1 **Organic structural properties of kerogen as predictors of source rock type and**
2 **hydrocarbon potential**

3 Todd L. Longbottom¹, William C. Hockaday¹, Kenneth S. Boling¹, Gaoyuan Li², Yohan
4 Letourmy¹, Hailiang Dong^{2,3}, Stephen I. Dworkin¹

5 ¹*Department of Geosciences, Baylor University, Waco, TX, 76706, USA*

6 ²*State Key Laboratory of Biogeology and Environmental Geology, China University of*
7 *Geosciences, Beijing 100083, China*

8 ³*Department of Geology and Environmental Science, Miami University, Oxford, Ohio, 45056,*
9 *USA*

10
11 *Keywords: Kerogen, ¹³C NMR, pyrolysis, Cretaceous source rock, coal*

Abstract

This study improves upon previously identified correlations between the chemical structure of kerogen and potential hydrocarbon (oil and gas) yields assayed by Rock-Eval pyrolysis. We propose a quantitative structure-catagenesis relationship that predicts the hydrocarbon generation potential of source rocks and of lacustrine, marine, and terrestrial origin (types I, II, and humic coals). We used one-dimensional solid-state ^{13}C Nuclear Magnetic Resonance (^{13}C NMR) spectroscopy with ^1H spectral editing to determine the abundance of carbon functional groups, including non-protonated and mobile groups. An NMR-based van Krevelen analysis readily separated the kerogen types. Single regression matrices of NMR-based structure parameters against Rock-Eval hydrocarbon yield revealed distinct dynamics of the kerogen types upon pyrolysis. Multiple regression showed that alkyl, oxygen-substituted alkyl, and carbonyl groups were strong contributors to hydrocarbon production, while oxygen-substituted aromatic carbons were strongly counterproductive. Catagenetic relationships established for kerogen provide insight into kerogen structure evolution upon pyrolysis, and can more closely constrain the mechanisms of hydrocarbon generation for use in sedimentary basin modelling.

1 - Introduction

Kerogen—the insoluble fraction of organic matter in sediments—is a vast and economically important reservoir of sedimentary organic carbon (Tissot and Welte, 1984). Kerogen is the dominant source of carbon for hydrocarbon generation in the subsurface. Therefore, the chemical characterization of kerogens and their hydrocarbon generation status, or hydrocarbon generation potential is a major concern of exploration and production. However, due to the insolubility and chemical heterogeneity of kerogen, there are few analytical techniques capable of providing quantitative molecular-level information on the chemical structure of kerogens (Boucher et al., 1990; Salmon et al., 2011). Kerogen characterization methods are often destructive, requiring chemolytic treatments, such as sequential alkaline permanganate degradation (Bajc et al., 2001; Rullkötter and Michaelis, 1990), ruthenium tetroxide oxidation (Boucher et al., 1990), or pyrolytic “cracking” of kerogen into more readily identifiable monomers (Eglinton et al., 1991; Horsfield, 1989; Philp and Gilbert, 1987). Advances in analytical pyrolysis, such as catalytic hydropyrolysis, can yield biomarker hydrocarbons from kerogen, asphaltenes, and recent sediments with little alteration to biomarker structure or stereochemistry (Bennett et al., 2004; Love et al., 1995; Russell et al., 2004; Snape et al., 1989). However, degradative techniques generate products that cannot be unambiguously related to the native kerogen structure. To gain a comprehensive understanding of the kerogen geomacromolecule, a combination of analytical techniques is beneficial (Bharati et al., 1995) while theoretical modeling efforts bolster characterization by constraining molecular formulae and multi-dimensional structure (Lille et al., 2003).

Solid-state spectroscopy techniques, such as, ^{13}C nuclear magnetic resonance (NMR) are powerful for elucidating the structure of kerogens without bond cleavage (Cao et al., 2013). When used in combination with hydrocarbon generation assays, quantitative NMR analyses of kerogen chemical structure can provide powerful insight to mechanisms of petroleum formation. For example, NMR spectral features of kerogen are highly-correlated with the oil and gas potential of the source rock (Maciel et al., 1978, 1979; Miknis et al., 1982; Qin et al., 1991; Witte et al., 1988). The genetic potential of oil shale is largely dependent on the abundance of aliphatic carbon (Takeda and Asakawa, 1988). The carbon functional groups in kerogen, as delineated by ^{13}C NMR, generally have prominent signals in the chemical shift regions corresponding to aliphatic (0 – 90 ppm), olefinic/aromatic (90 – 165 ppm), and carbonyl (165 – 215 ppm) carbon (Resing et al., 1978; Vitorović et al., 1978). The ratio of aliphatic to aromatic carbon (C) of kerogen has been widely employed in describing the effects of thermal maturity on kerogen structure (Barwise, 1984; Dennis et al., 1982; Patience et al., 1992; Smernik et al., 2006; Werner-Zwanziger et al., 2005), where sedimentary organic matter becomes increasingly aromatic and condensed with thermal maturation.

