to study compounds like coal and pitch materials with complex and stochastic nature. The chemical and physical data for the selected coals were published in earlier works [23,24]. We demonstrate the capability of this tool for the theoretical study of the chemical conversion process via solvent extraction with reactive MD simulations using ReaxFF parameters with and without explicit solvent molecules such as tetralin as hydrogen donor and naphthalene as a chemically more inert solvent. Our study highlights the rapid development and application of stochastic atomistic models that align with experimental characterizations, enabling the examination of chemical processes such as pyrolysis and solvent extraction.

## 2. Materials and methods

### 2.1. Coal samples

The experimental data of four coal samples were utilized to construct the atomistic model: A subbituminous coal from the Powder River Basin “*Monarch*,” two bituminous coals from the Illinois Basin: “*Herrin*” and “*Springfield*,” and a bituminous coal from Central Appalachian “*Blue Gem*”. The chemical and physical data for these samples have previously been published [23,24,61].

### 2.2. Molecular size estimation using HRTEM and lattice fringe analysis

We identified the molecular size of PAH molecules using HRTEM followed by the lattice fringe image analysis. The details are available in an earlier publication [21,62]. An FEI Talos<sup>TM</sup> F200X scanning/transmission electron microscope provided 0.12 nm resolution with the recorded micrographs generated in the bright field mode at 630 kX magnification operated at 200 kV. Image processing and analysis used Photoshop<sup>TM</sup> software in conjunction with the QIA plugin. The micrographs were cropped to be square and reduced in size to permit a fast Fourier transform with a filter and an inverse Fourier transform to reduce the noise. Prior to that stage, the contrast distribution was enhanced by histogram shaping. The threshold was manually selected to isolate the lattice fringes. Fringes that were  $<3$  or  $<9.4$  Å (larger than  $3\times 3$  rings) were rejected to be consistent with known coal chemistry. Sizes for PAH identification are the same as previously used [27]. Six to nine micrographs were used for each sample to determine the distribution of PAH molecules [27,63].

### 2.3. Solid state $^{13}\text{C}$ NMR measurement

The demineralized coal samples were characterized with solid state  $^1\text{H}$ - $^{13}\text{C}$  Cross-Polarization Magic-Angle Spinning (CPMAS) NMR spectroscopy. Data were collected with a custom-built 3.2 mm 1H-X-Y triple

resonance MAS probe on a 14.1 T magnet spectrometer (1H 600 MHz) managed by a Bruker Avance III console at the MagLab of Florida State University. The characterization was conducted as follows: Approximately 40 mg of powdered coal was loaded to a rotor and spun at a frequency of 16 kHz. The experiments utilized pulse powers of 50 and 100 kHz for the  $90^\circ$  pulses on the  $^1\text{H}$  and  $^{13}\text{C}$  channels, respectively. Applying a ramped (70–110%) spin-lock pulse for  $^1\text{H}$  channel and a rectangular pulse for  $^{13}\text{C}$  channel, an optimal  $^1\text{H}$ - $^{13}\text{C}$  CP state was identified with  $\nu_1$  ( $^1\text{H}$ )  $\approx$  65 kHz and  $\nu_1$  ( $^{13}\text{C}$ ) = 50 kHz. The  $^1\text{H}$  decoupling method utilized for the direct and indirect acquisition stages of  $^{13}\text{C}$  signals was SPINAL-64, employing a decoupling radiofrequency power of 90 kHz. Based on a method suggested by Solum et al. [15], the fractional quantification of carbon sites of coal samples was characterized with variable contact  $^1\text{H}$ - $^{13}\text{C}$  CPMAS and dipolar-dephasing experiments. The decay constants of each spectrum peak component, both the Lorentzian and the Gaussian type of initial magnetization were determined via the dipolar-dephasing experiments employing 24 delay times ranging from 0 to 30  $\mu\text{s}$ . 13 variable contact  $^1\text{H}$ - $^{13}\text{C}$  CPMAS experiments were conducted, with contact times ranging from 0.5 to 10 ms, to extract the time constants of  $^1\text{H}$ - $^{13}\text{C}$  polarization transfer and the proton spin-lattice relaxation. Fractional contents of carbon sites identified in these experiments were extracted from the molecular parameters.

### 2.4. Mass spectrum using LDI-MS measurement

LDI-MS analyses were conducted using a Bruker Ultraflextreme MALDI-TOF apparatus at Pennsylvania State University featuring a Smartbeam-II IR laser (NdYAG) with a maximum repetition rate of 2 kHz. The specific method parameters were fine-tuned accordingly. In this measurement, the examination was performed using unprocessed coal powders.

### 2.5. Fourier transform infrared (FTIR) spectroscopy

Demineralized coal samples were ground with bromide potassium (KBr) at a concentration of 50 vol% with a pestle and mortar and 1 mm diameter pellets were made using a Carver hydraulic press (pressure of 4 tons). A PerkinElmer Frontier Fourier transform infrared (FTIR) spectrometer was used in attenuated total reflection (ATR) mode. ATR mode with a high concentration in coal gave a better signal than transmission mode with a low concentration (typically 0.5–1 wt%). The background and the sample spectra were generated by cumulating 16 scans over a wavenumber range of 400 to 4000  $\text{cm}^{-1}$  with an accuracy of 2  $\text{cm}^{-1}$ . The baseline of all spectra was manually corrected.

### 2.6. Bottom-up construction workflow

A program was created in the Python programming language, the “*MolecularSolidBuilder*” (<https://github.com/MoleCraftHUB/MolecularSolidBuilder>) for hydrocarbons with a bottom-up construction approach to generate coal molecular structures using the RDKit package [64]. The code enables users to build a collection of molecules that match the experimental data suite encompassing elemental composition, molecular weight distributions,  $^{13}\text{C}$  NMR data, and HRTEM platelet-size data. The specific process is reviewed in detail in the Model construction section.

### 2.7. Molecular dynamics simulations

The thermal pyrolysis and solvent extraction process for the atomistic coal structures were examined through MD simulations performed via the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) code using ReaxFF [65–69]. It is widely used for chemical reactions of large-scale and complex atomistic models that can describe the bond formation and cleavage processes to predict intermediate

radicals and product molecules. Recently, ReaxFF has been intensively utilized for thermal conversion studies of coal with large-scale atomistic models [32,33,44,45], predicting processes that agree with experimental observations. A review of its application to coal has also recently been published [52]. The simulations here used a canonical ensemble with constant volume and constant temperature (NVT ensemble), where the temperature was controlled using a Nose-Hoover thermostat [70,71] with a temperature step of 0.1 fs and a damping parameter of 25 time-steps. The velocity-Verlet integration algorithm was used for all simulations. The initial volumes of a box containing coal molecules or coal molecules with solvent molecules were  $150^3 \text{ \AA}^3$  and  $200^3 \text{ \AA}^3$ , respectively. The molecules are randomly distributed in the large box without overlap of positions using PACKMOL [72]. The target density of solid phases was derived from experimental characterizations of each coal specimen, estimating the desired volume and box size with molecular mass. The electrostatic interactions were computed based on environment-dependent charge equilibration with the electronegativity equalization method (EEM) and tapered electrostatic energy using the cutoff radius of  $10 \text{ \AA}$  [73,74]. ReaxFF MD simulations often utilize high temperatures (2000–4000 K) to accelerate chemical reaction process by increasing the collision factor and the kinetic energy to overcome kinetic barriers in short simulation time ( $\sim 1 \text{ ns}$ ) as compared to the experiment process ( $\sim 1800 \text{ s}$  or  $\sim 30 \text{ min}$ ). Arrhenius equation can be used to estimate a desired temperature in MD simulations based on time scale difference if the experimental reaction constant can be obtained [52, 75]. These accelerated high-temperature simulations demonstrated consistency of chemical reaction process compared with experimental observations [76–78]. In this work, we performed high-temperature NVT simulations at 1073 and 2073 K for 1 ns for coals in explicit solvents. This approach enables us to address the limitation of the shorter process time of reactive MD simulations compared to a slow experimental process at  $\sim 673 \text{ K}$  (or  $400 \text{ }^\circ\text{C}$ ) for 5–30 min in the solvent

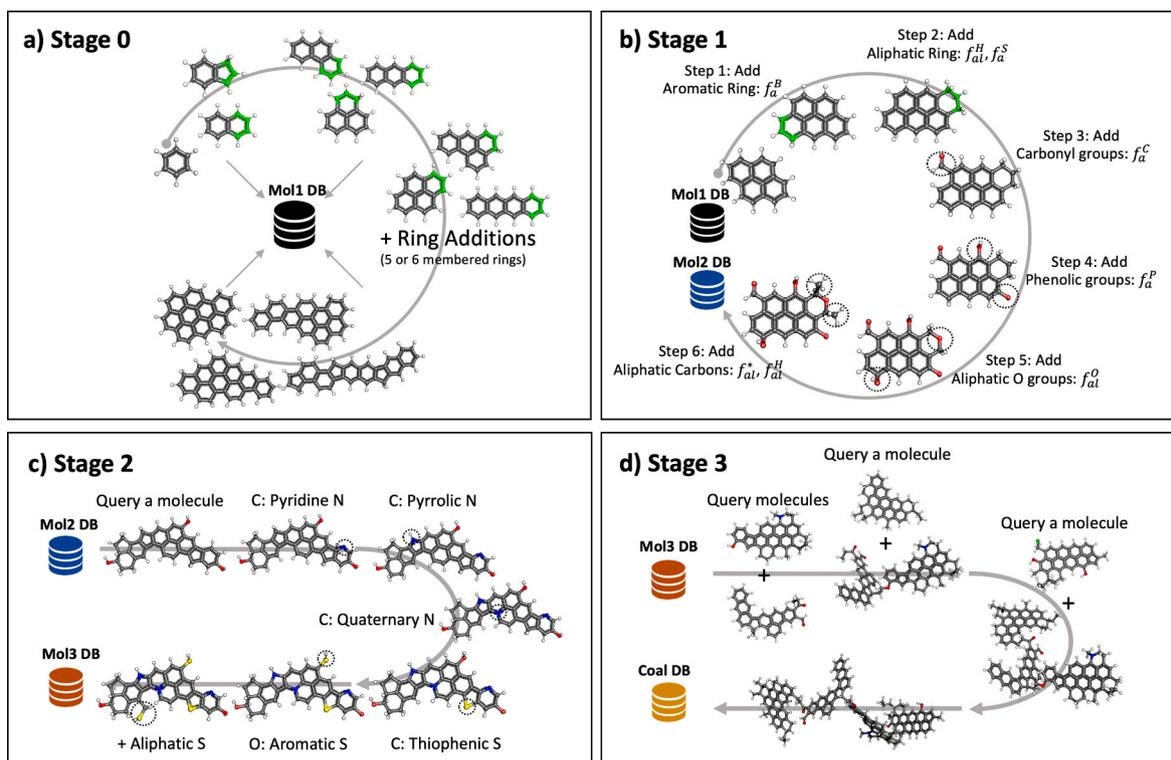
extraction process. Multiple versions of ReaxFF have been parameterized and available for hydrocarbon studies to predict different aspects and properties.

Along with a few available parameters, we tested the relative formation energy difference of 450 molecules using four ReaxFF parameterizations with respect to B3LYP/6-311G\* (Fig. S1). We observed that the two parameters by Mattsson et al. [79] and Nielson et al. [80] predict very similar relative formation energies and do not lead to bond formation/breaking of coal molecules in solid phases. The ReaxFF parameter by Mattsson et al. [79] has previously been used to study pyrolysis, coke formation of lignin, and coal [31,81–83]. We examined the solvent extraction and pyrolysis process using both parameter sets and showed the results obtained with the parameter set by Mattsson et al. [79] in the manuscript. The results obtained with the parameter set by Nielson et al. [80] were shown in the supporting information. We do not find statistically significant differences in describing these highly dynamic processes. For molecular visualization, the Ovito [84] and PyMol [85] tools were used to visualize coal structures and molecular species, respectively.

