The isotopic structures of geological organic compounds

JOHN M. EILER[1]*, MATTHIEU CLOG[2], MICHAEL LAWSON[3], MAX LLOYD[1], ALISON PIASECKI[4], CAMILO PONTON[1] & HAO XIE[1]

[1]*Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125*

[2]*University of Glasgow, Glasgow G12 8QQ, Scotland, UK*

[3]*ExxonMobil Upstream Research Company, Spring, TX 77389*

[4]*Department of Earth Science, University of Bergen, 5020 Bergen, Norway*

**Correspondence: eiler@gps.caltech.edu*

Abstract: Organic compounds are ubiquitous in the Earth's surface, sediments and many rocks, and preserve records of geological, geochemical and biological history; they are also critical natural resources and major environmental pollutants. The naturally occurring stable isotopes of volatile elements (D, ^{13}C, ^{15}N, $^{17,18}O$, $^{33,34,36}S$) provide one way of studying the origin, evolution and migration of geological organic compounds. The study of bulk stable isotope compositions (i.e. averaged across all possible molecular isotopic forms) is well established and widely practised, but frequently results in non-unique interpretations. Increasingly, researchers are reading the organic isotopic record with greater depth and specificity by characterizing stable isotope 'structures' – the proportions of site-specific and multiply substituted isotopologues that contribute to the total rare-isotope inventory of each compound. Most of the technologies for measuring stable isotope structures of organic molecules have been only recently developed and to date have been applied only in an exploratory way. Nevertheless, recent advances have demonstrated that molecular isotopic structures provide distinctive records of biosynthetic origins, conditions and mechanisms of chemical transformation during burial, and forensic fingerprints of exceptional specificity. This paper provides a review of this young field, which is organized to follow the evolution of molecular isotopic structure from biosynthesis, through diagenesis, catagenesis and metamorphism.

The stable isotope compositions of natural organic compounds can potentially be used toaddress a variety of significant questions, such as identifying the sources of hydrocarbon pollutants in near-surface waters or air (Elsner *et al.* 2012), characterizing the precursor biomolecules parental to geological hydrocarbons (Ferreira *et al.* 2012), recognizing and quantifying conditions and mechanisms of hydrocarbon generation (Whiticar *et al.* 1986; Chung *et al.* 1988; Tang *et al.* 2000), or of physical or biological destruction (Martini *et al.* 2003). There is an extensive scientific literature addressing such questions using measurements of the isotopic compositions – generally D/H, $^{13}C/^{12}C$, $^{15}N/^{14}N$ or $^{34}S/^{32}S$ ratios – of bulk oil or gas fractions or, more usefully, individual molecular constituents (i.e. compound-specific analysis). Virtually all such studies have focused on the average isotopic contents of these materials, summed across all of the many isotopologues of the compounds of interest. For instance, the $^{13}C/^{12}C$ ratio of an organic compound, averaged across isotopologues that may differ in their site of ^{13}C substitution and/or numbers of ^{13}C substitutions. Such measurements fail to observe the information potentially recorded by differences in proportions of the many site-specific and multiply substituted species. Note that small gaseous molecules (e.g. CO_2, CH_4, N_2O) commonly have around ten isotopologues, small metabolites or hydrocarbons (e.g. alanine or *iso*-octane) typically have thousands of isotopologues, and larger biomolecules (sugars, fatty acids) have millions, to millions of millions (e.g. hopanoids, hormones, proteins and other macromolecular materials). Nearly all such isotopologues can be thought of as unique chemical species, differing in their physical, thermodynamic and chemical-kinetic properties.

In this paper, we review the small but rapidly growing body of work that examines the stable isotope distributions of hydrocarbons at intramolecular scales – i.e. 'site-specific' differences in isotopic composition between structurally non-equivalent sites in the same molecule, and 'clumped-isotope' compositions, or probabilities that two or more rare isotopes are present in the same molecule. We refer to this family of isotopic properties collectively as

From: Lawson, M., Formolo, M. J. & Eiler, J. M. (eds) 2018. *From Source to Seep: Geochemical Applications in Hydrocarbon Systems*. Geological Society, London, Special Publications, **468**, 53–81.
First published online December 14, 2017, https://doi.org/10.1144/SP468.4

molecular 'isotopic structure'. There have been few papers that include measurements of the isotopic structures of natural hydrocarbons (and even fewer that include a significant number of samples with known origins). However, there is a larger body of work documenting isotopic structures of biomolecules that are likely precursors to petroleum and natural gas compounds, or examining fundamental fractionations controlling isotopic structures of organics; and, over the last 5 years, considerable effort has been put into advancing the technologies that permit such analyses (Appendix B). Taken together, this work suggests we are on the cusp of a new, potentially large sub-discipline of organic geochemistry.

This review aims to present a comprehensive, organized 'snap shot' of this field as it presently exists, and to predict some of the ways in which it could evolve over the next several years. We do not provide a detailed discussion of the clumped-isotope geochemistry of methane – arguably the largest and most mature part of our subject – because it is considered in its own review paper in this volume (see Stolper *et al.* 2017). The primary focus of this paper is on natural hydrocarbons containing two or more carbon atoms, or organic compounds that may be precursors to such natural hydrocarbons. Somewhat unusual for such a review, we have made an effort to include work that, so far, has been presented only in theses, conference proceedings and public reports to funding agencies, because significant contributions to analytical technologies and scientific findings have not yet appeared in the peer-reviewed literature. We doubtless have missed some such sources, and some of this material is likely to evolve by the time it reaches the peer-reviewed literature.

An organizing principle: the genealogy of molecular isotopic structures from biomass to petroleum and gas compounds

The body of work discussed in this review spans a great range in techniques, chemical compounds and motivating questions – many seemingly far removed from petroleum geochemistry. We attempt to lend this material a rational overarching structure by casting it as a narrative about the molecular isotopic structures of biomolecules and their geological transformation into petroleum and gas compounds. We begin this discussion with an overview of sedimentary organic matter and the changes it undergoes during diagenesis, catagenesis and metamorphism (see Fig. 1, adapted from Tissot & Welte 1984).

Sedimentary organic matter is dominated by four major components: lipids (including fatty acids and hopanoids); proteins and amino acids; carbohydrates

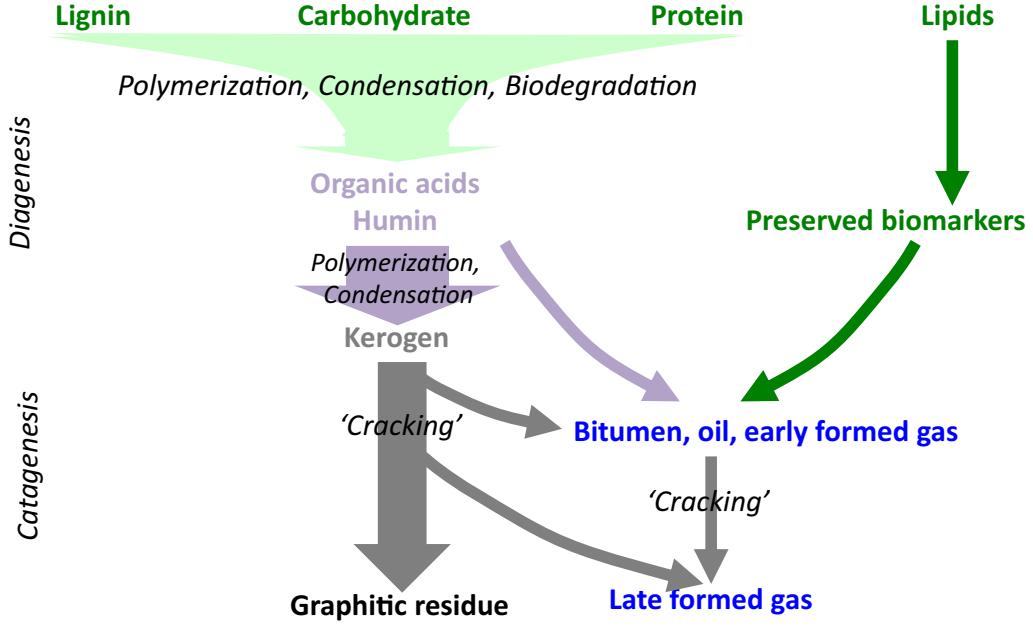

Fig. 1. Schematic representation of the evolution of sedimentary organic matter through formation, diagenesis and catagenesis, increasing in depth of burial and maximum temperature from top to bottom. Adapted from Tissot & Welte (1984).

(sugars, starches and cellulose); and lignin and other polymerized and structurally ill-defined components (Tissot & Welte 1984; Wakeham *et al.* 1997; Hedges *et al.* 2000; Freeman 2001). Cellulose and lignin are proportionately more abundant in terrestrial sedimentary organic matter, whereas proteins and lipids are proportionately more abundant in marine organic matter. The abundance ratios and chemical forms of these constituents change rapidly after formation through utilization by heterotrophs, environmental degradation, and early diagenesis during and just after burial. Soluble carbohydrate and protein decrease radically in proportion to cellulose, lignin and other polymerized components, and the remaining fraction undergoes polymerization and condensation to form humin – a refractory, structurally complex macromolecular matrix – with more reactive moieties transforming to small organic acids. Lipids and closely related compounds (e.g. hydrocarbons formed by decarboxylation of fatty acids) are relatively well preserved through these initial transformations.

Further burial and heating leads to the onset of catagenesis, marked by further polymerization, condensation and other reactions, transforming humin into kerogen and concentrating relatively labile constituents (fatty and other organic acids, alkanes) into the first generation of bitumen, oil and early-formed natural gas (which typically contains sub-equal amounts of methane and higher order volatile alkanes, along with CO_2, H_2, H_2S and other minor constituents; Tissot & Welte 1984; Marshall & Rodgers 2008). Further heating drives more extensive catagenesis, condensing the relatively open structures of immature kerogens to form progressively more graphitic networks, decomposing, or 'cracking' kerogen components to liberate oil and gas fraction compounds, and degrading existing oil and gas compounds to form smaller, simpler and more volatile species (Quigley & Mackenzie 1988; Burnham & Sweeney 1989; Lewan 1993; Mao *et al.* 2010). Overall, the proportion of gas to oil rises, methane makes up a greater fraction of all gas, and proportions of well-preserved biomarkers diminish (Price & Schoell 1995; Prinzhofer & Huc 1995). Where this process proceeds to near the onset of greenschist-facies metamorphism (*c.* 250 + °C), oil and early-formed gas will be nearly quantitatively expelled or thermally destroyed, and kerogen begins conversion to graphite, leaving only graphitic carbon and methane-rich late-formed gas (Wada *et al.* 1994). Two significant open questions about this family of processes are: could oil-fraction compounds persist for long times at relatively high temperatures (Mango 1991), and, when water is present, does kerogen cracking continue to produce gas components at high temperatures (200 + °C; Seewald *et al.* 1998)?

We ask several questions about the evolution in molecular isotopic structure through the preceding chain of events:

(1) What are the isotopic structures of the components of sedimentary organic matter that will ultimately transform to geological hydrocarbons?
(2) Which of these intramolecular isotopic properties (site-specific variations; proportions of clumped-isotope species) are preserved through diagenesis and catagenesis to appear in oil and gas components?
(3) How do the reactions of diagenesis and catagenesis subsample or modify the isotopic structures of original biomolecules?
(4) Can we recognize new compounds formed during thermal maturation of organic matter based on their isotopic structures?
(5) Do those compounds record the conditions (e.g. temperatures and/or pressures) of their formation?
(6) Do exchange reactions proceed rapidly enough in geological environments to fully equilibrate the isotopic structures of biomolecules and/or derived hydrocarbons?
(7) Do the reactions of advanced catagenesis in high-maturity systems leave some imprint on molecular isotopic structures of their residues?
(8) Can we see the effects of microbial degradation or production of organics in the subsurface (i.e. during biodegradation of oil and gas)?

The remainder of this paper proceeds through its subject from the beginning of biosynthesis to the end of catagenesis and onset of metamorphism. In the interests of brevity, much of the relevant technical material – the nomenclature used in the study of molecular isotopic structures, and the technologies for making such measurements – is consigned to the Appendices.

In the beginning: isotopic structures of biomolecules

Nearly all organic components of petroleum and natural gas are derived directly or indirectly from biomolecules in sedimentary organic matter, and so the isotopic structures in the products of biosynthesis must be the first step in our exploration of our subject.

Equilibrium controls?

In past decades, there was a protracted debate as to whether the isotopic contents (and, when discussed, isotopic structures) of biomolecules formed in or near isotopic equilibrium. If so, they should be understood through the thermodynamic isotope effects that control many inorganic heterogeneous

and homogeneous isotope exchange equilibria (Galimov 1973, 1974, 1985). If, instead, their isotopic compositions are controlled by irreversible reactions, then they can only be approached through some understanding of inheritance of isotopic compositions from precursors and the kinetic isotope effects (KIEs) associated with non-reversible chemistry (Ivlev *et al.* 1974; DeNiro & Epstein 1977; O'Leary *et al.* 1981).

This debate was long-since decided in favour of inheritance and kinetics (we will see below several good reasons why; see Hayes (2001) for a broader review of this issue). But there are at least three reasons for us to retain an interest in thermodynamically controlled isotopic distributions in biomolecules:

(1) Some biosynthetic reactions are reversible, at least for some elements, sites and conditions, and so could achieve something close to equilibrium isotope distributions (e.g. Valentine *et al.* 2004; Gilbert *et al.* 2012*a*, *b*). This possibility is re-enforced by the recent discovery that biogenic methane in marine seeps and most subsurface environments forms in clumped-isotope equilibrium at its formation temperature (Stolper *et al.* 2017).

(2) Buried organic matter may be susceptible to reversible isotopic exchange with water and other environmental constituents, meaning even some compounds that form by kinetically controlled processes may later achieve isotopic equilibrium during burial diagenesis (e.g. Schimmelmann *et al.* 2006).

(3) Equilibrium fractionations often provide useful approximations of KIEs because both are influenced by isotopic effects on vibrational energies of molecules (for example, the C isotope fractionation associated with CO_2 reduction on the RuBisCO enzyme is similar in direction and amplitude to the equilibrium fractionation between CO_2 and graphite (Guy *et al.* 1993; Chacko *et al.* 2001)).

Few experimental studies examine equilibrium isotopic fractionations in organic molecules (for the illuminating reason that they are difficult to produce in most compounds; see Wang & Sessions (2009*a*, *b*) for rare examples). But there is a much larger number of theoretical studies that predict site-specific and clumped-isotope structures of organics (Galimov & Shirinskii 1975; Galimov 1985; Rustad 2009; Webb & Miller 2014; Kubicki *et al.* 2016; Piasecki *et al.* 2016*a*; Webb *et al.* 2017). Several trends observed or predicted by these studies are as follows:

(1) The distribution of ^{13}C within and between organic molecules is predicted to follow a pattern where the $\delta^{13}C$ of carbon sites that participate in C—O bonds (carboxyl and carbonyl groups) is greater than carbon that is exclusively bonded to other carbon and hydrogen (e.g. methyl groups), both of which are greater than the $\delta^{13}C$ of carbons bonded to sulphur (e.g. $H_3C—S—\ldots$, such as in methionine). The amplitudes of these effects are predicted to be tens of per mille at earth-surface conditions (Galimov & Shirinskii 1975; Galimov 1985; Rustad 2009). Where multiple C—H bonding environments are present (such as the alkanes), ^{13}C is predicted to be preferentially partitioned into —C— and —CH— v. $—CH_2—$, and all these are enriched compared to $– CH_3$ sites (Webb & Miller 2014; Kubicki *et al.* 2016; Piasecki *et al.* 2016*a*), with amplitudes on the order of 10‰.

(2) The distribution of deuterium in hydrocarbons is generally predicted to follow a pattern where $—CH_2—$ are *c.* 70‰ higher in δD than $—CH_3—$ groups at ambient temperatures (Wang *et al.* 2009*a*, *b*; Webb & Miller 2014; Kubicki *et al.* 2016; Piasecki *et al.* 2016*a*). In branched alkanes, —CH— groups are predicted to be even more strongly enriched in D relative to methyl positions ($\varepsilon^{D}_{CH–CH3}$ *c.* 150‰ at 20°C). The D abundances in H-bearing sites adjacent to C═C double bonds (i.e. in alkenes) are somewhat lower than their structural equivalents in alkanes (by *c.* 20–40‰ at 100°C).

(3) Adjacent multiple ^{13}C substitutions in ethane and propane are expected to have Δ_i values (enrichments relative to a stochastic distribution) of *c.* 0.5‰ at 20°C (Piasecki *et al.* 2016*a*; Webb *et al.* 2017). Two non-adjacent ^{13}C substitutions in hydrocarbons are predicted to have negligible enrichments at equilibrium (Δ_i value $\ll$0.1‰). Immediately adjacent clumping of ^{13}C with D or D with D in hydrocarbons are predicted to have associated Δ_i values of *c.* 5‰ and *c.* 20‰, respectively, at earth-surface temperatures. Predicted enrichments (Δ_i values) of clumped-isotope species that have multiple but non-adjacent $^{13}C + D$ or $D + D$ substitutions (e.g. $H_2D^{12}C—^{13}CH_2—^{12}CH_3$) are about an order of magnitude smaller, placing them near or below the limits of precision for any current analytical method.

