Taking the Pulse of Life:

Intramolecular and Clumped Isotopic Perspectives on the Origins and Evolution of Hydrocarbons in Geological and Prebiotic Systems

Thesis by

Guannan Dong

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2024

(Defended May 21, 2024)

Guannan Dong
ORCID: 0000-0002-6844-2492

ACKNOWLEDGEMENTS

This thesis and the whole grad school experience would not have been possible without so many people, each of whom influenced my PhD work and life in unique and valuable ways.

I'm grateful for my advisors and thesis committee: John Eiler, Alex Sessions, Jess Adkins, Woody Fischer, and George Rossman. John, thank you for being the best advisor I could have wished for, for your sharp insights and scientific rigor, for supporting and providing resources to pursue my goals, and for empowering me to grow and evolve as a scientist and as a human being. Also, thanks to you and Jane for introducing me to otters :). Alex, thank you for welcoming me into your lab and group and for being a tremendous organic geochemistry guru. Thanks for taking us to New Zealand! Jess, thank you for taking me under your wings in my first year and continuing to keep your door open. Thanks for being an inspiration with your passion and teaching me box models that turned out to be a very sharp tool in my toolkit. Thank you, Woody, for inspiring me with your breadth and depth of knowledge. Thank you, George, for your curiosity and for all the welcoming and motivating conversations. I'm also grateful for learning from Andy Thompson, Ken Farley, Jared Leadbetter, and Mike Brown in various courses, and Joe kirschvink and Paul Asimow on diverse trips of GPS.

I would like to thank GPS administrative and support staff Kacey Gibson, Julie Lee, Jen Shechet, Julia Zuckerman, Leticia Calderon, Ruth Martinez, Mark Garcia, Ken Ou and Ricardo Valentino whose efforts have enabled me to do what I do and travel to various places.

My academic family has been a cornerstone of my journey: Hao Xie, Sarah Zeichner, Renee Wang, Nithya Thiagarajan, Tim Csernica, Gabriella Weiss, Noam Lotem, Elliott Mueller, Elle Chimiak, Elise Wilkes, Alex Meshoulam, Alex Phillips, Josh Anadu, Amy Hofmann, Surjyendu Bhattacharjee, Xavier Mangenot, Camilo Ponton, Uri Ryb, Max Lloyd, Lubna Shawar, Simon Andren, Shae Silverman, Makayla Betts, Hannah Dion-Kirschner, Yeosoon Park, and Madison Dunitz. Science can feel challenging and lonely sometimes, thank you all for being great company and making this journey fun, as well as pushing each other to grow. I'm grateful for the magician lab managers and engineers who make things work in the lab: Nami Kitchen, Fenfang Wu, Jonathan Treffkorn, Michael Mathuri, Issaku Kohl, Brody Guckenberger and Randy Delaney.

I appreciate friends from various groups that kept me sane and brought me joy in grad school. The Pit and folks in my cohort: Renee Wang, Madison Douglas, Eva Scheller, Daniela Osorio, Will Palfey, Newton Nguyen, Emma Sosa, Freya Morris, Rui Cheng, Jackie Dowling, and Olivia Pardo. I couldn't imagine going through these years with anyone else. I also cherish the friendships within the GPS community: Fenfang Wu, Leah Sabbeth, Sijia Dong, Nicole Nie, Tony Wang, Sang Chen, Sharon Newman, John Naviaux, Ted Present, Dan Johnson, Danielle Monteverde Potocek, Ren Marquez, Haoyu Li, and Peter Martin. I appreciate the good times spent with the badminton crew: Pond Sirorattanakul, Shuo Cao,

Jerry Wang, Yuexin Li, Yijian Zhou, Guanli Wang, Ruolin Deng, and Hao Zhang. I also want to thank my buddies from USTC: Mi Lei, Weiting Deng, Xiang Li, Shan Zhong, Jiajun Du, Shuoyan Xiong, and Weiyi Liu. Grad school would have been much less colorful without you all. I'm grateful for longtime friends from back home who kept me grounded: Dinghe Cui, Yuexin Li, Yuning Wu, Tian Wang, and Shaoheng Li. A special shout-out to Yuexin for reuniting in this very division and being my roommate last year! And thanks go to roommates from all eras: Mi Lei and Julia Arnold.

I want to thank my collaborators for helping me achieve things that would not have happened without you: Tom McCollom, Alexandre Ferreira, Juliana Andrade Iemini, Kate French, Justin Birdwell, Mike Formolo, Michael Lawson, Yanhua Shuai, Aivo Lepland, Xinchu Wang and Jaime Cesar.

I owe a great deal of achievements to my previous advisors and mentors: professors at USTC—Fang Huang, Liping Qin, and Yan'an Shen—and my summer research mentors at MIT—Tanja Bosak and Ana De Santiago Torio. Thank you for showing me this world of doing cool stuff and pushing the boundaries of human knowledge. I also want to acknowledge those who offered hands-on training and who went out of their way to make those exposures and connections enjoyable: Fei Wu, Xingchao Zhang, Yingnan Zhang, Zhen Zeng, Kelsey Moore, David Wang, Mirna Daye, Haitao Shang, Xiaolei Liu, and Jeemin Rhim.

I'm fortunate to have a community of excellent colleagues who inspire me to do great science everyday: Daniel Stolper, Kate Freeman, Alexis Gilbert, Jochen Brocks, Ilya Bobrovskiy, Ann Pearson, Alexis Templeton, Karin Öberg, Barbara Sherwood Lollar, Caj Neubauer, Huiming Bao, Mang Lin, Will Leavitt, Jihua Hao, Nils Kuhlbusch, Min Song, Johan Weijers, Selva Marroquin, and Miquela Ingalls.

I'm grateful for the people that improved my life at Caltech and during grad school: cardio dance teacher Mariam Sufi, Laura Kim-Flower and Daniel Yoder at ISP, gym front desk Josh, guitar teacher Matthew Elgart, James Berk and other staff at the career center, and career coach Aditya Mahara.

Thanks to friends and colleagues from Caltech C and entrepreneur community: Ziran Zhou, Ruolin Deng, Ding Zhong, Stephanie Yanchinski, Bryan Gerber and Wenyuan Chen. I joined both groups in the later part of grad school. Thank you for trusting me with responsibilities to host events and develop resources for Caltech community interested in applying science and technology in industry and business. These experiences have broadened my horizon.

I also want to acknowledge people who came before me, some I have never met, others I work with closely on daily basis: Zhengrong Wang, Weifu Guo, Ying Wang, Hao Xie, Sang Chen, and Sijia Dong. They made this PhD possible through quantum entanglement, I think.

Lastly, my deepest gratitude goes to my family for their unconditional love and support. My parents, grandparents, cousin, and aunts have always been my steadfast pillars, nurturing my curiosity and drive with their exemplary values and support.

ABSTRACT

Life's origins and fate are tightly intertwined. All life as we know it is composed of proteins, carbohydrates, lipids, and nucleic acids. These biomolecules are synthesized today by cellular machinery made of the same components, leaving open questions about the origins of life and prebiotic chemistry. After death, organic remnants are buried in sediments, undergoing microbial reworking, consolidation, and transformation into kerogen. As temperature and pressure increase with depth, kerogen matures, releasing oil and gas before ultimately transforming into graphite. The question remains: can we decipher the traces of life (and non-life) from somewhat altered organic matter from the past or on other planets? This thesis explores the origins and evolution of one of the most fundamental classes of compounds—hydrocarbons—across geological and prebiotic settings through novel applications of intramolecular and clumped isotope analysis.

Chapter 2 delves into the evolution of isotopic signatures in methane, the simplest hydrocarbon, during the maturation process. By studying the clumped isotope effects of thermogenic methane formation through pyrolysis experiments, this chapter challenges previously held assumptions about abiotic and microbial signatures. The findings offer new opportunities to constrain the thermal maturation of sedimentary organic matter and have implications for the search for extraterrestrial life.

To facilitate high-precision measurements of hydrocarbon isotopic structures, Chapter 3 presents hardware and software developments enabling automated, high-throughput analysis using Orbitrap mass spectrometry. Chapter 4 then introduces a novel method coupling gas chromatography and Orbitrap MS to simultaneously measure intramolecular ^{13}C and ^{2}H distributions in *n*-alkanes, validating the technique for forensic fingerprinting and natural sample characterization.

Turning to the impact of thermal maturation, Chapter 5 examines how *n*-alkane intramolecular isotope patterns evolve through pyrolysis experiments. Kinetic isotope effects control residual *n*-alkane isotopic compositions, with minimal alteration to intramolecular distributions under the studied conditions, suggesting preservation of primitive signatures.

Chapter 6 brings together these analytical developments to compare intramolecular isotope compositions of *n*-nonadecane from sedimentary, abiotic, and biological sources. Distinctive isotopic fingerprints are established for each source, with implications for interpreting organic matter histories and detecting potential signatures of extraterrestrial life.

Collectively, this thesis expands the "molecular detective" toolkit for tracing hydrocarbon origins across diverse environments, from deep petroleum systems to potential prebiotic reaction pathways. The findings illuminate key processes governing isotopic biosignatures and their preservation through geological time.

PUBLISHED CONTENT AND CONTRIBUTIONS

Dong G., Xie H., Formolo M., Lawson M., Sessions A. and Eiler J. (2021) Clumped isotope effects of thermogenic methane formation: Insights from pyrolysis of hydrocarbons. *Geochimica et Cosmochimica Acta* **303**, 159–183. DOI: 10.1016/j.gca.2021.03.009

G.D. participated in the conceptualization, methodology, validation, investigation, data curation, formal analysis, modeling, visualization, writing, and project administration.

TABLE OF CONTENTS

Chapter 1 — An introduction

All life forms are composed of proteins, carbohydrates, lipids, and nucleic acids and these polymers are constructed from amino acids, sugars, glycerol, fatty acids, and nucleotides. These biomolecules are synthesized from carbon, hydrogen, oxygen, nitrogen, phosphate, and sulfur by intricate cellular machinery, which is itself composed of the same biomolecular building blocks. This presents a classic chicken-and-egg problem: How were the first biomolecules synthesized in the absence of the very cellular machinery that now produces them?

Conversely, after the cessation of life, organic remnants such as lignin, carbohydrates, proteins, and lipids are buried in sediments. Through microbial reworking and sediment consolidation, these biomolecules degrade, condense, and transform into a macromolecular organic structure known as kerogen during diagenesis. As this journey continues deeper into the sedimentary column, where temperatures range from 50–150°C and pressures from 300–1000 bars, kerogen undergoes ‘slow cooking’, excreting ‘organic sweat’ in the form of crude oil and natural gas during the catagenesis stage. Eventually, it is completely transformed into geologic char—graphite—after metagenesis. Over millions to billions of years, this process of maturation and transformation from diverse and intricate life forms to a more homogeneous existence raises a question: can a discerning observer identify these traces of previous life, understand who they were, where they came from, and what they went through?

1. Exploring hydrocarbon isotope systematics

This thesis is a tale of these two themes (Figure 1.1), utilizing analytical tools bestowed by recent innovation to explore the high-dimensional isotope content of hydrocarbons—a fundamental class of compounds at the heart of these questions (Figure 1.2). By employing high-resolution magnetic sector isotope ratio mass spectrometry and Orbitrap-based isotope ratio mass spectrometry, we can examine clumped and intramolecular isotope systematics,

providing a more nuanced view of isotopic distributions than traditional methods. These previously unattainable capabilities offer new insights into a broad range of topics.

2. Hydrocarbons: Pivotal, problematic, prevalent

Hydrocarbons have played a pivotal role in shaping the modern world, providing the energy that fuels our societies and the chemical feedstocks that form the basis of countless products. However, the widespread use of hydrocarbons has also led to global warming and environmental pollution, prompting urgent calls for a transition to clean, sustainable energy sources. As we navigate this critical juncture, it is essential to recognize the diverse origins and roles of hydrocarbons in the context of both Earth's history and the search for life beyond our planet.

Hydrocarbons can be synthesized abiotically through mechanisms relevant to prebiotic chemistry and the origins of life. They are ubiquitous in the cosmos, detected in meteorites, on extraterrestrial bodies such as Mars, Europa, and Ceres, and even in staggering quantities on Saturn’s moon Titan. Within living organisms, hydrocarbons and their biological precursors serve crucial structural and metabolic functions, while also recording physiological and environmental conditions. Sedimentary hydrocarbons are considered time capsules that experience and survive the maturation processes, potentially preserving echoes of life through the ages.

3. Thesis overview

This thesis embarks on a journey to explore the isotopic signatures of hydrocarbons at various stages of their existence, from their abiotic and biological synthesis to their evolution in the geological record.

Chapter 2 delves into the evolution of isotopic signatures in methane, the simplest hydrocarbon, during the maturation process. By studying the clumped isotope effects of thermogenic methane formation through pyrolysis experiments, this chapter challenges

previously held assumptions about abiotic and microbial signatures. The findings offer new opportunities to constrain the thermal maturation of sedimentary organic matter and have implications for the search for extraterrestrial life.

Chapter 3 focuses on the development of hardware and software to streamline the analysis of hydrocarbons using Orbitrap-based isotope ratio mass spectrometry. By automating the sample introduction and online separation system, this work demonstrates a significant improvement in sample throughput and data quality, paving the way for high-precision measurements of intramolecular isotopic structures.

Chapter 4 presents a novel method for simultaneously probing the intramolecular carbon and hydrogen isotope structures of n-alkanes using gas chromatograph (GC)-Orbitrap mass spectrometry. This chapter addresses the challenges and opportunities of investigating these simple yet fundamental compounds, showcasing the method's capabilities for forensic fingerprinting and natural sample characterization.

Chapter 5 investigates the effects of thermal maturation on the intramolecular carbon and hydrogen isotope structures of n-alkanes through anhydrous pyrolysis experiments. The results suggest that kinetic isotope effects control the isotopic composition of residual n-alkanes, with minimal alteration of intramolecular isotope distribution under the studied conditions. These findings have implications for the preservation of isotopic signatures in sedimentary organic matter.

Chapter 6 tackles the foci of the two themes by comparing the intramolecular carbon and hydrogen isotope compositions of *n*-nonadecane (nC_{19}) from sedimentary, abiotic, and biological sources. This chapter sheds light on the distinctive characteristics of each source and their potential implications for the search for extraterrestrial life and the history of sedimentary organic matter. We chose to study *n*-nonadecane out of all the *n*-alkanes based on three criteria: (1) an odd number of carbon atoms for investigating a potential contrast in the odd- and even-numbered positions in *n*-alkanes (the site-specific contrast in odd- and

even-numbered positions cancels out in *n*-alkanes with an even number of carbon atoms due to the symmetry of the molecule), (2) high abundance in abiotic synthesis experiments for sample availability (the abundance of *n*-alkane decrease with increasing carbon number in Fischer–Tropsch synthesis), (3) low volatility for ease of sample handling, (4) avoiding potentially common contaminants nC_{15} and nC_{17} alkanes derived from abundant biological fatty acids, palmitic acid and stearic acid. Although subsequent laboratory studies showed no evidence for more significant nC_{15} and nC_{17} alkane contamination than alkanes of other carbon numbers.

The findings presented in this thesis contribute to our understanding of the origins and evolution of life's building blocks and the complex history of Earth's biogeochemical cycles. By expanding our "molecular detective toolkit", this work opens new avenues for detecting and interpreting signatures in both ancient and extraterrestrial contexts. As we continue to unravel the mysteries of life's emergence and evolution, the study of hydrocarbons and their isotopic signatures will undoubtedly play a central role in this fascinating endeavor.

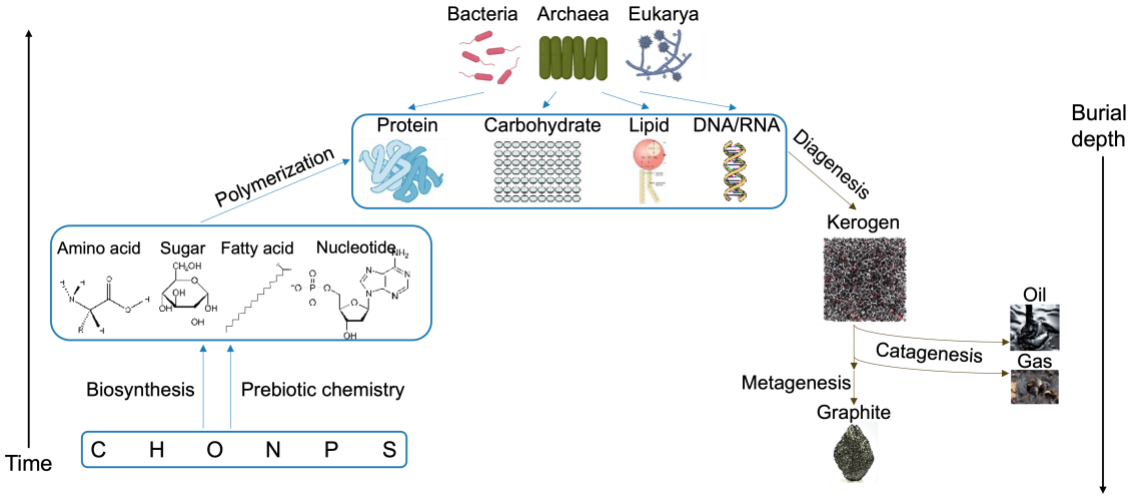


Figure 1.1 Illustration of the dual themes of this thesis: origins and records of life.

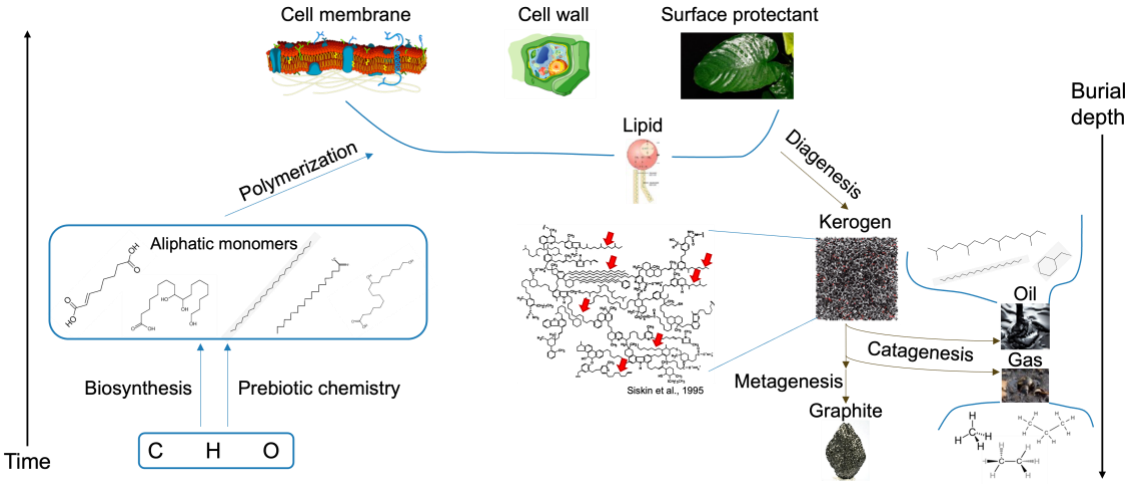


Figure 1.2 Aliphatic compounds are at the focal point of the origins and records of life.

Chapter 2 — Clumped isotope effects of thermogenic methane formation: Insights from pyrolysis of hydrocarbons

Dong G., Xie H., Formolo M., Lawson M., Sessions A. and Eiler J. (2021) Clumped isotope effects of thermogenic methane formation: Insights from pyrolysis of hydrocarbons. *Geochimica et Cosmochimica Acta* **303**, 159–183. DOI: 10.1016/j.gca.2021.03.009

Guannan Dong[a], Hao Xie[a], Michael Formolo[b], Michael Lawson[c], Alex Sessions[a], John Eiler[a]

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

[b] *ExxonMobil Upstream Integrated Solutions, Spring, TX 77389, USA*

[c] *ExxonMobil Upstream Business Development, Spring, TX 77389, USA*

Abstract

Methane clumped isotope analysis is a tool used to constrain the formation or equilibration temperatures of methane, or to differentiate methane of thermogenic, microbial or 'abiotic' origins. Geothermometry applications are based on the temperature dependence of relative abundances of multiply-substituted isotopologues in thermodynamic equilibrium, whereas assignments of biogenicity or 'abiogenicity' rely on kinetic isotope effects associated with synthesis, which disturb clumped isotope abundances away from expected equilibrium proportions. However, kinetic processes in thermogenesis or during post-generation storage of thermogenic gas may cause isotopic disequilibrium, confounding thermometry applications or leading to 'false positive' identifications of microbial or abiogenic gases. Non-equilibrated clumped isotope compositions have been observed in thermogenic gases including unconventional oil-associated gases and from coal pyrolysis experiments. The isotopic disequilibria might be caused by kinetic isotope effects expressed during gas migration (including extraction), or by irreversible chemical processes, such as breaking carbon-carbon bonds in an alkyl precursor. In this study, we performed controlled pyrolysis experiments at 400°C on n-octadecane ($C_{18}H_{38}$). We characterized the gas chemistry, and

compound-specific carbon and hydrogen isotope and methane clumped isotope compositions of the gas products. We found that $\Delta^{13}CH_3D$ values (anomalies relative to a stochastic distribution of isotopes) appear to be relatively close to equilibrium at the experimental temperature, whereas $\Delta^{12}CH_2D_2$ values are 30 to 40‰ lower than expected for equilibrium. The large deficit in $\Delta^{12}CH_2D_2$ can be explained by assembling hydrogen atoms affected by two distinct kinetic isotope effects into a methane molecule, previously referred to as a 'combinatorial effect'. We present a kinetic model that describes the full isotopic systematics, including anomalous $\Delta^{12}CH_2D_2$ deficits, of pyrolysis product methane. Finally, we propose a model for the isotope signatures of natural thermogenic methane where the non-equilibrium $\Delta^{12}CH_2D_2$ composition is a signature of the onset of catagenetic methane production. Our model also describes ways in which this signature disappears as further maturation drives $\Delta^{12}CH_2D_2$ to equilibrium through hydrogen exchange. Our findings demonstrate that anomalous depletion in $\Delta^{12}CH_2D_2$ is not a unique signature for microbial or putative abiotic methane, and specifically, it can be generated during pyrolytic chemistry.

1. Introduction

The multiply substituted, or 'clumped' isotopologues of methane include two that are analyzable at high precision, $^{13}CH_3D$ and $^{12}CH_2D_2$ (typically at errors of 0.3 and 1.0‰, 1SE, respectively), using high-resolution mass spectrometry or infrared absorption spectroscopy (D.A. Stolper et al., 2014; Ono et al., 2014; Young et al., 2017; Gonzalez et al., 2019; Eldridge et al., 2019a). When methane is in internal isotopic equilibrium, the proportions of these doubly substituted species must conform to equilibria of isotope exchange reactions:

$$^{12}CH_3D + {}^{13}CH_4 \rightleftharpoons {}^{12}CH_4 + {}^{13}CH_3D \quad \text{(R1)}$$

$$2\,{}^{12}CH_3D \rightleftharpoons {}^{12}CH_4 + {}^{12}CH_2D_2 \quad \text{(R2)}$$

The equilibrium constants for such reactions are functions of temperature; therefore, the relative abundances of non-, singly- and doubly-substituted isotopologues can constrain the

temperature of methane formation or last equilibration. This approach to methane geothermometry is particularly useful because it permits a temperature estimate using only the isotopic composition of methane, without knowledge of isotope compositions of co-existing substrates, e.g., water, molecular hydrogen or kerogen. Such clumped isotope temperature estimates can constrain temperatures of natural gas formation or storage, contribute to reconstructions of the thermal histories of basins containing petroleum-forming systems, and differentiate thermogenic versus microbial methane (based on their distinct formation temperatures). However, all such applications of clumped isotope analysis to quantitative geothermometry require that methane forms and retains internal isotopic thermodynamic equilibrium. If instead methane forms out of thermodynamic equilibrium or is driven from an equilibrium isotopic distribution by some post-formation process, proportions of clumped isotopologues could instead reflect chemical or physical kinetics, mixing, inheritance from a precursor, or so-called 'combinatorial' isotope effects (Yeung, 2016; Röckmann et al., 2016; Douglas et al., 2017; Stolper et al., 2018).

It has been argued that non-equilibrium methane clumped isotope abundances can be used to identify the mechanisms of microbial methanogenesis and methanotrophy, including the energetics and reversibility of metabolic reactions (Wang et al., 2015, 2016; Stolper et al., 2015; Gruen et al., 2018; Cao et al., 2019; Ash et al., 2019; Ono et al., 2021). Microbial methane can be produced by hydrogenotrophic and methylotrophic methanogenesis and consumed by methanotrophy through aerobic or anaerobic oxidation (Valentine et al., 2004; Valentine, 2011; Yoshinaga et al., 2014). These microbial metabolic processes can result in both equilibrated and non-equilibrated clumped isotope compositions in methane; both equilibrated and non-equilibrated compositions are common in natural environments, whereas all laboratory culture studies published to-date have produced non-equilibrium compositions (Wang et al., 2015; Stolper et al., 2015; Douglas et al., 2016; Young et al., 2017; Gruen et al., 2018; Giunta et al., 2019; Taenzer et al., 2020). Enzymatic reversibility has been proposed to control the expression of kinetic isotope effects during microbial methanogenesis. Specifically, the limitation of substrates in the deep subsurface and marine

settings promotes reversibility of the elementary steps in methanogenesis and leads to equilibrated proportions of clumped isotopes of methane; whereas non-equilibrium compositions are observed in freshwater environments and laboratory cultures grown with high activities of substrates that drive rapid net methane formation (Valentine et al., 2004; Wang et al., 2015; Stolper et al., 2015). It has also been suggested that methane can be driven to internal isotopic equilibrium in marine sediments by anaerobic oxidation (Ash et al., 2019). In that case, clumped isotope compositions record the temperature at which methane is oxidized rather than its initial production temperature (if different).

Abiotic methane can be produced by high-temperature reactions in the mantle or hydrothermal environments, as well as Fischer–Tropsch-Type (FTT) reactions (Etiope and Sherwood Lollar, 2013). Hydrothermal methane is thought to be produced in internal isotopic equilibrium in general, in which case its clumped isotope composition records its formation temperature (Wang et al., 2015, 2018; Douglas et al., 2017; Labidi et al., 2020); though some methane samples associated with hydrothermal vents are out of internal isotopic equilibrium and therefore suspected to be mixtures of microbial and abiotic gases (Young, 2019), or interpreted to be partially re-ordered at post-formation residence temperature (Labidi et al., 2020). In contrast, abiotic methane produced by laboratory Sabatier reaction and methane from crustal fluids that is interpreted to be abiotic and formed by FTT synthesis are both characterized by clumped isotope composition inconsistent with internal equilibrium; most notably, they exhibit pronounced deficits in $^{12}CH_2D_2$ compared to amounts expected for equilibrium at suspected gas generation temperatures, and in some cases compared to the stochastic distribution of all isotopes among all isotopologues (so-called ‘anti-clumping’ signatures). The non-equilibrated clumped isotope composition observed in the laboratory Sabatier reaction has been attributed to hydrogen quantum tunneling (Young et al., 2017), or ‘combinatorial’ isotope effect (Cao et al., 2019).

Thermogenic methane is generally recognized to form by catagenesis, or thermal ‘cracking’, of kerogen and other organic matter (Tissot and Welte, 1984). This process is generally

considered irreversible, involving isotopic fractionations controlled by kinetic isotope effects (KIE) (e.g., Chung et al., 1988; Tang et al., 2000; Ni et al., 2011). However, the clumped isotope compositions of most thermogenic methane seem to be consistent with thermodynamic equilibrium at independently-constrained formation temperatures (D. A. Stolper et al., 2014; Wang et al., 2015; Young et al., 2017; Douglas et al., 2017; Stolper et al., 2018; Giunta et al., 2019). This discrepancy could be taken as evidence for an alternative hypothesis that catagenetic reactions proceed through locally reversible metastable equilibria (Helgeson et al., 2009; Mango et al., 2009; Thiagarajan et al., 2020b). Alternatively, like interpretations of some natural microbial methane, thermogenic methane might be produced by irreversible, kinetically controlled reactions, but subsequently driven to equilibrium by some more reversible chemical process. Regardless, this contradiction between common views of catagenesis and the implications of methane clumped isotope analyses presents a quandary.

Some insight to this issue comes from exceptional natural thermogenic methanes with apparent clumped-isotope temperatures that significantly disagree with independent constraints on their conditions of formation. Unconventional oil-associated gases are formed and stored within the same geologic formation (i.e., shale gases), and are associated with liquid hydrocarbons (i.e., methane is dissolved in a liquid hydrocarbon phase). The clumped-isotope-based temperatures of these gases exceed the expected maximum temperatures experienced by the associated oils, and in many cases do not represent feasible geological temperatures at which methane could co-exist with a liquid hydrocarbon phase; therefore they appear to have formed or been driven out of isotopic equilibrium (Douglas et al., 2017; Stolper et al., 2018). The disequilibria observed in those gases could reflect a chemical or physical kinetic isotope effect associated with the extraction of hydrocarbons (Stolper et al., 2018). Specifically, the recovery technique of unconventional deposits is different from conventional reservoirs and requires hydraulic fracturing and production configurations that may impact the rate of gas desorption from the formation. Moreover, in the presence of liquid hydrocarbons, kinetic isotope effects could be expressed during dissolution/degassing of

methane in/out of oil, diffusion from shale through hydraulic fractures, and adsorption or desorption onto shale. These effects are currently poorly understood and need to be examined by studying natural samples with known production histories and laboratory experiments.

Alternatively, the disequilibria in these natural thermogenic methanes could reflect chemical kinetic isotope effects in methane generation (Douglas et al., 2017; Stolper et al., 2018). Some support for this possibility comes from coal pyrolysis experiments, which generate methane with clumped-isotope compositions that are inconsistent with the experimental temperatures under some conditions (Shuai et al., 2018). In these experiments, the transition from equilibrated to non-equilibrated distributions of methane clumped isotopes coincide with the onset of ethane cracking. This finding was interpreted as evidence that methane generated from kerogen catagenesis forms in local isotopic equilibrium, but that methane formed by kinetically controlled cracking of alkyl moieties (ethane in particular) initially forms out of isotopic equilibrium (Shuai et al., 2018). Xia and Gao (2019) constructed a kinetic reaction network model to investigate the effect of kinetic isotopic fractionations on methane clumped isotopes during catagenesis. They predicted that 'anti-clumping' (marked clumped isotope deficits) can result from a 'combinatorial effect' in methane generation. Specifically, they argue that a combination of both a D-rich hydrogen pool and a D-poor hydrogen pool, or a combination of hydrogen atoms that experienced different deuterium KIE's, to the four hydrogens in the methane molecule leads to deficits in $^{12}CH_2D_2$ compared to the stochastic abundance. Additionally, no study to-date has resolved the relative contributions of the two doubly substituted isotopologues ($^{13}CH_3D$ and $^{12}CH_2D_2$) to the observed disequilibria (in Δ_{18}, combined clumped isotope index) in unconventional oil-associated methane and methane product in coal pyrolysis experiments (Douglas et al., 2017; Stolper et al., 2018; Shuai et al., 2018).

Here we report results of controlled pyrolysis experiments on n-octadecane ($C_{18}H_{38}$), focusing on measurements of both the $^{13}CH_3D$ and $^{12}CH_2D_2$ proportions in product methane, as well as the carbon and hydrogen isotope compositions of C_1-C_4 hydrocarbon products.

We show that proportions of $^{13}CH_3D$ in product methane appear to be close to thermodynamic equilibrium at the temperatures of pyrolysis, or consistent with a random sampling of an apparently equilibrated or near equilibrated methyl group from the precursor(s), but that proportions of $^{12}CH_2D_2$ are strongly depleted relative to equilibrium at the experimental (or any) temperature. The distinctive depletion in $^{12}CH_2D_2$ associated with octadecane pyrolysis decreases in amplitude as pyrolysis proceeds, i.e., the composition of $^{12}CH_2D_2$ approaches equilibrium with maturation. We present a kinetic model to describe the behaviors of the five most abundant isotopologues of methane formed by octadecane pyrolysis. This model fits experimental results in all compositional properties we considered, and the fitted model parameters are consistent with independent constraints on isotope effects associated with carbon-carbon or carbon-hydrogen bond cleavage. The most important feature of the model is its explanation of the anomalous $^{12}CH_2D_2$ depletion as the result of a 'combinatorial effect' (Yeung, 2016; Röckmann et al., 2016; Xia and Gao, 2019), in this case attributable to the reaction of relatively deuterium-rich CH_3 radical intermediate with relatively deuterium-depleted capping hydrogen to form the product methane. The subsequent approach to equilibrium with respect to $^{12}CH_2D_2$ abundance can be explained by hydrogen exchange reactions in the system. These experiments demonstrate that exceptional depletion in $^{12}CH_2D_2$ is not a unique signature of biosynthesis or Sabatier reactions, but also can result from thermally-induced methane generation. We propose a model of natural thermogenic methane generation that predicts the value of $\Delta^{12}CH_2D_2$ as a function of thermal maturation.

2. Methods

2.1. Pyrolysis experiments

Aliquots of ~150 mg of n-octadecane (Alfa Aesar, LOT U28D017) were weighed and placed in quartz vessels. Each vessel is made of a bulb (1-inch outer diameter (OD), $1\frac{1}{4}$ inch long) joined to a 1/4-inch-OD, 7-inch-long tube. All parts of the vessel are made of 0.05-inch-thick

quartz. The vessels were evacuated to approximately 10^{-4} Torr using a mercury diffusion pump backed by a rotary pump, while cooled to approximately 77 K by immersion in a liquid nitrogen bath, and then flame-sealed. Then, the sealed quartz vessels were placed in a Mellen NACCI tube furnace at 400°C and heated for 72 to 397 hours (Table 2.1).

The equivalent maturity (Easy%RoV) of each experiment was calculated using the experimental times and temperature following Sweeney and Burnham (1990) and Burnham (2019). Easy%Ro values calculated from Sweeney and Burnham (1990) were also provided (Table 2.1).

Table 2.1 Pyrolysis experiment information (temperature 400°C)

experimental time (hours)	octadecane (mg)	Easy%RoV *
72	165.7	2.55
96	159.4	2.67
138	148.8	2.83
190	149.2	2.97
254	146.5	3.10
312	166.8	3.19
397	159.5	3.30

* Easy%Ro values for experiments from 72 to 397 hours are 1.75, 1.82. 1.92, 2.00, 2.09, 2.14, and 2.22, respectively.

2.2. Recovery of experimental products

After pyrolysis, vessels were removed from the furnace and quenched within ~1 min to room temperature in air. The vessels were then opened using a tube cracker attached to a vacuum line. The gas product was expanded into an evacuated glass volume and split into two aliquots by closing a valve in the middle of the volume. Methane in one aliquot was purified following the method described in D.A. Stolper et al. (2014). C_{2+} hydrocarbons (and a small amount of methane and hydrogen gas) in the other aliquot were condensed into a Pyrex®

tube by vapor transfer through a vacuum line, with the tube immersed in liquid nitrogen, and the tube was then flame-sealed for storage.

The amount of methane and the total gas product was measured by manometry and converted to molarity by the ideal gas law.

2.3. Compound-specific carbon and hydrogen isotope measurement

The $\delta^{13}C$ and δD of ethane, propane, propene, and n-butane of the recovered gas, and the starting substrate n-octadecane, were measured at Caltech by GC-combustion- isotope ratio mass spectrometry (IRMS) (for $\delta^{13}C$) and GC-pyrolysis-IRMS (for δD) on a Thermo-Scientific Delta^{+} XP. Gas products were introduced to a GS-GASPRO gas chromatography column (30m×0.320mm) through a programmed temperature vaporization (PTV) injector held at 160°C and operated in splitless mode, using He as a carrier gas flowing at 1.4 ml/min. The GC oven was held at 40°C for 5 min, and heated to 80°C, held for 3 min; then to 100°C, held for 4 min; and to 110°C, held for 9 min; and finally heated to 160°C, held for 8 min. All the ramping rates of temperatures were at 20°C/min. ~5 mg substrate n-octadecane was dissolved in 1.5 ml hexane then diluted by 60 times. The solution was injected to a ZB-5MS column (30m×0.25mm×0.25μm film thickness) at a flow rate of 1.4 mL/min. The oven was started at 100˚C, and heated from 100˚C to 220˚C at 20˚C/min then to 240˚C at 5˚C/min and held for 3min. The PTV injector was in PTV Large Volume mode. The injector was held at 60˚C for 0.1min, then heated to 350˚C at 14.5˚C/sec and held for 1min before cooling down to 300˚C at 1˚C/sec. Co-injected peaks of methane reference gas (δD = -161.5‰) were used for single-point isotopic calibration of analytes. That reference gas was in turn calibrated against a standard solution of eight fatty acid methyl esters ('F8 mix' of Arndt Schimmelmann, Indiana University). A working gas standard of C_2-C_5 gas mixture (including CIT-Ethane-1 and CIT-P1) is run along with the experimental samples to monitor instrumental status (Piasecki et al., 2016a; Clog et al., 2018; Xie et al., 2018).

$\delta^{13}C$ and δD are expressed using delta notation relative to Vienna Pee Dee Belemnite (VPDB) and Vienna Standard Mean Ocean Water (VSMOW), respectively:

$$\delta^{13}C = \frac{^{13}R_{sample}}{^{13}R_{VPDB}} - 1 \quad (1)$$

$$\delta D = \frac{^{D}R_{sample}}{^{D}R_{VSMOW}} - 1 \quad (2)$$

where $^{13}R = \frac{^{13}C}{^{12}C}$ and $^{D}R = \frac{D}{H}$; and $\delta^{13}C$ and δD values are reported in per mil notation. The long-term experimental reproducibility for measurements of standards and replicate samples on these instruments is 0.15‰ for $\delta^{13}C$ and 5‰ for δD (1SE).

The relative abundance of ethane, propane, and other hydrocarbons were calculated by the corresponding peak areas (reported by Isodat software of GC-combustion-IRMS) divided by the number of carbon atoms in each product molecule. A natural gas working standard with known relative concentrations for C_1-C_5 hydrocarbons (Table S1, characterized by Isotech Laboratories) is run to monitor relative sensitivities of compounds of interest.

The yield of each gas component was calculated as the mole ratio of carbon in that component to that of the carbon in starting material n-$C_{18}H_{38}$, e.g.:

$$yield_{methane} = \frac{molarity\ of\ carbon\ in\ the\ product\ methane}{molarity\ of\ carbon\ in\ the\ starting\ n\text{-}C_{18}H_{38}} \quad (3)$$

$$yield_{ethane} = \frac{molarity\ of\ carbon\ in\ the\ product\ ethane}{molarity\ of\ carbon\ in\ the\ starting\ n\text{-}C_{18}H_{38}} \quad (4)$$

2.4. Methane clumped isotope measurement

Values of $\delta^{13}C$ and δD and clumped-isotope compositions of methane were measured on a high-resolution isotope-ratio mass spectrometer, the Thermo Scientific Ultra, as described in

detail by Thiagarajan et al., (2020), see also D.A. Stolper et al., (2014). Clumped-isotope compositions are expressed using $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ (D.A. Stolper et al., 2014) where

$$\Delta^{13}CH_3D = \frac{^{13}CH_3D_R}{^{13}CH_3D_{R^*}} - 1 \quad (5)$$

$$\Delta^{12}CH_2D_2 = \frac{^{12}CH_2D_2R}{^{12}CH_2D_2R^*} - 1 \quad (6)$$

and

$$^{13}CH_3DR = \frac{[^{13}CH_3D]}{[^{12}CH_4]} \quad (7)$$

$$^{12}CH_2D_2R = \frac{[^{12}CH_2D_2]}{[^{12}CH_4]} \quad (8)$$

The brackets denote the molar concentration of an isotopologue as a fraction of all methane molecules. R* is the ratio expected for a random (stochastic) distribution of isotopes amongst all isotopologues, based on the sample's δD and $\delta^{13}C$ values. Values obtained using Equations (5), and (6) are reported in per mil (‰). Δ values are converted into apparent temperature (in Kelvin) via the equations (Eldridge et al., 2019a):

$$\Delta^{13}CH_3D = \frac{1.47348 \times 10^{19}}{T^7} - \frac{2.08648 \times 10^{17}}{T^6} + \frac{1.19810 \times 10^{15}}{T^5} - \frac{3.54757 \times 10^{12}}{T^4} + \frac{5.54476 \times 10^{9}}{T^3} - \frac{3.49294 \times 10^{6}}{T^2} + \frac{8.89370 \times 10^{2}}{T} \quad (9)$$

$$\Delta^{12}CH_2D_2 = -\frac{9.67634 \times 10^{15}}{T^6} + \frac{1.71917 \times 10^{14}}{T^5} - \frac{1.24819 \times 10^{12}}{T^4} + \frac{4.30283 \times 10^{9}}{T^3} - \frac{4.48660 \times 10^{6}}{T^2} + \frac{1.86258 \times 10^{3}}{T} \quad (10)$$

2.5. Numerical modeling

We integrated ordinary differential equations in the numerical models presented in the discussion section of this study using MATLAB ordinary differential equation solver ODE15s. We found the best fits of models to experimental results using MATLAB nonlinear least-squares solver lsqcurvefit with the trust-region-reflective algorithm.

3. Results

3.1. Gas chemistry

Methane, ethane, propane, propene, and n-butane are generated in all experiments (Figure 2.1, Table 2.2). Ethene and hydrogen gas are also produced but not abundant enough to determine isotopic compositions. Ethane is the most abundant recovered gaseous component in all experiments (35–41% of bulk gas by molarity), whereas propene and n-butane are always trace components (≤ 6% of bulk gas by molarity). The gas remains quite 'wet' (i.e., rich in C_{2+} hydrocarbons) at all experimental durations, with $C_1/(C_2+C_3)$ of only 0.3 to 0.425. In the experiment with the longest duration, about 4% of carbon in starting n-octadecane is transferred to methane, 13% is transferred to ethane, and 12% to propane.

Table 2.2 Time evolution of carbon-mole-ratio yield, δ13C and δD of gas products and clumped isotope compositions of methane

experimental time	carbon molar yield			$\delta^{13}C$ (‰)						$\delta^{2}H$ (‰)						clumped isotope compositions (‰)			
(hours)	CH_4	C_2H_6	C_3H_8	CH_4	C_2H_6	C_3H_8	C_3H_6	n-C_4H_{10}	$C_{18}H_{38}$	CH_4	C_2H_6	C_3H_8	C_3H_6	n-C_4H_{10}	$C_{18}H_{38}$	$\Delta^{13}CH_3D$	±	$\Delta^{12}CH_2D_2$	±
0									-32.9						-86				
72	0.013	0.050	0.051	-58.08	-42.8	-38.6	-32.7	-35.5		-309.2	-237	-161	-86	-129		1.4	0.9	-41.7	1.6
96	0.022	0.074	0.074	-55.59	-41.0	-37.1	-29.5	-33.4		-282.2	-215	-150	-64	-116		0.7	0.4	-37.7	1.5
138	0.028	0.078	0.077	-55.70	-41.3	-37.1	-28.0	-32.3		-283.8	-216	-147	-60	-113		1.6	0.4	-37.9	1.3
190	0.027	0.080	0.076	-55.71	-40.3	-36.0	-28.5	-33.3		-280.0	-200	-141	-46	-107		0.3	0.3	-35.8	1.3
254	0.036	0.110	0.099	-54.55	-41.4	-36.8	-28.8	-33.4		-260.1	-205	-142	-51	-107		1.9	0.4	-28.6	1.4
312	0.039	0.127	0.114	-53.61	-39.5	-35.6	-26.8	-31.2		-254.6	-196	-134	-41	-93		1.5	0.3	-33.7	1.2
397	0.042	0.128	0.117	-53.36	-39.1	-35.0	-26.0	-30.7		-250.3	-195	-132				1.5	0.3	-31.8	1.2

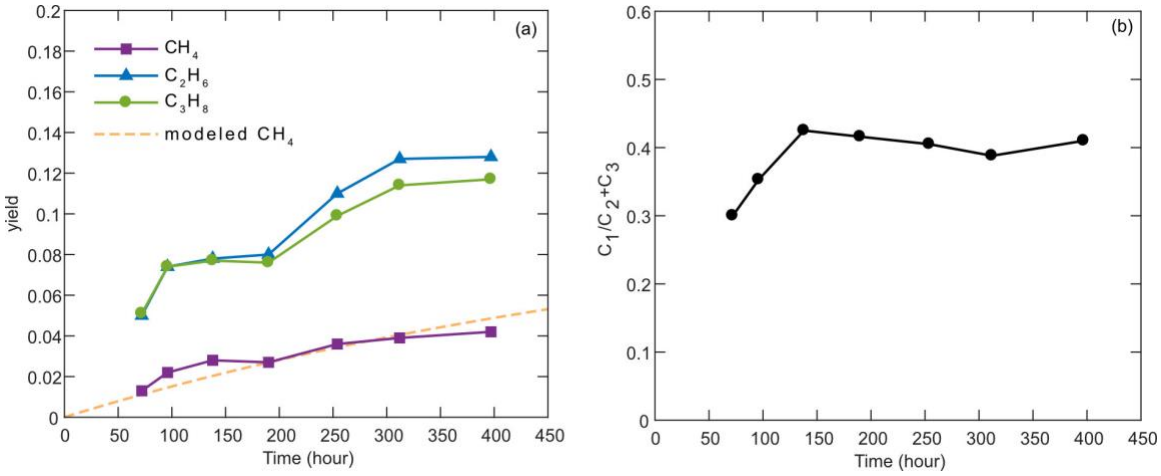


Figure 2.1 Time evolution of (a) yield and (b) dryness of gas products in pyrolysis of n-octadecane. The yield is expressed in carbon-mole-ratio, which is the molar ratio of carbon in the gas component (i.e., methane, ethane, or propane) divided by that in the starting n-octadecane. Methane yield curve modeled by the cracking and cracking-exchange model (yellow dashed line) fits the experimental data with R^2 of 0.73.

3.2. Compound-specific carbon and hydrogen isotope compositions of major gas products

The measured $\delta^{13}C$ and δD values of methane, ethane, propane, propene, and n-butane are given in Table 2.2 and Figure 2.2. $\delta^{13}C$ and δD of methane evolved from -58.1 to -53.4‰ and from -309.2 to -250.2‰, respectively, as experimental duration progressed from 72 hours to 397 hours (Figure 2.3a). Carbon and hydrogen isotope compositions of the other gas components all underwent similar increases with time throughout the experiments, maintaining an order of $\delta^{13}C$ and δD values: methane < ethane < propane < n-butane < propene.

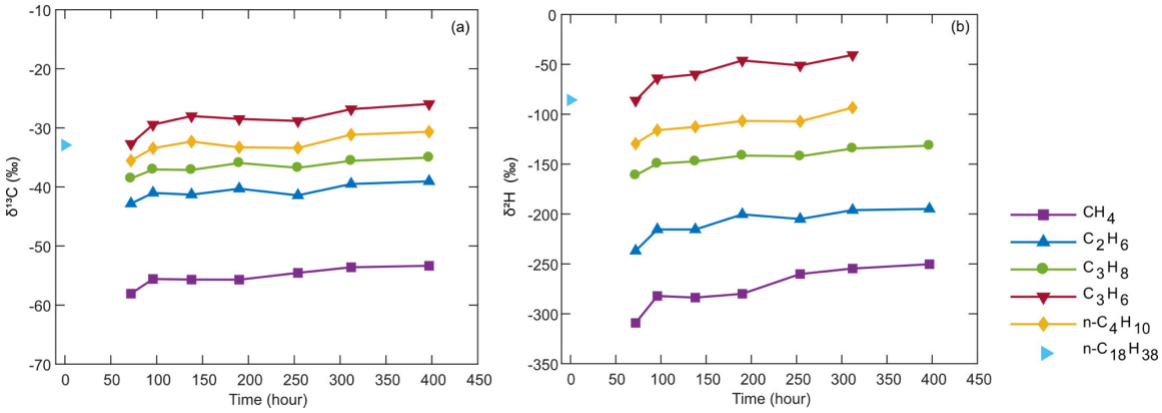


Figure 2.2 Time evolution of compound-specific (a) carbon and (b) hydrogen isotope compositions of gas products in pyrolysis of n-octadecane. Error bars are smaller than the symbols.

The carbon isotope fractionation we observe between product methane and substrate n-octadecane approximately agrees (within 5%) with the previous experimental studies (Sackett, 1978; Jeffrey, 1981) and theoretical prediction of the KIE's associated with formation of methane from substrates n-hexane or n-octane (Tang et al., 2000; Xiao, 2001; Ni et al., 2011) (Table 2.4). However, for hydrogen isotopes, the observed methane—n-octadecane fractionation also agrees with previous experiments (Sackett, 1978), but is more extreme than that predicted by Tang et al. (2005) and Ni et al. (2011) (Table 2.4). We discuss this discrepancy in section 4.2.

3.3. Carbon and hydrogen isotope composition of the starting substrate n-octadecane

The $\delta^{13}C$ of n-octadecane is -32.9‰ and δD is -86‰ (Table 2.2, Figure 2.2). Although we do not have clear information about the origin of the substrate from the supplier, this composition is consistent with that of petroleum materials (Schoell, 1984; Chung et al., 1992).

3.4. Clumped-isotope compositions of methane

The $\Delta^{13}CH_3D$ values of methane products are in the range of 0.2 to 1.9‰, which appear to be within 2–3 standard errors of the expected equilibrium composition of 1.2‰ at the experimental temperature (Table 2.2, Figure 2.3b). However, the $\Delta^{12}CH_2D_2$ values span from -41 to -28‰, much lower than the equilibrium value of 2.0‰ at the experimental temperature and more variable than can be explained by analytical uncertainties alone (propagated 1 standard error 1.2–1.6‰). The $\Delta^{12}CH_2D_2$ values become less negative as the pyrolysis proceeds and equivalent thermal maturity (Easy%RoV) increases (Figure 2.3b).

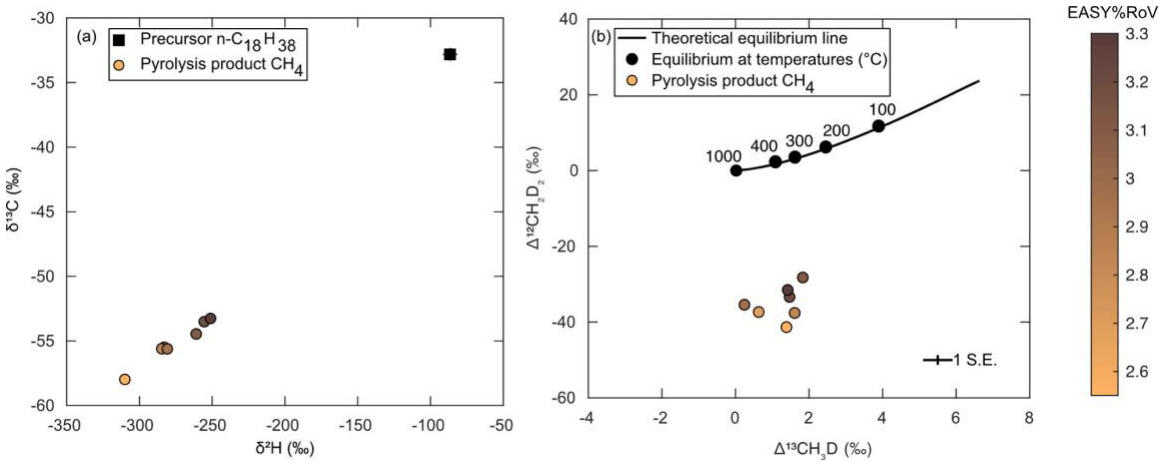


Figure 2.3. (a) carbon and hydrogen isotope and (b) clumped-isotope compositions of methane generated in pyrolysis of n-octadecane. The data points for pyrolysis products are color-coded with Easy%RoV (Burnham, 2019). In (a), the error bars of pyrolysis products are smaller than the marker. Both $\delta^{13}C$ and δD increase with maturation. In (b), the solid line shows the thermodynamic equilibrium $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ values from 2000 to 0°C (Eldridge et al., 2019a). The equilibrated value at labeled temperatures (°C) are marked as black dots. The $\Delta^{13}CH_3D$ values of methane products appear to be within 2 standard errors (2 SE) of the expected equilibrium composition, whereas the $\Delta^{12}CH_2D_2$ values are much more depleted than equilibrium and do not correspond to any possible temperature. Moreover, the $\Delta^{12}CH_2D_2$ values are progressively approaching the equilibrium value with maturation.

The $\Delta^{12}CH_2D_2$ values of methane produced in the pyrolysis of n-octadecane are among the lowest values observed in any prior study. Other comparable examples are abiotic methane

produced by the Sabatier reaction in the laboratory (-55‰), microbial methane produced by methanogen cultures (-55 to -12‰) and from the boreal lake (-40‰), and natural gases from Southwest Ontario (-20 to -14‰) and the Michigan Basin (-23 to -11‰) (Young et al., 2017; Giunta et al., 2019; Gonzalez et al., 2019; Young, 2019). Previous findings of strongly negative $\Delta^{12}CH_2D_2$ values have been interpreted as evidence for KIE's (possibly including hydrogen tunneling effects) associated with low-temperature irreversible reactions; it is unexpected, in the context of prior interpretations of such anomalies, that similar signatures would be associated with pyrolytic reactions occurring at 400°C.

4. Discussion

4.1. Kinetic isotope effects and compound-specific isotope compositions

The gas components produced from each experiment exhibit a negative linear correlation between $\delta^{13}C$ and 1/N, where N is the number of carbon atoms in each product molecule—the so-called 'Chung plot' (Figure 2.4, Table 3). These trends conform to a model that explains carbon isotope compositions of natural gas components as a result of kinetic isotope effects associated with breaking a single carbon-carbon bond in a straight side chain from a larger precursor organic molecule (Chung et al., 1988). If one interprets the linear correlations in Figure 2.4 in terms of this model, the vertical intercept is the $\delta^{13}C$ value of precursor organic molecule, whereas the slope represents the isotopic fractionation in bond breaking. However, the intercepts of trends defined by pyrolysis products of this study are higher in $\delta^{13}C$ than the starting substrate (Table 2.2, Table 2.3). It has been previously suggested that such discrepancies between the 'Chung diagram' intercepts and actual substrates reflect the fact that the precursor to gas components gets more enriched in heavy isotopes as reaction proceeds because of Rayleigh distillation (Rooney et al., 1995). This hypothesis could be supported by our observation that the 'Chung plot' intercepts generally increase as pyrolysis proceeds and methane yield increases. Alternatively, the evolution of these 'Chung plot' intercepts could be affected by secondary cracking of C_{2+} (Peterson et al., 2018). The slopes in Figure 2.4 do not change systematically throughout the experimental

period, presumably meaning the carbon isotopic fractionation associated with cracking reactions remains approximately constant at all experimental times. Additionally, the carbon isotope fractionations between C_1-C_4 and precursor octadecane are approximately consistent with theoretical predictions for homolytic cleavage on n-hexane and n-octane and β scission on n-octane (Tang et al., 2000; Xiao, 2001; Ni et al., 2011) (Table 2.4).

Table 2.3 Linear regression of 1/N and δ13C of product gas components, where N is the number of carbon atoms in each of the product molecule (Chung et al., 1988).

experimental time (hours)	slope (‰)	±	intercept (‰)	±	R^2
72	-29.8	0.6	-28.2	0.4	0.999
96	-28.9	1.1	-26.7	0.6	0.997
138	-29.9	2.0	-26.1	1.2	0.991
190	-29.8	0.6	-25.8	0.3	0.999
254	-27.6	1.1	-27.2	0.7	0.997
312	-28.8	1.8	-25.0	1.0	0.993
397	-29.2	1.7	-24.3	1.0	0.993

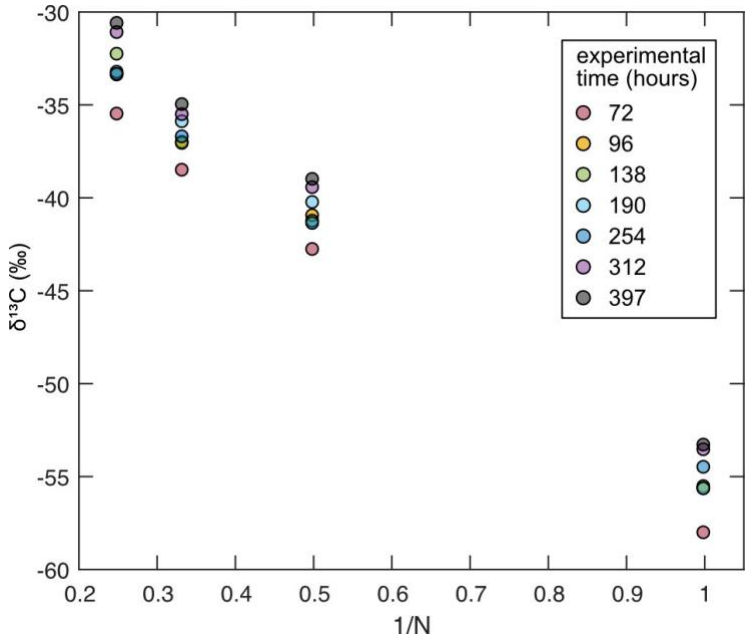


Figure 2.4 Chung plot for product gas components (methane, ethane, propane, and n-butane) in pyrolysis of n-octadecane (Chung et al., 1988). Components at each experimental time have a negative linear correlation between $\delta^{13}C$ and 1/N (linear regression data in Table 2.3), where N is the number of carbon atoms in each of the product molecules. The intercepts increase as pyrolysis proceeds (except for one outlier), whereas the slopes do not change systematically throughout the experimental period (Table 2.3).

Table 2.4 carbon and hydrogen isotope fractionations between gas products and precursors at 400°C

gas product	precursor	$^{13C}\alpha_{\text{gas product-precursor}}$	$^{D}\alpha_{\text{gas product-precursor}}$	data source
CH_4		0.9740	0.7556	a
	n-$C_{18}H_{38}$	0.9718	0.7629 or 0.7779	b
		0.9706	/	c
	n-C_6H_{14}	0.9931	/	d
	n-C_8H_{18}	0.9918	0.9156	e
		/	0.8423	f
		/	0.8432	g
		0.9897	/	h
	n-C_8H_{17}	0.9700	/	i
C_2H_6	n-$C_{18}H_{38}$	0.9897	0.8346	a
	n-C_8H_{18}	0.9866	0.9745	e
		/	0.9504	g
		0.9949	/	h
		0.9850	/	i
C_3H_8	n-$C_{18}H_{38}$	0.9942	0.9174	a
	n-C_8H_{18}	0.9901	0.9984	e
		/	1.0103	g
		0.9966	/	h
		0.9900	/	i
n-C_4H_{10}	n-$C_{18}H_{38}$	0.9973	0.9522	a
	n-C_8H_{18}	0.9933	0.9867	e
		0.9974	/	h
		0.9925	/	i

[a] 72 hour experiment in this study

[b] 17 hour experiment in Sackett (1978)

[c] 7 hour experiment in Jeffrey (1981)

[d] theoretical KIE of homolytical cleavage, Tang et al. (2000)

[e] theoretical KIE of homolytical cleavage, Ni et al. (2011), energies of two forming radicals were used to represent the transition barrier in calculations.

[f] theoretical KIE of homolytical cleavage, Tang et al. (2005), quantum chemistry density functional theory calculations were used to locate the transition state.

[g] theoretical KIE of homolytical cleavage, Ni et al. (2011), quantum chemistry density functional theory calculations were used to locate the transition state.

[h] theoretical KIE of homolytical cleavage, Xiao et al. (2001)

[i] theoretical KIE of beta scission, Xiao et al. (2001)

Notably, the $\delta^{13}C$ of propene does not follow the negative linear correlation defined by most C_1 to C_4 species in Figure 2.4 (see also Table 2.2). Instead, propene $\delta^{13}C$ is higher than

expected by such observed correlation and even higher than that of the starting n-octadecane. We speculate that the distinctive ^{13}C enrichment of propene reflects an inverse KIE (i.e., a faster net reaction for heavy isotopologues) that results from β-scission of a radical precursor to yield an alkene, for example,

$$CH_3CH_2CH_2\dot{C}HCH_3 \Leftrightarrow CH_3\dot{C}H_2 + CH_2{=}CHCH_3 \quad \text{(R3)}$$

where a new double bond is formed to the central carbon in propene. Reaction (R3) is predicted to have a product-like transition state, i.e., the transition state is late and has some of the character of the double bond in product propene (Jitariu et al., 2003). Therefore, the carbon destined for the central site of propene will experience a more confined bonding environment in the transition state than in the reactant. Transition state theory predicts the kinetic isotope effect of reaction (R3) to be the equilibrium isotope fractionation between the transition state and the reactant, which leads to an inverse kinetic isotope effect (Bigeleisen and Wolfsberg, 1957). Similarly, we note that propene also has a higher δD than the starting n-octadecane. We suspect the deuterium enrichment in propene sources from its radical precursor, where the radical precursor experiences hydrogen isotope fractionation in isomerization, for example,

$$CH_3CH_2CH_2CH_2\dot{C}H_2 \Leftrightarrow CH_3CH_2CH_2\dot{C}HCH_3 \quad \text{(R4)}$$

where an inverse secondary KIE may result from the change of orbital hybridization on the terminal carbon from sp^2 to sp^3 (e.g., Singleton and Thomas, 1995). Alternatively, secondary cracking of propene and its conversion to graphitic products may contribute to the enrichment of ^{13}C and D in the residual propene. Additionally, alkene and alkane can attain metastable equilibrium in their relative abundance in hundreds of hours in the presence of H_2 under hydrothermal conditions (Seewald, 1994). If propene and propane are also in isotopic equilibrium, we might expect higher $\delta^{13}C$ and δD of propene than propane in the experiments (Table 2.2). We are not aware of any first-principles theoretical studies of the carbon or hydrogen isotopic fractionations associated with propene generation by β-scission or

equilibrium fractionation between propene and propane, but such studies would be obvious ways to further explore these hypotheses.