### 3. Results and discussion

#### 3.1. Model construction

Following the bottom-up procedure demonstrated by Castro-Marcano et al. [29] and Gao et al. [81], we created a new Python-based tool, “MolecularSolidBuilder” (<https://github.com/MolecraftHUB/MolecularSolidBuilder>), to rapidly generate atomistic models containing an ensemble of hydrocarbon-rich structures to represent coals or pitches. The code is designed to create large-scale atomistic models of molecular solids containing molecules of arbitrary size from experimental input and involves multiple stages, as shown in Fig. 1. In



**Fig. 1.** Stage and step-wise bottom-up molecule construction approach. a) Stage 0 for PAH molecule generation collected as Mol1 DB. b) Stage 1 for adding rings and functional groups with 6 steps to match to  $^{13}\text{C}$  NMR data collected as Mol2 DB. c) Stage 2 for incorporation of heteroatoms (oxygen, nitrogen, and sulfur atoms) to match the elemental composition collected as Mol3 DB. d) Stage 3 for bridging multiple molecules to match mass distribution collected as Coal DB. The carbon, hydrogen, nitrogen, oxygen, and sulfur atoms are colored in grey, white, blue, and yellow, respectively. In addition, the green atoms are carbons for newly added rings. (A colour version of this figure can be viewed online.)

**Table 1**  
PAH molecule size distribution for the structural representations.

Name	Size (Å)	Monarch		Herrin		Springfield		Blue Gem	
		Frequency (%)	Number						
PAH1	≤3.0	31	310	21	189	28	280	13	100
PAH2	3.0–4.5	43	430	49	441	51	510	59	460
PAH3	4.5–6.0	9	90	10	90	9	90	10	80
PAH4	6.0–7.5	8	80	9	81	6	60	8	60
PAH5	7.5–11.5	7	70	8	72	5	50	7	54
PAH6	11.5–14.5	2	20	3	27	1	10	3	20
Total Number			1000		900		1000		774

coals, PAH molecules have various numbers of single or fused aromatic rings, typically with three to five fused rings, with different side chains and heteroatoms (the distribution of sizes is rank and maceral-dependent)

A common approach is utilizing a suite of idealized PAH molecules such as benzene, naphthalene, anthracene, phenanthrene, fluorene, pyrene, etc [29,32,86]. In our model builder, we increased the molecule complexity by adding and concatenating 5-membered and 6-membered rings starting from the idealized PAH molecules with various sites of PAH such as bay, cove, fjord, M, and K regions as shown in Fig. S2. Fig. 1a shows that a range of structures can be generated from benzene or cyclopentadiene, increasing the size of hydrocarbons based on chemical rules that consist of fused 5-membered and 6-membered rings. Various membered rings can be added to PAH molecules, but we are restricted to having 5-membered and 6-membered rings in this work. The tool enables the addition of various rings with 3–10 members, as illustrated in Fig. S2. Generated molecules were collected to the PAH database (DB). The list of PAH molecules for “which” coals or more general units used is shown in Fig. S3 (e.g., PAH1-1 is the benzene and PAH2-1 and PAH2-2 are the naphthalene and the indene, respectively). We randomly added functional groups and side chains to increase the structural diversity and incorporate experimental characteristics (although some bio-specific isomerization is likely retained from the biomass inputs). The functional groups for coal models in this work include methyl, ethyl, carbonyl, phenolic and naphthenic hydrocarbons. The tool can also be modified to incorporate additional functional groups using the Simplified Molecular-Input Line-Entry System (SMILES) notation. Molecular structures in all stages and steps in Fig. 1 are stored as three-dimensional geometries based on structural embedding and molecular mechanics using MMFF94s of RDKit [64].

The lattice fringe analysis of HRTEM from the previous study [21] was utilized to estimate the distribution of PAH molecules [62]. Table 1 shows the PAH molecule distributions for each cluster comprising the individual coal structures. The size of concatenated molecules was estimated from the length and breadth of the enclosing box of three-dimensional coordinates of atoms. The single number for each molecule was calculated from the average value of the width and the length of the rectangular bounding box. The total number of PAH

molecules for four coal models was 1000 for Monarch (subbituminous), 900 for Herrin (bituminous), 1000 for Springfield (bituminous), and 774 for Blue Gem (bituminous) to construct systems of ~40,000 atoms with the suggested distribution from the lattice fringe image analysis.

Consistent with the coal rank, Monarch coal has the highest percentage of smaller PAH molecules (PAH1 and PAH2) accounting for 74 % (Table 1). The three bituminous coals have a higher contribution of the larger PAH molecules (PAH3~PAH6). The total molecule numbers were initially defined with ~20,000 atoms (hydrogens and carbons) for each structure. Further concatenation of ring structures and adding aliphatic and functional groups in Fig. 1b make the final structure consist of ~40,000 atoms (size required to capture some molecular weight distribution).

Given the suggested number of the different sizes of PAH, we randomly selected duplicated PAH molecules in each size estimation for each coal model from the Mol1 database collected in stage 0 (Fig. 1a). The collection of simplest PAH molecules is decorated and modified in the following stage (Fig. 1b) by closely monitoring the experimental data of <sup>13</sup>C NMR characterization (Table 2), providing information on carbon skeletal structures. This is achieved by adding and modifying aromatic bridgehead carbons ( $f_a^B$ ), aliphatic carbon bonded to 2 hydrogens or 1 hydrogen ( $f_{al}^H$ ), aromatic alkylated carbon ( $f_a^S$ ), aromatic carbonyl carbon ( $f_a^C$ ), aromatic phenolic carbon ( $f_a^P$ ), aliphatic carbon bonded to oxygen ( $f_{al}^O$ ), aliphatic carbon bonded to 0 hydrogen or 3 hydrogens ( $f_{al}^*$ ).

The computational <sup>13</sup>C NMR parameters were collected from all molecules at every modification step. We randomly selected applicable sites for desired structural modifications and modified molecules until the set of molecules had consistent parameters for <sup>13</sup>C NMR data. In the step 1 and 2 of the stage 1, we expanded the number of rings for aromatic carbons for bridgeheads ( $f_a^B$ ) and aliphatic carbons for cycloalkanes, adjusting the fraction of alkylated aromatic carbons ( $f_a^H$  and  $f_a^S$ ) because selected PAH molecules from the size distribution did not satisfy the conditions for these parameters. In the following steps (3–6), functional groups were added to molecules with a carbonyl group ( $f_a^C$ ), phenolic groups ( $f_a^P$ ), aliphatic oxygen group ( $f_{al}^O$ ), and aliphatic carbon groups ( $f_{al}^*$  and  $f_{al}^H$ ). The resulting molecules of each coal model were added to Coal DB, as shown in Fig. 1.

**Table 2**  
<sup>13</sup>C NMR data for coals (Coal-experiment) and calculated data for atomistic models (Coal model).

Structural parameters	$f_a$	$f_a'$	$f_a^C$	$f_a^H$	$f_a^N$	$f_a^P$	$f_a^S$	$f_a^B$	$f_{al}$	$f_{al}^H$	$f_{al}^*$	$f_{al}^O$	$\chi_b$
Monarch-experiment	0.67	0.63	0.04	<b>0.23</b>	<b>0.40</b>	0.04	0.16	0.20	0.33	0.19	0.14	0.06	0.32
Monarch-model	0.67	0.63	0.04	<b>0.22</b>	<b>0.41</b>	0.04	0.16	0.20	0.33	0.19	0.14	0.06	0.32
Herrin-experiment	0.69	0.66	0.03	0.19	<b>0.46</b>	<b>0.04</b>	0.17	0.25	0.31	0.19	0.12	0.05	0.38
Herrin-model	0.69	0.66	0.03	0.19	<b>0.47</b>	<b>0.05</b>	0.17	0.25	0.31	0.19	0.12	0.05	0.38
Springfield-experiment	0.69	0.66	0.03	0.23	0.43	0.04	0.15	0.24	0.31	0.19	0.12	0.05	0.36
Springfield-model	0.69	0.66	0.03	0.23	0.43	0.04	0.15	0.24	0.31	0.19	0.12	0.05	0.36
Blue Gem-experiment	0.70	0.67	0.03	0.21	0.46	0.04	0.17	0.25	<b>0.30</b>	<b>0.20</b>	0.10	0.01	0.37
Blue Gem-model	0.70	0.67	0.03	0.21	0.46	0.04	0.17	0.25	<b>0.29</b>	<b>0.19</b>	0.10	0.01	0.37

Fractions of sp<sup>2</sup>-hybridized carbon (approximate error estimate):  $f_a$  = total carbon (±0.03);  $f_a'$  = in an aromatic ring (±0.04);  $f_a^C$  = carbonyl,  $\delta > 165$  ppm (±0.02);  $f_a^H$  = protonated and aromatic (±0.03);  $f_a^N$  = non-protonated and aromatic (±0.03);  $f_a^P$  = phenolic or phenolic ether,  $\delta = 150$ –165 ppm (±0.02);  $f_a^S$  = alkylated aromatic,  $\delta = 135$ –150 ppm (±0.03);  $f_a^B$  = aromatic bridgehead (±0.04);  $f_{al}$  sp<sup>3</sup>-hybridized carbon (±0.02);  $f_{al}^H$  = CH or CH<sub>2</sub> (±0.02);  $f_{al}^*$  = CH<sub>3</sub> or non-protonated (±0.03);  $f_{al}^O$  = bonded to oxygen,  $\delta = 50$ –90 ppm (±0.02). Entries with more than 1% difference between experimental results and models were bolded in the table.

**Table 3**  
Proximate, ultimate analysis, and helium pycnometry densities.

	Proximate analysis (wt.%, as-received basis)				Ultimate analysis (wt.%, dmmf basis)					Pycnometry (g/cm <sup>3</sup> , dry ash free basis)
	Moisture	Volatile Matter	Fixed carbon	Ash Yield	C	H	N	O <sup>a</sup>	S	Density
Monarch	9.58	36.59	39.01	14.82	72.99	4.86	1.46	20.04	0.65	1.50
Herrin	2.67	41.49	51.05	4.79	83.97	5.61	1.79	5.19	3.44	1.33
Springfield	6.43	39.13	51.68	2.76	86.99	5.59	1.98	2.61	2.83	1.35
Blue Gem	2.05	37.31	59.44	1.20	86.70	5.71	2.31	4.32	0.96	1.36

<sup>a</sup> Oxygen contents in Ultimate analysis measured by difference.

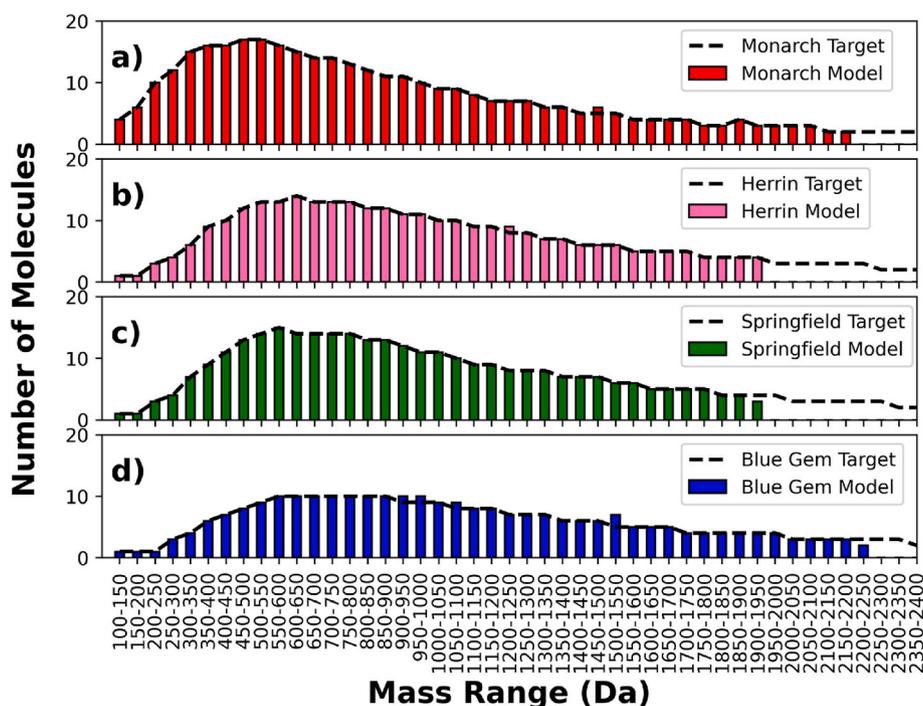
**Table 4**  
Elemental composition from experiment and model.