Hydrocarbon quantity and product distribution (e.g. oil/gas ratios, saturates, aromatics, acids, etc.) are influenced by the quantity and molecular composition of the kerogen, as determined by the source organisms and depositional environment. Therefore, source rocks are often classified (typed) as lacustrine algal (I), marine or planktonic (II), and terrestrial or humic (III). Ultimate chemical analyses (elemental and molecular) of thermally immature kerogens have shown that type classes differ primarily in the relative aromatic C content (Barwise, 1984; Mann et al., 1991; Vučelić et al., 1979). The mole fraction aromatic C (f_a) is initially much higher in sediments with

significant terrestrial influence, primarily due to the occurrence of lignin in higher plant tissue, while f_a is generally lower in aquatic (lacustrine or marine) organic matter that has not undergone thermal maturation (Mann et al., 1991). Ratios of aliphatic, aromatic, and carboxylic spectral regions have been shown to be diagnostic of differences in kerogen sources (i.e. kerogen types I-III), where aliphatic to aromatic C ratio decreases from type I to types II and III (Vučelić et al., 1979).

The primary aim of this study was to develop an empirical model that predicts the hydrocarbon potential of source rock types I, II, and coals from the molecular attributes of the kerogen fraction. The molecular features of kerogen were determined by solid-state ^{13}C NMR spectroscopy and hydrocarbon potential of whole rock samples were assayed by Rock Eval® pyrolysis. A secondary aim was to evaluate formal NMR-based algorithms for kerogen type analysis. Finally, we discuss the implications in the context of structure-catagenesis relationships.

3 - Materials & Methods

3.1 - Samples & kerogen isolation

Sediment samples of the Eagle Ford formation were collected from 9 outcrop localities near Waco, TX. These are marine (type II) muds deposited during the Cenomanian and Turonian stages of the late Cretaceous (Boling and Dworkin, 2015). Eleven samples of the Yixian Formation were collected from an outcrop in Liaoning Province, China. These are lacustrine, type I, sediments deposited in the early Cretaceous. Four coal samples were collected from an outcrop of the Carboniferous Joggins Formation, Bay of Fundy, Nova Scotia, Canada.

129

130 Bitumen was extracted for type II sediments and coals by Soxhlet apparatus, using an azeotrope
131 of dichloromethane and methanol to isolate the kerogen fraction. Five grams of each solid
132 residue was then demineralized using dilute (10 wt%) solutions of HCl and HF acids, using the
133 procedure recommended by Gélinas et al. (2001). Type I kerogens were acid demineralized
134 without prior bitumen extraction due to low organic matter content.

135

136 *3.2 - Elemental analysis & RockEval® pyrolysis*

137 Approximately 5 to 10 mg of pulverized bulk sediment was weighed into silver capsules and
138 titrated with 1 N HCl to remove carbonate C. Bulk sediment total organic carbon (TOC) and
139 nitrogen (N) were then determined by catalytic combustion in a Costech 4010 Elemental
140 Analyzer at Baylor University. Elemental H, C, N, and O were determined for the kerogen
141 isolates using a Thermo Finnigan Flash Elemental Analyzer at Galbraith laboratories in
142 Knoxville, TN.

143

144 RockEval® pyrolysis and maturity testing of bulk sediment was performed at Geomark Research
145 Ltd. source rock lab in Humble, TX using a Rock-Eval-2 system. The temperature program
146 began with a 300°C isotherm for 3 minutes allowing the free hydrocarbons to be volatilized (S1
147 peak), followed by a 25°C/minute temperature ramp to 550°C permitting the cracking and
148 volatilization of high molecular weight hydrocarbons (S2 peak), and finally a 600°C temperature
149 isotherm for 1 minute. The temperature at the maximum of the S2 peak is recorded as T_{max} ,
150 while the organic C-derived CO₂ evolved throughout pyrolysis (<600°C) was detected and

denoted as the S3 peak. Representative pyrograms for each sediment type are provided in the online supplementary materials (Figures S6- S8).

3.3 –Solid-State ^{13}C NMR spectroscopy

The NMR analyses were conducted on a standard bore 300 MHz Bruker Avance III spectrometer equipped with a 4 mm magic angle-spinning (MAS) probe, operating at a spin rate of 12 kHz. The cross-polarization (CP) data were acquired with a variable amplitude contact pulse (Metz et al., 1994) of 2 ms and composite pulse proton decoupling (TPPM15). The Hartman-Hahn matching conditions for ^1H - ^{13}C polarization transfer were optimized using crystalline glycine as an external standard. Dipolar-dephasing (non-quaternary suppression) spectral editing was conducted for both CP and DP experiments using the pulse sequence of Harbison et al. (1985), with a total dephasing delay of 77 μs (Figure S1). The optimal dephasing delay was determined experimentally as the shortest duration allowing for complete suppression of signal from non-quaternary carbons in glycine (N-Alkyl) and vanillyl alcohol (aromatic and methylenic) reference compounds (Sigma-Aldrich). Time-domain NMR signals were each exponentially-multiplied, zero filled to 16,384 data points, and Fourier transformed with 60 Hz line broadening. The resulting NMR spectra were fitted with a linear baseline and integrated using three previously-reported conventions, to facilitate the use of the NMR data for the purposes of: carbon functional group analysis (Smernik and Oades, 2001; Solum et al., 1989), analysis of hydrocarbon potential (Qin et al., 1991; Figure S2), and elemental H/C and O/C ratio calculation (Conte et al., 2002; Hockaday et al., 2009). The NMR algorithm for elemental ratio (C:H:O) determination was completed as an independent means of constructing van Krevelen diagrams for kerogen type analysis (van Krevelen, 1950; see sections S1.2. for further methodological detail). All NMR

peak areas were converted to units of mg/g using whole-rock and isolated kerogen TOC values for direct comparison to RockEval® parameters. The quantitative reliability of all NMR experiments was assessed using a spin counting procedure (Smernik and Oades, 2000a, b). These methods and results are described briefly in the supplemental online materials.