The hydrogen isotope fractionation between wet gases (C_2 to C_4) and precursor octadecane has an order of ethane > propane > n-butane. Following the logic of common interpretations of the Chung diagram (above), this order could reflect the proportions of hydrogen that was isotopically fractionated by secondary KIE's in carbon-carbon bond cleavage as a fraction of all hydrogen in the product molecule, i.e., relatively more hydrogen sites in ethane are fractionated than those in propane, than those in n-butane. However, the hydrogen isotope fractionations between wet gases and source substrate octadecane are 30–150‰ more negative than those predicted by Ni et al. (2011) (Table 4). We suggest this discrepancy could have the same origin(s) as the discrepancy between predicted and experimental results of methane-octadecane fractionations, as stated in 3.2. We return to this issue in more detail in 4.2.1 and 4.2.2.4.

The chemical and isotopic compositions of C_2 to C_4 gases in our experiments offer insights into the reactions by which they formed and are relevant to the interpretation of the isotopic composition of product methane, which we develop in detail below. However, the depth of analysis required to fully address the findings for wet gas compounds goes beyond the scope of this paper and will be addressed in a future paper.

4.2. Modeling the isotopic content of methane in pyrolysis of n-octadecane

We propose a model of the reactions and accompanying isotopic fractionations associated with methane formation by alkane cracking as an aid to understanding and quantifying the isotopic content of methane in the pyrolysis experiments of n-octadecane. Any such model should address the following features of our experimental results: values of $\Delta^{12}CH_2D_2$ of product methane that reach negative tens of per mil; ~ -200‰ fractionation between δD of product methane and precursor n-octadecane; and a trend of increasing $\delta^{13}C$, δD, and $\Delta^{12}CH_2D_2$ of methane with time and progress of pyrolysis reactions (Figure 2.3). We present

two separate models that attempt to quantitatively explain these data. Both can explain the depleted $\Delta^{12}CH_2D_2$ and δD values of product methane, but the first—a model of cracking reactions alone — is incapable of describing the rise of $\Delta^{12}CH_2D_2$ with time and the rate of increase in δD with time for methane, whereas the second—a coupled model of cracking reactions and hydrogen exchange of product methane—captures those features. We present the cracking model first despite its inadequacy to fully fit the experimental results because its discussion demonstrates the kinetic isotope effects in methane generation and the 'combinatorial effect', which are also key components of our more successful and complex second model. We then add hydrogen exchange reactions to the cracking model, thereby explaining all major features of the experimental data.

4.2.1. Cracking model

4.2.1.1. methane formation and 'capping' hydrogen

Pyrolytic chemistry, or 'cracking', of n-alkanes involves different types of elementary reactions. Cracking can initiate with a family of homolytic cleavage reactions on different carbon-carbon bonds of n-alkanes. Various alkyl radical intermediates can (1) go through further cracking through β scission reactions, producing alkene and alkyl radical, (2) rearrange the position of the single electron in the radical structure by alkyl isomerization, or (3) transform to a neutral, stable product by hydrogen abstraction. The stable product can experience further cracking in the same manner as the starting n-alkane. Such a reaction network involves alkanes, alkenes, and alkyls of various carbon numbers. Therefore, a chemical kinetic model of pyrolysis often incorporates hundreds of chemical species and thousands of reactions (e.g., Yi et al., 2011; Wang et al., 2017). Alternatively, previous studies have developed kinetic models for oil cracking and natural gas generation with overall reactions from stable precursors to stable products and/or using generalized compound classes to represent complex mixtures (Ungerer et al., 1988; Jackśon et al., 1995; Tang et al., 2000).

Methane formation in pyrolysis of n-alkane involves breaking a carbon-carbon bond through homolytic cleavage or β-scission of hydrocarbon precursors, creating a methyl radical as an immediate product (Figure 2.5). This radical is metastable and short-lived, transforming to a neutral, stable methane molecule by abstracting a hydrogen radical that 'caps' the single electron in the methyl-radical intermediate (Figure 2.5). The KIEs in such a process of methane generation include a primary ^{13}C KIE and a secondary deuterium KIE in breaking the carbon-carbon bond, and a primary deuterium KIE in breaking a carbon-hydrogen bond when abstracting the 'capping' hydrogen. A primary deuterium KIE in hydrogen abstraction by a methyl radical can be as large as ~ -700‰ at 400°C (Gray et al., 1971). However, isotope effects associated with extracting the capping hydrogen from its precursor(s) have long been ignored in theoretical studies on compound-specific hydrogen isotope compositions of natural gas (e.g., Ni et al., 2011), until recently Xia and Gao (2019) showed such a primary deuterium KIE may affect clumped isotope composition of methane.

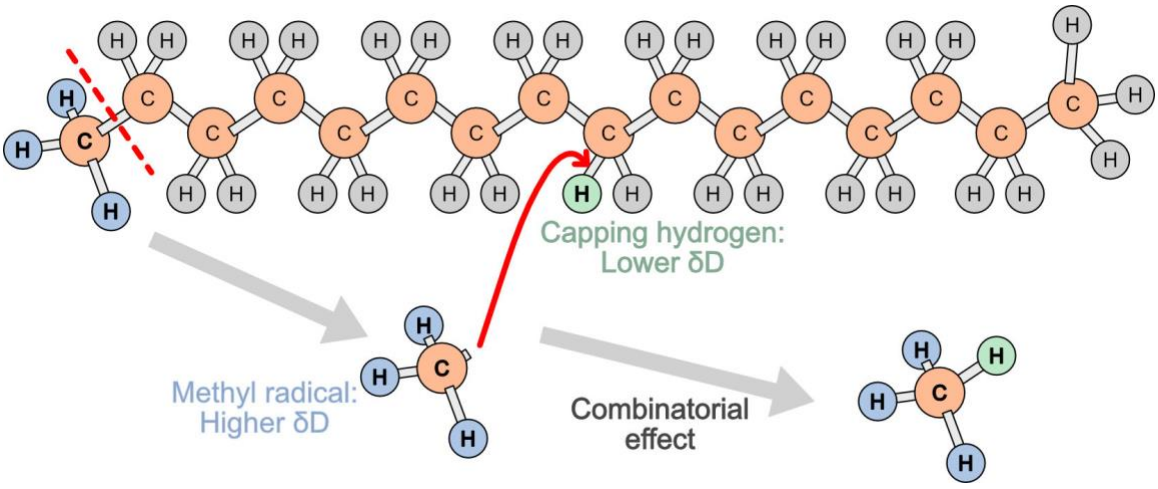


Figure 2.5 Cartoon representation of methane generation in pyrolysis of n-octadecane. Methane generation starts with producing a methyl radical by breaking a carbon-carbon bond through homolytic cleavage or β scission. Then the methyl radical attacks co-existing hydrogen bearing compounds (i.e., hydrocarbons including precursor n-octadecane) and abstracts the capping hydrogen. The three hydrogen atoms from the methyl precursor in each methane molecule constitute a pool with higher deuterium concentration, whereas the capping hydrogen makes up a second pool with lower deuterium concentration. Nevertheless, the four hydrogen atoms are symmetrically equivalent in a product methane molecule. This symmetric equivalence

leads one to predict a stochastic abundance of $^{12}CH_2D_2$ that is much higher than the true probability of randomly forming $^{12}CH_2D_2$—this phenomenon is the 'combinatorial effect'. When four hydrogens in methane are sourced from isotopically different pools or fractionated by distinct isotope effects, there will be a combinatorial effect, causing a deficit in $\Delta^{12}CH_2D_2$ but not affecting $\Delta^{13}CH_3D$.

4.2.1.2. Cracking model setup

In order to test the effect of capping hydrogen on the isotopic content of methane produced in pyrolysis experiments, we consider the following methane generation reaction in n-octadecane pyrolysis:

$$A\text{-}CH_3 + B\text{-}H \rightarrow A\cdot + B\cdot + CH_4 \quad (R5)$$

where $A\text{-}CH_3$ represents methyl precursors and B-H represents hydrogen bearing compounds (Xia and Gao, 2019). The reaction rate of (R3) is:

$$\frac{d[CH_4]}{dt} = k_{CH4Gen}[ACH_3][BH] \quad (11)$$

where [i] represents the concentration of species i and k_{CH4Gen} is the rate constant of (R5). We model the generation of the five most abundant isotopologues of methane numerically with 14 isotopic species and 8 isotope-specific reactions (RA1-RA8 in Table A1). We incorporate the primary ^{13}C KIE for $A\text{-}^{13}CH_3$ ($^{13C}KIE_{PRI}$), secondary deuterium KIE for $A\text{-}^{12}CH_2D$ ($^{D}KIE_{SEC}$), and primary deuterium KIE for B-D ($^{D}KIE_{PRI}$) on the rate constants of these reactions (Table A1, Appendix A1). We assume the kinetic isotope effect for multiply-substituted species is the product of the KIE's for the corresponding singly-substituted species (as would occur in the case of no clumped-isotope-specific kinetic isotope effects). This approximation is reasonable because clumped-isotope-specific chemical kinetic vibrational isotope effects in this system at 400°C are expected to be much smaller than the $^{12}CH_2D_2$ anomaly we wish to explain. All concentrations in the model are normalized by moles of starting n-octadecane in the experiments. The initial concentration of $A\text{-}CH_3$ is

constrained by the rate of increase in methane $\delta^{13}C$ (i.e., the slope of Figure 2.6a is determined by the size of the A-CH_3 pool that experiences Rayleigh distillation), and is consistent with the number of methyl groups in n-$C_{18}H_{38}$. The initial concentration of B-H is constrained by the number of hydrogen atoms in n-$C_{18}H_{38}$ excepting methyl groups (Table 5). The rate constant for $^{12}CH_4$ generation (k_{CH4Gen}) is fitted to the time evolution of the yield of methane (Figure 2.1a) with methods described in section 2.5. We use the measured $\delta^{13}C$ and δD values of n-$C_{18}H_{38}$ to set the D/H and $^{13}C/^{12}C$ ratios of A-CH_3 and B-H. And we fit $^{13C}KIE_{PRI}$, $^{D}KIE_{PRI}$, $^{D}KIE_{SEC}$ and initial clumped isotope compositions of the methyl group in A-CH_3 ($\Delta^{13}CH_2D$ and $\Delta^{12}CHD_2$) to the measured time evolutions in isotopic properties of methane using methods detailed in section 2.5 (Table 5).

Table 2.5 Pyrolysis model parameters and data sources. Fitted parameters are compared with data from computational and experimental studies.

parameter	cracking model	cracking-exchange model	data source	comparison	reference
precursor concentrations (dimensionless)					
ACH_3	2	2	fitting	2	number of methyl groups in n-C18H38
BH	32	32	number of H in n-$C_{18}H_{38}$ except precursor methyl		
kinetic parameters (s^{-1})					
k_{CH4Gen}	1.29E-08	1.29E-08	fitting		
k_{CH4-BH}	/	2.96E-07	fitting		
$k_{CH4-CH4}$	/	8.86E-08	fitting		
precursor isotope compositions (‰)					
$\delta^{13}C_{VPDB}$ (ACH_3)		-32.9	measurement		
δD_{SMOW} (ACH_3)		-85.7	measurement		
δD_{SMOW} (BH)					
$\Delta^{13}CH_2D$ (ACH_3)	1.6572	3.0195	fitting		
$\Delta^{12}CHD_2$ (ACH_3)	0.0000	0.0716	fitting		
Isotope effects					
$^{13C}KIE_{PRI}$	0.9732	0.9732	fitting	0.9931	Tang et al. (2000)[a]
				0.9918	Ni et al (2011)[a]
				0.9897	Xiao et al. (2001)[a]
				0.9700	Xiao et al. (2001)[b]
$^{D}KIE_{SEC}$	0.9278	0.9011	fitting	0.9156	Ni et al (2011)[ac]
				0.8432	Ni et al (2011)[ad]
				0.8423	Tang et al. (2005)[ad]
$^{D}KIE_{PRI}$	0.3567	0.2990	fitting	0.2603-0.3449	Gray et al. (1971)[e]
$^{13C}KIE_{PRI}{}^{C-H}$	/	0.9732	fitting	>0.7348	Gray et al. (1971)[f]
$^{D}KIE_{SEC}{}^{C-H}$	/	0.8037	fitting	0.9572-0.9863	Lu et al. (1971)[h]
				0.4722	Gray et al. (1990)[g]
$^{D}\alpha^{eq}_{CH4-BH}$	/	0.8859	fitting	0.8676-0.8810	Wang et al. (2009)[i]
				0.9752-0.9921	Piasecki et al. (2016)[j]

[a] homolytic cleavage

[b] β scission

[c] energies of two forming radicals were used to represent the transition barrier in calculations

[d] quantum chemistry density functional theory calculations were used to locate the transition state

[e] hydrogen abstraction from methyl and methylene groups in ethane, propane, n-butane and toluene

[f] hydrogen abstraction from $^{14}CH_4$ by methyl radical

[g] hydrogen abstraction from methane by methyl radical

[h] hydrogen abstraction from methane by hydrogen radical

[i] BH as n-undecane

[j] BH as propane

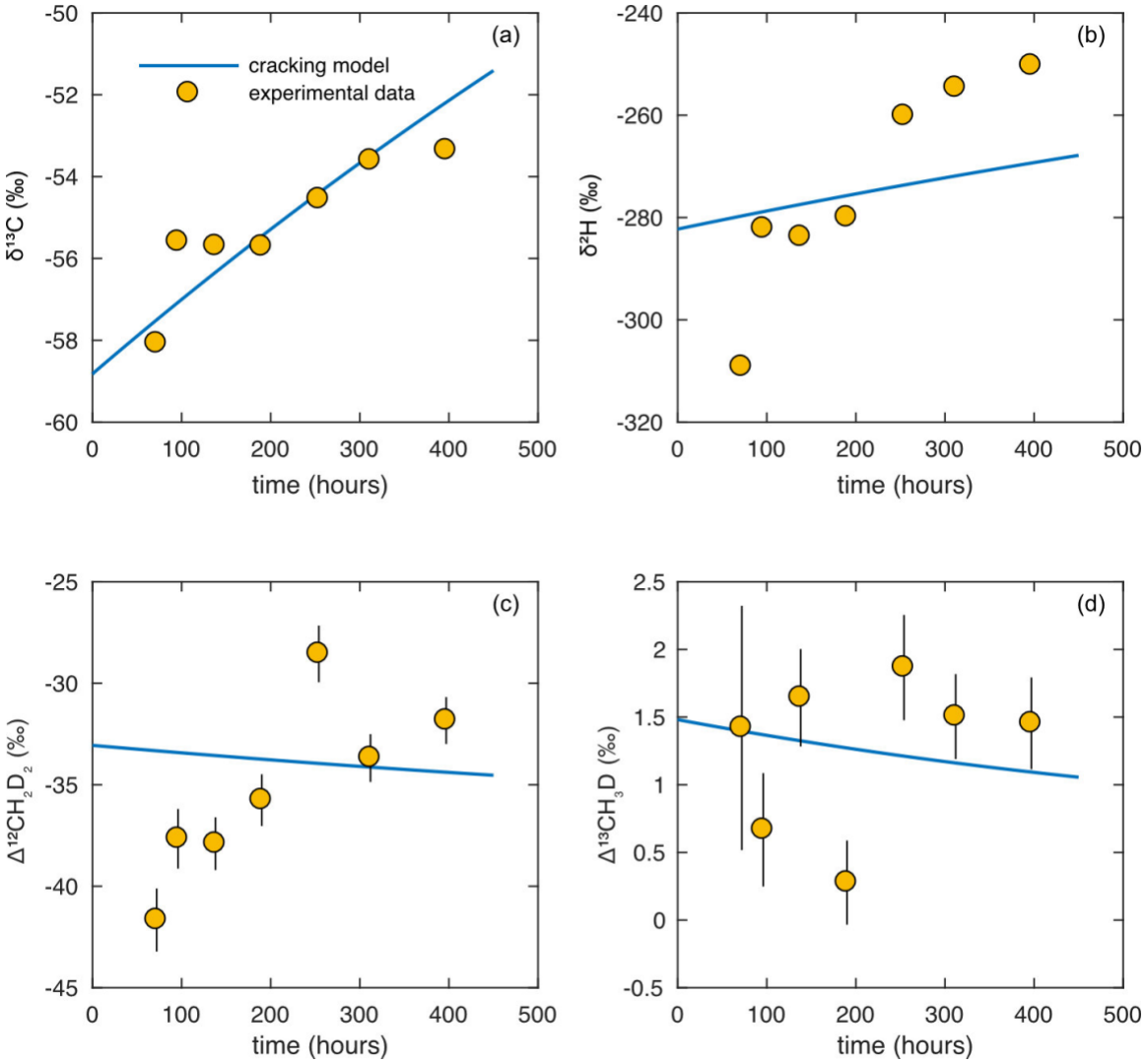


Figure 2.6 The optimal cracking model results (blue line) fitted to pyrolysis experiment results (yellow circles) (Table 2.5). The cracking model fits the trend in $\delta^{13}C$ with an R-square of 0.71 and fits the $\Delta^{13}CH_3D$ values to within 2 standard errors. However, this model was not capable of finding a solution that fit the trends for δD and $\Delta^{12}CH_2D_2$ to within small multiples of their respective errors (see also Figure 2.10). On the contrary, the model predicts a decreasing trend in $\Delta^{12}CH_2D_2$ with increasing reaction progress and a slower rate of increase in δD with time. Error bars are 1 standard experimental error.

4.2.1.3. Rayleigh distillations and the combinatorial effect

We find the cracking model fits the time evolution of the yield and $\delta^{13}C$ of methane with an R^2 of 0.73 and 0.71 (Figure 2.1a, Figure 2.6a), and fits measured $\Delta^{13}CH_3D$ values to within 2 standard errors (Figure 2.6d). However, this model fits δD and $\Delta^{12}CH_2D_2$ poorly with R^2 of 0.30 and -0.29, respectively (Figure 2.6b, c).

The modeled increase in $\delta^{13}C$ and δD with time reflects the fact that the isotopic compositions of the accumulated products of Rayleigh distillation trend toward that of the source substrate as reaction proceeds (Figure 2.6a, b). The enrichment of $\delta^{13}C$ of product methane with time agrees well with the modeled trend, however the increase in methane δD in experiments is more rapid than that expected for the effect of Rayleigh distillation during cracking (Figure 2.6b, see also Figure 2.10).

The $\Delta^{12}CH_2D_2$ value of methane produced by cracking are modeled to be 30–40‰ lower than the stochastic distribution (Figure 2.6c). In the cracking model, product methane inherits three out of its four hydrogen atoms from the methyl precursor (A-CH_3) with a fitted KIE of ~0.928 (Table 2.5), leading to a δD value of ~ -150‰. The fourth hydrogen in the product methane, the capping hydrogen, is abstracted from hydrogen bearing compound (B-H) with a fitted KIE of ~0.357 (Table 2.5), resulting in a δD value for this capping hydrogen as low as ~ -670‰. Therefore, the three hydrogen atoms from A-CH_3 in each methane molecule constitute a pool with higher deuterium concentration, $[D]_A$, whereas the capping hydrogen from BH makes up a second pool with lower deuterium concentration, $[D]_B$, the deuterium concentration of product methane is the weighted average of these two pools:

$$[D]_{CH4} = (3[D]_A + [D]_B)/4 \quad (12)$$

Nevertheless, the four hydrogen atoms are symmetrically equivalent in a product methane molecule. This symmetric equivalence leads one to predict a stochastic abundance of $^{12}CH_2D_2$ (the reference frame for $\Delta^{12}CH_2D_2$ values) as

$$[^{12}CH_2D_2]_{sto} \approx 6\times[D]_{CH4}\times[D]_{CH4} = 6\times[(3[D]_A + [D]_B)/4]^2 \quad (13)$$

However, the actual probability of randomly forming $^{12}CH_2D_2$ is:

$$[^{12}CH_2D_2] \approx 3\times[D]_A\times[D]_B + 3\times[D]_A^2 \quad (14)$$

And for all cases where $[D]_A$ and $[D]_B$ lie between 0 and 1 and are unequal:

$$6\times[(3[D]_A + [D]_B)/4]^2 > 3\times[D]_A\times[D]_B + 3\times[D]_A^2 \quad (15)$$

This inequality in (15) leads to negative $\Delta^{12}CH_2D_2$ (i.e., deficit relative to stochastic distribution). And larger differences between $[D]_A$ and $[D]_B$ lead to larger deficits in $\Delta^{12}CH_2D_2$ (Figure 2.10). Note that we leave out the [H] terms in the derivation above (though it is included in the numerical model) to make the inequality easier to prove. This approximation leads to no analytically resolvable systematic errors provided $0<[D]_A<0.001$ and $0<[D]_B<0.001$ (i.e., when δD of both pools is < ~5400‰—clearly true in the case of our experiments).

This mechanism for producing exceptional clumped isotope Δ values has been called the 'combinatorial effect' (Yeung, 2016; Röckmann et al., 2016; Clog et al., 2018; Cao et al., 2019; Xia and Gao, 2019; Taenzer et al., 2020). Such effects can arise any time two or more symmetrically equivalent molecular sites are created in such a way that they have different probabilities of containing a rare isotope substitution; these anomalies have been observed in microbial methane (Taenzer et al., 2020), and predicted for O_2, N_2 (Yeung, 2016) and ethane (Clog et al., 2018), and should be expected for emerging clumped isotope systems like sulfate (Ueno et al., 2019; Neubauer et al., 2020), H_2 (Popa et al., 2019), perchlorate (Martin, 2020), nitrate (Neubauer et al., 2020), and Δ_{48} in CO_2 (Fiebig et al., 2019a; Bajnai et al., 2020).

The cracking model predicts a 5‰ decrease in $\Delta^{12}CH_2D_2$ over the course of experiments (Figure 2.6c). This reflects the fact that B-H is a larger reservoir for capping hydrogen than

A-CH_3 is as a reservoir for methyl. So the increase in δD of residual B-H caused by Rayleigh distillation is smaller than that for residual A-CH_3. This effect reinforces the difference in δD between the methyl group and capping hydrogen, and the modeled $\Delta^{12}CH_2D_2$ value decreases further with experimental time rather than increases as we observe.

No combinatorial effect is expected for the $\Delta^{13}CH_3D$ value in the model. The stochastic abundance of $^{13}CH_3D$ is

$$[^{13}CH_3D]_{sto} \approx 4\times[^{13}C]_{CH4}\times[D]_{CH4} = [^{13}C]_{CH4}\times(3[D]_A+[D]_B) \quad (16)$$

In the model, $^{13}CH_3D$ forms by inheriting its ^{13}C from A-CH_3, whereas the deuterium may be inherited either from A-CH_3 or B-H (Table A1). The probability of randomly forming $^{13}CH_3D$ is

$$[^{13}CH_3D] \approx [^{13}C]_A\times3\times[D]_A+[^{13}C]_A\times[D]_B \quad (17)$$

And (16)=(17). The 0.5‰ decrease in modeled $\Delta^{13}CH_3D$ value over the course of experiments (Figure 2.6d) reflects a nonlinear effect of producing methane from a progressively fractionated precursor pool (A-CH_3) in a Rayleigh distillation process. This stems from clumped isotope mixing effects—when one mixing endmember has higher $\delta^{13}C$ and δD than the other endmember, the nonlinear variations in $\Delta^{13}CH_3D$ of the mixture will lead to a positive anomaly (Douglas et al., 2016). In our model, the initial methyl precursor (A-CH_3) separates into a product pool that experiences KIEs and therefore has depleted $\delta^{13}C$ and δD, and a residual reactant pool that experiences enrichment in $\delta^{13}C$ and δD. Following the nonlinearity of mixing in Δ values, the initial precursor will have a higher $\Delta^{13}CH_2D$ value than the weighted average of those of the product pool and the residual reactant pool. Since the product has the same $\Delta^{13}CH_2D$ value as the initial precursor (because of no clumped-isotope-specific KIEs), the residual reactant will have lower $\Delta^{13}CH_2D$ than the initial precursor. As the reaction proceeds, $\Delta^{13}CH_2D$ of the residual reactant gradually decreases, following a trend of Rayleigh distillation. Therefore, it causes $\Delta^{13}CH_3D$ of product methane

to decrease in the cracking model—but only subtly (by less than 2 standard errors of measurement; Figure 2.6d).

4.2.1.4. Summary of cracking model

The cracking model demonstrates that the combinatorial effect of combining methyl and the 'capping' hydrogen leads to the large deficit in $\Delta^{12}CH_2D_2$ (Figure 2.6c, A1) and draws down the δD of product methane (eqn. (12), Figure 2.10) —both noticeable features of our experimental products. Nevertheless, this model is incapable of capturing the time-evolving enrichments of δD and $\Delta^{12}CH_2D_2$ as pyrolysis proceeds. Consequently, as we consider these enrichments a robust observation (both trends have amplitudes that exceed 10 standard errors of measurements), we develop a model that considers additional chemical reactions with the purpose of explaining these observations.

4.2.2. Cracking-exchange model

In this section, we develop a model that can explain the rate of enrichment in $\delta^{13}C$, δD, and $\Delta^{12}CH_2D_2$ and lack of analytically significant change in $\Delta^{13}CH_3D$ with pyrolysis progress (Figure 2.3). The increase of $\delta^{13}C$, δD, and $\Delta^{12}CH_2D_2$ with thermal maturity (Ro) has also been observed in natural gases (Xie et al., 2019), so understanding this feature of our experimental results may be significant for the interpretation of data for natural samples.

4.2.2.1. hydrogen exchange reactions

The relatively rapid rise in δD and $\Delta^{12}CH_2D_2$ of methane in our experiments could reflect several processes, but we focus our attention on hydrogen exchange because it is known to be an important process both in pyrolysis experiments and in nature (Sessions et al., 2004; Schimmelmann et al., 2006; Sessions, 2016; Wei et al., 2019; Labidi et al., 2020; Xie et al., 2020).

Methane might exchange hydrogen during the pyrolysis with any of the co-existing hydrogen bearing compounds, i.e., hydrocarbons including precursor n-octadecane. However, we are not aware of experimental or theoretical constraints on the kinetics of hydrogen exchange between hydrocarbon compounds in absence of metal catalyst. This should be investigated in future research. Our best constraint on the rates of hydrogen exchange involving methane is perhaps its exchange with water, as water is a plausible substrate participating in hydrogen exchange in sedimentary environments (Sessions et al., 2004; Schimmelmann et al., 2006; Sessions, 2016). The half-life for hydrogen exchange between methane and water has been estimated to be on the time scale of tens of hours to a year at the temperature of our pyrolysis experiments, and millions of years in sedimentary basins (Wang et al., 2018 and references herein). Although the chemical mechanism of hydrogen exchange can be different for specific types of reactions (free-radical reactions versus reactions involving carbonium ions), the time scale of equilibration through hydrogen exchange observed in these previous studies seems broadly consistent with that of our pyrolysis experiments and natural gas formation at corresponding temperatures.

4.2.2.2. Cracking-exchange model setup

In order to test the effect of hydrogen exchange on the isotopic composition of methane produced in pyrolysis experiments, we add to the cracking model in 4.2.1 (1) hydrogen exchange reactions between the five most abundant isotopologues of methane and hydrogen bearing compound B-H or B-D

$$CH_xD_y + B\text{-}H \rightleftarrows CH_{x+1}D_{y-1} + B\text{-}D \quad \text{(R6)}$$

where CH_xD_y and $CH_{x+1}D_{y-1}$ represent isotopologues of methane participating in the exchange, and (2) hydrogen exchange reactions among the five most abundant isotopologues of methane, via (R1) and (R2) (see a full list of exchange reactions in Table A1). The rate of each methane-BH or methane-methane exchange reaction is:

$$r_n = \sigma_r\, k[i][j] \quad (11)$$

where r_n is the rate of reaction RAn (n = 9–20) in Table A1, σ_r and k are the symmetry number and the rate constant of RAn, respectively, and [i] and [j] are concentrations of reactants of RAn (Table A1, Appendix A1).

We develop a cracking-exchange model numerically with 14 isotopic species and 20 isotope-specific reactions (Table A1). In addition to the parameters in the cracking model, we incorporate a primary ^{13}C KIE for exchange involving $^{13}CH_4$ ($^{13C}KIE_{PRI}^{C-H}$), a secondary deuterium KIE for exchange involving $^{12}CH_3D$ ($^{D}KIE_{SEC}^{C-H}$), an equilibrium hydrogen isotope fractionation between methane and BH ($^{D}\alpha^{eq}_{CH4\text{-}BH}$), and equilibrium constants of reaction (R1) (K_{13CH3D}) and (R2) ($K_{12CH2D2}$) (Eldridge et al., 2019a) in the rate constants of these reactions (Table A1, Appendix A1). We assume the kinetic isotope effect for multiply-substituted species is the product of corresponding singly-substituted species, as in the cracking model. All concentrations in the model are normalized by moles of starting n-octadecane in experiments.

The initial concentrations of A-CH_3 and B-H, and the rate constant for $^{12}CH_4$ generation (k_{CH4Gen}) are the same as in the cracking model because hydrogen exchange reactions do not alter the molecular compositions of methyl precursors (A-CH_3), hydrogen bearing compounds (B-H) or methane. We use the measured $\delta^{13}C$ and δD values of n-$C_{18}H_{38}$ to set the D/H and $^{13}C/^{12}C$ ratios of A-CH_3 and B-H. And we fit $^{13C}KIE_{PRI}$, $^{13C}KIE_{PRI}^{C-H}$, $^{D}KIE_{PRI}$, $^{D}KIE_{SEC}$, $^{D}KIE_{SEC}^{C-H}$, $^{D}\alpha^{eq}_{CH4\text{-}BH}$, the initial clumped isotope compositions of the methyl group in A-CH_3 ($\Delta^{13}CH_2D$ and $\Delta^{12}CHD_2$) and the rate constants of the methane-BH hydrogen exchange reaction ($k_{CH4\text{-}BH}$) and the methane-methane hydrogen exchange reaction ($k_{CH4\text{-}CH4}$), where the fit minimized disagreement with time evolution of all measured isotopic properties of methane using methods given in section 2.5 (Table 2.5).

4.2.2.3. Partial equilibration of δD and Δ12CH2D2 in the cracking-exchange model

We find the cracking-exchange model fits the time evolution of yield, $\delta^{13}C$, δD and $\Delta^{12}CH_2D_2$ of methane with an R^2 of 0.73, 0.71, 0.87, and 0.58 (Figure 2.1, Figure 2.7a, b, c), and fits the $\Delta^{13}CH_3D$ values to within 2 to 3 standard errors (Figure 2.7d). We conclude this model is capable of simultaneously describing all major features of the experimental data.

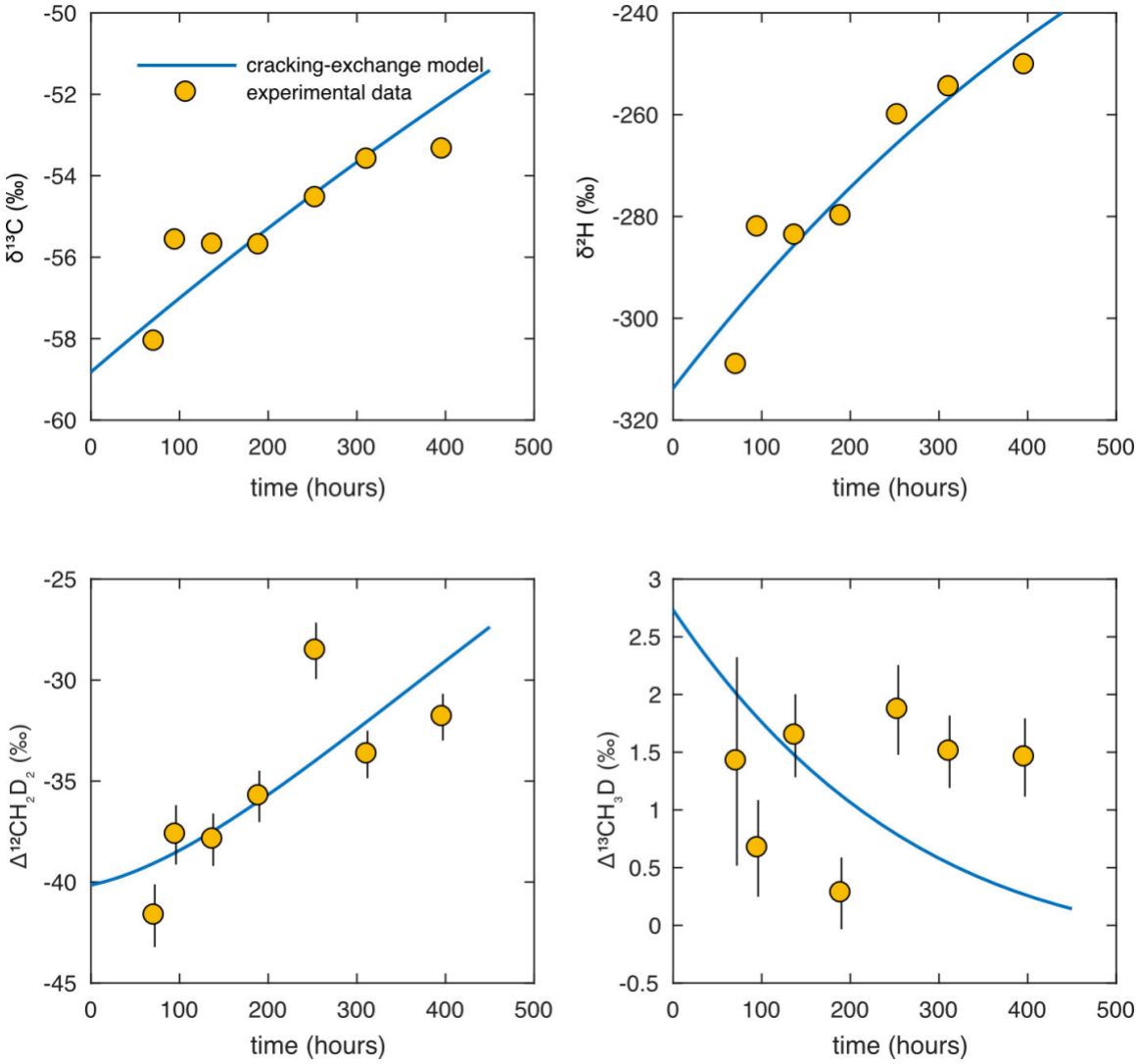


Figure 2.7 The optimal cracking-exchange model results (blue line) fitted to pyrolysis experiment results (yellow circles) (Table 2.5). The cracking-exchange model fits $\delta^{13}C$, δD and $\Delta^{12}CH_2D_2$ with an R-square of

0.71, 0.87, and 0.58, and fits the $\Delta^{13}CH_3D$ values to within 2–3 standard errors at all reaction progress values. All major features of the data can be explained, especially the enrichment in $\Delta^{12}CH_2D_2$ and δD as pyrolysis proceeds. Error bars are 1 standard experimental error.

In contrast to the cracking model, δD and $\Delta^{12}CH_2D_2$ of methane in the cracking-exchange model increase at a rate consistent with experimental results (Figure 2.7b, c). This reflects the fact that hydrogen exchange reactions between methane and co-existing hydrogen bearing compound B-H partially equilibrates the hydrogen isotope fractionation between methane and BH throughout our experimental period. This equilibration process leads to an increase in δD of methane with time more rapid than that expected for Rayleigh distillation alone (Figure 2.6b, 7b). The half-life of methane δD equilibration is 415 hours in the model (if approximated as a pseudo first-order reaction). This process also partially equilibrates clumped isotope compositions of methane as it affects all of methane's hydrogen atoms equally, and therefore partially erases the $\Delta^{12}CH_2D_2$ deficit (Figure 2.7c).

The experimental data are consistent with the cracking-exchange model results, with or without methane-methane hydrogen exchange reactions using our fitted parameters (Figure 2.11). This may reflect the fact that the hydrogen exchange between methane and other hydrogen bearing compounds (B-H) is more important to the equilibration of $\Delta^{12}CH_2D_2$ in the experiments than methane-methane hydrogen exchange reactions.

The modeled $\Delta^{13}CH_3D$ value decreases by ~2‰ over the course of our experiments (Figure 2.7d). This decrease results from the nonlinear effect of Rayleigh distillation (4.2.1.3) and hydrogen exchange reactions (Xia and Gao, 2019). Specifically, the primary ^{13}C KIE in breaking the carbon-hydrogen bond of methane in reactions (RA11) and (RA14) slows down the equilibration of $^{13}CH_3D$. This trend is not clearly evidenced by our data, but could be consistent with observations, given analytical uncertainties.

4.2.2.4. Model assumptions and comparison of fitted parameters with prior constraints

In the model, we assume the isotopic compositions of A-CH_3 and B-H are equal to $\delta^{13}C$ and δD values of n-$C_{18}H_{38}$ measured by GC-combustion/pyrolysis-irMS, i.e., the average values for all carbon and hydrogen sites in n-$C_{18}H_{38}$. So any difference in isotopic composition between the precursors of methyl groups (A-CH_3), precursors of capping hydrogens (B-H), and the rest of the sites of the substrate will be interpreted as contributing to the net KIE observed in the reaction. In a potentially extreme case that A-CH_3 and B-H have different δD values of -85.7 and -572.0‰, respectively (one of the largest measured intramolecular hydrogen isotope contrast in fatty acids from plants (Billault et al., 2001)), and assuming product methane inherits hydrogen atoms without any KIEs, we predict a deficit of -23‰ in $\Delta^{12}CH_2D_2$ and hydrogen isotope fractionation between methane and starting A-CH_3 of -133‰. We do not know the site-specific hydrogen isotopic composition of our reactant n-octadecane, so its contribution to the deficit in $\Delta^{12}CH_2D_2$ cannot be quantified (in fact, we are not aware of any measurements of natural-abundance, site-specific hydrogen isotope structures of alkanes larger than propane). But this problem should be revisited if and when it becomes possible to directly couple clumped isotope measurements of product methane with site-specific measurements of the D/H ratios of alkyl and methyl groups from methane's precursor.

The primary and secondary KIEs of breaking carbon-carbon and carbon-hydrogen bonds in n-alkanes through homolytic cleavage, β scission and hydrogen abstraction have been systematically studied with density function theory calculations (Canadell and Olivella, 1984; Lu et al., 1990; Tang et al., 2000, 2005; Xiao, 2001; Ni et al., 2011), or laboratory experiments (Jackson et al., 1962; Gray et al., 1971 and references herein). The values we fit for the KIEs of both carbon and hydrogen isotopes in the cracking-exchange model are within 10% of previous predictions and experimental results (Table 2.5). We conclude that the fitted KIEs are consistent with the range of these independent constraints.

Gray et al. (1971) noted that the experimentally-determined KIE's (k_D/k_H) for a series of hydrogen abstraction reactions by methyl radicals are smaller than predicted. At the temperature of our experiment (400°C), experimental primary deuterium KIE's are ~100–150‰ lower than predictions. These predictions assume the C-H stretching mode of the abstracted hydrogen is lost in the transition state (i.e., the stretching frequency becomes imaginary). Gray et al. proposed two hypotheses to explain the contradiction between experimental and predicted KIEs: first, in addition to the stretching mode, the bending mode may also be lost in the transition state. This change leads to a prediction in better agreement with the experimental value. A second possible explanation for the lower kinetic isotope effect is caused by quantum mechanical tunneling of hydrogen. However, previous calculations do not show significant tunneling contributions to the KIE in methyl radical attacking a methylene or methyl group (Johnston and Rapp, 1961; Salomon, 1964; Gray et al., 1971). In any case, we note that our fitted model is consistent with the experimental constraints on the KIE in hydrogen abstraction.

The equilibrium hydrogen isotope fractionation between methane and hydrogen-bearing alkyl compound B-H ($^{D}\alpha^{eq}_{CH4\text{-}BH}$) at 400°C can be estimated from reduced partition function ratios using Urey-Bigeleisen-Meyer (U-B-M) theory, with reduced partition functions extrapolated to our experimental temperature (Wang et al., 2009) or calculated from vibrational frequencies of fundamental modes computed by density function theory (Piasecki et al., 2016b). The fitted $^{D}\alpha^{eq}_{CH4\text{-}BH}$ value in our model is consistent with these estimated constraints within 2–3 standard errors of δD measurement by GC-pyrolysis-irMS (Table 2.5).

We are not aware of any published experimental studies that characterize the clumped isotope compositions of methyl groups in hydrocarbon compounds, or any computational studies on isotopic clumping in methyl in compounds larger than ethane. Nonetheless, the equilibrium $\Delta^{13}CH_2D$ and $\Delta^{12}CHD_2$ of the methyl group in ethane can be calculated using U-B-M reduced partition function theory (Piasecki et al., 2016b). The equilibrium $\Delta^{13}CH_2D$

and $\Delta^{12}CHD_2$ values are predicted to be 1.068‰ and 1.853‰ at 400°C, respectively. In our cracking model, the fitted initial clumped isotope compositions of methyl groups are within 2 standard errors of $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ measurements (Table 2.5). In the cracking-exchange model, the fitted initial $\Delta^{13}CH_2D$ and $\Delta^{12}CHD_2$ are 3.02‰ and 0.07‰, respectively. Sensitivity tests of our model show that $\Delta^{13}CH_3D$ is sensitive to the initial $\Delta^{13}CH_2D$ of the methyl group, which agrees with the findings in Xia and Gao (2019). Nevertheless, for the highly-depleted $\Delta^{12}CH_2D_2$ values observed in our experiments, any initial $\Delta^{12}CHD_2$ value in a plausible range (0–8‰) does not significantly affect the modeled $\Delta^{12}CH_2D_2$ value.

While the $\Delta^{13}CH_3D$ values of experimental products appear to be consistent with the thermodynamic equilibrium composition at 400°C, within two standard errors, we recognize that this does not necessarily indicate that $^{13}CH_3D$ forms in or conforms to thermodynamic equilibrium throughout the experimental period. In fact, our model predicts that methane initially inherits the $\Delta^{13}CH_2D$ of the methyl precursor, and both the generation and subsequent hydrogen exchange reactions can have kinetic effects on $\Delta^{13}CH_3D$ (Figure 2.6 6d, 4.2.1.3, Figure 2.7d, 4.2.2.3). Although the range of our experimental data constrains that such kinetic effects on $\Delta^{13}CH_3D$ values are limited throughout the experimental period, these kinetic effects in Rayleigh distillation and exchange reactions may be able to explain the non-equilibrated and, in some cases, negative Δ_{18} values of methane produced during pyrolysis of coal (Shuai et al., 2018; Xia and Gao, 2019).

A recent non-peer-reviewed presentation reports isotopic contents of methane produced by the high temperature (600°C) pyrolysis of an ethane substrate, finding results broadly similar to those reported here (i.e., apparently equilibrated $\Delta^{13}CH_3D$ and highly-depleted $\Delta^{12}CH_2D_2$ values; Eldridge et al., 2019b). It has also been shown recently that low-temperature *Pseudomonas stutzeri* cultures produce methane by combining a methyl group (formed by decomposing methylphosphonate, CH_3O_3P) and a hydrogen atom from water (Taenzer et al., 2020). The combinatorial effect can explain the observed $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ values

and is particularly clearly evidenced by variation of $\Delta^{12}CH_2D_2$ between methane formed in the presence of waters of various hydrogen isotope compositions. In that study, $\Delta^{13}CH_3D$ values are invariant in a range of -0.3 to 0.93‰—similar to those observed here and by Eldridge et al. (2019b), but out of equilibrium at culture temperature. While the role of combinatorial effects on $^{12}CH_2D_2$ seems clear in all of these studies, further insight into the controls of $^{13}CH_3D$ will require investigations of the clumped isotope composition of methyl groups in the precursors of methane (i.e., n-octadecane, ethane, and methylphosphonate in the studies discussed above), as well as further understanding of the kinetics of hydrogen exchange reactions.

4.3. Modeling the isotopic content of natural thermogenic methane

We extrapolate the cracking-exchange model to geological times and temperatures to test whether the processes we propose that controlled our experiments might also explain the known patterns of isotopic variation associated with the production of thermogenic methane. The aim of this modeling activity is not to provide fitted parameters universally applicable to thermogenic methane formation, nor to assert that the chemistry of our experiments must necessarily be relevant to natural thermogenesis; rather, we ask whether adapting the cracking-exchange model to data of natural thermogenic methane will lead to a plausible and testable mechanistic hypothesis that explains its observed isotopic properties, which are generally open to a range of interpretations (D. A. Stolper et al., 2014; Young et al., 2017; Stolper et al., 2018; Giunta et al., 2019; Xia and Gao, 2019).

We can predict the isotopic properties of methane if it forms by the mechanisms included in the model of our experiments, and the kinetics of the model follow the Arrhenius equation and experience kinetic isotope effects on the pre-exponential factors and activation energies (Gray et al., 1971; Tang et al., 2000; Ni et al., 2011; Xia and Gao, 2019). We impose a temperature-time history where the temperature range and the heating rate simulate the process of typical sedimentary burial (Table 6). A full description of the model setup is in Appendix A2.

For equivalent thermal maturity (Easy%RoV, Burnham, 2019) between 0.6 and 0.8, our model predicts that this low-maturity thermogenic methane will have $\delta^{13}C$ of -65 to -50‰ and δD of -440 to -300‰. Both $\delta^{13}C$ and δD are expected to then increase with subsequent maturation until δD reaches equilibrium with co-existing hydrogen-bearing compounds at Easy%RoV ~1.5 (Figure 2.8a). $\Delta^{12}CH_2D_2$ of low-maturity methane is predicted to start as low as ~ -60‰ because of the combinatorial effect, then rise to partially equilibrated $\Delta^{12}CH_2D_2$ values of -25 ~ 0‰ for maturities in the range of Easy%RoV ~0.8–1.2, and finally reach equilibrated $\Delta^{12}CH_2D_2$ values at maturity Easy%RoV > ~1.2 (Figure 2.8b, c). The $\Delta^{13}CH_3D$ composition is predicted to decrease relative to substrates when Easy%RoV < ~0.8 because of kinetic effects in Rayleigh distillation and exchange reactions (as discussed in 4.2.2.4). As thermal maturity rises above 0.8%, the $\Delta^{13}CH_3D$ value increases and approaches thermodynamic equilibrium at maturity Easy%RoV > ~1.1 (Figure 2.8b, d).

Table 2.6 Thermogenesis model parameters and data sources

parameter	value	data source
temperature range (°C)	50-200	Peters et al. (2004)
heating rate	3°C/Myr	
<u>precursor concentrations (dimensionless)</u>		
ACH_3	1	/
BH	100	/
<u>precursor isotope compositions (‰)</u>		
$\delta^{13}C_{VPDB}$ (ACH_3)	-20, -25, -30,	Schoell (1984); Chung et al. (1992)
δD_{SMOW} (ACH_3)	-100	
δD_{SMOW} (BH)	-100	
$\Delta^{13}CH_2D$ (ACH_3)	4.0	fitting
$\Delta^{12}CHD_2$ (ACH_3)	1.0	/
<u>kinetic parameter</u>		
A_{CH4Gen} (s^{-1})	1.00E+08	Xia and Gao (2019)
$A_{CH4\text{-}BH}$ (s^{-1})	2.00E-05	fitting
$A_{CH4\text{-}CH4}$ (s^{-1})	3.16E-05	fitting
E_{CH4Gen} (kJ/mol)	185	Xia and Gao (2019)
E^S_{CH4Gen} (kJ/mol)	25	
$E_{CH4\text{-}BH}$ (kJ/mol)	100	Metcalfe (1973), Koepp (1978)
$E^S_{CH4\text{-}BH}$ (kJ/mol)	13.3	
$E_{CH4\text{-}CH4}$ (kJ/mol)	60	Xia and Gao (2019)
<u>Isotope effect parameters</u>		
$^{13C}A^*/A_{PRI}$	1.03	Ni et al. (2011)
$^{13C}\Delta Ea_{PRI}$ (J/mol)	211.3	
$^{D}A^*/A_{SEC}$	1.07	
$^{D}\Delta Ea_{SEC}$ (J/mol)	1333	
$^{D}A/A^*_{PRI}$	1.0745	Gray et al. (1971)
$^{D}\Delta Ea_{PRI}$ (J/mol)	5409.1	

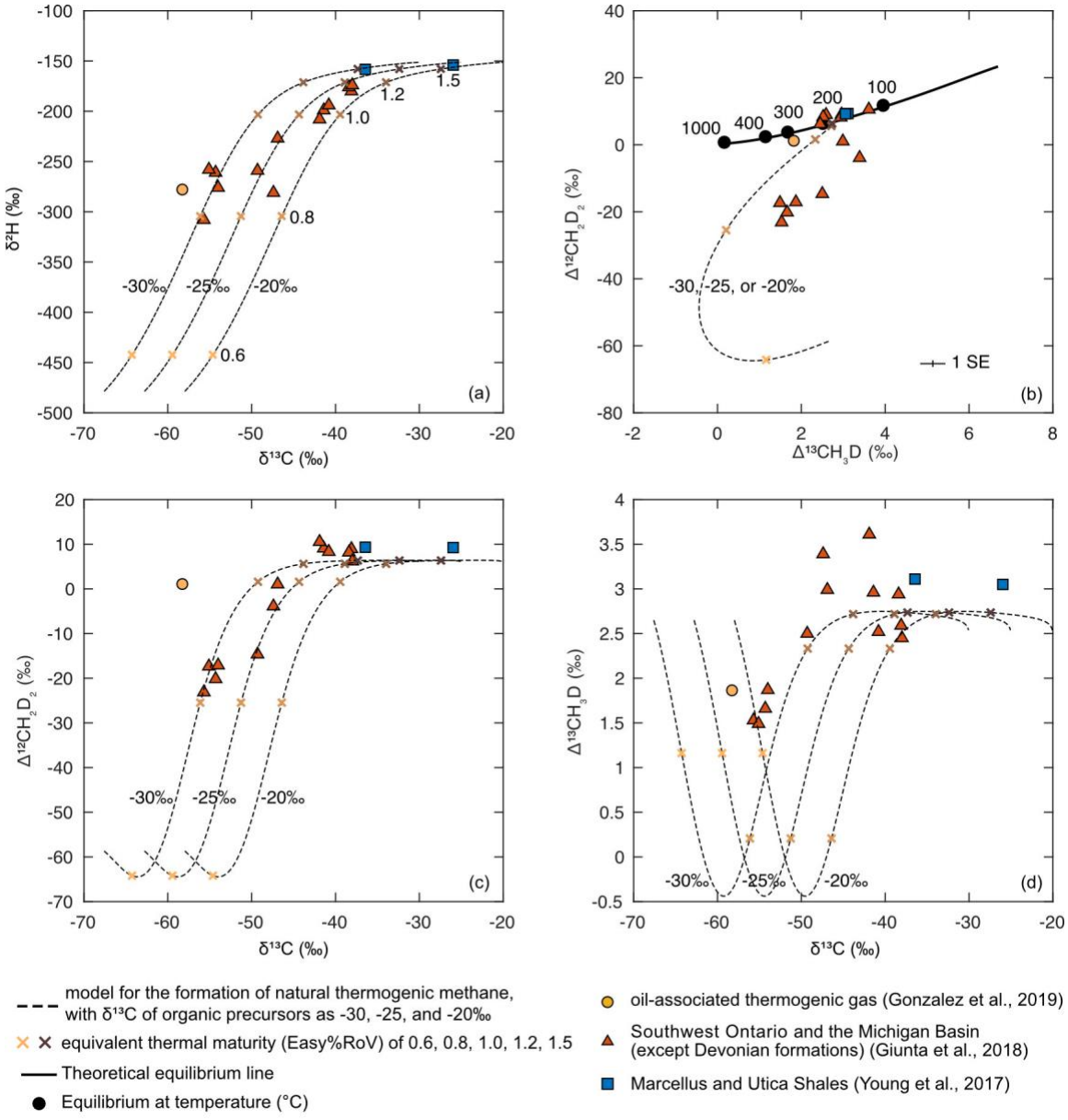


Figure 2.8 Modeling the generation, maturation and equilibration of natural thermogenic methane (Table 2.6, Appendix A2) with $\delta^{13}C$ of the organic precursor as -30, -25, and -20‰. This model predicts $\delta^{13}C$ and δD (a) increase with thermal maturity (%Ro) until δD reaches equilibrium with co-existing hydrogen-bearing compounds at ~1.5%Ro. $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ are insensitive to $\delta^{13}C$ of the organic precursor (b). $\Delta^{12}CH_2D_2$ (b, c) increases from ~ -60‰ at ~0.6%Ro, to ~0‰ at 1.0%Ro, and reaches equilibrium at burial temperature at maturity > ~1.1%Ro. $\Delta^{13}CH_3D$ (b, d) first decreases at maturity < ~0.8%Ro, then increases at 0.8–1.0%Ro and approach equilibrium at ~1.2%Ro. For mature gases with %Ro >1.5, clumped isotope

compositions are fully equilibrated. The predictions are consistent with the trend of data for natural thermogenic methanes and suggest an alternative thermogenic origin for the deficits in $\Delta^{12}CH_2D_2$ of some natural gases. Yellow and brown crosses indicate equivalent thermal maturity of 0.6, 0.8, 1.0, 1.2 and 1.5%Ro (color-coated with Easy%RoV). The black cross in (b) indicates 1 SE of analyses. Data from this study, Young et al. (2017), Giunta et al. (2019), Gonzalez et al. (2019).

The predictions of our model fit the $\delta^{13}C$, δD, $\Delta^{12}CH_2D_2$ and $\Delta^{13}CH_3D$ data of natural thermogenic methanes generally well (Young et al., 2017; Gonzalez et al., 2019) (Figure 2.8). $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ values of some of these methanes are higher than expected for their $\delta^{13}C$ values (Figure 2.8c, d). This may reflect the possible migration of methane to reservoirs colder than its production temperature, which may cause isotope compositions to re-equilibrate to the lower storage temperature (this effect is not included in our model). Additionally, our model predictions fit well with the trend in isotopic composition of some methanes from Southwest Ontario and the Michigan Basin, which were previously interpreted to be mixtures of thermogenic and microbial gases (Giunta et al., 2019). We discuss this observation in section 4.4.

There are a few caveats in extrapolating the cracking-exchange model of n-octadecane pyrolysis to modeling of natural gas formation. We modeled a closed system where all generated methane is accumulated in the product, whereas natural gases often have semi-open reservoirs (Xia and Gao, 2019). Additionally, the thermal cracking of a pure compound such as n-octadecane and hydrogen exchange among hydrocarbons may not adequately represent the range of reactions taking place during the cracking and maturation of kerogens or crude oils (Tang et al., 2000). Despite these potential shortcomings, the trend of isotopic compositions of methane predicted by the model is consistent with data obtained for natural thermogenic methanes.

4.4. Implications for the interpretation of clumped isotope compositions of methane in natural gases

Giunta et al. (2019) examined gases in Southwest Ontario and the Michigan Basin from Cambrian to Devonian strata. The high C_1/C_{2+} ratio in many Antrim Shale gases (30 to 300) (Upper Devonian strata in Michigan Basin) suggests significant microbial contributions, from either microbial methanogenesis or secondary microbial degradation of hydrocarbons. In this previous study, Silurian gases have negative and varied $\Delta^{12}CH_2D_2$ compositions (-23 to 1‰), whereas the Ordovician and Cambrian gases appear to have clumped isotope compositions consistent with thermodynamic equilibrium. Giunta and colleagues proposed a two-end-member mixing model, including a thermogenic equilibrium end member and a microbial disequilibrium end member. Note that some variations in the 'wetness' of thermogenic end members are required by this model to fit data from different suites.

Based on our experimental results and cracking-exchange model, we suggest an alternative hypothesis for the Silurian gases that exhibit deficits in $\Delta^{12}CH_2D_2$ values. We suggest that they are early-oil- to early-gas-window thermogenic methanes that retain some part of the signature of the combinatorial effect on their $\Delta^{12}CH_2D_2$ values during formation. This hypothesis is supported by the low thermal maturity of the in situ source and host rocks (Barker and Pollock, 1984; Sherwood Lollar et al., 1994), the low C_1/C_{2+} ratio of those gases, and the disequilibrium hydrogen isotope fractionation between methane and co-existing water. Consistent with the cracking-exchange model for natural methane formation in section 4.3, older gases from Cambrian and Ordovician strata in Southwest Ontario exhibit equilibrated $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ and equilibrated hydrogen isotope fractionations between methane and water. These equilibria may result from isotope exchange with water after a long enough period at their burial temperatures. $\delta^{13}C$, δD, $\Delta^{12}CH_2D_2$ and $\Delta^{13}CH_3D$ data of gases from Southwest Ontario and Michigan Basin (except for the Devonian Antrim Shale, where a microbial component seems clearly implicated (Martini et al., 1996)), can be explained by the cracking-exchange model for natural gases, especially considering a

variation in $\delta^{13}C$ values of organic precursors (Figure 2.8). $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ values higher than model prediction could result from cooling and re-equilibration after reaching the maximum burial temperature.

Gonzalez et al. (2019) measured an oil-associated thermogenic gas of early-oil-window maturity. Consistent with the cracking-exchange model for natural gases presented here, this oil-window gas has low $\delta^{13}C$ (−58.12±0.02‰), and non-equilibrated and depleted $\Delta^{12}CH_2D_2$ (0.8±0.8‰). This gas falls close to the trends predicted here in δD - $\delta^{13}C$ and $\Delta^{13}CH_3D$ - $\Delta^{12}CH_2D_2$ plots but deviates from the predicted $\delta^{13}C$ - $\Delta^{12}CH_2D_2$ and $\delta^{13}C$ - $\Delta^{13}CH_3D$ relationship (Figure 2.8). Although the model can qualitatively explain the disequilibrium in $\Delta^{12}CH_2D_2$, further quantitative discussion requires knowledge of geological burial history and precursors to methane in this gas.

Methane from the Beecher Island gas field shows a resolvable depletion (2.96‰) in $\Delta^{12}CH_2D_2$ relative to equilibrium at the apparent $\Delta^{13}CH_3D$ temperature of 142°C (Young et al., 2017). Although the disequilibrium in $\Delta^{12}CH_2D_2$ value resembles that of the oil-associated thermogenic gas in Gonzalez et al. (2019), the Beecher Island gas was attributed to a microbial origin as the $\delta^{13}C$ and δD of the methane fall in the microbial region on the Whiticar plot (Whiticar, 1999; Young et al., 2017; Milkov and Etiope, 2018). Future study may investigate whether the Beecher Island gas could be affected by mixing processes or by multiple thermogenic and microbial processes.

We also note that the depletion in $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ within early gas window (equivalent thermal maturity 0.8–1.2%Ro) will cause the Δ_{18} of methane to be depleted relative to equilibrium values. In our model, Δ_{18} is 1‰ lower than thermodynamic equilibrium at maturity Easy%RoV ~1, and Δ_{18} becomes indistinguishable from equilibrium at Easy%RoV > ~1.2. Therefore, the non-equilibrium Δ_{18} composition of unconventional oil-associated gases could be explained by kinetic effects in methane generation and hydrogen exchange and the combinatorial effects on $\Delta^{12}CH_2D_2$ (Douglas et al., 2017; Stolper et al., 2018; Xia and Gao, 2019).

4.5. Implications for the role of the combinatorial effect in abiogenesis and biogenesis of methane

Previous studies have argued that large depletions in $\Delta^{12}CH_2D_2$ of methane generated by the Sabatier reaction in the laboratory result from hydrogen quantum tunneling (Young et al., 2017; Young, 2019). However, the quantum tunneling model failed to explain the observed clumped isotope effects and δD values simultaneously (Young et al., 2017; Cao et al., 2019). Alternatively, the depleted $\Delta^{12}CH_2D_2$ could be explained by the combinatorial effect of assembling hydrogen of different isotopic compositions (Cao et al., 2019). According to the reaction scheme of Fischer–Tropsch type reactions (Young et al., 2017), methane production from CO and H_2 starts with forming a CH species through hydrogen addition to an oxygen-bonded carbon and dehydroxylation of an OH group; next, three hydrogen atoms are added to the CH successively. Therefore, it is possible that the first hydrogen is affected by isotope effects different from those affecting the last three added hydrogens (and those three might differ from one another for more subtle reasons relating to the evolving structures of the intermediates). Any such combinatorial effects will cause deficits in $\Delta^{12}CH_2D_2$. Cao et al. (2019) proposed a marginally different mechanism for this experiment where the addition of the first two hydrogens is considered reversible, while the addition of the last two hydrogens is considered irreversible. If distinct isotope effects are expressed in reversible and irreversible steps in this mechanism, the combinatorial effects will also cause deficits in $\Delta^{12}CH_2D_2$.

A similar combinatorial effect is also expected in the experimental decomposition of tetrakis(trimethylsilyl) silane in Young et al., (2017). Methane is produced by CH_3 escaping from the silane structure followed by the addition of hydrogen. If the CH_3 and the capping hydrogen have distinct D/H, the combinatorial effect will cause a deficit in $\Delta^{12}CH_2D_2$. This deficit may have been erased by hydrogen exchange during heating at 300–600°C for tens of

days, explaining why these higher temperature experiments produce methane that approaches equilibrium.