Elements Composition	C atomic. %	H atomic. %	N atomic. %	O atomic. %	S atomic. %	Number. of atoms
Monarch-experiment	49.4	39.5	0.84	10.1	0.16	–
Monarch-model	48.0	42.9	0.86	8.00	0.17	39,333
Herrin-experiment	53.2	42.6	0.97	2.46	0.82	–
Herrin-model	51.1	43.2	0.99	3.90	0.83	37,070
Springfield-experiment	54.8	42.2	1.08	1.24	0.68	–
Springfield-model	50.7	43.1	1.09	4.40	0.68	38,812
Blue Gem-experiment	53.9	42.6	1.24	2.03	0.23	–
Blue Gem-model	52.1	43.0	1.07	3.55	0.28	37,918

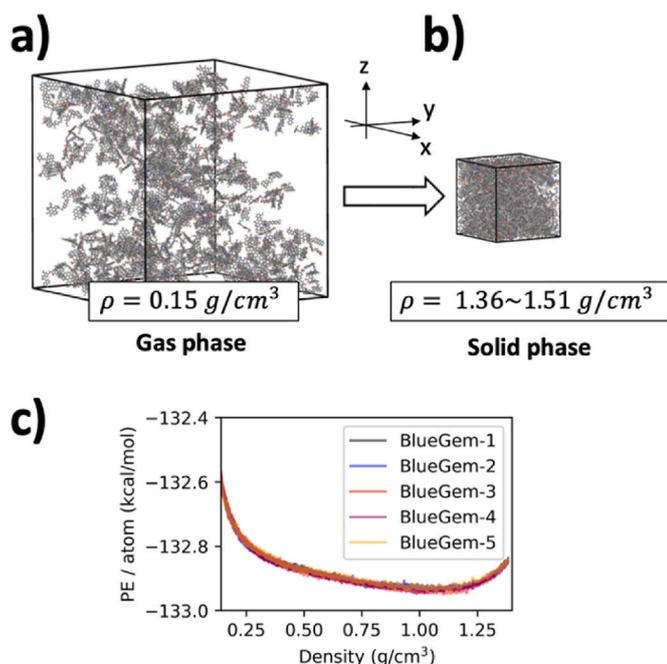
Oxygen contents are the most significant errors after matching the NMR parameters of  $f_{al}^C$ ,  $f_{al}^H$ , and  $f_{al}^O$

As a result of stage 1 (Fig. 1b), the collection of molecules of a model (Mol2 DB) has a good match for the experimental data with ~1 % difference for all <sup>13</sup>C NMR parameters except for  $f_{al}^O$ . The aliphatic carbon bonded to oxygen ( $f_{al}^O$ ) was deliberately matched by 1 % less than the experimental  $f_{al}^O$  in Table 2 because it increased further when the PAH molecules were bridged with O atoms. The heteroatoms were added in stage 2 (Fig. 1c) after modifying the carbon skeletal structure in stage 1 with reported types for nitrogen and sulfur functional groups. The heteroatom types were limited to pyridine, pyrrolic, quaternary nitrogen, thiophenic, aliphatic, and aromatic sulfur based on previous characterization [22,87].

We note that incorporating  $f_{al}^C$ ,  $f_{al}^H$ , and  $f_{al}^O$  leads to adding oxygen-forming aldehyde, ketone, carboxyl functional groups, phenolic oxygens, open ethers and oxygen in cyclic carbon and also within the side chains. Due to insufficient experimental characterization studies of heteroatom types and contents of these four coal samples, we utilized the relative proportion of nitrogen and sulfur from the literature. However, this can be further improved by analyzing the given coal samples. The amounts of pyridine, pyrrolic, and quaternary nitrogen are fixed to 26 %, 62 %, and 12 %, respectively, based on the XPS report by Kelemen et al. for Argonne Premium bituminous coal [22]. The amount of sulfur is distributed with a goal of 41 % as thiophenic sulfur, 26 %



**Fig. 2.** Mass distribution of molecules of coal models for a) Monarch, b) Herrin, c) Springfield, and d) Blue Gem. Experimental target values from LDI-MS data are shown in solid black lines. The number of molecules at each molecular weight bin (MWs) in atomistic models is shown in the bar chart up to 2400 Da. (A colour version of this figure can be viewed online.)



**Fig. 3.** a) Gas phase with coal molecular model and b) solid phase with coal molecular model compressed to the experimental densities. c) Compression curve of 5 Blue Gem models from different distribution of molecules (density vs potential energy (PE) per atom) with the MD simulation at 300 K using ReaxFF parameter. (A colour version of this figure can be viewed online.)

within aromatic structures, and 33 % aliphatic sulfur based on the XANES report by George et al. for Argonne Premium bituminous coal [87]. We did not include water molecules as a structural component in these atomistic representations.

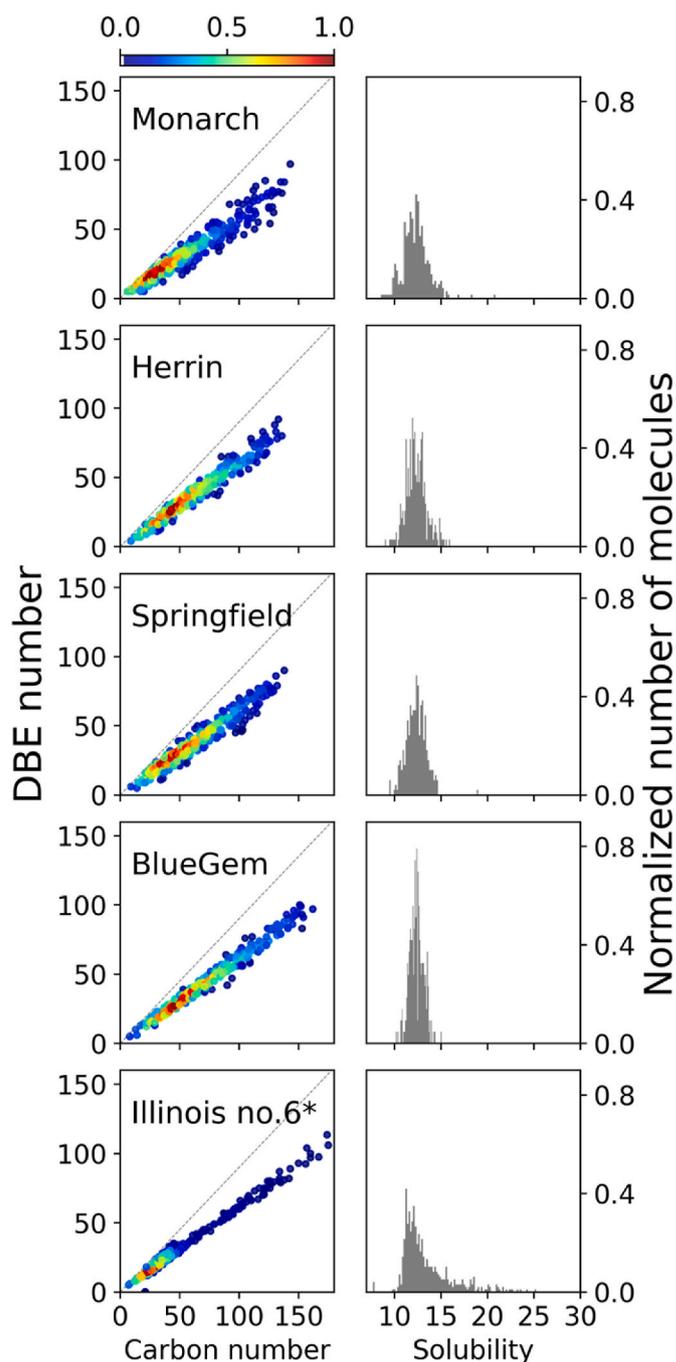
From stage 1 and stage 2, the individual PAH molecules were decorated with the number of aromatic rings and closed aliphatic chains (cyclic compounds), functional groups, and heteroatoms. Note that we matched  $^{13}\text{C}$  NMR with less than 1 % difference for all entries, but we still have errors for the elemental composition, with the most significant error for oxygen and carbon percentage difference in Springfield coal, with 3.2 and 4.1 % between the experimental data and the model (Table 4).

The errors for the elemental compositions of ultimate analysis (Table 3) originated from the utilization of the  $^{13}\text{C}$  NMR data over the elemental composition for the oxygen content because it provides the specific types of oxygens such as carbonyl ( $t_a^c$ ), phenolic ( $t_a^p$ ) and aliphatic carbon ( $t_a^o$ ) atoms. This will be further improved by optimizing the number of cross-linking with specific pairs of carbon and oxygen atoms instead of simply adding necessary amounts of oxygen and randomly linking molecules by bridging randomly selected functional groups.

Coal consists of molecules of different sizes and molecular weights; while some are isolated, many are cross-linked for these ranks. The molecular weight (MW) distribution estimates from LDI-MS guided the target range and the number of molecules in different MW of each coal sample (black lines of Fig. 2). In stage 3 (Fig. 1d), PAH molecules from Coal DB' were randomly selected and bridged with the available pair of side chains and aromatic atoms to generate molecules for the target distribution. The MW distributions were shown as a color bar for the number of constructed molecules and solid black lines for the initial target numbers of molecules obtained from experimental distributions. In our bottom-up construction method, we built an individual coal model with desired numbers of molecules from 100 up to 2400 Da with a 50 Da interval (using all individual PAH molecules for each coal model). Each stage of the entire molecular solid model generation is guided by

distinct experimental data, sequentially accumulating different structural factors. In Stage 0, the PAH size distribution is aligned with HRTEM lattice fringe data, while in Stage 1, the skeletal structure corresponds to  $^{13}\text{C}$  NMR data. Subsequently, Stage 2 aligns with the elemental composition derived from ultimate analysis, and Stage 3 corresponds to the molecular weight distribution.

With this systematic model construction process, we have a different distribution and number of molecules for coal models. The bottom-up approach allows one to generate an ensemble of molecules in different



**Fig. 4.** Carbon number vs DBE number (left column) and Solubility parameter vs Normalized number of molecules (right column) for different coal models. Illinois no. 6 model from Ref. [29] was also added to examine relative differences in DBE populations and the normalized distribution of solubility parameters. The relative number of molecules for DBE is colored with the normalized number of molecules in blue-green-red color spectrum. (A colour version of this figure can be viewed online.)

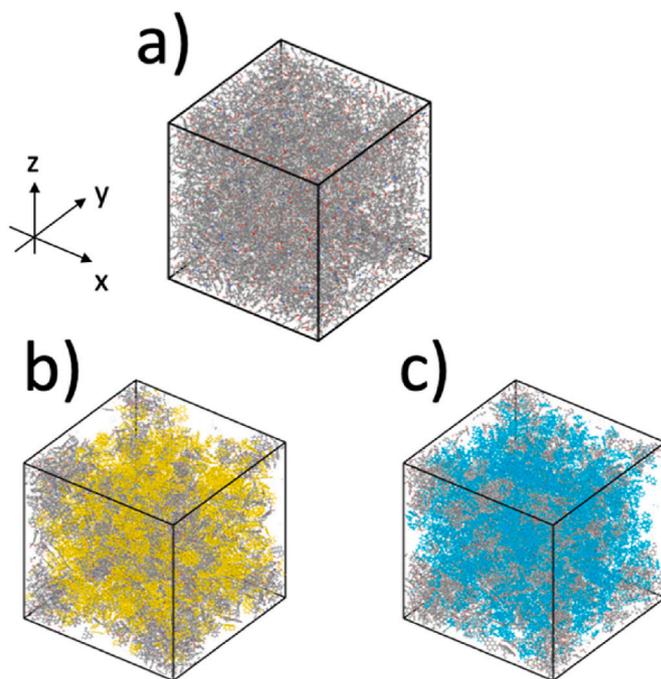
PAH core sizes and molecular weights, consistently matching experimental  $^{13}\text{C}$  NMR and elemental composition data. We note that characterization errors and variations will significantly impact the final structures.

### 3.1.1. Solid structures from the ensemble of molecules

From the bottom-up construction process for the given coal materials, an ensemble of individual macromolecules and small molecules can be generated for each coal sample. However, coals are solid porous structures. Thus, the ensemble must be compressed to match experimental densities, determined from the pycnometer measurement (Table 3) with the range of 1.36–1.51 g/cm<sup>3</sup> (dry ash free basis, daf), consistent with literature expectations for these ranks [88]. For the solid materials, the molecules for each coal material were all randomly placed in a three-dimensional box at gas density ( $\sim 0.15$  g/cm<sup>3</sup>) and compressed to the experimental solid density using the MD simulations at room temperature (300 K) during 40 ps by reducing the box size by 2 Å for all directions (x, y, z) every 1 ps (Fig. 3). Five gas phase samples of Blue Gem were compressed to investigate potential energy biases by the random distribution of molecules (Fig. 3c). There were no noticeable differences in the energy curves and the optimal density. We note that the MD simulations at 300 K were non-reactive without bond formation/breaking between molecules during the compression with two different sets of ReaxFF parameters (Mattsson et al. [79] and Nielson et al. [80]).