3.5 – Statistical analyses

The average response variable, S2, was predicted using two distinct multiple regression models. In the first model, S2 was predicted using CPMAS & DPMAS fractions (f_o , f_g , and f_a) defined by Qin et al. (1991) as predictor variables, while the second model incorporated the structural parameters listed in Table 1, measured by CPMAS and DPMAS with dipolar-dephasing according to methods of Solum et al. (1989). All statistical analyses were carried out using JMP Pro software version 11.0.

4 - Results & Discussion

4.1 - Elemental & RockEval analysis

Total organic carbon (TOC) concentrations of the bulk sediment samples reflect a range of organic-lean to organic-rich facies, with the majority of the samples being marked by “adequate” source potential (TOC >1%; Jarvie, 1991; Table 2). The S2 values follow a pattern similar to TOC, where most samples are characterized by a “good source potential” with values exceeding 5 mg HC/g (Espitalie, 1982). Thermal maturity was estimated by T_{max} , whereby the majority of type I and II kerogens are classified as immature with respect to petroleum generation, and several samples are characterized by maturities bordering the onset of oil generation (Espitalie et

al., 1985). Coal samples have experienced modest thermal maturity, as $T_{\max}$ values are within the operationally-defined oil window ($>430^{\circ}\text{C}$).

4.2 - NMR spectral properties of kerogens

In total, 49 kerogen samples were analyzed by ^{13}C CPMAS NMR with and without dipolar-dephasing, representing types I, II, and coal, and the carbon functional groups and structural parameters derived from NMR are summarized in Tables S4 and S5. Figure 1 shows typical ^{13}C CPMAS NMR spectra of the Yixian, Eagle Ford, and Joggins kerogens (types I, II, and coal, respectively), with important chemical differences related to the sourcing organisms and depositional environments.

4.3 - NMR-based kerogen typing

The structural characteristics that distinguish kerogen types, described qualitatively above, can be used as a basis for determining kerogen type in sediments containing organic matter of unknown origin. The NMR-based van Krevelen diagrams are compared to those based upon Rock-Eval and conventional elemental (C:H:O) analysis in Figure 2. Generally, the NMR-based method (Figures 2C) exhibits superior accuracy and precision relative to the combustion- and pyrolysis-based methods (Figures 2A and 2B; Table S6). The combustion elemental analysis of kerogen suffers from errors caused by incomplete removal of inorganic minerals, which appear in Figure 2A as elevated O/C ratios. Whereas kerogen typing errors in the Rock-Eval method (Figure 2B) may arise from the large “inert fraction” of kerogen (see S4 in Table 2) that is non-volatile, and therefore undetectable, under the programmed pyrolysis conditions.

219 The van Krevelen plot (Figure 2C) derived from the direct assignment (DA) NMR approach of
220 Hockaday et al. (2009) is consistent with the conventional elemental analysis (Figure 2A) to the
221 extent that type I kerogens of the Yixian formation were accurately predicted as such, and also
222 exhibit thermal immaturity due to the relative abundances of oxygen and hydrogen ($O/C > 0.5$,
223 $H/C > 1.0$). Kerogens approach the origin (100% C) of the van Krevelen diagram with
224 increasing maturity (van Krevelen, 1961). The CP-derived atomic ratios of the type II Eagle Ford
225 kerogens exhibit modest errors, and some were identified as predominantly algal (type I), while
226 others were identified as planktonic (type II) in origin. There were subtle differences between the
227 atomic ratios estimated from CPMAS and DPMAS NMR experiments (Table S7). This can be
228 attributed to a slight underestimation of non-protonated and oxygenated aromatic C in the
229 CPMAS spectra, causing a proportional overestimation of the H/C and underestimation of the
230 O/C ratios. This is evidenced by the placement of coals within the type II region of Figure 2C.
231

232 The thermal evolution of the various kerogen types, and their respective catagenetic products are
233 markedly different. Typically, type I kerogens have an abundance of aliphatic C, and a strong
234 oil-prone tendency upon maturation. Type II kerogens are mixed oil/gas-prone, and coals are
235 predominantly gas-prone, owing to more diverse (O-containing) molecular compositions (Tissot
236 and Welte, 1984). These generalities are inherently useful in rapid appraisals of source facies on
237 regional scales. However, understanding structure-catagenesis relationships for predicting
238 hydrocarbon yields requires statistical models that describe the relationships between pyrolysis
239 assay observations and kerogen molecular structure (Cao et al., 2013). Prior correlations between
240 structural parameters as determined by CPMAS ^{13}C NMR and hydrocarbon yields, such as those
241 established in Qin et al. (1991), could suffer from inaccuracy introduced by the broad signal

from aliphatic C, characteristic of aquatic kerogens (Cao et al., 2013). Therefore, more detailed characterization, such as the spectroscopic determination of protonated vs. nonprotonated structures, may be a promising means by which to forge more quantitative structure-catagenesis relationships.