For all of the generalizations made above, the magnitudes of these equilibrium fractionations are predicted to increase with decreasing temperature, without 'cross-overs' (changes in sign of fractionation) over the range of temperatures most relevant to natural hydrocarbons (*c.* 0–300°C). Only a small number of these predictions have been directly observed by experiment, and the level of theory used to calculate most of them has uncertainties

that can be as much as *c.* 50%, relative (Kubicki *et al.* 2016). What is called for here is a body of experimental work that is only now becoming technically feasible and has no real precedence in the literature.

The (mostly) kinetically controlled isotopic structures of common biomolecules

Despite the reasons for interest in isotope exchange equilibria between and within organic molecules, the weight of evidence strongly indicates that the isotopic structures of biomolecules dominantly reflect inheritance from their precursors, modified by the KIEs of irreversible biosynthetic reactions. If features of these biosynthetic isotopic structures are inherited in some recognizable form by petroleum or gas compounds, then we should be able to use stable isotope data to identify specific precursors of petroleum compounds and the mechanisms of reactions that converted them to oil and gas components. This vision has been recognized for more than 30 years (e.g. Monson & Hayes 1982*a*). Nevertheless, we face this large, complex subject with only a few constraints and little detailed understanding:

Fatty acids. We understand more about the isotopic structures of fatty acids than any other major class of biomolecules (although the title for best-described group might go instead to the sugars; below). An early indirect insight into this subject came from an experiment conducted by DeNiro & Epstein (1977), who examined the carbon isotope effect associated with decarboxylation of pyruvate to form acetyl coenzyme A – a step common to the biosynthetic pathways of fatty acids (and other metabolites). This study found that the KIE for pyruvate decarboxylation leads to a *c.* 15‰ difference in $\delta^{13}C$ between methyl (higher $\delta^{13}C$) and CO–R (lower $\delta^{13}C$) in product acetyl groups. It is worth noting that this is an important instance where the expectations of equilibrium isotope distributions are clearly not observed (i.e. at equilibrium CO_n groups should be higher in $\delta^{13}C$ than CH_n groups). This finding leads to the expectation that biological fatty acids should be lower in $\delta^{13}C$ than associated carbohydrate (as is commonly observed); and, because most sites in biological fatty acids are derived from head-to-tail attachment of acetyl groups ('elongation'), fatty acids should exhibit a distinctive 'even/odd' carbon isotope structure, where higher $\delta^{13}C$ positions derived from the methyl ends of acetyl groups alternate with lower $\delta^{13}C$ positions from the CO–R ends.

In a series of papers that are noteworthy both for their analytical innovation and depth of mechanistic interpretation, Monson & Hayes (1980, 1982*a*, *b*) examined the carbon isotope structures of fatty acids by chemical attack (decarboxylation and ozonolysis, followed by decarboxylation) to selectively liberate a subset of the carbon sites. They first studied *Escherichia coli* – a bacterium that lacks mitochondria – finding that terminal carboxyl groups of C14–18 fatty acids are generally lower in $\delta^{13}C$ than their bulk parent molecules (by anywhere from *c.* 1 to *c.* 14‰), and that the olefinic carbons in unsaturated fatty acids (i.e. the C=C double bond – the sites oxidized by ozonolysis) are also depleted in ^{13}C by several per mille relative to their bulk parent. (Note this is a second instance where carbon isotope variations within natural biomolecules are opposite those expected for equilibrium isotope distributions). Monson & Hayes (1980, 1982*a*) develop an argument based on our understanding of fatty acid biosynthesis, showing that these findings are consistent with an 'even/odd' carbon isotope ordering, where carbons from the CO–R sites carry a distinctive ^{13}C depletion (as one would expect from the experiment of Deniro & Epstein (1977)). We will see that this pattern could be responsible for one of the more distinctive findings to date regarding carbon isotope structures of alkanes.

However, as self consistent as these findings may be, the picture they paint is incomplete. In a follow-up study, Monson & Hayes (1982*b*) applied their approach to fatty acids from *Saccharomyces cerevisiae* (yeast; a eukaryote that possesses mitochondria), obtaining results suggesting an 'even/odd' isotopic ordering that is more subtle (approximately several per mille in amplitude) and opposite in sign, suggesting assembly from acetyl groups having ^{13}C-depleted methyl groups and ^{13}C-rich CO–R groups. A detailed explanation of this finding is beyond what we can explore here, but may generally be said to result from the citric acid cycle, which produces and consumes acetyl-CoA by means other than pyruvate decarboxylation.

High-sensitivity nuclear magnetic resonance (NMR) techniques have been used to study the hydrogen isotope structures of several natural fatty acids (Billault *et al.* 2001; Duan *et al.* 2002; Markai *et al.* 2002; Lesot *et al.* 2008). This work is complicated by the fact that hydrogen sites in the centres of saturated fatty acids are difficult to distinguish by NMR; nevertheless, through a combination of derivatization and focus on unsaturated and branched fatty acids and novel NMR approaches, a detailed picture of the pattern of natural variations has been reconstructed. Site-specific D/H variations generally have amplitudes of *c.* 300–400‰; one common pattern observed is strong (several hundred ‰) D depletion of hydrogen in methyl groups relative to CH_2 groups – at least broadly consistent in direction with predicted equilibrium controls (Wang & Sessions 2009*a*, *b*). However, several observations argue against such a simple interpretation. The amplitudes of natural variations are several times

greater than expected for equilibrium effects; and, CH_2 groups adjacent to sites of desaturation can also be strongly D depleted. Furthermore, pro-S and pro-R hydrogen sites can differ in D/H ratio. All of these findings suggest the importance of KIEs and inheritance from isotopically distinct pools of H during biosynthesis. The oxygen and clumped-isotope structures of natural fatty acids are unknown.

Amino acids. Some features of the site-specific carbon isotope structures of amino acids were constrained by Abelson & Hoering (1961) – this being the first attempt to systematically examine isotopic structures of an important class of biomolecules. This study compared the bulk $\delta^{13}C$ values of several common essential amino acids with the $\delta^{13}C$ of CO_2 released from those amino acids by decarboxylation, thus constraining the difference in $\delta^{13}C$ between carboxyl and non-carboxyl carbons. (We are aware of current efforts to revisit this subject using NMR and high resolution mass spectrometry, but neither body of work has been published.)

The first-order result of this work is that primary producers (autotrophs) contain amino acids with carboxyl positions that are generally higher in $\delta^{13}C$ than other carbon sites, by roughly 10‰ on average. This is an instance where the predictions of equilibrium stable isotope effects (^{13}C enrichments in the oxidized carbon sites) are met, at least approximately. Rustad (2009) explores this relationship in some detail by comparing theoretical equilibrium predictions to measured natural products for several amino acids, noting that despite differences in detail there is a general similarity in direction and magnitude of difference in $\delta^{13}C$ between carboxyl and non-carboxyl carbon sites. However, Abelson & Hoering (1961) also examined amino acids from two heterotrophs and found they exhibit the opposite pattern – carboxyl carbons lower in $\delta^{13}C$ than non-carboxyl carbons (i.e. similar to the carboxyl/non-carboxyl difference in fatty acids from *E.coli*). To the best of our knowledge, there are no observations of the site-specific H, O, N or S isotope geochemistry or clumped-isotope geochemistry of amino acids.

Carbohydrates. Early studies of carbon isotope structures of biosynthetic fatty acids and amino acids (above) assumed, explicitly or implicitly, that primary photosynthate (3-phosphoroglyceric acid) and derived carbohydrates have relatively simple carbon isotope structures with minimal site-specific variations. However, chemical degradations of sugars followed by Isotope Ratio Mass Spectrometry (IRMS) (Rossmann *et al.* 1991) and recent NMR studies of sugars and starches (Gilbert *et al.* 2012*a*, *b*) show this not to be the case. Site-specific carbon isotope variations in common C6 sugars have amplitudes up to *c.* 25‰, and differ systematically in the pattern of site-specific variation depending on photosynthetic pathway (C3 v. C4 v. crassulacean acid metabolism (CAM) photosynthesis). The patterns of carbon isotope variation in glucosyl and fructosyl groups are complex and diverse, but generally exhibit depletions in $\delta^{13}C$ in the C1 and (particularly) C6 positions relative to C2–5 positions. The largest systematic elaboration on this pattern is that the C2 positions of glucosyl groups are generally lower in $\delta^{13}C$ than C2 positions in fructosyl groups from the same source. These variations can be understood as consequences of primary carbon fixation (suggested to set the $\delta^{13}C$ of C6 carbons in sugars), steps in the pentose phosphate pathway, photorespiration, and reactions involving assembly of sucrose and larger carbohydrates, and breakdown of those compounds for export and utilization.

A long-standing and large body of work examines the D/H ratio of 'non-exchangeable hydrogen' in cellulose – that is, hydrogen that is bound to carbon and does not undergo isotopic exchange with water on laboratory timescales, as opposed to the highly exchangeable hydroxyl hydrogen also present in cellulose (Epstein *et al.* 1976, and many subsequent papers). While such data is not precisely a site-specific isotopic measurement, it does sample a subset of molecular sites having common properties and so has some bearing on our subject. The first-order findings of this work are that these non-exchangeable hydrogens are lower in δD than coexisting environmental waters, typically by *c.* 200‰ but with secondary dependence on aridity and temperature. This phenomenon is well explored as a palaeoclimate archive.

Detailed (i.e. truly site-specific) hydrogen isotope structures of carbohydrates are less well known in natural materials, but have been examined by NMR methods in soluble sugars, starch and cellulose from several sources (Zhang *et al.* 2002; Betson *et al.* 2006; Augusti 2007; Augusti *et al.* 2008). The most general finding is that site-specific D/H variations have high amplitudes (hundreds of per mille); that patterns of D/H variation differ by photosynthetic pathway (i.e. C3 v. C4 v. CAM plants); and that differences in these isotopic structures between sugars and starches from the same plants suggest a strong influence of reactions occurring after initial synthesis of C6 sugars. The deconvolution of primary biosynthetic signals from subsequent fractionation and/or exchange is best understood for wood cellulose, based on experiments involving introducing labelled water to plant tissues followed by NMR analysis of hydrogen isotope structures of extracted carbohydrates (Augusti *et al.* 2008). This work suggests site-specific data could be used to deconvolve climatic records from variable metabolic effects (although this strategy has not been widely employed as yet).

A second significant finding is that the difference in D/H ratio between the two hydrogens attached to the C6 position of glucose (designated $C6H^R$ and $C6H^S$) from C3 plants varies by *c.* 200‰, changing systematically with environmental pCO_2/pO_2 (Ehlers *et al.* 2015). This is believed to result from the fact that varying pCO_2/pO_2 changes the relative rates of photosynthesis and photorespiration, which changes the mechanism of hydrogen addition to the relevant carbon site of 3-phosphoroglycerate – a sugar precursor. It is unknown whether this or other similar site-specific hydrogen isotope variations can survive diagenesis and catagenesis to appear in petroleum or gas compounds. Nevertheless, it provides some sense of the possible amplitudes and environmental controls of the hydrogen isotope structures of biomolecules.

Despite the fact that the bulk oxygen isotope composition of cellulose is well explored as a palaeoclimate archive, we are not aware of any previous observations of the oxygen isotope structures of natural carbohydrates.

Terpenes. Martin *et al.* (2006) present a summary and interpretation of NMR measurements of the hydrogen isotope structures of geraniol and α-pinene of various origins, showing that site-specific D/H variations are very large (a factor of *c.* 4) and differ systematically between C3 and C4 plants. These findings can be interpreted as consequences of inheritance of hydrogens from metabolic precursors (pyruvate, glyceraldehyde 3-phosphate, dimethylallyl diphosphate and isopentenyl diphosphate) accompanied by large KIEs. The resulting hydrogen isotope structures differ from common patterns of equilibrium hydrogen isotope distribution by a factor of two or more and so present another instance of strong violations of equilibrium in molecular-scale isotopic distributions in biomolecules. The terpenoids are abundant in petroleum, and, if their biosynthetic hydrogen isotope structures survive diagenesis and catagenesis, they may present an opportunity to isotopically fingerprint the biosynthetic origins of an oil component. Alternatively, if isotopic exchange and fractionation during burial overprint these high-amplitude biosynthetic signatures, it is imaginable that the extent of preservation v. homogenization of such effects could serve as a proxy for oil maturity.

Lignin. The isotopic structures of biopolymers and other large biomolecules present great challenges to existing technologies for site-specific and clumped-isotope analyses (see Appendix B). Nevertheless, there have been significant efforts to observe at least some features of the isotopic anatomy of lignin, which has been the subject of a measurement technique involving chemical extraction of CH_3 from methoxy groups, followed by hydrogen and/or carbon isotope analysis (Keppler *et al.* 2007; Feakins *et al.* 2013). This technique reveals that methoxy group hydrogen has a D/H ratio that is similar to the non-exchangeable hydrogens in cellulose; i.e. it is *c.* 150–200‰ lower than but correlated with the δD of environmental water. This relationship is of interest primarily because of its potential as a palaeoclimate archive. The carbon isotope compositions of methoxy groups from lignin are less well explored because they do not obviously add any information beyond a bulk analysis of the global $\delta^{13}C$ of woody materials, and there is evidence that they are vulnerable to changes during biological degradation of plant litter (Anhauser *et al.* 2015). We revisit this issue below, in our discussion of the carbon isotope structures of lignites and coals that have been subjected to diagenetic and catagenetic changes.

General observations

We conclude this discussion with a few general observations: The carbon isotope structures of fatty acids are perhaps the firmest ground on which we can stand when trying to predict the isotopic structures of petroleum and gas components, because lipids are relatively abundant and well preserved after early diagenesis, and because they exhibit relatively simple systematic isotopic variations (i.e. the 'even/odd ordering' carbon isotope motif). However, the differences in isotopic structures between fatty acids from bacteria and Eukarya reveal complexity that complicates our task. It will not be possible to develop a predictive framework for the isotopic structures of petroleum compounds derived from fatty acids until we have a more complete understanding of the diversity in isotopic structures of these precursors, and some sense of how that diversity maps onto different forms of sedimentary organic matter. The same could be said for the amino acids, with their systematic differences between autotrophs and heterotrophs, and carbohydrates, with their systematic differences among C3, C4 and CAM plants. Perhaps more problematically, the isotopic structures of primary biological straight-chain hydrocarbons are unknown, despite their abundance and obvious relevance to interpreting structurally similar petroleum compounds. It seems plausible that they are closely related to the fatty acids (i.e. simply through decarboxylation); nevertheless, observations of the isotopic structures of these important compounds should be a high priority. Similarly, little is known about the isotopic structures of hopanoids; these compounds represent a diverse and potentially information-rich target for future studies of isotopic structure. Finally, we note that there are no observations of any kind that constrain

the abundances of any of the multiply substituted forms of any of the abundant classes of biomolecules (excepting biogenic methane; see Douglas *et al.* (2017) and Stolper *et al.* (2017)).

Kerogen

Essentially all of the work summarized in this review concerns the isotopic structures of biomolecules (above) or components of oil and natural gas (below). It is natural to wish to explain the compositions of the latter through derivation from the former, but there is a third major class of natural organic compounds that stand between these two: the humins, lignins and kerogens in sedimentary rocks (Fig. 1). These structurally complex organic polymers are derived from buried organic matter, they contain components that are recognizable biomolecules, and their thermal degradation produces most petroleum and natural gas components. It will probably be impossible to reach a complete, detailed and coherent understanding of the isotopic structures of petroleum and gas components until we come to terms with the isotopic structures of the materials from which many of them are derived. The key question is whether these organic polymers formed during diagenesis and catagenesis mostly preserve the isotopic structures of biomolecules, or are instead dominated by isotopic fractionations during decarboxylation, reduction, deamination, polymerization, condensation (ring closure) and other diagenetic and catagenetic reactions.

There are vanishingly few direct observations of isotopic structures of kerogens and other sedimentary organic polymers that can guide us, but our understanding of kerogen synthesis (Tissot & Welte 1984; Burnham & Sweeney 1989; Hedges *et al.* 2000; Mao *et al.* 2010) leads to several expectations. First, the increasing proportions of straight-chain hydrocarbons, fatty acids, lignins and cellulose (relative to amino acids and soluble sugars) during remineralization and early diagenesis (Wakeham *et al.* 1997; Hedges *et al.* 2000; Freeman 2001; Marshall & Rodgers 2008) should lead us to expect that kerogens have isotopic structures more similar to these preserved compounds. Specifically, it is reasonable to predict the type I kerogens (rich in aliphatic hydrocarbons) should be most similar in isotopic structure to the fatty acids, the type II kerogens (rich in aromatic hydrocarbons) should be more similar to terpinoids and hopanoids (unfortunately uncharacterized for their isotopic structures at present), and type III kerogens (derived from sources rich in woody land plants) should more closely resemble the isotopic structures of lignin and cellulose.

The most obvious caveat to these predictions is that the chemical transformations that occur during formation and thermal maturation of kerogen may be associated with KIEs having significant amplitudes, potentially modifying the isotopic structures in biomolecules into new patterns with which we are currently unfamiliar (Schimmelmann *et al.* 1999, 2001). Furthermore, it is known that at least some of the hydrogen in kerogens is exchangeable with water during burial diagenesis (Schimmelmann *et al.* 2006). This veil of diagenetic and catagenetic processes that separates the biomolecules from the geological kerogens may be the greatest hurdle to really characterizing and understanding the isotopic structures of kerogen-derived petroleum hydrocarbons.