After we built four coal systems, we checked their features and compared them with experimental characterizations. First, we computed infrared (IR) spectra with our coal models using both individual molecules with the normal mode analysis using the density-functional tight-binding (DFTB) method and condensed coal system with the fast Fourier transform dipole autocorrelation function using ReaxFF (see Fig. S5). These computational IR spectra were used to confirm the consistency of constructed molecules with experimental spectra of coal samples. We observe that the experimental IR spectra of the three bituminous coals showed no distinguishable differences in peak positions and intensities. The experimental IR spectrum of Monarch, classified as subbituminous coal, exhibits slight variations with peaks around  $\sim 750$  and  $\sim 3200$  cm<sup>-1</sup>, corresponding to the vibration of the aromatic nucleus and the -OH stretching, respectively. Computational IR spectra at the molecular level for the four coals (Fig. S6) illustrate that they comprise distinct sets of molecules. However, the resulting spectra become indistinguishable when averaging the computational IR spectra of these molecular sets (Fig. S5) in both DFTB and ReaxFF. The relative amounts of chemical structural units and functional groups of these coal molecules are comparable, making integrated spectra indistinguishable. The vibrational peak positions of integrated IR spectra from those of individual molecules with DFTB were in better agreement with experimental results than those of condensed systems with ReaxFF. The inability to distinguish between IR spectra of various coals, despite using distinct sets of molecules in the computational IR spectra, suggests that IR spectra can only assess the relative quantities of different vibrational modes rather than their specific molecular structures.

Next, the double bond equivalent (DBE) number and solubility parameters ( $\delta$ ) of individual molecules for each coal model were analyzed to verify their resemblance/disparity, as illustrated in Fig. 4. The DBE number is calculated with eq (1), which is calculated with the number of hydrogen ( $n_H$ ), carbon ( $n_C$ ), and nitrogen ( $n_N$ ) of each molecule. The solubility parameter ( $\delta$ ) is calculated based on eq (2) where  $n_H$ ,  $n_C$ ,  $n_N$ ,  $n_O$ , and  $n_S$  are corresponding to the number of hydrogen, carbon, nitrogen, oxygen and sulfur together with the aromatic carbon ratio ( $f_a$ ) of each molecule, respectively. It had been formulated with the group contribution approach and developed by Painter et al. [10] The populations of DBE numbers were represented by a color gradient in the scatter plot, ranging from blue to red, corresponding to their relative frequencies in the region. We also determined the analysis for the



**Fig. 5.** a) Blue Gem model, b) Blue Gem coal mixed with naphthalene (NAP) model, and c) Blue Gem coal mixed with tetralin (TET) model. All systems were compressed to the density of  $\sim 1.00$  g/cm<sup>3</sup> for simulation consistency. The coal, naphthalene, and tetralin molecules were colored differently to show the overall distribution of solvent molecules: grey for coal atoms, yellow for naphthalene atoms, and blue for tetralin atoms. (A colour version of this figure can be viewed online.)

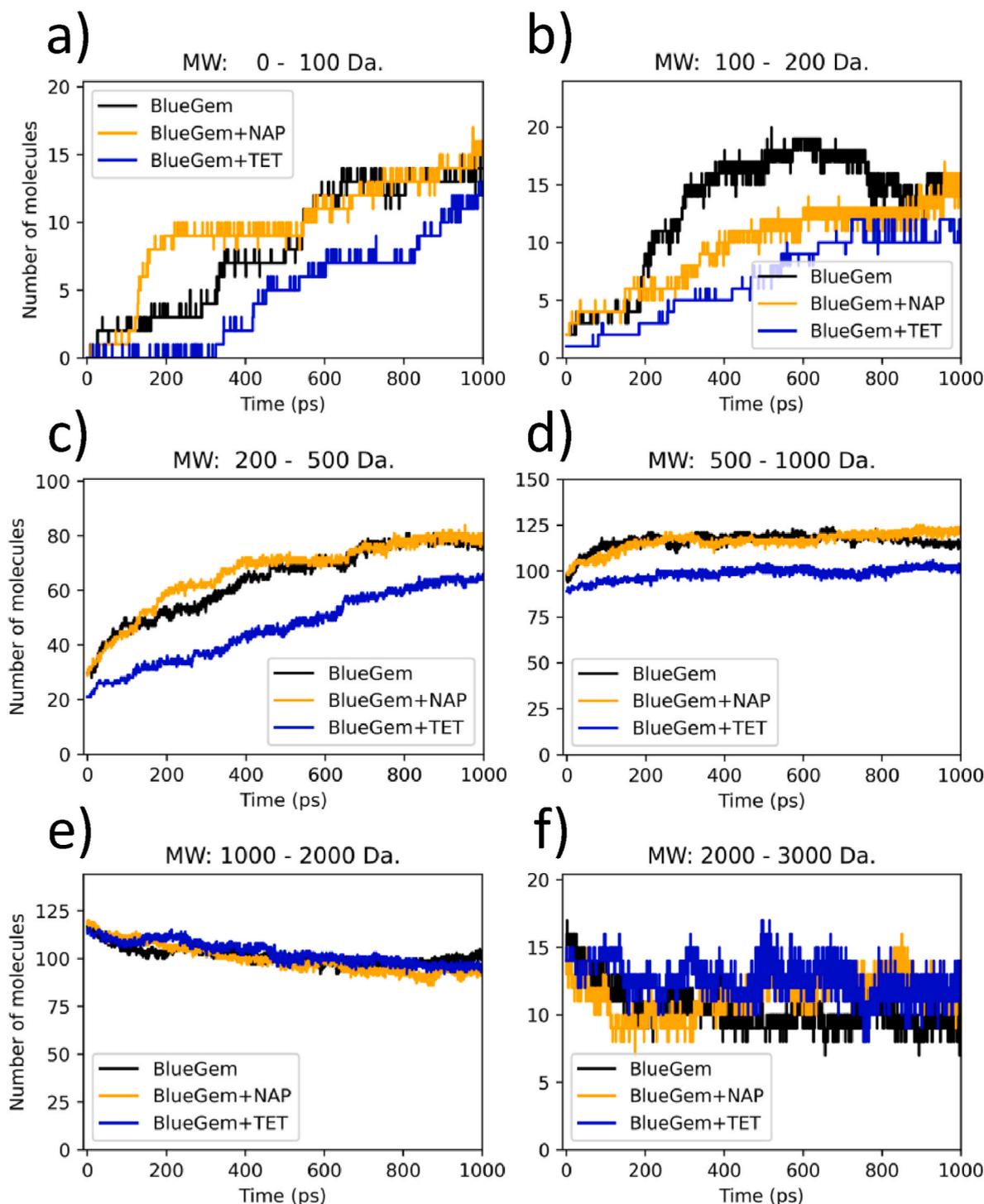
molecular model for Illinois no. 6 coal [29], as this is the same seam as the Herron sample. In this work, coal molecules constructed for Monarch, Herrin, Springfield, and Blue Gem are distributed widely with DBE number and narrowly with the solubility parameters compared to the Illinois no. 6 coal model. The differences are related to the larger scale of the Illinois No. 6 (50,789 atoms within 728 molecules) atomistic representation and a higher inertinite contribution (10 vs. 6.5 %) [24,51].

$$\text{DBE} = n_C - \frac{n_H}{2} + \frac{n_N}{2} + 1 \quad (1)$$

$$\delta = \frac{7.0 + 63.5f_a + 63.5\left(\frac{n_H}{n_C}\right) + 106\left(\frac{n_O}{n_C}\right) + 51.8\left(\frac{n_N+n_S}{n_C}\right)}{-10.9 + 12f_a + 13.9\left(\frac{n_H}{n_C}\right) + 5.5\left(\frac{n_O}{n_C}\right) - 2.8\left(\frac{n_N+n_S}{n_C}\right)} \quad (2)$$

### 3.2. Chemical decomposition simulations

We examined the high-temperature chemical reactions of our coal models with two common solvents (naphthalene and tetralin) to examine the role of hydrogen donation capability for the chemical conversion of coals. We also built the system with individual coal molecules well dispersed in solvent molecules, assuming that solvent molecules would well disperse a solid coal particle into individual molecules by breaking associate intermolecular interactions before the chemical reactions in the solvent extraction process. The weight ratio for the coal and the solvent was  $\sim 1:1$  by utilizing all molecules of the Blue Gem coal model: 37,918 atoms from 263 molecules (Table 4) and 2000 molecules of naphthalene (C<sub>10</sub>H<sub>8</sub>, NAP) or tetralin (C<sub>10</sub>H<sub>12</sub>, TET). To consistently study the chemical reaction dynamics, we conducted three dynamics simulations with the Blue Gem model: thermal pyrolysis of Blue Gem, solvent extraction with NAP, and solvent extraction with TET. Following the same condensation process in the solid coal model construction method, the coal and the solvent molecules were randomly placed in a



**Fig. 6.** The number of molecules at different molecular weight ranges from dynamics simulations at 1073 K with ReaxFF for a) 0–100 Da, b) 100–200 Da, c) 200–500 Da, d) 500–1000 Da, e) 1000–2000 Da, and f) 2000–3000 Da. (A colour version of this figure can be viewed online.)

$200^3 \text{ \AA}^3$  box at  $\sim 0.15 \text{ g/cm}^3$  density as the gas phase. Then, high-density systems were prepared with the gas phase compressing to the density at  $\sim 1.0 \text{ g/cm}^3$  at 300 K for 50 ps by reducing the box size by  $2 \text{ \AA}$  every 1 ps.

Fig. 5 displays the three systems after compression: a) Blue Gem model, b) Blue Gem + NAP, and c) Blue Gem + TET. The solvent molecules were colored grey for Blue Gem, yellow for NAP, and blue for TET to illustrate their volumetric distributions. The reactive MD simulations were performed at two temperatures (low and high), 1073 K (or 800 °C) and 2073 K (or 1800 °C) for 1 ns to observe the chemical reactions between coal and solvents. A direct comparison of MW changes over

time for the Blue Gem model without solvent, for Blue Gem + NAP, and for Blue Gem + TET are shown in Figs. 6 and 7.

Molecules in each MW range were tracked instead of showing the exact chemical species with the same chemical formula (because it was highly complex due to the many unique intermediate species in the pyrolysis and extraction processes). Fig. S7 shows the time evolution for the number of molecules from 50 to 3000 Da counting molecules with 50 Da bins. In the pyrolysis of Blue Gem, we observed that gas and small molecules (MW less than 500 Da) were formed (Fig. 6a–c and 7a–c). The low MW molecules such as  $\text{H}_2$ ,  $\text{CH}_4$ ,  $\text{H}_2\text{O}$ , and  $\text{NH}_3$  were chemically