4.4 – Contribution of functional groups to genetic potential

Simple linear regression models correlating hydrocarbon potential (S2) with the absolute abundances of oil-prone (f_o , 25 – 45 ppm), gas-prone (f_g , 0 – 25 and 45 – 90 ppm), mobile methyl (f_{al}^* , 0 – 20 ppm), and oxygen-substituted aliphatic (f_{al}^O , 60 – 90 ppm) carbons show strong coefficients of determination for type I kerogen ($r^2 > 0.90$). However, these same correlations are not as strong for kerogen types II and coal ($r^2 > 0.68$), and the slopes of the regression lines differ substantially for types I, II, and coal (Table S1, Figures S3 – S5). This indicates that during the Rock-Eval simulation of the catagenesis process, the pyrolytic formation of hydrocarbons (S2) is not strictly determined by the relative abundance of a given carbon functional group or spectral region (e.g. methylenic CH₂, 25 – 45 ppm). Furthermore, the difference in the slopes of the f_o vs S2 and f_g vs S2 trendlines across the kerogen types are an indication that structural controls on catagenesis operate at the molecular scale, rather than the functional group scale. For example, methylene CH₂ carbons present in an algenan-derived structure may have a different hydrocarbon generation potential than methylene CH₂ carbons in a lignin-derived structure. To overcome this shortcoming of simple linear regression models, we applied a multiple-regression approach that included all of the carbon functional groups and NMR spectral regions listed in Table 1. Note that f_a^S and f_a^B may not be fully resolved in certain chemical structures (Table 1, footnote).

265

266 Multiple regression models provide fundamental insight to the process of catagenesis because
267 they identify the relative contributions (positive and negative) of multiple structures in kerogen
268 to the total hydrocarbon yield (S2). Two multiple regression models were developed. The first
269 model is based upon NMR spectral regions (Table 1 (f_o , f_g , f_a)), shown by Qin et al. (1991) to be
270 correlated with oil and gas potentials. The second regression model was based upon interpreted
271 structural parameters of kerogen (Table 1 (f_a^C , f_a^P , f_a^S , f_a^B , f_{al} , and f_{al}^O)) shown by Solum et al.
272 (1989) to predict the maturity (rank) of coals. The first multiple regression model (Equation (1))
273 confirms that oil-prone (f_o) and gas-prone (f_g) carbons contribute substantially to S2, and
274 inert/aromatic carbons show a negative relationship with S2 (Figure 3A; Table S2). The
275 coefficient of determination using these predictor variables was 0.83 and the RMS error of
276 prediction was ± 11.7 mg hydrocarbon per gram of sediment. The p-values for each of the three
277 parameters (<0.05), indicate that f_o , f_g , and f_a are significantly correlated with S2. The model
278 coefficients indicate that f_g has the strongest positive effect on S2, where a one-unit increase in
279 S2 corresponds to a 0.69-unit increase in f_g .

280

281 Hydrocarbon potential (S2) = $10.18 + (f_o * 0.50) + (f_g * 0.69) + (f_a * -0.20)$ (1)

282

283 The second multiple regression model (Equation (2)) had a higher coefficient of determination
284 (0.93) and lower RMS error (± 7.78 mg hydrocarbon / g sediment), indicating that including
285 structural properties determined by dipolar-dephasing NMR can more accurately predict S2
286 (Figure 3B; Table S3). The p-values for all parameters are <0.05 , apart from alkyl-substituted
287 aromatic C (f_a^S), which was included in the model because it did not detract from model

accuracy. We also found that separation of mobile methyl (f_{al}^*) groups from other methylenic groups and quaternary alkyl C did not improve accuracy of model predictions.

$$\text{Hydrocarbon potential (S2)} = 3.47 + (f_a^C * 4.21) + (f_a^P * -4.11) + (f_a^S * -0.08) + (f_a^B * 0.63) + (f_{al} * 0.43) + (f_{al}^O * 2.74) \quad (2)$$

Equation (2) indicates that carbonyl C (f_a^C), oxygen-substituted alkyl C (f_{al}^O) are the strongest contributors (on a mass-normalized basis) to the hydrocarbon potential of all three kerogen types. The smaller magnitude of the regression coefficients for structural parameters f_a^S and f_a^B is likely due to the relative thermal stability of aromatic structures. The largest overall contribution to S2 comes from the alkyl C term ($f_{al} * 0.43$). The strong negative contribution of oxygen-substituted C (f_a^P) suggests phenolics and aromatic ethers are highly antagonistic to hydrocarbon formation. Overall, Equations (1) and (2) agree to the extent that terms containing alkyl carbons (f_o , f_g , f_{al} , and f_{al}^O) make the largest contributions to total hydrocarbon yield (S2), while contributions from aromatic C are minor or negative.