The combinatorial effect could also cause a deficit in $\Delta^{12}CH_2D_2$ of microbial methane if hydrogen atoms source from two or more isotopically different pools, e.g., from water, a reduced hydrogen pool, sulfur-bonded hydrogen, or carbon-bonded hydrogen (Valentine et al., 2004; Scheller et al., 2013; Greening et al., 2016; Young et al., 2017; Giunta et al., 2019; Cao et al., 2019). In fact, a combinatorial effect has been observed and studied in methane production by *Pseudomonas stutzeri* cultures, as discussed in 4.2.2.4 (Taenzer et al., 2020).

In conclusion, when four hydrogens in methane are sourced from isotopically different pools or fractionated by distinct isotope effects, there will be a combinatorial effect, irrespective of the conditions, mechanisms, and substrates of methanogenesis; the commonality of this mechanism provides a ready explanation for the similarities in clumped isotope systematics (particularly large depletion in $\Delta^{12}CH_2D_2$) in microbial, (putative) abiotic and thermogenic methane generated at early-oil to early-gas window maturities. An important implication of this conclusion is that $\Delta^{12}CH_2D_2$ alone is not a unique signature for abiotic or microbial methane.

4.6. Distinguishing thermogenic, microbial and abiotic methane by isotopic compositions

We compiled published $\delta^{13}C$, δD, $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ of natural methanes and examined whether data are organized into forensically diagnostic fields in plots of $\Delta^{13}CH_3D$ vs. $\Delta^{12}CH_2D_2$ and $\delta^{13}C$ - $\Delta^{12}CH_2D_2$ (Figure 2.9, b). We also plotted these natural methanes on the Whiticar plot with genetic fields from Milkov and Etiope (2018) (Figure 2.9c).

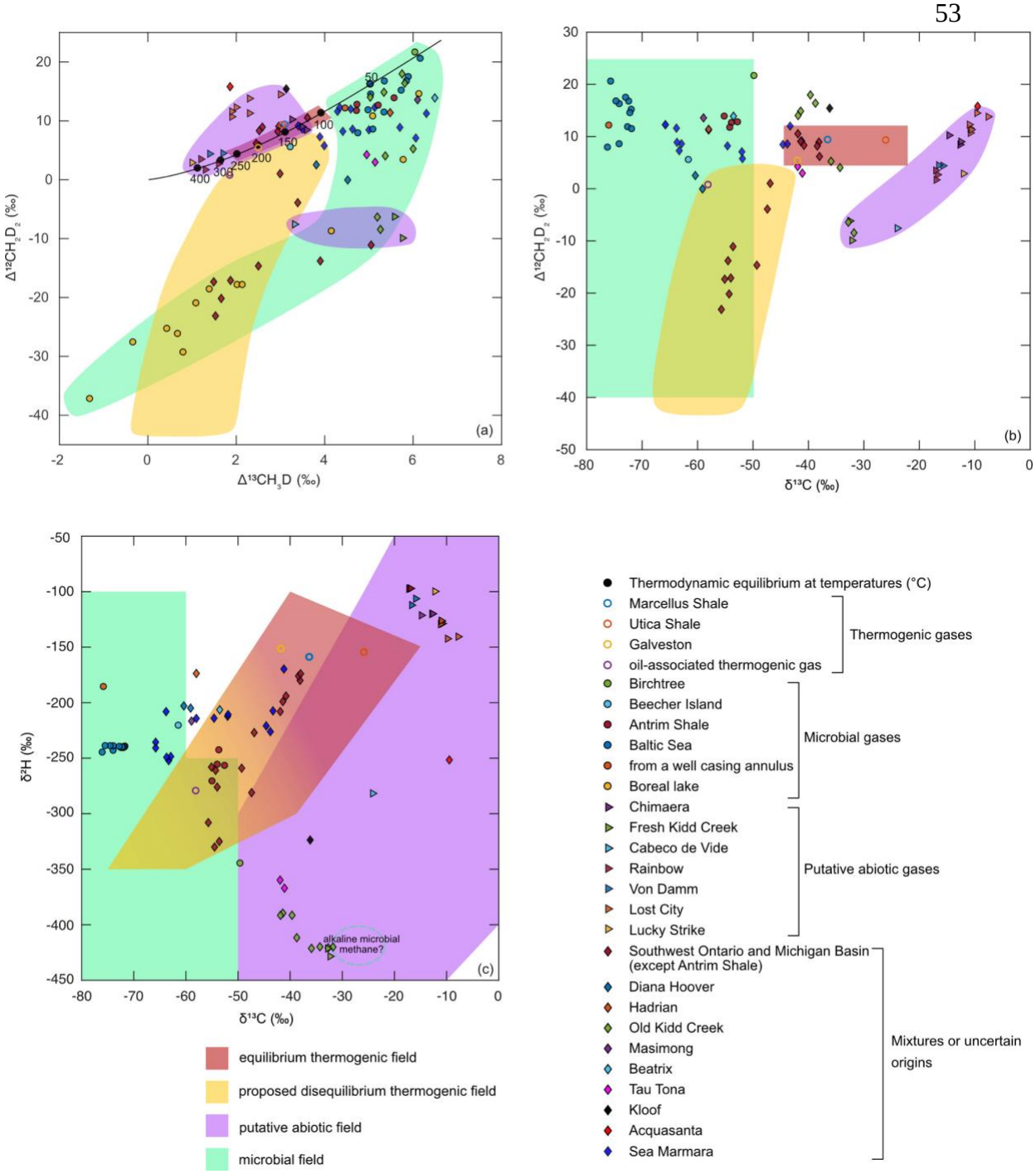


Figure 2.9 Diagrams of (a) $\Delta^{13}CH_3D$ - $\Delta^{12}CH_2D_2$, (b) $\delta^{13}C$ - $\Delta^{12}CH_2D_2$ and (c) $\delta^{13}C$ - δD with forensically diagnostic fields. We propose a disequilibrium thermogenic field based on a cracking-exchange model for natural thermogenic methane. Clumped isotope data from Young et al. (2017), Giunta et al. (2019), Ash et al. (2019), Gonzalez et al. (2019), Young (2019), Thiagarajan et al. (2020a), Labidi et al. (2020), and Giunta

et al. (2021). Original data of boreal lake were not provided and thus cannot be plotted on (b). Genetic fields on (c) from Milkov and Etiope (2018) and Miller et al. (2018) (see also Fiebig et al. (2019b) and Reeves and Fiebig (2020) for $\delta^{13}C$ - δD values of volcanic-hydrothermal methane).

When both $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ are in equilibrium at temperatures >100°C, and $\delta^{13}C$ and δD values are consistent with the widely recognized thermogenic range (Milkov and Etiope, 2018) (Figure 2.9c), it seems robust to identify such methane as thermogenic gas (Figure 2.9a). However, we propose a second, disequilibrium thermogenic field based on our cracking-exchange model (Figure 2.8) for natural gases (Figure 2.9a, b). For this disequilibrium thermogenic field, we propose a wide range in $\Delta^{12}CH_2D_2$ values (starting with ~ -40‰ and approaching equilibrium values expected for 100 to 250°C). The $\delta^{13}C$ of the proposed field extends from ~ -65 to -45‰ and co-evolves with $\Delta^{12}CH_2D_2$ (Figure 2.9b). We suggest a relatively broad range of $\Delta^{13}CH_3D$ values (~0 to 4‰) (Figure 2.9a), which is consistent with the prediction by cracking-exchange model for natural thermogenic methane (Figure 2.8b). Thermogenic gas in this field is interpreted to have been generated at early-oil to early-gas window maturities (est. %Ro ~ 0.6–1.2, Fig 8).

Microbial methane in a plot of $\Delta^{13}CH_3D$–$\Delta^{12}CH_2D_2$ extends from the equilibrated composition at temperatures <100°C to negative values ($\Delta^{13}CH_3D$ ~ -1.2‰ and $\Delta^{12}CH_2D_2$ ~ -37‰, Figure 2.9a). The wide distribution of microbial methane across this field likely reflects variations in substrate limitation and reversibility of microbial methanogenesis, and other equilibration processes, such as anaerobic oxidation of methane (Martini et al., 1996; Wang et al., 2015; Stolper et al., 2015; Gruen et al., 2018; Ash et al., 2019; Ono et al., 2021).

The compositional fields for microbial and disequilibrium thermogenic gases overlap between $\Delta^{13}CH_3D$ of ~0 to 3‰ and $\Delta^{12}CH_2D_2$ of -10 to -35‰, as well as between $\delta^{13}C$ of ~ -65 to -50‰ and $\delta^{13}D$ of ~ -350 to -250‰ (Figure 2.9). These overlapping fields mean that subsets of thermogenic methane and microbial methane cannot be distinguished using bulk or clumped isotopes. In these cases, other information, such as gas chemistry and contextual geological data, may be the only useful discriminants. However, the subset of microbial

methane having $\Delta^{13}CH_3D$ values < ~0‰ or > 4‰, or $\delta^{13}C$ less than ~ -70 ‰ appears to have no equivalents among methanes produced by catagenesis.

Methane of a putative abiotic origin is so far most distinct from other methane sources in the $\delta^{13}C - \Delta^{12}CH_2D_2$ plot (Figure 2.9b). The putative abiotic fields are characterized by relatively high $\delta^{13}C$, sometimes accompanied by a deficit in $\Delta^{12}CH_2D_2$ (e.g., Kidd Creek), or with $\Delta^{13}CH_3D$ or $\Delta^{12}CH_2D_2$-based temperatures of ~60–400°C (e.g., Chimaera, and hydrothermal methanes). Additionally, note that microbial processes could affect the pristine abiotic signature and drive clumped isotopes to equilibrium at environmental temperatures (e.g., fresh Kidd Creek versus old Kidd Creek in Figure 2.9, Ash et al., 2019).

Overall, $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ compositions provide additional information to sourcing methane in supplement to $\delta^{13}C$ and δD. However, initial suggestions that these clumped isotope proxies commonly provide unique signatures of biogenesis or abiogenesis appear to be untrue. Early catagenesis is a widespread process that yields abundant methane, potentially having isotopic signatures resembling some microbial and putative abiotic gases. In any case, we expect the genetic fields in Figure 2.9a and Figure 2.9b to expand as future researches characterize clumped isotope compositions of more natural methanes.

5. Summary and conclusions

We performed controlled pyrolysis experiments on n-octadecane ($C_{18}H_{38}$) at 400°C. We found that $\Delta^{13}CH_3D$ values of generated methane appear to be consistent with thermodynamic equilibrium at the experimental temperature, whereas $\Delta^{12}CH_2D_2$ values are 30 to 40‰ depleted compared with equilibrium. Values of $\Delta^{12}CH_2D_2$ increase with advancing pyrolysis progress. We quantitatively demonstrated that the large deficit in $\Delta^{12}CH_2D_2$ can result from the combinatorial effect in assembling hydrogen from isotopically-different pools into a methane molecule. The increase of $\Delta^{12}CH_2D_2$ with experimental time can be explained by partial equilibration of methane through hydrogen exchange with co-existing hydrocarbons.

We propose a cracking-exchange model for isotope signatures of thermogenic methane, including singly and doubly-substituted species, during generation, maturation and equilibration of natural gas. This model can explain the isotopic compositions of thermogenic methane that have been published to date. However, there is at present an underappreciation of the potential for such signatures among thermogenic natural gases, and more natural samples, particularly methane generated at early-gas window maturities, should be examined and compared with the model.

Our model for thermogenic methane also reveals that early-oil to early-gas window thermogenic methane has isotopic compositions ($\delta^{13}C$, δD, $\Delta^{13}CH_3D$, and $\Delta^{12}CH_2D_2$) similar to microbial methane. Although this similarity leads to difficulty in distinguishing early thermogenic methane from microbial methane using clumped isotopes, one could still recognize those early thermogenic gases by $\delta^{13}C$ of C_1-C_5 gas compounds, which generally plot in straight lines on the 'Chung plot' (Chung et al., 1988), or by the presence of higher alkanes, like C_{4+} , which are not recognized as possibly produced by microbial metabolisms (Milkov and Etiope, 2018).

More generally, the combination of our findings with those from prior studies suggests a common theme to the isotopic structures of methanes formed by important natural processes: If the methane generation pathway uses methyl as a precursor and is irreversible, $\Delta^{13}CH_3D$ of methane will partially inherit the initial $\Delta^{13}CH_2D$ of the methyl group, while $\Delta^{12}CH_2D_2$ will be controlled by the difference in δD between the methyl group and the capping hydrogen; if δD of the methyl group and the capping hydrogen are very close, $\Delta^{12}CH_2D_2$ will inherit the initial $\Delta^{12}CHD_2$ value of the methyl group, whereas when these two pools differ in δD, large $\Delta^{12}CH_2D_2$ deficits will result. This model for the interpretation of methane clumped isotope compositions suggests that we should prioritize future studies of the clumped isotope compositions of methyl groups in potential methane precursors (Lloyd et al., 2020).

Last but not least, hydrogen exchange reactions have been proposed or predicted to cause non-equilibrium signatures in clumped isotope compositions of methane in this study, Xia and Gao (2019), Labidi et al. (2020) and Giunta et al. (2021). These discussions would fundamentally benefit from systematic experimental investigation on the kinetics of hydrogen exchange among isotopologues of methane, and between methane and other hydrogen-bearing substrates (e.g., water and hydrocarbons).

ACKNOWLEDGEMENT

This work was supported by Caltech, ExxonMobil Upstream Integrated Solutions (formerly Upstream Research Company), the Caltech Joint Industry Partnership for Petroleum Geochemistry, Thermo Fischer Scientific, and the Department of Energy BES program. We thank Nami Kitchen, Nithya Thiagarajan, Yanhua Shuai, Newton Nguyen, and David Wang for helpful discussions. We thank Xinyu Xia and two anonymous reviewers for helpful comments.

REFERENCES

Ash J. L., Egger M., Treude T., Kohl I., Cragg B., Parkes R. J., Slomp C. P., Sherwood Lollar B. and Young E. D. (2019) Exchange catalysis during anaerobic methanotrophy revealed by 12CH2D2 and 13CH3D in methane. *Geochemical Perspectives Letters*, 26–30.

Bajnai D., Guo W., Spötl C., Coplen T. B., Methner K., Löffler N., Krsnik E., Gischler E., Hansen M., Henkel D., Price G. D., Raddatz J., Scholz D. and Fiebig J. (2020) Dual clumped isotope thermometry resolves kinetic biases in carbonate formation temperatures. *Nature Communications* **11**.

Barker J. F. and Pollock S. J. (1984) The geochemistry and origin of natural gases in Southern Ontario. *Bull. Can. Pet. Geol.* **32**, 313–326.

Bigeleisen J. and Wolfsberg M. (1957) Theoretical and Experimental Aspects of Isotope Effects in Chemical Kinetics. In *Advances in Chemical Physics* (eds. I. Prigogine and P. Debye). John Wiley & Sons, Inc., Hoboken, NJ, USA. pp. 15–76.

Billault I., Guiet S., Mabon F. and Robins R. (2001) Natural Deuterium Distribution in Long-Chain Fatty Acids Is Nonstatistical: A Site-Specific Study by Quantitative2H NMR Spectroscopy. *ChemBioChem* **2**, 425–431.

Burnham A. K. (2019) Kinetic models of vitrinite, kerogen, and bitumen reflectance. *Organic Geochemistry* **131**, 50–59.

Canadell E. and Olivella S. (1984) Theoretical analysis of kinetic isotope effects. Deuterium isotope effects in the reaction of methyl with methane. *The Journal of Physical Chemistry* **88**, 3545–3549.

Cao X., Bao H. and Peng Y. (2019) A kinetic model for isotopologue signatures of methane generated by biotic and abiotic CO2 methanation. *Geochimica et Cosmochimica Acta* **249**, 59–75.

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

Chung H. M., Rooney M. A., Toon M. B. and Claypool G. E. (1992) Carbon Isotope Composition of Marine Crude Oils. *AAPG Bulletin* **76**.

Clog M., Lawson M., Peterson B., Ferreira A. A., Santos Neto E. V. and Eiler J. M. (2018) A reconnaissance study of 13C–13C clumping in ethane from natural gas. *Geochimica et Cosmochimica Acta* **223**, 229–244.

Douglas P. M. J., Stolper D. A., Eiler J. M., Sessions A. L., Lawson M., Shuai Y., Bishop A., Podlaha O. G., Ferreira A. A., Santos Neto E. V., Niemann M., Steen A. S., Huang L., Chimiak L., Valentine D. L., Fiebig J., Luhmann A. J., Seyfried W. E., Etiope G., Schoell M., Inskeep W. P., Moran J. J. and Kitchen N. (2017) Methane clumped isotopes: Progress and potential for a new isotopic tracer. *Organic Geochemistry* **113**, 262–282.

Douglas P. M. J., Stolper D. A., Smith D. A., Walter Anthony K. M., Paull C. K., Dallimore S., Wik M., Crill P. M., Winterdahl M., Eiler J. M. and Sessions A. L. (2016) Diverse origins of Arctic and Subarctic methane point source emissions identified with multiply-substituted isotopologues. *Geochimica et Cosmochimica Acta* **188**, 163–188.

Eldridge D. L., Korol R., Lloyd M. K., Turner A. C., Webb M. A., Miller T. F. and Stolper D. A. (2019a) Comparison of Experimental vs Theoretical Abundances of $^{13}CH_3D$ and $^{12}CH_2D_2$ for Isotopically Equilibrated Systems from 1 to 500 °C. *ACS Earth and Space Chemistry* **3**, 2747–2764.

Eldridge D. L., Lloyd M. K. and Stolper D. A. (2019b) Methane Clumped Isotope Compositions From Ethane Pyrolysis Experiments. *Goldschmidt Abstracts* **2019**.

Etiope G. and Sherwood Lollar B. (2013) ABIOTIC METHANE ON EARTH: ABIOTIC METHANE ON EARTH. *Reviews of Geophysics* **51**, 276–299.

Fiebig J., Bajnai D., Löffler N., Methner K., Krsnik E., Mulch A. and Hofmann S. (2019a) Combined high-precision Δ48 and Δ47 analysis of carbonates. *Chemical Geology* **522**, 186–191.

Fiebig J., Stefánsson A., Ricci A., Tassi F., Viveiros F., Silva S., Lopez T. M., Schreiber C., Hofmann S. and Mountain B. W. (2019b) Abiogenesis not required to explain the origin of volcanic-hydrothermal hydrocarbons. *Geochemical Perspectives Letters*, 23–27.

Giunta T., Labidi J., Kohl I. E., Ruffine L., Donval J. P., Géli L., Çağatay M. N., Lu H. and Young E. D. (2021) Evidence for methane isotopic bond re-ordering in gas reservoirs sourcing cold seeps from the Sea of Marmara. *Earth and Planetary Science Letters* **553**, 116619.

Giunta T., Young E. D., Warr O., Kohl I., Ash J. L., Martini A., Mundle S. O. C., Rumble D., Pérez-Rodríguez I., Wasley M., LaRowe D. E., Gilbert A. and Sherwood Lollar B. (2019) Methane sources and sinks in continental sedimentary systems: New insights from paired clumped isotopologues 13CH3D and 12CH2D2. *Geochimica et Cosmochimica Acta* **245**, 327–351.

Gonzalez Y., Nelson D. D., Shorter J. H., McManus J. B., Dyroff C., Formolo M. J., Wang D. T., Western C. M. and Ono S. (2019) Precise measurements of 12CH2D2 by tunable infrared laser direct absorption spectroscopy. *Analytical Chemistry*.

Gray P., Herod A. A. and Jones A. (1971) Kinetic data for hydrogen and deuterium atom abstraction by methyl and trifluoromethyl radicals in the gaseous phase. *Chemical Reviews* **71**, 247–294.

Greening C., Ahmed F. H., Mohamed A. E., Lee B. M., Pandey G., Warden A. C., Scott C., Oakeshott J. G., Taylor M. C. and Jackson C. J. (2016) Physiology, Biochemistry, and Applications of F_{420} - and F_{o} -Dependent Redox Reactions. *Microbiology and Molecular Biology Reviews* **80**, 451–493.

Gruen D. S., Wang D. T., Könneke M., Topçuoğlu B. D., Stewart L. C., Goldhammer T., Holden J. F., Hinrichs K.-U. and Ono S. (2018) Experimental investigation on the controls of clumped isotopologue and hydrogen isotope ratios in microbial methane. *Geochimica et Cosmochimica Acta* **237**, 339–356.

Helgeson H. C., Richard L., McKenzie W. F., Norton D. L. and Schmitt A. (2009) A chemical and thermodynamic model of oil generation in hydrocarbon source rocks. *Geochimica et Cosmochimica Acta* **73**, 594–695.

Jackśon K. J., Burnham A. K., Braun R. L. and Knauss K. G. (1995) Temperature and pressure dependence of n-hexadecane cracking. *Organic Geochemistry* **23**, 941–953.

Jackson W. M., McNesby J. R. and Darwent B. deB. (1962) Reaction of Methyl- d_3 Radicals with Isobutane, Isobutane-2- *d*, and Propane. *The Journal of Chemical Physics* **37**, 1610–1615.

Jeffrey A. W. A. (1981) Thermal and clay catalysed cracking in the formation of natural gas. Ph.D. dissertation, Texas A&M University.

Jitariu L. C., Jones L. D., Robertson S. H., Pilling M. J. and Hillier I. H. (2003) Thermal Rate Coefficients via Variational Transition State Theory for the Unimolecular Decomposition/Isomerization of 1-Pentyl Radical: Ab Initio and Direct Dynamics Calculations. *The Journal of Physical Chemistry A* **107**, 8607–8617.

Johnston H. S. and Rapp D. (1961) Large Tunnelling Corrections in Chemical Reaction Rates. [1] II. *Journal of the American Chemical Society* **83**, 1–9.

Koepp M. (1978) D/H isotope exchange reaction between petroleum and water: A contributory determinant for D/H-isotope ratios in crude oils? *Short Papers of the Fourth International Conference, Geochronology, Cosmochronology, Isotope Geology 1978* **USGS Open-File Report 78-701 (ed. R. E. Zartman)**, 2.

Labidi J., Young E. D., Giunta T., Kohl I. E., Seewald J., Tang H., Lilley M. D. and Früh-Green G. L. (2020) Methane thermometry in deep-sea hydrothermal systems: Evidence for re-ordering of doubly-substituted isotopologues during fluid cooling. *Geochimica et Cosmochimica Acta* **288**, 248–261.

Lloyd M. K., Eldridge D. L. and Stolper D. A. (2020) Clumped 13CH2D and 12CHD2 compositions of methyl groups from wood and synthetic monomers: methods, experimental and theoretical calibrations, and initial results. *Geochimica et Cosmochimica Acta*.

Lu D. H., Maurice D. and Truhlar D. G. (1990) What is the effect of variational optimization of the transition state on .alpha.-deuterium secondary kinetic isotope effects? A prototype: CD3H + H .dblarw. CD3 + H2. *Journal of the American Chemical Society* **112**, 6206–6214.

Mango F. D., Jarvie D. and Herriman E. (2009) Natural gas at thermodynamic equilibrium Implications for the origin of natural gas. *Geochemical Transactions* **10**.

Martin P. E. (2020) Detection and Analysis of Martian Low-Temperature Geochemistry. Ph.D. dissertation, California Institute of Technology. Available at: https://resolver.caltech.edu/CaltechTHESIS:12112019-142555725.

Martini A. M., Budai J. M., Walter L. M. and Schoell M. (1996) Microbial generation of economic accumulations of methane within a shallow organic-rich shale. *Nature* **383**, 155–158.

Milkov A. V. and Etiope G. (2018) Revised genetic diagrams for natural gases based on a global dataset of >20,000 samples. *Organic Geochemistry* **125**, 109–120.

Miller H. M., Chaudhry N., Conrad M. E., Bill M., Kopf S. H. and Templeton A. S. (2018) Large carbon isotope variability during methanogenesis under alkaline conditions. *Geochimica et Cosmochimica Acta* **237**, 18–31.

Neubauer C., Crémière A., Wang X. T., Thiagarajan N., Sessions A. L., Adkins J. F., Dalleska N. F., Turchyn A. V., Clegg J. A., Moradian A., Sweredoski M. J., Garbis S. D. and Eiler J. M. (2020) Stable Isotope Analysis of Intact Oxyanions Using Electrospray Quadrupole-Orbitrap Mass Spectrometry. *Analytical Chemistry*. Available at: https://pubs.acs.org/doi/abs/10.1021/acs.analchem.9b04486 [Accessed February 15, 2020].

Ni Y., Ma Q., Ellis G. S., Dai J., Katz B., Zhang S. and Tang Y. (2011) Fundamental studies on kinetic isotope effect (KIE) of hydrogen isotope fractionation in natural gas systems. *Geochimica et Cosmochimica Acta* **75**, 2696–2707.

Ono S., Rhim J. H., Gruen D. S., Taubner H., Kölling M. and Wegener G. (2021) Clumped isotopologue fractionation by microbial cultures performing the anaerobic oxidation of methane. *Geochimica et Cosmochimica Acta* **293**, 70–85.

Ono S., Wang D. T., Gruen D. S., Sherwood Lollar B., Zahniser M. S., McManus B. J. and Nelson D. D. (2014) Measurement of a Doubly Substituted Methane Isotopologue, $^{13}CH_3D$, by Tunable Infrared Laser Direct Absorption Spectroscopy. *Analytical Chemistry* **86**, 6487–6494.

Peters K. E., Walters C. C. and Moldowan J. M. (2004) *The Biomarker Guide*. 2nd ed., Cambridge University Press.

Peterson B. K., Formolo M. J. and Lawson M. (2018) Molecular and detailed isotopic structures of petroleum: Kinetic Monte Carlo analysis of alkane cracking. *Geochimica et Cosmochimica Acta* **243**, 169–185.

Piasecki A., Sessions A., Lawson M., Ferreira A. A., Neto E. V. S. and Eiler J. M. (2016a) 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., Peterson B. and Eiler J. (2016b) Prediction of equilibrium distributions of isotopologues for methane, ethane and propane using density functional theory. *Geochimica et Cosmochimica Acta* **190**, 1–12.

Popa M. E., Paul D., Janssen C. and Röckmann T. (2019) H 2 clumped isotope measurements at natural isotopic abundances. *Rapid Communications in Mass Spectrometry* **33**, 239–251.

Reeves E. P. and Fiebig J. (2020) Abiotic Synthesis of Methane and Organic Compounds in Earth's Lithosphere. *Elements* **16**, 25–31.

Reeves E. P., Seewald J. S. and Sylva S. P. (2012) Hydrogen isotope exchange between n-alkanes and water under hydrothermal conditions. *Geochimica et Cosmochimica Acta* **77**, 582–599.

Röckmann T., Popa M. E., Krol M. C. and Hofmann M. E. G. (2016) Statistical clumped isotope signatures. *Scientific Reports* **6**.

Rooney M. A., Claypool G. E. and Moses Chung H. (1995) Modeling thermogenic gas generation using carbon isotope ratios of natural gas hydrocarbons. *Chemical Geology* **126**, 219–232.

Sackett W. M. (1978) Carbon and hydrogen isotope effects during the thermocatalytic production of hydrocarbons in laboratory simulation experiments. *Geochimica et Cosmochimica Acta* **42**, 571–580.

Salomon M. (1964) ISOTOPE EFFECTS IN METHYL RADICAL ABSTRACTION REACTIONS. *Canadian Journal of Chemistry* **42**, 610–613.

Scheller S., Goenrich M., Thauer R. K. and Jaun B. (2013) Methyl-Coenzyme M Reductase from Methanogenic Archaea: Isotope Effects on the Formation and Anaerobic Oxidation of Methane. *Journal of the American Chemical Society* **135**, 14975–14984.

Schimmelmann A., Sessions A. L. and Mastalerz M. (2006) HYDROGEN ISOTOPIC (D/H) COMPOSITION OF ORGANIC MATTER DURING DIAGENESIS AND THERMAL MATURATION. *Annual Review of Earth and Planetary Sciences* **34**, 501–533.

Schoell M. (1984) Recent advances in petroleum isotope geochemistry. *Organic Geochemistry* **6**, 645–663.

Seewald J. S. (1994) Evidence for metastable equilibrium between hydrocarbons under hydrothermal conditions. *Nature* **370**, 285–287.

Sessions A. L. (2016) Factors controlling the deuterium contents of sedimentary hydrocarbons. *Organic Geochemistry* **96**, 43–64.

Sessions A. L., Sylva S. P., Summons R. E. and Hayes J. M. (2004) Isotopic exchange of carbon-bound hydrogen over geologic timescales 1 1Associate editor: J. Horita. *Geochimica et Cosmochimica Acta* **68**, 1545–1559.

Sherwood Lollar B., Weise S. M., Frape S. K. and Barker J. F. (1994) Isotopic constraints on the migration of hydrocarbon and helium gases of southwestern Ontario. *Bull. Can. Pet. Geol.* **42**, 283–295.

Shuai Y., Douglas P. M. J., Zhang S., Stolper D. A., Ellis G. S., Lawson M., Lewan M. D., Formolo M., Mi J., He K., Hu G. and Eiler J. M. (2018) Equilibrium and non-equilibrium controls on the abundances of clumped isotopologues of methane during thermogenic formation in laboratory experiments: Implications for the chemistry of pyrolysis and the origins of natural gases. *Geochimica et Cosmochimica Acta* **223**, 159–174.

Singleton D. A. and Thomas A. A. (1995) High-Precision Simultaneous Determination of Multiple Small Kinetic Isotope Effects at Natural Abundance. *Journal of the American Chemical Society* **117**, 9357–9358.

Stolper D. A., Lawson M., Davis C. L., Ferreira A. A., Neto E. V. S., Ellis G. S., Lewan M. D., Martini A. M., Tang Y., Schoell M., Sessions A. L. and Eiler J. M. (2014) Formation temperatures of thermogenic and biogenic methane. *Science* **344**, 1500–1503.

Stolper D. A., Lawson M., Formolo M. J., Davis C. L., Douglas P. M. J. and Eiler J. M. (2018) The utility of methane clumped isotopes to constrain the origins of methane in natural gas accumulations. *Geological Society, London, Special Publications* **468**, 23–52.

Stolper D. A., Martini A. M., Clog M., Douglas P. M., Shusta S. S., Valentine D. L., Sessions A. L. and Eiler J. M. (2015) Distinguishing and understanding thermogenic and biogenic sources of methane using multiply substituted isotopologues. *Geochimica et Cosmochimica Acta* **161**, 219–247.

Stolper D.A., Sessions A. L., Ferreira A. A., Santos Neto E. V., Schimmelmann A., Shusta S. S., Valentine D. L. and Eiler J. M. (2014) Combined 13C–D and D–D clumping in methane: Methods and preliminary results. *Geochimica et Cosmochimica Acta* **126**, 169–191.

Sweeney J. J. and Burnham A. K. (1990) Evaluation of a simple model of vitrinite reflectance based on chemical kinetics. *AAPG Bulletin* **74**, 1559–1570.

Taenzer L., Labidi J., Masterson A. L., Feng X., Rumble D., Young E. D. and Leavitt W. D. (2020) Low Δ 12 CH 2 D 2 values in microbialgenic methane result from combinatorial isotope effects. *Geochimica et Cosmochimica Acta* **285**, 225–236.

Tang Y., Huang Y., Ellis G. S., Wang Y., Kralert P. G., Gillaizeau B., Ma Q. and Hwang R. (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.

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

Thiagarajan N., Kitchen N., Xie H., Ponton C., Lawson M., Formolo M. and Eiler J. (2020a) Identifying thermogenic and microbial methane in deep water Gulf of Mexico Reservoirs. *Geochimica et Cosmochimica Acta* **275**, 188–208.

Thiagarajan N., Xie H., Ponton C., Kitchen N., Peterson B., Lawson M., Formolo M., Xiao Y. and Eiler J. (2020b) Isotopic evidence for quasi-equilibrium chemistry in thermally mature natural gases. *Proceedings of the National Academy of Sciences* **117**, 3989–3995.

Tissot B. P. and Welte D. H. (1984) *Petroleum formation and occurrence*. 2nd, rev.enl. ed ed., Springer-Verlag, Berlin ; New York.

Ueno Y., Katsuta T., Ishimaru T. and Yoshida N. (2019) A new method for measuring 34S-18O clumping of sulfate. *Goldschmidt Abstracts* **2019**.

Ungerer P., Behar F., Villalba M., Heum O. R. and Audibert A. (1988) Kinetic modelling of oil cracking. *Organic Geochemistry* **13**, 857–868.

Valentine D. L. (2011) Emerging Topics in Marine Methane Biogeochemistry. *Annual Review of Marine Science* **3**, 147–171.

Valentine D. L., Chidthaisong A., Rice A., Reeburgh W. S. and Tyler S. C. (2004) Carbon and hydrogen isotope fractionation by moderately thermophilic methanogens 1

1Associate editor: N. E. Ostrom. *Geochimica et Cosmochimica Acta* **68**, 1571–1590.

Wang D. T., Gruen D. S., Lollar B. S., Hinrichs K.-U., Stewart L. C., Holden J. F., Hristov A. N., Pohlman J. W., Morrill P. L., Konneke M., Delwiche K. B., Reeves E. P., Sutcliffe C. N., Ritter D. J., Seewald J. S., McIntosh J. C., Hemond H. F., Kubo M. D., Cardace D., Hoehler T. M. and Ono S. (2015) Nonequilibrium clumped isotope signals in microbial methane. *Science* **348**, 428–431.

Wang D. T., Reeves E. P., McDermott J. M., Seewald J. S. and Ono S. (2018) Clumped isotopologue constraints on the origin of methane at seafloor hot springs. *Geochimica et Cosmochimica Acta* **223**, 141–158.

Wang D. T., Welander P. V. and Ono S. (2016) Fractionation of the methane isotopologues 13 CH 4 , 12 CH 3 D, and 13 CH 3 D during aerobic oxidation of methane by Methylococcus capsulatus (Bath). *Geochimica et Cosmochimica Acta* **192**, 186–202.

Wang Y., Sessions A. L., Nielsen R. J. and Goddard W. A. (2009) Equilibrium 2H/1H fractionations in organic molecules. II: Linear alkanes, alkenes, ketones, carboxylic acids, esters, alcohols and ethers. *Geochimica et Cosmochimica Acta* **73**, 7076–7086.

Wang Y., Zhao Y., Liang C., Chen Y., Zhang Q. and Li X. (2017) Molecular-level modeling investigation of n-decane pyrolysis at high temperature. *Journal of Analytical and Applied Pyrolysis* **128**, 412–422.

Wei L., Gao Z., Mastalerz M., Schimmelmann A., Gao L., Wang X., Liu X., Wang Y. and Qiu Z. (2019) Influence of water hydrogen on the hydrogen stable isotope ratio of methane at low versus high temperatures of methanogenesis. *Organic Geochemistry* **128**, 137–147.

Whiticar M. J. (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology* **161**, 291–314.

Xia X. and Gao Y. (2019) Kinetic clumped isotope fractionation during the thermal generation and hydrogen exchange of methane. *Geochimica et Cosmochimica Acta* **248**, 252–273.

Xiao Y. (2001) Modeling the Kinetics and Mechanisms of Petroleum and Natural Gas Generation: A First Principles Approach. *Reviews in Mineralogy and Geochemistry* **42**, 383–436.

Xie H., Dong G., Thiagarajan N., Shuai Y., Mangenot X., Formolo M. J., Lawson M. and Eiler J. M. (2019) Methane Clumped Isotopologues With High-resolution Gas Source Isotope Ratio Mass Spectrometry. In *AGU Fall Meeting Abstracts* pp. V14B-05.

Xie H., Ponton C., Formolo M. J., Lawson M., Ellis G. S., Lewan M. D., Ferreira A. A., Morais E. T., Spigolon A. L. D., Sessions A. L. and Eiler J. M. (2020) Position-specific distribution of hydrogen isotopes in natural propane: Effects of thermal cracking, equilibration and biodegradation. *Geochimica et Cosmochimica Acta* **290**, 235–256.

Xie H., Ponton C., Formolo M. J., Lawson M., Peterson B. K., Lloyd M. K., Sessions A. L. and Eiler J. M. (2018) Position-specific hydrogen isotope equilibrium in propane. *Geochimica et Cosmochimica Acta* **238**, 193–207.

Yeung L. Y. (2016) Combinatorial effects on clumped isotopes and their significance in biogeochemistry. *Geochimica et Cosmochimica Acta* **172**, 22–38.

Yi J., Jun L., Jing-Bo W., Jian-Li W., Quan Z., Yao-Qiang C., Xiang-Yuan L., 四川大学化学学院, 成都 610065;,College of Chemistry, Sichuan University, Chengdu 610065, P. R. China;, and 四川大学化工学院, 成都 610065,College of Chemical Engineering, Sichuan University, Chengdu 610065, P. R. China (2011) Experiment and Kinetics Simulation on the Pyrolysis of n-Decane. *Acta Physico-Chimica Sinica* **27**, 1061–1067.

Yoshinaga M. Y., Holler T., Goldhammer T., Wegener G., Pohlman J. W., Brunner B., Kuypers M. M. M., Hinrichs K.-U. and Elvert M. (2014) Carbon isotope equilibration during sulphate-limited anaerobic oxidation of methane. *Nature Geoscience* **7**, 190–194.

Young E. D. (2019) A Two-Dimensional Perspective on CH4 Isotope Clumping: Distinguishing Process from Source. In *Deep Carbon* (eds. B. N. Orcutt, I. Daniel, and R. Dasgupta). Cambridge University Press. pp. 388–414.

Young E. D., Kohl I. E., Lollar B. S., Etiope G., Rumble D., Li (李姝宁) S., Haghnegahdar M. A., Schauble E. A., McCain K. A., Foustoukos D. I., Sutclife C., Warr O., Ballentine C. J., Onstott T. C., Hosgormez H., Neubeck A., Marques J. M., Pérez-Rodríguez I., Rowe A. R., LaRowe D. E., Magnabosco C., Yeung L. Y., Ash J. L. and Bryndzia L. T. (2017) The relative abundances of resolved l2 CH 2 D 2 and 13 CH 3 D and mechanisms controlling isotopic bond ordering in abiotic and biotic methane gases. *Geochimica et Cosmochimica Acta* **203**, 235–264.

6. Appendix

6.1. Derivation of rate constants of exchange reactions (RA9–RA20) in Table A1

Each of the exchange reactions (RA9-RA20) includes more than one elementary steps. As a kinetic treatment, the bulk rate constant can be regarded as the product of the rate constants of two rate determining steps (breaking carbon-hydrogen bond(s) in methane isotopologues and, if applicable, breaking the B-hydrogen bond in hydrogen bearing compound BH). In this way, the symmetry number of the overall exchange reaction can be treated as the product of the symmetry numbers of the above two rate determining steps (Table A1), where the symmetry number of an elementary step is the quotient between the product of rotational symmetry numbers of reactants and that of the transition state (Xia and Gao, 2019).

We treat RA9 and RA16 as reference reactions.

$$r_9 = 4\times k_{CH4\text{-}BH}[^{12}CH_4][BH] \quad (A1)$$

$$r_{16} = 16\times k_{CH4\text{-}CH4}[^{12}CH_4][^{12}CH_4] \quad (A2)$$

where r_i is the rate of reaction RAi, 4, 16, and $k_{CH4\text{-}BH}$ and $k_{CH4\text{-}CH4}$ are the symmetry numbers and the rate constants of RA9 and RA16, respectively.

RA10 has deuterium substitution in reactant B-D and its kinetic isotope effect (KIE) is $^{D}KIE_{PRI}$ (as defined in Section 4.2.1.2). So

$$r_{10} = 4\times k_{CH4\text{-}BH}\times {}^{D}KIE_{PRI}\,[^{12}CH_4][BD] \quad (A3)$$

RA11 and RA12, respectively, have additional ^{13}C and D substitution in reactant methane and therefore have primary ^{13}C KIE ($^{13C}KIE_{PRI}^{C\text{-}H}$) and secondary deuterium KIE ($^{D}KIE_{SEC}^{C\text{-}H}$) on breaking a carbon-hydrogen bond in methane. And thus their rate constants in Table A1.

RA13 is the reverse reaction of RA10. When RA10 and RA13 reach equilibrium,

$$r_{10} = r_{13} \tag{A4}$$

And the equilibrium constant of this pair of reactions is

$$K_{eq10-13} = \frac{[^{12}\mathrm{CH_3D}][\mathrm{BH}]}{[^{12}\mathrm{CH_4}][BD]} = \frac{4\,{}^{D}R_{CH4}}{{}^{D}R_{BH}} = 4\,{}^{D}\alpha^{eq}_{CH4-BH} \tag{A5}$$

where ${}^{D}\alpha^{eq}_{CH4\text{-}BH}$ is the equilibrium hydrogen isotope fractionation between methane and BH. Plug equation (A3) and (A5) to (A4), we get:

$$r_{13} = k_{CH4\text{-}BH} \times {}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}[^{12}CH_3D][BH] \tag{A6}$$

Therefore, ${}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}$ is equal to the primary deuterium KIE in breaking a carbon-hydrogen bond in methane.

Similarly, RA14 is the reverse reaction of RA11 and their equilibrium constant is

$$K_{eq11-14} = \frac{[^{13}\mathrm{CH_3D}][\mathrm{BH}]}{[^{13}\mathrm{CH_4}][BD]} = 4\,{}^{D}\alpha^{eq}_{CH4-BH} \cdot K_{13CH3D} \tag{A7}$$

where

$$K_{13CH3D} = \frac{[^{13}\mathrm{CH_3D}][^{12}\mathrm{CH_4}]}{[^{13}\mathrm{CH_4}][^{12}\mathrm{CH_3D}]} \tag{A8}$$

And the equilibrium constant of the pair of reactions RA12 and RA15 is

$$K_{eq12-15} = \frac{[^{12}\mathrm{CH_2D_2}][\mathrm{BH}]}{[^{12}\mathrm{CH_3D}][BD]} = 4\,{}^{D}\alpha^{eq}_{CH4-BH} \cdot K_{12CH2D2} \tag{A9}$$

where

$$K_{12CH2D2} = \frac{[^{12}\mathrm{CH_2D_2}][^{12}\mathrm{CH_4}]}{[^{12}\mathrm{CH_3D}]^2} \quad \text{(A10)}$$

And thus rate constants of RA14 and RA15 in Table A1.

The rate constants and reaction rates of RA17–RA20 can be calculated similarly.

6.2. Model setup for the formation of natural thermogenic methane

We extrapolate the cracking-exchange model with 14 isotopic species and 20 isotope-specific reactions (Table A1) to geological times and temperatures. We model a temperature range of 50 to 200°C at a heating rate of 3°C/Myr (and 1 and 10°C/Myr in Fig S3). We assume the normalized concentrations of precursors ACH_3 and BH are 1 and 100, respectively (BH concentration as 10 and 1000 in figure S4). We model a range of $\delta^{13}C$ values of initial organic precursor A-CH_3 (-20, -25, -30‰) based on the natural variation in kerogens and oils (Schoell, 1984; Chung et al., 1992) (Figure 2.8). As for hydrogen isotope composition of ACH_3 and BH, conventional wisdom holds that hydrogen from water frequently participates in the formation of hydrocarbons (Lewan, 1997; Seewald et al., 1998; Schimmelmann et al., 1999; Schimmelmann et al., 2006). But recent low-temperature and long-term heating experiments of shales and coals showed that hydrogen transferred from water to methane only contributes to ~1 to 13% of organic hydrogen at T≤200°C (Wei et al., 2019). Acknowledging the complexity and uncertainty of the hydrogen source in the geological settings, we use the principle of parsimony and assume the methyl precursor A-CH_3 and the hydrogen-bearing compound B-H have the same initial δD value, -100‰ (and -80 and -120‰ in figure S2) (Schoell, 1984).

For isotope effects in the model, we utilize predicted Arrhenius parameters for $^{13C}KIE_{PRI}$ and $^{D}KIE_{SEC}$ (Ni et al., 2011). And we approximate $^{13C}KIE_{PRI}{}^{C\text{-}H}$ and $^{D}KIE_{SEC}{}^{C\text{-}H}$ with $^{13C}KIE_{PRI}$ and $^{D}KIE_{SEC}$, respectively. We calculate the average Arrhenius parameters for $^{D}KIE_{PRI}$ of abstraction on carbon, oxygen and nitrogen-bound hydrogen by methyl radical (Gray et al., 1971) (Table 2.6). Our pyrolysis experiment results suggest that the clumped isotope effect

in the methane generation step is small, so we assume the KIE for multiply-substituted species is the product of corresponding singly-substituted species. We calculate the equilibrium hydrogen isotope fractionation $^{D}\alpha^{eq}_{CH4\text{-}BH}$ by U-B-M reduced partition function theory using vibrational frequencies of methane and the methylene group in propane (Piasecki et al., 2016).

We use the kinetic model for methane generation from Xia and Gao (2019). We assume the activation energy of methane-BH exchange reactions is normally distributed. We parameterize the normal distribution with the experimentally determined activation energies of hydrogen exchange between methane and water with and without the presence of metal oxide catalyst (Metcalfe, 1973; Koepp, 1978) (Table 2.6). We utilize the activation energy of the propagation step in the hydrogen isotope exchange network (Xia and Gao, 2019) for our methane-methane exchange reaction (RA16). We fit the pre-exponential factors of reactions RA9 and RA16 (Table A1) and initial $\Delta^{13}CH_2D$ of A-CH_3 to the isotopic properties of methanes from Southwest Ontario and the Michigan Basin (except Devonian formations) (Giunta et al., 2019) and from Marcellus Shale and Utica Shale (Young et al., 2017) (Table 2.6, Figure 2.8). The model parameters predict a half-life for the hydrogen exchange between methane and B-H that is consistent with experimental constraints on that between methane and water (Koepp, 1978). The equivalent thermal maturity (Easy%RoV) of modeled time-temperature was calculated following Sweeney and Burnham (1990) and Burnham (2019).

Table A1. reactions in the cracking and cracking-exchange models

reaction number	reaction equation	symmetry number (σ_r)	rate constants [ab]
methane generation reactions			
RA1	$A\,^{12}CH_3 + BH \rightarrow {}^{12}CH_4 + A + B$	1	k_{CH4Gen}
RA2	$A\,^{13}CH_3 + BH \rightarrow {}^{13}CH_4 + A + B$	1	$k_{CH4Gen}\times{}^{13C}KIE_{PRI}$
RA3	$A\,^{12}CH_2D + BH \rightarrow {}^{12}CH_3D + A + B$	1	$k_{CH4Gen}\times{}^{D}KIE_{SEC}$
RA4	$A\,^{12}CH_3 + BD \rightarrow {}^{12}CH_3D + A + B$	1	$k_{CH4Gen}\times{}^{D}KIE_{PRI}$
RA5	$A\,^{13}CH_2D + BH \rightarrow {}^{13}CH_3D + A + B$	1	$k_{CH4Gen}\times{}^{13C}KIE_{PRI}\times{}^{D}KIE_{SEC}$
RA6	$A\,^{13}CH_3 + BD \rightarrow {}^{13}CH_3D + A + B$	1	$k_{CH4Gen}\times{}^{13C}KIE_{PRI}\times{}^{D}KIE_{PRI}$
RA7	$A\,^{12}CHD_2 + BH \rightarrow {}^{12}CH_2D_2 + A + B$	1	$k_{CH4Gen}\times{}^{D}KIE_{SEC}\times{}^{D}KIE_{SEC}$
RA8	$A\,^{12}CH_2D + BD \rightarrow {}^{12}CH_2D_2 + A + B$	1	$k_{CH4Gen}\times{}^{D}KIE_{SEC}\times{}^{D}KIE_{PRI}$
methane-BH exchange			
RA9	$^{12}CH_4 + BH \rightarrow {}^{12}CH_4 + BH$	4	$k_{CH4\text{-}BH}$
RA10	$^{12}CH_4 + BD \rightarrow {}^{12}CH_3D + BH$	4	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}$
RA11	$^{13}CH_4 + BD \rightarrow {}^{13}CH_3D + BH$	4	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}\times{}^{13C}KIE_{PRI}{}^{C\text{-}H}$
RA12	$^{12}CH_3D + BD \rightarrow {}^{12}CH_2D_2 + BH$	3	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}\times{}^{D}KIE_{SEC}{}^{C\text{-}H}$
RA13	$^{12}CH_3D + BH \rightarrow {}^{12}CH_4 + BD$	1	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}$ [c]
RA14	$^{13}CH_3D + BH \rightarrow {}^{13}CH_4 + BD$	1	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}\times{}^{13C}KIE_{PRI}{}^{C\text{-}H}/K_{13CH3D}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}$ [d]
RA15	$^{12}CH_2D_2 + BH \rightarrow {}^{12}CH_3D + BD$	2	$k_{CH4\text{-}BH}\times{}^{D}KIE_{PRI}\times{}^{D}KIE_{SEC}{}^{C\text{-}H}/(8/3\times K_{12CH2D2})/{}^{D}\alpha^{eq}_{CH4\text{-}BH}$ [e]
methane-methane exchange			
RA16	$^{12}CH_4 + {}^{12}CH_4 \rightarrow {}^{12}CH_4 + {}^{12}CH_4$	16	$k_{CH4\text{-}CH4}$
RA17	$^{13}CH_4 + {}^{12}CH_3D \rightarrow {}^{13}CH_3D + {}^{12}CH_4$	4	$k_{CH4\text{-}CH4}\times{}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}\times{}^{13C}KIE_{PRI}{}^{C\text{-}H}$
RA18	$^{12}CH_3D + {}^{12}CH_3D \rightarrow {}^{12}CH_2D_2 + {}^{12}CH_4$	3	$k_{CH4\text{-}CH4}\times{}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}\times{}^{D}KIE_{SEC}{}^{C\text{-}H}$
RA19	$^{13}CH_3D + {}^{12}CH_4 \rightarrow {}^{13}CH_4 + {}^{12}CH_3D$	4	$k_{CH4\text{-}CH4}\times{}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}\times{}^{13C}KIE_{PRI}{}^{C\text{-}H}/K_{13CH3D}$
RA20	$^{12}CH_2D_2 + {}^{12}CH_4 \rightarrow {}^{12}CH_3D + {}^{12}CH_3D$	8	$k_{CH4\text{-}CH4}\times{}^{D}KIE_{PRI}/{}^{D}\alpha^{eq}_{CH4\text{-}BH}\times{}^{D}KIE_{SEC}{}^{C\text{-}H}/(8/3\times K_{12CH2D2})$

[a] see also derivations in Appendix A1

[b] KIEs as defined in 4.2.1.2 and 4.2.2.2

[c] ${}^{D}\alpha^{eq}_{CH4\text{-}BH}$ is the equilibrium hydrogen isotope fractionation between methane and B-H

[d] K_{13CH3D} is the equilibrium constant of reaction (R1)

[e] $K_{12CH2D2}$ is the equilibrium constant of reaction (R2)

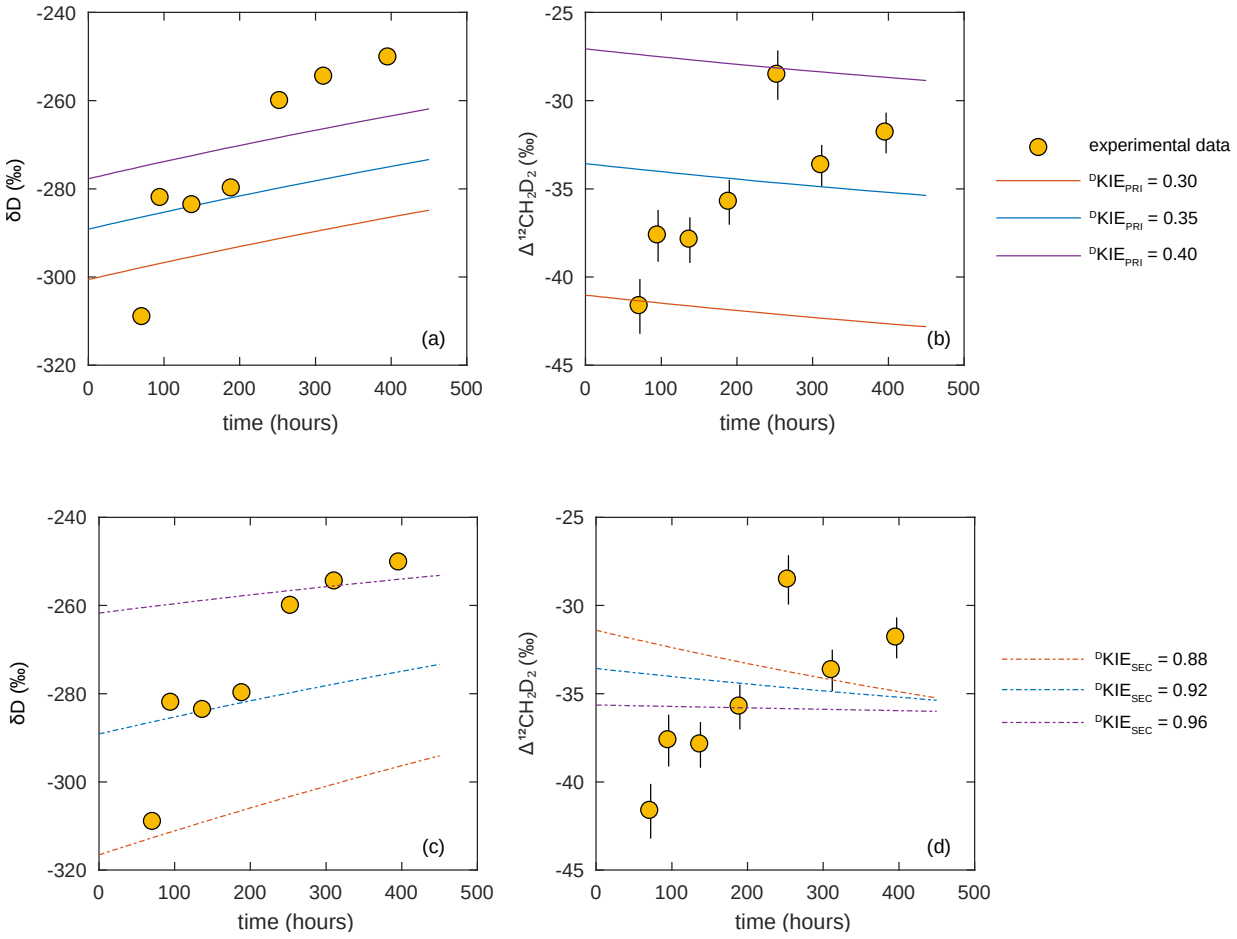


Figure 2.10 Sensitivity tests on the primary deuterium KIE of B-D ($^{D}KIE_{PRI}$), and secondary deuterium KIE of A-$^{12}CH_2D$ ($^{D}KIE_{SEC}$) on the cracking model. The change of δD (a) and $\Delta^{12}CH_2D_2$ (b) when varying $^{D}KIE_{PRI}$ from 0.30 to 0.40 ($^{D}KIE_{SEC}$ = 0.92) and the change of δD (c) and $\Delta^{12}CH_2D_2$ (d) when varying $^{D}KIE_{SEC}$ from 0.88 to 0.96 ($^{D}KIE_{PRI}$ = 0.35) demonstrate that the absolute values of $^{D}KIE_{PRI}$ and $^{D}KIE_{SEC}$ controls δD, and that the relative difference in $^{D}KIE_{PRI}$ and $^{D}KIE_{SEC}$ controls $\Delta^{12}CH_2D_2$. $\delta^{13}C$ and $\Delta^{13}CH_3D$ are not sensitive to $^{D}KIE_{PRI}$ and $^{D}KIE_{SEC}$. Values of other parameters in Table 5.

REFERENCES

Burnham A. K. (2019) Kinetic models of vitrinite, kerogen, and bitumen reflectance. *Organic Geochemistry* **131**, 50–59.

Chung H. M., Rooney M. A., Toon M. B. and Claypool G. E. (1992) Carbon Isotope Composition of Marine Crude Oils. *AAPG Bulletin* **76**.

Giunta T., Young E. D., Warr O., Kohl I., Ash J. L., Martini A., Mundle S. O. C., Rumble D., Pérez-Rodríguez I., Wasley M., LaRowe D. E., Gilbert A. and Sherwood Lollar B. (2019) Methane sources and sinks in continental sedimentary systems: New insights from paired clumped isotopologues 13CH3D and 12CH2D2. *Geochimica et Cosmochimica Acta* **245**, 327–351.

Gray P., Herod A. A. and Jones A. (1971) Kinetic data for hydrogen and deuterium atom abstraction by methyl and trifluoromethyl radicals in the gaseous phase. *Chemical Reviews* **71**, 247–294.

Koepp M. (1978) D/H isotope exchange reaction between petroleum and water: A contributory determinant for D/H-isotope ratios in crude oils? *Short Papers of the Fourth International Conference, Geochronology, Cosmochronology, Isotope Geology 1978* **USGS Open-File Report 78-701 (ed. R. E. Zartman)**, 2.

Metcalfe A. (1973) The exchange of hydrogen between methane and deuterium oxide over nickel. *Journal of Catalysis* **30**, 250–254.

Ni Y., Ma Q., Ellis G. S., Dai J., Katz B., Zhang S. and Tang Y. (2011) Fundamental studies on kinetic isotope effect (KIE) of hydrogen isotope fractionation in natural gas systems. *Geochimica et Cosmochimica Acta* **75**, 2696–2707.

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

Schimmelmann A., Lewan M. D. and 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., Sessions A. L. and Mastalerz M. (2006) HYDROGEN ISOTOPIC (D/H) COMPOSITION OF ORGANIC MATTER DURING DIAGENESIS AND THERMAL MATURATION. *Annual Review of Earth and Planetary Sciences* **34**, 501–533.

Schoell M. (1984) Recent advances in petroleum isotope geochemistry. *Organic Geochemistry* **6**, 645–663.

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

Sweeney J. J. and Burnham A. K. (1990) Evaluation of a simple model of vitrinite reflectance based on chemical kinetics. *AAPG Bulletin* **74**, 1559–1570.

Wei L., Gao Z., Mastalerz M., Schimmelmann A., Gao L., Wang X., Liu X., Wang Y. and Qiu Z. (2019) Influence of water hydrogen on the hydrogen stable isotope ratio of methane at low versus high temperatures of methanogenesis. *Organic Geochemistry* **128**, 137–147.

Xia X. and Gao Y. (2019) Kinetic clumped isotope fractionation during the thermal generation and hydrogen exchange of methane. *Geochimica et Cosmochimica Acta* **248**, 252–273.

Young E. D., Kohl I. E., Lollar B. S., Etiope G., Rumble D., Li (李姝宁) S., Haghnegahdar M. A., Schauble E. A., McCain K. A., Foustoukos D. I., Sutclife C., Warr O., Ballentine C. J., Onstott T. C., Hosgormez H., Neubeck A., Marques J. M., Pérez-Rodríguez I., Rowe A. R., LaRowe D. E., Magnabosco C., Yeung L. Y., Ash J. L. and Bryndzia L. T. (2017) The relative abundances of resolved l2 CH 2 D 2 and 13 CH 3 D and mechanisms controlling isotopic bond ordering in abiotic and biotic methane gases. *Geochimica et Cosmochimica Acta* **203**, 235–264.

Supplementary Materials

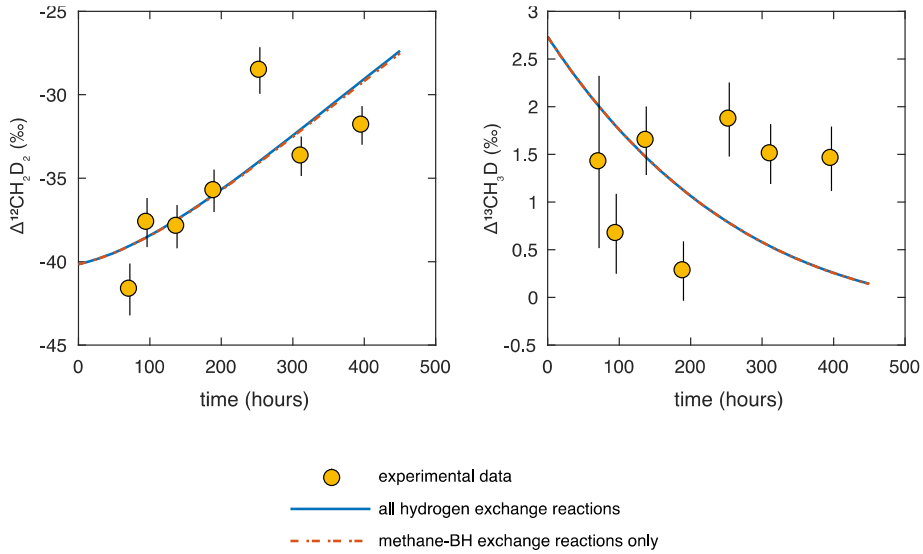


Figure 2.11 Test on methane-methane hydrogen exchange reactions in the cracking-exchange model. There is minimal difference on modeled Δ12CH2D2 and Δ13CH3D with and without hydrogen exchange reactions among methane (rate constant of methane-methane exchange kCH4-CH4 as 8.86×10-8/s and 0, respectively). Values of other parameters in Table 5.