What is known about these processes is that they do not result in dramatic differences in bulk $\delta^{13}C$ between recently deposited sedimentary organic matter and more deeply buried kerogens (Schimmelmann *et al.* 1999, 2001; Nabbefeld *et al.* 2010). Even relatively mature kerogens generally have bulk $\delta^{13}C$ and δD values in the range of bulk organic-rich sediment. Thus, it seems reasonable to expect that fractionations and exchange reactions accompanying diagenesis and catagenesis are likely to be restricted to particular sites that are relatively reactive, labile and/or exchangeable.

We know of only two related observation that speak directly to these issues: a study of the C and H isotope compositions of methoxyl groups in lignin subjected to microbial attack during burial in leaf litter (Anhauser *et al.* 2015), and a more recent study of the same moieties in lignites from the Japan margin, sampled by cruise IODP-337, as described by Lloyd *et al.* (2016) as well as Inagaki *et al.* (2015). Anhauser *et al.* showed that moderate amounts (*c.* 50%) of removal of methoxyl groups from lignin through microbial attack have no effect on δD, but drive $\delta^{13}C$ of the residual methoxyl groups up by *c.* 5‰. Lloyd *et al.* (2016) found that the trace of residual methoxyl groups in ancient, deeply buried lignites (estimated as *c.* 1% of the initial methoxy inventory remains) have δD_{SMOW} values of −220 to −230‰ (similar to modern plant lignins) but extraordinary $\delta^{13}C_{PDB}$ values of +25–+40‰, which are among the highest $\delta^{13}C$ values known for terrestrial geological materials and *c.* 50‰ enriched relative to methoxyl groups in common plant lignins. One explanation for these findings is that the methoxyl groups are degraded during deposition, diagenesis and earliest catagenesis (microbially and/or possibly by other mechanisms), and that this loss is rate limited by cleavage of the C—O bond in the methoxyl group, which is accompanied by a *c.* 10‰ primary KIE for the methoxyl carbon but no or only negligible secondary KIEs for the adjacent methoxyl-group hydrogens. (One prediction of this hypothesis is that the oxygens in the residual methoxy groups should be similarly

enriched by tens of per mille in ^{18}O and ^{17}O, although it is not obvious at present how they might be measured for their isotopic composition.) These findings, sparse as they are, suggest that the site-specific isotope geochemistry of kerogens hold potential as a record of the mechanisms and extent of kerogen maturation.

KIEs associated with 'cracking' reactions

The formation of geological bitumen, oil and hydrocarbon gas is mostly due to thermal degradation of kerogen; formation of light oil and gas components by thermal degradation of oil and bitumen components is widely suspected but more controversial (Tissot & Welte 1984; Lewan 1985, 1993; Quigley & Mackenzie 1988; Mango 1991; Price & Schoell 1995; Seewald *et al.* 1998). The liquid products (bitumens and oils) are dominated by naphthenes and paraffins – that is, aromatic and straight- or branched-chain hydrocarbons, respectively – whereas the gas products are overwhelmingly composed of low-molecular weight (C1–5) alkanes. These reactions generally proceed by cleavage of kerogen components, either at the sites of functional groups or at C—C bonds in chain- or ring-structures, producing diverse petroleum compounds, abundant H_2O and CO_2, and leaving residues that are depleted in H and O and condense, or aromatize, towards a progressively more graphitic structure.

If we are to predict the isotopic structure of a particular petroleum compound that is a product of these processes, the key questions are: what specific components of kerogen are its precursors; what are the isotopic structures of those precursors; what specific reactions occurred (e.g. in a simple case, what bond in a precursor broke to liberate the compound of interest); and what are the KIEs associated with that reaction? This list is straightforward and necessary, but probably cannot be answered rigorously for any petroleum or gas compound. Kerogens are structurally complex, containing many compounds and sites that potentially serve as reactants for any given product (particularly the relatively small, structurally simple hydrocarbons that are most amenable to isotopic studies). The mechanisms and stoichiometries of 'cracking' reactions have been the subject of much prior study, but are generally described only at a schematic level intended to represent some average across a broad family of reactants and products.

Two ways of making progress with this challenging problem have been suggested. The first way is to conduct experimental or theoretical studies that examine KIEs associated with simplified model systems (e.g. production of methane by breaking the CH_3—CH_{2a} bond in an *n*-alkane). These studies have the advantage of constraining potentially generalizable atomistic mechanisms, and the disadvantage that they obviously oversimplify the reality of geological petroleum and gas formation. The second way is to conduct experimental studies that empirically relate the isotopic composition of a petroleum or gas product to the isotopic content of a kerogen reactant. These studies have the advantage of being conceptually straightforward and obviously relevant to at least the kerogens and experimental conditions that were studied; their disadvantage is that they provide few mechanistic insights that might let one generalize their findings to other source materials, conditions and products.

Idealized models

The first widely used mechanistic model for the KIEs associated with oil and gas formation was presented by Chung *et al.* (1988). (Note our discussion is also generally consistent with the elaboration of this model to include Rayleigh effects in Rooney *et al.* (1995).) This model is inspired by the observation that the $\delta^{13}C$ values of individual molecular components of natural gases (methane, ethane, propane, butanes) often display a linear negative correlation with the inverse of their carbon number (i.e. one-quarter for butane, one-third for propane, etc.), with a vertical intercept similar to the $\delta^{13}C$ values of inferred source kerogens for those gases (Fig. 2). One interpretation of such trends is that these gas components are derived from pyrolysis of an isotopically homogeneous pool of kerogen carbon, with a carbon isotope fractionation that affects only one carbon atom in the product molecule. In this case, the methane–kerogen fractionation equals the KIE (usually inferred to be *c.* 20‰), the ethane–kerogen fractionation half this value (i.e. because only one of the two carbons expresses the fundamental KIE), and so forth. Chung *et al.* (1988) conceptualized this model by suggesting the source substrate can be thought of as a straight-chain hydrocarbon, and that the 'cracking' reaction involves severing one C—C bond to release the product alkane (e.g. breaking CH_3—$CH_{2a,}$ where CH_3 is the terminal methyl group and CH_{2a} is the methylene group immediately adjacent to it, yields methane; breaking CH_{2a}—CH_{2b} yields ethane, etc.). It follows that the carbon isotope composition of the carbon adjacent to the broken bond (e.g. one of the two terminal carbons in propane) should have a $\delta^{13}C$ equal to that for kerogen minus the KIE, and that all other carbons in the product molecule should have a $\delta^{13}C$ equal to that for source kerogen.

Although Chung *et al.* (1988) do not address site-specific or clumped-isotope variations, their model can be extrapolated to predict the possible consequences for such indices. Site-specific

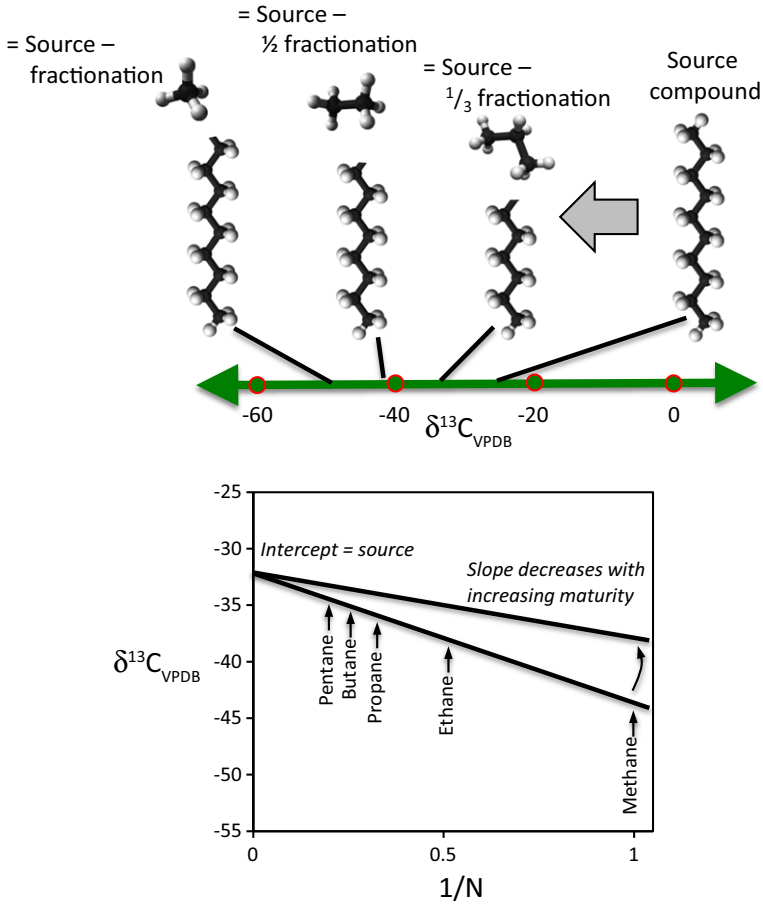

Fig. 2. Conceptual model for the carbon isotope compositions of natural gas components formed by cracking a larger organic molecule (represented here by an *n*-alkane), adapted from Chung *et al.* (1988). Each product molecule is created by cleaving a single C—C bond in the source compound, with a fractionation that may vary with temperature but is assumed to be a constant for all reactions having different product compounds. The smallest product (methane) expresses this full fractionation, and larger products dilute that fractionation by combining the fractionated site with adjacent unfractionated sites. The $\delta^{13}C_{VPDB}$ of the products will exhibit a negative linear correlation with $1/N$, where N is the number of carbon atoms in each product molecule, with a slope that flattens with increasing temperature (and thus diminishing fractionation) and with increasing reaction progress. $\delta^{13}C_{VPDB} = (R^{13}_{sample}/R^{13}_{VPDB} - 1) \times 1000$, where R^{13} is the $^{13}C/^{12}C$ abundance ratio and VPDB stands for the Vienna Pee Dee Belemnite reference standard.

compositions of *n*-alkanes produced according to this model are relatively easy to predict: in all cases, the average $\delta^{13}C$ of the two terminal CH_3 groups should be equal to that of the source kerogen (assumed for the time being to be homogeneous) minus half of the KIE; e.g. if kerogen has a $\delta^{13}C_{PDB}$ of −25‰ and C—C bond cleavage involves a 20‰ normal KIE, then product *n*-butane should have an average $\delta^{13}C_{PDB}$ of −35‰ for the two indistinguishable terminal methyl groups, and of −25‰ for the two indistinguishable interior CH_2 groups. We are not aware of any effort to extend the Chung *et al.* (1988) model to consider hydrogen isotope compositions of natural gas components, or as a means of predicting isotopic compositions of oil compounds.

The clumped-isotope effects of the Chung *et al.* (1988) model are also easy to predict, though somewhat counterintuitive. Consider the case of a precursor straight-chain hydrocarbon having a random distribution of ^{13}C throughout its structure, 'cracking' at one of the CH_{2a}–CH_{2b} bonds to produce ethane, with a KIE of 20‰. Immediately prior to the cracking reaction, the probabilities of ^{13}C being present on each of the two carbon sites that will eventually be transferred to ethane are equal to each other

and to the bulk ^{13}C concentration in the precursor ('$[^{13}C]_i$'). Thus, in this case the probability that both of those two sites contain ^{13}C is $[^{13}C]_i^2$, and the Δ_{13C2H6} value (enrichment of the clumped-isotope species relative to a random distribution) of that population of sites is 0‰. An infinitesimal progress of the cracking reaction will produce a population of product ethane molecules, which will consist of one carbon atom inherited from the terminal CH_3 site of the precursor and one carbon atom inherited from the CH_{2a} site of that precursor. According to the Chung model, the first of these atoms will have a probability of inheriting ^{13}C exactly equal to the bulk $[^{13}C]_i$ of the precursor, but the second will have a probability equal to that for the precursor, $[^{13}C]_i$, multiplied by a number closely related to the fractionation factor for the cleavage reaction (if the $\alpha_{KIE} = 0.98$, then this multiplier is 0.9802; the difference is a simple result of the difference between $^{13}C/^{12}C$ ratios and ^{13}C concentrations). Once ethane forms, these two carbon atoms are symmetrically equivalent, and the concentration of ^{13}C in the ethane pool as a whole, averaged across the molecule, will be $([^{13}C]_i \times (0.5 + [0.9802]/2))$. If that ^{13}C were randomly distributed among all carbon atoms, then the probability of finding an ethane with two ^{13}Cs would just be this number squared: $([^{13}C]_i \times (0.5 + [0.9802]/2)^2$. However, these two carbon atoms come from two different sources and have two different probabilities of containing ^{13}C: one of them is unfractionated and has the composition of the precursor, whereas the other is fractionated and contains less than the substrate's $[^{13}C]_i$ value. As a result, the actual probability of encountering a $^{13}C_2H_6$ molecule is: $([^{13}C]_i^2 \times 0.9802)$, which is lower than suggested by the expression for the random distribution, above. That is, even without considering the distinctive fractionation behaviours of the clumped-isotope species, or any possible non-random isotopic structure of the precursor, the Chung *et al.* (1988) model predicts that cracking reactions should produce ethane with a negative Δ_{13C2H6} value. This is just one example of many possible clumped-isotope effects that could arise from cracking reactions that subsample organic molecules accompanied by KIEs.

Several subsequent models add significant sophistication to the treatment of stable isotope fractionation during natural gas formation but the researchers in question do not clearly comment on the expected molecular-scale isotopic structures of products or residual reactants (e.g. Berner *et al.* 1995). An exception that does add new details regarding isotopic structure is the work of Tang *et al.* (2000, 2005), who present the most sophisticated treatment of KIEs associated with hydrocarbon cracking published to date. Similar to the Chung *et al.* (1988) model, Tang *et al.* (2000) consider a relatively large *n*-alkane ($C_{16}H_{34}$, hexadecane) that undergoes cleavage reactions at or near one of its ends to produce methane (breaking the $CH_3—CH_{2a}$ bond), ethane (breaking the $CH_{2a}—CH_{2b}$ bond) or propane (breaking the $CH_{2b}—CH_{2c}$ bond). However, rather than simply making a first-order assumption about the sites and magnitudes of the associated KIEs, Tang *et al.* (2000) performed a density functional theory (DFT) model of the vibrational energetics of the reactants and products and calculated the relative reaction rates of the singly ^{13}C-substituted isotopologues. Their findings are loosely consistent with the assumptions of Chung *et al.* (1988) (a 'normal' KIE, which favours transfer of ^{12}C to products, and which is strongest at the carbon sites adjacent to the breaking bond, and *c.* 10–20‰ in amplitude). However, this more rigorous treatment also reveals two features that could be common to the KIEs associated with cracking reactions. First, 'secondary isotope effects', or isotope effects associated with ^{13}C substitution at sites not adjacent to the bond being broken (e.g. a difference between the rate of forming propane from $^{12}CH_3—^{12}CH_2—^{12}CH_2\mathbf{—}^{12}CH_2\ldots$ and $^{13}CH_3—^{12}CH_2—^{12}CH_2\mathbf{—}^{12}CH_2$, where the bold line indicates the bond to be broken). These secondary isotope effects have amplitudes *c.* one-quarter to one-half of those of the primary isotope effects (i.e. associated with substitution adjacent to the breaking bond), which are surprisingly large, and significant if one is to understand the isotopic compositions of petroleum compounds at the level of a few per mille or better. Second, the amplitudes of KIEs (both primary and secondary) depend on reaction temperature; in the cases Tang *et al.* (2000) consider, fractionations generally decrease in amplitude with increasing temperature. This second finding raises a first-order problem with the empirical interpretation of gas stable isotope data: Chung *et al.* (1988) emphasize the importance of reaction *progress* in controlling the $\delta^{13}C$ of a given gas component (with products approaching compositions of substrates as reaction progress increases). Tang *et al.* (2000) demonstrate that the same changes can reflect reaction to similar extents but at higher *temperature*.

Tang *et al.* (2005) experimentally examine the isotopic compositions of residues of pyrolysis of oil-fraction compounds (C13–21 *n*-alkanes) and extend the model of Tang *et al.* (2000) to explain their findings. The most important new points made by this second paper are (1) larger alkanes crack more rapidly than smaller ones at a given thermal maturity, meaning the progress of pyrolysis reactions with thermal maturity can vary markedly from small to large compounds; and (2) oil-fraction compounds of intermediate size are both the products of cracking larger compounds (tending to reduce their $\delta^{13}C$ and δD) and residues of their own cracking reactions (tending to raise their $\delta^{13}C$ and δD values). These

arguments could be extended to encompass condensate and gas fraction compounds (though Tang *et al.* (2005) do not present detailed arguments regarding these species).

The treatment by Tang *et al.* (2000, 2005) of alkane-cracking reactions is an advance on that presented by Chung *et al.* (1988), both because of the use of first-principles chemical physics to describe KIEs and because of the consideration of the consequences of simultaneous cracking of a family of related compounds. However, for our purposes these two approaches are at least qualitatively similar in their predictions regarding isotopic structures of hydrocarbons, because they both examine the relatively narrow case of *n*-alkanes that react only be cleaving C—C bonds to produce smaller *n*-alkanes. While this is probably an important process in natural petroleum systems, it provides little insight into the expected consequences of cracking branched alkanes, aromatic compounds, fatty acids and other organic acids, kerogen components, or any of the many other organic components of petroleum and its source rocks. (In the interests of brevity, we do not explore in detail the quantitative differences between the Tang and Chung model approaches in the predicted isotopic structures of oil and gas compounds, but see Piasecki *et al.* (2018) for further discussion.)