4.5 – Structure-catagenesis relationships

The hydrocarbon yield (S2) from temperature-programed Rock-Eval pyrolysis is a widely-used assay of source rock catagenetic potential. Simple linear regression of carbon functional group abundances and hydrocarbon yields (Table S1) showed that kerogen type was modulating the genetic potential of a given carbon functional group, evident in the kerogen type-dependent slopes (Figures S3, S4, S5). Since kerogen type determines molecular structure (Figure 2C), we interpret this as evidence that molecular structure acts as a control on catagenesis, beyond the

commonly-identified controls of C abundance (TOC), elemental ratios (H/C, O/C, N/C), and functional group relative abundance (CH₂, CH₃, COOH, etc). The multiple regression analysis of kerogen functional groups found empirical evidence for the existence of structure-catagenesis relationships that act universally across kerogen types. Interestingly, the multiple regression procedure that produced Equation (2) identified functional groups in kerogen that have the following roles in hydrocarbon generation: [1] productive, to varying degrees, [2] nonproductive (inert), and [3] counterproductive (sinks). The ensuing discussion speculates upon the potential mechanisms underlying these three structure-catagenesis relationships.

[1] Among the hydrocarbon producing functional groups, alkyl C made the largest contribution to hydrocarbon yield (S₂), due to the high abundance of alkyl C in immature kerogens. In Equation (2) this is represented by the term ($0.43 * f_{al}$). However, on a mass-normalized basis, hydrocarbon contributions followed the order: Carbonyl C > oxygen-substituted alkyl C > aromatic bridgehead C > alkyl C. The ordering of mass-normalized contributions of these functional groups to S₂ can potentially be explained by a number of factors, most notably the enthalpy of reaction for the hydrogen (H₂) addition across O-bonds (-C-O, O-H), and the thermodynamic stability of O-containing leaving groups (e.g. formaldehyde, formic acid, water). Qin et al. (1991) showed that the oxygen-substituted alkyl groups are particularly gas-prone, based upon the Fisher assay. The large relative contribution of carbonyl and oxygen-substituted alkyl C to S₂ (oil and gas) is also parsimonious with the higher accessibility or lower steric hindrance of terminal acids and/or alcohols to hydrogenation reactions. Ester linkages may have a similar importance in kerogen types I & II. General trends during the onset of catagenesis mirror this finding. Carbonyl and carboxyl groups are progressively eliminated before ester or

ether linkages, and heterocycles (Tissot and Welte, 1984). The bridgehead C contribution to hydrocarbon (S2) may be related to the bicyclic aromatic structures (e.g. alkyl naphthalenes) that are common constituents of petroleum, as opposed to more inert polyaromatic structures.

[2] The regression coefficient (-0.08) of alkyl-substituted aromatic structures (f_a^S) is not significantly different from zero, indicating that cross-linked aromatic structures had no appreciable contribution to hydrocarbon genesis. We interpreted this as a relatively inert carbon pool, which is consistent with the high bond enthalpy of aromatic C–C=C bonds, due to resonance stabilization.

[3] The existence of counter-catagenic oxygen functionality has been previously noted in studies of coal pyrolysis (Boudou et al., 1994). However, it is unclear whether the oxygen-substituted aromatic structures (f_a^P) are acting as hydrogen sinks (preventing hydrocracking) or carbon sinks (promoting condensation reactions, semi-coke formation), or both. Phenolic and aromatic ether C could counteract hydrocarbon formation/expulsion, because they serve as hydrogen (H_2) scavengers, thereby inhibiting hydrogenation reactions, or because they act as sites for hydrocarbon condensation. Mechanisms of kerogen formation have previously invoked phenolic C as a possible site for oxidative condensation during early diagenesis (see review by Vandenbroucke and Largeau, 2007), and it is possible that phenolic C behaves similarly upon maturation. Evidence supporting this observation can be found in the biomass pyrolysis literature, wherein bio-oil yields decline and char yields increase with increasing lignin phenol abundance (e.g. Zhong and Wei, 2004; Serapiglia et al., 2015).

Conclusions

This paper presents methods ^{13}C NMR spectroscopy-based determination of kerogen source type (terrestrial, marine, lacustrine), and hydrocarbon potential of thermally immature sediments. Chemical structure information from ^{13}C NMR with ^1H spectral editing provides evidence for the existence of a universal structure-catagenesis relationship. We have proposed a structure-based empirical model for kerogen type-independent predictions of catagenic hydrocarbon potential.

Acknowledgements

We thank E. Suuberg, the Fuel editorial board, and two anonymous reviewers for their constructive comments that greatly improved this manuscript. We thank the departments of Geosciences and Chemistry at Baylor University, R. Zhang for help with sample analysis, GeoMark LTD, Galbraith laboratories, the Gulf Coast Association of Geological Societies (GCAGS) student research grant program, the C. Gus Glasscock Jr. Fund for Excellence in Environmental Science, and NSF EAR-IF 1132124.

Table 1. Structural parameters estimated using ^{13}C NMR experiments.