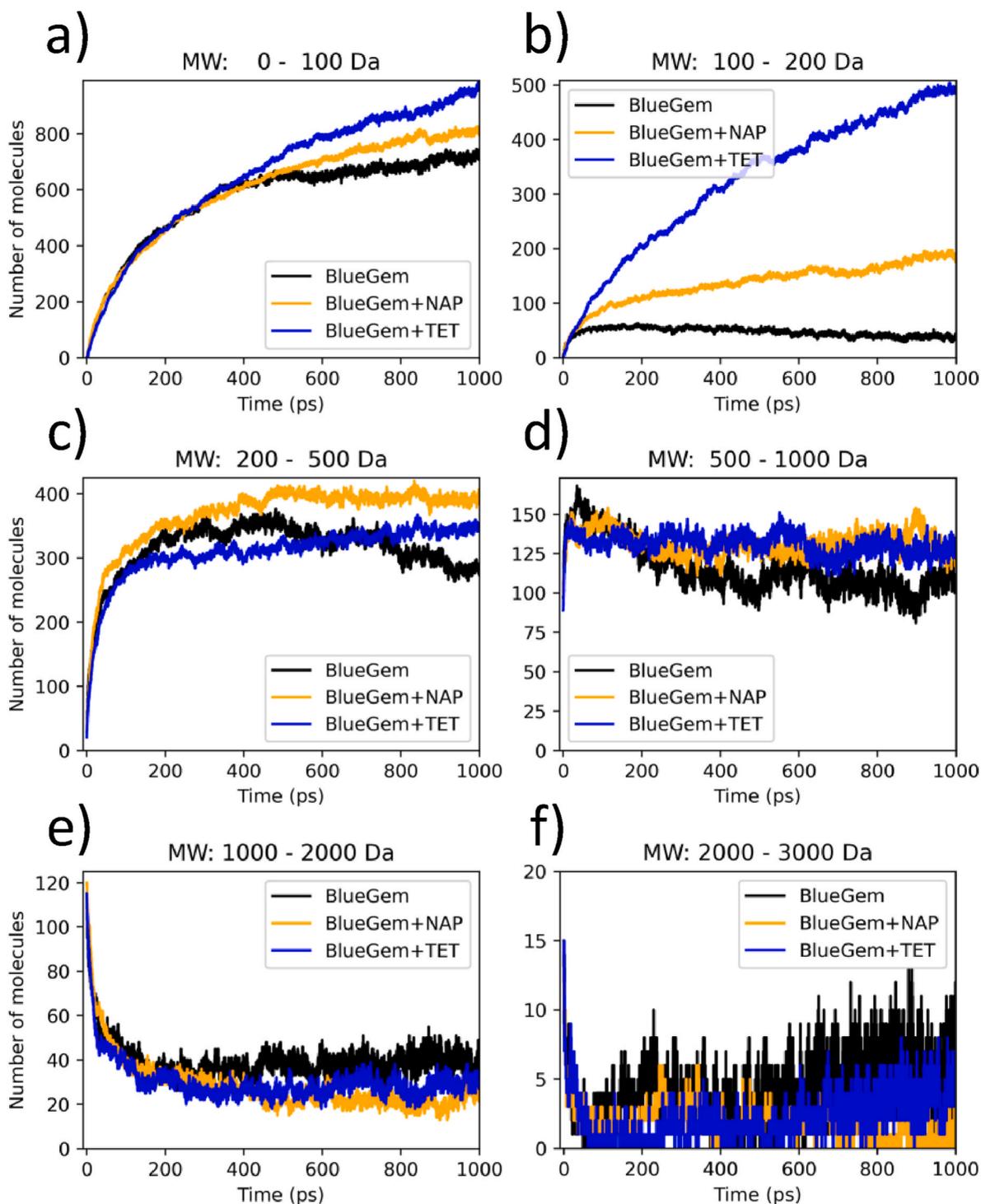
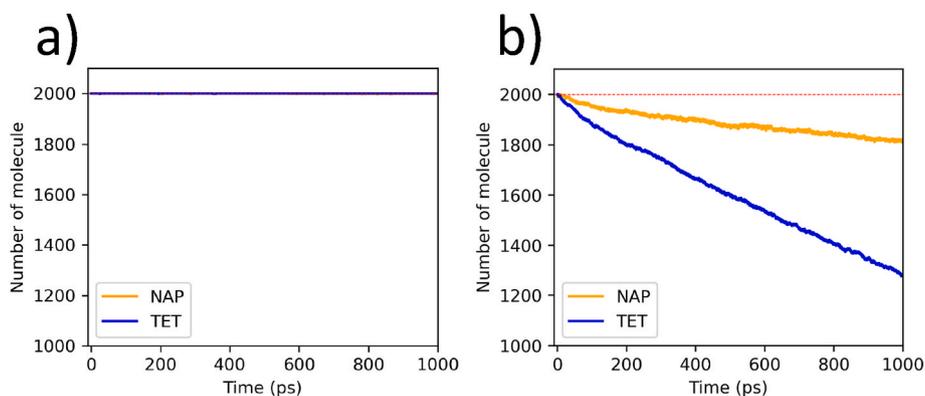


Fig. 7. The number of molecules at different molecular weight ranges from dynamics simulations at 2073 K with ReaxFF for a) 0–100 Da, b) 100–200 Da, c) 200–500 Da, d) 500–1000 Da, e) 1000–2000 Da, and f) 2000–3000 Da. (A colour version of this figure can be viewed online.)

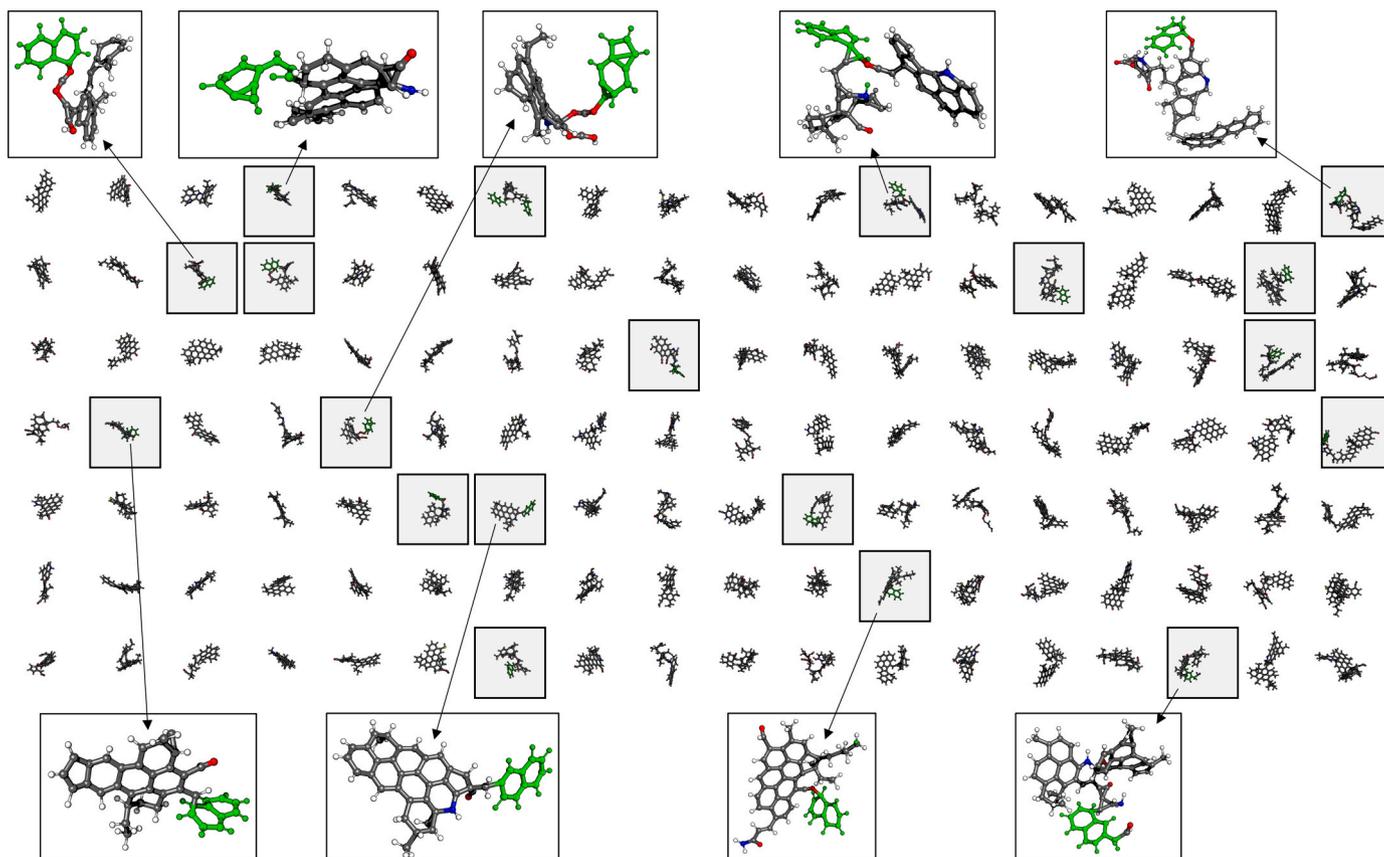
produced from the decomposition of coal at 1073 K (Fig. 6a) and coal with solvents at 2073 K (Figs. 7a and 8b). The production of gas molecules of pyrolysis of lignite and bituminous coal was a consistent outcome of both experimental and ReaxFF simulations reported by Mingjie Gao et al. [32] The increasing number of molecules in 200–500 Da in Figs. 6c and 7c was the result of breaking crosslinking between PAH molecules with decreasing number of molecules with high molecular weights (500–3000 Da). Limiting the reaction rate, the lower temperature simulations (at 1073 K for 1 ns) in Fig. 6e–f showed that coal macromolecules were first broken with an increasing number of

species (263–328 molecules without the formation of gas molecules).

Solvent extraction dynamics (Blue Gem + NAP or Blue Gem + TET) differed from the pyrolysis process of Blue Gem in forming more gas molecules and small molecules from solvent decomposition and coal dissociation, as shown in Fig. 7a–c. To see the evolution of intermediate and product molecules in the range of 100–200 Da for solvent extraction cases, the intact solvent molecules (NAP,  $C_{10}H_8$ , and TET,  $C_{10}H_{12}$ ) were not plotted together in Figs. 6 and 7. Instead, the decomposition of the solvent molecules is separately shown in Fig. 8. The two solvents behaved differently. The NAP solvent has no significant effect on the



**Fig. 8.** The number of solvent molecules during the thermal solvent extraction process at (a) 1073 K and (b) 2073 K from dynamics simulations with ReaxFF. (A colour version of this figure can be viewed online.)



**Fig. 9.** The 119 molecules in the range of 500–1000 Da from Blue Gem + NAP products at 2073 K. The hydrogen, carbon, nitrogen, oxygen, and sulfur atoms are colored in white, grey, blue, red, yellow, respectively, while the carbon and hydrogen atoms from solvent molecules are colored in green. Molecules in the grey box are coal molecules that react with solvent radicals. (A colour version of this figure can be viewed online.)

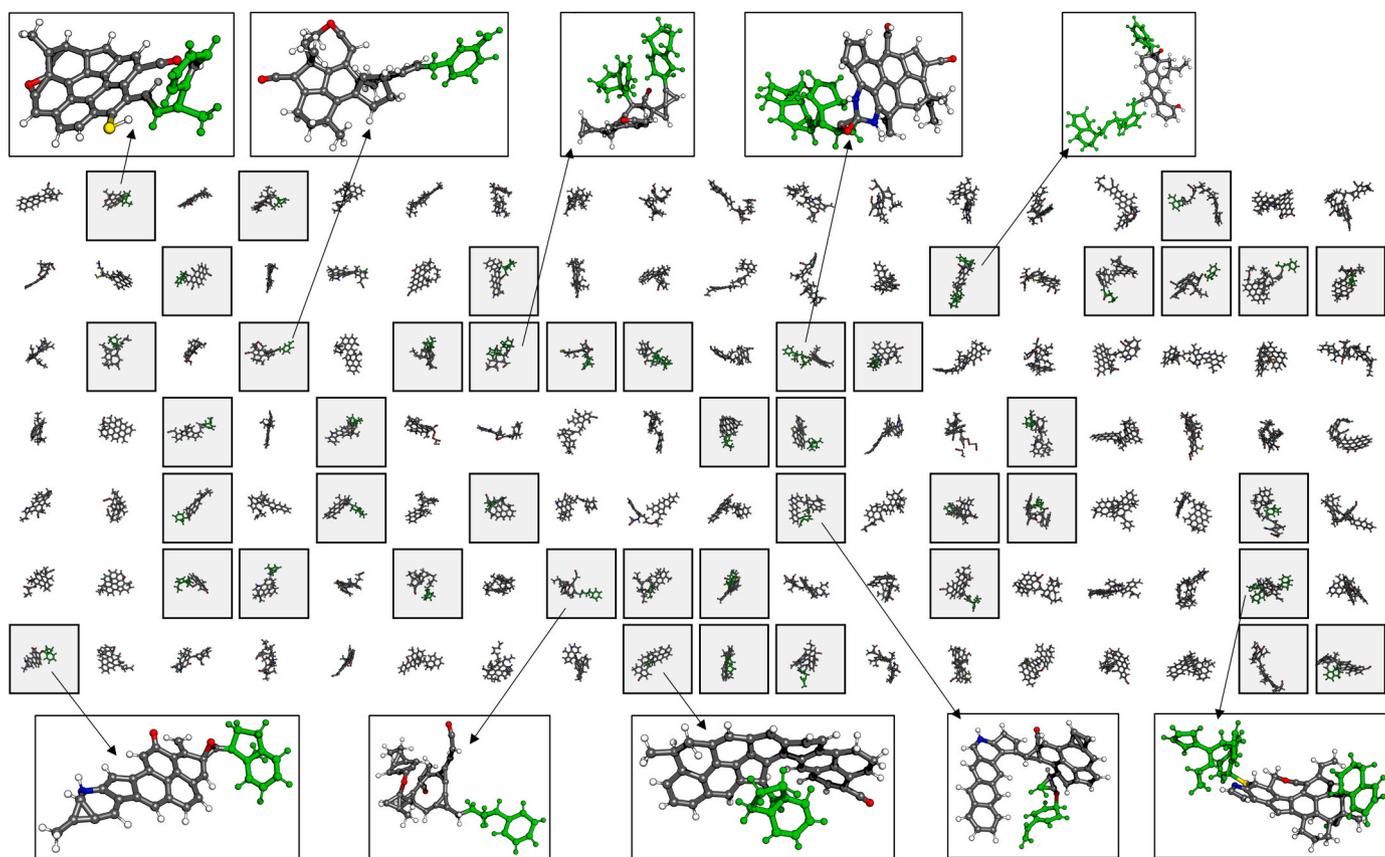
formation of gas molecules, while TET produces more than twice as many gas molecules due to the decomposition of TET and coal molecules (Fig. 7a). In Blue Gem + NAP, the number of gas molecules was very close to the pyrolysis of Blue Gem, implying that the hydrogen donation capability of NAP is limited and radicals from solvents are challenging to form.