Experimental simulations

There have been numerous previous studies of bulk stable isotope fractionations (i.e. averaged across all isotopologues of a given molecule) associated with pyrolysis and hydropyrolysis of kerogen, bitumen and oil (e.g. Lewan 1983; Berner *et al.* 1995; Schimmelmann *et al.* 1999, 2001; Tang *et al.* 2000, 2005). These demonstrate that carbon isotope effects are generally 'normal' in direction (favouring transfer of light isotopes from reactants to products), have amplitudes of *c.* 10–20‰ for $\delta^{13}C$ and *c.* 100‰ for δD, and decrease in amplitude with increasing size of the product molecule, temperature of reaction, and integrated progress of the reaction. All of these trends are broadly consistent with the theoretical expectations reviewed in the previous section.

We are aware of only three previous studies that document the clumped-isotope and position-specific isotope effects associated with experimental pyrolysis of kerogen to produce hydrocarbons larger than methane (see Stolper *et al.* (2017), for a review of experimental constraints on methane clumped-isotope compositions from such experiments). Results from these studies are currently in press, in review or available only in abstract form, but we consider them here in the interests of making this review as current and complete as possible.

Ethane $\Delta^{13}C_2H_6$. Clog *et al.* (2013, 2014) and Clog & Eiler (2014) present analyses of the bulk and clumped-isotope ($\Delta^{13}C_2H_6$) composition of ethane produced by hydropyrolysis of two kerogen-rich sedimentary rocks: Woodford Shale and Albian/Aptian lacustrine shale from the Araripe basin, Brazil (Fig. 3a). The $\Delta^{13}C_2H_6$ index is currently reported relative to an arbitrary reference standard – a commercial ethane gas that has a $\delta^{13}C$ and δD broadly similar to that of common thermogenic gases (below). For this reason, it is not known whether ethane produced by kerogen cracking has a $^{13}C_2H_6$ abundance more or less than expected for a random distribution (i.e. $\Delta^{13}C_2H_6$ above or below zero in an absolute reference frame). Nevertheless, these experimental products are 1–2‰ lower in $\Delta^{13}C_2H_6$ than the reference gas and *c.* 1–3‰ lower than *c.* 75% of the natural ethanes analysed to date. Thus, kerogen cracking produces ethane that is at least relatively low in $\Delta^{13}C_2H_6$.

Increasing the temperature and time of hydropyrolysis of Woodford Shale increases the $\Delta^{13}C_2H_6$ of product ethane subtly but significantly (by a few tenths of per mille) – consistent with our expectations based on extrapolation of the Chung *et al.* (1988) model (above). However, increasing temperature and duration of hydropyrolysis of Araripe shale has the opposite effect, further decreasing $\Delta^{13}C_2H_6$ by several tenths of per mille. The differences between these two experimental series could have several explanations. Perhaps the two source rocks in question contain ethane precursors with different carbon isotope structures, or perhaps the relative contributions of kerogen, bitumen and oil cracking over the temperature ranges of these experiments were not the same (presuming ethane formed by cracking these different precursors could differ in $\Delta^{13}C_2H_6$). It would be premature to prefer one of these or other plausible ideas before conducting more experiments of this kind. Nevertheless, these findings demonstrate that kerogen cracking can produce ethane with diverse clumped-isotope composition, varying by source rock and maturity.

Clog *et al.* (2013, 2014) and Clog & Eiler (2014) also present results of experiments in which ethane is destroyed by pyrolysis, either exposed to glass at 600°C or in the presence of glass and Ni catalyst at 200°C. Analysis of the residues of these experiments show them to be higher in $\delta^{13}C$ relative to starting materials (as expected for kinetically controlled pyrolysis that is faster for ^{12}C than ^{13}C), but substantially reduced in Δ_{13C2H6} (Fig. 3b). The fact that Δ_{13C2H6} values decrease as $\delta^{13}C$ increases is not surprising – this is the sign of the clumped-isotope effect anticipated for residues of molecular diffusion in most cases (Eiler 2007; 2013), and simple models of chemical-KIEs can follow mass laws similar to molecular diffusion (Bigeleisen 1949). However,

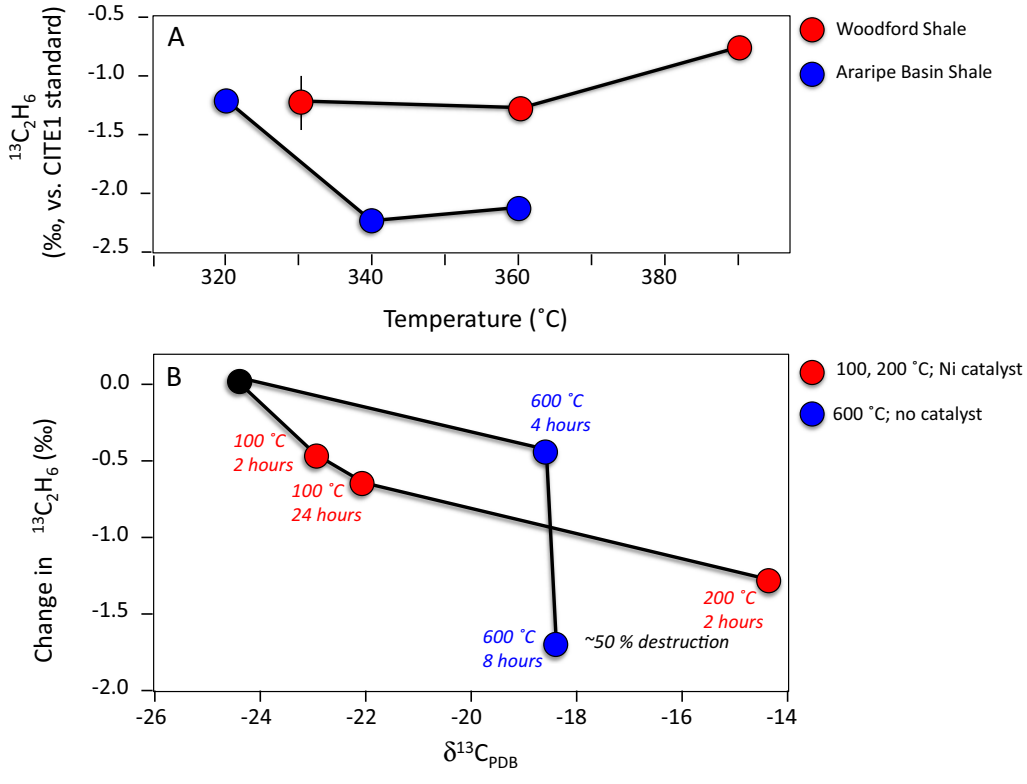

Fig. 3. $2 \times {}^{13}C$ clumped-isotope compositions ($\Delta^{13}C_2H_6$ values) of ethane: (**a**) produced by hydropyrolysis of shales; or (**b**) residual to ethane cracking. In (a), Δ_i values are reported relative to an arbitrary ethane standard; in (b) the vertical axis reflects the change in Δ_i from the starting composition. The highest $\delta^{13}C$ points in panel (b) reflect *c.* 50% ethane consumption. Data from Clog *et al.* (2014).

the magnitude of the decrease in $\Delta^{13}C_2H_6$ observed in ethane-cracking experiments is greater than can be explained by these sorts of transport-limited mass laws (which are generally restricted to *c.* 0.5‰ or less for tens of per mille ranges in $\delta^{13}C$). One possibility is that the KIE for ethane pyrolysis violates simple diffusion-like behaviour (e.g. if fractionations are controlled by differences in vibrational energy between reactant and transition states, or if the reaction mechanism has several steps). Another possibility is that the reactant has a large excess in $\Delta^{13}C_2H_6$ and that ethane consumption is slightly reversible, and thus is accompanied by a re-equilibration process that reduces that excess.

Propane and higher order hydrocarbon site-specific ^{13}C. Piasecki (2015) and Piasecki *et al.* (2016*b*, 2018) present analyses of the site-specific ^{13}C composition of propane (i.e. difference in $\delta^{13}C$ between terminal CH_3— and central —CH_2— groups) for products of hydropyrolysis of Woodford Shale (Fig. 4a). Their findings indicate that terminal carbon positions increase in $\delta^{13}C$ with increasing extent and temperature of pyrolysis, as expected by both the Chung *et al.* (1988) and Tang *et al.* (2000, 2005) models of isotope effects associated with *n*-alkane cracking to form propane. However, these experiments also reveal that the central carbon position also rises over the course of cracking, by several per mille. This finding clearly disagrees with the simplified assumptions implicit in the Chung *et al.* (1988) treatment; it is consistent in direction with the predictions of Tang *et al.* (2000, 2005), though different in magnitude (*c.* 100% greater than the predicted change in central position $\delta^{13}C$). There are several possible explanations of this result. First, the various temperature steps of this hydropyrolysis experiment produce propane from different proportions of precursors – kerogen dominating early and oil dominating late – and these precursors could differ in the $\delta^{13}C$ for the carbon that is transferred to the central position of propane (this is the explanation favoured by Piasecki (2015)). Second, propane formation may involve secondary KIEs that are

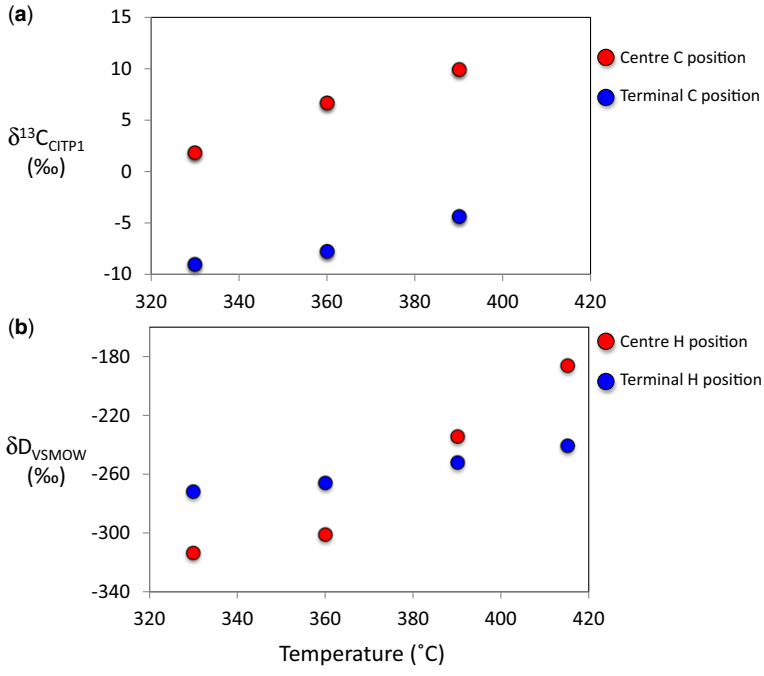

Fig. 4. Site-specific isotopic composition of propane produced by hydropyrolysis of Woodford Shale: (**a**) presents carbon isotope compositions, reported relative to the intralaboratory standard, CITP1 (data from Piasecki (2015)); (**b**) presents hydrogen isotope compositions reported relative to VSMOW, meaning the plotted centre-terminal difference reflects the actual value of the site-specific fractionation (data from Ponton *et al.* (2016)).

qualitatively similar to those predicted by Tang *et al.* (2000) but quantitatively stronger (possible, though perhaps unlikely given that these data imply the secondary isotope effect on the central C position is similar in strength to the primary isotope effect on the terminal C position). Third, it may be that propane forms not by cleavage of the CH_{2b}—CH_{2c} bond of an *n*-alkane, as assumed in the Chung and Tang models, but rather by decomposition of a branched or aromatic compound, involving a primary isotope effect on the site that donates carbon to the central position of product propane.

Julien *et al.* (2015) present the results of a study of site-specific carbon isotope fractionations associated with evaporation of several potential organic environmental contaminants. Although they find significant site-specific signatures of evaporation for ethanol and other small, polar compounds, the two naturally significant hydrocarbons they examined (toluene and *n*-heptane) show only insignificant or subtle site-specific differences in effective vapour pressure isotope effect (*c.* 1‰ or less). These findings suggest isotopic fractionations associated with volatility may be significant for some oil compounds (e.g. organic acids) but are unlikely to be large for the major components of oil and natural gas.

Propane site-specific D. There are no published details regarding the site-specific H isotope fractionations associated with pyrolysis of kerogen or other natural hydrocarbon precursors, and it is not obvious that they can be reasonably guessed at by extrapolation of the proposed models for ^{13}C effects. Thus, our only guide to this subject is the long-standing finding that δD values of natural gas and petroleum compounds often rise with increasing maturity in products from the same or similar source rocks, implying some sort of 'normal' KIE (i.e. H-bearing species react faster than D-bearing species, and increasing the progress of the reaction slowly diminishes this effect as products approach the composition of original substrates).

A recent conference presentation (Ponton *et al.* 2016) includes preliminary results documenting the site-specific D/H variation in propane generated by hydropyrolysis of Woodford Shale. These experiments (reproduced in Fig. 4b) show that the δD of both the terminal and central positions of propane rise by tens of per mille over the course of heating from 330 to 415°C (roughly corresponding to the transition from primary to secondary cracking in natural petroleum-generating systems), and that the difference in δD between these two sites also rises by

tens of per mille, from a non-equilibrium fractionation where the central position is lower in δD than the ends at low temperature/maturity, to an approximately equilibrated signature where the central position δD is greater than the terminal at high temperature/maturity (Fig. 4b). One interpretation of this finding is that increasing thermal stress is associated with an approach to equilibrium intramolecular hydrogen isotope structure, either through internal exchange of hydrogens among different sites in the same molecule or through heterogeneous exchange between the molecule of interest and some other compound (water, clay, etc.).

Natural bitumen and oil-fraction compounds

A single study has been published documenting the isotopic structures of samples of compounds that are major components of natural bitumens and oils: Gilbert *et al.*'s (2013) study of the carbon isotope compositions of the CH_3, CH_{2a} and CH_{2b} positions (i.e. terminal and adjacent two carbon sites) of C11–31 *n*-alkanes. Before we discuss these measurements, one key point must be made first: the sample set examined in this study consists of pure compounds purchased from chemical supply companies. Thus, while it seems possible that they are related in some way to natural biological and/or petroleum sources, there is little concrete information that ties them to any particular natural material or process. While we will attempt to relate Gilbert *et al.*'s (2013) findings to our understanding of the isotopic structures of fatty acids and isotope effects associated with kerogen and oil cracking (as those authors did), it is possible that these measurements are not directly relevant to any natural petroleum compounds.

Gilbert *et al.* (2013) find three patterns of carbon isotope variation in the C11–31 *n*-alkanes (Fig. 5a; note these measurements were made by NMR, and that they do not constrain differences between the three plotted sites and the bulk molecules – i.e. only site-to-site differences in each molecule are constrained): (1) all compounds having an odd number of carbons and 17 or more carbons in all exhibit a pattern where both CH_3 and CH_{2b} are several per mille lower in $\delta^{13}C$ than the CH_{2a} position; (2) all compounds having an even number of carbons and

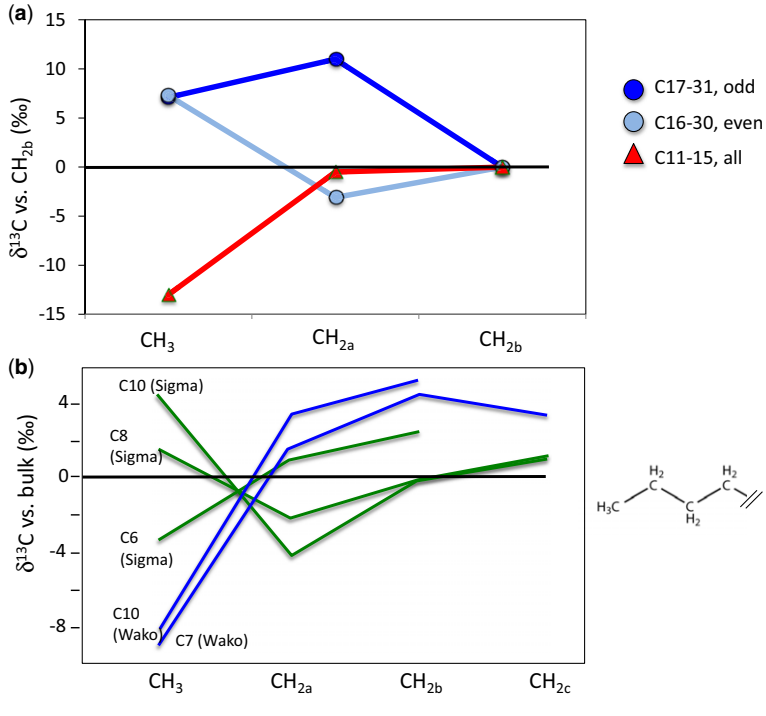

Fig. 5. Site-specific carbon isotope structures of the terminal 3 or 4 sites of *n*-alkanes from (**a**) oil-fraction compounds (from Gilbert *et al.* (2013)) or (**b**) condensate-fraction compounds (from Gilbert *et al.* (2016*b*)). In (a) data are reported relative to the CH_{2b} site (third from the end) and it is not known how any of these sites relate to the bulk $\delta^{13}C_{PDB}$; in (b), data are reported relative to the known molecule-average ('bulk').

16 or more carbons in all have a pattern where both CH_3 and CH_{2b} are several per mille higher in $\delta^{13}C$ than the CH_{2a} position (curiously, the difference between CH_3 and CH_{2b} is the same, on average, for these two groups); and (3) all compounds having 15 or fewer total carbons (whether even- or odd-carbon number) display a pattern where the $\delta^{13}C$ of the CH_{2a} and CH_{2b} sites are the same as one another but the CH_3 site is substantially ^{13}C-depleted relative to both CH_{2a} and CH_{2b}.