Chemical Shift (ppm)	Interpretation	Symbol	NMR	ref
<i>functional group and elemental analyses</i>				
0 – 45	alkyl	f _{al}	a	1
0 – 20	mobile methyl	f _{al} [*]	b	2
20 – 45	mobile and quaternary alkyl	-	b	2
45 – 60	methoxyl and amine	-	a	1
45 – 60	methoxyl	-	b	2
60 – 90	O-substituted alkyl	f _{al} ^O	a	1
90 – 110	O ₂ -substituted alkyl	-	a	1
110 – 145	total aromatic	-	a	1
110 – 135	non-protonated aromatic bridgehead [†]	f _a ^B	b	3
110 – 135	protonated aromatic	-	b	3
135 – 145	alkyl-substituted aromatic [†]	f _a ^S	b	3
145 – 165	O-substituted aromatic	f _a ^P	a	1
165 – 190	carboxyl and amide	f _a ^C	a	1
190 - 220	aldehyde and ketone		b	3
190 - 220	ketone		b	3
<i>hydrocarbon potential</i>				
0 – 25, 45 – 90	Gas prone fraction	f _g	a	4
25 – 45	Oil prone fraction	f _o	a	4
90 – 165	Inert fraction	f _a	a	4
NMR method a = cross polarization or direct polarization, b= dipolar-dephasing; reference 1(Baldock et al., 2004); 2(Smernik et al., 2001); 3(Solum et al., 1989); 4(Qin et al., 1991).				
† note that some alkyl-substituted aromatic C in ortho-position to O-substitutions may overlap some bridgehead C signals.				

Table 2. Comparison of whole rock samples on the basis of organic carbon and Rock-Eval pyrolysis

Formation	Kerogen Type	<u>Source potential</u>		<u>Thermal maturity</u>		<u>Oxygen index</u>	<u>Inert kerogen</u>
		TOC (%)	S2 (mg HC/g)	oil content (S1*100/TOC)	T_{max} (Celcius)	S3 (mg CO ₂ /g)	S4 (mg HC/g)
Yixian	I	0.7 – 4.9	0.5 - 32	12 – 25	432 - 437	0.1 – 1.2	7 - 28
Eagle Ford	II	0.9 – 9.5	0.9 - 61	1.6 - 61	408 - 430	0.1 – 3.3	9 - 68
Joggins	Coal	26 – 62	32 - 131	1.6 – 3.8	430 - 437	1.4 – 3.5	284 - 609

*range of measured values

371 **Figure 1.** Representative ^{13}C CPMAS NMR spectra of kerogen types I, II, and coal from the
 372 Yixian formation, Eagle Ford Formation, and Joggins formation, respectively. The suite of type
 373 I kerogens are characterized by high aliphatic to aromatic ratio, owing to a lipid-rich algal
 374 biomass source. Type II kerogens are less aliphatic, with marginally higher aromatic C content,
 375 with an increase in o-alkyl and ketone C. Coals predominantly aromatic, resulting from a
 376 terrestrial plant source and higher thermal maturity.

377 **Figure 2.** Comparison of kerogen typing by conventional (A, B) and NMR-based (C) methods:
 378 A) H/C and O/C atomic ratios determined by elemental analysis (van Krevelen, 1950) and B)
 379 Hydrogen and Oxygen indices from Rock-Eval 2, and C) Direct assignment (DA) approach
 380 (Hockaday et al., 2009). The Hydrogen Index, $HI = (100 \times S2 / TOC)$ and Oxygen Index, $OI =$
 381 $(100 \times S3 / TOC)$.

382 **Figure 3.** Multiple regression models for S2 value prediction using structural parameters derived
 383 from 1-dimensional ^{13}C NMR experiments: A) oil-prone (f_o), gas-prone (f_g), and inert C (f_a); and
 384 derived from, dipolar-dephasing NMR experiments: B) carbonyl C (f_a^C), phenolic or aromatic
 385 ether C (f_a^O), alkyl-substituted aromatic C (f_{al}^S), aromatic bridgehead C (f_a^B), alkyl (f_{al}), and
 386 oxygen substituted aliphatic C (f_{al}^O).