The number of small molecules (100–200 Da in Fig. 7b) was the largest in Blue Gem + TET, with the decomposition of TET and liberation of small PAH molecules bridged in large molecules of the coal. The lower number in Blue Gem + NAP was attributed to the higher stability of NAP compared to TET. Note that most of the aromatic rings of TET remained intact, while the aliphatic fractions were easily decomposed at

this high-temperature condition. We observed that TET decomposes faster than NAP, with hydrogen donation and dissociation of the cycloalkanes chain of TET at the same temperature condition.

The molecules in the range of 200–500 Da (Fig. 7c) were significantly increased in two cases: in one case, the crosslinking of coal molecules was dissociated, liberating the small PAH molecules. There was also bond formation between radicals of PAH molecules either from dissociated solvent molecules or small PAH molecules from coal. This was consistent with decreasing the number of heavier nonvolatile fractions between 500 and 3000 Da.

We investigated the average variations and deviations in the evolution of molecular species by conducting an additional set of simulations



**Fig. 10.** The 126 molecules in the range of 500~1000 Da from Blue Gem + TET products at 2073 K. The hydrogen, carbon, nitrogen, oxygen, and sulfur atoms are colored in white, grey, blue, red, yellow, respectively, while the carbon and hydrogen atoms from solvent molecules are colored in green. Molecules in the grey box are coal molecules that react with solvent radicals. (A colour version of this figure can be viewed online.)

with solids featuring molecules distributed differently. These simulations were carried out at a high temperature of 2073 K (see Fig. S8). The results are comparable, and no significant deviation among molecular species across all molecular weight ranges was observed.

The bond order analysis also tracked the molecular structure information, chemical formula, and origin of each atom. We referenced the bond orders and the connectivity to obtain Cartesian coordinates of individual molecules from large-scale simulations. Figs. 9 and 10 display product molecules of solvent extraction with NAP and TET at 1 ns in the range of 500~1000 Da. We colored the carbon and hydrogen atoms from NAP or TET molecules in green to distinguish them from the atoms of the original coal molecules. In addition, the coal molecules that reacted with solvent molecules were highlighted with grey boxes, and a few cases were magnified to examine their structures. The structure files are provided in the database for further examination.

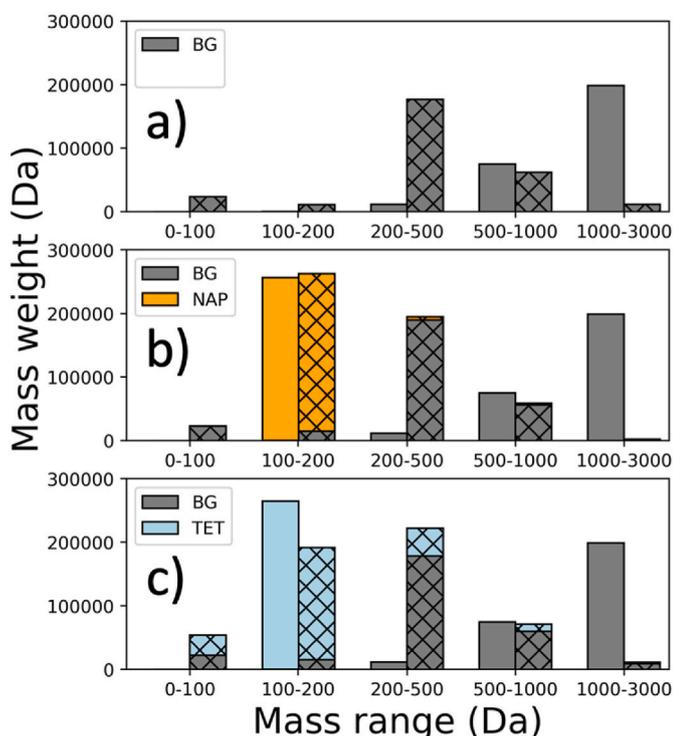
We observed that 45 coal molecules reacted with TET and 19 coal molecules reacted with NAP, respectively. This demonstrated that the high reactivity and hydrogen donation capability of TET provide many radicals and proton transfer to radicals of coal molecules and form aryl bonding with radicals of TET. In particular, the cyclic aliphatic chains of TET at 2073 K were broken and reacted with coals or other TET molecules while the aromatic carbons of NAP were mainly dehydrogenated and reacted with functional groups of coal molecules. The aromatic hydrocarbons are inert, and the naphthenic hydrocarbons are easier to crack but not released as gas molecules. High reactivity and cracking of cyclic aliphatic chains or naphthenic hydrocarbons are consistent with an experimental report about preparing desired precursors for carbon fiber production [89].

This implies that the final products of the coal conversion with solvent extraction can depend on the composition and properties of

solvents. A complete list of product gas molecules of the solvent extraction simulations is provided in Figs. S12 and S13 for the molecular structures and Tables S3 and S4 for the chemical formula, mass, and numbers with NAP and TET, respectively.

The conversion of large coal fractions would be advantageous for producing high-valued carbonaceous products (assuming all the molecules are desirable). Solvent extraction is a more effective procedure for obtaining valuable low molecular weight fractions than pyrolysis and combustion, which result in a significant amount of material being converted into char and coke. NAP had limited reactivity to form radicals and hydrogen donation in our simulations, which is consistent with experimental data [90]. At the same time, TET was significantly more reactive, quickly forming many radicals and hydrogen donations to create nonvolatile fractions. The formation of isomers and radicals from TET was reported in an evaluation focused on coal liquefaction [91,92]. Consistent with their experimental observation of naphthalene and the formation of other radicals, the smaller molecules were formed from the hydrogen donation and decomposition of tetralin in the simulations.

Fig. 11 and Table S2 show the origin of fraction in the range of MW by tracking the atomic indices at the end of the simulations. The fraction at  $t = 1$  ns exhibited minor variations when derived from coal through pyrolysis, solvent extractions with NAP, and TET. Notably, a distinction arose in the case of NAP and TET, wherein additional fractions underwent conversion from solvents to products for 200~1000 Da. Assuming gas and volatiles (0~200 Da), including solvents, could be quickly evaporated and removed from residues of higher MW (200~3000 Da) at the experimental temperature condition (300~400 °C) after the solvent extraction process, the coal extraction process with solvents could result in an exchange of materials that contribute to the volatiles fraction after breaking cross-links between coal and solvent. Specifically, the gas



**Fig. 11.** The fractions at each MW range for  $T = 2073$  K at 0 ns (left bar chart) and 1 ns (right bar chart with hash pattern) for a) Blue Gem (BG), b) Blue Gem + NAP, and c) Blue Gem + TET. (A colour version of this figure can be viewed online.)

molecules with NAP were derived from the dissociation of coal molecules functional groups and cyclic alkanes by 99 %. The gas molecules comprised atoms originating from coal and solvent at 41 % and 59 %, respectively. The high hydrogen donation capability enabled functional groups to protonate and break from coal molecules easily, leading to TET decomposition. Concurrently, the radicals formed with proton transfer were occasionally bonded to radicals of coal molecules in the high MW range. This led to the exchange of solvent and coal for nonvolatile fractions (200–3000 Da). The exchanged amount with NAP was 3–12 %, while that with TET was 16–20 %.

In this work, we report results obtained with the reactive MD simulations using ReaxFF parameter for hydrocarbon systems reported by Mattsson et al. [79], which they used to study polymer shock reactions. We also conducted simulations using the same conditions with another ReaxFF parameterized by Nielsen et al. [80], which was utilized to study carbon nanotube formation using a transition metal catalyst such as Ni nanoparticles, and the results were included in the supplementary materials (Figs. S9–S11). Even if there were differences in the molecular structures and chemical formula of the chemical intermediates and products, a consistent outcome for thermolysis of coal and solvent extraction was obtained, that is, breaking cross-linking and generating gaseous products.

#### 4. Conclusions

In this study, using data from experimental characterization, we constructed a large-scale atomistic solid coal model ([https://github.com/MoleCraftHUB/MSB\\_Coal\\_DB](https://github.com/MoleCraftHUB/MSB_Coal_DB)) containing ~40,000 atoms using our newly developed open-source Python program, “MolecularSolidBuilder” (<https://github.com/MoleCraftHUB/MolecularSolidBuilder>). The computational IR spectra through DFTB and ReaxFF calculations showed comparable peak positions to the experimental IR spectra. Systematic reactive MD simulations through ReaxFF were further performed to investigate the solvent extraction process of Blue

Gem coal with two types of aromatic solvents, naphthalene, and the hydrogen donor solvent, tetralin, at two temperature conditions (1073 and 2073 K). The products of solvent extraction were significantly different due to the solvent types in terms of ratios of aliphatic chains coming from solvent molecules. Naphthalene was relatively stable, rarely forming radicals and donating hydrogen to coal and other naphthalene molecules. Tetralin was consumed rapidly by forming radicals with alkanes and donating hydrogens to coal or other tetralin molecules.

Additionally, we observed that the fraction of molecules in different molecular weight ranges evolved differently with naphthalene or tetralin. More fractions of tetralin molecules were part of coal molecules, and significant amounts of volatile components were generated by the scissions of aliphatic bonds, mainly oxygen bridges or methyl cross-linking. The coal fraction involved in a molecular exchange with the naphthalene solvent was ~3–12 %, compared to ~15–20% with tetralin. Our results provide a better understanding of the role of hydrogenation in the solvent extraction process and will be valuable in guiding the preparation of precursors for value-added products obtained from pitches made by coal solvent extraction.

This study focuses on the construction and modeling of experimentally characterized coal samples. The developed software aims to elucidate the initial phase and structure of coals in their solid state, specifically to investigate the solvent extraction process at elevated temperatures. We propose that this methodology shows potential for expanding its application to cover the gradual conversion of coal, thereby facilitating theoretical understanding relevant to producing tar and pitches as crucial precursors for value-added products. This extension entails the integration of both reactive MD simulations and exploring hypothetical modifications to molecular structures, guided by in-situ characterizations. Our software tightly integrates experiment and theory, which is expected to advance fundamental understanding of coal conversion processes and allow insights for developing valuable industrial products.

#### CRediT authorship contribution statement

**Pilsun Yoo:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – original draft, Writing – review & editing. **Gang Seob Jung:** Writing – review & editing. **Matthew R. Ryder:** Formal analysis, Methodology, Writing – review & editing. **Frederic Vautard:** Data curation, Formal analysis, Methodology, Writing – review & editing. **Ercan Cakmak:** Data curation, Formal analysis, Methodology. **Sungsool Wi:** Data curation. **Matthew C. Weisenberger:** Data curation, Formal analysis. **Edgar Lara-Curzio:** Funding acquisition, Project administration, Writing – review & editing. **Jonathan P. Mathews:** Data curation, Formal analysis, Methodology, Resources, Writing – review & editing. **Stephan Irl:** Formal analysis, Supervision, Writing – review & editing.

#### 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.

#### Acknowledgment

This research work was sponsored by the U.S. Department of Energy Fossil Energy and Carbon Management Program, Advanced Coal Processing Program, C4WARD project (FEAA155). This research used the resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. A portion of this work was performed at the National High Magnetic Field Laboratory NMR User Facility, which is supported by the

National Science Foundation Division of Chemistry and Division of Materials Research through DMR-2128556 and the State of Florida. The authors also acknowledge the University of Kentucky Center for Applied Energy Research analytical group for the proximate and ultimate analyses of the coals.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.carbon.2024.118939>.