Gilbert *et al.* (2013) suggest that the carbon isotope patterns of the C16+ *n*-alkanes can be understood to be products of decarboxylation of fatty acids, which are known to exhibit even/odd ordering of alternately high and low $\delta^{13}C$ values (though in detail it is not clear how the two patterns they observe in *n*-alkanes relate to the diverse carbon isotope structures of fatty acids from *E. coli* and yeast; see above). The C11–15 *n*-alkanes could be understood to be products of cracking longer-chain alkanes and fatty acids, where mixing between fragments derived from even- and odd-carbon number precursors averages out to produce a mixture with no significant 'even/odd' isotopic ordering in the interiors of carbon chains, but a KIE associated with the cracking reaction leaves its imprint as ^{13}C depletion of the terminal carbons (at least one of which was presumably adjacent to the bond that broke in the longer compound).

Gilbert *et al.* (2013) clearly demonstrate the existence of systematic, high-amplitude carbon isotope variations in the *n*-alkanes; and, although one could propose other explanations of their findings, the hypotheses they present are easily understood and sensibly derived from an understanding of fatty acid synthesis. For these reasons, this study suggests large alkanes inherit their isotopic structures from their fatty acid precursors, whereas small ones mostly express 'mean reversion' through mixing, plus the fractionations associated with cracking (Gilbert *et al.* 2013). But, we must recall that none of this argument is based on measurements of recognized natural petroleum compounds. Studying natural geochemical materials and processes will be much more technically challenging because the NMR techniques used by Gilbert *et al.* (2013) currently require *c.* 0.1–1 g of pure samples. Each of the *n*-alkanes generally make up only *c.* 0.1–0.5 wt % of natural crude oil (*Oil in the Sea*, National Research Council 2003), meaning one analysis would require isolation out of one or more litres of crude oil, hundreds of litres of oil-contaminated waters, or kilograms of oil-containing rock. This is feasible but presents a so-far unaddressed technical challenge. Alternatively, it may be possible to study isotopic structures of oil-fraction compounds by pyrolysis followed by mass spectrometry (Gilbert *et al.* 2016*a*, *b*) or direct mass spectrometry (Piasecki *et al.* 2016*b*), with significantly reduced sample sizes. However, the instruments and methods capable of such measurements are still in development (see Appendix B).

Natural condensate-fraction compounds

Gilbert *et al.* (2016*b*) present (in abstract form) the initial results of an NMR study of the carbon isotope structures of the C6,7,8 and C10 *n*-alkanes – species that are most strongly concentrated into the condensate fraction of petroleum systems (i.e. species less volatile than natural gas constituents, found in light oils, and more volatile than most components of bitumen and heavy oils). Their findings (Fig. 5b) loosely resemble observations for the C11–15 *n*-alkanes in their previous study (Fig. 5a; Gilbert *et al.* 2013): the interior CH_2 sites are relatively uniform in $\delta^{13}C$ and conform to a single, relatively simple pattern (CH_{2a} 1–2‰ lower in $\delta^{13}C$ than CH_{2b}, which is within *c.* 1‰ of CH_{2c} when present), whereas the terminal sites are strongly fractionated relative to the average of all CH_2 sites. However, there are two discrepancies between this pattern and that for the average C11–15 pattern in Figure 5a: (1) terminal CH_3 groups may be strongly ^{13}C depleted *or* enriched relative to CH_2 sites; and (2) the same compound obtained from different sources (C10 from Waco or Sigma Aldrich suppliers) can have opposite signs of the CH_3—CH_{2a} fractionation. It is difficult to imagine how these findings can be reconciled with the concept that these compounds are formed by mixing the products of cleaving fatty acid or longer alkane chains, superimposed on a normal primary KIE at the terminal sites. Furthermore, as for the C11–31 alkanes discussed above, these data were generated on pure alkanes purchased from chemical suppliers that generally do not provide detailed information about sources and formation mechanisms. For this reason, we can only speculate as to how these findings might speak to the carbon isotope structures of natural low-molecular-weight hydrocarbons.

Natural gas constituents

Up to this point, our focus on our subject has gradually blurred as we have moved from the long-standing body of work on relatively well-understood biosynthetic compounds, through the veil of barely known isotopic structures of humins, lignins and kerogens, to a sparse sampling of compounds that may or may not be related to components of natural bitumens and oils. However, there is some promise at the end of our story, as the isotopic structures of constituents of natural gases have been the subject of a relatively large, rapidly growing and diverse body of

analytical work. The largest part by far concerns the abundances of doubly substituted forms of methane ($^{13}CH_3D$ and, more recently, $^{12}CH_2D_2$). This subject has grown so large in the last year that it warrants its own review (Douglas *et al.* 2017; Stolper *et al.* 2017). We focus here on the smaller but still significant body of recent work on the carbon and hydrogen isotope structures of natural propane and ethane.

Site-specific ^{13}C in natural propane

Currently, four independent research groups using four different analytical technologies are exploring the difference in $\delta^{13}C$ between the central and terminal C positions in propane, including: (1) high resolution gas source mass spectrometry (Piasecki *et al.* 2016*b*, 2018); (2) pyrolysis followed by gas chromatography (GC) separation and combustion of the products and IRMS of the resulting CO_2 peaks (Gilbert *et al.* 2016*a*); (3) sequential chemical degradation of propane followed by separate combustion of the products and IRMS on the resulting CO_2 (Gao *et al.* 2016); and (4) NMR on propane condensed in high pressure tubes (Liu *et al.* 2015). Only the first of these techniques has been applied to suites of natural gases having known origins, but all four provide important complementary constraints: NMR data and selective chemical degradations yield site-specific data that can be directly anchored to absolute scales (e.g. Pee Dee Belemnite (PDB)) but their sample sizes make it challenging to create large datasets on natural materials, whereas the mass spectrometric and pyrolitic techniques permit smaller sample sizes but are essentially relative methods, constraining only differences in site-specific $\delta^{13}C$ between samples and reference propane. Taken together, these various methods provide the following picture of intramolecular ^{13}C distribution in natural propane.

Gao *et al.* (2016) show that propane in the widely used natural gas reference standard, 'NG3', has a difference: $\delta^{13}C_{centre} - \delta^{13}C_{ends}$ of 19.2‰, where both quantities are known on the PDB scale. This is in the same direction and has the same order of magnitude as the equilibrium carbon isotope fractionation predicted by first-principles theory at geologically relevant temperatures (above). Liu *et al.* (2015) report an NMR measurement showing a similar centre–end fractionation of 15.8‰ for an unknown (or at least unreported) propane gas; considering that nearly all readily available propane is from natural thermogenic gas, it seems likely that this is a material generally similar to propane in NG3. Gilbert *et al.* (2016*a*) report the centre–end carbon isotope fractionation for three propanes: NGS2 (another widely available natural gas standard) = 12.8‰, whereas two other commercial propanes of uncertain origin have much lower values of 8.2 and 1.8‰. This result is compromised by the fact that the pyrolysis method used in this study involves unknown fractionations that potentially effect accuracy; thus, the fact that they find smaller absolute fractionations than Gao *et al.* (2016) and Liu *et al.* (2015) may not be significant. Nevertheless, it is significant that Gilbert *et al.* (2016*a*) report a large (11‰) range in the centre–end fractionation.

Piasecki (2015) and Piasecki *et al.* (2016*b*, 2018) report the site-specific carbon isotope compositions of ten natural propanes from the Antrim Shale (Michigan), Eagle Ford Shale (Texas), and Potiguar Basin (Brazil). These data were measured by mass spectrometry and (as for the Gilbert *et al.* (2016*a*) observations) constrain *relative* differences in $\delta^{13}C_{centre}$ and $\delta^{13}C_{ends}$ between samples but not the *absolute* centre–end fractionation. The finding of this work (Fig. 6a) is that, as one moves from lower maturity wet gases (Antrim Shale and lower maturity oil-associated Potiguar gases) to higher maturity gases (higher maturity Potiguar and Eagle Ford suites), the $\delta^{13}C$ of both terminal and centre-position carbons rise. Assuming that the rise in bulk $\delta^{13}C$ is a generally useful proxy for increasing thermal maturity (Prinzhofer & Huc 1995), the data trend in Figure 6a suggests that this rise is first concentrated exclusively in the terminal C position for the first half of the maturity range (a trend consistent with the Chung *et al.* (1988) model), after which the centre position becomes the dominant site of further increases in $\delta^{13}C$. One interpretation of this finding is that the first half of the trend reflects primary kerogen cracking following the expectations of the Chung *et al.* model, whereas the second half of the trend marks a shift to another precursor, such as bitumen or oil cracking (so-called secondary cracking). This interpretation is consistent with the findings for Woodford Shale cracking experiments (above; Piasecki *et al.* 2018), but remains speculative and should be examined with further experiments and studies of closely related natural gases that vary in thermal maturity.

Site-specific D in natural propane

Liu *et al.* (2015) report NMR data documenting site-specific hydrogen isotope fractionation in propane, $\delta D_{centre} - \delta D_{ends} = -26 \pm 10‰$ for an undescribed propane. This contrasts with the fractionation of *c.* +50–80‰ theoretically predicted for equilibrium at relevant geological temperatures, and, taken at face value, suggests that propane can be generated with an unequilibrated hydrogen isotope structure.

Ponton *et al.* (2016) present preliminary results of mass spectrometric measurements of the site-specific hydrogen isotope composition of 12 natural propanes from several petroleum-forming systems:

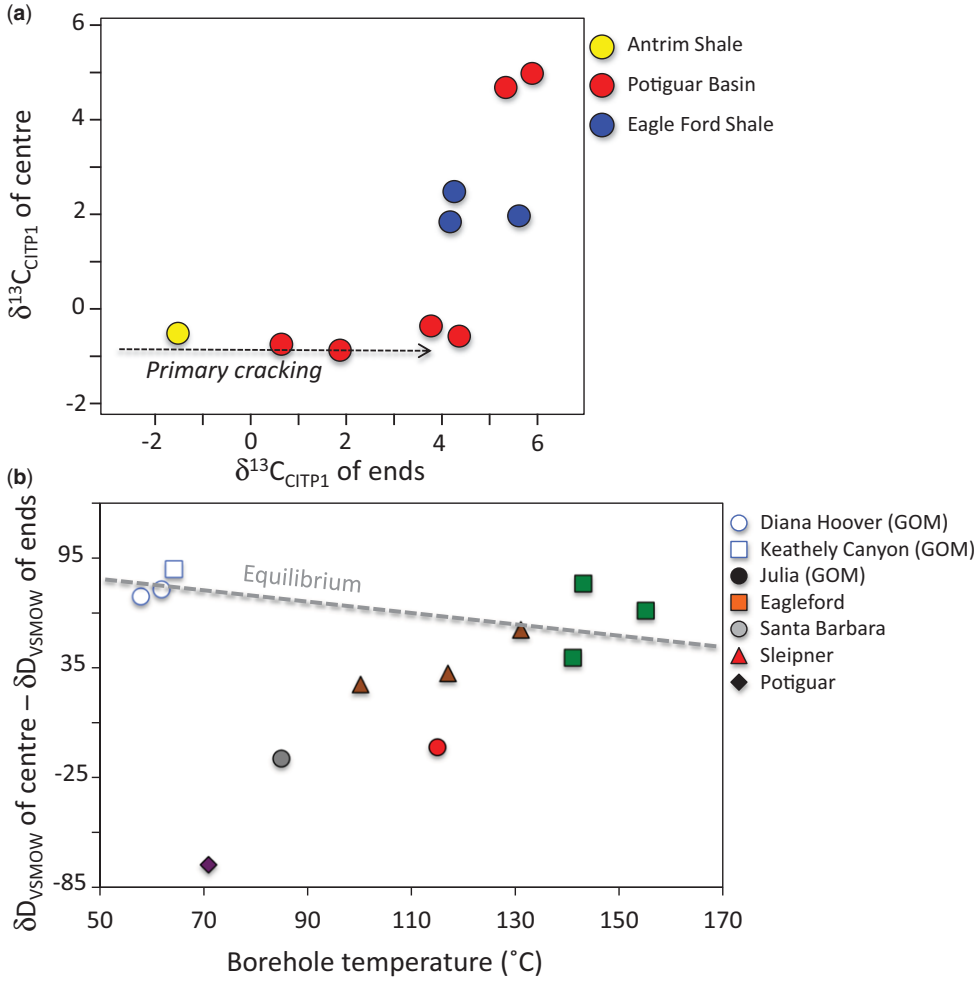

Fig. 6. Site-specific isotopic structures of natural propane. (**a**) carbon isotopes, from Piasecki (2015) (reported v. the intralaboratory standard, CITP1); (**b**) hydrogen isotopes, from Ponton *et al.* (2016) (on an absolute reference frame scale). In (a) the arrow represents the path through this composition space predicted by the model of Chung *et al.* (1988). In (b), the dashed line represents the equilibrium site-specific hydrogen isotope fractionation predicted by Piasecki *et al.* (2016*a*) (similar to that in Webb & Miller (2014)).

three wells in the Gulf of Mexico, the Eagle Ford Shale (Texas), Sleipner (North Sea), Potiguar Basin (Brazil) and Sacate field (coastal California). These data were generated by mass spectrometry but, unlike the ^{13}C data described above, can be tied to an absolute δD_{VSMOW} scale because the same techniques were applied to propane that had been experimentally driven to an equilibrium isotopic structure by exposure to Pd catalyst (i.e. propane can be standardized using a 'heated gas' or 'absolute reference frame' of the kind previously used for clumped-isotope and/or site-specific stable isotope measurements of CO_2, O_2, CH_4 and N_2O; Eiler 2007, 2011). The commercially obtained propane gas cylinder used by Ponton *et al.* (2016) as an intralaboratory standard (a gas with an unknown origin, but probably purified from thermogenic natural gas) yields a site-specific hydrogen isotope fractionation, $\delta D_{centre}–\delta D_{ends} = -25 \pm 6‰$, which is essentially identical to the finding by Liu *et al.* (2015) on their intralaboratory reference material. It is tempting to imagine that commercial suppliers provided the same or similar propane to both labs, although this is not known to be the case and could equally well be a coincidence (or conspiracy of errors!).

The results of Ponton *et al.* (2016) for natural propanes should be considered preliminary, but define a

trend where the bulk molecular δD rises from the lowest to highest values commonly seen in thermogenic propane (−170 to −115‰) while the centre–end hydrogen isotope fractionation rises from −72 to +89 (±6)‰. This finding is similar to results for propane generated by hydropyrolysis of Woodford Shale (centre–end fractionation rises from −52 to +67‰ over a 69‰ rise in δD; above). One interpretation of this result is that the site-specific hydrogen isotope fractionation in propane is a maturity indicator, analogous to common interpretations of the bulk $\delta^{13}C$ or δD of natural gas components. However, a different interpretation is suggested by two observations: (1) the high end of the range of centre–end fractionations is similar to the expected equilibrium fractionation at shallow crustal temperatures; and (2) the centre–end hydrogen isotope fractionation exhibits a relationship with the borehole temperature of the well from which each sample was collected (Fig. 6b). In particular, 9 of the 12 studied samples exhibit a positive trend, the high end of which is indistinguishable from the predicted equilibrium relationship, whereas the remaining three samples (from the Dianna Hoover and Keathley Canyon wells in the Gulf of Mexico) also lie on the predicted equilibrium trend but at lower temperatures (60 v. 150°C). This finding suggests the possibility that thermogenic propane may be formed out of equilibrium with respect to its site-specific hydrogen isotope structure but then evolve towards equilibrium.

The most obvious process that might promote such equilibration is intra- or intermolecular hydrogen isotope exchange, perhaps catalyzed by some coexisting material (e.g. water, clay or oil). This process would probably proceed more quickly at higher temperature and thus would provide an explanation for the fact that the gases from the highest temperature wells approach the equilibrium distribution. It is less obvious why the group of three Gulf of Mexico gases also have an equilibrium hydrogen isotope structure despite their low borehole temperatures. Four possibilities occur to us. First, there could have been prior storage of propane at greater temperatures in some deeper reservoir. Second, some material may be present in the reservoirs for these samples (that is absent in the other reservoirs) and particularly effective at promoting re-equilibration down to low temperatures. Third, reservoir microbial ecosystems (known to be present and actively generating methane in the Keathley Canyon field) may be able to catalyze hydrogen isotope exchange of petroleum compounds. This idea is inspired by prior work of Valentine *et al.* (2004), who showed that the enzymatic pathway responsible for biological methanogenesis operates nearly reversibly when the partial pressure of H_2 is low; thus, microbes could act as a kind of specialized catalyst for hydrogen isotope exchange that short-circuits the thermally-activated, abiological pathways. Fourth, perhaps low maturity gases are generated by a mechanism that promotes intramolecular equilibrium, whereas higher temperature cracking of propane precursors is kinetically controlled (perhaps counter-intuitive, but we would say possible). All four of these suggestions are speculative at this stage, but illustrate the sorts of processes that may be revealed by further exploration of this proxy.