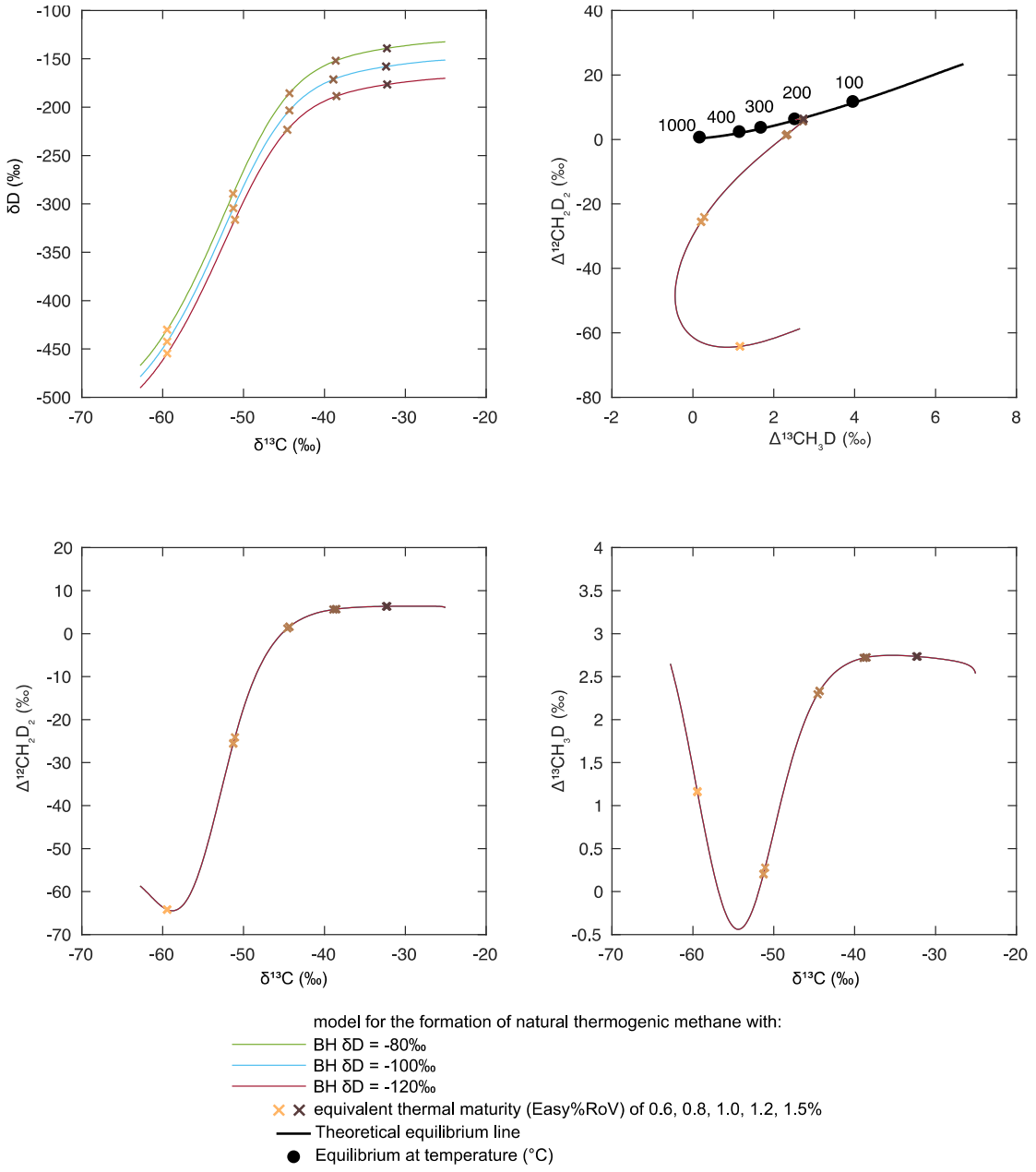


Figure 2.12 Sensitivity tests on the initial δD of organic precursors A-CH3 and B-H in the cracking-exchange model for natural thermogenic methane, with δD of B-H as -80, -100 and -120‰. δD of methane changes proportionally with δD of B-H. δ13C, Δ13CH3D and Δ12CH2D2 are not sensitive. Values of other parameters in Table 2.6.

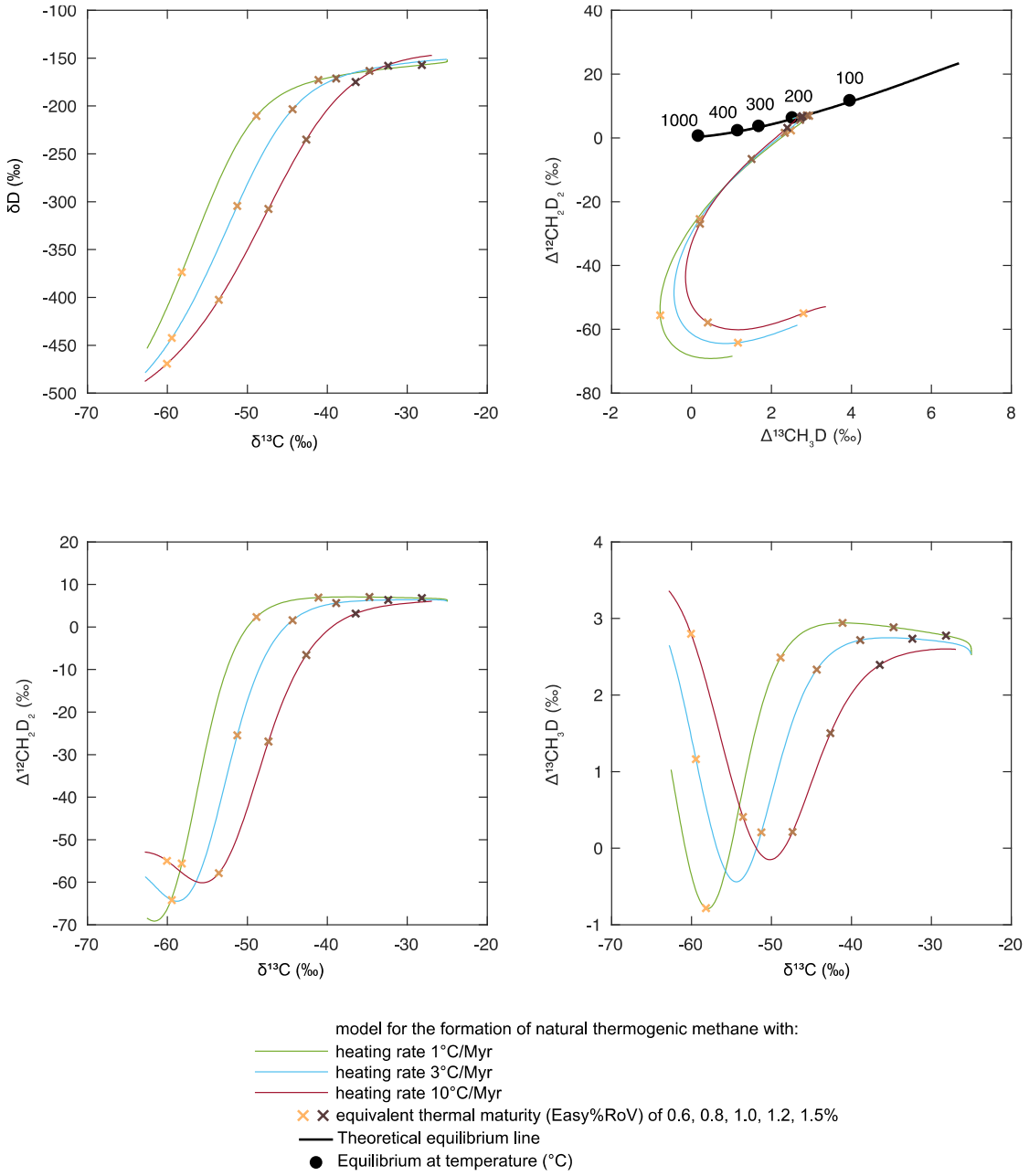


Figure 2.13 Sensitivity tests on the heating rate in the cracking-exchange model for natural thermogenic methane, with heating rate as 1, 3 and 10°C/Myr. At a same thermal maturity, methane produced at a faster heating rate has stronger non-equilibrium signature, whereas methane produced at a slower heating rate is more advances in its progress to approaching thermodynamic equilibrium. The effect of hydrogen exchange is manifest at slower heating rates. The heating rate also affects the half-life of equilibration of methane. Values of other parameters in Table 2.6.

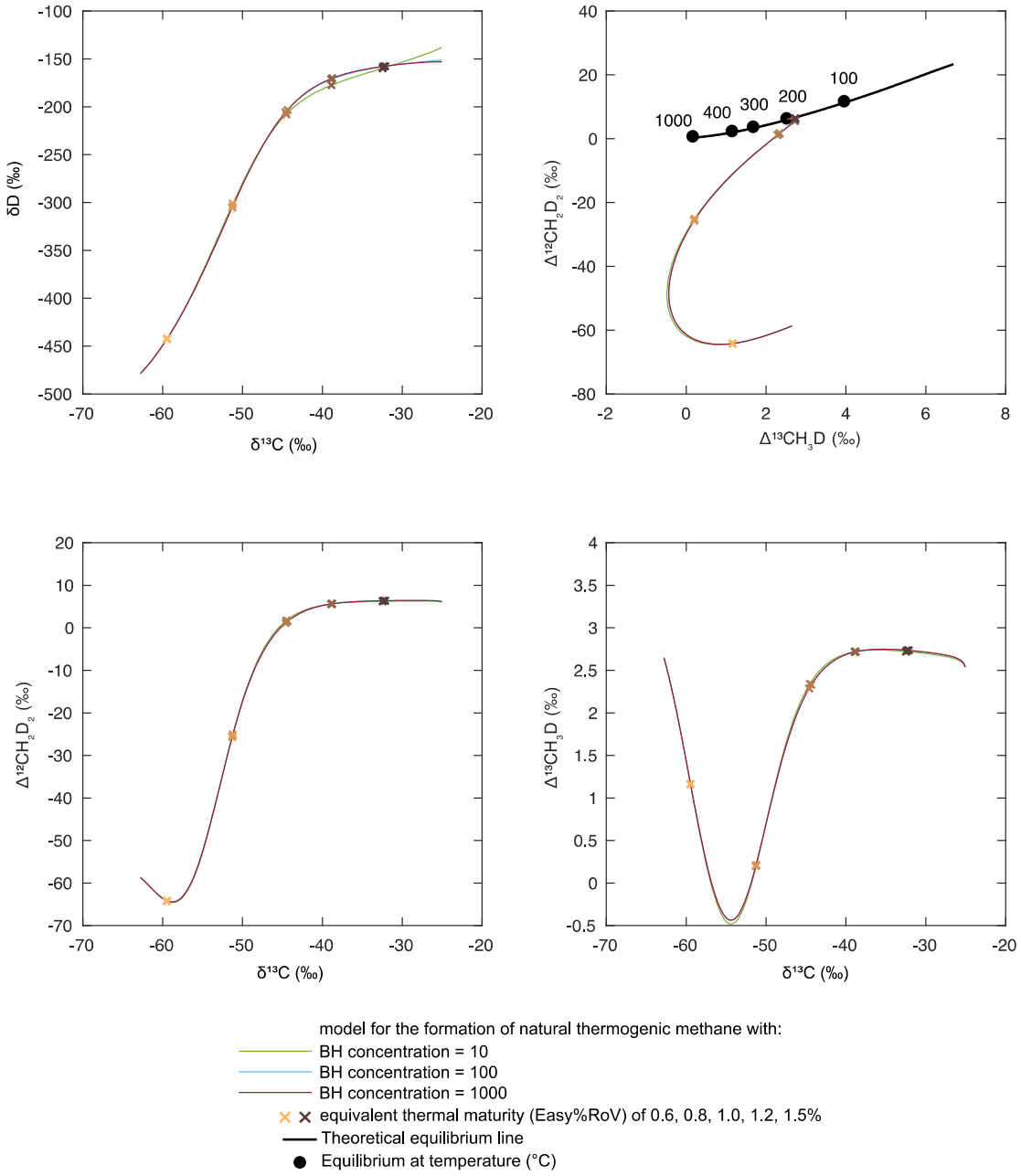


Figure 2.14 Sensitivity tests on the concentration of B-H in the cracking-exchange model for natural thermogenic methane, with concentration of B-H as 10, 100 and 1000. Note that k_{CH4Gen}[BH] and $k_{CH4\text{-}BH}$[BH] are held constant. A reservoir effect on δD of methane becomes obvious when [BH] decreases to 10. $\delta^{13}C$, $\Delta^{13}CH_3D$ and $\Delta^{12}CH_2D_2$ are not sensitive to the concentration of B-H. Values of other parameters in Table 6.

Table S1. Molecular composition of the natural gas working standard

compound	relative concentration (%)
CH_4	82.99
C_2H_6	9.19
C_3H_8	2.06
i-C_4H_{10}	0.71
n-C_4H_{10}	0.56
i-C_5H_{12}	0.28
n-C_5H_{12}	0.17
C_{6+}	0.44
others	3.59

Chapter 3 — Automating sample introduction and inline separation for analysis of molecular isotopic structures by GC-orbitrap mass spectrometry

Guannan Dong, Jonathan Treffkorn, Kenneth Farley, John Eiler

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA

ABSTRACT

Gas chromatography (GC)-Orbitrap mass spectrometry enables analyses of intramolecular isotopic structures (i.e., site-specific isotopic contents and proportions of multiply substituted isotopologues) of a variety of volatile compounds. Many applications require custom-built sample introduction devices and in previous studies these devices have been manually operated for each analysis. High precision (sub-per mil) isotope ratio measurements require replicate analyses under identical instrumental conditions and operational procedures. Therefore, a manually operated component has presented a bottleneck for the development of this mass spectrometry technique. Here we automate the sample introduction and online separation system on a GC-Orbitrap. Analyses can now be scheduled to run automatically for hours to days. And we report for the first time that the external experimental reproducibility over the time of tens of hours for Orbitrap-based GC-irMS measurements (using a 'peak broadener' sample delivery system) is close to the shot-noise-limited internal precision of each analysis.

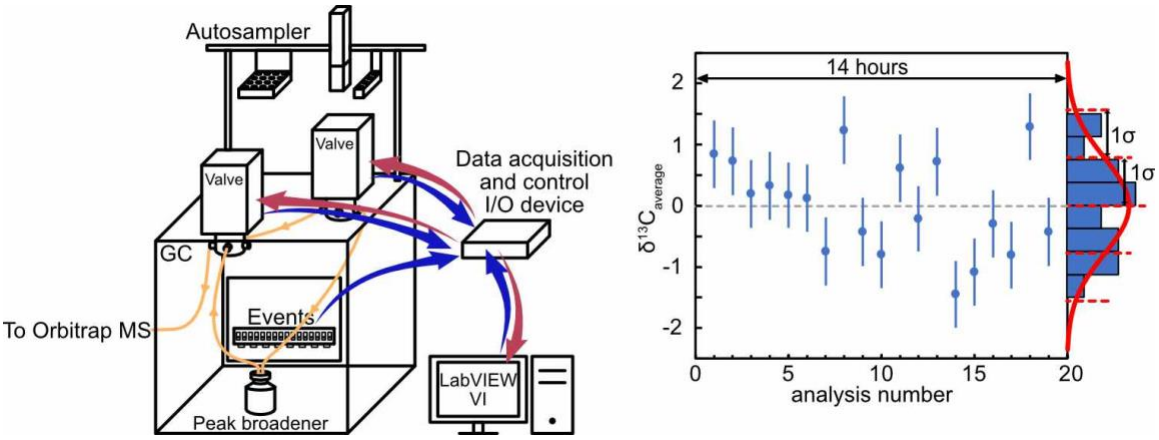


Figure 3.1 graphical abstract

1. Introduction

Variations in the natural abundances of stable isotopes are widely used to fingerprint the sources of chemical substances, and to constrain physical, chemical and biological conditions and processes those substances have experienced, with applications in geological, environmental, biological, archaeological, forensic and cosmological sciences (Valley et al., 2001; Sharp, 2017). The isotopic properties of a bulk material, a specific compound, or a particular position within a molecular structure provide different levels of information on a system, an organism, a pathway, or an elementary reaction (Eiler, 2013; Gilbert, 2021). Recently, Orbitrap-based isotope ratio mass spectrometry (IRMS) (Eiler et al., 2017) has been proven to be useful for determining the compound-specific and intramolecular isotope compositions of amino acids (Neubauer et al., 2018; Chimiak et al., 2021; Wilkes et al., 2022), a carboxylic acid (Mueller et al., 2022), oxyanions (Martin, 2020; Neubauer et al., 2020; Hilkert et al., 2021) and aromatic hydrocarbons (Cesar et al., 2019) and the stable and radioactive isotope compositions of metals (Hoegg et al., 2019a, 2021), and likely can be used for this purpose on other diverse analytes.

In an Orbitrap-based IRMS measurement, an analyte is first ionized in the ion source (possibly accompanied by fragmentation), then the ions are focused, prefiltered, and on many

analytical platforms are then transmitted by a series of ion optics to a quadrupole mass filter. The quadrupole filters the transmitted ions according to their mass to charge ratios, permitting only a specified subset of the full mass range of those ions to be used for further steps in the analysis. The filtered ions are either deflected (i.e., lost) or allowed into a curved linear ion trap referred to as the 'C-trap', controlled by the Automatic Gain Control (AGC) system, which monitors ion current and automatically regulates the period of time during which that current is allowed to pass from the quadrupole to C-trap. The ions held in the C-trap are then injected through a lens system into the Orbitrap analyzer that detects the image current induced by the motion of the ions within. Each 'scan' of an injected packet of ions by the Orbitrap analyzer typically lasts tens to hundreds of milliseconds (higher with increasing mass resolution). Finally, a fast-Fourier-transform analysis of the time-series of the image current (the 'transient') produces a mass spectrum that includes the various isotopically substituted ion peaks of interest, from which isotope ratios can be calculated. A more detailed description can be found in (Eiler et al., 2017).

Among the three different platforms used to-date for Orbitrap-based IRMS (GC-Orbitrap (Eiler et al., 2017), ESI-Orbitrap (Eiler et al., 2017) and LS-APGD-Orbitrap (Hoegg et al., 2016)), GC-Orbitrap is most suitable for analyses of volatile, gas-chromatography (GC)-amenable compounds (including compounds that are rendered GC amenable through derivatization). GC allows for separation of a compound of interest from other compounds (e.g., if the sample is a complex mixture, as is often the case in organic analysis) based on differences in chromatography retention times. A compound typically elutes from the GC column over a period of seconds. But the Orbitrap analyzer has a limit on the number of ions that it can observe in such a short time, because of the need to mitigate space charge effects on each ion packet (limiting and controlling the sizes of these packets is a primary purpose of the AGC system). I.e., only a finite number of scans can occur in a period of a few seconds, and each of those scans is significantly limited in the number of the ions it can contain. In order to resolve this quandary, a custom-built sample introduction system has been developed (Eiler et al., 2017).

This sample delivery system consists of, in addition to the standard components of a commercial, automated gas chromatography system, two four-port Valco valves, a ‘peak broadener’ volume and various connecting capillaries (Figure 3.2). Each of the Valco valves has two positions, and the various ways these valves can be positioned controls the route of gas flow from the GC column to the peak broadener, connecting capillaries, and the ion source of the mass spectrometer. Specifically, in one valve configuration (Figure 3.2a) GC effluent is directed through capillaries connecting the two valves directly to the transfer line and the ion source for analysis, without passing through the peak broadener; this configuration is principally used to precisely gauge elution times, but can be used for low-precision isotopic analysis (with special conditions and data processing requirements) (Zeichner et al., 2022). Alternatively, a second valve configuration (Figure 3.2b) is used for a ‘peak capture’ experiment, whereby a specific, selected GC effluent is directed to and mixed in the peak broadener volume (which is prefilled with helium); a further valve switch to create a third configuration after a compound of interest has been captured (Figure 3.2c) prevents unwanted later-eluting compounds from entering the peak broadener. Subsequently, a flow of pure helium carrier gas purges the peak broadener and delivers the trapped compound of interest through the transfer line for analysis over the relatively long period of the residence time in the peak broadener volume (typically minutes to tens of minutes).

In previous studies using this system, samples have been manually injected into the GC and the Valco valves have been manually operated using handles that protrude from unused, insulated detector/injector ports on the top of GC oven. These operations need to be timed precisely to coordinate with the elution time of compound of interest. And for many applications, in order to achieve high precision (sub-per mil) isotope ratio measurements, these procedures need to be repeated under identical instrumental conditions and operational procedures for multiple replicate analyses, in some cases over periods of hours to tens of hours. This practice has a high demand for the skill and attention of the analyst and introduces the potential for human errors. Therefore, the manual nature of these methods has presented a bottleneck for the development of Orbitrap-based GC-IRMS technique. Recently, the

automatic alternating introduction of sample and reference materials on an ESI-Orbitrap has improved the fidelity of isotope-ratio measurements of nitrate (Hilkert et al., 2021). LabVIEW-based programs have been widely used to control instruments, collect data and automate analytical systems (Wagner et al., 2010), including for a variety of isotope ratio measurements (Farley et al., 1999; Passey et al., 2010; Ono et al., 2014). These prior developments present models for potential automated approaches to Orbitrap GC-IRMS.

In this study, we automate the sample introduction and online separation system on a GC-Orbitrap using LabVIEW virtual instruments (VIs) and a digital input/output (I/O) device to control electric actuators of the Valco valves and an autosampler to inject samples. We also evaluate the performance of the automated system, allowing us to conduct the first study of the experimental reproducibility of Orbitrap-based GC-IRMS over tens of hours, specifically by way of analyses of site-specific carbon and hydrogen isotope compositions of p-xylene.

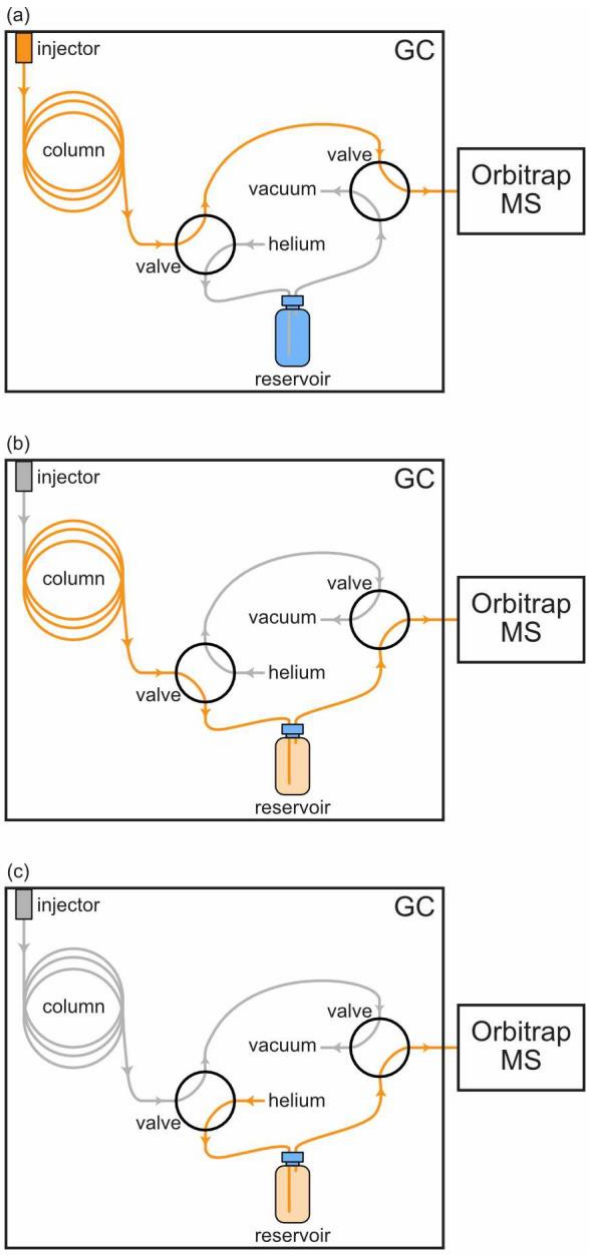


Figure 3.2 Schematic illustration of a custom-built sample delivery system on GC-Orbitrap and its different configurations: (a) GC effluent directed through capillaries connecting two valves directly to the transfer line; (b) GC effluent directed to and mixed in the peak broadener volume; (c) a helium flow purging the peak broadener and the effluent from it directed to the transfer line.

2. Materials and methods

2.1. Instruments and automation methods

The automation of Orbitrap GC-IRMS was done on a Thermo Scientific Q Exactive Orbitrap mass spectrometer with a Trace 1310 GC. We integrated with this base system a Thermo Scientific Triplus RSH autosampler, an Aux Temperature/Cryo Module on the GC, two four-way Valco valves with universal electric actuators (EUH-4N4WT, VICI), a peak broadener volume (as per (Eiler et al., 2017)), a data acquisition and control I/O device (USB-6525, NI) and LabVIEW VIs (available at doi.org/10.22002/p6wqg-g3t24) (Figure 3.3). The Triplus RSH autosampler can perform liquid or headspace injection of the analytical sample to the GC. The Valco valves are driven by electric actuators that protrude from insulated unused detector/injector ports on the top of the GC.

There are two modes to command the positions of the Valco valves: 1. Direct control by toggling switch icons in the front panel of the LabVIEW VI computer interface; 2. Pre-programming run-time events that control valve positions over the course of a GC method execution, i.e., a 'run' (Figure 3.3). In the first mode, when the user toggles a switch icon in the LabVIEW VI interface, the I/O device signals the corresponding electric actuator to move the valve to the designated position. In the second mode, the user programs run-time events in the run table of the GC method so that the 'Events' ports on the Aux Temperature/Cryo Module of GC output signals to the I/O device at certain times during a GC run. The LabVIEW VI program processes the received signal and commands the I/O device to signal the electric actuators to move the valves at the programmed times. The electric actuators also output signals regarding the current valve positions to the I/O device, and the LabVIEW VI program uses these readouts to check that the valves have been moved to the designated positions in both modes. Utilizing the second mode, multiple analyses can be scheduled to run sequentially using Thermo Scientific Xcalibur software. Such automated sequences provide the first means of performing highly reproducible Orbitrap GC-IRMS measurements over periods of hours to days.

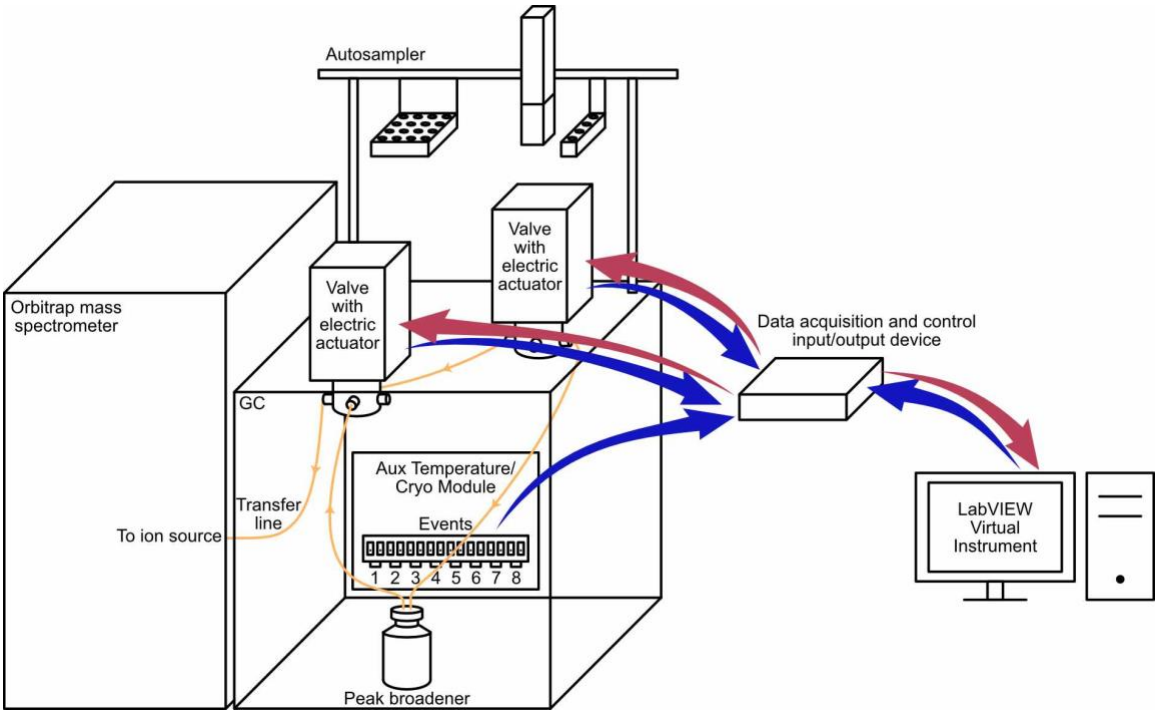


Figure 3.3 Schematic illustration of the automated sample delivery system on GC-Orbitrap and its signal flow. The computer hosting the LabVIEW VI achieves data acquisition and control of the hardware through an input/output (I/O) device. The 'Events' ports on the Aux Temperature/Cryo Module of GC output programmed signals to the I/O device at certain times during a GC run. The I/O device signals the electric actuators to move the valves to certain positions; the electric actuators also output signals regarding the current valve positions to the I/O device. The injector, GC column and helium flow are left out for simplicity.

2.2. Material and method for p-xylene analyses

We analyzed the $C_7H_7^+$ fragment ion of p-xylene (formed by loss of a methyl group during electron-impact ionization, Figure 3.4a) by 'peak capture' experiments (Eiler et al., 2017) to evaluate the performance of the automated sample delivery system and the experimental reproducibility of the Orbitrap-based GC-IRMS over the time of hours. Although our purpose here is to document long-term experimental reproducibility of this system, this measurement could serve as a component of a site-specific C and H isotope analysis (Cesar et al., 2019).

A solution of perfluorotributylamine (PFTBA, Sigma-Aldrich, LOT# 77299) was used to calibrate the mass scale and mass isolation for positive ions and to auto-tune the ion source, injection flatapole and inter-flatapole lens at 70 eV electron energy, maximizing the signal intensity on the Q Exactive GC.

P-xylene (Alfa Aesar, 99% purity, LOT# Q28B033) was diluted in n-hexane to make a 0.02% (v/v) (or 1.62 mM) solution. 1.5 μL of the p-xylene solution was injected by the autosampler to a TG-5SilMS column (30 m × 0.25 mm × 0.25 μm film thickness) through a split/splitless injector held at 250°C and operated with a split ratio of 10, using helium as a carrier gas with a flow rate of 1.2 mL/min. The GC oven was started at 50°C, heated from 50°C to 78°C at 5°C/min and held for 34 min. We used a 10 cm^3 passivated stainless-steel peak broadener (Eiler et al., 2017). The positions of the Valco valves started as in Figure 3.2c, changed to those in Figure 3.2a at 5 min, then to those in the Figure 3.2b at 5.5 min, and back to Figure 3.2c configuration at 6.4 min. The transfer line and the electron-impact ion (EI) source were set at 250°C. The filament was turned on at 5 min. We analyzed positive ions and used a quadrupole scan range of 86 to 97 m/z (no 'lock masses' were used). We used the prescan AGC mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 10 to 200 ms for scans during which sample was present). The nominal mass resolution was set at 120,000 (full-width-half-maximum, FWHM, at m/z = 200) and the number of microscans to 1. We ran the acquisition for 40 min, which we established was the time required for the mass spectra of effluent from the peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., p-xylene-free) measurements. Each analysis, starting from the sample injection and lasting for 40 min, was repeated continuously for 19 times over the time span of 14 hours.

2.3. Data analysis

The Thermo Scientific FTStatistic proprietary software (Eiler et al., 2017; Hilkert et al., 2021) was used for extraction of scan-by-scan signal intensities for mass spectral peaks having mass-to-charge ratios corresponding to the ions of interest and other relevant data

(such as noise intensity for each mass spectral peak) from RAW files. FTStatistic-extracted data are processed and converted to isotope ratios using a python script (available at doi.org/10.22002/d1mnz-0bs34). Briefly, we cull the scan-by-scan data of a 'peak capture' experiment by a threshold of 10% of the maximum intensity of $^{12}C_7H_7^+$ ion (i.e., accept a continuous set of scans for which the ion intensity of $^{12}C_7H_7^+$ consistently exceeds 10% of the maximum of that value during p-xylene elution from the peak broadener) (Figure 3.4b). We then drop any scans for which TIC×IT (total ion current × injection time) is more than three standard deviations away from the mean of the set of scans accepted in the previous step. Next, the scan-by-scan counts of different ions, N_I are calculated by

$$N_I = (S/N)*(C_N/z)*(R_N/R)^{1/2} \quad (1)$$

where S is the reported absolute intensity for the mass-to-charge ratio of the ion of interest (in the arbitrary units of 'NL score' intensity), N is the peak noise specific to that mass-to-charge ratio (also reported in 'NL score' units), both measured at the resolution setting R; C_N is an experimental constant related to the number of charges corresponding to the Orbitrap noise band at a reference resolution R_N (Makarov and Denisov, 2009; Eiler et al., 2017), and z is the charge number of the ion of interest. Finally, we calculate the carbon isotope ratio of the $C_7H_7^+$ fragment by:

$$^{13}R_{C_7H_7^+} = \frac{\frac{\sum_i {}^{13}N_i \times {}^{13}I_i}{\sum_i {}^{13}I_i}}{\frac{\sum_i {}^{12}N_i \times {}^{12}I_i}{\sum_i {}^{12}I_i}} \quad (2)$$

where $^{12}N_i$ and $^{13}N_i$ are the counts of $^{12}C_7H_7^+$ and $^{13}C_7H_7^+$ ions in scan i, respectively, and $^{12}I_i$ and $^{13}I_i$ are the intensities of $^{12}C_7H_7^+$ and $^{13}C_7H_7^+$ ions for scan i. This calculation is essentially a ratio of integrated counts of $^{13}C_7H_7^+$ to $^{12}C_7H_7^+$, weighted by the fraction of the total analyte in each scan (nominally proportional to the signal intensities of ion peaks of interest) eluting during the period in which each scan was made. Importantly, the number of ions that reached the Orbitrap was relatively constant across scans (due to the action of the AGC system), so this weighting mostly reflects differences in the concentration of analyte delivered to the ion

source over time, and the weighting correction is principally intended to compensate for any time-dependent isotopic variations over the course of the sample's elution into the ion source (e.g., due to chromatographic and other isotopic fractionations that might occur during sample transfer). Similarly, the hydrogen isotope ratio can be calculated by

$$^{2}R_{C_7H_7{}^+} = \frac{\frac{\sum_i {}^{2}N_i \times {}^{2}I_i}{\sum_i {}^{2}I_i}}{\frac{\sum_i {}^{1}N_i \times {}^{1}I_i}{\sum_i {}^{1}I_i}} \quad (3)$$

where $^{1}N_i$ and $^{2}N_i$ are the counts of $^{12}C_7H_7{}^+$ and $^{12}C_7H_6D^+$ ions in scan i, respectively, and $^{1}I_i$ and $^{2}I_i$ are the intensities of $^{12}C_7H_7{}^+$ and $^{12}C_7H_6D^+$ ions for scan i. The carbon and hydrogen isotope compositions of the $C_7H_7{}^+$ ion from each analysis are expressed using delta notation relative to the average of the total 19 analyses:

$$\delta^{13}C = \frac{^{13}R_{analysis\,N}}{^{13}R_{average}} - 1 \quad (4)$$

$$\delta D = \frac{^{D}R_{analysis\,N}}{^{D}R_{average}} - 1 \quad (5)$$

3. Results and discussion

The internal precisions of each analysis of p-xylene from one sample injection are 0.53–0.56‰ and 4.9–5.2‰ for carbon and hydrogen isotope compositions of the $C_7H_7{}^+$ fragment ion, respectively (i.e., the relative standard error of all accepted scans in each analysis) (Table 3.1). These internal precisions conform to shot-noise limits. The external relative standard deviations of carbon and hydrogen isotope analyses of the $C_7H_7{}^+$ fragment ions from 19 injections of the p-xylene solution are 0.79 and 6.5‰, respectively (Table 3.1). These values are ~1.3–1.4 times the respective internal precisions. We think this relatively modest discrepancy between standard errors of single analysis and standard deviations of multiple analyses reflects instrumental biases that are common to conventional GC-IRMS or GC-MS platforms (e.g., (Ricci et al., 1994)), involving subtle irreproducibility in complex processes

like sample vaporization in the injector, elution from gas chromatography, ionization and quadrupole mass filtering. Importantly, the measured $\delta^{13}C$ and $\delta^{2}H$ values show no clear systematic drift over the course of 19 analyses in 14 hours and all are within 2–3 respective internal standard errors from each other (Figure 3.5). Additionally, we examine the conformance of the measured $\delta^{13}C$ and $\delta^{2}H$ values to a normal distribution in the following ways:

(1) Histogram: The distributions of $\delta^{13}C$ and $\delta^{2}H$ values follow the normal distribution curves well (Figure 3.5a–b).
(2) Quantile-quantile plot: We plot the theoretical quantiles of the normal distribution against the actual quantiles of measured $\delta^{13}C$ or $\delta^{2}H$ values (Figure 3.5c–d, scripts available at doi.org/10.22002/mrqvr-ypb85). The data points follow the 1:1 reference line closely and lie within the confidence envelope, indicating measurement results follow a normal distribution.
(3) Spiegelhalter test for normality: We perform a Bayesian test on measured $\delta^{13}C$ or $\delta^{2}H$ values for normality against symmetric alternatives (the uniform distribution and the double exponential distribution) using the Spiegelhalter test in R package 'normtest' (Spiegelhalter, 1977). The p-values for the $\delta^{13}C$ or $\delta^{2}H$ values are 0.75 and 0.63, respectively. The result indicates that the distributions of the $\delta^{13}C$ or $\delta^{2}H$ values have insignificant deviations from the normal distribution.
(4) Shapiro–Wilk test: We also perform a test of normality in frequentist statistics on the $\delta^{13}C$ or $\delta^{2}H$ values using the Shapiro–Wilk test in R package 'dplyr'. The p-values for $\delta^{13}C$ or $\delta^{2}H$ results are 0.80 and 0.55, respectively. This also indicates that the measured $\delta^{13}C$ or $\delta^{2}H$ values have insignificant deviations from the normal distribution.

Therefore, we conclude that the measured $\delta^{13}C$ and $\delta^{2}H$ values follow normal distributions.

Table 3.1 Carbon and hydrogen isotope compositions of the C7H7+ fragment of p-xylene from replicate analyses

Analysis number[a]	$^{13}C^{12}C_6H_7/^{12}C_7H_7$	$\delta^{13}C_{average}$ (‰)	RSE (‰)[b]	$^{12}C_7H_6D/^{12}C_7H_7$	$\delta^{2}H_{average}$ (‰)	RSE (‰)[b]
1	0.07854	0.84	0.55	0.0009498	3.3	5.0
2	0.07853	0.73	0.55	0.0009532	6.8	5.1
3	0.07849	0.19	0.56	0.0009383	-9.0	5.1
4	0.07850	0.33	0.55	0.0009371	-10.2	5.1
5	0.07849	0.17	0.54	0.0009507	4.2	4.9
6	0.07849	0.13	0.55	0.0009430	-4.0	5.1
7	0.07842	-0.75	0.56	0.0009459	-0.9	5.0
8	0.07857	1.24	0.56	0.0009524	5.9	5.1
9	0.07844	-0.42	0.56	0.0009546	8.3	5.0
10	0.07841	-0.80	0.55	0.0009435	-3.5	5.0
11	0.07852	0.62	0.56	0.0009422	-4.7	5.0
12	0.07846	-0.21	0.53	0.0009435	-3.4	5.2
13	0.07853	0.72	0.55	0.0009535	7.2	5.0
14	0.07836	-1.45	0.55	0.0009473	0.6	5.0
15	0.07839	-1.09	0.55	0.0009556	9.3	5.1
16	0.07845	-0.30	0.55	0.0009342	-13.2	5.1
17	0.07841	-0.81	0.54	0.0009469	0.2	5.0
18	0.07858	1.29	0.54	0.0009462	-0.6	5.0
19	0.07844	-0.43	0.56	0.0009502	3.7	5.0
average	0.07848			0.0009467		
RSD (‰)[c]	0.79			6.5		

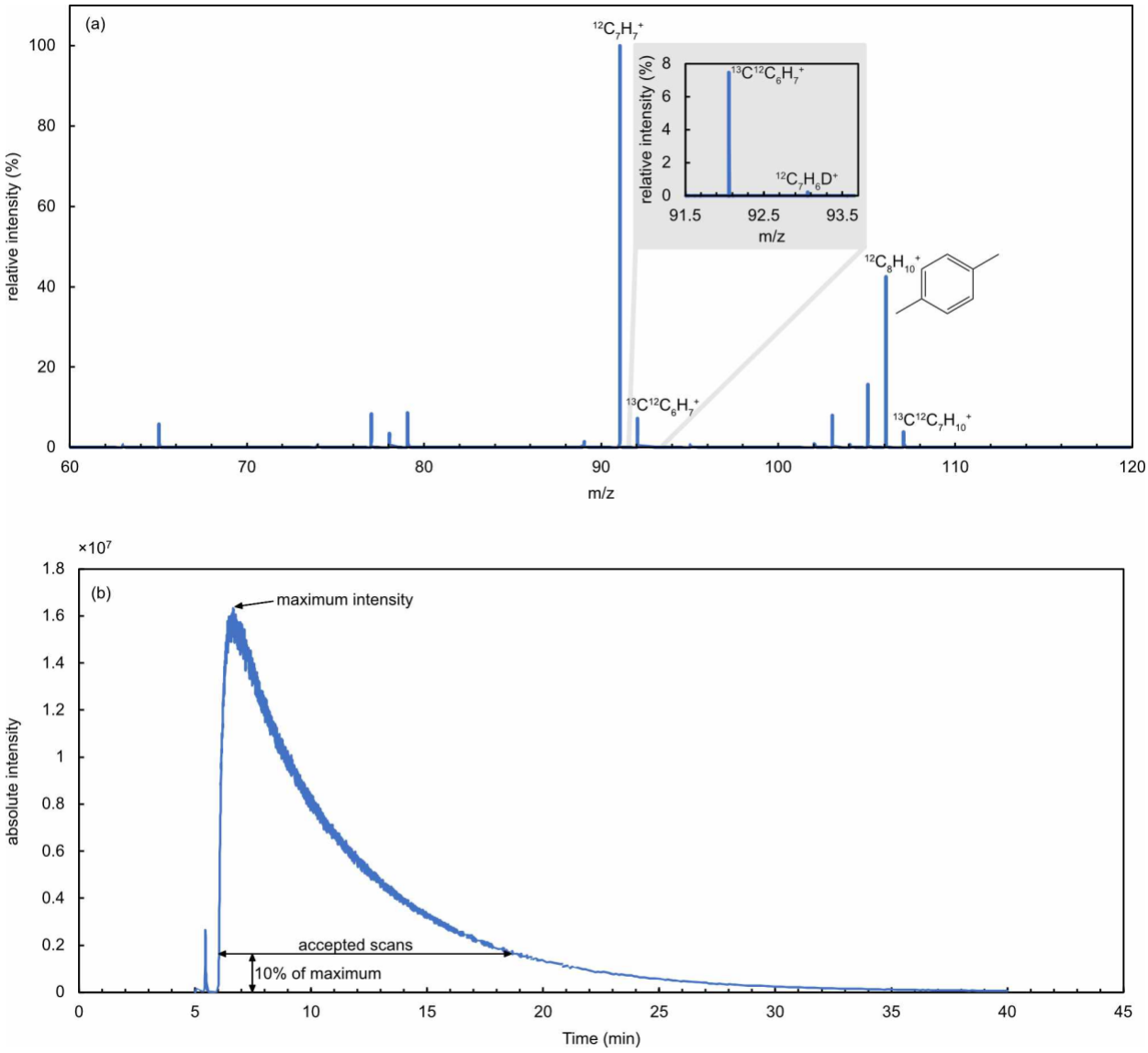


Figure 3.4 (a) A mass spectrum of p-xylene in this study with chemical formulas and/or structure of the ions labeled for a subset of peaks. The inset shows a blown-up mass spectrum where both 13C12C6H7+ and 12C7H6D+ peaks are visible. (b) A selected ion chromatogram for 12C7H7+ of a 'peak capture' experiment of p-xylene. We accept a continuous set of scans where the intensity of 12C7H7+ consistently exceeds 10% of the maximum value for calculation of isotope ratios. Another compound (likely m-xylene from impurity of the lab chemical) with the 12C7H7+ fragment elutes from the GC column at 5.4 min before the 'capture' of p-xylene in the peak broadener.

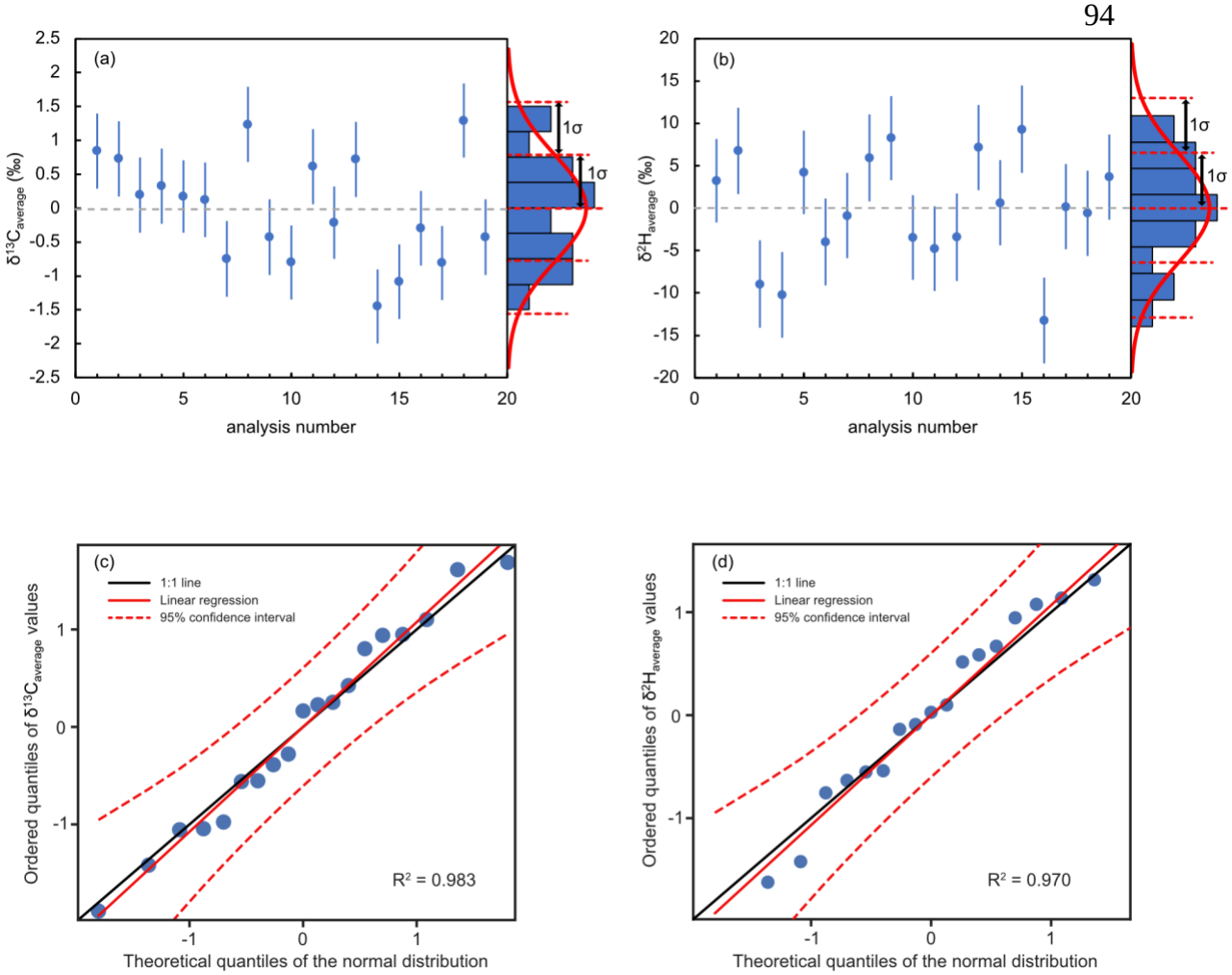


Figure 3.5 (a) carbon and (b) hydrogen isotope compositions of the C7H7+ fragment of p-xylene from replicate analyses and quantile-quantile plots of the measured δ13C (c) and δ2H (d) values. The δ13C (a) and δ2H (b) values show no clear systematic drift over the course of 19 analyses and 14 hours and are within 2–3 internal standard errors from each other, respectively (error bars are 1 internal standard error). The histograms of the measured δ13C (a) and δ2H (b) values are shown with normal distributions overlay. The δ13C (c) and δ2H (d) values follow normal distributions as data points follow the 1:1 reference line and lie within the confidence envelope when plotting the theoretical quantiles of the normal distribution against the actual quantiles of measured δ13C (c) or δ2H (d) values.

We interpret the observation of drift-free and normally distributed $\delta^{13}C$ and $\delta^{2}H$ values as evidence that there no obvious drift in the instrumental performance that could affect the

measured isotope ratios during the period of our analyses, and therefore no drift correction is necessary for this set of analyses.

More applied experience will be required to establish whether this drift-free performance over such a time period is truly typical. But this observed performance mirrors a previous finding that a measurement of the PFTBA reference standard can conform to shot noise limits for a single extended acquisition across close to 24 hours, where the PFTBA was introduced directly to the EI source using a reference-compound reservoir and stainless steel transfer line (Eiler et al., 2017). So we think it reasonable to expect that such drift-free behavior could be typical of such measurements and time spans.

The novel feature of the data in Table 3.1 Carbon and hydrogen isotope compositions of the C7H7+ fragment of p-xylene from replicate analyses and Figure 3.5 is that we now have shown that similar continuity of overall instrumental fractionations can be obtained despite the numerous and complex changes in sample delivery and starts and stops of acquisitions that occurred across an extended session of separate analyses. However, note that system drift in measured isotope ratios has been observed in prior Orbitrap IRMS measurements made on nitrate using an electrospray ionization (ESI) source (Hilkert et al., 2021) or during experimental sessions when an EI source was dirty or potentially damaged (e.g., unexpected venting of the instrument or improper cleaning of the ion source). It is possible that the difference in performance between this study and (Hilkert et al., 2021) reflects time-dependent variations in instrumental isotopic fractionation in an ESI source, and the general absence of such drift in a helium-buffered, well-functioning EI source. Nevertheless, overall, the results of p-xylene analyses demonstrate that automation of sample delivery and online separation and isotopic analysis enables Orbitrap-based IRMS measurements spanning across tens of hours and having external reproducibility comparable to the best-documented standard errors of individual experiments with such systems.

There are four impacts of the new capabilities provided by automated sample delivery, online separation, and Orbitrap-based isotopic analyses:

1. This automation increases the time utilization of an Orbitrap-based GC-IRMS instrument by 2–3 times by enabling analyses during time periods in a day when there typically will not be an analyst onsite.
2. The increased instrument utilization grants the practical possibility of high-precision experimental reproducibility, e.g., the standard error in $\delta^{13}C$ of the average of the set of replicate measurements of the $C_7H_7^+$ fragment of p-xylene presented here is 0.1‰. Time management is currently the central concern in the design and execution of isotope ratio measurements using the Orbitrap-based instruments (Eiler et al., 2017). The increased available beam time provides more leverage in the trade-off between precision and analytical duration.
3. The practical high-precision experimental reproducibility broadens the potential applications to study of isotope geochemistry in the context of geological, environmental and biological problems in terrestrial environments where isotopic variations may be subtle (e.g., (Wilkes et al., 2022)), i.e., beyond the demonstrated applications to extraterrestrial meteoritic samples where, for example, carbon isotope fractionations are often on the magnitude of tens to hundreds of per mil (Chimiak et al., 2021).
4. Finally, the automation of sample introduction and online separation averts potential human errors in manual operations for any precious samples including meteorites or samples returned from space missions.

The automation of Orbitrap-based GC-IRMS measurements allow high-precision study of carbon, hydrogen, oxygen, nitrogen, sulfur, and other isotope systematics of all GC amenable compounds for applications in geological, environmental, biological, archaeological, forensic, and cosmochemical sciences. This automated system is readily adoptable on other GC-Orbitrap instruments: the hardware we used is commercially available and the LabVIEW scripts are available at doi.org/10.22002/p6wqg-g3t24. Additional sketches are provided in the supplemental material.

4. Conclusions

We automated the sample introduction and online separation system on a GC-Orbitrap that had been modified for Orbitrap IRMS measurements. The automated system can perform scheduled analyses for hours to days without human intervention. The automation has greatly improved the sample throughput for Orbitrap-based GC-IRMS measurements for applications in cosmochemistry, biogeochemistry, petroleum geochemistry, among other

disciplines We expect this automation scheme to be readily adoptable on other GC-Orbitrap instruments. And, the principle of using LabVIEW VIs and digital input/output (I/O) devices to interface the GC method and electric actuators is applicable for developing other modular sample delivery systems for peak trapping, dual-inlet sample-standard comparison and diverting solvent peaks in conventional GC-MS instruments (Eiler et al., 2017; Hoegg et al., 2019b).

Over tens of hours, the external precision of measurements using the automated 'peak broadener' sample delivery system is close to the internal precision of each measurement and to shot-noise limits of precision. We did not observe any systematic drift in measured isotope ratios over the course of 19 separate analyses and 14 hours in this study. We suggest that the factors leading to system drifts on the Orbitrap IRMS may be subtle variations in isotopic discrimination in the ion source.

ACKNOWLEDGEMENTS

This work was supported by Thermo Fischer Scientific and sponsored by the US Department of Energy, the Simons Foundation, and the Caltech Joint Industry Partnership for Petroleum Geochemistry. We thank Brody Guckenberger and Alex Sessions for helpful discussion and Nami Kitchen and Fenfang Wu for laboratory support.

REFERENCES

Cesar J., Eiler J., Dallas B., Chimiak L. and Grice K. (2019) Isotope heterogeneity in ethyltoluenes from Australian condensates, and their stable carbon site-specific isotope analysis. *Organic Geochemistry* **135**, 32–37.

Chimiak L., Elsila J. E., Dallas B., Dworkin J. P., Aponte J. C., Sessions A. L. and Eiler J. M. (2021) Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system. *Geochimica et Cosmochimica Acta* **292**, 188–202.

Eiler J., Cesar J., Chimiak L., Dallas B., Grice K., Griep-Raming J., Juchelka D., Kitchen N., Lloyd M., Makarov A., Robins R. and Schwieters J. (2017) Analysis of

molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry. *International Journal of Mass Spectrometry* **422**, 126–142.

Eiler J. M. (2013) The Isotopic Anatomies of Molecules and Minerals. *Annu. Rev. Earth Planet. Sci.* **41**, 411–441.

Farley K. A., Reiners P. W. and Nenow V. (1999) An Apparatus for High-Precision Helium Diffusion Measurements from Minerals. *Anal. Chem.* **71**, 2059–2061.

Gilbert A. (2021) The Organic Isotopologue Frontier. *Annu. Rev. Earth Planet. Sci.* **49**, 435–464.

Hilkert A., Böhlke J. K., Mroczkowski S. J., Fort K. L., Aizikov K., Wang X. T., Kopf S. H. and Neubauer C. (2021) Exploring the Potential of Electrospray-Orbitrap for Stable Isotope Analysis Using Nitrate as a Model. *Anal. Chem.* **93**, 9139–9148.

Hoegg E. D., Barinaga C. J., Hager G. J., Hart G. L., Koppenaal D. W. and Marcus R. K. (2016) Preliminary Figures of Merit for Isotope Ratio Measurements: The Liquid Sampling-Atmospheric Pressure Glow Discharge Microplasma Ionization Source Coupled to an Orbitrap Mass Analyzer. *J. Am. Soc. Mass Spectrom.* **27**, 1393–1403.

Hoegg E. D., Godin S., Szpunar J., Lobinski R., Koppenaal D. W. and Marcus R. K. (2019a) Coupling of an atmospheric pressure microplasma ionization source with an Orbitrap Fusion Lumos Tribrid 1M mass analyzer for ultra-high resolution isotopic analysis of uranium. *J. Anal. At. Spectrom.* **34**, 1387–1395.

Hoegg E. D., Godin S., Szpunar J., Lobinski R., Koppenaal D. W. and Marcus R. K. (2021) Resolving Severe Elemental Isobaric Interferences with a Combined Atomic and Molecular Ionization Source–Orbitrap Mass Spectrometry Approach: The ^{87}Sr and ^{87}Rb Geochronology Pair. *Anal. Chem.* **93**, 11506–11514.

Hoegg E. D., Godin S., Szpunar J., Lobinski R., Koppenaal D. W. and Marcus R. K. (2019b) Ultra-High Resolution Elemental/Isotopic Mass Spectrometry ($m/\Delta m > 1{,}000{,}000$): Coupling of the Liquid Sampling-Atmospheric Pressure Glow Discharge with an Orbitrap Mass Spectrometer for Applications in Biological Chemistry and Environmental Analysis. *J. Am. Soc. Mass Spectrom.* **30**, 1163–1168.

Makarov A. and Denisov E. (2009) Dynamics of ions of intact proteins in the Orbitrap mass analyzer. *J. Am. Soc. Mass Spectrom.* **20**, 1486–1495.

Martin P. E. (2020) Detection and Analysis of Martian Low-Temperature Geochemistry. Ph.D. dissertation, California Institute of Technology.

Mueller E. P., Sessions A. L., Sauer P. E., Weiss G. M. and Eiler J. M. (2022) Simultaneous, High-Precision Measurements of $\delta^{2}H$ and $\delta^{13}C$ in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry. *Anal. Chem.* **94**, 1092–1100.

Neubauer C., Crémière A., Wang X. T., Thiagarajan N., Sessions A. L., Adkins J. F., Dalleska N. F., Turchyn A. V., Clegg J. A., Moradian A., Sweredoski M. J., Garbis S. D. and Eiler J. M. (2020) Stable Isotope Analysis of Intact Oxyanions Using Electrospray Quadrupole-Orbitrap Mass Spectrometry. *Analytical Chemistry* **92**, 3077–3085.

Neubauer C., Sweredoski M. J., Moradian A., Newman D. K., Robins R. J. and Eiler J. M. (2018) Scanning the isotopic structure of molecules by tandem mass spectrometry. *International Journal of Mass Spectrometry* **434**, 276–286.

Ono S., Wang D. T., Gruen D. S., Sherwood Lollar B., Zahniser M. S., McManus B. J. and Nelson D. D. (2014) Measurement of a Doubly Substituted Methane Isotopologue, $^{13}CH_3D$, by Tunable Infrared Laser Direct Absorption Spectroscopy. *Analytical Chemistry* **86**, 6487–6494.

Passey B. H., Levin N. E., Cerling T. E., Brown F. H. and Eiler J. M. (2010) High-temperature environments of human evolution in East Africa based on bond ordering in paleosol carbonates. *Proceedings of the National Academy of Sciences* **107**, 11245–11249.

Ricci M. P., Merritt D. A., Freeman K. H. and Hayes J. M. (1994) Acquisition and processing of data for isotope-ratio-monitoring mass spectrometry. *Organic Geochemistry* **21**, 561–571.

Sharp Z. (2017) Principles of Stable Isotope Geochemistry, 2nd Edition.

Spiegelhalter D. J. (1977) A test for normality against symmetric alternatives. *Biometrika* **64**, 415–418.

Valley J. W., Cole D. R. and Mineralogical Society of America eds. (2001) *Stable isotope geochemistry.*, Mineralogical Society of America, Washington, D.C.

Wagner C., Armenta S. and Lendl B. (2010) Developing automated analytical methods for scientific environments using LabVIEW. *Talanta* **80**, 1081–1087.

Wilkes E. B., Sessions A. L., Zeichner S. S., Dallas B., Schubert B., Jahren A. H. and Eiler J. M. (2022) Position-specific carbon isotope analysis of serine by gas chromatography/Orbitrap mass spectrometry, and an application to plant metabolism. *Rapid Comm Mass Spectrometry* **36**.

Zeichner S. S., Wilkes E. B., Hofmann A. E., Chimiak L., Sessions A. L., Makarov A. and Eiler J. M. (2022) Methods and limitations of stable isotope measurements via direct elution of chromatographic peaks using gas chromotography-Orbitrap mass spectrometry. *International Journal of Mass Spectrometry*, 116848.

Chapter 4 — Characterizing intramolecular carbon and hydrogen isotope structures of *n*-alkanes by GC-Orbitrap

Guannan Dong[a], Nithya Thiagarajan[a], Tim Cscernia[a], Alexandra Ferreira[b], Michael Lawson[c], Julia van Winden[d], Michael Formolo[e], Alex Sessions[a], John Eiler[a]

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

[b] *Petrobras-CENPES, Ilha do Fundão, Rio de Janeiro RJ 21941-909, Brazil*

[c] *AkerBP, Lysaker 1366, Norway*

[d] *Shell Global Solutions International, 2288 GS Rijswijk, The Netherlands*

[e] *ExxonMobil Upstream Integrated Solutions, Spring, TX 77389, USA*

Abstract

Intramolecular isotope analysis of *n*-alkanes offers insights into their origins and evolution and the conditions and processes that influence them. Existing techniques have limited applications due to sample size requirements and the lack of hydrogen isotope measurement. This study presents a novel method using GC-Orbitrap mass spectrometry to simultaneously probe the intramolecular carbon and hydrogen isotope structures of *n*-alkanes with nanomole sensitivity. We measured the abundances of ^{13}C- and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11} and C_6H_{13} fragment ions generated by electron impact ionization, achieving precisions of <1‰ and <5‰ for $\Delta^{13}C$ and $\Delta^{2}H$, respectively. This approach is generalizable to virtually all *n*-alkanes. We investigated the provenance of carbon and hydrogen atoms in fragment ions using isotopically labeled standards and modeled the fragmentation process to relate isotopic contents of fragments to site-specific isotopic structures of parent molecules. Proof-of-concept analyses demonstrate the method's capabilities for forensic fingerprinting and natural sample characterization. Quantitative constraints on the fragmentation schemes of *n*-hexane, *n*-heptane, *n*-nonane, *n*-tetradecane, *n*-hexadecane and *n*-nonadecane enable testing hypotheses regarding their site-specific isotope compositions, with results suggesting large isotope variations at certain molecular positions can be retained in fragment-specific isotope

compositions. This study lays the groundwork for diverse applications of fragment-specific isotope analysis of n-alkanes in geochemistry, environmental studies, cosmochemistry, paleoclimatology, and biochemistry.