## References

- [1] M.W. Haanel, Recent progress in coal structure research, *Fuel* 71 (11) (1992) 1211–1223.
- [2] H. Gagarin, S. Sridhar, I. Lange, M.D. Bazilian, Considering non-power generation uses of coal in the United States, *Renew. Sustain. Energy Rev.* 124 (2020) 109790, <https://doi.org/10.1016/j.rser.2020.109790>.
- [3] R. Atkins, in: *Coal in a New Carbon Age: Powering a Wave of Innovation in Advanced Products & Manufacturing*, 2019. <https://www.nationalcoalcoalouncil.org/>. (Accessed 18 October 2023).
- [4] D05 Committee, Classification of Coals by Rank, ASTM International, West Conshohocken, PA, 2019, <https://doi.org/10.1520/d0388-19a>.
- [5] Y. Chen, M. Mastalerz, A. Schimmelmann, Characterization of chemical functional groups in macerals across different coal ranks via micro-FTIR spectroscopy, in: 28th Int. Pittsburgh Coal Conf., vol. 104, 2012, pp. 22–33, <https://doi.org/10.1016/j.coal.2012.09.001>.
- [6] M. Steller, W. Kalkreuth, I. Wieschenkämper, Effect of Vitrinite and Inertinite Fluorescence Properties on Combustion and Hydrogenation Reactivities of Coals, Elsevier, 1991, pp. 90–94.
- [7] M. Ozer, O.M. Basha, G. Stiegel, B. Morsi, 7 - effect of coal nature on the gasification process, in: T. Wang, G. Stiegel (Eds.), *Integrated Gasification Combined Cycle (IGCC) Technologies*, Woodhead Publishing, 2017, pp. 257–304, <https://doi.org/10.1016/B978-0-08-100167-7.00007-X>.
- [8] R.A. Keogh, D.N. Taulbee, J.C. Hower, B. Chawla, B.H. Davis, Liquefaction characteristics of the three major maceral groups separated from a single coal, *Energy Fuels* 6 (5) (1992) 614–618.
- [9] J.W. Larsen, T.K. Green, J. Kovac, The nature of the macromolecular network structure of bituminous coals, *J. Org. Chem.* 50 (24) (1985) 4729–4735, <https://doi.org/10.1021/jo00224a014>.
- [10] P.C. Painter, J. Graf, M.M. Coleman, Coal solubility and swelling. 3. A model for coal swelling, *Energy Fuels* 4 (4) (1990) 393–397, <https://doi.org/10.1021/ef00022a010>.
- [11] M. Nishioka, The associated molecular nature of bituminous coal, *Fuel* 71 (8) (1992) 941–948.
- [12] D.W. Van Krevelen, *Coal: Typology-Physics-Chemistry-Constitution*, 1993.
- [13] J. Milligan, K. Thomas, J. Crelling, Solvent swelling of maceral concentrates, *Energy Fuels* 11 (2) (1997) 364–371.
- [14] A.L. Chaffee, C.J.R. Fookes, Polycyclic aromatic hydrocarbons in Australian coals—III. Structural elucidation by proton nuclear magnetic resonance spectroscopy, *Org. Geochem.* 12 (3) (1988) 261–271, [https://doi.org/10.1016/0146-6380\(88\)90263-X](https://doi.org/10.1016/0146-6380(88)90263-X).
- [15] M.S. Solum, R.J. Pugmire, D.M. Grant, Carbon-13 solid-state NMR of argonne-premium coals, *Energy Fuels* 3 (2) (1989) 187–193, <https://doi.org/10.1021/ef00014a012>.
- [16] D. Bodzek, A. Marzec, Molecular components of coal and coal structure, *Fuel* 60 (1) (1981) 47–51, [https://doi.org/10.1016/0016-2361\(81\)90030-2](https://doi.org/10.1016/0016-2361(81)90030-2).
- [17] J. Pajares, J.M.D. Tascón, *Coal Science*, Elsevier, 1995.
- [18] A. Marzec, H.-R. Schulten, Structure and reactivity of coals: field ionization mass spectrometry viewpoint, *Fuel* 73 (8) (1994) 1294–1305.
- [19] D. Van Niekerk, R.J. Pugmire, M.S. Solum, P.C. Painter, J.P. Mathews, Structural characterization of vitrinite-rich and inertinite-rich permian-aged South African bituminous coals, *Int. J. Coal Geol.* 76 (4) (2008) 290–300.
- [20] M.R. Narkiewicz, J.P. Mathews, Improved low-volatile bituminous coal representation: incorporating the molecular-weight distribution, *Energy Fuels* 22 (5) (2008) 3104–3111, <https://doi.org/10.1021/ef700779j>.
- [21] Q. Zhong, Q. Mao, L. Zhang, J. Xiang, J. Xiao, J.P. Mathews, Structural features of qingdao petroleum coke from HRTEM lattice fringes: distributions of length, orientation, stacking, curvature, and a large-scale image-guided 3D atomistic representation, *Carbon* 129 (2018) 790–802.
- [22] S.R. Kelemen, M.L. Gorbaty, P.J. Kwiatek, Quantification of nitrogen forms in Argonne Premium coals, *Energy Fuels* 8 (4) (1994) 896–906, <https://doi.org/10.1021/ef00046a013>.
- [23] E. Cakmak, A. Annamraju, J.P. Mathews, L. He, N. Gallego, E. Lara-Curzio, Characterization of porosity and pore accessibility of vitrinite-rich bituminous and subbituminous coals by small-angle neutron scattering, mercury intrusion porosimetry, and low-pressure N<sub>2</sub> adsorption, *Energy Fuels* 37 (1) (2023) 191–203, <https://doi.org/10.1021/acs.energyfuels.2c02814>.
- [24] E. Cakmak, J.C. Hower, J.P. Mathews, M.C. Weisenberger, R. Kaplan, J. Lacy, Y. Zhang, E. Lara-Curzio, Microstructural diversity and digestion yields of select bituminous and subbituminous coals as raw material candidates for carbon fiber precursor production, *Fuel* 348 (2023) 128545, <https://doi.org/10.1016/j.fuel.2023.128545>.
- [25] J.H. Shinn, From coal to single-stage and two-stage products: a reactive model of coal structure, *Fuel* 63 (9) (1984) 1187–1196, [https://doi.org/10.1016/0016-2361\(84\)90422-8](https://doi.org/10.1016/0016-2361(84)90422-8).
- [26] T. Takanohashi, M. Iino, K. Nakamura, Simulation of interaction of coal associates with solvents using the molecular dynamics calculation, *Energy Fuels* 12 (6) (1998) 1168–1173, <https://doi.org/10.1021/ef980042x>.
- [27] D.V. Niekerk, J.P. Mathews, Molecular representations of permian-aged vitrinite-rich and inertinite-rich South African coals, *Fuel* 89 (1) (2010) 73–82, <https://doi.org/10.1016/j.fuel.2009.07.020>.
- [28] D. Van Niekerk, J.P. Mathews, Molecular dynamic simulation of coal-solvent interactions in permian-aged South African coals, *Sel. Pap. Int. Conf. Coal Sci. Technol.* 92 (4) (2011) 729–734, <https://doi.org/10.1016/j.fuproc.2010.08.011>, 2009.
- [29] F. Castro-Marcano, V.V. Lobodin, R.P. Rodgers, A.M. McKenna, A.G. Marshall, J. P. Mathews, A molecular model for Illinois No. 6 Argonne Premium coal: moving toward capturing the continuum structure, *Fuel* 95 (2012) 35–49, <https://doi.org/10.1016/j.fuel.2011.12.026>.
- [30] M. Zheng, Z. Wang, X. Li, X. Qiao, W. Song, L. Guo, Initial reaction mechanisms of cellulose pyrolysis revealed by ReaxFF molecular dynamics, *Fuel* 177 (2016) 130–141, <https://doi.org/10.1016/j.fuel.2016.03.008>.
- [31] M. Zheng, X. Li, F. Nie, L. Guo, Investigation of molecular scale effects on coal pyrolysis using ReaxFF MD simulation, *Mol. Simulat.* 43 (13–16) (2017) 1081–1088, <https://doi.org/10.1080/08927022.2017.1356456>.
- [32] M. Gao, X. Li, C. Ren, Z. Wang, Y. Pan, L. Guo, Construction of a multicomponent molecular model of fugu coal for ReaxFF-MD pyrolysis simulation, *Energy Fuels* 33 (4) (2019) 2848–2858, <https://doi.org/10.1021/acs.energyfuels.8b04434>.
- [33] X. Li, M. Zheng, C. Ren, L. Guo, ReaxFF molecular dynamics simulations of thermal reactivity of various fuels in pyrolysis and combustion, *Energy Fuels* 35 (15) (2021) 11707–11739, <https://doi.org/10.1021/acs.energyfuels.1c01266>.
- [34] H. Yan, B. Nie, C. Peng, P. Liu, X. Wang, F. Yin, J. Gong, Y. Wei, S. Lin, Molecular model construction of low-quality coal and molecular simulation of chemical bond energy combined with materials studio, *Energy Fuels* 35 (21) (2021) 17602–17616.
- [35] Y. Zhang, Applications of noncontact atomic force microscopy in petroleum characterization: opportunities and challenges, *Energy Fuels* 35 (18) (2021) 14422–14444, <https://doi.org/10.1021/acs.energyfuels.1c02193>.
- [36] P. Chen, J.N. Metz, A.S. Gross, S.E. Smith, S.P. Rucker, N. Yao, Y. Zhang, Ex situ and in situ thermal transformations of M-50 pitch revealed by non-contact atomic force microscopy, *Energy Fuels* 35 (22) (2021) 18210–18219, <https://doi.org/10.1021/acs.energyfuels.1c02487>.
- [37] J.P. Mathews, A.C. Van Duin, A.L. Chaffee, The utility of coal molecular models, *Fuel Process. Technol.* 92 (4) (2011) 718–728.
- [38] J.P. Mathews, A.L. Chaffee, The molecular representations of coal – a review, *Fuel* 96 (2012) 1–14, <https://doi.org/10.1016/j.fuel.2011.11.025>.
- [39] M. Gao, X. Li, C. Ren, Z. Wang, Y. Pan, L. Guo, Construction of a multicomponent molecular model of fugu coal for ReaxFF-MD pyrolysis simulation, *Energy Fuels* 33 (4) (2019) 2848–2858, <https://doi.org/10.1021/acs.energyfuels.8b04434>.
- [40] Y.N. Zhuravlev, A. Porokhnov, Computer simulation of coal organic mass structure and its sorption properties, *Int. J. Coal Sci. Technol.* 6 (2019) 438–444.
- [41] Y. Zhang, S. Hu, Q. Zhong, J. Zhuo, J.P. Mathews, A large-scale molecular model of fenghuangshan anthracite coal, *Fuel* 295 (2021) 120616, <https://doi.org/10.1016/j.fuel.2021.120616>.
- [42] P. Hou, Y. Zhou, W. Guo, P. Ren, Q. Guo, H. Xiang, Y.-W. Li, X.-D. Wen, Y. Yang, Rational design of hydrogen-donor solvents for direct coal liquefaction, *Energy Fuels* 32 (4) (2018) 4715–4723.
- [43] G. Cheng, B. Tan, Z. Zhang, S. Fu, W. Haiyan, F. Wang, Characteristics of coal-oxygen chemisorption at the low-temperature oxidation stage: DFT and experimental study, *Fuel* 315 (2022) 123120.
- [44] G.-Y. Li, F. Wang, J.-P. Wang, Y.-Y. Li, A.-Q. Li, Y.-H. Liang, ReaxFF and DFT study on the sulfur transformation mechanism during the oxidation process of lignite, *Fuel* 181 (2016) 238–247, <https://doi.org/10.1016/j.fuel.2016.04.068>.
- [45] F. Castro-Marcano, A.M. Kamat, M.F. Russo, A.C.T. van Duin, J.P. Mathews, Combustion of an Illinois No. 6 coal char simulated using an atomistic char representation and the ReaxFF reactive force field, *Combust. Flame* 159 (3) (2012) 1272–1285, <https://doi.org/10.1016/j.combustflame.2011.10.022>.
- [46] M. Zheng, X. Li, L. Guo, Dynamic trends for char/soot formation during secondary reactions of coal pyrolysis by large-scale reactive molecular dynamics, *J. Anal. Appl. Pyrolysis* 155 (2021) 105048, <https://doi.org/10.1016/j.jaap.2021.105048>.
- [47] D. Van Niekerk, J.P. Mathews, Simulation of solvent extraction of South African vitrinite- and inertinite-rich coals, *Energy Fuels* 24 (12) (2010) 6393–6399, <https://doi.org/10.1021/ef101055f>.
- [48] D. Van Niekerk, F. Castro-Marcano, C.M. Colina, J.P. Mathews, Solvent swelling extent of permian-aged vitrinite- and inertinite-rich coals: experiments and modeling using the perturbed-chain statistical associating fluid theory (PC-SAFT), *Energy Fuels* 25 (6) (2011) 2559–2564, <https://doi.org/10.1021/ef2001508>.
- [49] D. Van Niekerk, G.D. Mitchell, J.P. Mathews, Petrographic and reflectance analysis of solvent-swelled and solvent-extracted South African vitrinite-rich and inertinite-rich coals, *Int. J. Coal Geol.* 81 (1) (2010) 45–52, <https://doi.org/10.1016/j.coal.2009.10.021>.
- [50] D.V. Niekerk, P.M. Halleck, J.P. Mathews, Solvent swelling behavior of permian-aged South African vitrinite-rich and inertinite-rich coals, *Fuel* 89 (1) (2010) 19–25, <https://doi.org/10.1016/j.fuel.2009.06.028>.
- [51] F. Castro-Marcano, J.P. Mathews, Constitution of Illinois No. 6 Argonne Premium coal: a review, *Energy Fuels* 25 (3) (2011) 845–853.