$^{13}C_2H_6$ in natural ethane

Clog *et al.* (2013, 2014) and Clog & Eiler (2014) present $\Delta^{13}C_2H_6$ values for ethane from 25 natural gases from the Marcellus Shale (Pennsylvania), Haynesville Shale (Texas and Louisiana), Eagleford Shale (Texas) and the Potiguar and Sergipe Alagoas petroleum fields (Brazil). To the best of our knowledge, this constitutes the largest and most diverse set of observations constraining the intramolecular isotopic structure of a hydrocarbon larger than methane recovered from natural samples. The most striking finding of this work is the relatively large range of variability in $\Delta^{13}C_2H_6$ (5 ‰, or a factor of *c.* 10–20 greater can be easily explained by equilibrium clumped-isotope effects or simple physical processes like diffusion or mixing). The measured range is also *c.* five-fold greater than that observed in products of shale pyrolysis experiments (above). This suggests that most of the variance in this isotopic proxy must be driven by something other than common kerogen and oil-cracking reactions, temperature-dependent equilibrium or simple processes like diffusive transport. One process that might explain these findings is ethane cracking: the experiments summarized above produced a 1.8‰ range in $\Delta^{13}C_2H_6$; if this process operated as a Rayleigh distillation, the residue left over after *c.* 90% ethane loss could manifest the full range in $\Delta^{13}C_2H_6$ values Clog *et al.* (2014) observed in natural samples. If this process is in fact responsible for most of the measured range, then $\Delta^{13}C_2H_6$ might be most useful as a proxy for secondary cracking of wet gas components. Two observations offer support for this possibility: (1) relatively dry shale gases are consistently in the lower half of the range of measured $\Delta^{13}C_2H_6$ values; and (2) a plot of gas wetness v. $\Delta^{13}C_2H_6$ (Fig. 7) reveals a positive overall trend, and better-defined positive trends for each of three suites of related samples – the Potiguar, Eagle Ford and Sergipe Alagoas. These trends are broadly consistent with the relationship between wetness and $\Delta^{13}C_2H_6$ predicted based on the ethane-cracking experiments (solid curve in Fig. 7).

A word on the doubly substituted methanes

The clumped-isotope compositions of natural methanes are reviewed in Stolper *et al.* (2017); see also

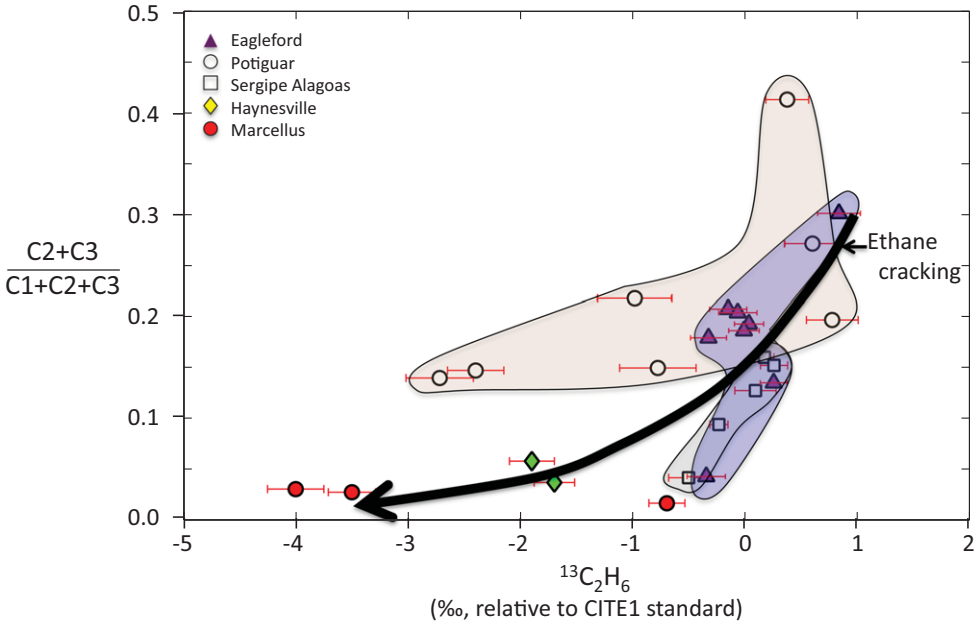

Fig. 7. Clumped-isotope composition of ethane from natural gas samples, plotted v. the 'gas wetness' (proportion of ethane + propane to the total of methane + ethane + propane) in those samples; data from Clog *et al.* (2013, 2014) and Clog & Eiler (2014) (reported v. the intralaboratory standard, CITE1). For comparison, the heavy black arrow shows the trend predicted for ethane cracking, assuming Rayleigh distillation and isotopic fractionation based on experiments depicted in Figure 3b.

the review by Douglas *et al.* 2017). We do not discuss this topic in detail here, but point out one finding that is more striking and meaningful when contrasted with the subjects covered in this paper: methane is commonly, even typically, in clumped-isotope equilibrium at its temperature of formation, across a wide range of conditions and formation mechanisms. Recent papers on methane clumped-isotope geochemistry have focused attention on non-equilibrium clumped-isotope compositions. However, these findings have been observed only in a few settings and formation mechanisms: a subset of biogenic and serpentine-related methanes; experimental thermogenic methane formed by alkyl cracking; and possibly a subset of oil-associated unconventional gases. The more common finding (estimated at *c.* 80+ % of natural samples analysed to date) is that methane apparent temperatures are similar to known or expected temperatures of methanogenesis. For this reason, the methane clumped-isotope proxies may be a rare instance where Galimov's (1974, 1985) hypothesis of an organic stable isotope geochemistry dominated by equilibrium thermodynamics seems to be usually true. Why does methane stand in contrast to the other groups of organic compounds that have been explored for their isotope structures – lipids, sugars, amino acids, oil and condensate alkanes, propane, ethane? The question is particularly puzzling when one considers that the bulk stable isotope fractionations associated with methanogenesis have been successfully described through KIEs (Tang *et al.* 2000, 2005). A reasonable conclusion is that methanogensis often occurs in chemical environments that promote local reversible H exchange among the moieties that will become methane (and/or with some other, unrecognized, hydrogen pool, such as H radicals in kerogen structures). It is striking that this seems to be the case both for most natural biogenic methanogenesis in subsurface environments and for thermogenic methanogenesis involving kerogen cracking. An important question for future research is whether this is also true of the hydrogen isotope structures of larger organic compounds, which are essentially unknown outside of sugars and propane.

The end game: graphitization

Burial or subduction of organic matter to conditions where rocks undergo macroscopic metamorphic reactions (*c.* 300+ °C) leads to dramatic reorganization of the H/C/N/O/S volatiles from organics into just a few stable forms: graphite (or, at high pressure,

diamond), pyrite or pyrrhotite, ammonium ion in feldspar or sheet silicates, and simple molecular gases (Harrison 1976). However, this transformation is gradual and severely limited by kinetics, preserving carbonaceous material that is structurally intermediate between kerogen and crystalline graphite well into high metamorphic grades (Mao *et al.* 2010). (There is also evidence for traces of C2+ hydrocarbons in metamorphic rocks (Mango 1991), although one might ask whether some such evidence reflects contamination rather than persistence of complex organic molecules.) Nevertheless, as temperature rises through the metamorphic range, eventually NH_3, H_2S, CO_2 and even CH_4 become unstable, leaving crystalline graphite and the inorganic molecular gases (CO_2, CO, H_2O, H_2) as the only abundant remnants of buried organic matter.

These processes raise three questions about the stable isotope contents of organics in the rock record. (1) How long do vestiges of the original isotopic structures of biomolecules or fingerprints of catagenetic 'cracking' reactions persist? (2) At what stage can one no longer distinguish between the products of metamorphosing biomolecules and non-biological sources of graphite and simple molecular gases? (3) Do any organic materials have thermodynamically controlled isotopic structures under such extreme conditions? There is a long history of approaching these questions using bulk stable isotope data (e.g. the $\delta^{13}C$ value of graphite, or the hydrogen isotope fractionation between methane and H_2; Dunn & Valley 1992; Wada *et al.* 1994; Horibe & Craig 1995; Schimmelmann *et al.* 2001). Such constraints are useful, but must contend with ambiguities that arise from the wide range of geological processes that can raise or lower bulk isotopic contents. We ask: what insights could molecular isotopic structure (site-specific and clumped-isotope compositions) bring to these questions?

Graphitization of kerogen – transformation of its non-periodic polymerized structure into crystalline graphite, accompanied by loss of H, N, O and S – begins in the 'gas window' (*c.* 200°C) and occurs very gradually, with increases in crystallinity and coarsening in grain size persisting well into the amphibolite facies (Harrison 1976). The $\delta^{13}C$ and δD values of kerogens do not evolve radically over the oil and gas windows (*c.* 100–250°C), but studies of the bulk carbon isotope fractionation between graphitic carbon and coexisting carbonate in metasediments show that subsequent graphitization is accompanied by carbon isotope exchange (Dunn & Valley 1992). This exchange starts around 400°C (Wada *et al.* 1994) but is slow and does not reach equilibrium until temperatures reach *c.* 500–600°C. A key insight is that even at amphibolite-facies metamorphic conditions, small graphite grains are further from equilibrium with coexisting carbonate than are large graphite grains (Dunn & Valley 1992); this implies that carbon isotope exchange requires grain growth – i.e. little or no exchange occurs by swapping carbon atoms in and out of the sites of existing graphitic lattices. Taken together, these findings suggest the carbon isotope structures of components of kerogens may be preserved to great depths and temperatures in the Earth. If so, preservation of large site-specific carbon isotope variations in kerogens and 'graphitic' carbon might provide a more conclusive proof of biogenicity than bulk $\delta^{13}C$ values – a widely used but controversial and often ambiguous biomarker (e.g. Van Zuilen *et al.* 2002; Ohtomo *et al.* 2014). On the other hand, it is also possible that graphitization involves little heterogeneous carbon isotope exchange but significant amounts of molecular- and lattice-scale isotope rearrangement that could erase original biogenic isotopic structures.

Conclusions and prospects

We have attempted to provide a comprehensive and up-to-date overview of what is known about the molecular-scale isotopic structures of natural organic compounds, from their origins in biosynthesis through their transformation by catagenesis, to their destruction by metamorphism. This is a topic of tremendous complexity and with great potential to record information of scientific and practical importance. However, the technologies that enable site-specific and clumped-isotope analyses of organic compounds have only recently emerged (though they are advancing rapidly; Appendix B), there has been little inter-calibration of the diverse instruments and labs, and the application of these technologies has been sparse and in some cases difficult to relate to environmental and geological samples. For these reasons, the most lasting value of this review may be to organize and illuminate the scattered elements of an emerging field, in hopes that we can lend the field structure, pose initial hypotheses about its most important processes and phenomena, and make clear some of its most pressing needs.

The work summarized in this paper, and in the accompanying paper by Stolper *et al.* (2017) provide evidence that the isotopic structures of geological organic compounds can reflect any combination of the following factors: inheritance from primary biomolecules (e.g. the carbon isotope structures of aliphatic hydrocarbons, presumably derived from lipids); KIEs associated with maturation of refractory organics into humin, lignin and kerogen (e.g. extreme ^{13}C enrichment in methoxyl groups of lignins); KIEs during kerogen cracking (e.g. ^{13}C depletion in terminal sites of propane and other alkanes); thermodynamic equilibria at conditions of

thermogenic reactions (the best example of which may be the clumped-isotope compositions of the most thermogenic methanes); KIEs during secondary cracking (e.g. the $\Delta^{13}C_2H_6$ value of ethane, particularly in dry gases); equilibration after formation due to protracted time at geological temperatures (e.g. the site-specific hydrogen isotope compositions of some propanes); KIEs during biological production and consumption (e.g. clumped-isotope compositions of some biogenic methanes in earth-surface environments); and biologically catalyzed equilibration (e.g. clumped-isotope compositions of methanes in subsurface environments). This rich set of processes has been revealed by just the last several years of exploratory work, promising much more to come as our analytical tools mature and see broader application.

We close our discussion with a list of conclusions and suggestions regarding future work that will be particularly important for near-term progress,

(1) *The interplay of chemical kinetics and equilibria*: Our first explorations of the isotopic structures of geological organics indicate that both chemical kinetics and thermodynamic equilibria are capable of controlling intramolecular isotopic properties. For example, carbon isotope structures of common biomolecules are clearly dominated by kinetic factors and the same seems to be true of hydrogen isotope structures of freshly synthesized sugars and at least some geological propanes. Yet methanes often have equilibrated distributions of isotopologues, propanes from some natural environments seem to have equilibrated hydrogen isotope structures, and at least some moieties of kerogens undergo hydrogen isotope exchange with their environments during burial diagenesis. The fact that both kinetic and equilibrium processes seem to be important in natural organics is sure to lead to confusion, at least in the near term, and raising the following questions. (1) When should an isotopic structure be interpreted as a biosignature v. a thermometer. (2) What are we to make of samples that have only partially equilibrated some initially kinetically controlled structure (or kinetically modified an initially equilibrated structure). (3) How will we parse the isotopic structures of molecules that have some properties controlled by kinetics and other properties controlled by thermodynamic equilibria? However, a longer-term view suggests these phenomena will be a rich archive rather than a problem. The combined site-specific and clumped-isotope properties of organic molecules include a large number of analysable species, and it seems plausible that we will learn to read complex multi-property measurements as records of sources, formation mechanism and environmental conditions. Furthermore, once the kinetics controlling the approach to equilibrium are better understood, it is possible that molecular isotopic structures will present opportunities for reconstructing temperature–time histories (i.e. a form of 'geospeedometry'). The most promising opportunities of this last sort may be the hydrogen isotope structures of relatively volatile and structurally simple compounds, such as the low-molecular-mass alkanes.

(2) *Defining isotopic equilibrium*: Considering the geological, environmental and economic importance of hydrocarbons, and the major role stable isotope analysis plays in their study, it is surprising how little concrete information constrains isotope exchange equilibria involving these compounds. This generalization is true both for the homogeneous equilibria that potentially control molecular isotopic structure and for the more familiar heterogeneous isotope exchange reactions between different chemical species. Other than vapour pressure isotope effects measured on low-temperature liquids (Jansco & Van Hook 1974), there are few concrete data that confidently constrain isotope exchange equilibria involving hydrocarbons larger than methane (Julien *et al.* 2015). This is an extraordinary gap in our understanding of the fundamentals of this subject, and should serve as a call for future research.

(3) *First-principles models*: Given the importance of irreversible reactions in petroleum formation and the great number and diversity of petroleum compounds, it is clear that the conceptual and theoretical models summarized in section 'KIEs associated with 'cracking' reactions' are too simple and narrow to serve as the basis for a rigorous interpretation of the stable isotope compositions of oil and gas constituents. The highest-priority needs include: extension of rigorous models of KIEs to encompass precursors other than *n*-alkanes and products larger than propane (particularly condensate and oil-fraction compounds); consideration of hydrogen isotope effects associated with diagenetic and catagenetic reactions; and a theoretical exploration of KIEs for clumped-isotope species (i.e. beyond the simple sampling-statistics effects discussed above with reference to ethane).

(4) *The kerogen problem*: The greatest weakness of our subject is that we have little insight into the detailed isotopic structures of kerogens, and it is not obvious how any of the existing

analytical technologies (Appendix B) could be adapted to study them. We suggest that the best way to approach this problem may be by analysis of the isotopic structures of oil and gas compounds produced by controlled heating of kerogens – that is, to approach it as an inverse problem rather than by direct measurements. Such studies might be made more insightful by studying the products of cracking experiments performed on compounds that are model systems for kerogenous material, with sites of interest isotopically labelled to aid characterization of reaction mechanisms.

(5) *Technology development*: The subject of this review is an emerging field that has come into being as a result of technology experiments aimed at expanding the capabilities of stable isotope geochemistry (Appendix B). Most of those technologies are relatively young and rapidly evolving, such that we should expect the next several years to result in dramatic increases in the numbers of active laboratories, the volume and quality of observations, and the numbers and types of molecular isotopic properties that can be observed. The most pressing need is for technological innovations that will permit analyses of small samples (of the order of micromols and less) of large molecules (of the order of ten and more atoms per molecule), preferably constraining diverse site-specific and clumped-isotope properties. NMR is the most mature of the relevant techniques, but it requires a dramatic reduction in sample size and the development of the ability to quantify multiply substituted species – something that will only become possible with a very large improvement in sensitivity. Mass spectrometry seems to have the greatest promise as a highly sensitive technique that is applicable to diverse properties of diverse compounds, but demonstrated instruments lack the mass resolution or number of detectors that would be required for a truly general approach to the problem. Fourier transform mass spectrometry holds promise as a way past these hurdles, but is as yet little understood as a quantitative precise tool for measuring natural isotope abundances.

The authors are grateful to the editor and two anonymous reviewers for their insightful comments and suggestions on an earlier draft of this work. Work presented in this paper was supported by the National Science Foundation programmes in Petrology and Geochemistry and EAR Instrumentation, by sponsored research support from Exxon Mobil and Petrobras, and by the California Institute of Technology. We thank Nami Kitchen for her contributions to analytical and experimental work presented here.

Appendix A

Nomenclature

The molecular-scale isotopic structures of organics is an emerging, relatively specialized field, and therefore uses a language and quantitative units that may be unfamiliar to some readers. A summary of this subject is challenging because its nomenclature and units are currently mixed (and in some respects confused). However, the following synopsis should provide a useful guide to most recent papers.

The isotopic forms of molecules have been described using the words: 'isotopologue', which we take to mean a version of a molecule that is unique in its number and/or location of rare-isotope substitutions; and 'isotopomer', which we take to mean one of some number of isotopologues that share a common isotopic stoichiometry (e.g. the same number of ^{13}C atoms) but differ from one another in the sites of those rare-isotope substitutions. Thus, by the usage adopted here (and in many recent papers) all isotomomers are also unique isotopologues of the same molecule. At least one recent study suggests the term 'isotopocule' (Toyoda *et al.* 2015), which is effectively equivalent to our use of the term isotopologue.