1. Introduction

Analyzing the natural abundance variations of stable isotopes is a powerful tool for identifying the origins of chemical substances and constraining the physical, chemical, and biological conditions and processes that have influenced them. This approach has found widespread applications across diverse fields, including forensics, environmental science, geology, biology, archaeology and cosmology (Valley et al., 2001; Sharp, 2017). Isotopic characteristics can be examined at various levels, such as bulk materials, specific compounds, or even individual positions within molecular structures, each offering unique insights into systems, organisms, pathways, or elementary reactions (Eiler, 2013; Gilbert, 2021).

The carbon and hydrogen isotope composition of higher *n*-alkanes (C_{5+}) has extensive applications in petroleum geochemistry (Pedentchouk and Turich, 2018), paleoclimate studies, and environmental forensics. In petroleum geochemistry, these isotopic signatures are employed for oil-oil and oil-source rock correlations (Peters et al., 2007), distinguishing the sources of organic matter (Bjorøy et al., 1991; Li et al., 2001), providing insights into thermal maturation processes (Tang et al., 2005). In paleoclimate studies, n-alkane isotopic signatures offer information about the sources and processing of organic matter, distinguishing between terrestrial and aquatic inputs (Mead et al., 2005) and identifying contributions from different plant types (Tanner et al., 2007). These proxies reflect changes in vegetation, temperature, as well as hydrological conditions, enabling paleoclimate reconstructions (Sachse et al., 2012). In environmental forensics, n-alkane isotopes are used to trace contaminants and aid in remediation efforts (Philp, 2007). The diverse applications of n-alkane isotopes highlight their significance in understanding the origin, transformation, and fate of organic matter.

However, these applications are sometimes limited by the multiple sources and numerous processes influencing the carbon and hydrogen isotopic composition of n-alkanes (Castañeda and Schouten, 2011) and post-depositional alteration (Pedentchouk et al., 2006) and degradation in the environment (Zech et al., 2011; Brittingham et al., 2017). Careful interpretation and multi-proxy approaches are necessary to overcome these challenges.

Intramolecular or site-specific carbon isotope analysis of n-alkanes is an emerging approach to develop more informative proxies for their origins and transformations in the environment. Isotopic ^{13}C NMR spectroscopy (Gilbert et al., 2013; Julien et al., 2016) and on-line pyrolysis coupled with GC-C-IRMS (Corso and Brenna, 1999; Gilbert et al., 2016) are two techniques being explored to measure position-specific $^{13}C/^{12}C$ ratios. NMR provides direct, precise analysis but requires large samples. Pyrolysis-IRMS is more sensitive and partially preserves the isotopic signature of specific n-alkane positions in the pyrolysis fragments. But demands on sample size and potential matrix effects have limited applications to some terminal carbon positions of commercially available *n*-alkanes instead of natural samples, and no prior data of this kind is available for hydrogen isotopes.

In this study, we document the methods and performance characteristics of an analytical approach to simultaneously probe the intramolecular carbon and hydrogen isotope structures of *n*-alkanes from complex natural samples with nanomole sensitivity by GC-Orbitrap. We use the Orbitrap mass analyzer to measure the abundances of isotopologues of C_4H_9, C_5H_{11} and C_6H_{13} fragment ions of *n*-alkanes generated by electron-impact ionization (EI), to precisions of <1‰ and 3‰ for ^{13}C and ^{2}H substituted fragments, respectively. This analytical approach is generalizable to virtually all higher *n*-alkanes of various carbon number. We investigated the provenance of carbon and hydrogen atoms in the targeted fragment ions using isotopically labeled standards and built a model for EI chemistry to relate isotopic contents of fragments to isotopic structures of parent molecules.

2. Methods and materials

2.1. N-alkane samples

2.1.1. *N*-heptane

5 uL of *n*-heptane, respectively, from two chemical stocks (SK: Sigma Aldrich, LOT# MBQ8992V, made in South Korea. Ge: Honeywell, LOT# SZBG232S, made in Germany) were injected into two 15 mL gas-tight vials pre-flushed with helium and fully vaporized under laboratory temperature. 70 or 200 uL vapor was sampled from the gas-tight vials using a gas-tight syringe (Hamilton, LOT#631656) for analyses by GC-Orbitrap mass spectrometry. The petrochemical or refining and distillation origins of the *n*-heptanes are unknown.

2.1.2. *N*-hexadecane

We obtained *n*-hexadecane (TCI, LOT# 078XF) and site-specific labeled hexadecanes, 1,2-$^{13}C_2$-hexadecane (Sigma Aldrich, LOT# CX1002, 99 atom % ^{13}C) and 1-^{2}H-hexadecane (Sigma Aldrich, LOT# OI2335, 98 atom % ^{2}H). Hexadecanes were diluted in *n*-hexane to make 10 mg/ml (44 mM) solutions. Natural abundance *n*-hexadecane solution was volumetrically mixed with solutions of site-specific labeled hexadecanes to prepare standards with 0, 5, 10, and 20% labels (Table 4.8). Additional standards were made by mixing natural abundance *n*-hexadecane with 0.14% 1,2-$^{13}C_2$ label and 0.05% 1-^{2}H label (Table 4.10). The various labeled solutions were diluted to concentrations specified below for analyses on GC-Orbitrap mass spectrometry.

2.1.3. N-nonadecane

N-nonadecane (TCI, % purity, LOT# OGO01) was diluted in *n*-hexane to make 10.2 mg/mL (37.99 mM) stock solutions, CIT-C19. Working standards were created by diluting the CIT-C19 stocks to 0.1 mg/mL with *n*-hexane for GC Orbitrap analyses.

2.1.4. Rose essential oil and n-alkane extraction

Rose essential oils have been reported to contain *n*-nonadecane produced by the plant naturally (Verma et al., 2011; Zgheib et al., 2020). We obtained a commercial rose essential oil (doTERRA, LOT# 213007A) and extracted the *n*-alkane fraction by solid-phase extraction (SPE) following the method described in (Sessions, 2006; Dion-Kirschner et al., 2020). Briefly, aminopropyl SPE separations used 0.5g of Sepra NH2 (50um, 65A) stationary phase (Phenomenex, batch# S209-098 or S209-100) hand-packed into 8 mL glass syringe barrels. 24 uL of rose essential oil sample was eluted in four fractions: F1 (hydrocarbons), 4 mL hexane; F2 (ketones and ethers), 7 mL 4:1 hexane/dichloromethane (DCM); F3 (alcohols), 7 mL 9:1 DCM/acetone; F4 (carboxylic acids), and 9mL 49:1 DCM/formic acid (all solvent ratios in v/v). Following the aminopropyl SPE, Ag-ion columns were used to remove alkenes from *n*-alkanes. Briefly, 0.5 g of silver nitrate on silica gel stationary phase (Sigma Aldrich, batch# MKCP3740) hand-packed into 8 mL glass syringe barrels. The fraction of hydrocarbons (F1) was eluted in three fractions: F1 (alkanes), 5 mL hexane; F2 (alkenes), 5 mL 4:1 hexane/DCM; F3 (blank), 5 mL acetone (all solvent ratios in v/v).

2.2. Instrumentation and analytical method

2.2.1. Instrumentation

We used a Thermo Scientific Q Exactive Orbitrap mass spectrometer with a Trace 1310 GC and a Thermo Scientific Triplus RSH autosampler. The sample introduction and online separation of the GC-Orbitrap system has been customarily modified (Eiler et al., 2017), Chapter 3). Briefly, two types of measurements were conducted with the custom sample introduction and online separation system:

1. Direct elution measurement: isotope ratios of target ions were measured as a compound of interest eluted from a GC over a period of seconds (Zeichner et al., 2022).

2. Peak capture measurement: a compound of interest was first directed to and mixed in a peak broadener volume (which is prefilled with helium) as it elutes from GC. The isotope ratios of target ions were then measured as a flow of pure helium carrier gas purges the peak broadener and delivers the trapped compound of interest over tens of minutes (Eiler et al., 2017).

Sample introduction and online separation for liquid samples was done with the autosampler and the automated custom system (Chapter 3), whereas vapor samples were injected manually using a gas-tight syringe, and the peak capturing and broadening procedure was also executed manually by turning two Valco valves in the custom system.

We used a TG-5SilMS column (30 m × 0.25 mm × 0.25 μm film thickness). The electron-impact ion (EI) source was set at 250°C. We used the positive ion mode and no 'lock masses' were used.

A solution of perfluorotributylamine (PFTBA, Sigma-Aldrich, LOT# 77299) was used to calibrate the mass scale and mass isolation for positive ions and to auto-tune the ion source, injection flatapole and inter-flatapole lens with electron energy set at 70 eV and C-trap energy offset at 0 V (unless stated otherwise), maximizing the signal intensity on the Q Exactive GC.

2.2.2. *N*-heptane analysis

We analyzed the ^{13}C-substituted isotopologues of C_4H_9, C_5H_{11}, C_6H_{13} and C_7H_{16} fragment ions of *n*-heptane. Vaporized *n*-heptane was injected manually to the GC column through a split/splitless injector held at 50°C and operated with a split ratio of 20, using helium as a carrier gas with a flow rate of 1.5 mL/min. The GC oven was set at 50°C. The transfer line was set at 250°C. We set the electron energy at 32 eV to improve the yield of target fragment ions. We used a quadrupole scan range of 47 to 82 m/z for $C_4H_9^+$ and $C_5H_{11}^+$ analyses and 76 to 104 m/z for $C_6H_{13}^+$ and $C_7H_{16}^+$. We used the prescan AGC mode with an AGC target

of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 10 to 200 ms for scans during which sample was present). The number of microscans was set to 1.

2.2.2.1. Comparison of n-heptane from two commercial sources

70 uL of sample of SK or Ge *n*-heptane was injected in each analysis. We used a 10 cm^3 passivated stainless-steel peak broadener. We set the C-trap energy offset at 0 V for $C_4H_9^+$ and $C_5H_{11}^+$ analyses and 2.3 V for $C_6H_{13}^+$ and $C_7H_{16}^+$ for best sensitivity for the corresponding ions. The nominal mass resolution was set at 120,000 (full-width-half-maximum, FWHM, at m/z = 200). We ran the acquisition for 20 min, which we established was the time required for the mass spectra of effluent from this peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., heptane-free) measurements. Each analysis of Ge sample was followed by an analysis of SK sample and each pair of Ge-SK analyses was repeated 5 times.

2.2.2.2. Experiment to improve measurement precision

We conducted another experiment to test on improvement of measurement precision. 200 uL of sample from Ge was injected in each measurement. We used a 20 cm^3 passivated stainless-steel peak broadener. We set the C-trap energy offset at 1.5 V. The nominal mass resolution was set at 60,000. We ran the acquisition for 40 min and the analyses was replicated.

2.2.3. *N*-hexadecane analysis

2.2.3.1. Direct elution measurement

We analyzed the ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions of 0, 5, 10, and 20% 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecane to study the provenance of carbon and hydrogen atoms in the fragments. The analyses were via direct elution measurement (Zeichner et al., 2022) out of precaution for potential memory effect in peak broadener volume by peak capture measurements.

1 uL of 0.003 mg/mL (13.2 uM) hexadecane solutions of the above different label levels was injected to the GC column through a split/splitless injector held at 250°C and operated in the splitless mode, using helium as a carrier gas with a flow rate of 1.2 mL/min. The GC oven was started with a 1.5 min hold at 100°C, heated from 100°C to 215°C at 20°C/min and 215°C to 225°C at 5°C/min, and held at 225°C for 1 min. The transfer line was set at 265°C. The filament was turned on at 5 min. We used a quadrupole scan range of 50 to 500 m/z. We used the prescan AGC mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 0.1 to 30 ms for scans during which sample was present). The nominal mass resolution was set at 120,000 (FWHM, at m/z = 200). The number of microscans was set to 1. We ran the acquisition for the duration of the GC oven temperature program and the hexadecane peak eluted from 8.35 to 8.6 min. An analysis sequence of unlabeled, 5, 10, and 20% 1-^{2}H-labeled, 5, 10, and 20% 1,2-$^{13}C_2$-labeled, and unlabeled hexadecane was repeated 3–6 times.

2.2.3.2. Peak capture measurement

We analyzed the ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions of unlabeled n-hexadecane and hexadecanes with 0.14% 1,2-$^{13}C_2$ label and 0.05% 1-^{2}H label via peak capture measurement.

0.7-1 uL of 0.2 mg/mL (0.88 mM) hexadecane solutions of the above different label levels was injected to the GC column through a split/splitless injector held at 250°C and operated in the splitless mode, using helium as a carrier gas with a flow rate of 1.4 mL/min. The GC oven was started with a 1.5 min hold at 100°C, heated from 100°C to 215°C at 20°C/min and 215°C to 225°C at 5°C/min, and held at 225°C for 40 min. We used a 10 cm^3 passivated stainless-steel peak broadener. The transfer line was set at 265°C. The filament was turned on at 5 min. We measured one out of the three fragments in each analysis and used a quadrupole scan range of 52–63 mz for C_4H_9 fragment, 66.5–76.5 mz for C_5H_{11}, and 81–90 m for C_6H_{13}. We used the prescan AGC mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 10 to 100 ms for scans during

which sample was present). The nominal mass resolution was set at 120,000 (FWHM, at m/z = 200). The number of microscans was set to 1. We ran the acquisition for 50 min, which we established was the time required for the mass spectra of effluent from the peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., hexadecane-free) measurements. Each analysis of labeled hexadecane was bracketed by two analyses of the unlabeled *n*-hexadecane and each bracketed analysis of labeled hexadecane was repeated 4 times for each fragment and each label type.

2.2.4. *N*-nonadecane analysis

We analyzed the ^{13}C- and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions of *n*-nonadecane CIT-C19 and from rose essential oil via peak capture measurement spanning seven sessions from February 2022 to February 2024.

About 1 uL of 0.1 mg/mL (0.37 mM) CIT-C19 solution or rose essential oil solution with similar nonadecane concentration was injected to the GC column through a split/splitless injector held at 250℃ and operated in the splitless mode, using helium as a carrier gas with a flow rate of 1.2 mL/min. The GC oven was started with a 1.5 min hold at 100℃, heated from 100℃ to 220℃ at 20℃/min, from 220℃ to 230℃ at 3℃/min, and from 230 to 320℃ at 3.2℃/min, and held at 320℃ for 3.5 min. We used a 10 cm^3 passivated stainless-steel peak broadener. The transfer line was set at 265℃. The filament was turned on at 5 min. We measured one out of the three fragments in each analysis and used a quadrupole scan range of 52–63 mz for C_4H_9 fragment, 66.5–76.5 mz for C_5H_{11}, and 81–90 m for C_6H_{13}. We used the prescan AGC mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 1 to 200 ms for scans during which sample was present). The nominal mass resolution was set at 120,000 (FWHM, at m/z = 200). The number of microscans was set to 1. We ran the acquisition for the duration of the GC oven temperature program, which we established was the time required for the mass spectra of effluent from the peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., hexadecane-free) measurements and for constituents of the alkane

fraction in the rose essential oil to finish eluting from the GC column. Each analysis of the rose essential oil sample was bracketed by two analyses of CIT-C19, and each bracketed analysis was repeated 4 times for each fragment.

2.3. Calculation for isotope data

The data analysis method has been detailed in Chapter 3. Briefly, The Thermo Scientific FTStatistic software and Python scripts (available at doi.org/10.22002/catyq-fxd56) are used to process mass spectral data from peak capture and direct elution experiments. The process involves culling scan-by-scan data based on a 10% threshold of base ion intensity for peak capture measurements or based on elution time for direct elution measurements, removing scans that are 3σ outliers in TIC×IT (total ion current × injection time), and calculating scan-by-scan ion counts using a formula that incorporates ion intensity, noise, resolution, and charge. Finally, we calculate the carbon isotope ratio of the fragment ion.

For peak capture measurements, we calculate by (Chapter 3)

$$ {}^{J}R_{frag} = \frac{\frac{\sum_i {}^{j1}N_i \times {}^{j1}I_i}{\sum_i {}^{j1}I_i}}{\frac{\sum_i {}^{j2}N_i \times {}^{j2}I_i}{\sum_i {}^{j2}I_i}} \qquad (4\text{-}1) $$

where ${}^{j1}N_i$ and ${}^{j2}N_i$ are the counts of different isotopologues of fragment ions in scan i, respectively, and ${}^{j1}I_i$ and ${}^{j2}I_i$ are the intensities of different isotopologues for scan i.

For direct elution measurement, we calculate by (Zeichner et al., 2022)

$$ {}^{J}R_{frag} = \frac{\sum_i {}^{j1}N_i}{\sum_i {}^{j2}N_i} \qquad (4\text{-}2) $$

Site-specific isotope composition refers to the isotope ratios (e.g., $^{13}C/^{12}C$ or $^{2}H/^{1}H$) at individual molecular positions within a compound, whereas fragment-specific isotope

composition describes the isotope ratios of molecular fragments generated during the ionization and fragmentation of a parent molecule in a mass spectrometer.

The isotope fractionation of the fragment ions between *n*-heptanes are expressed using an epsilon notation:

$$^{J}\varepsilon_{subtance\ A-subtance\ B} = \frac{^{J}R_{frag,substance\ A}}{^{J}R_{frag,substance\ B}} - 1 \qquad (4\text{-}3)$$

The fragment-specific isotope composition of *n*-nonadecane in reference to CIT-C19 are calculated using a delta notation after offsetting the molecular average isotope fractionation:

$$\Delta J_{\mathrm{CIT-C19}} = \frac{^{J}R_{frag,substance\ A}}{^{J}R_{frag,substance\ B}} \Big/ \frac{^{J}R_{mol,substance\ A}}{^{J}R_{mol,substance\ B}} - 1 \quad (4\text{-}4)$$

2.4. Molecular average carbon and hydrogen isotope measurements

The $\delta^{13}C_{VPDB}$ and $\delta^{2}H_{VMOW}$ of *n*-heptane chemical stocks, 0 and 0.14% 1,2-$^{13}C_2$ labeled and 0.05% 1-^{2}H label *n*-hexadecanes, and rose and CIT-C19 *n*-nonadecanes were measured at Caltech by GC- combustion-IRMS (for $\delta^{13}C_{VPDB}$) and GC-pyrolysis-IRMS (for $\delta^{2}H_{VMOW}$) on a Thermo-Scientific Delta+ XP or MAT 253. The GC methods are the same as on the GC-Orbitrap analysis.

Co-injected peaks of CO_2 ($\delta^{13}C_{VPDB}$ = -12.04‰) and H_2 ($\delta^{2}H_{VSMOW}$ = -254‰) reference gases were used for single-point isotopic calibration of analytes. The reference gases were in turn calibrated against a standard solution of eight fatty acid methyl esters ('F8 mix' of Arndt Schimmelmann, Indiana University). The F8 mix was also run along with the samples to monitor the instrumental status. The long-term experimental reproducibility for measurements of standards and replicate samples on these instruments is 0.15‰ for δ^{13}C and 5‰ for δ^{2}H (1SE).

2.5. Numerical modeling

We constructed the fragmentation models in section 3.2 using MATLAB. We found the best fits of fragmentation scheme to experimental data using MATLAB linear least-squares solver lsqlin with the interior-point algorithm and the OptimalityTolerance as 1e-14.

3. Results and discussion

3.1. Performance characteristics of intramolecular carbon and hydrogen isotope analyses of n-alkanes

3.1.1. Precision and reproducibility of peak capture measurements of *n*-heptane, *n*-hexadecane, and *n*-nonadecane

The parameters (including sample size) and statistics of various peak capture measurements on *n*-heptane, *n*-hexadecane, and *n*-nonadecane are listed in Table 4.1. A representative measurement of C_5H_{11} fragment of *n*-nonadecane and its statistics are illustrated in Figure 4.1. *N*-hexadecane and *n*-nonadecane analyses have sensitivity of sub-nanomole for each acquisition, whereas *n*-heptane analyses require hundreds of nanomoles of sample to achieve good chromatographic peak shape using split mode in the injector. ^{13}C- and ^{2}H-substituted isotopologues can be simultaneously measured in a single acquisition. Carbon and hydrogen isotope ratios of each scan from the representative acquisition produce normal distributions. Throughout the acquisition, the relative standard error of isotope ratios in consecutive scans scales with shot noise errors, reaching tenths of ‰ in tens of minutes. All carbon isotope analyses achieve ‰ or sub-‰ level internal precision and external reproducibility whereas hydrogen isotope analyses have precision and reproducibility < 5‰. The measurement accuracy of will be discussed in section 3.1.2 and 3.2.4-3.2.4.2.

Although hydrogen isotope analysis was not a target in the design of *n*-heptane measurements, we analyzed the data post measurements to understand the performance of the measurement parameters. About 25–50% of all scans did not detect ^{2}H-substituted

isotopologues. These non-detections are possibly due to lower mass resolution used and lower ion count for each isotopologue (measuring two fragments together within the quadrupole scan range and splitting the ion load allowed by AGC target between them). But from the successful experience measuring ^{2}H-substituted fragments from *n*-hexadecane and *n*-nonadecane, similar measurement can probably be achieved for *n*-heptane with mass resolution of 120000 and quadrupole scan range observing single fragment in each scan.

Table 4.1 Metrics of peak capture measurements of fragments of *n*-heptane, *n*-hexadecane, and *n*-nonadecane

	measurement parameters				precision								reproducibility							
					carbon isotope				hydrogen isotope				carbon isotope				hydrogen isotope			
compound	mass resolution	AGC target	acquisition duration (min)	sample size per injection (nmol)	C4H9	C5H11	C6H13	C7H16	C4H9	C5H11	C6H13	C7H16	C4H9	C5H11	C6H13	C7H16	C4H9	C5H11	C6H13	C7H16
n-heptane	120000	2.00E+05	20	160	1.7	1.7	1.4	0.9					1.9	2.7	1.6	1.4				
n-heptane	60000	2.00E+05	40	450	0.7	0.6	0.6	0.5												
n-hexadecane	120000	2.00E+05	35	0.6	0.6	0.5	0.5		3	3	3		0.7	0.7	0.7		4	5	4	
n-nonadecane	120000	2.00E+05	35	0.4	0.6	0.5	0.5		4	3	3		0.8	0.8	0.8		4	3	3	

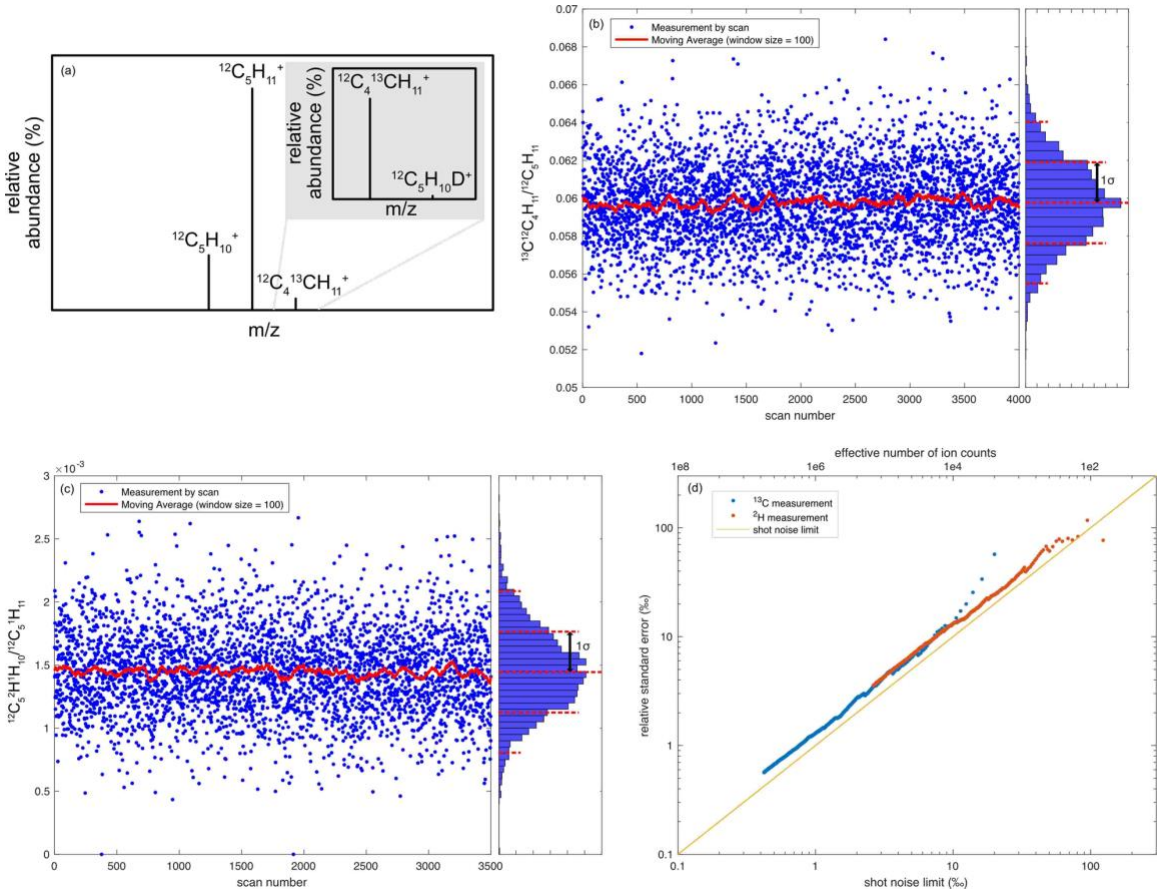


Figure 4.1 Peak capture measurement metrics of C_5H_{11} fragment from *n*-nonadecane. (a) Mass spectrum showing simultaneous measurement of all three isotopologues of C_5H_{11} fragment during a single acquisition. (b) Carbon and (c) hydrogen isotope ratios of each scan from a representative acquisition produce normal distributions. Moving averages of 100 scans are shown as red lines. (d) Throughout an acquisition, the relative standard error of isotope ratios in consecutive scans scales with shot noise errors, reaching tenths of ‰ in tens of minutes. The effective ion counts of isotopologues a and b: $n_{eff} = n_a * n_b/(n_a+n_b)$. Shot noise limit: $\sigma_{shot\ noise} = \sqrt{1/n_{eff}}$.

3.1.2. Proof-of-concept forensic isotope analysis of commercial *n*-heptanes

Commercial *n*-heptanes sourced from Germany and South Korea have similar molecular average $\delta^{13}C$ but the carbon isotope composition of C_4H_9 and C_6H_{13} fragments are distinguishable from each other, with fragment-specific isotope fractionations larger than 2–

3 standard errors of the measurements (Table 4.2, Figure 4.2). Additionally, the carbon isotope fractionation of molecular ions measured by GC-Orbitrap is consistent with molecular average $\delta^{13}C$ measurement by GC-c-IRMS (Table 4.2). These data support the accuracy of isotope analysis by GC-Orbitrap and demonstrate that fragment-specific isotope composition of *n*-heptane can provide additional information and constraints to distinguish sources by isotope forensics.

Table 4.2 Carbon isotope analyses of C_4H_9, C_5H_{11}, C_6H_{13} and C_7H_{16} fragments and molecular average of commercial *n*-heptanes

sample	Analysis number[a]	$^{13}C^{12}C_3H_9/^{12}C_4H_9$	RSE (‰)[b]	$^{13}C^{12}C_4H_{11}/^{12}C_5H_{11}$	RSE (‰)[b]	$^{13}C^{12}C_5H_{13}/^{12}C_6H_{13}$	RSE (‰)[b]	$^{13}C^{12}C_6H_{16}/^{12}C_7H_{16}$	RSE (‰)[b]	Molecular average (‰)
Ge	1	0.03544	1.57	0.04219	1.67	0.06451	1.36	0.07425	0.85	
	2	0.03557	1.61	0.04209	1.70	0.06430	1.34	0.07414	0.91	
	3	0.03562	1.64	0.04191	1.68	0.06421	1.34	0.07433	0.91	
	4	0.03558	1.62	0.04209	1.70	0.06450	1.36	0.07429	0.94	
	5	0.03554	1.71	0.04192	1.74	0.06443	1.34	0.07451	0.90	
average		0.03555		0.04204		0.06439		0.07431		-26.27
RSE (‰)[c]		0.83		1.30		0.89		0.80		0.15
sample	Analysis number[a]	$^{13}C^{12}C_3H_9/^{12}C_4H_9$	RSE (‰)[b]	$^{13}C^{12}C_4H_{11}/^{12}C_5H_{11}$	RSE (‰)[b]	$^{13}C^{12}C_5H_{13}/^{12}C_6H_{13}$	RSE (‰)[b]	$^{13}C^{12}C_6H_{16}/^{12}C_7H_{16}$	RSE (‰)[b]	Molecular average (‰)
SK	1	0.03566	1.70	0.04210	1.71	0.06419	1.35	0.07419	0.87	
	2	0.03558	1.69	0.04227	1.71	0.06427	1.36	0.07416	0.86	
	3	0.03575	1.69	0.04200	1.72	0.06410	1.35	0.07423	0.89	
	4	0.03571	1.73	0.04219	1.76	0.06425	1.35	0.07432	0.86	
	5	0.03572	1.69	0.04206	1.67	0.06429	1.37	0.07434	0.89	
average		0.03568		0.04213		0.06422		0.07425		-26.88
RSE (‰)[c]		0.84		1.12		0.53		0.48		0.15
	$^{13}C^{12}C_3H_9$	SE	$^{13}C^{12}C_4H_{11}$	SE	$^{13}C^{12}C_5H_{13}$	SE	$^{13}C^{12}C_6H_{16}$	SE	Molecular average	SE
ε_{Ge-SK} (‰)[d]	-3.76	1.18	-2.09	1.71	2.62	1.04	0.77	0.93	0.61	0.21

[a] analyses in Ge-SK pairs
[b] internal relative standard error
[c] external relative standard error
[d] carbon isotope fractionation between fragments of Ge and SK

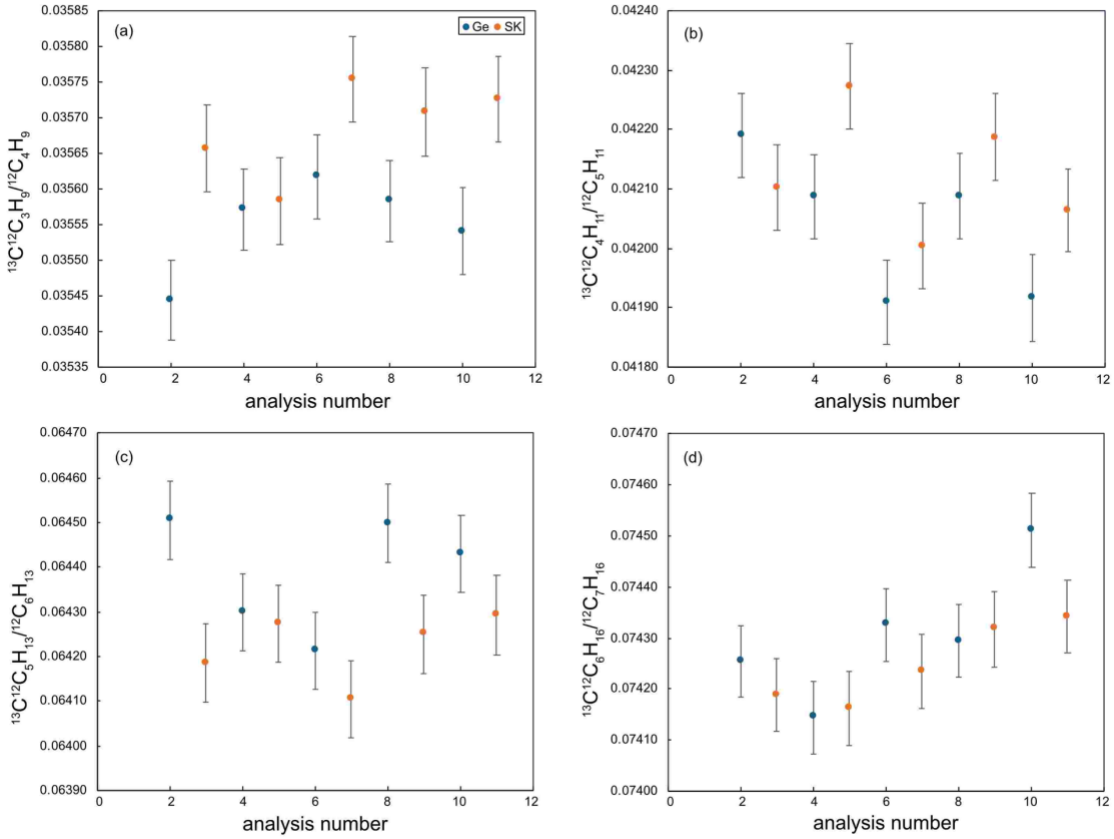


Figure 4.2. Carbon isotope analysis data of (a) C_4H_9, (b) C_5H_{11}, (c) C_6H_{13} and (d) C_7H_{16} fragments of commercial *n*-heptanes from Germany (Ge) and South Korea (SK).

3.1.3. Long-term precision of *n*-nonadecane measurements

Fragment-specific carbon and hydrogen isotope composition of rose essential oil were analyzed in seven sessions in reference to working standard CIT-C19 spanning from 2022 to 2024 (Table 4.3, Figure 4.3). The replicate measurements are within 3 standard errors from each other. The external standard deviation of the measurements over 2 years was comparable to the internal standard error of each measurement. The fragment-specific $\Delta^2H_{CIT\text{-}C19}$ of C_4H_9, C_5H_{11}, and C_6H_{13} fragments from the rose *n*-nonadecane demonstrated a

V-shape. These data support the long-term precision and stability of isotope analysis by GC-Orbitrap and demonstrate that fragment-specific isotope composition of *n*-nonadecane can provide characteristic constraints forensically.

Table 4.3 Long term carbon and hydrogen isotope measurement results of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of rose *n*-nonadecane in reference to the *n*-nonadecane working standard, CIT-C19 and their molecular average carbon and hydrogen isotope composition. The external standard deviation of measurements over 2 years is comparable to the internal standard error of each measurement. The fragment-specific $\Delta^{13}C_{CIT\text{-}C19}$ and $\Delta^{2}H_{CIT\text{-}C19}$ values are calculated after offsetting the molecular average isotope fractionation (equation (4-4)).

	$\Delta^{13}C_{nC19stock}$ (‰)			SE (‰)			$\Delta^{2}H_{nC19stock}$ (‰)			SE (‰)		
measurement date	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13
20220215	-1.3	0.7	2.1	1.1	0.6	0.8	9	-5	8	3	4	4
20220327	-1.1	0.2	-0.7	1.1	0.5	0.6	8	4	17	7	11	7
20220601	-0.8	0.6	-0.7	0.7	0.8	0.7	16	10	7	4	4	2
20221213	0.2	-0.5	1.2	1.0	0.6	0.7	15	2	8	5	6	4
20231222	-3.0	-1.9	-2.2	0.7	0.7	0.6	10	10	10	6	5	3
20231228	-2.6	0.2	-0.6	0.5	0.4	0.8	12	7	16	4	4	4
20240110	0.0	-1.3	-0.6	0.4	0.7	0.5	15	5	14	3	5	1
20240117	-0.7	-0.2	-0.3	0.5	0.7	0.5	16	8	18	2	3	3
20240214	-1.7	-0.5	-1.0	0.8	0.6	0.7	8	1	0	3	7	6
average	-1.2	-0.3	-0.3				12	5	11			
SD	1.1	0.9	1.2				3	5	6			
SE	0.4	0.3	0.4				1	2	2			

molecular average	$\delta^{13}C$ (‰)	$\delta^{2}H$ (‰)
rose *n*-nonadecane	-28.6	-240
CIT-C19	-30.4	-103

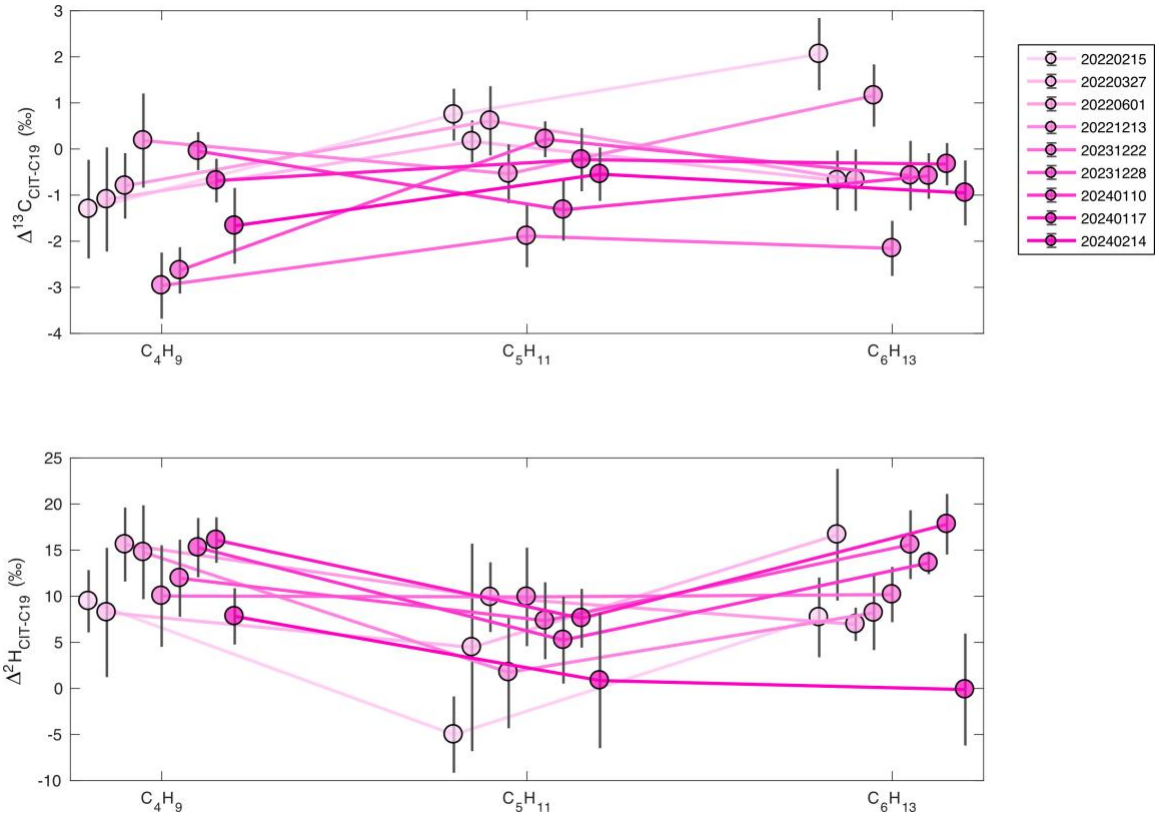


Figure 4.3 Long term carbon and hydrogen isotope measurement results of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of rose *n*-nonadecane in reference to the *n*-nonadecane working standard, CIT-C19. The replicate measurements over 2 years are within 3 standard errors from each other. The external standard deviation of the measurements was comparable to the internal standard error of each measurement. The fragment-specific $\Delta^2H_{CIT-C19}$ of C_4H_9, C_5H_{11}, and C_6H_{13} fragments from the rose *n*-nonadecane demonstrated a V-shape. The fragment-specific $\Delta^{13}C_{CIT-C19}$ and $\Delta^2H_{CIT-C19}$ values are calculated after offsetting the molecular average isotope fractionation (equation (4-4)). Error bars are standard errors of the measurements.

3.2. Developing the transfer function from site-specific isotope composition to fragment-specific isotope composition

We have demonstrated the performance characteristics of isotopologue measurements of fragment ions from various *n*-alkanes. In this section, we attempted to bridge the fragment-specific isotopic measurements to site-specific isotopic composition of *n*-alkanes through a hypothesis testing approach. We developed transfer functions to forward predict the

fragment-specific isotope composition of *n*-alkanes from their site-specific isotope composition by modeling the fragmentation in electron impact ion (EI) sources.

3.2.1. Provenance of carbon and hydrogen atoms in the measured fragments

The ionization and fragmentation of *n*-alkanes in EI sources involve complex, competing processes influenced by factors such as molecular structure, ionization energy, and electron energy (McLafferty and Tureček, 1993; Gross, 2004; Watson and Sparkman, 2007). High-energy electrons cause ionization and extensive fragmentation, generating molecular ions with varying internal energies that undergo further unimolecular decomposition (Wendelboe et al., 1981b, a; Wendelboe and Bowen, 1982). Key pathways include direct C-C bond cleavage (Meyerson, 1965), hydrogen shift (Howe, 1975; Wolkoff and Holmes, 1978), carbon skeleton isomerization (Meyerson, 1965; Wolkoff and Holmes, 1978; Wendelboe et al., 1981b, a; Holmes et al., 1982; Wolkoff et al., 1983), and secondary fragmentations (Meyerson, 1965; Lavanchy et al., 1978). Relative contributions depend on alkane size (Lavanchy et al., 1979) and ion internal energy (Lavanchy et al., 1978; Maccoll, 1982). Formation of characteristic fragment ions ($C_nH_{2n+1}^+$, $C_nH_{2n-1}^+$) involves multi-step processes. Labeling studies (^{13}C, D) have elucidated some mechanisms and carbon and hydrogen provenance (Meyerson, 1965; Wolkoff and Holmes, 1978; Wendelboe et al., 1981b, a; Holmes et al., 1982; Wolkoff et al., 1983), but understanding larger alkanes remains challenging due to isomeric structures and rearrangements (McLafferty and Tureček, 1993).

3.2.2. A framework for modeling the EI fragmentation of *n*-alkanes

We adapted a framework for electron impact fragmentation of *n*-alkanes to develop a quantitative model (Lavanchy et al., 1978, 1979). Lavanchy et al. conducted a detailed quantitative investigation of the electron impact induced fragmentation of *n*-alkanes (specifically *n*-hexane, *n*-heptane, *n*-nonane, and *n*-tetradecane) using site-specific ^{13}C- and ^{2}H-labeled compounds with an electron energy of 70 eV. Taking *n*-heptane as an example, they utilized *n*-heptanes labeled at the following positions: [1-^{13}C], [2-^{13}C], [3-^{13}C], [4-^{13}C],

[1,7-$^{13}C_2$], [1,2-$^{13}C_2$], [1,3-$^{13}C_2$], [2,3-$^{13}C_2$], [1,2,3-$^{13}C_3$]; [4-d_1], [3,3-d_2], [3,5-d_4], [1,7-d_6]. Measuring the mass spectra of unlabeled and these site-specific labeled *n*-alkanes and deriving the abundances of ^{13}C- and ^{2}H-containing fragment ions provided adequate constraints on the fate of carbon and hydrogen atoms at each position (Lavanchy, 1977).

They proposed two competing modes of decompositions: (a) Fragmentation of the molecular ion to form an alkyl ion and an alkyl radical

$$[C_mH_{2m+2}]^+ \rightarrow [C_nH_{2n+1}]^+ + C_{(m-n)}H_{2(m-n)+1} \qquad (4\text{-}5)$$

(b) Subsequent fragmentation of the resulting alkyl ions, forming a lower homologue and an alkene neutral species

$$[C_nH_{2n+1}]^+ \rightarrow [C_kH_{2k+1}]^+ + C_{(n-k)}H_{2(n-k)} \qquad (4\text{-}6)$$

where each carbon or hydrogen in the alkyl ion $[C_nH_{2n+1}]^+$ has the same probability to be lost to the neutral fragment, i.e., the carbon and hydrogen atoms are equivalent to being fully scrambled through carbon skeleton rearrangement or hydrogen shift. The formation of each product fragment ion has been considered in terms of reaction sequences made up by (4-5) and (4-6) (Figure 4.4).

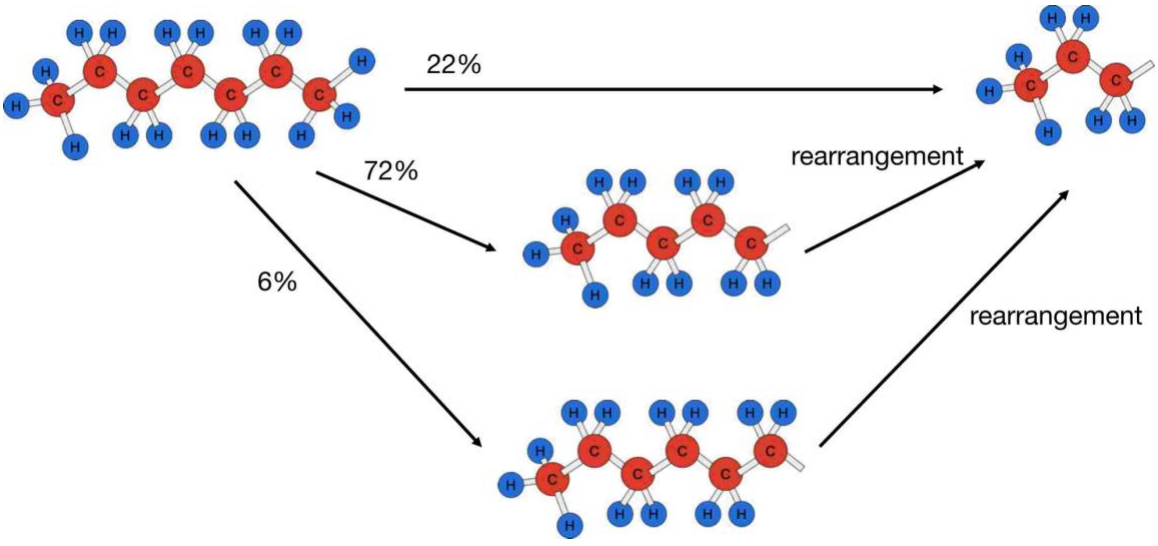


Figure 4.4 Fragmentation starts with the molecular ion to form an alkyl ion through a direct cleavage of carbon-carbon bond. Carbon and hydrogen atoms in the resulting alkyl ions are equivalent to being fully scrambled through carbon skeleton rearrangement or hydrogen shift. Subsequent fragmentation of the resulting alkyl ions forms a lower homologue product ion.

The first-order finding was that product fragment ions containing at least half of the number of carbon atoms in the molecular ions have about 80% probability to come from direct cleavage of the molecular ions, whereas this probability for smaller product fragment ions is about 15% and consecutive pathways of reaction (4-5) and (4-6) contribute to most of the smaller product fragment ion formation (Figure 4.4).

Quantitative fragmentation schemes were fully determined for *n*-hexane, *n*-heptane, and *n*-nonane by fitting of the contribution of each specific pathway to the abundance of ^{13}C- and ^{2}H-containing fragments (Table 4.4) (Lavanchy et al., 1978, 1979).

Table 4.4 Fragmentation schemes for *n*-hexane, *n*-heptane, and *n*-nonane. Relative contribution from decomposition pathways through primary (intermediate) product and secondary (final) product ions are listed in units of %TIC (total ion current). In each row, the primary (intermediate) fragment ions are products from fragmentation of the molecular ion. For each column, secondary (final) fragment ions are products from fragmentation of the primary (intermediate) fragment ion. The number in row $C_nH_{2n+1}^+$ and column $C_kH_{2k+1}^+$ corresponds to the contribution of pathway $C_mH_{2m+2}^+$ (molecular ion) $\rightarrow C_nH_{2n+1}^+ \rightarrow C_kH_{2k+1}^+$ to TIC. The total of relative abundance of all product ions for each n-alkane is not 100% because fragment ions other than CnH2n+1 also contribute to TIC (e.g., CnH2n-1, CnH2n-3). The lower left parts of the scheme matrices are empty because recombination reactions are not considered.

n-C6H14		secondary (final) product				
		C6H13+	C5H11+	C4H9+	C3H7+	C2H5+
primary (intermediate) product	C6H13+	0	0	1.9	0	0
	C5H11+		1.3	0	3.9	0
	C4H9+			17.4	0	7
	C3H7+				10.1	0
	C2H5+					0.9

n-C7H16		secondary (final) product					
		C7H15+	C6H13+	C5H11+	C4H9+	C3H7+	C2H5+
primary (intermediate) product	C7H15+	0	0	3.7	0	0	0
	C6H13+		0.7	0	2.2	1.2	0
	C5H11+			7.9	0	14.9	0
	C4H9+				8.7	0	5.5
	C3H7+					4.6	0
	C2H5+						0.9

n-C9H20		secondary (final) product						
		C9H19+	C8H17+	C7H15+	C6H13+	C5H11+	C4H9+	C3H7+
primary (intermediate) product	C9H19+	0	0	1.7	0	0	0	0
	C8H17+		0	0	1.5	2.2	2	0
	C7H15+			1	0	0	12.8	0
	C6H13+				6.8	0	0.5	10.2
	C5H11+					3.3	0	6.3
	C4H9+						2.5	0
	C3H7+							2.7
	C2H5+							

Each row in Table 4.4 represents a primary (intermediate) fragment ion that is a product from fragmentation of the molecular ion. For each column, secondary (final) fragment ions are products from fragmentation of the primary (intermediate) fragment ion. I.e., a number in row $C_nH_{2n+1}^+$ and column $C_kH_{2k+1}^+$ corresponds to the contribution of pathway $C_mH_{2m+2}^+$ (molecular ion) → $C_nH_{2n+1}^+$ → $C_kH_{2k+1}^+$. For example, for n-C6H14, 24.4% (17.4%+7%) of molecular ion was fragmented to C4H9+ initially. 17.4% ended up as C4H9+ and 7% continued to get fragmented to C2H5+.

From the fragmentation scheme matrices for *n*-hexane, *n*-heptane, and *n*-nonane, we can make five inferences through inductive reasoning:

1. The superdiagonal: the set of elements directly above the elements comprising the diagonal are zero, i.e., the pathways $C_mH_{2m+2}^+$ (molecular ion) → $C_nH_{2n+1}^+$ → $C_{n-1}H_{2n-1}^+$ are forbidden.
2. The top row: the elements in the top row are zero except for the second element, i.e., the pathways $C_mH_{2m+2}^+$ (molecular ion) → $C_mH_{2m+1}^+$ → $C_kH_{2k-1}^+$ are forbidden except for $C_mH_{2m+2}^+$ (molecular ion) → $C_mH_{2m+1}^+$ → $C_{m-2}H_{2m-3}^+$.
3. The second row: the elements in the second row for secondary fragments with less than about a third of the number of carbon atoms in the molecular ion are zero.
4. The sum of each column: the elements in each column sum up to the relative abundance of final product fragment ions, which can be quantified in mass spectra.
5. The sum of each row: the elements of each row sum up to the relative abundance of primary fragment ions from the initial fragmentation of the molecular ion. The relative abundance of primary fragment product appears to follow normal distribution (Figure 4.5). We fitted the fragmentation scheme data of *n*-hexane, *n*-heptane, and *n*-nonane from Table 4.4 with

$$I\left(C\#_{frag}/C\#_{mol}\right) = A(C\#_{mol}) \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (4\text{-}7)$$

where I is the ion intensity of primary fragment products, $C\#_{frag}$ and $C\#_{mol}$ are carbon numbers of primary fragment products and molecular ions, σ and μ are the standard deviation and mean of the normal distribution, and A is a prefactor that is a function of the carbon number of molecular ions. The fitted parameters are listed in Table 4.5.

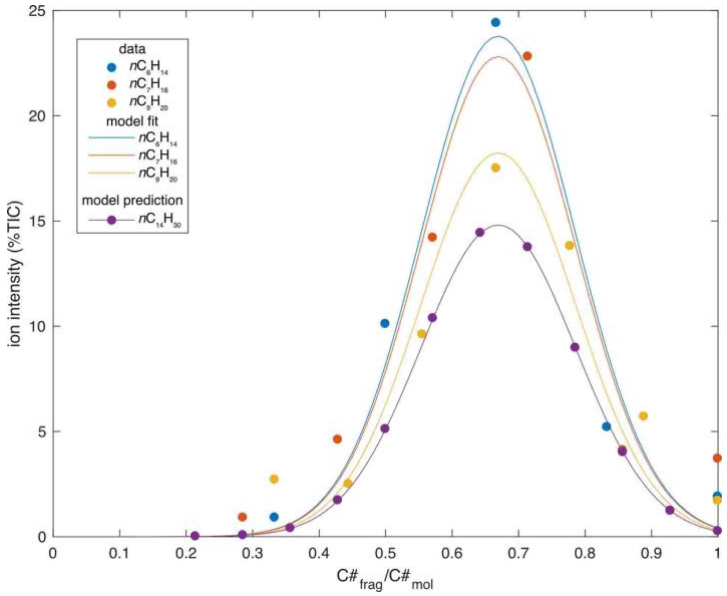


Figure 4.5 The distribution of primary fragment products for *n*-hexane, *n*-heptane, *n*-nonane, and *n*-tetradecane, showing the relative abundance of primary fragment ions ($C_nH_{2n+1}^+$) produced directly from the initial fragmentation of the molecular ions ($C_mH_{2m+2}^+$). The x-axis represents the ratio of the carbon number in the primary fragment ion (n) to the carbon number in the parent molecular ion (m). The y-axis shows the relative abundance of each primary fragment ion as %TIC. The data points are fitted with a normal distribution curve (Equation (4-7)), demonstrating a consistent fragmentation pattern across different *n*-alkane chain lengths.

Table 4.5 Parameters of the normal distribution of primary fragment products for *n*-hexane, *n*-heptane, *n*-nonane, and n-tetradecane

parameter		value
σ		0.1161
μ		0.6694
A	*n*C6	6.9193
	*n*C7	6.6377
	*n*C9	5.3056
	*n*C14	4.3101

3.2.3. Modeling the EI fragmentation of *n*-tetradecane

Lavanchy and colleagues measured the mass spectra of site-specific labeled *n*-tetradecane [1,14-$^{13}C_2$], [5,10-$^{13}C_2$], [6,9-$^{13}C_2$], [7,8-$^{13}C_2$], [5,7,8,10-$^{13}C_4$], [6,7,8,9-$^{13}C_4$], [1,7,8,14-$^{13}C_4$], and [5,6,7,8,9,10-$^{13}C_6$] and calculated the relative abundance of fragment ions with total retention of ^{13}C labels (Lavanchy et al., 1979) but only derived a partial fragmentation scheme for product ions $C_{14}H_{29}^+$ to $C_8H_{17}^+$ due to lack of constraints (Table 4.7) (Lavanchy, 1977). Here we built on their effort and derived the full fragmentation scheme for *n*-tetradecane using their fitting approach and the inferences above as additional constraints. The code to achieve this is available here (doi.org/10.22002/y21ya-4xb50)

3.2.3.1. Fit the contribution of each pathway to labeled n-tetradecane data

For a given labeled *n*-tetradecane, many decomposition pathways may result in the formation of a particular fragment retaining the total ^{13}C labels. The intensity of this fragment is the sum of the contribution of the different reaction pathways that form this fragment in different abundance. We constructed an array of the unknown contributions of these pathways to fit. We modeled each decomposition pathway that may result in the formation of fragment $C_7H_{15}^+$ to $C_3H_7^+$ and obtained the relative abundance of ions with total ^{13}C label retentions. The experimentally determined relative abundance of total retention for each fragment and each labeled *n*-tetradecane (Table 4.6) constitutes 34 constraints to fit.

Table 4.6 Relative abundance (%) of final fragment products with total label retention from site-specific ^{13}C-labeled *n*-tetradecane (Lavanchy et al., 1979)

13C-labeled positions of	Relative abundance of total label retention in fragment products (%)				
tetradecane	C7H15	C6H13	C5H11	C4H9	C3H7
1, 14	0	2	1	0	0
5, 10	28	14	8	3	1
6, 9	22	22	7	6	2
7, 8	25	26	22	12	5
5, 7, 8, 10	8	4	1	0	
6, 7, 8, 9	8	0	4	1	
1, 7, 8, 14	0	0	0	0	
5, 6, 7, 8, 9, 10	4	1			

3.2.3.2. Additional constraints: zero contribution pathways

We applied inferences 1-3 and set the elements in the superdiagonal of the fragmentation scheme matrix of *n*-tetradecane and the top two rows for final products $C_7H_{15}^+$ to $C_3H_7^+$ as zero (Table 4.7). With these setting, we reduced the number of unknown contributions to final products $C_7H_{15}^+$ to $C_3H_7^+$ to 35 pathways.

3.2.3.3. Additional constraints: distribution of final products

We applied inference 4 and set the sum of each column of final products $C_7H_{15}^+$ to $C_3H_7^+$ according to mass spectrum measurement (Table 4.7) (Lavanchy et al., 1979). This constitutes 5 constraints.

3.2.3.4. Additional constraint: modeling the distribution of primary fragment products

We applied inference 5 and fitted the sum of relative abundance of primary products from *n*-tetradecane molecular ion to their total %TIC from mass spectrum measurement (Lavanchy et al., 1979) by fitting the prefactor A in equation (4-7) (Table 4.5, Figure 4.5). The relative abundance of primary products set the sum of each row in the fragmentation scheme matrix. This constitutes 10 constraints.

3.2.3.5. Fit the pathway contribution array

We fitted the 35 unknown pathway contributions to a total of 49 constraints and obtained the least-square fitting (Table 4.7). The resulting fragmentation scheme fitted the labeled *n*-tetradecane data to a standard deviation of 1.8%, which is comparable to that of previous studies (Lavanchy et al., 1978). We took this result to study fragmentation of longer *n*-alkanes.

Table 4.7 Fragmentation scheme for *n*-tetradecane. Relative contribution from decomposition pathways of *n*-tetradecane through primary (intermediate) product and secondary (final) product ions are listed in units of %TIC. The tabulating logic is similar to Table 4.4. The sum of each column and each row are listed below the column head and next to the row head. The fragmentation scheme for final products $C_{14}H_{29}^+$ to $C_8H_{17}^+$ and the sum of each column are from (Lavanchy, 1977). The sum of each row and the bolded elements for final products $C_7H_{15}^+$ to $C_3H_7^+$ are fitted in section 3.2.3.

n-C14H30			secondary (final) product											
			C14H29	C13H27	C12H25	C11H23	C10H21	C9H19	C8H17	C7H15	C6H13	C5H11	C4H9	C3H7
		sum	0	0	0.4	0.8	1.2	1.4	1.6	2.1	7.5	11.5	18.5	15.3
primary (intermediate) product	C14H29	0.26	0	0	0.12	0	0	0	0	0	0	0	0	0
	C13H27	1.23		0	0	0.16	0	0.32	0	0	0	0	0	0
	C12H25	4.01			0.29	0	0.17	0	0.5	**0.22**	**2.91**	**0.00**	**0.00**	**0.00**
	C11H23	8.97				0.59	0	0	0	**0.00**	**1.70**	**0.00**	**2.92**	**3.84**
	C10H21	13.75					1.01	0	0	**0.97**	**0.00**	**4.34**	**3.53**	**3.98**
	C9H19	14.43						1.08	0	**0.00**	**1.14**	**0.07**	**5.25**	**6.97**
	C8H17	10.37							1.12	0	**0.68**	**4.49**	**3.75**	**0.41**
	C7H15	5.11								**0.90**	0	**2.12**	**2.17**	**0.00**
	C6H13	1.72									**1.07**	0	**0.74**	**0.00**
	C5H11	0.40										**0.48**	0	**0.00**
	C4H9	0.06											**0.15**	0
	C3H7	0.01												**0.09**

3.2.4. EI fragmentation of *n*-hexadecane: Site-specific labeling measurements and modeling

We investigated the fragmentation of *n*-hexadecane on GC-Orbitrap by measuring the mass spectra of site-specific labeled hexadecanes and modeling the fragmentation scheme.

3.2.4.1. Site-specific labeling measurement results

The fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ion of 0, 5, 10, and 20% 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes is listed in Table 4.8 and shown in Figure 4.6. The linear regression of fractional abundance represents a mixing line between two endmembers: natural abundance *n*-hexadecane at 0% label level and pure site-specific labeled hexadecane at 100% label level. The intercept of the mixing line is the natural fractional abundance of the corresponding isotopologue in the unlabeled *n*-hexadecane and the slope is the difference in the fractional abundance of the corresponding isotopologues in the pure site-specific labeled hexadecane and the unlabeled *n*-hexadecane.

Similar mixing line can be plotted for modeled fractional abundance by a single cut model, where all fragment ions are assumed as products of a single C-C cleavage from the molecular ions, i.e., from reaction (4-5) only. We calculated the intercept, i.e., the natural fractional abundance of isotopologues, based on the molecular average $\delta^{13}C$ and $\delta^{2}H$ of the unlabeled *n*-hexadecane (Table 4.10) and assuming a stochastic distribution (rare isotopes are randomly distributed at different positions of the molecule). We derived the slope by modeling reaction (1–4) for 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes. The single cut model is not consistent with the measurement results (Figure 4.6).