387 **References**

- 388 Bajc, S., Amblès, A., Largeau, C., Derenne, S. and Vitorović, D. (2001) Precursor biostructures
 389 in kerogen matrix revealed by oxidative degradation: oxidation of kerogen from Estonian
 390 kukersite. *Org. Geochem.* **32**, 773-784.
- 391 Barwise, A.J.G., Mann, A.L., Eglinton, G., Gowar, A.P., Wardroper, A.M.K., Gutteridge, C.S.,
 392 (1984) Kerogen characterisation by ^{13}C NMR spectroscopy and pyrolysis-mass
 393 spectrometry. *Org. Geochem.* **6**, 343-349.
- 394 Bennett, B., Lager, A., Russell, C.A., Love, G.D. and Larter, S.R. (2004) Hydropyrolysis of
 395 algae, bacteria, archaea and lake sediments; insights into the origin of nitrogen
 396 compounds in petroleum. *Org. Geochem.* **35**, 1427-1439.
- 397 Bharati, S., Patience, R.L., Larter, S.R., Standen, G. and Poplett, I.J.F. (1995) Elucidation of the
 398 Alum Shale kerogen structure using a multi-disciplinary approach. *Org. Geochem.* **23**,
 399 1043-1058.
- 400 Boling, K.S. and Dworkin, S.I. (2015) Origin of organic matter in the Eagle Ford Formation.
 401 *Interpretation* **3**, SH27-SH39.
- 402 Boucher, R.J., Standen, G., Patience, R.L. and Eglinton, G. (1990) Molecular characterisation of
 403 kerogen from the Kimmeridge clay formation by mild selective chemical degradation and
 404 solid state ^{13}C -NMR. *Org. Geochem.* **16**, 951-958.
- 405 Boudou, J.P., Espitalie, J., Bimer, J. and Salbut, P.D. (1994) Oxygen groups and oil suppression
 406 during coal pyrolysis. *Energy & Fuels* **8**, 972-977.
- 407 Cao, X.Y., Yang, J. and Mao, J.D. (2013) Characterization of kerogen using solid-state nuclear
 408 magnetic resonance spectroscopy: A review. *Int. J. Coal Geol.* **108**, 83-90.
- 409 Conte, P., Piccolo, A., van Lagen, B., Buurman, P. and Hemminga, M.A. (2002) Elemental
 410 quantitation of natural organic matter by CP/MAS ^{13}C NMR spectroscopy. *Solid State*
 411 *Nucl. Magn. Reson.* **21**, 158-170.
- 412 Dennis, L.W., Maciel, G.E., Hatcher, P.G. and Simoneit, B.R.T. (1982) ^{13}C Nuclear magnetic
 413 resonance studies of kerogen from Cretaceous black shales thermally altered by basaltic
 414 intrusions and laboratory simulations. *Geochim. Cosmochim. Acta* **46**, 901-907.
- 415 Eglinton, T.I., Larter, S.R. and Boon, J.J. (1991) Characterisation of kerogens, coals and
 416 asphaltenes by quantitative pyrolysis—mass spectrometry. *Journal of Analytical and*
 417 *Applied Pyrolysis* **20**, 25-45.
- 418 Espitalie, J. (1982) Institut Francais du Petrole, Syntheses Geologiques et Geochimie 7020 dated
 419 April 28, 1982. doi.
- 420 Espitalie, J., Deroo, G. and Marquis, F. (1985) Rock-Eval pyrolysis and its applications. *Rev.*
 421 *Inst. Fr. Pet.* **40**, 563-579.
- 422 Gélinas, Y., Baldock, J.A. and Hedges, J.I. (2001) Demineralization of marine and freshwater
 423 sediments for CP/MAS ^{13}C NMR analysis. *Org. Geochem.* **32**, 677-693.
- 424 Harbison, G.S., Raleigh, D.P., Herzfeld, J. and Griffin, R.G. (1985) High-field 2D exchange
 425 spectroscopy in rotating solids. *Journal of Magnetic Resonance* **64**, 284-295.
- 426 Hockaday, W.C., Masiello, C.A., Randerson, J.T., Smernik, R.J., Baldock, J.A., Chadwick, O.A.
 427 and Harden, J.W. (2009) Measurement of soil carbon oxidation state and oxidative ratio
 428 by ^{13}C nuclear magnetic resonance. *Journal of Geophysical Research G: Biogeosciences*
 429 doi:10.1029/2008JG000803.
- 430 Horsfield, B. (1989) Practical criteria for classifying kerogens: Some observations from
 431 pyrolysis-gas chromatography. *Geochim. Cosmochim. Acta* **53**, 891-901.