Any isotopologue that contains two or more rare isotopes (^{13}C, D, ^{15}N, ^{34}S, etc.) may be termed a 'multiply substituted' or (colloquially) 'clumped' isotopologue. Thus, $^{13}C_2H_6$ is a clumped isotopologue of ethane. 'Site-specific' or 'position-specific' isotopic differences generally refer to differences in proportions of two or more isotopologues that share a common number and type of rare isotopes but differ in the molecular sites of isotopic substitution. Thus, $^{12}CH_2D—^{12}CH_2—^{12}CH_3$ and $^{12}CH_3—^{12}CHD—^{12}CH_3$ are two isotopologues of propane that differ in their sites of deuteration; thus they are also isotopomers of one another. Samples that differ in proportions of these two species are said to exhibit site-specific isotopic fractionation or variation.

The units used to report clumped-isotope variations are relatively uniform and easily explained, though to date they have only been applied to relatively simple molecules (CO_2, N_2O, CO_3^{-2}, N_2, CH_4, C_2H_6), and it is not clear that current practice will be appropriate for discussion of the multiply substituted isotopologues of larger, more complex molecules (particularly those where clumped-isotope effects are themselves site-specific). Nevertheless, there is only one common practice in this field: clumped-isotope compositions are reported as Δ_i values, where i refers to an isotopologue of interest (or sometimes a collection of isotopologues that all share the same cardinal mass; thus i can be either a chemical and isotopic formula or a cardinal number). Δ_i values are calculated as:

$$\Delta_i = \left(\frac{R_i}{R_i^*} - 1\right) \times 1000 \qquad \text{(A1)}$$

where R_i is the ratio of the isotopologue of interest to the unsubstituted isotopologue of that same molecule, and R_i^*

is the value that ratio would have if all isotopes were randomly distributed among all isotopologues of the compound of interest. A less commonly used alternate to this nomenclature defines Δ_i values to equal the deviation, in per mille, of the equilibrium constant for an isotope exchange reaction from the value of that equilibrium constant in the case of random distribution of isotopes among all isotopologues (generally equal to the high temperature limit of that equilibrium constant; Wang *et al.* 2004; Piasecki *et al.* 2016*a*).

The nomenclature used to describe site-specific isotopic variations is more variable, in many cases unique to each paper (even when considering multiple papers from the same research group). Previously proposed units can be broadly defined as falling into two categories.

First, reports of the isotopic composition of a single indicated site are presented as: a concentration of a rare isotope (e.g. [^{13}C]); or, an isotope ratio (e.g. $^{13}C/^{12}C$, or R^{13}); or, a 'delta' value relative to an absolute reference scale (e.g. $\delta^{13}C_{VPDB}$); or, a 'delta' value relative to an arbitrary standard ($\delta^{13}C_{STD}$). Second, reports of an isotopic difference between two non-equivalent sites, A and B, are usually given as: a ratio of ratios (e.g. R_A^{13}/R_B^{13}, or equivalently α_{A-B}^{13C}); or, as an ε value (where $\varepsilon_{A-B} = 1000 \ln(\alpha^i_{A-B})$; or, as a difference between two 'delta' values (e.g. $\delta^{13}C_{VPDB}$ of site A minus $\delta^{13}C_{VPDB}$ of site B). A dizzying diversity of superscripts, subscripts and other symbols are used to indicate which molecular site is being discussed, meaning one generally must approach the units and terms of each paper as a unique kind of jargon.

Clearly, there is a need to improve the consistency of units used to report site-specific isotopic variations. We suggest the best nomenclature would avoid arithmetic artifacts arising from non-linearities in δ^iX_{STD} scales, and would focus on the unique information recorded by site-specific variations as distinct from that recorded by the more familiar indices of bulk stable isotope composition; ε^i_{A-B} values may be best suited to this purpose (i.e. a fractionation factor, expressed in per mille, of isotope ratio R^i between sites A and B).

Appendix B

Analytical technologies

There can be little question that the isotopic structures of organic molecules record a rich archive constraining sources, conditions and geochemical evolution. The question is, how can this archive be read? The technologies of measurements of molecular isotopic structure were recently reviewed by Eiler (2013); here we update this review and briefly summarize the technologies that have been applied to the study of petroleum and natural gas compounds.

(1) *Chemical degradation*: The oldest well-demonstrated approaches to measuring the site-specific isotopic compositions of organic molecules involve selective chemical attack, such as decarboxylation (Abelson & Hoering 1961), oxidative cleaving by ozonolysis followed by decarboxylation (Monson & Hayes 1980, 1982*a*, *b*), or release of methyl groups from methoxy-bearing compounds through reaction with iotic acid (Keppler *et al.* 2007; Feakins *et al.* 2013). We are aware of only one instance where such methods have been applied to a petroleum or natural gas compound: Gao *et al.* (2016) demonstrate that propane can be converted to propanol, then to acetone and finally to acetic acid. Combustion followed by IRMS of the acetic acid yields the average $\delta^{13}C$ of the terminal and central carbon sites, whereas combustion followed by IRMS of the original propane constrains the quantity: $(2 \times \delta^{13}C_{end} + \delta^{13}C_{centre})/3$. Thus, these two measurements can be used to solve for the difference, $\delta^{13}C_{centre} - \delta^{13}C_{end}$.

(2) *NMR:* NMR spectroscopy permits quantification of the abundances of nuclides having net nuclear spin (including H, D and ^{13}C), and in many instances discriminates between signals from non-equivalent molecular sites. Its most common use is characterizing structures of unknown organic molecules, but it has been used to observe the relative abundances of stable isotopes at natural abundances and per mille precisions since the early 1980s (Martin & Martin 1981; Caer *et al.* 1991). However, only in the last 10 years has NMR emerged as a broadly applied technique. These methods have been primarily applied to food science forensics and studies of natural products (Martin *et al.* 2006). Three studies use NMR to examine the carbon or hydrogen isotope structures of petroleum compounds (Gilbert *et al.* 2013, 2016*a*, *b*; Liu *et al.* 2015). A significant challenge of these techniques is that they generally require pure samples in approximately tens to hundreds of milligram quantities. For this reason, all work on hydrocarbons published or presented so far has examined the isotopic structures of commercial chemicals (and/or a few widely distributed reference standards), making it difficult to understand how the results might relate to natural geological materials. Nevertheless, NMR is currently the most productive and well-demonstrated technique for site-specific analysis of organic compounds larger than propane, and, even if other techniques are better suited to natural samples, it is likely to remain an essential tool for characterizing reference standards. It has not yet been shown that NMR can observe 'clumped' isotope species at their natural abundances and useful precision.

(3) *Pyrolytic degradation*: Pyrolysis of organic molecules can yield products that selectively sample specific molecular sites, such that a site-specific isotopic analysis can be performed by separating the products of pyrolysis (such as by GC) and then analysing them separately for an isotopic property of interest. Past efforts to do this have generally involved the passing of the products of pyrolysis through a

gas chromatographic column, followed by combustion and continuous flow IRMS of each eluted peak (Corso & Brenna 1997, 1999). The primary use of such methods is to analyse the CH_3 site of ethanol (which is sampled by CH_4 produced by pyrolysis). This technique has been applied to propane to discriminate between the $\delta^{13}C$ value of the terminal, methyl carbon site and the $\delta^{13}C$ of the bulk molecule, thus constraining the difference between the terminal and central positions (Gilbert *et al.* 2016*a*, *b*).

(4) *Mass spectrometry*: Mass spectrometric measurements of clumped-isotope compositions can be made simply by quantifying proportions of singly and multiply substituted ion beams, where the highest precisions can be obtained by simultaneous detection in a multi-collector sector mass spectrometer (Eiler 2007, 2013). It is also well established that site-specific isotopic variations can be constrained by mass spectrometric analysis of the isotopic compositions of two or more ion species that differ in their proportions of the sites of interest (e.g. the ^{15}N fractionation between α v. β nitrogen in N_2O can be distinguished by comparing the isotopic compositions of NO^+ and N_2O^+ ions; Yoshida & Toyoda 2000). Until recently, such measurements were restricted to simple C—N—O gases because of the presence of complex interferences in the mass spectra of other compounds. However, with the advent of high resolution isotope ratio mass spectrometry (Eiler *et al.* 2013), these measurements can be made on a variety of low-molecular-weight organic and inorganic molecular gases (Magyar *et al.* 2016; Piasecki *et al.* 2016*b*; Stolper *et al.* 2017). The recent exploration of stable isotope ratio analysis using the exceptionally high mass resolution of Fourier transform mass spectrometry is expanding this capability into diverse higher molecular weight organics (Eiler *et al.* 2017).

(5) *Infrared spectroscopy*: Infrared spectroscopy should be an ideal basis for site-specific and clumped-isotope analyses because the vibrational spectra of isotopologues are effectively unique, and sensitivity can be exceptional. Such techniques are demonstrated for site-specific ^{15}N analysis of N_2O (Waechter *et al.* 2008) and clumped-isotope analysis of methane (Ono *et al.* 2014). However, no infrared spectroscopic measurements have been made constraining isotopic structures of organic molecules more complex than methane, so no such data appear in this review. It is possible that spectroscopic measurements of clumped-isotope compositions of ethane and clumped- and position-specific measurements of propane could be developed, though it is unlikely they will ever be applied more broadly than this due to the complexity of vibrational spectra of larger molecules, and the difficulty of performing such measurements on species that are not room-temperature gases.

References

Abelson, P.H. & Hoering, T.C. 1961. Carbon isotope fractionation in formation of amino acids by photosynthetic organisms. *Proceedings of the National Academy of Sciences*, **47**, 623–632.

Anhauser, T., Greule, M., Zech, M., Kalbitz, K., McRoberts, C. & Keppler, F. 2015. Stable hydrogen and carbon isotope ratios of methoxyl groups during plant litter degradation. *Isotopes in Environmental and Health Studies*, **51**, 143–154.

Augusti, A. 2007. *Monitoring climate and plant physiology using deuterium isotopomers of carbohydrates*. PhD thesis, Umea University, Sweden.

Augusti, A., Betson, T.R. & Schleucher, J. 2008. Deriving correlated climate and physiological signals from deuterium isotopomers in tree rings. *Chemical Geology*, **252**, 1–8.

Berner, U., Faber, E., Scheeder, G. & Panten, D. 1995. Primary cracking of algal and landplant kerogens: kinetic models of isotope variations in methane, ethane and propane. *Chemical Geology*, **126**, 233–245.

Betson, T.R., Augusti, A. & Schleucher, J. 2006. Quantification of deuterium isotopomers of tree-ring cellulose using nuclear magnetic resonance. *Analytical Chemistry*, **78**, 8406–8411.

Bigeleisen, J. 1949. The relative reaction velocities of isotopic molecules. *Journal of Chemical Physics*, **17**, 675–678.

Billault, I., Guiet, S., Mabon, F. & Robins, R. 2001. Natural deuterium distribution in long-chain fatty acids is nonstatistical: a site-specific study by quantitative ^2H NMR spectroscopy. *ChemBioChem*, **2**, 425–431.

Burnham, A.K. & Sweeney, J.J. 1989. A chemical kinetic model of vitrinite maturation and reflectance. *Geochimica et Cosmochimica Acta*, **53**, 2649–2657.

Caer, V., Trierweiler, M., Martin, G.J. & Martin, M.L. 1991. Determination of site-specific carbon isotope ratios at natural abundance by carbon-13 nuclear magnetic resonance spectroscopy. *Analytical Chemistry*, **63**, 2306–2313.

Chacko, T., Cole, D.R. & Horita, J. 2001. Equilibrium oxygen, hydrogen and carbon isotope fractionation factors applicable to geologic systems. *In*: Valley, J.W. & Cole, D.R. (eds) *Stable Isotope Geochemistry*. Reviews in Mineralogy and Geochemistry, **43**. Mineralogical Society of America, Chantilly, VA, 1–81.

Chung, H.M., Gormly, J.R. & Squires, R.M. 1988. Origin of gaseous hydrocarbons in subsurface environments: theoretical considerations of carbon isotope distribution. *Chemical Geology*, **71**, 97–104.

Clog, M. & Eiler, J. 2014. C-H and C-C clumping in ethane by high-resolution mass spectrometry. Abstract presented at the 2014 Fall Meeting of the American Geophysical Union, San Francisco, USA.

Clog, M., Eiler, J., Guzzo, J.V.P., Moraes, E.T. & Souza, I.V.A. 2013. Doubly ^{13}C-substituted ethane. Abstract presented at the 2013 Goldschmidt Meeting, Florence, Italy. *Mineralogical Magazine*, **77**, 897.

Clog, M.D., Ferreira, A.A., Santos Neto, E.V., Eiler, J. M. 2014. Ethane C-C clumping in natural gas: a proxy for cracking processes? Abstract presented at the 2014 Fall Meeting of the American Geophysical Union, San Francisco, USA.

Corso, T.N. & Brenna, J.T. 1997. High-precision position-specific isotope analysis. *Proceedings of the National Academy of Sciences*, **94**, 1049–1053.

Corso, T.N. & Brenna, J.T. 1999. On-line pyrolysis of hydrocarbons coupled to high-precision carbon isotope ratio analysis. *Analytica Chimica Acta*, **397**, 217–224.

DeNiro, M.J. & Epstein, S. 1977. Mechanism of carbon isotope fractionation associated with the lipid synthesis. *Science*, **197**, 261–363.

Douglas, P., Stolper, D. *et al.* 2017. Methane clumped isotopes: progress and potential for a new isotopic tracer. *Organic Geochemistry*. First published online 16 August, 2017, https://doi.org/10.1016/j.orggeochem.2017.07.016

Duan, J.-R., Billault, I., Mabon, F. & Robins, R. 2002. Natural deuterium distribution in fatty acids isolated from peanut seed oil: a site-specific study by quantitative 2H NMR spectroscopy. *Chembiochem*, **3**, 752–759.

Dunn, S.R. & Valley, J.W. 1992. Calcite-graphite thermometry: a test for polymetamorphism in marble, Tudor gabbro, Ontario. *Journal of Metamorphic Geology*, **10**, 487–501.

Ehlers, I., Augusti, A., Betson, T.R., Nilsson, M.B., Marshall, J.D. & Schleucher, J. 2015. Detecting long-term metabolic shifts using isotopomers: CO_2-driven suppression of photorespiration in C-3 plants over the 20th century. *Proceedings of the National Academy of Sciences*, **112**, 15585–15590.

Eiler, J.M. 2007. 'Clumped-isotope' geochemistry – The study of naturally-occurring, multiply-substituted isotopologues. *Earth and Planetary Science Letters*, **262**, 309–327.

Eiler, J.M. 2011. Paleoclimate reconstruction using carbonate clumped isotope thermometry. *Quaternary Science Reviews*, **30**, 3575–3588.

Eiler, J.M. 2013. The isotopic anatomies of molecules and minerals. *Annual Reviews of Earth and Planetary Sciences*, **41**, 411–441.

Eiler, J.M., Clog, M. *et al.* 2013. A high-resolution gas-source isotope ratio mass spectrometer. *International Journal of Mass Spectrometry*, **335**, 45–56.

Eiler, J., Cesar, J. *et al.* 2017. Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry. *International Journal of Mass Spectrometry*, **422**, 26–142, https://doi.org/10.1016/j.ijms.2017.10.002

Elsner, M., Jochmann, M.A., Hofstetter, T.B., Hunkeler, D., Bernstein, A., Schmidt, T.C. & Schimmelmann, A. 2012. Current challenges in compound-specific stable isotope analysis of environmental organic contaminants. *Analytical and Bioanalytical Chemistry*, **403**, 2471–2491.

Epstein, S., Yapp, C.J. & Hall, J.H. 1976. Determination of D/H ratio of non-exchangeable hydrogen in cellulose extracted from aquatic and land plants. *Earth and Planetary Science Letters*, **30**, 241–251.

Feakins, S.J., Ellsworth, P.V. & Sternberg, L.da S.L. 2013. Lignin methoxyl hydrogen isotope ratios in a coastal ecosystem. *Geochimica et Cosmochimica Acta*, **121**, 54–66.

Ferreira, A.A., Santos Neto, E.V., Sessions, A.L., Schimmelmann, A. & Neto, F.R.A. 2012. $^{2}H/^{1}H$ ratio of hopanes, tricyclic and tetracyclic terpanes in oils and source rocks from the Potiguar Basin, Brazil. *Organic Geochemistry*, **51**, 13–16.

Freeman, K.H. 2001. Isotopic biogeochemistry of marine organic carbon. *In*: Valley, J.W. & Cole, D.R. (eds) *Stable Isotope Geochemistry*. Reviews in Mineralogy and Geochemistry, **43**. Mineralogical Society of America, Chantilly, VA, 579–605.

Galimov, E.M. 1973. *Izotopy ugleroda v heftegazovoy geologii* [Carbon Isotopes in Oil-Gas Geology]. NASA Technical Translation F-682.

Galimov, E.M. 1974. Organic geochemistry of carbon isotopes. *In*: Tissot, B. & Blenner, F. (eds) *Advances in Organic Geochemistry, 1973; Proceedings of the 6th International Meeting on Organic Geochemistry*. Éditions Technip, Paris, 439–452.

Galimov, E.M. 1985. *The Biological Fractionation of Isotopes*. Academic Press, London.

Galimov, E.M. & Shirinskii, V.G. 1975. Ordered distribution of carbon isotopes in individual compounds and components of lipid fraction of organisms. *Geokhimiya*, **4**, 503–528.