Table 4.8 Fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ion of 0, 5, 10, and 20% 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes by direct elution measurements. The fractional abundance is calculated as the ratio of the ion count of the substituted fragment to the sum of the ion counts of the unsubstituted, ^{13}C-substituted, $^{13}C_2$-substituted, and ^{2}H-substituted fragments.

label type	measurement date	label level (%)	analysis number	fractional abundance					
				$[^{13}C^{12}C_3H_9]$	$[^{13}C^{12}C_4H_{11}]$	$[^{13}C^{12}C_5H_{13}]$	$[^{13}C_2{}^{12}C_2H_9]$	$[^{13}C_2{}^{12}C_3H_{11}]$	$[^{13}C_2{}^{12}C_4H_{13}]$
1,2-$^{13}C_2$	20220525	0	1	0.042562	0.050047	0.057174	0.000280	0.000484	0.000523
			2	0.044592	0.051922	0.058649	0.000383	0.000774	0.001334
			3	0.043624	0.051643	0.058940	0.000334	0.000793	0.001030
			4	0.043097	0.052465	0.058897	0.000393	0.000669	0.001392
		5	1	0.051469	0.062637	0.066956	0.003413	0.007513	0.009837
			2	0.050274	0.062641	0.066769	0.003259	0.007113	0.009540
			3	0.050728	0.062040	0.067097	0.003534	0.006911	0.010143
		10	1	0.062545	0.075660	0.081344	0.007982	0.015757	0.021785
			2	0.060360	0.074634	0.083314	0.007607	0.015686	0.021459
			3	0.060924	0.076041	0.079818	0.007489	0.015552	0.020856
		20	1	0.079441	0.101340	0.104941	0.015005	0.029750	0.043514
			2	0.078982	0.100957	0.102649	0.015190	0.030973	0.042815
			3	0.076787	0.102347	0.101808	0.015149	0.030836	0.041146
	20220612	0	1	0.043585	0.052423	0.059797	0.000234	0.000395	0.000656
			2	0.044081	0.051876	0.060360	0.000385	0.000458	0.000900
			3	0.044837	0.052579	0.059340	0.000300	0.000682	0.000678
			4	0.044265	0.052339	0.060480	0.000215	0.000532	0.000869
			5	0.044326	0.050368	0.058682	0.000329	0.000546	0.000894
			6	0.045579	0.051753	0.058985	0.000305	0.000702	0.000928
			7	0.044374	0.051985	0.058254	0.000278	0.000498	0.000950
		5	1	0.052015	0.062737	0.068463	0.003909	0.007287	0.009846
			2	0.053435	0.063318	0.068495	0.003795	0.007242	0.009871
			3	0.053012	0.064018	0.067800	0.004115	0.006912	0.010308
			4	0.051596	0.061911	0.069035	0.003855	0.006967	0.009195
			5	0.053584	0.061822	0.069290	0.004082	0.006782	0.009775
			6	0.052744	0.062355	0.069854	0.003417	0.007174	0.009766
		10	1	0.062963	0.076496	0.080610	0.009537	0.014958	0.020904
			2	0.064237	0.079305	0.081269	0.008988	0.016697	0.022646
			3	0.063007	0.075879	0.079661	0.008983	0.016854	0.020674
			4	0.064775	0.076976	0.082160	0.009628	0.015859	0.021024
			5	0.063606	0.076969	0.081934	0.009426	0.016128	0.022179
			6	0.065051	0.076881	0.079208	0.009255	0.016118	0.022448
		20	1	0.083683	0.102997	0.104827	0.017800	0.032045	0.042797
			2	0.084133	0.099771	0.103180	0.017725	0.030312	0.041162
			3	0.086764	0.102733	0.106193	0.018145	0.033139	0.043764
			4	0.084952	0.102801	0.107252	0.017905	0.032206	0.042408
			5	0.084721	0.104827	0.105854	0.018437	0.031745	0.043017
			6	0.086087	0.103199	0.104466	0.018660	0.031311	0.043548

label type	measurement date	label level (%)	analysis number	fractional abundance [$^{12}C_4{}^{2}H^{1}H_8$]	[$^{12}C_5{}^{2}H^{1}H_{10}$]	[$^{12}C_6{}^{2}H^{1}H_{12}$]
1-^{2}H	20220525	0	1	0.000811	0.000898	0.001456
			2	0.000965	0.001196	0.001404
			3	0.000859	0.000804	0.001306
			4	0.000719	0.001154	0.001208
		5	1	0.008873	0.014054	0.014982
			2	0.008591	0.014865	0.015789
			3	0.008302	0.014239	0.015969
		10	1	0.017442	0.028858	0.031644
			2	0.016723	0.029110	0.031179
			3	0.016724	0.028332	0.032133
		20	1	0.030064	0.052777	0.058633
			2	0.032129	0.054474	0.060994
			3	0.031821	0.053987	0.059781
	20220612	0	1	0.000716	0.000754	0.000979
			2	0.000772	0.000946	0.001275
			3	0.000737	0.001123	0.001048
			4	0.000658	0.000887	0.001259
			5	0.000722	0.001054	0.001278
			6	0.000665	0.001058	0.001662
			7	0.001069	0.000928	0.001343
		5	1	0.008692	0.012383	0.013171
			2	0.010283	0.014751	0.016671
			3	0.009620	0.014507	0.016608
			4	0.010120	0.014443	0.016022
			5	0.009455	0.014483	0.016070
			6	0.010215	0.014693	0.016656
		10	1	0.017840	0.027112	0.030435
			2	0.019339	0.028270	0.031530
			3	0.018101	0.028227	0.031706
			4	0.019222	0.028229	0.032598
			5	0.019241	0.028894	0.032363
			6	0.019757	0.027242	0.031103
		20	1	0.031198	0.045590	0.052047
			2	0.031866	0.050600	0.055567
			3	0.034924	0.052408	0.060068
			4	0.035176	0.054093	0.061650
			5	0.035899	0.054470	0.060535
			6	0.035446	0.052510	0.059504

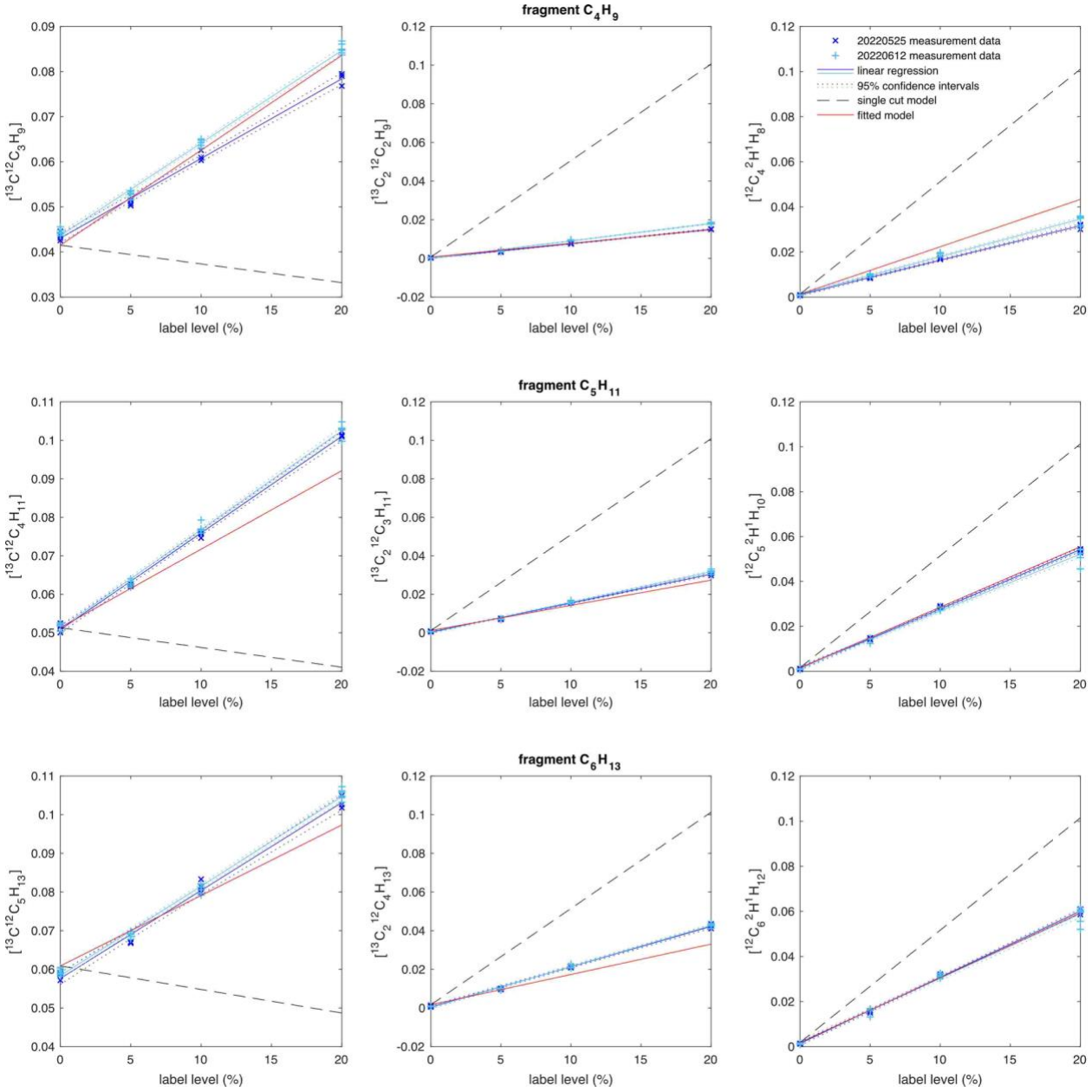


Figure 4.6 Linear regression of measured fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ion of 0, 5, 10, and 20% 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes, compared with modeled results. The single cut model assumes all the fragment ions are products of a single C-C cleavage from the molecular ion (reaction (4-5)), without considering secondary fragmentation processes. The fitted model incorporates the full fragmentation scheme derived Table 4.9 and is fitted in section 3.2.4-3.2.4.2. The fractional abundance is calculated as the ratio of the ion count of the

substituted fragment to the sum of the ion counts of the unsubstituted, ^{13}C-substituted, $^{13}C_2$-substituted, and ^{2}H-substituted fragments.

3.2.4.2. Modeling the EI fragmentation of *n*-hexadecane

We modeled the *n*-hexadecane fragmentation pathways that contribute to C_4H_9, C_5H_{11}, and C_6H_{13} fragments by adapting the *n*-tetradecane fragmentation scheme, applying the inferences, and fitting to the labeling measurement results in the same approach to section 3.2.3. The code to achieve this is available here (doi.org/ 10.22002/j6sxx-03b98).

Briefly, we set the contribution of fragmentation pathways through primary products $C_{12}H_{25}^+$ to $C_4H_9^+$ the same as the fragmentation scheme of *n*-tetradecane. We also set the top two rows as zero according to inference 2 and 3. We constructed an array of the unknown contributions of pathways $C_{16}H_{34} \rightarrow C_{14}H_{29}/C_{13}H_{27} \rightarrow C_4H_9/C_5H_{11}/C_6H_{13}$ to fit. We modeled these pathways for 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes and obtained the relative abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted fragment products. The experimentally determined relative abundance of isotopologues of each fragment for two labeled hexadecane constitutes 9 independent constraints. We added two constraints on the proportions of C_4H_9, C_5H_{11}, and C_6H_{13} final product ions based on their relative abundance in NIST Mass Spectral Library. We fitted the six unknown pathway contributions to a total of 11 constraints and obtained the least-square fitting (Table 4.8, Figure 4.6). The resulting fragmentation scheme fitted the labeled hexadecane data to a standard deviation of 3.5%, which is generally comparable to that of *n*-tetradecane fitting and previous studies.

Table 4.9 Contribution of *n*-hexadecane decomposition pathways to C_4H_9, C_5H_{11}, and C_6H_{13} fragments in units of %TIC through primary (intermediate) products. The tabulating logic is similar to Table 4.4 and Table 4.7. The relative abundance of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions in *n*-hexadecane mass spectrum from NIST Mass Spectral Library is listed below the column head, normalized to the abundance of C_4H_9. The contribution of fragmentation pathways through primary products $C_{12}H_{25}^+$ to $C_4H_9^+$ is adapted from the fragmentation scheme of *n*-tetradecane. The bolded elements are fitted in section 3.2.4-3.2.4.2.

n-C16H34		secondary (final) product		
		C6H13	C5H11	C4H9
	relative abundance	39	60	100
primary (intermediate) product	C16H33	0	0	0
	C15H31	0	0	0
	C14H29	**0.00**	**0.00**	**11.34**
	C13H27	**5.23**	**7.60**	**0.00**
	C12H25	2.91	0.00	0.00
	C11H23	1.70	0.00	2.92
	C10H21	0.00	4.34	3.53
	C9H19	1.14	0.07	5.25
	C8H17	0.68	4.49	3.75
	C7H15	0	2.12	2.17
	C6H13	1.07	0	0.74
	C5H11		0.48	0
	C4H9			0.15

We compared the fitted model result with measurements of 0.14% 1,2-$^{13}C_2$-hexadecane (molecular average $\delta^{13}C$ +15‰) and 0.05% 1-^{2}H-hexadecane (molecular average $\delta^{2}H$ +100‰), which are more comparable to natural isotopic variations (Table 4.10). The slope of the fitted model is consistent with that of linear regressions for 0 and 0.14% 1,2-$^{13}C_2$- and 0 and 0.05% 1-^{2}H-labeled hexadecanes (Figure 4.7). The discrepancy in the intercept may be due to different instrument fractionation or rare isotope discrimination between direct elution and peak capture measurements.

The slopes for sub % labeled hexadecanes also agree with those for tens of % labels. While we lack more rigorous constraints to test the accuracy of the measurements, this agreement is consistent with the measurements being accurate.

Table 4.10 Fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ion of 0 and 0.14% 1,2-$^{13}C_2$- and 0 and 0.05% 1-^{2}H-labeled hexadecanes by peak capture measurements. The fractional abundance is calculated as the ratio of the ion count of the substituted fragment to the sum of the ion counts of the unsubstituted, ^{13}C-substituted, $^{13}C_2$-substituted, and ^{2}H-substituted fragments.

sample	molecular average $\delta^{13}C$ (‰)	label level (%)	analysis number	$[^{13}C^{12}C_3H_9]$	$[^{13}C^{12}C_4H_{11}]$	$[^{13}C^{12}C_5H_{13}]$	$[^{13}C_2{}^{12}C_2H_9]$	$[^{13}C_2{}^{12}C_3H_{11}]$	$[^{13}C_2{}^{12}C_4H_{13}]$
n-$C_{16}H_{34}$	-29.9	0	1	0.046951	0.056036	0.060811	0.000759	0.001194	0.001530
			2	0.046911	0.056086	0.060901	0.000763	0.001190	0.001526
			3	0.046961	0.056057	0.060834	0.000754	0.001196	0.001533
			4	0.046859	0.056107	0.060879	0.000750	0.001190	0.001531
			5	0.046876	0.056050	0.060928	0.000752	0.001199	0.001531
n-$C_{16}H_{34}$ + 15‰ $\delta^{13}C$	-15.4	0.14	1	0.047177	0.056308	0.061075	0.000936	0.001403	0.001750
			2	0.047195	0.056351	0.061017	0.000944	0.001397	0.001760
			3	0.047155	0.056376	0.061130	0.000938	0.001403	0.001759
			4	0.047142	0.056370	0.061108	0.000948	0.001404	0.001753

sample	molecular average $\delta^{2}H$ (‰)	label level (%)	analysis number	$[^{12}C_4{}^{2}H^{1}H_8]$	$[^{12}C_5{}^{2}H^{1}H_{10}]$	$[^{12}C_6{}^{2}H^{1}H_{12}]$
n-$C_{16}H_{34}$	-22	0	1	0.001313	0.001502	0.001621
			2	0.001326	0.001502	0.001623
			3	0.001307	0.001497	0.001619
			4	0.001312	0.001490	0.001629
			5	0.001312	0.001486	0.001615
n-$C_{16}H_{34}$ + 100‰ $\delta^{2}H$	81	0.05	1	0.001475	0.001680	0.001798
			2	0.001461	0.001663	0.001789
			3	0.001459	0.001679	0.001785
			4	0.001471	0.001660	0.001793

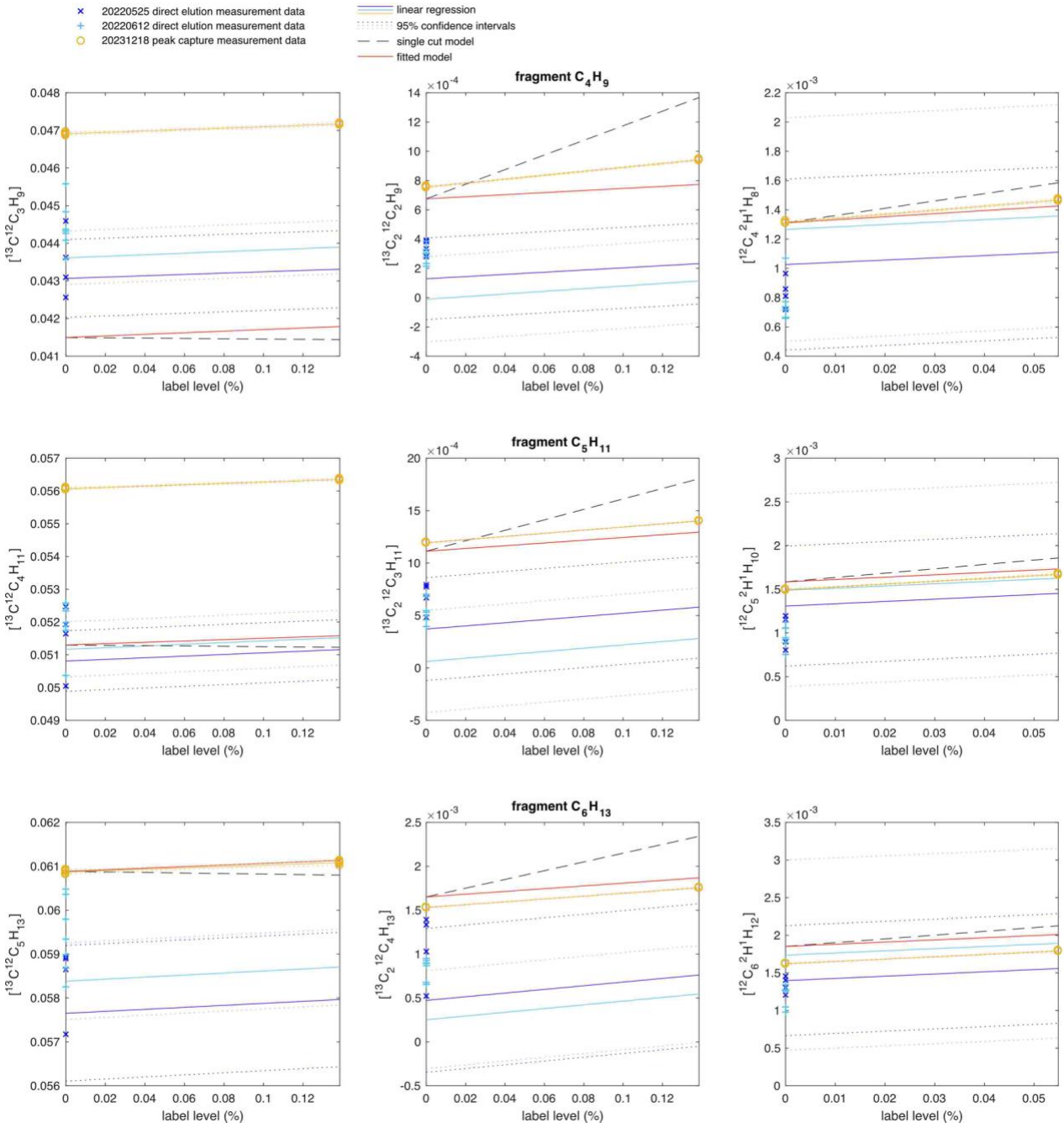


Figure 4.7 Linear regression of measured fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ion of 0 and 0.14% 1,2-$^{13}C_2$- and 0 and 0.05% 1-^{2}H-labeled hexadecanes. 0–20% 1,2-$^{13}C_2$- and 1-^{2}H-labeled data, linear regressions and single cut and fitted models from Figure 4.6 are also shown. The slope of the linear regressions for 0 and 0.14% 1,2-$^{13}C_2$- and 0

and 0.05% 1-^{2}H-labeled hexadecanes are consistent with the fitted model. The fractional abundance is calculated as the ratio of the ion count of the substituted fragment to the sum of the ion counts of the unsubstituted, ^{13}C-substituted, $^{13}C_2$-substituted, and ^{2}H-substituted fragments.

We ran Monte-Carlo simulation on the fragmentation scheme of *n*-hexadecane. Figure 4.8 shows probability distribution of the standard deviation of the fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted C_4H_9, C_5H_{11}, and C_6H_{13} fragments from 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes. A standard deviation of 3.5% by the fitted *n*-hexadecane fragmentation model ranks top 0.3% of the distribution, very close to the minimum possible standard deviation. The standard deviation is not very sensitive to the contribution of decompositions pathways in the fragmentation scheme as many different fragmentation schemes can achieve standard deviation comparable to previous studies (Lavanchy et al., 1978). Future work may add more constraints from other site-specific labeled hexadecanes.

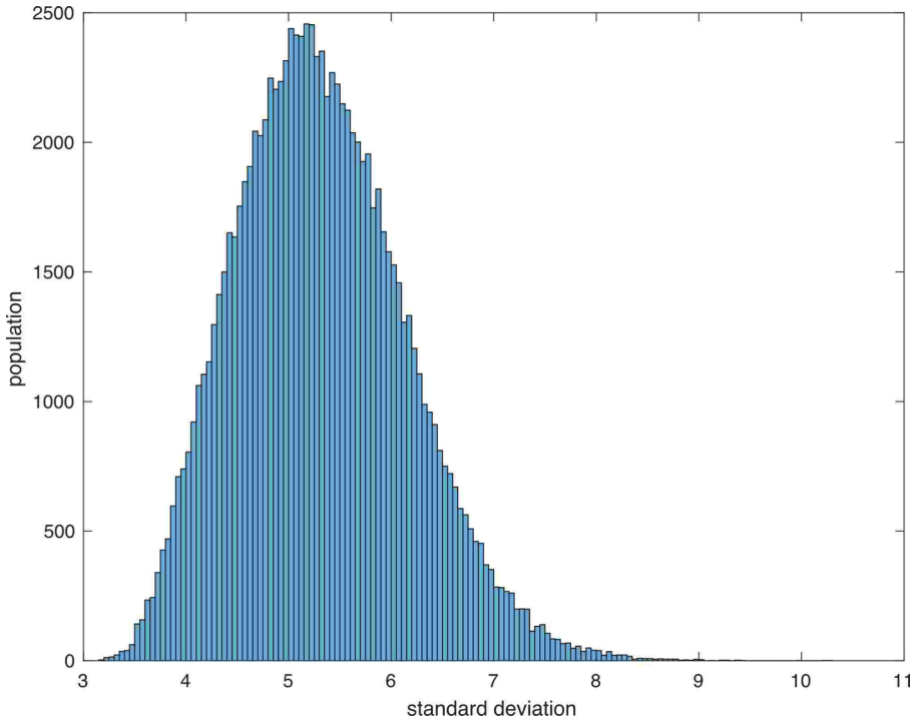


Figure 4.8 Probability distribution of the standard deviation of fractional abundance of ^{13}C-, $^{13}C_2$-, and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragments from 1,2-$^{13}C_2$- and 1-^{2}H-labeled hexadecanes,

obtained from a Monte Carlo simulation of the n-hexadecane fragmentation scheme. A standard deviation of 3.5% by the fitted *n*-hexadecane fragmentation model ranks top 1% of the distribution. The sensitivity of the standard deviation to the fragmentation scheme provides insight into the robustness of the derived scheme and the potential for alternative schemes to fit the experimental data.

Possible sources of error in the fitted fragmentation scheme of *n*-hexadecane include: (1) instrument fractionation on fragment fractional abundance data. For example, KIE in ionization/fragmentation, especially for ^{2}H (Howe, 1975; Wolkoff and Holmes, 1978; Holmes et al., 1982), and isotope fractionation in mass filtering and transmission processes (Eiler et al., 2017). (2) Scale stretch or compression in fragment fractional abundance between pure site-specific isotope labels and labels of different % levels, possibly from non-linearity of Orbitrap mass analyzer (Eiler et al., 2017). However, the fitted fragmentation scheme of *n*-hexadecane has overall good metrics (standard deviation) and is internally consistent with those of shorter *n*-alkanes. So we took this result to study *n*-nonadecane.

3.2.5. Fragmentation of *n*-nonadecane

Here we proposed a possible fragmentation scheme for C_4H_9, C_5H_{11}, and C_6H_{13} fragment products of *n*-nonadecane using a Bayesian approach without access to site-specific labeled *n*-nonadecane standards and used it as a transfer function to test hypotheses on site-specific isotopic structure of rose *n*-nonadecane. We also tested the sensitivity of possible fragmentation schemes by Monte-Carlo simulation.

3.2.5.1. Proposing a possible fragmentation scheme for *n*-nonadecane

We adapted the *n*-hexadecane fragmentation scheme by setting the contribution of fragmentation pathways through primary products $C_{14}H_{29}^{+}$ to $C_4H_9^{+}$ the same as the fragmentation scheme of *n*-hexadecane. We also set the contribution of fragmentation pathways through primary products $C_{19}H_{39}^{+}$ and $C_{18}H_{37}^{+}$ as zero according to inference 2 and 3. Next, we proposed the possible contribution of pathways $C_{19}H_{40} \rightarrow C_{17}H_{35}/C_{16}H_{33}/C_{15}H_{31}$

→ $C_4H_9/C_5H_{11}/C_6H_{13}$ (Table 4.11). We noted that the only constraints we have on the proposed contribution were the proportions of C_4H_9, C_5H_{11}, and C_6H_{13} final product ions based on their relative abundance in NIST Mass Spectral Library (Table 4.11). But we tested the sensitivity of the contribution in section.

Table 4.11 Contribution of *n*-nonadecane decomposition pathways to C_4H_9, C_5H_{11}, and C_6H_{13} fragments in units of %TIC through primary (intermediate) products. The tabulating logic is similar to Table 4.4, Table 4.7 and Table 4.9. The relative abundance of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions in *n*-nonadecane mass spectrum from NIST Mass Spectral Library is listed below the column head, normalized to the abundance of C_4H_9. The contribution of fragmentation pathways through primary products $C_{14}H_{29}^+$ to $C_4H_9^+$ is adapted from the fragmentation scheme of *n*-hexadecane. The bolded elements are proposed.

n-C19H40		secondary (final) product		
		C6H13	C5H11	C4H9
relative abundance		43	66	100
	C19H39	0	0	0
	C18H37	0	0	0
	C17H35	**0.00**	**0.00**	**5.00**
	C16H33	**3.00**	**4.00**	**0.00**
	C15H31	**0.00**	**0.00**	**0.00**
	C14H29	0.00	0.00	11.34
	C13H27	5.23	7.60	0.00
	C12H25	2.91	0.00	0.00
primary	C11H23	1.70	0.00	2.92
(intermediate)	C10H21	0.00	4.34	3.53
product	C9H19	1.14	0.07	5.25
	C8H17	0.68	4.49	3.75
	C7H15	0	2.12	2.17
	C6H13	1.07	0	0.74
	C5H11		0.48	0
	C4H9			0.15

3.2.5.2. Testing hypothetical site-specific isotope composition of rose n-nonadecane with the proposed fragmentation scheme

We tested five hypothetical site-specific carbon and hydrogen isotope compositions of rose *n*-nonadecane and three of those for the *n*-nonadecane working standard CIT-C19 (Table 4.12, Figure 4.9). The site-specific isotope structures are all symmetrical due to the symmetry

of *n*-nonadecane molecule. Biological aliphatic compounds are thought to have a zigzag pattern in the intramolecular isotopic structure due to assembling from two carbon unit monomers, i.e., the odd-numbered and even-numbered carbon or hydrogen positions have contrasting isotope composition (Hayes, 2001). $\delta^{13}C_{VPDB}$ of rose *n*-nonadecane structure #1 was constructed based on site-specific $\delta^{13}C_{VPDB}$ of plant fatty acids (Julien et al., 2022), #2 based on *E. coli* fatty acids (Monson and Hayes, 1980), and #3 on *S. cerevisiae* fatty acids (Monson and Hayes, 1982). Rose *n*-nonadecane structure #4 has constant $\delta^{13}C_{VPDB}$ values across all positions whereas #5 is the same except for carbon position 6 and 14 (symmetrical equivalent positions) that are -49‰. Site-specific $\delta^{13}C_{VPDB}$ of CIT-C19 #1 was hypothesized to have flat and constant values whereas #2 was constructed with an upside-down V-shaped curve for carbon position number versus site-specific $\delta^{13}C$, possibly like a residue of thermal maturation in petroleum system. CIT-C19 #3 has a V-shaped carbon position number versus site-specific $\delta^{13}C_{VPDB}$ curve. Site-specific $\delta^{2}H_{VSMOW}$ structure #1, #2, #3, and #4 of rose *n*-nonadecane mirror those of $\delta^{13}C_{VPDB}$ after adjusted for different scales of isotopic variations. Rose *n*-nonadecane structure #5 has constant $\delta^{2}H_{VSMOW}$ values across all positions except for hydrogen position 6 and 14 that are -40‰. Site-specific $\delta^{2}H_{VSMOW}$ of CIT-C19 #1 was hypothesized to be constant across all hydrogen positions, #2 was constructed as a V-shape and #3 an upside-down V-shape, possibly like a snapshot of partial hydrogen isotope equilibration in petroleum system.

Table 4.12 Hypothetical site-specific carbon and hydrogen isotope composition of rose *n*-nonadecane and CIT-C19 working standard. For description see Figure 4.9 and section 3.2.5-3.2.5.2.

			carbon or hydrogen position number																		
isotope composition	sample	structure #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$\delta^{13}C$	rose	1	-29.4	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-29.4
		2	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9
		3	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9	-36.9	-21.9
		4	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0
		5	-29.0	-29.0	-29.0	-29.0	-29.0	-49.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-29.0	-49.0	-29.0	-29.0	-29.0	-29.0	-29.0
	CIT-C19	1	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0	-30.0
		2	-39.9	-36.3	-33.1	-30.7	-28.7	-27.2	-26.1	-25.5	-25.0	-24.9	-25.0	-25.5	-26.1	-27.2	-28.7	-30.7	-33.1	-36.3	-39.9
		3	-20.1	-23.7	-26.9	-29.3	-31.3	-32.8	-33.9	-34.5	-35.0	-35.1	-35.0	-34.5	-33.9	-32.8	-31.3	-29.3	-26.9	-23.7	-20.1
$\delta^{2}H$	rose	1	-243	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-243
		2	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303
		3	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183	-303	-183
		4	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240	-240
		5	-240	-240	-240	-240	-240	-40	-240	-240	-240	-240	-240	-240	-240	-40	-240	-240	-240	-240	-240
	CIT-C19	1	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103	-103
		2	-235	-186	-145	-113	-86	-66	-51	-42	-37	-35	-37	-42	-51	-66	-86	-113	-145	-186	-235
		3	29	-20	-61	-93	-120	-140	-155	-164	-169	-171	-169	-164	-155	-140	-120	-93	-61	-20	29

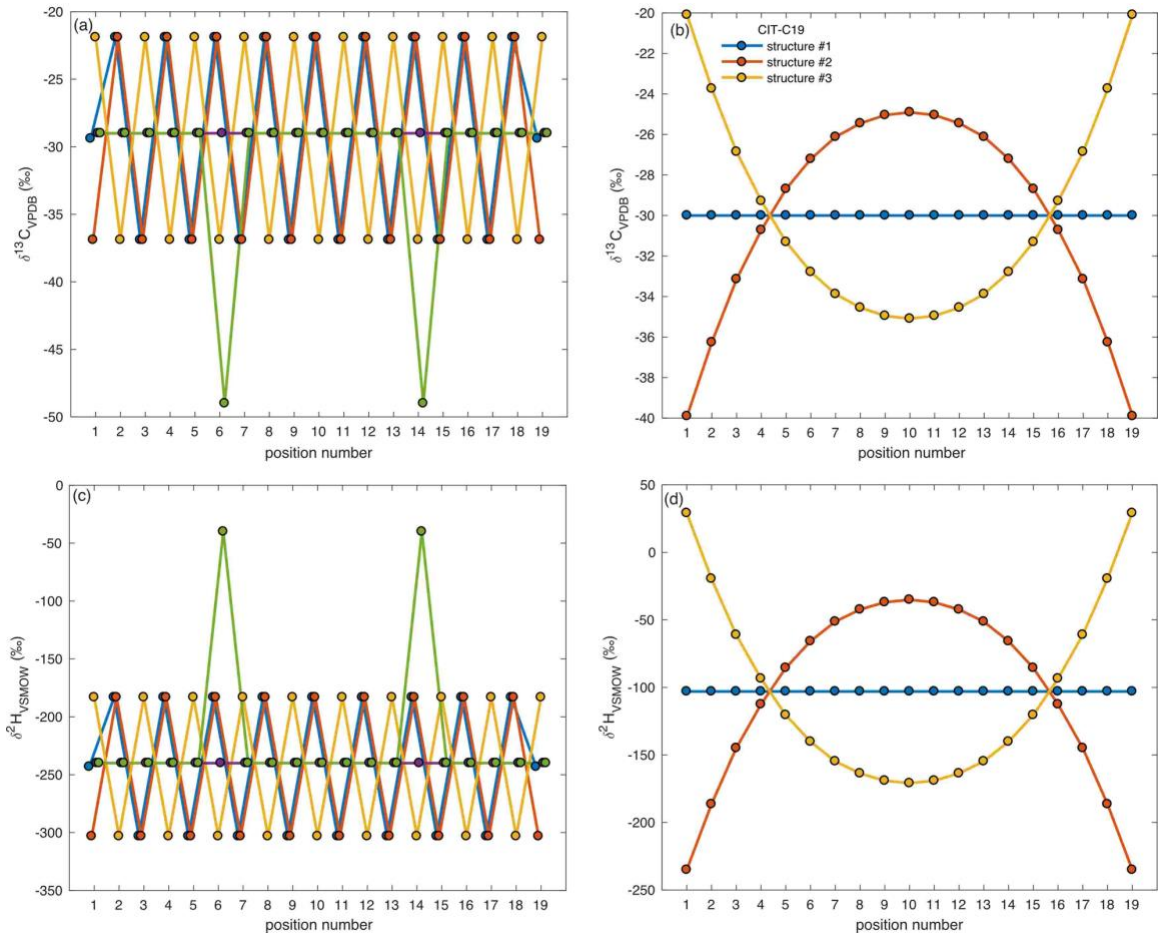


Figure 4.9 Hypothetical site-specific carbon (a, b) and hydrogen (c, d) isotope structures of rose n-nonadecane (a, c) and CIT-C19 working standard (b, d). The 'zigzag' patterns in the rose n-nonadecane structures (#1-3) are based on known site-specific $\delta^{13}C_{VPDB}$ values of plant and microbial fatty acids, which are thought to arise from the biosynthetic process of assembling fatty acids from 2-carbon units. Structures #4 and #5 serve as comparisons, with #5 having large variations at specific positions. The CIT-C19 structures #2 and #3 represent hypothetical scenarios related to thermal maturation and hydrogen exchange in petroleum systems.

We ran these hypothetical structures through the transfer function defined by the proposed fragmentation scheme for *n*-nonadecane (Table 4.13, Figure 4.10). The code to achieve this is available here (doi.org/10.22002/jezb3-3ne16). We found:

1. The hypothetical zigzag patterns of intramolecular isotope structure of rose *n*-nonadecane are obscured in the fragmentation process, resulting in fragment-specific isotope composition indistinguishable from each other and from a flat pattern with the current measurement precision.
2. The site-specific isotope structure of CIT-C19 determines the average offset between rose *n*-nonadecane and CIT-C19 across different fragments. Although similar isotope structure variations in rose *n*-nonadecane could affect that offset too.
3. The best fit to measurement results is rose structure #5 for both carbon and hydrogen isotopes, which have large variations at position 6 and 14 ($\delta^{13}C_{VPDB}$ = -49‰, versus -29 at other positions, and $\delta^{2}H_{VSMOW}$= -40‰ versus -240‰ for other positions).

The large variations at specific sites within *n*-nonadecane can result from site-specific kinetic isotope effect of fatty acid desaturases that convert saturated fatty acids into unsaturated fatty acids. There was a nonadecene compound of high abundance in the rose essential oil but we were unable to locate the position of the double bond without a standard. If the proposed large variations at specific sites were true, position 6 or 14 was likely part of a double bond in the nonadecene compound.

The best-fit resulted from highest sensitivity of C_4H_9 fragment isotope composition to that of position 6 and 14, due to the highest contribution of pathway $C_{19}H_{40} \rightarrow C_{14}H_{29} \rightarrow C_4H_9$ among all pathways that produce C_4H_9 fragment (Table 4.11). The intermediate product $C_{14}H_{29}$ contains both position 6 and 14 and passes the variations from these two positions onto C_4H_9 fragment with minimum dilution from other positions. The impact of these large variations on C_5H_{11} and C_6H_{13} fragment are not as sensitive because of the lack of contribution of pathway $C_{19}H_{40} \rightarrow C_{14}H_{29} \rightarrow C_5H_{11}/C_6H_{13}$ to C_5H_{11} and C_6H_{13} fragments. A similar best fit result is also possible with a zigzag pattern overlaid on top of the large variations at position 6 and 14.

The contribution of pathways $C_{19}H_{40} \rightarrow C_{14}H_{29} \rightarrow C_4H_9/C_5H_{11}/C_6H_{13}$, or lack thereof, in the proposed fragmentation scheme for *n*-nonadecane remains to be tested by site-specific labels.

But we test the sensitivity of contribution from pathways in the fragmentation scheme in the next section 3.2.5.3.

Table 4.13 Fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} of rose *n*-nonadecane in reference to CIT-C19 with hypothetical site-specific isotope composition. The bolded scenarios fit best with the long-term measurement values within 1–2 standard errors.

CIT-C19 structure #	rose structure #	$\Delta^{13}C_{CIT-C19}$ (‰) C4H9	C5H11	C6H13	CIT-C19 structure #	rose structure #	$\Delta^{2}H_{CIT-C19}$ (‰) C4H9	C5H11	C6H13
1	1	0.0	0.0	-0.1	1	1	0	0	-1
	2	0.1	0.1	0.1		2	1	1	1
	3	-0.1	-0.1	-0.1		3	-1	-1	-1
	4	0.0	0.0	0.0		4	0	0	0
	5	-0.4	0.0	0.1		5	5	0	-1
2	1	-0.3	-0.2	-0.7	2	1	-5	-3	-10
	2	-0.3	-0.2	-0.6		2	-4	-3	-9
	3	-0.4	-0.3	-0.7		3	-6	-4	-10
	4	-0.3	-0.2	-0.7		4	-5	-4	-10
	5	**-0.7**	**-0.2**	**-0.6**		5	0	-4	-11
3	1	0.3	0.3	0.6	3	1	5	4	9
	2	0.4	0.3	0.7		2	6	4	11
	3	0.3	0.2	0.6		3	4	3	9
	4	0.3	0.2	0.7		4	5	4	10
	5	0.0	0.3	0.8		5	**10**	**3**	**8**
long term measured average		-1.2	-0.3	-0.3	long term measured average		12	5	11
SD		1.1	0.9	1.2	SD		3	5	6
SE		0.4	0.3	0.4	SE		1	2	2

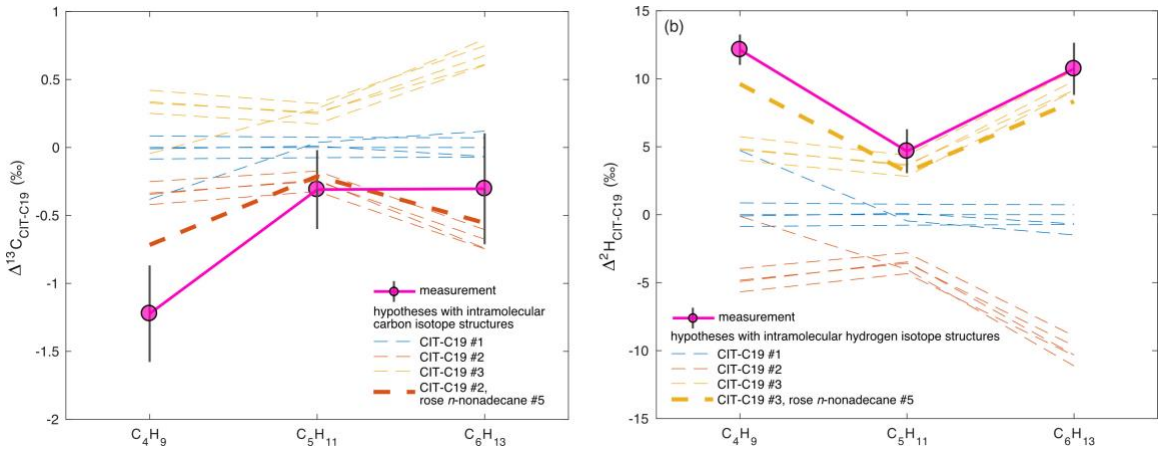


Figure 4.10 Fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} of rose *n*-nonadecane, calculated relative to the CIT-C19 standard, using hypothetical site-specific isotope compositions. The bolded lines represent scenarios that best fit the long-term measured values within 1–2 standard errors, suggesting the presence of large, site-specific isotopic variations in the parent molecule.

Error bars represent standard errors of the long-term measured fragment-specific carbon and hydrogen isotope composition of rose *n*-nonadecane.

3.2.5.3. Testing the sensitivity of fragmentation schemes of n-nonadecane by Monte-Carlo simulation

We conducted two Monte-Carlo simulations on: (1) the contribution of pathways $C_{19}H_{40} \rightarrow C_{17}H_{35}/C_{16}H_{33}/C_{15}H_{31} \rightarrow C_4H_9/C_5H_{11}/C_6H_{13}$ and (2) the contribution of all pathways from *n*-nonadecane to $C_4H_9/C_5H_{11}/C_6H_{13}$ fragments, using the best-fit site-specific isotope composition of rose *n*-nonadecane and CIT-C19 from the previous section 3.2.5-3.2.5.2. We found:

1. The fragment-specific isotope composition are not sensitive to the contribution of pathways $C_{19}H_{40} \rightarrow C_{17}H_{35}/C_{16}H_{33}/C_{15}H_{31} \rightarrow C_4H_9/C_5H_{11}/C_6H_{13}$ (Figure 4.11). The variations in fragment-specific isotope composition by Monte-Carlo simulation are smaller than 1–2 measurement errors.
1. The variations in fragment-specific isotope composition resulted from various fragmentation schemes span 1–2‰ for $\Delta^{13}C$ and 15–30‰ for $\Delta^{2}H$ (Figure 4.12). The sensitivity to the fragmentation scheme decreases from C_4H_9 to C_5H_{11} to C_6H_{13}, possibly due to better retention of site-specific signals in larger fragments.

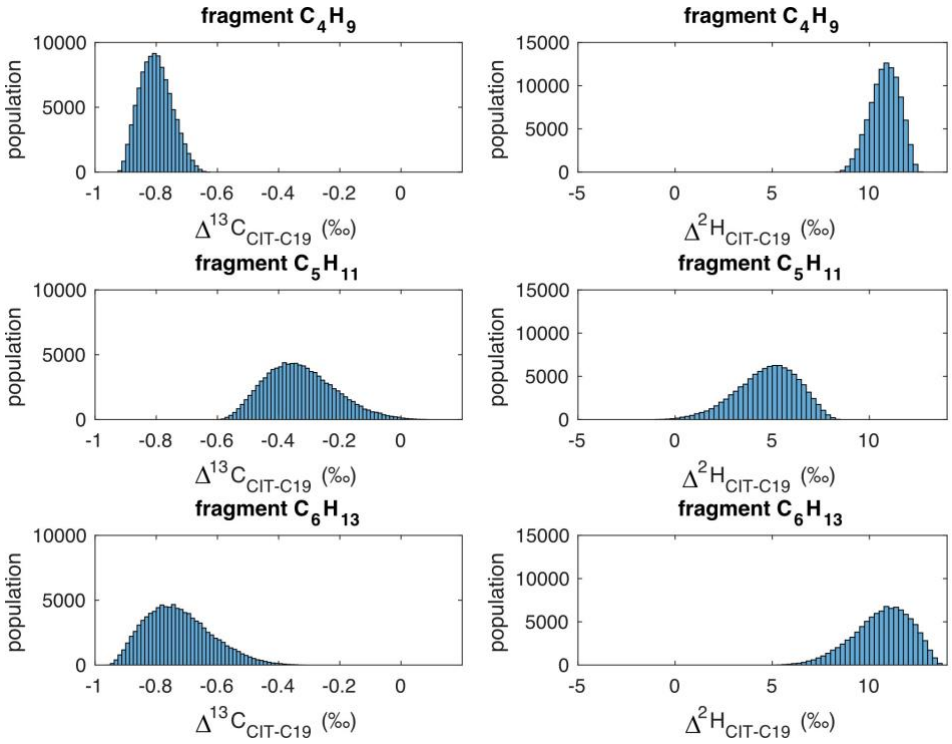


Figure 4.11 Monte-Carlo simulation results of fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of rose *n*-nonadecane in reference to CIT-C19. The contribution of pathways $C_{19}H_{40} \rightarrow C_{17}H_{35}/C_{16}H_{33}/C_{15}H_{31} \rightarrow C_4H_9/C_5H_{11}/C_6H_{13}$ are simulated with best-fit site-specific isotope composition of rose *n*-nonadecane and CIT-C19 from section 3.2.5-3.2.5.2. Total simulation number is 10,000.

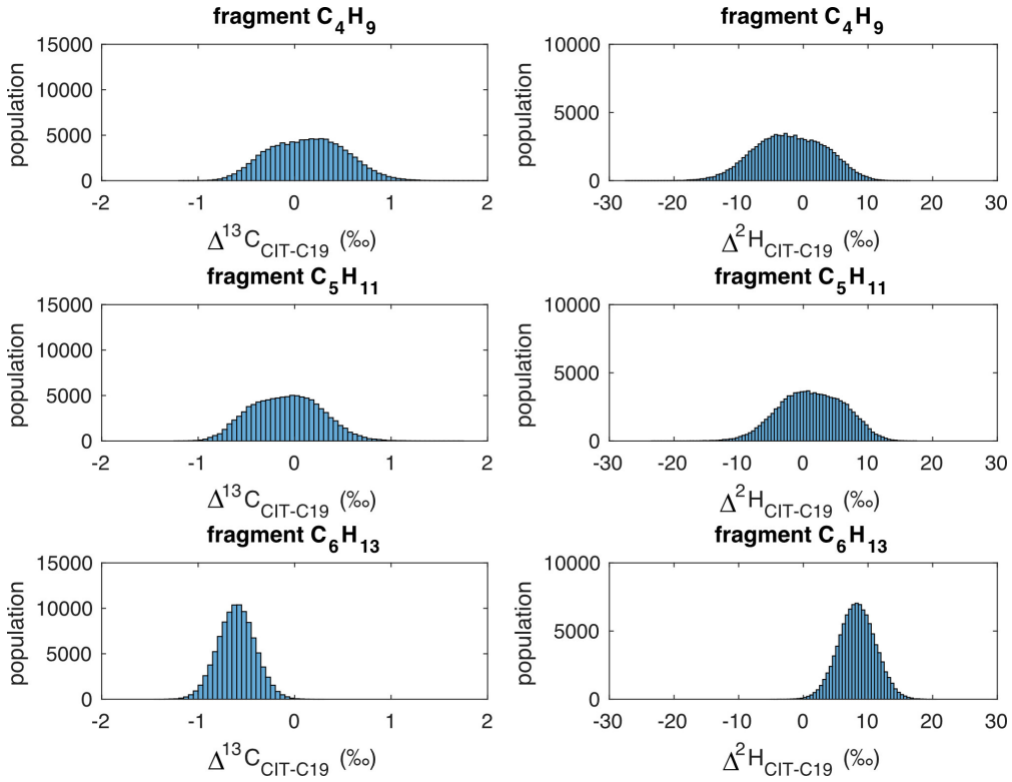


Figure 4.12 Monte-Carlo simulation results of fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} fragments from rose *n*-nonadecane, calculated relative to CIT-C19. The simulation explores the sensitivity of the calculated isotope compositions to variations in the full fragmentation scheme, using the best-fit site-specific isotope compositions of rose n-nonadecane and CIT-C19 derived in section 3.2.5-3.2.5.2. The decreasing range of variation from C_4H_9, C_5H_{11} and C_6H_{13} suggests that larger fragments may retain more information about the original site-specific isotope composition of the parent molecule, while smaller fragments are more affected by averaging or scrambling of isotopic signals during the fragmentation process.

4. Conclusion and future implications

In this study, we developed analytical methods using GC-Orbitrap to simultaneously probe the intramolecular carbon and hydrogen isotope structures of n-alkanes with nanomole sensitivity. Through measurements of site-specific labeled standards and modeling of the EI fragmentation process, we constructed transfer functions to relate the isotopic composition of fragment ions to the site-specific isotopic structures of parent *n*-alkanes.

The performance tests demonstrate that this approach can achieve ‰ level precision for $\Delta^{13}C$ and <5‰ precision for $\Delta^{2}H$ of C_4H_9, C_5H_{11} and C_6H_{13} fragment ions, with long-term reproducibility comparable to the internal precision of a single measurement. Proof-of-concept analysis of *n*-heptanes from different commercial sources and n-nonadecane from a natural sample compared to a working standard illustrates the capability of this method for isotope forensics and natural sample characterization.

Through measurements of isotopically labeled standards and modeling of the EI fragmentation process, we constructed quantitative transfer functions that relate the isotopic composition of fragment ions to the site-specific isotopic structures of parent n-alkanes. These transfer functions allowed us to test hypotheses regarding the site-specific isotope compositions of n-alkanes from natural samples. Our results suggest that large (tens of ‰ for carbon and hundreds of ‰ for hydrogen) isotope variations at certain molecular positions can be retained in the fragment-specific isotope compositions, while smaller, position-to-position variations are largely obscured during the fragmentation process.

This study lays the groundwork for future applications of fragment-specific isotope analysis of n-alkanes to study their origins and transformations in nature and in the laboratory. Establishing an absolute reference frame by synthesizing additional site-specific isotopic standards (Heusler et al., 1975) or equilibrating hydrogen isotopes in the presence of Pt/Pd catalyst will enhance the capabilities of this approach. Exploring softer ionization techniques, such as chemical ionization or low energy and low temperature electron impact ionization (Maccoll, 1982; Bowen and Maccoll, 1983) may provide more favorable fragmentation schemes. Leveraging NMR analytical techniques to characterize site-specific isotope composition of fatty acids (Julien et al., 2022), which can then be converted to n-alkanes through decarboxylation and hydrogenation, offers a promising route to produce calibrated isotope standards for this method.

With refinement, this method could find diverse applications, including aiding oil-oil or oil-source rock correlation in petroleum exploration, fingerprinting environmental

contaminants, studying thermal maturation of organic matter in sedimentary basins, reconstructing paleoclimate using sedimentary *n*-alkanes, elucidating biochemical reaction networks with or without isotope labeling, and tracing abiogenic hydrocarbon formation in the solar system and beyond, such as through clumped ^{13}C-^{13}C measurements (Zeichner et al., 2023). The combination of high-resolution mass spectrometry and intramolecular isotope analysis opens up exciting possibilities for studying n-alkanes and other organic molecules with unprecedented isotopic and structural detail.

Acknowledgements

This work was supported by the Caltech Joint Industry Partnership for Petroleum Geochemistry (Petrobras, Shell, ExxonMobil), the Simons Foundation, the Department of Energy BES program and ThermoFischer Scientific. We thank Nami Kitchen, Fenfang Wu and Michael Mathuri for laboratory operation support, Shae Silverman for assistance with GC-IRMS analysis, and Hannah Dion-Kirschner, Gabriella Weiss, and Alex Phillips for SPE procedures. We are grateful for helpful discussions with Kate Freeman, Elliott Mueller, Yanhua Shuai, Issaku Kohl, Sarah Zeichner, Amy Hoffman, Elle Chimiak, Elise Wilkes, Johan Weijers, and David Wang. We thank Sarah Zeichner, Tim Csernica, Elise Wilkes, Peter Martin, and Max Lloyd for their contributions to the communal Orbitrap data analysis code. During the preparation of this work, the author used Claude 3 Opus (large language model developed by Anthropic) in order to improve readability and language. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Bibliography

Bjorøy M., Hall K., Gillyon P. and Jumeau J. (1991) Carbon isotope variations in n-alkanes and isoprenoids of whole oils. *Chemical Geology* **93**, 13–20.

Bowen R. D. and Maccoll A. (1983) Low energy, low temperature mass spectra: I—selected derivatives of *n* -octane. *Org. Mass Spectrom.* **18**, 576–581.

Brittingham A., Hren M. T. and Hartman G. (2017) Microbial alteration of the hydrogen and carbon isotopic composition of n-alkanes in sediments. *Organic Geochemistry* **107**, 1–8.

Castañeda I. S. and Schouten S. (2011) A review of molecular organic proxies for examining modern and ancient lacustrine environments. *Quaternary Science Reviews* **30**, 2851–2891.

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

Dion-Kirschner H., McFarlin J. M., Masterson A. L., Axford Y. and Osburn M. R. (2020) Modern constraints on the sources and climate signals recorded by sedimentary plant waxes in west Greenland. *Geochimica et Cosmochimica Acta* **286**, 336–354.

Eiler J., Cesar J., Chimiak L., Dallas B., Grice K., Griep-Raming J., Juchelka D., Kitchen N., Lloyd M., Makarov A., Robins R. and Schwieters J. (2017) Analysis of molecular isotopic structures at high precision and accuracy by Orbitrap mass spectrometry. *International Journal of Mass Spectrometry* **422**, 126–142.

Eiler J. M. (2013) The Isotopic Anatomies of Molecules and Minerals. *Annu. Rev. Earth Planet. Sci.* **41**, 411–441.

Gilbert A. (2021) The Organic Isotopologue Frontier. *Annu. Rev. Earth Planet. Sci.* **49**, 435–464.

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

Gilbert A., Yamada K. and Yoshida N. (2013) Exploration of intramolecular 13C isotope distribution in long chain n-alkanes (C11–C31) using isotopic 13C NMR. *Organic Geochemistry* **62**, 56–61.

Gross J. H. (2004) *Mass spectrometry: a textbook*., Springer, Berlin New York.

Hayes J. M. (2001) Fractionation of Carbon and Hydrogen Isotopes in Biosynthetic Processes. *Reviews in Mineralogy and Geochemistry* **43**, 225–277.

Heusler A., Ganz P. and Gäumann T. (1975) Synthesis of N-alkane derivatives labelled with several ^{13}C. *J. Labelled Cpd. Radiopharm.* **11**, 37–42.

Holmes J. L., Burgers P. C., Mollah M. Y. A. and Wolkoff P. (1982) Isomerization of alkane molecular ions. *J. Am. Chem. Soc.* **104**, 2879–2884.

Howe I. (1975) Hydrogen shifts in the decomposition of branched chain alkane radical cations. *Org. Mass Spectrom.* **10**, 767–769.

Julien M., Nun P., Höhener P., Parinet J., Robins R. J. and Remaud G. S. (2016) Enhanced forensic discrimination of pollutants by position-specific isotope analysis using isotope ratio monitoring by 13C nuclear magnetic resonance spectrometry. *Talanta* **147**, 383–389.

Julien M., Zhao Y., Ma R., Zhou Y., Nakagawa M., Yamada K., Yoshida N., Remaud G. S. and Gilbert A. (2022) Re-evaluation of the 13C isotope fractionation associated with lipids biosynthesis by position-specific isotope analysis of plant fatty acids. *Organic Geochemistry* **174**, 104516.

Lavanchy A. (1977) FRAGMENTATION DES ALCANES EN SPECTROMETRIE DE MASSE. Thesis, ÉCOLE POLYTECHNIQUE FÉDERALE DE LAUSANNE.

Lavanchy A., Houriet R. and Gäumann T. (1978) The mass spectrometric fragmentation of *n* -heptane. *Org. Mass Spectrom.* **13**, 410–416.

Lavanchy A., Houriet R. and Gäumann T. (1979) The mass spectrometric fragmentation ofn-alkanes. *Org. Mass Spectrom.* **14**, 79–85.

Li M., Huang Y., Obermajer M., Jiang C., Snowdon L. R. and Fowler M. G. (2001) Hydrogen isotopic compositions of individual alkanes as a new approach to petroleum correlation: case studies from the Western Canada Sedimentary Basin. *Organic Geochemistry* **32**, 1387–1399.

Maccoll A. (1982) Plenary lecture: Ion enthalpies and their application in mass spectrometry. *Org. Mass Spectrom.* **17**, 1–9.

McLafferty F. W. and Tureček F. (1993) *Interpretation of mass spectra*. 4th ed., University Science Books, Mill Valley, Calif.

Mead R., Xu Y., Chong J. and Jaffé R. (2005) Sediment and soil organic matter source assessment as revealed by the molecular distribution and carbon isotopic composition of n-alkanes. *Organic Geochemistry* **36**, 363–370.

Meyerson S. (1965) Organic Ions in the Gas Phase. XV. Decomposition of *n* -Alkanes under Electron Impact. *The Journal of Chemical Physics* **42**, 2181–2186.

Monson K. D. and 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. *J Biol Chem* **255**, 11435–11441.

Monson K. D. and Hayes J. M. (1982) Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Saccharomyces cerevisiae. Isotopic fractionation in lipid synthesis as evidence for peroxisomal regulation. *J Biol Chem* **257**, 5568–5575.

Pedentchouk N., Freeman K. H. and Harris N. B. (2006) Different response of δD values of n-alkanes, isoprenoids, and kerogen during thermal maturation. *Geochimica et Cosmochimica Acta* **70**, 2063–2072.

Pedentchouk N. and Turich C. (2018) Carbon and hydrogen isotopic compositions of *n* -alkanes as a tool in petroleum exploration. *SP* **468**, 105–125.

Peters K. E., Walters C. C. and Moldowan J. M. (2007) Biomarkers and isotopes in the environment and human history. In *The biomarker guide* Cambridge University Press.

Philp R. P. (2007) The emergence of stable isotopes in environmental and forensic geochemistry studies: a review. *Environ Chem Lett* **5**, 57–66.

Sachse D., Billault I., Bowen G. J., Chikaraishi Y., Dawson T. E., Feakins S. J., Freeman K. H., Magill C. R., McInerney F. A., Van Der Meer M. T. J., Polissar P., Robins R. J., Sachs J. P., Schmidt H.-L., Sessions A. L., White J. W. C., West J. B. and Kahmen A. (2012) Molecular Paleohydrology: Interpreting the Hydrogen-Isotopic Composition of Lipid Biomarkers from Photosynthesizing Organisms. *Annu. Rev. Earth Planet. Sci.* **40**, 221–249.

Sessions A. L. (2006) Seasonal changes in D/H fractionation accompanying lipid biosynthesis in Spartina alterniflora. *Geochimica et Cosmochimica Acta* **70**, 2153–2162.

Sharp Z. (2017) Principles of Stable Isotope Geochemistry, 2nd Edition.

Tang Y., Huang Y., Ellis G. S., Wang Y., Kralert P. G., Gillaizeau B., Ma Q. and Hwang R. (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.

Tanner B. R., Uhle M. E., Kelley J. T. and Mora C. I. (2007) C3/C4 variations in salt-marsh sediments: An application of compound specific isotopic analysis of lipid biomarkers to late Holocene paleoenvironmental research. *Organic Geochemistry* **38**, 474–484.

Valley J. W., Cole D. R. and Mineralogical Society of America eds. (2001) *Stable isotope geochemistry*., Mineralogical Society of America, Washington, D.C.

Verma R. S., Padalia R. C., Chauhan A., Singh A. and Yadav A. K. (2011) Volatile constituents of essential oil and rose water of damask rose (*Rosa damascena* Mill.) cultivars from North Indian hills. *Natural Product Research* **25**, 1577–1584.

Watson J. T. and Sparkman O. D. (2007) *Introduction to mass spectrometry: instrumentation, applications and strategies for data interpretation*. 4th ed., John Wiley & Sons, Chichester, England ; Hoboken, NJ.

Wendelboe J. F., Bowen R. D. and Williams D. H. (1981a) Unimolecular reactions of ionized alkanes. *J. Am. Chem. Soc.* **103**, 2333–2339.

Wendelboe J. F., Bowen R. D. and Williams D. H. (1981b) Unimolecular reactions of isolated organic ions: some isomers of C6H14 +? *J. Chem. Soc., Perkin Trans. 2*, 958.

Wendelboe J. S. F. and Bowen R. D. (1982) Unimolecular reactions of isolated organic ions: Some isomers of $[C_7H_{16}]^{+\cdot}$. *Org. Mass Spectrom.* **17**, 439–443.

Wolkoff P., Hammerum S. and Holmes J. L. (1983) The behavior of C6H14+· ions. *International Journal of Mass Spectrometry and Ion Physics* **47**, 343–346.

Wolkoff P. and Holmes J. L. (1978) Fragmentations of alkane molecular ions. *J. Am. Chem. Soc.* **100**, 7346–7352.

Zech M., Pedentchouk N., Buggle B., Leiber K., Kalbitz K., Marković S. B. and Glaser B. (2011) Effect of leaf litter degradation and seasonality on D/H isotope ratios of n-alkane biomarkers. *Geochimica et Cosmochimica Acta* **75**, 4917–4928.