- [52] S. Liu, L. Wei, Q. Zhou, T. Yang, S. Li, Q. Zhou, Simulation strategies for ReaxFF molecular dynamics in coal pyrolysis applications: a review, *J. Anal. Appl. Pyrolysis* (2023) 105882.
- [53] Q. Mao, S. Rajabpour, M. Kowalik, A.C.T. van Duin, Predicting cost-effective carbon fiber precursors: unraveling the functionalities of oxygen and nitrogen-containing groups during carbonization from ReaxFF simulations, *Carbon* 159 (2020) 25–36, <https://doi.org/10.1016/j.carbon.2019.12.008>.
- [54] Q. Mao, S. Rajabpour, M.K. Talkhoncheh, J. Zhu, M. Kowalik, A.C. Duin, Cost-effective carbon fiber precursor selections of polyacrylonitrile-derived blend polymers: carbonization Chemistry and structural characterizations, *Nanoscale* 14 (17) (2022) 6357–6372, <https://doi.org/10.1039/D2NR00203E>.
- [55] S. AlAreeqi, D. Bahamon, K. Polychronopoulou, L.F. Vega, Insights into the thermal stability and conversion of carbon-based materials by using ReaxFF reactive force field: recent advances and future directions, *Carbon* 196 (2022) 840–866, <https://doi.org/10.1016/j.carbon.2022.05.035>.
- [56] A. Pott, H. Broche, The solution of coal by extraction under pressure—the hydrogenation of the extract, *Fuel* 13 (3) (1934) 91–96.
- [57] H. Hao, T. Chang, L. Cui, R. Sun, R. Gao, Theoretical study on the mechanism of hydrogen donation and transfer for hydrogen-donor solvents during direct coal liquefaction, *Catalysts* 8 (12) (2018) 648.
- [58] E.J. Kuhlmann, D.Y. Jung, R.P. Gupta, C.A. Dyke, H.K. Zang, Coal liquefaction using a hydrogenated creosote oil solvent: H-atom transfer from hydrogen donor components in the solvent, *Fuel* 64 (11) (1985) 1552–1557.
- [59] I. Mochida, K. Kudo, N. Fukuda, K. Takeshita, R. Takahashi, Carbonization of pitches—IV: carbonization of polycyclic aromatic hydrocarbons under the presence of aluminum chloride catalyst, *Carbon* 13 (2) (1975) 135–139.
- [60] Y. Yamada, S. Matsumoto, K. Fukuda, H. Honda, Optically anisotropic texture in tetrahydroquinoline soluble matter of carbonaceous mesophase, *Tanso* 1981 (107) (1981) 144–146.
- [61] E. Cakmak, J.P. Mathews, S. Wi, M.R. Ryder, M.L. Chacon-Patino, A.M. McKenna, F. Vautard, M. Arnould, H. Meyer III, E. Lara-Curzio, Detailed characterization of vitrinite-rich subbituminous and bituminous coals for utilization in carbon fiber precursor production, *Energy Fuels* (2024). *To be published*.
- [62] J.P. Mathews, V. Fernandez-Also, A. Daniel Jones, H.H. Schobert, Determining the molecular weight distribution of pocahontas No. 3 low-volatile bituminous coal utilizing HRTEM and laser desorption ionization mass spectra data, 17th Int. Symp. Alcohol, *Fuel* 89 (7) (2010) 1461–1469, <https://doi.org/10.1016/j.fuel.2009.10.014>.
- [63] J.-M. Leyssale, P. Weisbecker, G. Chollon, B. Farbos, G.-L. Vignoles, J.-P. Da Costa, Analysis and Molecular Modeling of Pyrolytic Carbons Nano-Textures, John Wiley and Sons, Inc, United States, 2014.
- [64] P. Tosco, N. Stiefl, G. Landrum, Bringing the MMFF force field to the RDKit: implementation and validation, *J. Cheminform.* 6 (2014) 37, <https://doi.org/10.1186/s13321-014-0037-3>.
- [65] A.C.T. van Duin, S. Dasgupta, F. Lorant, W.A. Goddard, ReaxFF: a reactive force field for hydrocarbons, *J. Phys. Chem. A* 105 (41) (2001) 9396–9409, <https://doi.org/10.1021/jp004368u>.
- [66] K. Chenoweth, A.C.T. van Duin, W.A. Goddard, ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation, *J. Phys. Chem. A* 112 (5) (2008) 1040–1053, <https://doi.org/10.1021/jp709896w>.
- [67] H.M. Aktulga, J.C. Fogarty, S.A. Pandit, A.Y. Grama, Parallel reactive molecular dynamics: numerical methods and algorithmic techniques, *Parallel Comput.* 38 (4) (2012) 245–259, <https://doi.org/10.1016/j.parco.2011.08.005>.
- [68] L. Liu, Y. Liu, S.V. Zybin, H. Sun, W.A.I. Goddard, ReaxFF-Ig: correction of the ReaxFF reactive force field for London dispersion, with applications to the equations of state for energetic materials, *J. Phys. Chem. A* 115 (40) (2011) 11016–11022, <https://doi.org/10.1021/jp201599t>.
- [69] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P. S. Crozier, P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Tritt, S.J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* 271 (2022) 108171, <https://doi.org/10.1016/j.cpc.2021.108171>.
- [70] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.* 81 (1) (1984) 511–519.
- [71] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, *Phys. Rev. A* 31 (3) (1985) 1695.
- [72] L. Martínez, R. Andrade, E.G. Birgin, J.M. Martínez, PACKMOL: a package for building initial configurations for molecular dynamics simulations, *J. Comput. Chem.* 30 (13) (2009) 2157–2164.
- [73] W.J. Mortier, S.K. Ghosh, S. Shankar, Electronegativity-equalization method for the calculation of atomic charges in molecules, *J. Am. Chem. Soc.* 108 (15) (1986) 4315–4320.
- [74] G.O. Janssens, B.G. Baekelandt, H. Toufar, W.J. Mortier, R.A. Schoonheydt, Comparison of cluster and infinite crystal calculations on zeolites with the electronegativity equalization method (EEM), *J. Phys. Chem.* 99 (10) (1995) 3251–3258.
- [75] C. Jian, J.J. Adams, J.C. Grossman, N. Ferralis, Carbon fiber synthesis from pitch: insights from ReaxFF based molecular dynamics simulations, *Carbon* 176 (2021) 569–579, <https://doi.org/10.1016/j.carbon.2021.01.151>.
- [76] S. Rajabpour, Q. Mao, Z. Gao, M. Khajeh Talkhoncheh, J. Zhu, Y. Schwab, M. Kowalik, X. Li, A.C.T. van Duin, Low-temperature carbonization of polyacrylonitrile/graphene carbon fibers: a combined ReaxFF molecular dynamics and experimental study, *Carbon* 174 (2021) 345–356, <https://doi.org/10.1016/j.carbon.2020.12.038>.
- [77] Y. Chen, H. Chen, J. Wang, Y. Huang, Chemical kinetics of hexamethyldisiloxane pyrolysis: a ReaxFF molecular dynamics simulation study, *Int. J. Chem. Kinet.* 54 (7) (2022) 413–423, <https://doi.org/10.1002/kin.21570>.
- [78] F. Niefind, Q. Mao, N. Nayir, M. Kowalik, J.-J. Ahn, A.J. Winchester, C. Dong, R. A. Maniyara, J.A. Robinson, A.C. Duin, S. Pookpanratana, Watching (De) Intercalation of 2D metals in epitaxial graphene: insight into the role of defects, *Small* (2023) 2306554, <https://doi.org/10.1002/sml.202306554>.
- [79] T.R. Mattsson, J.M.D. Lane, K.R. Cochrane, M.P. Desjarlais, A.P. Thompson, F. Pierce, G.S. Grest, First-principles and classical molecular dynamics simulation of shocked polymers, *Phys. Rev. B* 81 (5) (2010) 054103.
- [80] K.D. Nielson, A.C.T. van Duin, J. Oxgaard, W.-Q. Deng, W.A. Goddard, Development of the ReaxFF reactive force field for describing transition metal catalyzed reactions, with application to the initial stages of the catalytic formation of carbon nanotubes, *J. Phys. Chem. A* 109 (3) (2005) 493–499, <https://doi.org/10.1021/jp046244d>.
- [81] M. Gao, X. Li, L. Guo, Pyrolysis simulations of fugu coal by large-scale ReaxFF molecular dynamics, *Fuel Process. Technol.* 178 (2018) 197–205, <https://doi.org/10.1016/j.fuproc.2018.05.011>.
- [82] M. Zheng, X. Li, F. Nie, L. Guo, Investigation of overall pyrolysis stages for liulin bituminous coal by large-scale ReaxFF molecular dynamics, *Energy Fuels* 31 (4) (2017) 3675–3683.
- [83] T. Zhang, X. Li, X. Qiao, M. Zheng, L. Guo, W. Song, W. Lin, Initial mechanisms for an overall behavior of lignin pyrolysis through large-scale ReaxFF molecular dynamics simulations, *Energy Fuels* 30 (4) (2016) 3140–3150, <https://doi.org/10.1021/acs.energyfuels.6b00247>.
- [84] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool, *Model. Simulat. Mater. Sci. Eng.* 18 (1) (2009) 015012.
- [85] L.L.C. Schrödinger, The PyMOL Molecular Graphics System, Version 1.8, 2015.
- [86] Y. Xiong, L. Jin, H. Yang, Y. Li, H. Hu, Insight into the aromatic ring structures of a low-rank coal by step-wise oxidation degradation, *Fuel Process. Technol.* 210 (2020) 106563, <https://doi.org/10.1016/j.fuproc.2020.106563>.
- [87] G.N. George, M.L. Gorbaty, S.R. Kelemen, M. Sansone, Direct determination and quantification of sulfur forms in coals from the Argonne Premium sample program, *Energy Fuels* 5 (1) (1991) 93–97, <https://doi.org/10.1021/ef00025a016>.
- [88] H. Gan, S. Nandi, P. Walker Jr., Nature of the porosity in American coals, *Fuel* 51 (4) (1972) 272–277.
- [89] J. Guo, Z. Li, B. Li, P. Chen, H. Zhu, C. Zhang, B. Sun, Z. Dong, X. Li, Hydrogenation of coal tar pitch for improved mesophase pitch molecular orientation and carbon fiber processing, *J. Anal. Appl. Pyrolysis* 174 (2023) 106146, <https://doi.org/10.1016/j.jaap.2023.106146>.
- [90] P. Hao, Z.-Q. Bai, Z.-T. Zhao, Z.-F. Ge, R.-R. Hou, J. Bai, Z.-X. Guo, L.-X. Kong, W. Li, Role of hydrogen donor and non-donor binary solvents in product distribution and hydrogen consumption during direct coal liquefaction, *Fuel Process. Technol.* 173 (2018) 75–80.
- [91] H. Chang, J. Li, S. Du, K. Shen, Q. Yang, H. Yi, J. Zhang, Transformation characteristics of hydrogen-donor solvent tetralin in the process of direct coal liquefaction, *Front. Chem.* 7 (2019) 737.
- [92] C. Koyunoglu, O. Eksioğlu, O. Yıldırım, H. Karaca, Co-liquefaction of yatağan lignite and waste tire under catalytic conditions. Part 1. Effect of fresh tetraline and recycled tetraline on the conversion, *Energy Sources, Part A Recovery, Util. Environ. Eff.* 40 (9) (2018) 1068–1075.