Gao, L., He, P., Jin, Y., Zhang, Y., Wang, X., Zhang, S. & Tang, Y. 2016. Determination of position-specific carbon isotope ratios in propane from hydrocarbon gas mixtures. *Chemical Geology*, **435**, 1–9.

Gilbert, A., Silvestre, V., Robins, R.J., Remaud, G.S. & Tcherkez, G. 2012*a*. Biochemical and physiological determinants of intramolecular isotope patterns in sucrose from C3, C4 and CAM plants accessed by isotopic ^{13}C NMR spectrometry: a viewpoint. *National Product Reports*, **29**, 476–486.

Gilbert, A., Robins, R.J., Remaud, G.S. & Tcherkez, G.G. B. 2012*b*. Intramolecular ^{13}C pattern in hexoses from autotrophic and heterotrophic C3 plant tissues. *Proceedings of the National Academy of Sciences*, **109**, 18204–18209.

Gilbert, A., Yamada, K. & Yoshida, N. 2013. Exploration of intramolecular ^{13}C isotope distribution in long chain n-alkanes (C_{11}–C_{31}) using isotopic ^{13}C NMR. *Organic Geochemistry*, **62**, 56–61.

Gilbert, A., Yamada, K., Suda, K., Ueno, Y. & Yoshida, N. 2016*a*. Measurement of position-specific ^{13}C isotopic composition of propane at the nanomole level. *Geochimica et Cosmochimica Acta*, **177**, 205–216.

Gilbert, A., Yamada, K. & Yoshida, N. 2016*b*. Evaluation of on-line pyrolysis coupled to isotope ratio mass spectrometry for the determination of position-specific ^{13}C isotope composition of short chain n-alkanes (C6-C12). *Talanta*, **153**, 158–162.

Guy, R.D., Fogel, M.L. & Berry, J.A. 1993. Photosynthetic fractionation of the stable isotopes of oxygen and carbon. *Plant Physiology*, **101**, 37–47.

Harrison, W.E. 1976. Laboratory graphitization of a modern estuarine kerogen. *Geochimica et Cosmochimica Acta*, **40**, 247–248.

Hayes, J.M. 2001. Fractionation of carbon and hydrogen isotopes in biosynthetic processes. *In*: Valley, J.W. & Cole, D.R. (eds) *Stable Isotope Geochemistry*. Reviews in Mineralogy and Geochemistry, **43**. Mineralogical Society of America, Chantilly, VA, 225–277.

Hedges, J.I., Eglington, G. *et al.* 2000. The molecularly-uncharacterized component of nonliving organic matter in natural environments. *Organic Geochemistry*, **31**, 945–958.

Horibe, Y. & Craig, H. 1995. D/H fractionation in the system methane-hydrogen-water. *Geochimica et Cosmochimica Acta*, **59**, 5209–5217.

Inagaki, F., Hinrichs, K.U. *et al.* 2015. Exploring deep microbial life in coal-bearing sediment down to *c.* 2.5 km below the ocean floor. *Science*, **349**, 420–424.

Ivlev, A.A., Loroleva, M.Y. & Kaloshin, A.G. 1974. Possible mechanisms of carbon isotope effect appearance in autotrophic organisms. *Doklady Akademii Nauk SSSR*, **217**, 224–227.

Jansco, G. & Van Hook, W.A. 1974. Condensed phase isotope effects (especially vapor pressure isotope effects). *Chemical Reviews*, **74**, 689–750.

Julien, M., Parinet, J., Nun, P., Bayle, K., Hohener, P., Robins, R.J. & Remaud, G.S. 2015. Fractionation in position-specific isotope composition during vaporization of environmental pollutants measured with isotope ratio monitoring by C-13 nuclear magnetic resonance spectrometry. *Environmental Pollution*, **205**, 299–306.

Keppler, F., Harper, D.B. *et al.* 2007. Stable hydrogen isotope ratios of lignin methoxyl groups as a paleoclimate proxy and constraint on the geographic origin of wood. *New Phytologist*, **176**, 600–609.

Kubicki, J.D., LaCroce, M.V. & Trout, C.C. 2016. H-D fractionation factors at individual sites on model petroleum compounds. Abstract presented at the 2016 San Diego Meeting of the American Chemical Society, April, San Diego, USA.

Lesot, P., Baillif, V. & Billault, I. 2008. Combined analysis of C-18 unsaturated fatty acids using natural abundance deuterium 2D NMR spectroscopy in chiral oriented solvents. *Analytical Chemistry*, **80**, 2963–2972.

Lewan, M.D. 1983. Effects of thermal maturation on stable organic carbon isotopes as determined by hydrous pyrolysis of Woodford Shale. *Geochimica et Cosmochimica Acta*, **47**, 1471–1479.

Lewan, M.D. 1985. Evaluation of petroleum generation by hydrous pyrolysis experimentation. *Philosophical Transactions of the Royal Society of London*, **315**, 123–134.

Lewan, M.D. 1993. Laboratory simulation of petroleum formation – hydrous pyrolysis. *In*: Engel, M.H. & Macko, S.A. (eds) *Organic Geochemistry*. Plenum Press, New York, 419–442.

Liu, C., McGovern, G.P. & Horita, J. 2015. Position-specific hydrogen and carbon isotope fractionations of light hydrocarbons by quantitative NMR. Abstract presented at the 2015 Fall Meeting of the American Geophysical Union, December, San Francisco, USA.

Lloyd, M., Sessions, A., Schimmelmann, A., Feakins, S. & Eiler, J. 2016. Determination of clumped ^{13}C-2H-H2 compositions of methoxyl groups in wood, lignin and simple organic monomers. Abstract presented at the 2016 Organic Geochemistry Gordon Conference, June, Tokyo.

Magyar, P.M., Orphan, V.J. & Eiler, J.M. 2016. Measurement of rare isotopologues of nitrous oxide by high-resolution multi-collector mass spectrometry. *Rapid Communications in Mass Spectrometry*, **30**, 1923–1940.

Mango, F. 1991. The stability of hydrocarbons under the time-temperature conditions of petroleum genesis. *Nature*, **352**, 146–148.

Mao, J., Fang, X., Lan, Y.Q., Schimmelmann, A., Mastalerz, M., Xu, L. & Schmidt-Rohr, K. 2010. Chemical and nanometer-scale structure of kerogen and its change during thermal maturation investigated by advanced solid-state ^{13}C NMR spectroscopy. *Geochimica et Cosmochimica Acta*, **74**, 2110–2127.

Markai, S., Marchand, P.A., Mabon, F., Baguet, E., Billault, I. & Robins, R.J. 2002. Natural deuterium distribution in branched-chain medium-length fatty acids is nonstatistical: a site-specific study by quantitative ^{2}H NMR spectroscopy of the fatty acids of capsaicinoids. *ChemBioChem*, **3**, 212–218.

Marshall, A.G. & Rodgers, R.P. 2008. Petroleomics: chemistry of the underworld. *Proceedings of the National Academy of Sciences*, **105**, 18090–18095.

Martin, G.J. & Martin, M.L. 1981. Isotopic labeling in natural abundance – application of high-resolution deuterium NMR to the study of vinyl compounds. *Comptes Rendus de L'Academie des Sciences Serie II*, **293**, 31–33.

Martin, G.J., Martin, M.L. & Remaud, G. 2006. SNIF-NMR – Part 3: From mechanistic affiliation to origin inference. *In*: Webb, G.A. (ed.) *Modern Magnetic Resonance*. Springer, 1669–1680.

Martini, A.M., Walter, L.M., Ku, T.C., Budai, J.M., McIntosh, J.C. & Schoell, M. 2003. Microbial production and modification of gases in sedimentary basins: a geochemical case study from a Devonian shale gas play, Michigan basin. *AAPG Bulletin*, **87**, 1355–1375.

Monson, K.D. & Hayes, J.M. 1980. Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Escherichia Coli: evidence regarding the coupling of fatty acid and phospholipid synthesis. *Journal of Biological Chemistry*, **255**, 11435–11441.

Monson, K.D. & Hayes, J.M. 1982*a*. Carbon isotopic fractionation in the biosynthesis of bacterial fatty acids. Ozonolysis of unsaturated fatty acids as a meansof determining the intramolecular distribution of carbon isotopes. *Geochimica et Cosmochimica Acta*, **46**, 139–149.

Monson, K.D. & Hayes, J.M. 1982*b*. Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Saccharomyces cerevisiae: isotopic fractionations in lipid synthesis as evidence for peroxisomal regulation. *Journal of Biological Chemistry*, **257**, 5568–5575.

Nabbefeld, B., Grice, K., Schimmelmann, A., Sauer, P.E., Bottcher, M.E. & Twitchett, R. 2010. Significance of $\delta D_{kerogen}$, $\delta^{13}C_{kerogen}$ and $\delta^{34}S_{pyrite}$ from several Permian/Triassic (P/Tr) sections. *Earth and Planetary Science Letters*, **295**, 21–29.

National Research Council 2003. *Oil in the Sea*. National Academies Press, Washington, DC.

Ohtomo, Y., Kakegawa, T., Ishida, A., Nagase, T. & Rosing, M.T. 2014. Evidence for biogenic graphite in early Archean Isua metasedimentary rocks. *Nature Geoscience*, **7**, 25–28.

O'Leary, M.H., Rife, J.E. & Slater, J.D. 1981. Kinetic and isotope effect studies of maize phosphoenolpyruvate carboxylase. *Biochemistry*, **20**, 7308–7314.

Ono, S., Wang, D.T. *et al.* 2014. Measurement of a doubly substituted methane isotopologue, $^{13}CH_3D$, by tunable

infrared laser direct absorption spectroscopy. *Analytical Chemistry*, **86**, 6487–6494.

Piasecki, A. 2015. *Site specific isotopes in small organic molecules*. PhD thesis, California Institute of Technology, Pasadena, CA, USA.

Piasecki, A., Sessions, A., Peterson, B. & Eiler, J. 2016*a*. Prediction of equilibrium distributions of isotopologues for methane, ethane and propane using density functional theory. *Geochimica et Cosmochimica Acta*, **190**, 1–12.

Piasecki, A., Sessions, A., Lawson, M., Ferreira, A.A., Neto, E.V.S. & Eiler, J.M. 2016*b*. Analysis of the site-specific carbon isotope composition of propane by gas source isotope ratio mass spectrometer. *Geochimica et Cosmochimica Acta*, **188**, 58–72.

Piasecki, A., Sessions, A. *et al.* 2018. Position-specific ^{13}C distributions within propane from experiments and natural gas samples. *Geochimica et Cosmochimica Acta*, **220**, 110–124. First published online October 6, 2017, https://doi.org/10.1016/j.gca.2017.09.042

Ponton, C., Xie, H. *et al.* 2016. Experiments constraining blocking temperatures of H isotope exchange in propane and ethane. Abstract presented at the 2016 Meeting of the International Clumped Isotope Workshop, January, Saint Petersburg USA.

Price, L.C. & Schoell, M. 1995. Constraints on the origins of hydrocarbon gas from compositions of gases at their site of origin. *Nature*, **378**, 368–371.

Prinzhofer, A.A. & Huc, A.Y. 1995. Genetic and post-genetic molecular and isotopic fractionations in natural gases. *Chemical Geology*, **126**, 281–290.

Quigley, T.M. & MacKenzie, A.S. 1988. The temperatures of oil and gas formation in the sub-surface. *Nature*, **333**, 549–552.

Rooney, M.A., Claypool, G.E. & Chung, H.M. 1995. Modeling thermogenic gas generation using carbon isotope ratios of natural gas hydrocarbons. *Chemical Geology*, **126**, 291–232.

Rossmann, A., Butzenlechner, M. & Schmidt, H.-L. 1991. Evidence for a nonstatistical carbon isotope distribution in natural glucose. *Plant Physiology*, **96**, 609–614.

Rustad, J.R. 2009. Ab initio calculation of the carbon isotope signatures of amino acids. *Organic Geochemistry*, **40**, 720–723.

Schimmelmann, A., Lewan, M.D. & Wintsch, R.P. 1999. D/H isotope ratios of kerogen, bitumen, oil, and water in hydrous pyrolysis of source rocks containing kerogen types I, II, IIS, and III. *Geochimica et Cosmochimica Acta*, **63**, 3751–3766.

Schimmelmann, A., Boudou, J.-P., Lewan, M.D. & Wintsch, R.P. 2001. Experimental controls on D/H and $^{13}C/^{12}C$ ratios of kerogen, bitumen and oil during hydrous pyrolysis. *Organic Geochemistry*, **32**, 1009–1018.

Schimmelmann, A., Sessions, A.L. & Mastalerz, M. 2006. Hydrogen isotopic (D/H) composition of organic matter during diagenesis and thermal maturation. *Annual Reviews of Earth and Planetary Sciences*, **34**, 501–533.

Seewald, J.S., Benitez-Nelson, B.C. & Whelan, J.K. 1998. Laboratory and theoretical constraints on the generation and composition of natural gas. *Geochimica et Cosmochimica Acta*, **62**, 1599–1617.

Stolper, D.A., Lawson, M., Formolo, M.J., Davis, C.L., Douglas, M.J. & Eiler, J.M. 2017. The utility of methane clumped isotopes to constrain the origins of methane in natural gas accumulations. *In*: Lawson, M., Formolo, M.J. & Eiler, J.M. (eds) *From Source to Seep: Geochemical Applications in Hydrocarbon Systems*. Geological Society, London, Special Publications, **468**. First published online December 14, 2017, https://doi.org/10.1144/SP468.3

Tang, Y., Perry, J.K., Jenden, P.D. & Schoell, M. 2000. Mathematical modeling of stable carbon isotope ratios in natural gases. *Geochimica et Cosmochimica Acta*, **64**, 2673–2687.

Tang, Y., Huang, Y. *et al.* 2005. A kinetic model for thermally induced hydrogen and carbon isotope fractionation of individual n-alkanes in crude oil. *Geochimica et Cosmochimica Acta*, **69**, 4505–4520.

Tissot, B.P. & Welte, D.H. 1984. *Petroleum Formation and Occurrence*. Springer Verlag, Berlin.

Toyoda, S., Yoshida, N. & Koba, K. 2015. Isotopocule analysis of biologically produced nitrous oxide in various environments. *Mass Spectrometry Reviews*, https://doi.org/10.1002/mas.21459

Valentine, D.L., Chidthaisong, A., Rice, A., Reeburgh, W. S. & Tyler, S.C. 2004. Carbon and hydrogen isotope fractionation by moderately thermophilic methanogens. *Geochimica et Cosmochimica Acta*, **68**, 1571–1590.

Van Zuilen, M., Lepland, A. & Arrhenius, G. 2002. Re-assessing the evidence for the earliest traces of life. *Nature*, **418**, 627–630.

Wada, H., Tomita, T., Matsuura, K., Iuchi, K., Ito, M. & Morikiyo, T. 1994. Graphitization of carbonaceous matter during metamorphism with references to carbonate and politic rocks of contact and regional metamorphisms, Japan. *Contributions to Mineralogy and Petrology*, **118**, 217–228.

Waechter, H., Mohn, J., Tuzson, B., Emmenegger, L. & Sigrist, M.W. 2008. Determination of N_2O isotopomers with quantum cascade laser based absorption spectroscopy. *Optics Express*, **16**, 9239–9244.

Wakeham, S.G., Lee, C., Hedges, J.I., Hernes, P.J. & Peterson, M.L. 1997. Molecular indicators of diagenetic status in marine organic matter. *Geochimica et Cosmochimica Acta*, **61**, 5363–5369.

Wang, Y. & Sessions, A.L. 2009*a*. Equilibrium $^{2}H/^{1}H$ fractionations in organic molecules. I. Experimental calibration of ab initio calculations. *Geochimica et Cosmochimica Acta*, **73**, 7060–7075.

Wang, Y. & Sessions, A.L. 2009*b*. Equilibrium $^{2}H/^{1}H$ fractionations in organic molecules. II. Linear alkanes, alkenes, ketones, carboxylic acids, esters, alcohols and ethers. *Geochimica et Cosmochimica Acta*, **73**, 7076–7086.

Wang, Z., Schauble, E.A. & Eiler, J.M. 2004. Equilibrium thermodynamics of multiply-substituted isotopologues of molecular gases. *Geochimica et Cosmochimica Acta*, **68**, 4779–4797.

Webb, M.A. & Miller, T.F. 2014. Position-specific and clumped stable isotope studies: comparison of the Urey and path-integral approaches for carbon dioxide, nitrous oxide, methane, and propane. *The Journal of Chemical Physics A*, **118**, 467–474.

Webb, M.A., Wang, Y., Braams, B.J., Bowman, J.M. & Miller II, T.F., III 2017. Equilibrium clumped-isotope

effects in doubly substituted isotopologues of ethane. *Geochimica et Cosmochimica Acta*, **197**, 14–26, https://doi.org/10.1016/j.gca.2016.10.001

Whiticar, M.J., Faber, E. & Schoell, M. 1986. Biogenic methane formation in marine and freshwater environments: CO_2 reduction vs acetate fermentation – Isotope evidence. *Geochimica et Cosmochimica Acta*, **50**, 693–709.

Yoshida, N. & Toyoda, S. 2000. Constraining the atmospheric N_2O budget from intramolecular site preference in N_2O isotopomers. *Nature*, **405**, 330–334.

Zhang, B.-L., Billault, I., Li, X., Mabon, F., Remaud, G. & Martin, M.L. 2002. Hydrogen isotopic profile in the characterization of sugars. Influence of the metabolic pathway. *Journal of Agricultural and Food Chemistry*, **50**, 1574–1580.