- Jarvie, D.M. (1991) Total organic carbon (TOC) analysis, in: Merrill, R.K. (Ed.), Treatise of petroleum geology: Handbook of petroleum geology, source and migration processes and evaluation techniques. AAPG.
- Lille, Ü., Heinmaa, I. and Pehk, T. (2003) Molecular model of Estonian kukersite kerogen evaluated by ^{13}C MAS NMR spectra☆. *Fuel* **82**, 799-804.
- Love, G.D., Snape, C.E., Carr, A.D. and Houghton, R.C. (1995) Release of covalently-bound alkane biomarkers in high yields from kerogen via catalytic hydropyrolysis. *Org. Geochem.* **23**, 981-986.
- Maciel, G.E., Bartuska, V.J. and Miknis, F.P. (1978) Correlation between oil yields of oil shales and ^{13}C nuclear magnetic resonance spectra. *Fuel* **57**, 505-506.
- Maciel, G.E., Bartuska, V.J. and Miknis, F.P. (1979) Improvement in correlation between oil yields of oil shales and ^{13}C n.m.r. spectra. *Fuel* **58**, 155-156.
- Mann, A.I., Patience, R.I. and Poplett, I.J.F. (1991) Determination of molecular structure of kerogens using ^{13}C NMR spectroscopy: I. The effects of variation in kerogen type. *Geochim. Cosmochim. Acta* **55**, 2259-2268.
- Metz, G., Wu, X.L. and Smith, S.O. (1994) Ramped-amplitude cross polarization in magic-angle-spinning NMR. *Journal of Magnetic Resonance, Series A* **110**, 219-227.
- Miknis, F.P., Netzel, D.A., Smith, J.W., Mast, M.A. and Maciel, G.E. (1982) ^{13}C NMR measurements of the genetic potentials of oil shales. *Geochim. Cosmochim. Acta* **46**, 977-984.
- Patience, R.L., Mann, A.L. and Poplett, I.J.F. (1992) Determination of molecular structure of kerogens using ^{13}C NMR spectroscopy: II. The effects of thermal maturation on kerogens from marine sediments. *Geochim. Cosmochim. Acta* **56**, 2725-2742.
- Philp, P.R. and Gilbert, T.D. (1987) A review of biomarkers in kerogens as determined by pyrolysis-gas chromatography and pyrolysis-gas chromatography-mass spectrometry. *Journal of Analytical and Applied Pyrolysis* **11**, 93-108.
- Qin, K., Chen, D. and Li, Z. (1991) A new method to estimate the oil and gas potentials of coals and kerogens by solid state ^{13}C NMR spectroscopy. *Org. Geochem.* **17**, 865-872.
- Resing, H.A., Garroway, A.N. and Hazlett, R.N. (1978) Determination of aromatic hydrocarbon fraction in oil shale by ^{13}C n.m.r. with magic-angle spinning. *Fuel* **57**, 450-454.
- Rullkötter, J. and Michaelis, W. (1990) The structure of kerogen and related materials. A review of recent progress and future trends. *Org. Geochem.* **16**, 829-852.
- Russell, C.A., Snape, C.E., Meredith, W., Love, G.D., Clarke, E. and Moffatt, B. (2004) The potential of bound biomarker profiles released via catalytic hydropyrolysis to reconstruct basin charging history for oils. *Org. Geochem.* **35**, 1441-1459.
- Salmon, E., Behar, F. and Hatcher, P.G. (2011) Molecular characterization of Type I kerogen from the Green River Formation using advanced NMR techniques in combination with electrospray ionization/ultrahigh resolution mass spectrometry. *Org. Geochem.* **42**, 301-315.
- Serapiglia, M.J., Mullen, C.A., Smart, L.B. and Boateng, A.A. (2015) Variability in pyrolysis product yield from novel shrub willow genotypes. *Biomass and Bioenergy* **72**, 74-84.
- Smernik, R.J. and Oades, J.M. (2000a) The use of spin counting for determining quantitation in solid state ^{13}C NMR spectra of natural organic matter: 1. Model systems and the effects of paramagnetic impurities. *Geoderma* **96**, 101-129.

- Smernik, R.J. and Oades, J.M. (2000b) The use of spin counting for determining quantitation in solid state ^{13}C NMR spectra of natural organic matter: 2. HF-treated soil fractions. *Geoderma* **96**, 159-171.
- Smernik, R.J. and Oades, J.M. (2001) Solid-state ^{13}C -NMR dipolar dephasing experiments for quantifying protonated and non-protonated carbon in soil organic matter and model systems. *European Journal of Soil Science* **52**, 103-120.
- Smernik, R.J., Schwark, L. and Schmidt, M.W.I. (2006) Assessing the quantitative reliability of solid-state C-13 NMR spectra of kerogens across a gradient of thermal maturity. *Solid State Nucl. Magn. Reson.* **29**, 312-321.
- Snape, C.E., Bolton, C., Dosch, R.G. and Stephens, H.P. (1989) High liquid yields from bituminous coal via hydropyrolysis with dispersed catalysts. *Energy & Fuels* **3**, 421-425.
- Solum, M.S., Pugmire, R.J. and Grant, D.M. (1989) Carbon-13 solid-state NMR of Argonne-premium coals. *Energy & Fuels* **3**, 187-193.
- Takeda, N. and Asakawa, T. (1988) Study of petroleum generation by pyrolysis—I. Pyrolysis experiments by Rock-Eval and assumption of molecular structural change of kerogen using ^{13}C -NMR. *Applied Geochemistry* **3**, 441-453.
- Tissot, B.P. and Welte, D.H. (1984) Petroleum Formation and Occurrence. Springer-Verlag.
- van Krevelen, D.W. (1950) Graphical-statistical method for the study of structure and reaction processes of coal. *Fuel* **29**, 269-284.
- van Krevelen, D.W. (1961) Coal: typology-chemistry-physics-constitution. Elsevier.
- Vandenbroucke, M. and Largeau, C. (2007) Kerogen origin, evolution and structure. *Org. Geochem.* **38**, 719-833.
- Vitorović, D., Vučelić, D., Gašić, M.J., Juranić, N. and Macura, S. (1978) Analysis of the organic matter of oil shales by nuclear magnetic resonance. *Org. Geochem.* **1**, 89-91.
- Vučelić, D., Juranić, N. and Vitorović, D. (1979) Potential of proton-enhanced ^{13}C n.m.r. for the classification of kerogens. *Fuel* **58**, 759-764.
- Werner-Zwanziger, U., Lis, G., Mastalerz, M. and Schimmelmann, A. (2005) Thermal maturity of type II kerogen from the New Albany Shale assessed by C-13 CP/MAS NMR. *Solid State Nucl. Magn. Reson.* **27**, 140-148.
- Witte, E.G., Schenk, H.J., Muller, P.J. and Schwochau, K. (1988) Structural modifications of kerogen during natural evolution as derived from ^{13}C CP MAS NMR, IR spectroscopy and Rock-Eval pyrolysis of Toarcian shales. *Org. Geochem.* **13**, 1039-1044.
- Zhong, C. and Wei, X. (2004) A comparative experimental study on the liquefaction of wood. *Energy* **29**, 1731-1741.