Zeichner S. S., Wilkes E. B., Hofmann A. E., Chimiak L., Sessions A. L., Makarov A. and Eiler J. M. (2022) Methods and limitations of stable isotope measurements via direct elution of chromatographic peaks using gas chromotography-Orbitrap mass spectrometry. *International Journal of Mass Spectrometry*, 116848.

Zgheib R., Najm W., Azzi-Achkouty S., Sadaka C., Ouaini N. and Beyrouthy M. E. (2020) Essential Oil Chemical Composition of *Rosa corymbifera* Borkh., *Rosa phoenicia* Boiss. and *Rosa damascena* Mill. from Lebanon. *Journal of Essential Oil Bearing Plants* **23**, 1161–1172.

Chapter 5 — Thermal maturation effect on intramolecular carbon and hydrogen isotope structure of *n*-alkanes

Guannan Dong[a], Alexandra Ferreira[b], Michael Lawson[c], Julia van Winden[d], Michael Formolo[e], Alex Sessions[a], John Eiler[a]

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

[b] *Petrobras-CENPES, Ilha do Fundão, Rio de Janeiro RJ 21941-909, Brazil*

[c] *AkerBP, Lysaker 1366, Norway*

[d] *Shell Global Solutions International, 2288 GS Rijswijk, The Netherlands*

[e] *ExxonMobil Upstream Integrated Solutions, Spring, TX 77389, USA*

Abstract

Intramolecular isotope distributions in organic molecules provide insights into their origins and formation mechanisms. However, applying this tool to sedimentary organic matter requires understanding how thermal maturation affects these isotope structures. We conducted anhydrous pyrolysis experiments on *n*-nonadecane at 400°C for 8–22 hours, equivalent to vitrinite reflectance maturities of 1.3–2%Ro to study the thermal cracking process. Using GC-P/C-IRMS and Orbitrap mass spectrometry, we tracked molecular average and fragment-specific (C_4H_9, C_5H_{11}, and C_6H_{13}) carbon and hydrogen isotope compositions of the residual n-nonadecane.

With increasing thermal maturity, the residual n-nonadecane fraction decreased exponentially, following first-order kinetics. Molecular average $\delta^{13}C$ and $\delta^{2}H$ values increased by 3.5‰ and 48‰, respectively, consistent with preferential cleavage of ^{12}C-^{12}C and ^{12}C-^{1}H bonds. Fragment-specific isotope compositions exhibited shifts in lockstep with molecular average changes. After correcting for molecular average isotope fractionations, fragment-specific $\Delta^{13}C$ and $\Delta^{2}H$ values show no statistically significant variations with pyrolysis extent.

Kinetic analysis suggests radical chain propagation reactions dominate the decomposition process. Quantum chemical calculations predict kinetic isotope effects consistent with observed isotopic compositions, supporting isotope effects associated with beta scission as the primary control. These results indicate that intramolecular isotope signatures in sedimentary n-alkanes may preserve those inherited from their precursors, with minimal alteration during thermal maturation under the experimental conditions studied.

1. Introduction

Intramolecular isotopic structure has emerged as a powerful tool in organic geochemistry, providing a detailed understanding of the distribution of isotopes within individual molecules and shedding light on the origins, histories, and processes that influence organic compounds. This approach has been applied to unravel metabolic pathways and fluxes (Abelson and Hoering, 1961; Monson and Hayes, 1980, 1982; Robins et al., 2003; Gilbert et al., 2012; Romek et al., 2015; Ehlers et al., 2015; Mueller et al., 2022; Wilkes et al., 2022; Julien et al., 2022; Weiss et al., 2023), identify substrates and synthetic mechanisms for extraterrestrial organics (Chimiak et al., 2021; Zeichner et al., 2023b, a; Chimiak and Eiler, 2024), investigate physical processes such as evaporation (Julien et al., 2015) and sorption (Fox et al., 2021), forensically distinguish sources of various chemical compounds (McKelvie et al., 2010; Hattori et al., 2011; Gilbert et al., 2013; Diomande et al., 2015; Julien et al., 2016), and differentiate between abiotic and biological amino acids (Chimiak et al., 2022; Weiss et al., 2023). Intramolecular isotope signatures have also been proposed to distinguish different sources of *n*-alkanes or aliphatic lipids, such as biological and abiotic origins (McCollom and Seewald, 2006). The recent development of the analytical techniques has greatly expanded the scope and practicality of intramolecular isotopic analysis in *n*-alkanes and fatty acids with more accessible carbon positions, improved precision, accuracy, and sensitivity (Julien et al., 2022; Zhao et al., 2023); Dong et al., Chapter 4).

However, the interpretation of intramolecular isotopic data is complicated by the potential for post-generation alteration during burial and preservation. Diagenetic and catagenetic

processes can alter the isotopic composition of organic compounds (Zech et al., 2011; Nguyen Tu et al., 2011). For example, thermal maturation can lead to the preferential cracking of isotopically lighter bonds, resulting in enrichment of heavy isotopes in the residual organic matter (Clayton, 1991; Schimmelmann et al., 1999; Tang et al., 2005). To fully harness the potential of intramolecular isotopic structure as a tool for unraveling the origins and histories of organic compounds, it is crucial to understand how thermal destruction alters these isotopic signals. This knowledge will enable us to determine whether intramolecular isotopic structure can still be utilized despite catagenesis or if pre-catagenesis isotope compositions can be reconstructed when the processes are well constrained.

In addition to its implications for understanding the origins of organic compounds, any systematic changes in intramolecular isotope structure with increasing maturity may serve as a novel maturity parameter. Whereas existing parameters based on biomarkers are limited by the complete destruction of these compounds at high maturities, n-alkanes persist over a wider range of thermal conditions (Peters et al., 2004; Killops and Killops, 2013). As such, an intramolecular isotope-based maturity parameter derived from *n*-alkanes could extend our ability to assess the thermal histories of sedimentary basins.

To investigate the effects of thermal maturation, specifically thermal cracking, on the intramolecular isotopic structure of *n*-alkanes, we conducted anhydrous, closed-system pyrolysis experiments on *n*-nonadecane (*n*C_{19}). We report the variations in carbon and hydrogen isotope structure of the residual *n*-nonadecane as a function of thermal maturity. We propose that the carbon and hydrogen isotope effects associated with hydrogen abstraction reactions are the primary control on the isotope composition of residual *n*-nonadecane in our experiments and, by extension, preserved organic matter in natural settings. To quantitatively describe the evolution of intramolecular carbon and hydrogen isotope structure during pyrolysis, we present a kinetic model that captures the key processes responsible for isotopic fractionation. Finally, we extrapolate our model to sedimentary basin conditions and make predictions regarding the intramolecular isotope effects of thermal

destruction on preserved organics across a range of thermal maturities. Our findings contribute to a more robust understanding of the effects of thermal maturation on the intramolecular isotopic structure of n-alkanes and lay the groundwork for the development of novel isotope-based maturity parameters.

2. Methods and materials

2.1. Pyrolysis experiments and product recovery

Aliquots of 40–70 mg of *n*-nonadecane (TCI, % purity, LOT# OGO01; same material as the working standard CIT-C19 (Dong et al., Chapter 4)) were weighed and placed in precombusted quartz tubes (8–9 inch long, ¼ inch outer diameter, 0.05 inch thick, a total volume of about 1 cc). The tubes were evacuated to approximately 10-4 Torr using a mercury diffusion pump backed by a rotary pump, while cooled to approximately 77 K by immersion in a liquid nitrogen bath, and then flame-sealed. Then, the sealed quartz tubes were placed in a Mellen NACCI tube furnace at 400 °C and heated for 8–22 h (Table 5.1).

After pyrolysis, tubes were removed from the furnace and quenched vertically within 1 min to room temperature in air. The fluid products concentrated to the bottom tip of the tube within tens of minutes. The tubes were then cracked open and ~10 mL of DCM was added to both pieces of the cracked tube to extract the non-volatile products. The cracked tubes containing pyrolysis product and DCM were then sonicated 15–20 min to facilitate the extraction process. The DCM extract was filtered with a precombusted glass fiber filter to remove any insoluble product. Aliquots of the extract were evaporated and redissolved in *n*-hexane for mass spectrometry analysis.

The equivalent thermal maturity (Easy%RoV) of each pyrolysis experiment was calculated for the temperature and duration of the experiment using a kinetic model for vitrinite reflectance (Burnham, 2019) (Table 5.1). The equivalent thermal maturity Easy%Ro

calculated using a previous version of the vitrinite reflectance kinetic model (Sweeney and Burnham, 1990) was also provided.

Table 5.1 pyrolysis experiment information (temperature 400°C). The equivalent thermal maturity (Easy%RoV) of each pyrolysis experiment was calculated for the temperature and duration of the experiment using a kinetic model for vitrinite reflectance (Burnham, 2019).

Experimental time (hr)	initial *n* -nonadecane (mg)	final *n* -nonadecane (mg)	residual fraction	Easy%RoV*
8	57.8	18.3	0.32	1.73
12	40.4	8.6	0.21	1.87
22	68.5	7.1	0.10	2.09

* Easy%Ro (Sweeney and Burnham, 1990) for experiments from 8 to 22 h are 1.25, 1.34, and 1.46, respectively.

2.2. GC-MS analysis

We conducted GC/MS analyses on the extracted products from the pyrolysis experiments to quantify the abundance of the residual n-nonadecane (Table 5.1). Analyses were performed a Thermo Scientific TSQ 9000 mass spectrometer with a Trace 1300 GC. The samples were injected by an AI/AS 1310 autosampler to a ZB-5ms column (30 m × 0.25 mm × 0.25 μm film thickness) through a PTV injector operated in splitless mode, using helium as a carrier gas with a flow rate of 1.2 mL/min. The injector was held at 40 °C for 0.3 min, then heated to 310°C at 12°C/s and held for 5 min before cooling down to 300 °C at 2 °C/s and held for 50 min . The GC oven was started with a 1 min hold at 70°C, heated from 70°C to 310°C at 5°C/min, and held at 310°C for 10 min. *N*-nonadecane was identified by comparison of mass spectra to library data and retention time to that of the working standard CIT-C19. *N*-nonadecane abundances were calculated by reference to an internal standard (50 μg of *n*-tricosane) using total ion count (TIC) peak areas.

2.3. Compound-specific carbon and hydrogen isotope measurement

The $\delta^{13}C_{VPDB}$ and $\delta^{2}H_{VMOW}$ of the starting and the residual *n*-nonadecane in the pyrolysis experiments were measured at Caltech by GC-combustion-IRMS (for $\delta^{13}C_{VPDB}$) and GC-pyrolysis-IRMS (for $\delta^{2}H_{VMOW}$) on a Thermo-Scientific Delta+ XP.

About 1 uL of 0.1 mg/mL (0.37 mM) CIT-C19 *n*-hexane solution or pyrolysis product *n*-hexane solution with similar *n*-nonadecane concentration was injected to a ZB-5MS (30 m × 0.25 mm × 0.25 μm film thickness) column through a PTV injector operated in Large Volume mode, using helium as a carrier gas with a flow rate of 1.4 mL/min.

Co-injected peaks of CO_2 ($\delta^{13}C_{VPDB}$ = -12.04‰) and H_2 ($\delta^{2}H_{VSMOW}$ = -224‰) reference gases were used for single-point isotopic calibration of analytes. The reference gases were in turn calibrated against a standard solution of eight fatty acid methyl esters ('F8 mix' of Arndt Schimmelmann, Indiana University). The F8 mix was also run along with the samples to monitor the instrumental status.

Molecular average carbon and hydrogen isotope composition is expressed using delta notation relative to standards including Vienna Pee Dee Belemnite (VPDB), Vienna Standard Mean Ocean Water (VSMOW), and CIT-C19 *n*-nonadecane (CIT-C19):

$$\delta^{13}C_{VPDB} = \frac{^{13}R_{sample}}{^{13}R_{VPDB}} - 1 \quad (5\text{-}1)$$

$$\delta^{2}H_{VSMOW} = \frac{^{2}R_{sample}}{^{2}R_{VSMOW}} - 1 \quad (5\text{-}2)$$

where $^{13}R = \frac{^{13}C}{^{12}C}$ and $^{2}R = \frac{^{2}H}{^{1}H}$; and $\delta^{13}C$ and $\delta^{2}H$ values are reported in per mil notation. The long-term experimental reproducibility for measurements of standards and replicate samples on these instruments is 0.15‰ for $\delta^{13}C$ and 5‰ for $\delta^{2}H$ (1SE).

2.4. Fragment-specific carbon and hydrogen isotope analysis

We analyzed the ^{13}C- and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions of starting and residual *n*-nonadecane via peak capture measurement on a Thermo Scientific Q Exactive Orbitrap mass spectrometer with a Trace 1310 GC and a Thermo Scientific Triplus RSH autosampler (Dong et al., Chapter 4).

Briefly, about 1 uL of 0.1 mg/mL (0.37 mM) CIT-C19 solution or pyrolysis product solution with similar nonadecane concentration was injected to a TG-5SilMS column (30 m × 0.25 mm × 0.25 μm film thickness) through a split/splitless injector held at 250℃ and operated in the splitless mode, using helium as a carrier gas with a flow rate of 1.2 mL/min. The GC oven was started with at 50℃, heated to 250℃ at 15℃/min, and held at 250℃ for 36 min. During the elution of the entire chromatographic peak of *n*-nonadecane, we directed and mixed the elute into a 10 cm^3 passivated stainless-steel peak broadener prefilled with helium. Then a flow of pure helium carrier gas (1.2 ml/min) purged the peak broadener and delivered the trapped *n*-nonadecane to the electron impact (EI) Orbitrap mass spectrometer over about 30 minutes. The transfer line was set at 250℃. The EI source was set at 250℃. The filament was turned on at 5 min. We measured one out of the three fragments in each analysis and used a quadrupole scan range of 52–63 mz for C_4H_9 fragment, 66.5–76.5 mz for C_5H_{11}, and 81–90 m for C_6H_{13}. We used the prescan AGC mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 1 to 200 ms for scans during which sample was present). We used the positive ion mode and no 'lock masses' were used. The nominal mass resolution was set at 120,000 (FWHM, at m/z = 200). The number of microscans was set to 1. We ran the acquisition for the duration of the GC oven temperature program, which we established was the time required for the mass spectra of effluent from the peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., nonadecane-free) measurements. Each analysis of the pyrolysis product sample was bracketed by two analyses of CIT-C19, and each bracketed analysis was repeated 5 times for each fragment.

A solution of perfluorotributylamine (PFTBA, Sigma-Aldrich, LOT# 77299) was used to calibrate the mass scale and mass isolation for positive ions and to auto-tune the ion source, injection flatapole and inter-flatapole lens with electron energy set at 70 eV and C-trap energy offset at 0 V, maximizing the signal intensity on the Q Exactive GC.

The fragment-specific isotope composition of *n*-nonadecane relative to CIT-C19 are calculated using a delta notation:

$$\delta^{13}C_{CIT-C19} = \frac{^{13}R_{frag,sample}}{^{13}R_{frag,CIT-C19}} - 1 \tag{5-3}$$

$$\delta^{2}H_{CIT-C19} = \frac{^{2}R_{frag,sample}}{^{2}R_{frag,CIT-C19}} - 1 \tag{5-4}$$

where $\delta^{13}C$ and $\delta^{2}H$ values are reported in per mil notation. To isolate the intramolecular isotope structure variation and eliminate the influence of differences in molecular average isotope composition, normalized fragment-specific isotope composition was also derived, which was achieved by offsetting the molecular average isotope fractionation:

$$\Delta^{13}C_{CIT-C19} = \frac{^{13}R_{frag,sample}}{^{13}R_{frag,CIT-C19}} \Big/ \frac{^{13}R_{mol,sample}}{^{13}R_{mol,CIT-C19}} - 1 \tag{5-5}$$

$$\Delta^{2}H_{CIT-C19} = \frac{^{2}R_{frag,sample}}{^{2}R_{frag,CIT-C19}} \Big/ \frac{^{2}R_{mol,sample}}{^{2}R_{mol,CIT-C19}} - 1 \tag{5-6}$$

where $\Delta^{13}C$ and $\Delta^{2}H$ values are also reported in per mil notation. The long-term experimental reproducibility for measurements of standards and replicate samples is about 1‰ for $\Delta^{13}C$ and 5‰ for $\Delta^{2}H$ (1SE).

2.5. Computational analysis

We fitted the fractionation factors for molecular average and fragment-specific isotope composition of *n*-nonadecane by nonlinear regression using MATLAB function nlinfit, found 95% confidence intervals for the fitted fractionation factors using MATLAB function nlparci, and made predictions with 95% confidence intervals using the fitted Rayleigh distillation model and MATLAB function nlpredci (Table 5.4, Figure 5.2). The code is available at doi.org/10.22002/catyq-fxd56). The isotope fractionation factors are expressed using an epsilon notation:

$$^{13C}\varepsilon = \frac{^{13C}R_{residual}}{^{13C}R_{product}} - 1 \quad (5\text{-}7)$$

$$^{2H}\varepsilon = \frac{^{2H}R_{residual}}{^{2H}R_{product}} - 1 \quad (5\text{-}8)$$

where $^{13C}\varepsilon$ and $^{2H}\varepsilon$ values are reported in per mil notation.

We constructed a numerical model of the kinetics and isotope systematics of the pyrolysis experiment using MATLAB and integrated ordinary differential equations using ordinary differential equation solver ODE45s.

2.6. Prediction of kinetic isotope effect (KIE) of hydrogen abstraction

We predicted carbon and hydrogen KIE for hydrogen abstraction by methyl radical from methyl and methylene groups in ethane and propane, respectively, using transition state theory and quantum chemical calculations. Transition state structures for the reactions

$$CH_3\bullet + C_2H_6 \rightarrow CH_4 + C_2H_5\bullet \quad (5\text{-}9)$$

and

$$CH_3\bullet + CH_3CH_2CH_3 \rightarrow CH_4 + CH_3CH\bullet CH_3 \quad (5\text{-}10)$$

were optimized using density functional theory at the B3LYP/aug-cc-pVTZ level in Entos Qcore. Harmonic vibrational frequencies were computed for un-, ^{13}C-, and ^{2}H-substituted reactants (Table S 5-1) and transition states to obtain reduced partition function ratios (Bigeleisen and Mayer, 1947):

$$\frac{\beta'}{\beta} = \prod \frac{\nu_i'}{\nu_i} \frac{e^{-U_i'/2}}{e^{-U_i/2}} \frac{1 - e^{-U_i}}{1 - e^{-U_i'}} \quad (5\text{-}11)$$

where ν_i and ν_i' are the harmonic frequencies of vibrational mode i in the unsubstituted and isotope-substituted versions of the molecule, and $U_i = h\nu_i$ /kT, where h is Planck constant, k is Boltzmann constant, and T is temperature. The KIEs were then determined using the semiclassical theory (Bigeleisen, 1949):

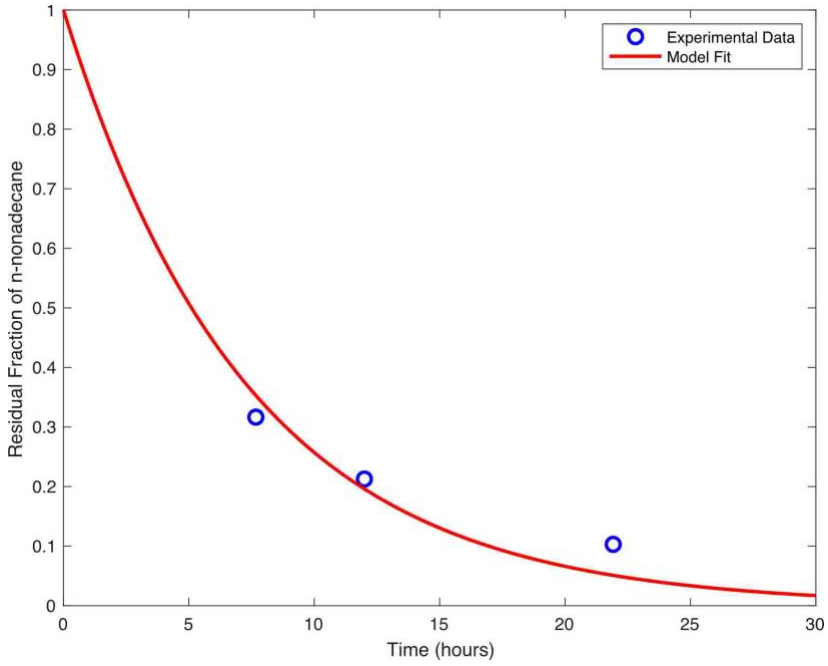
$$KIE = \frac{k'}{k} = \left(\frac{\nu_{im}'}{\nu_{im}}\right)_{TS} \cdot \frac{(\beta'/\beta)_{TS}}{(\beta'/\beta)_{reactant}} \quad (5\text{-}12)$$

where k and k' are the reaction rate constants of hydrogen abstraction for unsubstituted and isotope-substituted species, ν_{im} and ν_{im}' are the imaginary frequencies of the stretching mode along the coordinate of the carbon-hydrogen bond breaking for the unsubstituted and isotope-substituted versions of the transition state, $(\beta'/\beta)_{TS}$ is the reduced partition function ratio of the transition state, considering only its harmonically oscillating real modes, and $(\beta'/\beta)_{reactant}$ is the reduced partition function ratio of the reactant ethane or propane, considering all of its fundamental modes.

3. Results

3.1. Chemical composition of the product of pyrolysis experiment

GC-MS quantitation showed that the residual fraction of *n*-nonadecane decreasing exponentially with time during pyrolysis at 400°C, appearing to follow first-order kinetics (Figure 5.1). The residual fraction drops from 1.0 at the start to 0.10 after 22 hours (Table 5.1).



nC_{10}-C_{16} alkanes and 1-alkenes were produced at trace levels. The concentration of these lower molecular weight n-alkanes increased from 1.5% at 8 hours to 6% at 22 hours, normalized to the signal intensity of n-nonadecane, while 1-alkenes increased to 1% by 22 hours. *n*C_{17} and C_{18} alkanes were also detected, increasing from below 1% at 8 hours to 1–2% by 22 hours. In contrast to these lower molecular weight products, *n*$C_{20–22}$ alkanes were also detected at about 1% level at 22 hours. These higher molecular weight n-alkanes are longer than the starting n-nonadecane and are likely produced by recombination reactions, rather than by fragmentation of the starting material.

3.2. Evolution of molecular average carbon and hydrogen isotope composition of *n*-nonadecane

The molecular average $\delta^{13}C$ and $\delta^{2}H$ values of the starting *n*-nonadecane (CIT-C19) are -30.4‰ and -103‰, respectively (Table 5.2). Although we do not have clear information about the origin of the substrate from the supplier, this composition is consistent with that of petroleum materials (Schoell, 1984; Chung et al., 1992). The measured molecular average $\delta^{13}C$ and $\delta^{2}H$ of residual *n*-nonadecane in reference to CIT-C19 increased by 3.5‰ and 48‰, respectively, as the pyrolysis experiment progressed for 22 hours and the equivalent thermal maturity increased (Table 5.1, Table 5.2, Figure 5.1, Figure 5.2). The molecular average carbon and hydrogen isotope fractionation between the residual and the starting *n*-nonadecane is consistent with experimental result and theoretical prediction of thermal decomposition of *n*-icosane and *n*-heneicosane with similar equivalent thermal maturity (Easy%Ro) (Table 5.1) (Tang et al., 2005).

Table 5.2 Time evolution of molecular average $\delta^{13}C$ and $\delta^{2}H$ of *n*-nonadecane and fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} from *n*-nonadecane in the pyrolysis experiments. Fragment-specific carbon and hydrogen isotope composition $\delta^{13}C$, $\delta^{2}H$, $\Delta^{13}C$, and $\Delta^{2}H$ are defined in equation (5-3)-(5-6). $\Delta^{13}C$ and $\Delta^{2}H$ calculated after offsetting the molecular average isotope fractionation.

Experimental time (hr)	residual fraction	molecular average $\delta^{13}C_{CIT\text{-}C19}$ (‰)	$\delta^{2}H_{CIT\text{-}C19}$ (‰)	$\delta^{13}C_{CIT\text{-}C19}$ (‰)			SE (‰)			$\delta^{2}H_{CIT\text{-}C19}$ (‰)			SE (‰)		
				C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13
0	1.00	0.0	0	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0	0	0	0
8	0.32	1.5	19	0.9	2.3	1.5	0.5	0.6	0.5	15	20	18	3	3	2
12	0.21	1.6	26	4.4	2.1	-1.0	0.4	0.5	0.4	24	18	24	5	3	4
22	0.10	3.5	48	1.8	5.8	1.9	0.3	0.5	0.3	37	48	39	5	2	3

	molecular average		Experimental time (hr)	$\Delta^{13}C_{CIT\text{-}C19}$ (‰)			SE (‰)			$\Delta^{2}H_{CIT\text{-}C19}$ (‰)			SE (‰)		
	$\delta^{13}C_{VPDB}$	$\delta^{2}H_{VSMOW}$ (‰)		C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13
CIT-C19	-30.4	-103	0	0.0	0.0	0.0	0.5	0.5	0.5	0	0	0	3	3	3
			8	-0.6	0.8	0.0	0.5	0.6	0.5	-3	1	-1	3	3	2
			12	2.8	0.5	-2.6	0.4	0.5	0.4	-1	-7	-2	5	3	4
			22	-1.7	2.3	-1.6	0.3	0.5	0.3	-11	0	-9	4	2	3

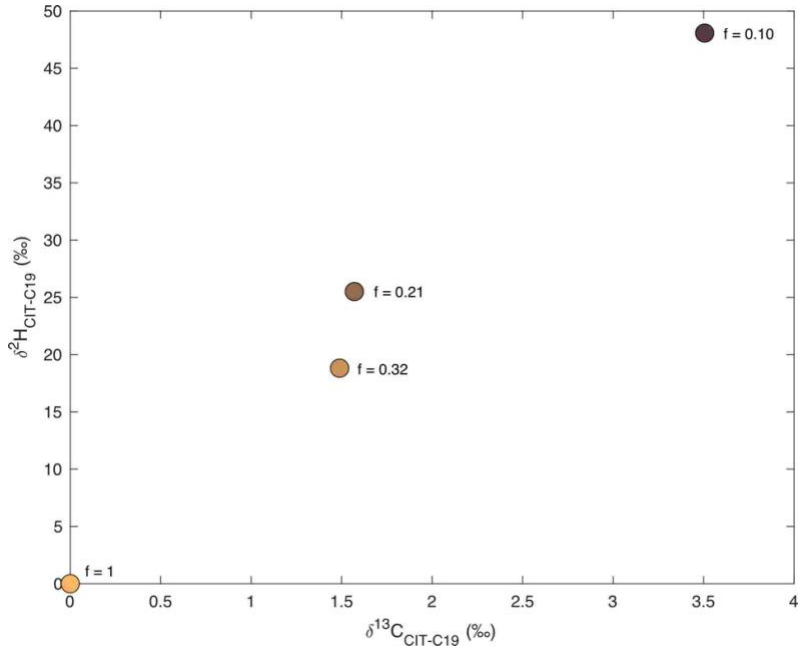
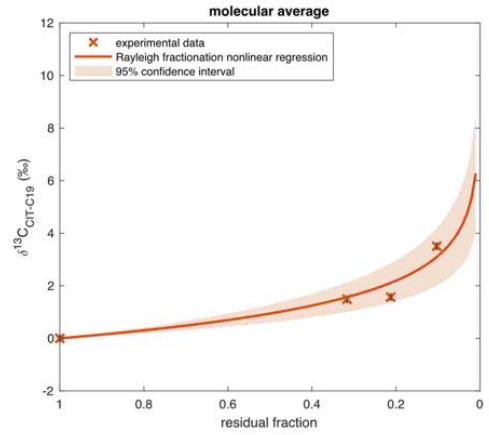
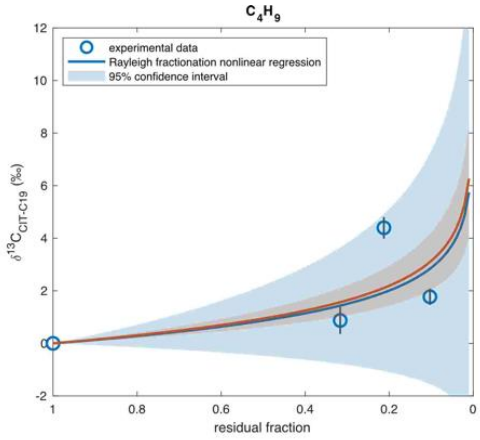


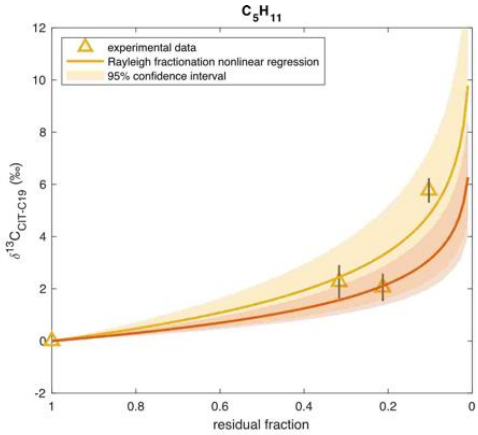
Figure 5.1 Evolution of molecular average $\delta^{13}C$ and $\delta^{2}H$ of *n*-nonadecane with the residual fraction (f) of *n*-nonadecane in the pyrolysis experiment.

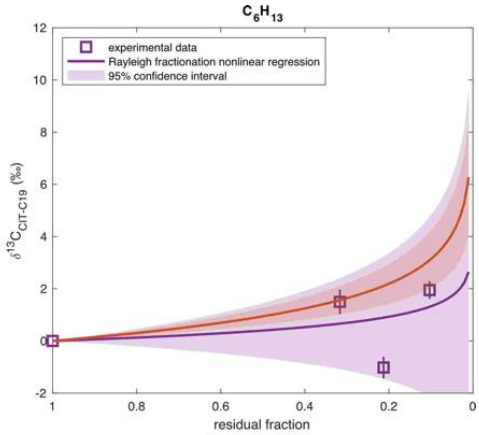
3.3. Evolution of carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions from *n*-nonadecane

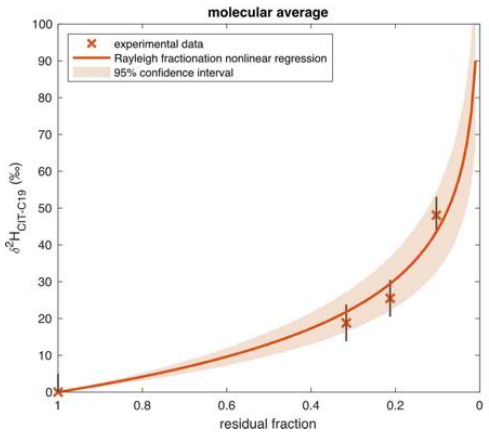
The fragment-specific carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} of the residual *n*-nonadecane also becomes more enriched during the pyrolysis experiment (Table 5.2, Figure 5.2). The evolution of fragment-specific carbon isotope composition appears to demonstrate non-monotonic trend over the course of pyrolysis. Such behavior may stem from different types of reactions at different stages in the cracking experiment, e.g., breaking down of aliphatic compounds into shorter chain alkanes versus ring formation and aromatization.

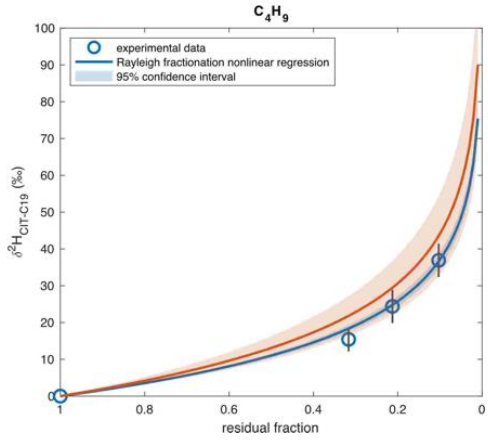

molecular average
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{13}C_{CIT\text{-}C19}$ (‰)
residual fraction


C_4H_9
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{13}C_{CIT\text{-}C19}$ (‰)
residual fraction


C_5H_{11}
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{13}C_{CIT\text{-}C19}$ (‰)
residual fraction


C_6H_{13}
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{13}C_{CIT\text{-}C19}$ (‰)
residual fraction


molecular average
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{2}H_{CIT\text{-}C19}$ (‰)
residual fraction


C_4H_9
experimental data
Rayleigh fractionation nonlinear regression
95% confidence interval
$\delta^{2}H_{CIT\text{-}C19}$ (‰)
residual fraction

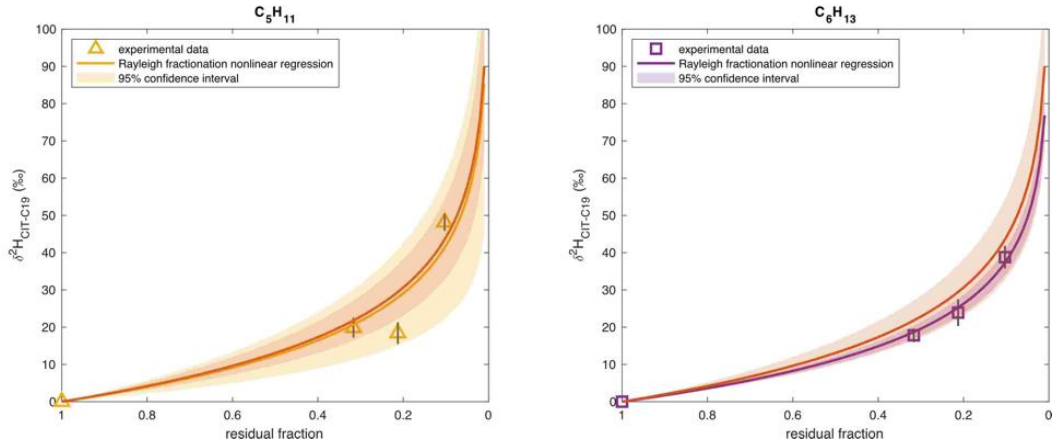


Figure 5.2 Evolution of molecular average and fragment-specific $\delta^{13}C$ and $\delta^{2}H$ values of *n*-nonadecane and its C_4H_9, C_5H_{11}, and C_6H_{13} fragments during pyrolysis, plotted against the residual fraction (f) of *n*-nonadecane. The isotopic evolution is fitted with a Rayleigh distillation model using nonlinear regression. Pyrolysis enriches both molecular average and fragment-specific isotopic composition of the residual *n*-nonadecane. Error bars represent ±1 standard measurement error.

After offsetting the molecular average isotope fractionation, the normalized fragment-specific isotope composition solely reflects the intramolecular isotope variation (Figure 5.3). The $\Delta^{13}C$ values of each fragment span a range of 3–5‰ throughout the pyrolysis experiments, close to 2–3 times of the long-term precision. The $\Delta^{2}H$ values of each fragment are within 2 standard measurement errors from each other and generally become more depleted as the thermal decomposition comes close to completion (residual fraction 0.21 and 0.10).

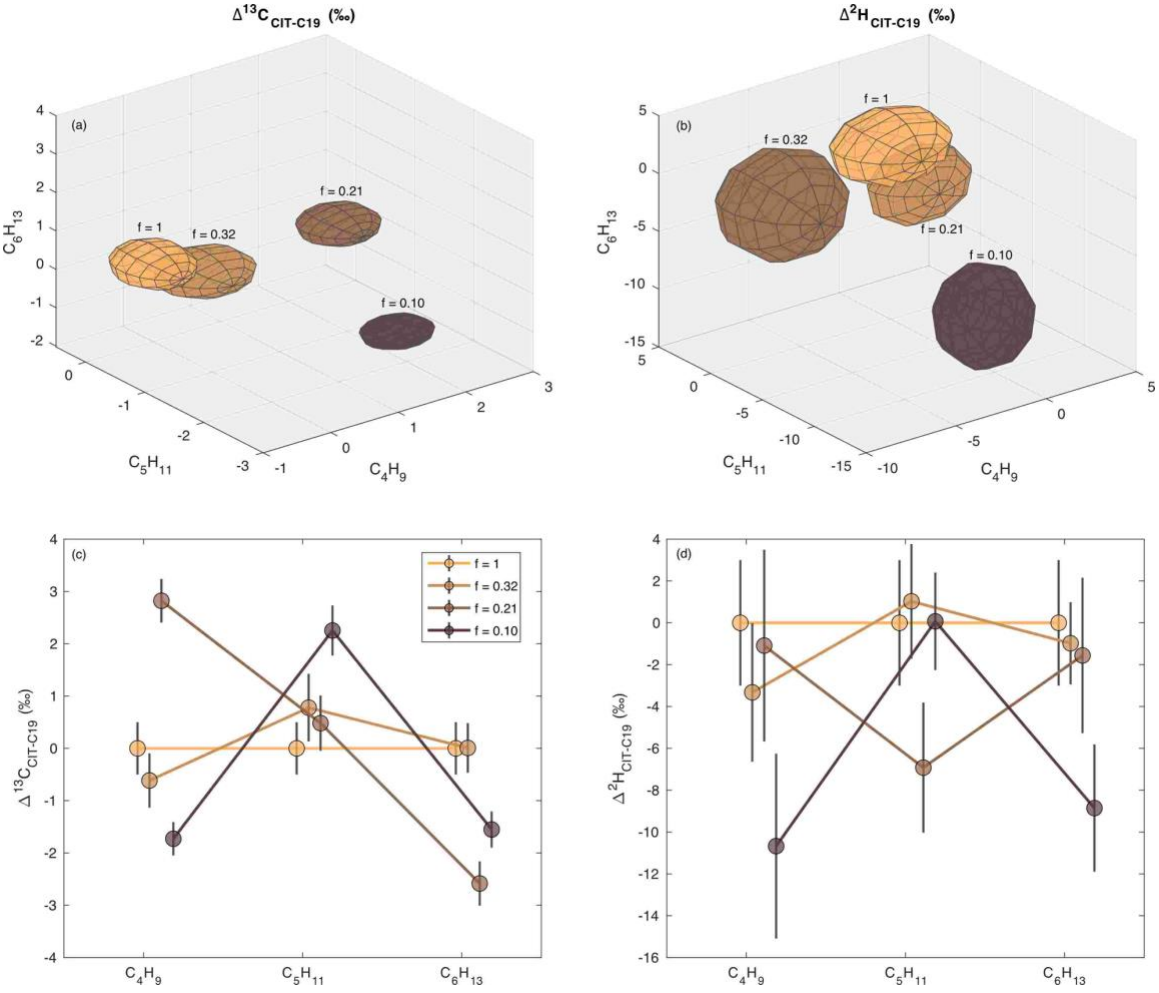


Figure 5.3 Evolution of the normalized fragment-specific $\Delta^{13}C$ ((a), (c)) and $\Delta^{2}H$ ((b), (d)) of C_4H_9, C_5H_{11}, and C_6H_{13} of *n*-nonadecane with the residual fraction (f) of *n*-nonadecane in the pyrolysis experiment. After offsetting the molecular average isotope fractionation, the normalized fragment-specific isotope composition solely reflects the intramolecular isotope variation. The $\Delta^{13}C$ values of each fragment (c) span a range of 3–5‰ across the pyrolysis experiments with no clear correlation to the residual fraction of *n*-nonadecane. And the total range of $\Delta^{13}C$ is close to 2–3 times long term precision. The $\Delta^{2}H$ values of each fragment are within 2 standard measurement errors from each other and generally become more depleted as the thermal decomposition comes close to completion (residual fraction 0.21 and 0.10). Elliptic radii of ellipsoids ((a), (b)) and error bars ((c), (d)) are 1 standard measurement error. No significant alteration on the intramolecular isotope distribution was observed in the thermal destruction process in the pyrolysis experiments by fragment-specific isotope analysis. Long term measurement reproducibility is about 1‰ and 5‰ for carbon and hydrogen isotope composition, respectively (Dong et al., Chapter 4).

3.4. Carbon and hydrogen KIE of hydrogen abstraction by methyl radical

The optimized transition state structures of reactions (5-9) and (5-10) are consistent with previous work by (Xiao, 2001) (Figure 5.4), where the hydrogen atom being transferred is positioned midway between the attacking methyl radical and the hydrogen donor (ethane or propane). This symmetric transition state suggests significant bonding between the methyl radical and the hydrocarbons in the transition state.

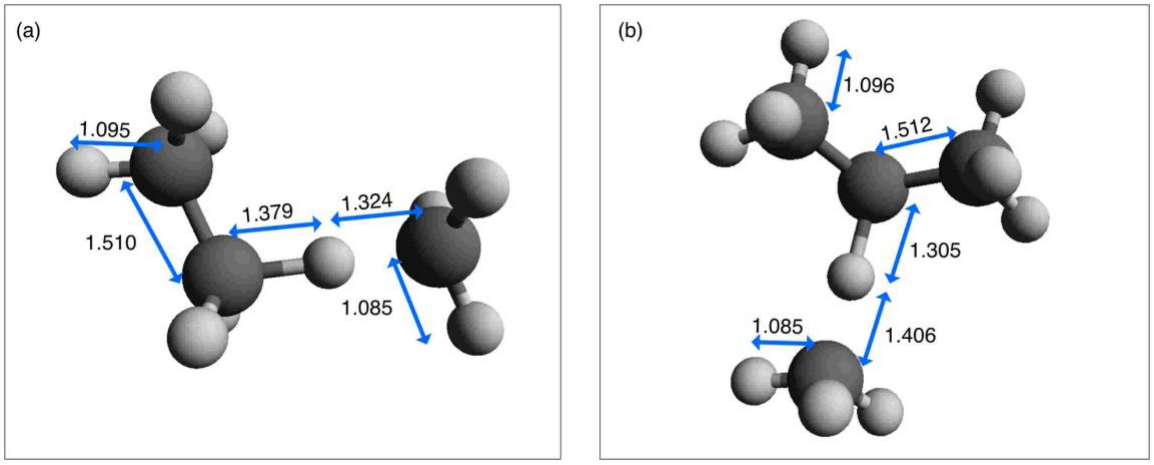


Figure 5.4 Optimized transition state structures for hydrogen abstraction by methyl radical from (a) ethane and (b) propane, based on ab initio calculations. The transferring hydrogen atom is positioned symmetrically between the methyl radical and the hydrogen donor molecule.

Table 5.3 Predicted carbon and hydrogen KIEs for hydrogen abstraction by methyl radical at 0-600°C

reaction	$CH_3\bullet + C_2H_6 \rightarrow CH_4 + C_2H_5\bullet$			$CH_3\bullet + CH_3CH_2CH_3 \rightarrow CH_4 + CH_3CH\bullet CH_3$		
temperature (°C)	primary 13C	primary 2H	secondary 2H	primary 13C	primary 2H	secondary 2H
0	0.9831	0.159	0.921	0.9846	0.154	0.923
100	0.9869	0.253	0.942	0.9881	0.251	0.944
200	0.9895	0.332	0.957	0.9903	0.332	0.958
300	0.9913	0.398	0.967	0.9920	0.399	0.968
400	0.9927	0.451	0.974	0.9932	0.453	0.974
500	0.9937	0.495	0.979	0.9941	0.498	0.979
600	0.9945	0.531	0.983	0.9948	0.534	0.983

The calculated kinetic carbon isotope fractionations for hydrogen abstraction at 400°C range from -6.8 to -7.3‰ (Table 5.3). These values are smaller than those expected for a transition state with a free hydrogen radical, likely due to the harmonic interaction in the transition state between the attacking methyl radical and the hydrogen donor molecules (Table S 5-1).

The predicted primary and secondary deuterium isotope fractionations for hydrogen abstraction at 400°C are -548‰ and -26‰, respectively (Table 5.3). These predictions are smaller than the KIEs extrapolated from experimental data on methyl radicals attacking deuterium-labeled ethane and n-butane, which yield values around -650‰ (Gray et al., 1971). This discrepancy could arise from additional effects, such as ^{2}H-^{2}H clumping in the multiply deuterated ethane and n-butane substrates used in the experiments. Quantum tunneling effects may also contribute to the observed differences between the predicted and experimental KIEs.

4. Discussion

4.1. Fit Rayleigh fractionation factors

Assuming the thermal destruction of *n*-nonadecane follows a Rayleigh distillation process, we fitted the fractionation factors for molecular average and fragment-specific isotope composition of *n*-nonadecane (Table 5.4), using the Rayleigh equation:

$$\delta_r + 1 = (\delta_i + 1) \cdot f^{\varepsilon} \quad (5\text{-}13)$$

where δ_r and δ_i are the isotopic composition of the residual (r) and initial (i) *n*-nonadecane, *f* is the fraction of the residual *n*-nonadecane, and ε is the isotope fractionation factor for thermal destruction of *n*-nonadecane.

Table 5.4 Fitted fractionation factors and their confidence intervals for molecular average and fragment-specific isotope composition of *n*-nonadecane assuming the thermal decomposition of n-nonadecane as a Rayleigh distillation process

$^{13C}\varepsilon$ (‰) (95% confidence interval)			
molecular average	C4H9	C5H11	C6H13
-1.4 (-1.8, -0.9)	-1.3 (-3.2, 0.7)	-2.1 (-3.2, -1.0)	-0.5 (-2.1, 0.9)
$^{2H}\varepsilon$ (‰) (95% confidence interval)			
molecular average	C4H9	C5H11	C6H13
-19 (-23, -14)	-15 (-17, -15)	-18 (-26, -10)	-16 (-18, -15)

The Rayleigh distillation fractionation factors for both carbon and hydrogen isotopes are indistinguishable among molecular average and fragment specific isotope compositions by 95% confidence. The confidence intervals for the evolution of isotope composition during thermal decomposition of *n*-nonadecane overlap among molecular average and each fragment partially due to the large confidence envelope (Figure 5.2). We conclude that the thermal cracking studied in our pyrolysis experiments enriched the molecular average isotope composition of *n*-nonadecane following a Rayleigh distillation process but did not monotonically alter the intramolecular isotope distribution observed by fragment-specific isotope analysis.

4.2. Modeling the isotopic composition of *n*-nonadecane during pyrolysis

Understanding the mechanisms governing the isotopic composition of n-nonadecane during pyrolysis is crucial for accurately interpreting stable isotope data in geochemical studies. To quantitatively elucidate these mechanisms, we developed a kinetic model that incorporates the isotope fractionation effects associated with the thermal decomposition of n-nonadecane (the code is available at doi.org/10.22002/8pw9w-pdt50). This model provides insights into the relative contributions of different reaction pathways and their impact on the observed isotopic signatures.

The rate of n-nonadecane decomposition was modeled using first-order kinetics:

$$\frac{d[C19]}{dt} = -k_{C19}[C19] \tag{5-14}$$

where [C19] represents the concentration of n-nonadecane and k_{C19} denotes the first-order rate constant. The fitting of the experimental data yielded a k_{C19} value of 3.77×10^{-5}/s.

Notably, this rate constant is significantly greater than the rate constant for homolytic cleavage reactions, which is approximately 1.20×10^{-9}/s for n-nonadecane at 400°C (Johnson et al., 2022). This discrepancy suggests that alternative reaction pathways, such as chain propagation reactions, play a dominant role in the decomposition of n-nonadecane under the experimental conditions. Specifically, the rate constant of hydrogen abstraction from n-nonadecane by methyl radical is $3.82\times10^{-3}cm^3/(mol\cdot s)$, and that of beta scission is 1.39×10^{-2}/s (Johnson et al., 2022). These additional reactions contributing significantly to the destruction of n-nonadecane is consistent with previous studies on the mechanism and kinetics of hydrocarbon pyrolysis (Savage, 2000; Xiao, 2001).

To investigate the isotope fractionations, we derived KIEs of homolytic cleavage and beta scission based on previous theoretical studies (Table 5.5). Additionally, we incorporated KIEs of homolytic cleavage, beta scission, and hydrogen abstraction obtained from our study (Table 5.3) into a Rayleigh fractionation model. By considering these Rayleigh fractionations by different KIEs, we were able to assess their relative control on the isotope composition of residual n-nonadecane. As shown in Figure 5.5, the enrichment of molecular average $\delta^{13}C$ and $\delta^{2}H$ of n-nonadecane observed in the pyrolysis experiments is consistent with the modeled fractionation caused by homolytic cleavage and beta scission. However, the experimental enrichment in carbon isotope composition of *n*-nonadecane is larger than the modeled fractionation caused by hydrogen abstraction alone.

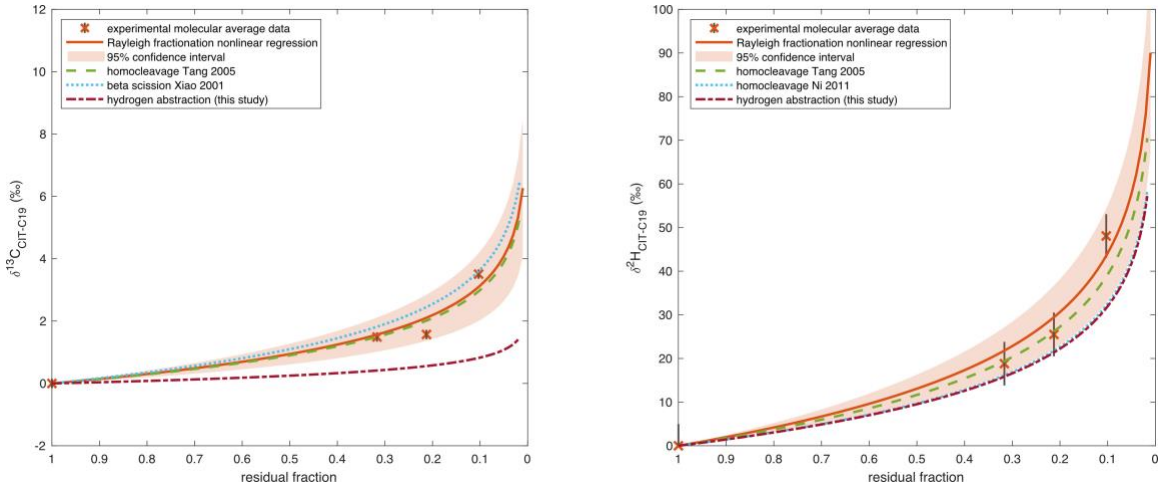


Figure 5.5 A comparison of the modeled fractionation of molecular average $\delta^{13}C$ and $\delta^{2}H$ of n-nonadecane caused by the KIEs of homolytic cleavage, beta scission, and hydrogen abstraction (dashed line) with the experimental results of pyrolysis (solid line and data points). The consistency between the experimental data and the modeled fractionation patterns for homolytic cleavage and beta scission supports their influence on the isotopic composition of residual *n*-nonadecane during thermal decomposition. In contrast, the discrepancy between the experimental data and the modeled fractionation in $\delta^{13}C$ of n-nonadecane caused by hydrogen abstraction suggests this reaction pathway has less impact on the observed isotopic signatures.

Table 5.5 Pyrolysis model parameters. The prefactor and activation energy difference of homolytic cleavage are adapted from (Xiao, 2001; Tang et al., 2005; Ni et al., 2011).

first order rate constant (s^{-1})	$3.77x10^{-5}$								
carbon KIE parameters					hydrogen KIE parameter				
homolytic cleavage, Tang 2005			beta scission, Xiao 2001		homolytic cleavage, Tang 2005			homolytic cleavage, Ni 2011	
prefactor (A*/A)	1.02			site specific	prefactor (A*/A)	1.07		prefactor (A*/A)	1.07
	activation energy difference at carbon site (J/mol)		T (°C)	KIE		activation energy difference at carbon site (J/mol)			activation energy difference at carbon site (J/mol)
C-C bond to break	Ci	Ci+1	0	0.9351	C-C bond to break	Ci	Ci+1	C-C bond to break	Ci	Ci+1
C1–C2	230.1	171.5	100	0.9516	C1–C2	1607	1339	C1–C2	1333	1172
C2–C3	200.8	175.7	200	0.9614	C2–C3	1372	1343	C2–C3	1179	1172
C3–C4	184.1	175.7	300	0.9667	C3–C4	1364	1343	C3–C4	1172	1172
C4–C5	184.1	175.7	400	0.9700	C4–C5	1347	1343	C4–C5	1172	1172
C5–C6	184.1	175.7	500	0.9719	C5–C6	1347	1343	C5–C6	1172	1172
C6–C7	184.1	175.7	600	0.9736	C6–C7	1347	1343	C6–C7	1172	1172
C7–C8	184.1	175.7			C7–C8	1347	1343	C7–C8	1172	1172
C8–C9	184.1	175.7			C8–C9	1347	1343	C8–C9	1172	1172
C9–C10	184.1	175.7			C9–C10	1347	1343	C9–C10	1172	1172

Combining the kinetic and isotopic fractionation analyses, we propose a mechanistic understanding of n-nonadecane pyrolysis. Although homolytic cleavage is an important initiation step, its slow kinetics limit its quantitative contribution to the overall thermal decomposition of n-nonadecane. Hydrogen abstraction, on the other hand, plays a crucial role in activating n-nonadecane molecules as part of the radical chain propagation process. However, due to its fast reaction rate, the KIE associated with hydrogen abstraction does not significantly express itself in the observed isotopic fractionation. Instead, we hypothesize that beta scission is the rate-limiting step in the decomposition process, allowing its KIE to be expressed in the isotopic signature of the residual n-nonadecane. To test this hypothesis, future work should focus on predicting the hydrogen KIE of beta scission and comparing it with the experimental results presented here.

In summary, our kinetic modeling approach, which incorporates isotope fractionation effects, provides valuable insights into the mechanisms controlling the isotopic composition of n-nonadecane during pyrolysis. The model highlights the dominant role of chain propagation reactions, particularly beta scission, in the thermal decomposition process and their influence on the observed isotopic signatures. These findings contribute to a better understanding of the factors shaping the stable isotope composition of hydrocarbons in geochemical systems and have implications for the interpretation of isotopic data in petroleum exploration and environmental studies.

5. Conclusion and future implications

Our pyrolysis experiments and fragment-specific isotope analysis show that the intramolecular carbon and hydrogen isotope distribution in n-nonadecane, as expressed by fragment-specific $\delta^{13}C$ and $\delta^{2}H$ values, is not significantly altered by thermal decomposition. The variations in fragment-specific isotope composition may stem from different types of reactions at different stages in the cracking experiment, e.g., breaking down of aliphatic compounds into shorter chains versus ring formation and aromatization.

To further understand the mechanisms, the next step could be to develop a model for the evolution of position-specific isotope composition controlled by homolytic cleavage, beta scission, and hydrogen abstraction reactions, and then calculate the resulting fragment-specific isotope compositions using a transfer function. We expect the model results to be largely consistent with the molecular average compositions.

It is crucial to acknowledge that our laboratory experiments simulate accelerated maturation conditions and focused on the thermal cracking process characterized by high temperatures and rapid reaction rates. In order to extrapolate to natural, slow maturation processes, we could derive the evolution of intramolecular isotope composition as a function of thermal maturity (e.g., Easy%Ro) at geologic heating rates (~3°C/Myr). We predict that thermal decomposition likely does not significantly affect intramolecular isotope structure under natural conditions, particularly as reflected in fragment-specific isotope compositions, because of the homogeneous chemical structure of n-alkanes.

To apply these findings to organic geochemistry, future work should address some key gaps between our pyrolysis experiments and natural conditions:

1. Study the process of n-alkane release from kerogen. If n-alkane precursors with different intramolecular isotopic structures are released at different thermal maturity stages due to varying bond strengths within the kerogen, one may observe variations in intramolecular isotope distributions with increasing thermal maturity.

2. Understand the effect of hydrogen exchange. High intramolecular heterogeneity of hydrogen isotope distribution has been observed for lipids (Billault et al., 2001), and H exchange drives it closer to equilibrium with increasing maturation (Schimmelmann et al., 2006). While competing with the rapid thermal decomposition of n-nonadecane, H exchange likely had a small effect over the short duration of our pyrolysis experiments (Dong et al., 2021). The extent to which our starting material (CIT-C19 n-nonadecane) deviates from equilibrium intramolecular

hydrogen isotope distribution is also unknown. Fortunately, the thermodynamics and kinetics of hydrogen exchange between organic molecules and water have been extensively studied (Larson and Hall, 1965; Metcalfe and Vickers, 1973; Koepp, 1978; Sessions et al., 2004; Wang et al., 2009; Xie et al., 2018; Turner et al., 2021). Building on this knowledge base, conducting isotopic labeling exchange experiments coupled with fragment-specific GC-Orbitrap analysis could test hypotheses about the homogeneity of H exchange for CH_2 vs. CH_3 groups in n-alkanes, laying the groundwork to develop intramolecular H isotope compositions as thermal maturity indicators.

By integrating an understanding of these processes with our pyrolysis findings, intramolecular isotope compositions could become powerful tools to decipher the sources and histories of sedimentary organic matter, shedding light on outstanding puzzles such as the cryptic behavior of compound-specific hydrogen isotope composition of n-alkanes in the Barney Creek Formation (Vinnichenko et al., 2021).

The findings of this study contribute to our understanding of how thermal maturation processes affect the intramolecular isotopic structure of n-alkanes, with potential implications for petroleum geochemistry and paleoclimate reconstruction. In petroleum geochemistry, a better understanding of the behavior of intramolecular isotopic signatures during catagenesis could ultimately lead to the development of new tools for reconstructing the thermal histories of source rocks and reservoirs. However, further research is needed to test the applicability of these findings to natural systems and to develop robust n-alkane-based maturity parameters. In the field of paleoclimatology, our results suggest that intramolecular isotopic signatures in sedimentary n-alkanes might be preserved during diagenesis and catagenesis, which is encouraging for their use as paleoclimate proxies. Nevertheless, additional studies are necessary to validate these findings across a wider range of geological conditions and processes. Overall, our study highlights the potential of

intramolecular isotopic analysis as a tool for investigating the fate of organic molecules in sedimentary systems, laying the groundwork for future research in this area.

Acknowledgements

This work was supported by the Caltech Joint Industry Partnership for Petroleum Geochemistry (Petrobras, Shell, ExxonMobil), the Simons Foundation, the Department of Energy BES program, and AkerBP. We thank Nami Kitchen and Fenfang Wu for laboratory operation support. We are grateful for helpful discussions with Hao Xie, Kate Freeman, Yanhua Shuai, Ilya Bobrovskiy, Jochen Brocks, Nithya Thiagarajan, and Johan Weijers. We thank Sarah Zeichner, Tim Csernica, Elise Wilkes, Peter Martin, and Max Lloyd for their contributions to the communal Orbitrap data analysis code. During the preparation of this work, the author used Claude 3 Opus (large language model developed by Anthropic) in order to improve readability and language. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Supplementary material

Table S 5-1 Harmonic frequencies of vibrational modes of reactants and transition states in the reactions $CH_3\bullet + C_2H_6 \rightarrow CH_4 + C_2H_5\bullet$ and $CH_3\bullet + CH_3CH_2CH_3 \rightarrow CH_4 + CH_3CH\bullet CH_3$

reactant vibrational frequency
v (cm-1)

CH3	CH3CH3	13CH3CH3	CH3CH2D	CH3-CH2-CH3	CH3-13CH2-CH3	CH3-CHD-CH3
541.0120	301.8916	301.8916	283.2889	216.5912	216.5909	216.5844
1406.5248	825.2717	824.4217	718.3661	265.9606	265.0107	262.1699
1406.6116	825.8127	824.9629	809.6069	366.1039	363.5829	364.0244
3107.0374	993.8138	978.7499	979.0068	749.5014	747.5269	665.6623
3281.4131	1221.5682	1213.4352	1139.0394	868.3195	861.6171	804.3752
3281.5016	1221.6569	1213.5293	1184.0728	910.5619	910.5619	864.0100
	1411.1678	1404.3115	1325.6547	931.1450	930.1988	926.6396
	1420.6755	1417.4981	1334.4999	1053.5118	1038.5478	1001.3064
	1501.4878	1500.9482	1416.0874	1176.5606	1166.2498	1148.5286
	1501.5673	1501.0263	1480.0024	1214.5640	1203.6121	1166.2068
	1504.5668	1503.2349	1502.9652	1317.5060	1317.5059	1185.3548
	1504.6270	1503.2958	1503.0300	1369.0256	1355.4438	1345.5140
	3023.6449	3020.6771	2243.7257	1402.1630	1401.5386	1349.0281
	3024.3319	3024.0342	3023.9834	1420.4046	1420.3532	1401.5316
	3066.0736	3059.5382	3042.1176	1489.3186	1488.1734	1420.3757
	3066.2718	3059.7355	3066.0522	1489.9744	1489.3202	1488.9255
	3091.0754	3086.7034	3080.2699	1495.4491	1494.7404	1494.4023
	3091.2659	3086.8942	3091.1027	1506.1807	1505.6625	1504.2306
				1511.8436	1509.0302	1505.5169
				3010.7379	3004.9754	2222.5231
				3013.8715	3013.8693	3013.8679
				3017.5378	3017.4446	3017.1738
				3028.8362	3019.6191	3020.7981
				3067.2252	3067.2235	3067.2110
				3077.6359	3076.4896	3072.8954
				3078.2308	3077.6342	3077.6389
				3080.0472	3079.8487	3079.8697

transition state vibrational frequency
v (cm-1)

CH3---C2H6	CH3---13C12CH6	CH3---C2H5D pri	CH3---C2H5D sec	CH3---C3H8	CH3---CH3-13CH2-CH3	CH3---CH3-CHD-CH3 pri	CH3---CH3-CHD-CH3 sec
-1646.6132	-1644.0764	-1230.9658	-1641.9720	-1591.3736	-1589.0143	-1193.0509	-1585.1398
29.0572	28.9950	29.0496	28.7304	19.3603	19.3569	15.5569	15.4644
138.1854	137.3884	136.1606	136.7529	116.4588	116.1175	115.3179	115.8437
161.2217	160.9594	159.9502	160.4830	150.7406	150.1340	148.7182	148.1254
365.9928	365.9831	339.4256	337.2908	194.1074	194.1056	194.0195	194.0351
472.0181	468.0228	459.7172	463.3053	238.4256	237.9234	235.7562	235.4095
544.4028	540.8939	531.4808	542.1973	353.4624	351.3087	347.5399	350.8807
674.8959	672.5055	674.6084	626.3041	463.4668	460.3912	454.4529	447.5087
855.3711	855.3569	845.3924	728.9119	531.6679	531.5215	511.3300	531.0408
906.4937	906.2904	882.5216	882.5523	544.4036	539.4758	537.1522	541.6622
1029.1460	1012.3072	1002.9075	1014.2609	857.9952	855.8805	834.3322	715.0458
1137.8787	1136.3450	1094.0693	1115.9063	890.8133	885.7718	875.1296	812.1653
1176.5309	1168.6150	1095.0513	1139.5017	928.7040	928.5246	923.5313	878.2371
1228.4624	1222.0070	1115.9112	1163.0702	951.9580	951.0266	939.8207	950.4294
1380.7430	1378.9917	1177.6566	1327.9747	1106.3596	1089.4320	1022.2196	1031.7746
1381.1345	1380.6135	1234.6965	1357.0146	1126.0927	1125.7812	1086.1369	1125.5238
1407.6392	1407.6169	1407.2807	1381.0991	1173.8364	1161.6054	1109.4096	1139.5110
1444.4231	1444.2983	1433.6547	1409.5382	1195.8543	1186.3369	1140.4489	1170.5325
1445.3130	1445.1648	1434.0943	1441.8114	1338.2960	1334.0475	1185.2343	1210.2087
1465.1702	1459.5397	1464.8424	1445.1022	1387.2122	1382.5901	1196.8003	1361.2380
1490.9102	1490.6685	1490.8475	1490.7288	1394.4407	1393.1295	1358.3337	1380.6913
1491.7699	1491.4600	1491.6637	1491.2978	1403.9674	1401.1758	1400.5281	1399.4739
3002.5117	3002.5030	3002.4489	2270.2152	1415.3688	1415.3510	1415.1583	1415.2711
3049.5425	3048.1039	3049.4754	3002.5607	1442.8649	1441.9215	1431.3467	1438.8609
3055.9065	3055.8869	3055.8164	3051.6191	1443.9253	1443.5716	1431.9030	1441.8938
3062.5427	3058.5013	3062.4981	3055.9188	1480.3894	1480.3642	1480.3663	1480.1841
3074.9272	3074.2736	3074.9227	3074.1797	1486.1161	1485.2880	1485.4664	1483.8363
3129.1143	3117.7665	3129.0745	3098.5900	1496.1152	1495.5205	1495.6906	1495.0433
3182.0468	3182.0462	3182.0287	3182.0411	1497.2552	1496.6118	1497.1543	1496.8144
3182.8649	3182.8648	3182.8486	3182.8646	2992.6162	2992.6161	2992.7848	2246.6936
				2996.3824	2996.3627	2996.4264	2992.7788
				3040.5177	3036.1649	3040.3813	2996.6092
				3045.3649	3045.3639	3045.3388	3045.3323
				3054.9770	3051.5424	3054.9909	3046.5759
				3057.3857	3057.3381	3057.2684	3057.2945
				3077.5543	3077.5541	3077.4948	3077.4818
				3082.3168	3081.2578	3082.2385	3078.0347
				3184.8197	3184.8196	3184.7783	3184.7865
				3185.6262	3185.6262	3185.5839	3185.5954

Bibliography

Abelson P. H. and Hoering T. C. (1961) CARBON ISOTOPE FRACTIONATION IN FORMATION OF AMINO ACIDS BY PHOTOSYNTHETIC ORGANISMS. *Proc. Natl. Acad. Sci. U.S.A.* **47**, 623–632.

Bigeleisen J. (1949) The Relative Reaction Velocities of Isotopic Molecules. *The Journal of Chemical Physics* **17**, 675–678.

Bigeleisen J. and Mayer M. G. (1947) Calculation of Equilibrium Constants for Isotopic Exchange Reactions. *The Journal of Chemical Physics* **15**, 261–267.

Billault I., Guiet S., Mabon F. and Robins R. (2001) Natural Deuterium Distribution in Long-Chain Fatty Acids Is Nonstatistical: A Site-Specific Study by Quantitative2H NMR Spectroscopy. *ChemBioChem* **2**, 425–431.

Burnham A. K. (2019) Kinetic models of vitrinite, kerogen, and bitumen reflectance. *Organic Geochemistry* **131**, 50–59.

Chimiak L. and Eiler J. (2024) Prebiotic synthesis on meteorite parent bodies: Insights from hydrogen and carbon isotope models. *Chemical Geology* **644**, 121828.

Chimiak L., Eiler J., Sessions A., Blumenfeld C., Klatte M. and Stoltz B. M. (2022) Isotope effects at the origin of life: Fingerprints of the Strecker synthesis. *Geochimica et Cosmochimica Acta* **321**, 78–98.

Chimiak L., Elsila J. E., Dallas B., Dworkin J. P., Aponte J. C., Sessions A. L. and Eiler J. M. (2021) Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system. *Geochimica et Cosmochimica Acta* **292**, 188–202.

Clayton C. (1991) Carbon isotope fractionation during natural gas generation from kerogen. *Marine and Petroleum Geology* **8**, 232–240.

Diomande D. G., Martineau E., Gilbert A., Nun P., Murata A., Yamada K., Watanabe N., Tea I., Robins R. J., Yoshida N. and Remaud G. S. (2015) Position-Specific Isotope Analysis of Xanthines: A ^{13} C Nuclear Magnetic Resonance Method to Determine the ^{13} C Intramolecular Composition at Natural Abundance. *Anal. Chem.* **87**, 6600–6606.

Dong G., Xie H., Formolo M., Lawson M., Sessions A. and Eiler J. (2021) Clumped isotope effects of thermogenic methane formation: Insights from pyrolysis of hydrocarbons. *Geochimica et Cosmochimica Acta* **303**, 159–183.

Ehlers I., Augusti A., Betson T. R., Nilsson M. B., Marshall J. D. and Schleucher J. (2015) Detecting long-term metabolic shifts using isotopomers: CO_2-driven suppression of photorespiration in C_3 plants over the 20th century. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15585–15590.

Fox A. C., Martineau E., Remaud G. S. and Freeman K. H. (2021) Position-specific isotope fractionation in amino acids sorbed to ice: Implications for the preservation of isotopologue biosignatures. *Geochimica et Cosmochimica Acta* **309**, 45–56.

Gilbert A., Robins R. J., Remaud G. S. and Tcherkez G. G. B. (2012) Intramolecular ^{13}C pattern in hexoses from autotrophic and heterotrophic C_3 plant tissues. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18204–18209.

Gilbert A., Yamada K. and Yoshida N. (2013) Exploration of intramolecular 13C isotope distribution in long chain n-alkanes (C11–C31) using isotopic 13C NMR. *Organic Geochemistry* **62**, 56–61.

Gray P., Herod A. A. and Jones A. (1971) Kinetic data for hydrogen and deuterium atom abstraction by methyl and trifluoromethyl radicals in the gaseous phase. *Chemical Reviews* **71**, 247–294.

Hattori R., Yamada K., Kikuchi M., Hirano S. and Yoshida N. (2011) Intramolecular Carbon Isotope Distribution of Acetic Acid in Vinegar. *J. Agric. Food Chem.* **59**, 9049–9053.

Johnson M. S., Dong X., Grinberg Dana A., Chung Y., Farina D., Gillis R. J., Liu M., Yee N. W., Blondal K., Mazeau E., Grambow C. A., Payne A. M., Spiekermann K. A., Pang H.-W., Goldsmith C. F., West R. H. and Green W. H. (2022) RMG Database for Chemical Property Prediction. *J. Chem. Inf. Model.* **62**, 4906–4915.

Julien M., Nun P., Höhener P., Parinet J., Robins R. J. and Remaud G. S. (2016) Enhanced forensic discrimination of pollutants by position-specific isotope analysis using isotope ratio monitoring by 13C nuclear magnetic resonance spectrometry. *Talanta* **147**, 383–389.

Julien M., Nun P., Robins R. J., Remaud G. S., Parinet J. and Höhener P. (2015) Insights into Mechanistic Models for Evaporation of Organic Liquids in the Environment Obtained by Position-Specific Carbon Isotope Analysis. *Environ. Sci. Technol.* **49**, 12782–12788.

Julien M., Zhao Y., Ma R., Zhou Y., Nakagawa M., Yamada K., Yoshida N., Remaud G. S. and Gilbert A. (2022) Re-evaluation of the 13C isotope fractionation associated with lipids biosynthesis by position-specific isotope analysis of plant fatty acids. *Organic Geochemistry* **174**, 104516.

Killops S. D. and Killops V. J. (2013) *Introduction to Organic Geochemistry*. 2nd ed., Wiley, Somerset.

Koepp M. (1978) D/H isotope exchange reaction between petroleum and water: A contributory determinant for D/H-isotope ratios in crude oils? *Short Papers of the Fourth International Conference, Geochronology, Cosmochronology, Isotope Geology 1978* **USGS Open-File Report 78-701 (ed. R. E. Zartman)**, 2.

Larson J. G. and Hall W. K. (1965) Studies of the Hydrogen Held by Solids. VII. The Exchange of the Hydroxyl Groups of Alumina and Silica-Alumina Catalysts with Deuterated Methane. *J. Phys. Chem.* **69**, 3080–3089.

McCollom T. and Seewald J. (2006) Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions. *Earth and Planetary Science Letters* **243**, 74–84.

McKelvie J. R., Elsner M., Simpson A. J., Sherwood Lollar B. and Simpson M. J. (2010) Quantitative Site-Specific ^{2}H NMR Investigation of MTBE: Potential for Assessing Contaminant Sources and Fate. *Environ. Sci. Technol.* **44**, 1062–1068.

Metcalfe A. and Vickers D. E. (1973) The exchange of hydrogen between methane and deuterium oxide over nickel. *Journal of Catalysis* **30**, 250–254.

Monson K. D. and 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. *J Biol Chem* **255**, 11435–11441.

Monson K. D. and Hayes J. M. (1982) Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Saccharomyces cerevisiae. Isotopic fractionation in lipid synthesis as evidence for peroxisomal regulation. *J Biol Chem* **257**, 5568–5575.

Mueller E. P., Sessions A. L., Sauer P. E., Weiss G. M. and Eiler J. M. (2022) Simultaneous, High-Precision Measurements of δ^{2}H and δ^{13}C in Nanomole Quantities of Acetate Using Electrospray Ionization-Quadrupole-Orbitrap Mass Spectrometry. *Anal. Chem.* **94**, 1092–1100.

Nguyen Tu T. T., Egasse C., Zeller B., Bardoux G., Biron P., Ponge J.-F., David B. and Derenne S. (2011) Early degradation of plant alkanes in soils: A litterbag experiment using 13C-labelled leaves. *Soil Biology and Biochemistry* **43**, 2222–2228.

Ni Y., Ma Q., Ellis G. S., Dai J., Katz B., Zhang S. and Tang Y. (2011) Fundamental studies on kinetic isotope effect (KIE) of hydrogen isotope fractionation in natural gas systems. *Geochimica et Cosmochimica Acta* **75**, 2696–2707.

Peters K. E., Walters C. C. and Moldowan J. M. (2004) *The Biomarker Guide*. 2nd ed., Cambridge University Press.

Robins R. J., Billault I., Duan J.-R., Guiet S., Pionnier S. and Zhang B.-L. (2003) Measurement of 2H distribution in natural products by quantitative 2H NMR: An approach to understanding metabolism and enzyme mechanism? *Phytochemistry Reviews* **2**, 87–102.

Romek K. M., Nun P., Remaud G. S., Silvestre V., Taïwe G. S., Lecerf-Schmidt F., Boumendjel A., De Waard M. and Robins R. J. (2015) A retro-biosynthetic approach to the prediction of biosynthetic pathways from position-specific isotope analysis as shown for tramadol. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 8296–8301.

Savage P. E. (2000) Mechanisms and kinetics models for hydrocarbon pyrolysis. *Journal of Analytical and Applied Pyrolysis* **54**, 109–126.

Schimmelmann A., Lewan M. D. and 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., Sessions A. L. and Mastalerz M. (2006) HYDROGEN ISOTOPIC (D/H) COMPOSITION OF ORGANIC MATTER DURING DIAGENESIS AND THERMAL MATURATION. *Annual Review of Earth and Planetary Sciences* **34**, 501–533.

Sessions A. L., Sylva S. P., Summons R. E. and Hayes J. M. (2004) Isotopic exchange of carbon-bound hydrogen over geologic timescales 1 1Associate editor: J. Horita. *Geochimica et Cosmochimica Acta* **68**, 1545–1559.

Sweeney J. J. and Burnham A. K. (1990) Evaluation of a simple model of vitrinite reflectance based on chemical kinetics. *AAPG Bulletin* **74**, 1559–1570.

Tang Y., Huang Y., Ellis G. S., Wang Y., Kralert P. G., Gillaizeau B., Ma Q. and Hwang R. (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.

Turner A. C., Korol R., Eldridge D. L., Bill M., Conrad M. E., Miller T. F. and Stolper D. A. (2021) Experimental and theoretical determinations of hydrogen isotopic

equilibrium in the system CH4H2H2O from 3 to 200 °C. *Geochimica et Cosmochimica Acta* **314**, 223–269.

Vinnichenko G., Jarrett A. J. M., Van Maldegem L. M. and Brocks J. J. (2021) Substantial maturity influence on carbon and hydrogen isotopic composition of n-alkanes in sedimentary rocks. *Organic Geochemistry* **152**, 104171.

Wang Y., Sessions A. L., Nielsen R. J. and Goddard W. A. (2009) Equilibrium 2H/1H fractionations in organic molecules. II: Linear alkanes, alkenes, ketones, carboxylic acids, esters, alcohols and ethers. *Geochimica et Cosmochimica Acta* **73**, 7076–7086.

Weiss G. M., Sessions A. L., Julien M., Csernica T., Yamada K., Gilbert A., Freeman K. H. and Eiler J. M. (2023) Analysis of intramolecular carbon isotope distributions in alanine by electrospray ionization Orbitrap mass spectrometry. *International Journal of Mass Spectrometry* **493**, 117128.

Wilkes E. B., Sessions A. L., Zeichner S. S., Dallas B., Schubert B., Jahren A. H. and Eiler J. M. (2022) Position-specific carbon isotope analysis of serine by gas chromatography/Orbitrap mass spectrometry, and an application to plant metabolism. *Rapid Comm Mass Spectrometry* **36**.

Xiao Y. (2001) Modeling the Kinetics and Mechanisms of Petroleum and Natural Gas Generation: A First Principles Approach. *Reviews in Mineralogy and Geochemistry* **42**, 383–436.

Xie H., Ponton C., Formolo M. J., Lawson M., Peterson B. K., Lloyd M. K., Sessions A. L. and Eiler J. M. (2018) Position-specific hydrogen isotope equilibrium in propane. *Geochimica et Cosmochimica Acta* **238**, 193–207.

Zech M., Pedentchouk N., Buggle B., Leiber K., Kalbitz K., Marković S. B. and Glaser B. (2011) Effect of leaf litter degradation and seasonality on D/H isotope ratios of n-alkane biomarkers. *Geochimica et Cosmochimica Acta* **75**, 4917–4928.

Zeichner S. S., Aponte J. C., Bhattacharjee S., Dong G., Hofmann A. E., Dworkin J. P., Glavin D. P., Elsila J. E., Graham H. V., Naraoka H., Takano Y., Tachibana S., Karp A. T., Grice K., Holman A. I., Freeman K. H., Yurimoto H., Nakamura T., Noguchi T., Okazaki R., Yabuta H., Sakamoto K., Yada T., Nishimura M., Nakato A., Miyazaki A., Yogata K., Abe M., Okada T., Usui T., Yoshikawa M., Saiki T., Tanaka Satoshi, Terui F., Nakazawa S., Watanabe S., Tsuda Y., Hamase K., Fukushima K., Aoki D., Hashiguchi M., Mita H., Chikaraishi Y., Ohkouchi N., Ogawa N. O., Sakai S., Parker E. T., McLain H. L., Orthous-Daunay F.-R., Vuitton V., Wolters C., Schmitt-Kopplin P., Hertkorn N., Thissen R., Ruf A., Isa J., Oba Y., Koga T., Yoshimura T., Araoka D., Sugahara H., Furusho A., Furukawa Y.,

Aoki J., Kano K., Nomura S. M., Sasaki K., Sato H., Yoshikawa T., Tanaka Satoru, Morita M., Onose M., Kabashima F., Fujishima K., Yamazaki T., Kimura Y. and Eiler J. M. (2023a) Polycyclic aromatic hydrocarbons in samples of Ryugu formed in the interstellar medium. *Science* **382**, 1411–1416.

Zeichner S. S., Chimiak L., Elsila J. E., Sessions A. L., Dworkin J. P., Aponte J. C. and Eiler J. M. (2023b) Position-specific carbon isotopes of Murchison amino acids elucidate extraterrestrial abiotic organic synthesis networks. *Geochimica et Cosmochimica Acta* **355**, 210–221.

Zhao Y., Ma R., Qi Y., He R., Zhu Z., Wang B., Wang Y., Yan Q., Julien M. and Zhou Y. (2023) A GC/C/IRMS-based method for position-specific carbon isotopic analysis of saturated long chain fatty acids. *Organic Geochemistry* **183**, 104652.

Chapter 6 — Intramolecular carbon and hydrogen isotopic signatures of sedimentary, abiotic, and biological n-alkanes

Guannan Dong[a], Alexandre Ferreira[b], Juliana Andrade Iemini[b], Tom McCollom[c], Katherine French[d], Justin Birdwell[d], Aivo Lepland[e,f], Jon Halvard Pedersen[g], Michael Moldowan[h,i], Alex Sessions[a], John Eiler[a]

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

[b] *Petrobras-CENPES, Ilha do Fundão, Rio de Janeiro RJ 21941-909, Brazil*

[c] *Laboratory for Atmospheric and Space Physics, Campus Box 600, University of Colorado, Boulder, CO 80309-0600, USA*

[d] *Central Energy Resources Science Center, U.S. Geological Survey, Denver Federal Center, Denver, Colorado 80225, USA*

[e] *Geological Survey of Norway, Trondheim 7491, Norway*

[f] *Department of Geology, University of Tartu, Tartu 50411, Estonia*

[g] *Lundin Norway AS, Lysaker 1366, Norway*

[h] *Stanford University, Department of Geological & Environmental Sciences, Stanford, CA 94305, USA*

[i] *Biomarker Technologies, Inc., 638 Martin Avenue, Rohnert Park, CA 94928, USA*

Abstract

Compound-specific stable isotope analysis of *n*-alkanes is widely used in organic geochemistry to identify sources of organic matter, evaluate thermal maturity, and reconstruct paleoenvironments. However, this approach can be limited by the co-occurrence of n-alkanes from different origins in many geological samples. To better constrain the sources and formation mechanisms of n-alkanes, we analyzed the intramolecular isotopic structures of n-nonadecane from a global suite of sedimentary, abiotic, and biological samples. We observe that oils from lacustrine source rocks have more ^{13}C-depleted n-alkanes (nC$_{18+}$) compared to marine oils, consistent with different sources of organic matter. The hydrogen isotope compositions of n-alkanes are generally more D-depleted in marine oils and bitumens compared to lacustrine oils, likely reflecting differences in source water.

Analysis of carbon and hydrogen isotopes in n-alkane fragments (C_4H_9, C_5H_{11}, C_6H_{13}), which reflect the intramolecular isotope distribution, reveals that the fragment-specific hydrogen isotope compositions (Δ^2H) of *n*-nonadecane vary over a 40‰ range. Brazilian lacustrine oils span this entire range, with Δ^2H values correlating with thermal maturity. In contrast, n-nonadecane in marine oils and bitumens exhibits homogeneous intramolecular hydrogen isotope distributions. We propose a model where kinetic fractionations during oil generation and hydrogen exchange produce large positive Δ^2H anomalies in n-alkanes, which are later erased by thermal maturation and equilibration, potentially catalyzed by clay minerals abundant in marine source rocks. The equilibrated state corresponds to a $\Delta^2H_{CIT\text{-}C19}$ value of about +5‰ based on our data and models. n-Nonadecane from abiotic hydrothermal Fischer–Tropsch synthesis has an intramolecular hydrogen isotope distribution close to this inferred equilibrium value, but is distinct from biological *n*-nonadecane. These findings demonstrate that the intramolecular isotopic structures of n-alkanes are powerful tools for elucidating their origins and formation mechanisms, with diverse applications to sedimentary systems, biogeochemical cycles, and the search for extraterrestrial life.

1. Introduction

Higher *n*-alkanes (C_5^+) are among the most abundant organic molecules in sedimentary rocks and are considered to be molecular fossils of aliphatic lipids and polymers. The aliphatic lipids and polymers are thought to be transformed to *n*-alkanes after deposition and during the burial of sedimentary organic matter. Higher n-alkanes and aliphatic lipids could also be synthesized abiotically through Fischer–Tropsch type (FTT) reactions in prebiotic environments for compartmentalization of protocells, or in industrial settings for clean fuel. The carbon and hydrogen isotope compositions of the biological lipid and polymer precursors reflect substrates, metabolic, physiological, and ecological processes, and environmental factors. Diagenetic and catagenetic processes can change the isotopic compositions of *n*-alkanes preserved in the rock record. Moreover, abiotic organic synthesis may generate higher n-alkanes in some geological settings, through reactions with distinctive

isotopic signatures that are influenced by factors such as reaction mechanisms and conversion rates.

Compound-specific isotope analysis (CSIA) of n-alkanes has been widely applied in various fields for source identification, thermal maturity assessment, oil-oil and oil-source rock correlation, and paleoclimate reconstruction. The carbon and hydrogen isotopic compositions of n-alkanes provide valuable insights into the origin of organic matter, such as distinguishing between C3 and C4 plants, terrestrial and aquatic sources, marine and lacustrine sources (Bjorøy et al., 1991; Collister et al., 1994; Seki et al., 2010). In sedimentary systems, the isotopic compositions of n-alkanes have been used to assess the thermal maturity of source rocks and to correlate oils with their potential source rocks (Guthrie et al., 1996; Peters et al., 2004; Tang et al., 2005; Pedentchouk and Turich, 2018). Furthermore, the carbon and hydrogen isotope composition of n-alkanes preserved in sedimentary records have been employed as proxies for reconstructing past climatic conditions, such as temperature and humidity (Sachse et al., 2012).

However, the application of CSIA to n-alkanes has several limitations. One major challenge is the presence of multiple sources and processes that can influence the isotopic composition of n-alkanes, including source input, biosynthetic pathways, environmental conditions, diagenesis, catagenesis, migration, and biodegradation (Mead et al., 2005; Sessions, 2016). The complexity of these factors can make it difficult to distinguish the specific origins and processes affecting n-alkanes in a given sample. Additionally, conventional CSIA techniques may not always be able to differentiate between biotic and abiotic sources of n-alkanes, particularly when the isotopic compositions overlap or the context is unclear (McCollom and Seewald, 2006). This limitation is especially relevant for extraterrestrial life detection applications, where distinguishing between biotic and abiotic origins of organic compounds is crucial.

To address these limitations, the analysis of intramolecular isotopic variations within n-alkanes has emerged as a promising approach. Intramolecular isotope analysis provides

additional constraints on the sources and processes affecting n-alkanes by examining the isotopic composition of individual carbon and hydrogen positions within the molecule (Gilbert et al., 2013; Julien et al., 2016) (Dong et al., Chapter 4&5). This technique can offer a more detailed understanding of the isotopic fractionation mechanisms and help to distinguish between different sources of n-alkanes. Notably, biosynthesis of aliphatic compounds, including algaenan and fatty acids, which are precursors of sedimentary hydrocarbons, elongate the carbon chains in two carbon units and typically results in a "zigzag" pattern in their intramolecular isotopic composition, i.e., a contrast in the isotope composition of neighboring positions. Previous studies have observed such intramolecular isotope variations for fatty acids from vegetable oils and leaf waxes (Monson and Hayes, 1980, 1982; Billault et al., 2001; Julien et al., 2022). In contrast, abiotic processes like FTT synthesis, which involve the elongation of carbon chains by one unit at a time, are expected to produce a more homogeneous intramolecular isotopic pattern, distinct from the "zigzag" pattern of biotic sources.

In this study, we surveyed the intramolecular isotopic structures of n-nonadecane (nC_{19}) from bitumen, oils, and Fischer–Tropsch synthesis to understand the intramolecular isotope signatures of different sources and their variations in correlation to process parameters like thermal maturity. By employing this advanced approach, we aim to provide new insights into the origins and processes affecting n-alkanes in various environmental and geological contexts, and to enhance our ability to differentiate between biotic and abiotic sources of these compounds.

2. Methods and materials

2.1. Sedimentary samples and geological background

2.1.1. Bitumen in US shales

Bitumen were extracted from three USGS shale geochemical reference materials (GRM) (Boquillas Shale, Marcellus Shale, Niobrara Shale) (Birdwell and Wilson, 2019) by Soxhlet extraction using chloroform.

Boquillas Shale bitumen is from the Late Cretaceous (Cenomanian–Turonian) strata of the Western Gulf Coast Basin in Del Rio, TX, which is equivalent to lower Eagle Ford Shale (Birdwell and Wilson, 2019). Boquillas Shale is primarily composed of organic-rich, calcareous shale, marl and limestone and was interpreted to be in upper continental slope depositional conditions of marine environment (Lock and Peschier, 2006; Lock et al., 2010; Donovan, 2016). Organic matter in Boquillas shale has high organic sulfur content (S_{org}/C = 0.41) that is generally consistent with Type II-S kerogen (Birdwell et al., 2021).

Marcellus Shale bitumen is from the Middle Devonian strata in the Appalachian Basin in LeRoy, NY (Birdwell and Wilson, 2019). Marcellus Shale is composed of black shale with interbedded limestone and siltstone and was interpreted to be deposited in a deep marine environment or on a shallower distant edge of the continental slope (Soeder et al., 2014). The Marcellus Shale reference material used in this study was from an argillaceous mudstone.

Niobrara Shale bitumen is from the Late Cretaceous strata in the Denver-Julesburg Basin in Lyons, CO. Niobrara Shale is composed of calcareous shale, marl, and limestone and was deposited in a shallow marine shelf environment (El Attar and Pranter, 2016).

The thermal maturity (determined by reflectance of solid bitumen and vitrinite macerals) and elemental composition of kerogen in the samples are listed in Table 6.1 (Birdwell and Wilson,

2019). Chromatograms of the saturate hydrocarbon fraction (analyzed by GC-FID at USGS) are shown in Figure 6.1.

2.1.2. North Sea oil

A conventional oil sample was collected from the Morkel discovery field in the North Sea. The oil was produced from the Upper Triassic Lunde Formation. The Lunde Formation is stratigraphically positioned below the Statfjord Formation; however, at this location, the Statfjord Formation is absent due to non-deposition or erosion (Norwegian Offshore Directorate, 2024a). The Lunde Formation is an interbedded sequence of fine to coarse-grained sandstones (2 to 10 m thick), claystones, marls and shales and was dominantly of continental origin, deposited in lacustrine and fluvial environments (Norwegian Offshore Directorate, 2024d). Lunde Formation is part of a structural high which is surrounded by several source kitchens capable of expelling petroleum at several different stages from the Late Cretaceous to present (Horstad et al., 1995). Specifically, the oil source rocks are mainly the Upper Jurassic Draupne Formation and also possibly the underlying Upper Jurassic Heather Formation. The Draupne Formation consists of the dark grey-brown to black, usually non-calcareous, carbonaceous, occasionally fissile claystone and was deposited in a marine environment with restricted bottom circulation and often with anaerobic conditions (Norwegian Offshore Directorate, 2024b). The Heather Formation consists mainly of grey silty claystone with thin streaks of limestone and was deposited in an open marine environment (Norwegian Offshore Directorate, 2024c).

The API gravity and biomarker analysis of the Morkel oil was orchestrated by Lundin Energy (Table 6.1). The thermal maturity was estimated based on the ββ/(ββ+αα) ratio of C_{29} steranes (isomerization at C-14 and C-17 in the C_{29} 20S and 20R regular steranes with thermal maturation) (Peters et al., 2004) (Table 6.1). A chromatogram of the saturate hydrocarbon fraction (analyzed by GC-MS at Caltech) is shown in Figure 6.1.

2.1.3. Brazilian oils

Six oils of marine source and five oils from lacustrine origins were collected from Espirito Santo Basin, Sergipe Basin, Santos Basin and Alagoas Basin in Brazil. The age, geologic formation and biomarker analysis result (performed at Petrobras) are listed in Table 6.1. The thermal maturity was estimated based on the ββ/(ββ+αα) ratio of C_{29} steranes (Peters et al., 2004) (Table 6.1).

Table 6.1 Sedimentary samples and their source and thermal maturity information, biomarker parameters and other properties.

Region	Sample name	Basin	Sample type	Age		Formation		Reservoir rock type	Organic matter source	Kerogen type
				Reservoir	Source rock	Reservoir	Source rock			
Brazil	26406	offshore Espirito Santo	oil	Lower Eocene	Turonian	Urucutuca Fm	Urucutuca Fm	Siliciclastic	marine	type II
	26404	offshore Espirito Santo	oil	Upper Paleocene	Turonian	Urucutuca Fm	Urucutuca Fm	Siliciclastic	marine	type II
	26405	offshore Espirito Santo	oil	Upper Paleocene	Turonian	Urucutuca Fm	Urucutuca Fm	Siliciclastic	marine	type II
	26397	offshore Espirito Santo	oil	Lower Eocene	Turonian	Urucutuca Fm	Urucutuca Fm	Siliciclastic	marine	type II
	26401	offshore Sergipe	oil	Paleocene	Aptian	Calumbi Fm.	Maceió Fm	Siliciclastic	marine	type II
	26396	Santos	oil	Albian	Albian	Guarujá Fm	Guarujá Fm	Carbonate	marine	type II
	26399	onshore Sergipe	oil	Barremian/Aptian	Barremian/Aptian	Morro do Chaves Fm.	Coqueiro-Seco Fm.	Carbonate	lacustrine	type I
	26400	onshore Alagoas	oil	Barremian/Aptian	Hauterivian-Barremian	Morro do Chaves Fm.	Barra de Itiúba Fm	Carbonate	lacustrine	type I
	26403	onshore Alagoas	oil	Barremian/Aptian	Hauterivian-Barremian	Coqueiro-Seco Fm.	Barra de Itiúba Fm	Siliciclastic	lacustrine	type I
	26402	onshore Alagoas	oil		Hauterivian-Barremian	Serraria Fm	Barra de Itiúba Fm	Siliciclastic	lacustrine	type I
	26398	onshore Alagoas	oil	Hauterivian-Barremian	Hauterivian-Barremian	Barra de Itiúba Fm	Barra de Itiúba Fm	Siliciclastic	lacustrine	type I
North Sea	Morkel	North Sea	oil	Upper Triassic	Upper Jurassic	Lunde Formation	Draupne Formation	Siliciclastic	marine	type II
USA	Boquillas	Western Gulf Coast	bitumen	Cenomanian–Turonian		Boquillas Formation		Shale	marine	type II-S
	Marcellus	Appalachian	bitumen	Middle Devonian		Marcellus Formation		Shale	marine	type II
	Niobrara	Denver-Jules	bitumen	Late Cretaceous		Niobrara Formation		Shale	marine	type II

Sample name	Thermal maturity		Maturity-related biomarker parameters		Source-related biomarker parameters					
	Vitrinite reflectance (%Ro)[a]	Hydrocarbon generation window	C_{29} ββ/(ββ+αα)	Ts/(Ts+Tm)[b]	TPP[c]	Tricyclic Terpane C_{26}/C_{25}	Steranes/ Hopanes	Tricyclic Terpane C_{21}/C_{23}	Kerogen H/C	API gravity
26406	0.75	peak oil	0.48	0.46	0.23	0.77	0.49	0.62		
26404	0.76	peak oil	0.49	0.50	0.24	0.81	0.44	0.59		
26405	0.76	peak oil	0.50	0.50	0.24	0.84	0.43	0.60		
26397	0.76	peak oil	0.50	0.47	0.22	0.82	0.51	0.59		
26401	0.83	peak oil	0.59	0.59	0.40	0.84	0.41	0.52		
26396	0.84	peak oil	0.61	0.81	0.36	1.02	1.79	0.69		
26399	0.67	early oil	0.35	0.37	0.64	1.35	0.15	0.76		
26400	0.71	peak oil	0.41	0.61	0.67	1.58	0.06	0.94		
26403	0.72	peak oil	0.43	0.62	0.61	1.77	0.11	0.98		
26402	0.85	peak oil	0.62	0.74	0.82	1.22	0.89	0.82		
26398	0.91	late oil	0.71	0.40	0.51	1.53	0.81	0.97		
Morkel	0.68	peak oil	0.38	0.46						41
Boquillas	0.54 ± 0.07	early oil							1.21 ± 0.10	
Marcellus	0.73 ± 0.09	peak oil							1.03 ± 0.17	
Niobrara	0.99 ± 0.09	late oil							0.82 ± 0.08	

[a] Vitrinite reflectance (%Ro) data source: Boquillas: reflectance of vitrinite macerals. Niobrara and Marcellus: reflectance of solid bitumen. Morkel and Brazilian oils: estimation based on the ββ/(ββ+αα) ratio of C_{29} steranes

[b] Ts/(Ts+Tm) parameter has a strong dependence on source

[c] TPP: Tetracyclic polyprenoid ratio

2.2. Fischer–Tropsch synthesis under hydrothermal conditions

Fischer–Tropsch synthesis was conducted under hydrothermal conditions to produce *n*-alkanes abiotically at University of Colorado, Boulder. The methods employed are essentially identical to those used previously to study Fischer–Tropsch-type (FTT) reactions by (McCollom and Seewald, 2006; McCollom et al., 2010). The experiments were conducted using a flexible-cell reaction apparatus where the reaction cell is composed of a gold bag with titanium fittings, contained inside a stainless-steel pressure housing. Prior to the experiments, the reaction cell components were heated for two hours at 450°C in air to remove organic contaminants and to create an inert TiO_2 layer on the surface of the titanium fittings. Previous experiments with the same reaction system have demonstrated that the reactor itself is not catalytic for organic synthesis under the experimental conditions (McCollom and Seewald, 2006; McCollom et al., 2010).

The experiments were performed by heating carbon monoxide (CO) and native iron (Fe^0) in the presence of an aqueous solution at 215–220°C (Table 6.2). The experiments were initiated by placing 1.5 g of Fe^0 (Fe powder, -22 mesh; Alfa Aesar) and ~25 g of deionized (DI) water or deuterated DI water in the gold reaction cell (cell volume was approximately 55 ml). The native Fe is included both as catalyst and as a source of molecular hydrogen (H_2) since it rapidly reacts with water at elevated temperatures to form magnetite (Fe_3O_4) plus H_2 according to the reaction:

$$3Fe^0 + 4H_2O \rightarrow Fe_3O_4 + 4H_2 \qquad (6\text{-}1)$$

To promote reaction of Fe^0, the water was acidified to a pH of about 2 with HCl. The target pressure for the experiments was 15 MPa, but in practice the pressures varied up to 3 MPa from this value during the course of some experiments.

After sealing the reaction cell with the Ti closure piece, it was placed in the pressure containment vessel, pressurized at room temperature, and the entire apparatus was placed

horizontally and heated in a furnace. Once the target temperature was attained, CO was injected through the sampling valve using the following procedure. First, a loop of stainless-steel tubing (12.6 ml) was filled with CO at a pressure of 0.93 MPa (the maximum of the pressure regulator on the CO tank) and attached to the sampling valve. A solution with the same composition as that in the reaction cell was then pumped into the tubing until it was equal to that within the heated pressure vessel. The valve was then opened and the CO along with 13 ml of solution was pumped into the reaction cell, with an equal amount of water bled out of the pressure containment vessel to maintain constant pressure during injection. In practice, some amount of CO remained in the injection loop after this procedure, so it is not possible to determine exactly how much CO was contained in the reaction cell following injection. However, assuming 90% injection efficiency, the amount of CO injected would have been about 4.3 millimoles, or 113 mmol/kg H_2O. The tank of CO used in the experiments was the same as that used in (McCollom et al., 2010) and its carbon isotope composition $\delta^{13}C_{VPDB}$ is -28‰.

The experiments were heated for about 40 hours. Following the reaction, the reactors were gradually cooled to room temperature while pumping water into the containment vessel to maintain elevated pressure. The pressure was then reduced to ~5 MPa by releasing water from the pressure containment vessel to allow volatile gases to exsolve and accumulate at the top of the reaction cell. Within 24 h of experiment termination, the gas phase was vented through a sampling valve and transferred to a stainless-steel cylinder for other studies. ~3 mL of dichloromethane (DCM) was then injected into the reaction cell and shaken for ~15 min to extract non-volatile organic products.

D_2O (99.9 atom % D, Sigma Aldrich, LOT # MKCP5525) was added to the DI water used in experiments FT22-1, FT22-2, and FT22-7 to increase the abundance of ^2H in the synthetic products for better analytical performance. The hydrogen isotope composition of the deuterated DI water substrate was measured on a Picarro L2140-i isotopic water analyzer at Caltech and calibrated against 4–6 working standards with δ^2H values ranging from −138‰

to 1273‰. These standards were in turn calibrated against the VSMOW, Greenland Ice Sheet Precipitation, and Standard Light Antarctic Precipitation international standards (Coplen, 1988) (Table 6.2). Experiments FT22-1 and FT22-2 were conducted in May 2022 and FT22-7 in August 2022. The bottles of deuterated DI water were stored with closed lids on lab bench between the experiments.

Table 6.2 Hydrothermal Fischer–Tropsch synthesis experiment conditions

analytical sample	synthesis experiment	temperature (°C)	duration (hr)	final pressure (psi)	substrate water δ^2H_{VSMOW} (‰)[a]		
					before experiments	after experiments[b]	analytical error
FT22-1&2	FT22-1	218	43	1770	731	721	0.3
	FT22-2	217	43	1988			
FT22-7	FT22-7	220	48	2388			
FT21	FT21-1	215	40	2000	NA[a]		
	FT21-2	222	45	2100			

[a] Not measured for FT21-1 and FT21-2 experiments. The substrate water in those experiments was DI water purified from tap water in Boulder, Colorado (~ −119‰ in April 2022 and April 2023)

[b] This measurement was on residual substrate water not used in the synthesis experiments to monitor any change in the substrate isotope composition across different experiments at different times

2.3. Purification

2.3.1. Sedimentary samples

The n-alkanes were isolated from the sedimentary samples to ensure clean chromatographic separation in compound-specific and fragment-specific isotope analysis. Samples of Morkel oil, Boquillas, Marcellus and Niobrara bitumen were processed at Caltech and the Biomarker Technologies, Inc. and Brazilian oils were processed at Petrobras.

10–100 mg whole oil and bitumen samples were dissolved in dichloromethane (DCM). Elemental sulfur was removed from the dissolved samples by percolating through a cupper powder-packed Pasteur pipette column (pre-rinsed with hydrochloric acid, water, acetone and DCM). Asphaltenes were precipitated from the sulfur-free samples by evaporating and re-dissolving in 1–4 mL of n-hexane and filtering through a pre-combusted glass bead column. The asphaltene-free fraction underwent column chromatography to isolate the

saturated hydrocarbon fraction. Samples were loaded onto a preactivated-silica gel column (20 cm length, 1 cm inner diameter) and eluted with n-hexane. The saturated hydrocarbon fraction was collected, and the solvent removed under a gentle stream of N_2 gas.

n-Alkanes were isolated from the saturated hydrocarbon fraction via urea adduction. A saturated solution of urea in methanol (~1 ml) was added dropwise to 10 mg of the saturate fraction dissolved in 6 ml of a pentane:acetone (2:1) mixture. The resulting white urea precipitate was transferred to a 10 ml column plugged with filter paper. Non-adducted materials were removed by rinsing the vial containing the precipitate twice with pentane and adding the rinses to the column. The crystals were isolated by expelling the solvent from the column with gentle N_2. The adducted compounds were released by dissolving the precipitate with 3 ml of H_2O, followed by a 1 ml pentane rinse. The H_2O/pentane mixture was expelled from the column with gentle N_2, collected in a vial, and additional pentane was added to achieve a ~50:50 pentane:H_2O ratio. The pentane layer was extracted twice by vortexing and pipetting, with the combined n-pentane extracts containing the isolated n-alkanes.

2.3.2. Hydrothermal Fischer–Tropsch synthesis products

We separated *n*-alkanes from the DCM extracts of hydrothermal Fischer–Tropsch synthesis experiments by solid-phase extraction (SPE) following the method described in (Sessions, 2006; Dion-Kirschner et al., 2020). Briefly, aminopropyl SPE separations used 0.5g of Sepra NH2 (50um, 65A) stationary phase (Phenomenex, batch# S209-098 or S209-100) hand-packed into 8 mL glass syringe barrels. DCM extracts were evaporated and redissolved in ~0.3 mL *n*-hexane. The solutions were eluted in four fractions: F1 (hydrocarbons), 4 mL hexane; F2 (ketones and ethers), 7 mL 4:1 hexane/dichloromethane (DCM); F3 (alcohols), 7 mL 9:1 DCM/acetone; F4 (carboxylic acids), and 9mL 49:1 DCM/formic acid (all solvent ratios in v/v). Following the aminopropyl SPE, Ag-ion columns were used to remove alkenes from *n*-alkanes. Briefly, 0.5 g of silver nitrate on silica gel stationary phase (Sigma Aldrich, batch# MKCP3740) hand-packed into 8 mL glass syringe barrels. The fraction of

hydrocarbons (F1) was eluted in three fractions: F1 (alkanes), 5 mL hexane; F2 (alkenes), 5 mL 4:1 hexane/DCM; F3 (blank), 5 mL acetone (all solvent ratios in v/v).

2.4. Compound-specific carbon and hydrogen isotope measurement

The isolated *n*-alkanes were analyzed for compound-specific carbon and hydrogen isotope composition. Samples of Morkel oil, Boquillas, Marcellus and Niobrara bitumen, FTT synthesis products, and rose essential oil and rose essential oil (Dong et al., chapter 4) were analyzed at Caltech by GC-combustion-isotope ratio mass spectrometry (IRMS) (for $\delta^{13}C_{VPDB}$) and GC-pyrolysis-IRMS (for $\delta^{2}H_{VMOW}$) on a Thermo-Scientific Delta+ XP and Brazilian oils were analyzed at Petrobras.

About 1 uL of 0.1 mg/mL (0.37 mM) CIT-C19 *n*-hexane solution or isolated n-alkanes with similar *n*-nonadecane concentration was injected to a ZB-5MS (30 m × 0.25 mm × 0.25 μm film thickness) column through a PTV injector operated in Large Volume mode, using helium as a carrier gas with a flow rate of 1.4 mL/min.

Co-injected peaks of CO_2 ($\delta^{13}C_{VPDB}$ = -12.04‰) and H_2 ($\delta^{2}H_{VSMOW}$ = -224‰) reference gases were used for single-point isotopic calibration of analytes. The reference gases were in turn calibrated against a standard solution of eight fatty acid methyl esters ('F8 mix' of Arndt Schimmelmann, Indiana University). The F8 mix was also run along with the samples to monitor the instrumental status.

Molecular average carbon and hydrogen isotope composition is expressed using delta notation relative to standards including Vienna Pee Dee Belemnite (VPDB), Vienna Standard Mean Ocean Water (VSMOW):

$$\delta^{13}C_{VPDB} = \frac{^{13}R_{sample}}{^{13}R_{VPDB}} - 1 \quad (6\text{-}2)$$

$$\delta^2 H_{VSMOW} = \frac{^2R_{sample}}{^2R_{VSMOW}} - 1 \quad (6\text{-}3)$$

where $^{13}R = \frac{^{13}C}{^{12}C}$ and $^{2}R = \frac{^{2}H}{^{1}H}$; and $\delta^{13}C$ and $\delta^{2}H$ values are reported in per mil notation. The long-term experimental reproducibility for measurements of standards and replicate samples on these instruments is 0.15‰ for $\delta^{13}C$ and 5‰ for $\delta^{2}H$ (1SE).

2.5. Fragment-specific carbon and hydrogen isotope analysis

We analyzed the ^{13}C- and ^{2}H-substituted isotopologues of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions of *n*-nonadecane via peak capture measurement on a Thermo Scientific Q Exactive Orbitrap mass spectrometer with a Trace 1310 GC and a Thermo Scientific Triplus RSH autosampler (Dong et al., Chapter 4).

Briefly, about 1 uL of 0.1 mg/mL (0.37 mM) CIT-C19 solution or isolated *n*-alkane solution with similar nonadecane concentration was injected to a TG-5SilMS column (30 m × 0.25 mm × 0.25 μm film thickness) through a split/splitless injector held at 250℃ and operated in the splitless mode, using helium as a carrier gas with a flow rate of 1.2 mL/min. The GC oven was started with at 50 or 100℃, heated to 320℃, and held at 320℃ for tens of minutes. During the elution of the entire chromatographic peak of *n*-nonadecane, we directed and mixed the elute into a 10 cm^3 passivated stainless-steel peak broadener prefilled with helium. Then a flow of pure helium carrier gas (1.2 ml/min) purged the peak broadener and delivered the trapped *n*-nonadecane to the electron impact (EI) Orbitrap mass spectrometer over about 30 minutes. The transfer line was set at 265℃. The EI source was set at 250℃. The filament was turned on at 5 min. We measured one out of the three fragments in each analysis and used a quadrupole scan range of 52–63 mz for C_4H_9 fragment, 66.5–76.5 mz for C_5H_{11}, and 81–90 m for C_6H_{13}. We used the prescan AGC (automatic gain control) mode with an AGC target of 2e5 and a maximum injection time of 3000 ms (though the injection time is within 1 to 200 ms for scans during which sample was present). We used the positive ion mode and no 'lock masses' were used. The nominal mass resolution was set at 120,000 (FWHM, at

m/z = 200). The number of microscans was set to 1. We ran the acquisition for the duration of the GC oven temperature program, which we established was the time required for the mass spectra of effluent from the peak broadener to reach peak intensities indistinguishable from those observed in blank (i.e., nonadecane-free) measurements. Each analysis of the pyrolysis product sample was bracketed by two analyses of CIT-C19, and each bracketed analysis was repeated 5 times for each fragment.

A solution of perfluorotributylamine (PFTBA, Sigma-Aldrich, LOT# 77299) was used to calibrate the mass scale and mass isolation for positive ions and to auto-tune the ion source, injection flatapole and inter-flatapole lens with electron energy set at 70 eV and C-trap energy offset at 0 V, maximizing the signal intensity on the Q Exactive GC.

The data analysis method has been detailed in Chapter 3 and 4. Briefly, The Thermo Scientific FTStatistic software and Python scripts (available at https://doi.org/10.22002/catyq-fxd56) are used to process mass spectral data from peak capture experiments. The process involves culling scan-by-scan data based on a 10% threshold of base ion intensity for peak capture measurements or based on elution time for direct elution measurements, removing scans that are 3σ outliers in TIC×IT (total ion current × injection time), and calculating scan-by-scan ion counts using a formula that incorporates ion intensity, noise, resolution, and charge. Finally, we calculate the isotope ratio of the fragment ion by

$$ {}^{J}R_{frag} = \frac{\frac{\sum_i {}^{j1}N_i \times {}^{j1}I_i}{\sum_i {}^{j1}I_i}}{\frac{\sum_i {}^{j2}N_i \times {}^{j2}I_i}{\sum_i {}^{j2}I_i}} \qquad (6\text{-}4) $$

where ${}^{j1}N_i$ and ${}^{j2}N_i$ are the counts of different isotopologues of fragment ions in scan i, respectively, and ${}^{j1}I_i$ and ${}^{j2}I_i$ are the intensities of different isotopologues for scan i.

Site-specific isotope composition refers to the isotope ratios (e.g., $^{13}C/^{12}C$ or $^{2}H/^{1}H$) at individual molecular positions within a compound, whereas fragment-specific isotope composition describes the isotope ratios of molecular fragments generated during the ionization and fragmentation of a parent molecule in a mass spectrometer.

The fragment-specific isotope composition of *n*-nonadecane relative to CIT-C19 are calculated using a delta notation:

$$\delta^{13}C_{CIT-C19} = \frac{^{13}R_{frag,sample}}{^{13}R_{frag,CIT-C19}} - 1 \quad (6\text{-}5)$$

$$\delta^{2}H_{CIT-C19} = \frac{^{2}R_{frag,sample}}{^{2}R_{frag,CIT-C19}} - 1 \quad (6\text{-}6)$$

where $\delta^{13}C$ and $\delta^{2}H$ values are reported in per mil notation. To isolate the intramolecular isotope structure variation and eliminate the influence of differences in molecular average isotope composition, normalized fragment-specific isotope composition was also derived, which was achieved by offsetting the molecular average isotope fractionation:

$$\Delta^{13}C_{CIT-C19} = \frac{^{13}R_{frag,sample}}{^{13}R_{frag,CIT-C19}} / \frac{^{13}R_{mol,sample}}{^{13}R_{mol,CIT-C19}} - 1 \quad (6\text{-}7)$$

$$\Delta^{2}H_{CIT-C19} = \frac{^{2}R_{frag,sample}}{^{2}R_{frag,CIT-C19}} / \frac{^{2}R_{mol,sample}}{^{2}R_{mol,CIT-C19}} - 1 \quad (6\text{-}8)$$

where $\Delta^{13}C$ and $\Delta^{2}H$ values are also reported in per mil notation. The long-term experimental reproducibility for measurements of standards and replicate samples is about 1‰ for $\Delta^{13}C$ and 5‰ for $\Delta^{2}H$ (1SE).

3. Result and discussion

3.1. Source and thermal maturity of North Sea and Brazilian oils

The source- and maturity-related biomarker results of Brazilian and North Sea oils are plotted in Figure 6.2. The Morkel oil is interpreted to be a mix of a high thermal maturity fraction (expressed by high API gravity (Table 6.1) resulted from the light hydrocarbons, from the latest stage of charge) and a lower maturity fraction (expressed by the biomarkers, from an early stage of charge). The origins of the Brazilian oils were characterized using saturated biomarker parameters TPP (tetracyclic polyprenoid ratio), tricyclic terpane ratios, and steranes/hopanes ratio (Figure 6.2), along with the compound-specific isotopic composition of *n*-alkanes (Section 3.3). There is no evidence of biodegradation. The two maturity parameters ββ/(ββ+αα) ratio of C_{29} steranes and Ts/(Ts+Tm) corroborate with each other.

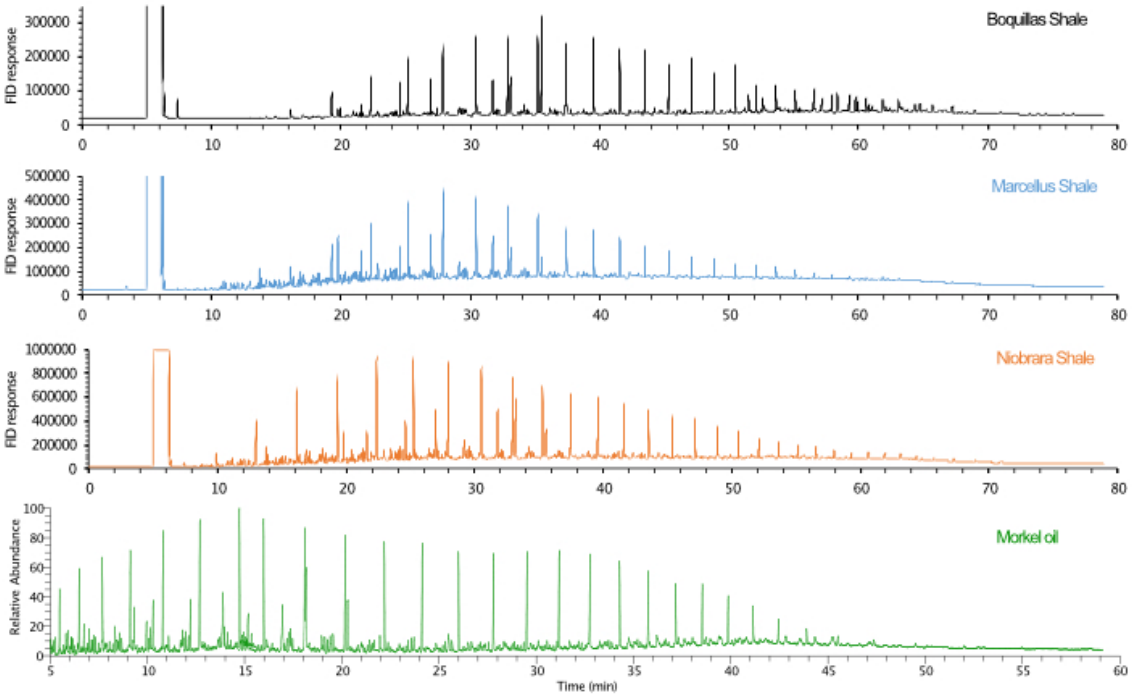


Figure 6.1 Chromatograms of the saturate hydrocarbon fraction analyzed by GC-FID or GC-MS for US marine shale extracts and Morkel oil.

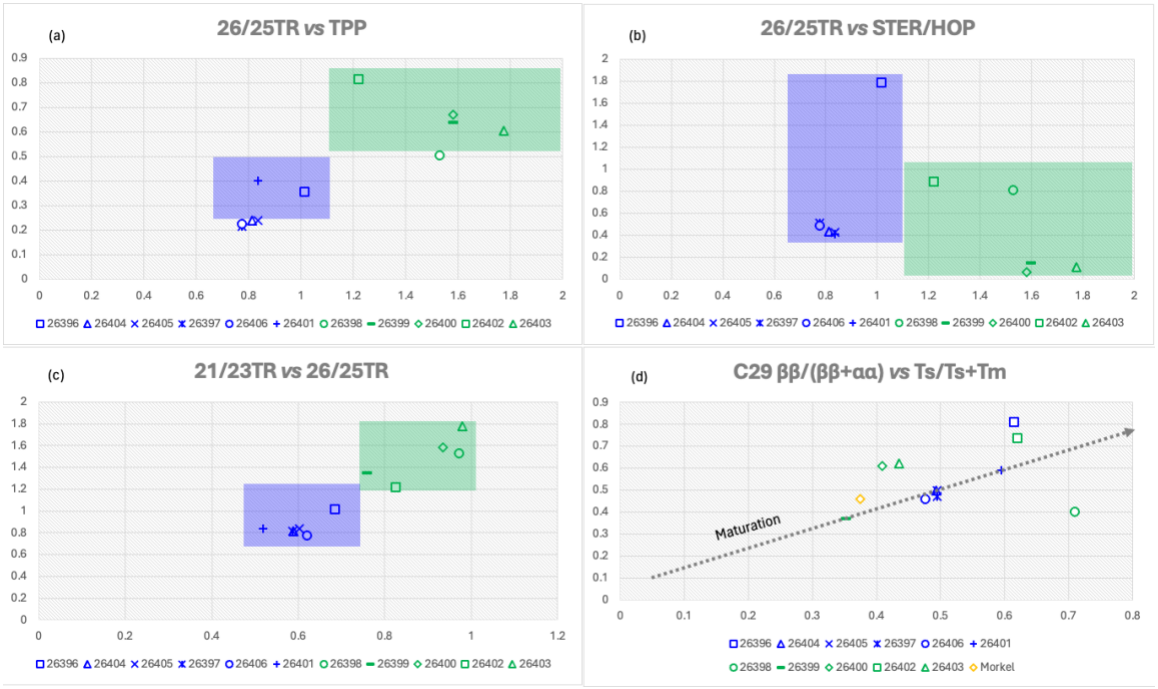


Figure 6.2 Source- (a, b, c) and maturity- (d) related biomarker analysis of Brazilian and North Sea oils (Table 6.1). The origins of the Brazilian oils were characterized by source-related biomarker parameters TPP (tetracyclic polyprenoid ratio), tricyclic terpane ratios, and steranes/hopanes ratio. The two maturity parameters ββ/(ββ+αα) ratio of C_{29} steranes and Ts/(Ts+Tm) corroborate with each other. The blue symbols and field are from marine source whereas the green symbols and field indicate lacustrine source.

3.2. Chemical composition of FTT synthesis products

n-Alkanes, alkenes, alcohols, and fatty acids are produced from FTT synthesis and generally follow an Anderson-Schulz-Flory distribution, i.e., a logarithmic relationship where the mole fraction of products decreases as the carbon number increases (Table 6.3, Figure 6.3). This observation is consistent with previous studies and is a result of carbon chain growth controlled by a constant probability factor α independent of chain length (Mccollom et al., 1999; McCollom and Seewald, 2006; McCollom et al., 2010). The chain growth probability factor p is defined as

$$X_n = (1-p) \times p^{n-1} \quad (6\text{-}9)$$

Where X_n is the mole fraction of the product with carbon number n and chain growth probability factor $0 < p < 1$ where higher value favors the formation of longer chain hydrocarbons and indicates a lower probability of chain termination. The chain growth probability factor p can be calculated by plotting the logarithm of n-alkane concentrations against carbon number and determining the slope, which equals ln(p). We calculated p in our experiments and found it generally consistent with previous studies (McCollom and Seewald, 2006). Experiment FT22-7 appears to produce more abundant higher *n*-alkanes than FT22-1 and FT22-2 experiments (Figure 6.3).

Table 6.3 Abundance and chain growth probability of *n*-alkanes in hydrothermal Fischer–Tropsch synthesis experiments.

Experiment	TIC area of n-alkane											chain growth probablity		
	nC14	nC15	nC16	nC17	nC18	nC19	nC20	nC21	nC22	nC23	nC24	ln (p)	SE	p
FT22-1	1.61E+07	6.49E+07	9.79E+07	1.08E+08	1.02E+08	8.67E+07	6.90E+07	5.05E+07	3.37E+07	2.20E+07		-0.12	0.01	0.88
FT22-2		1.93E+07	4.01E+07	5.79E+07	6.24E+07	5.96E+07	5.18E+07	4.13E+07	3.06E+07	2.18E+07	1.44E+07	-0.12	0.01	0.88
FT22-7	8.39E+06	3.59E+07	5.41E+07	6.27E+07	5.89E+07	5.04E+07	4.11E+07	3.12E+07	2.15E+07	1.43E+07		-0.15	0.01	0.86

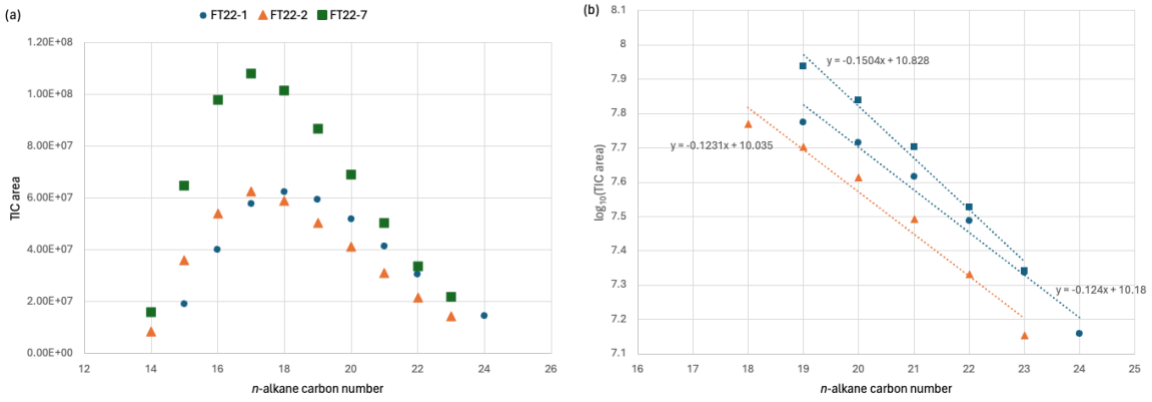


Figure 6.3 Abundance (a) and chain growth probability (b) of *n*-alkanes in hydrothermal Fischer–Tropsch synthesis experiments. The abundance of *n*-alkanes generally follows an Anderson-Schulz-Flory distribution. Evaporation of of the DCM extracts resulted in partial loss of some more volatile, lower molecular weight hydrocarbons (C14~C18), but loss of heavier compounds was probably negligible. The chain growth probability factor α is defined as $X_n = (1-\alpha) \times \alpha^{n-1}$ where X_n is the mole fraction of the product with carbon

number n and chain growth probability factor $0 < \alpha < 1$. Higher α value favors the formation of longer chain hydrocarbons and indicates a lower probability of chain termination. The chain growth probability factor α is calculated in (b) by plotting the logarithm of n-alkane concentrations against carbon number and determining the slope, which equals ln(α).

3.3. Compound-specific isotope composition of *n*-alkanes

3.3.1. Carbon isotope profile of sedimentary *n*-alkanes

The compound-specific carbon isotope composition (Table 6.4) versus the carbon number of n-alkanes constitutes a series of profiles for the samples (Figure 6.4). For sedimentary samples, profiles of oils from lacustrine origin stand out with more depleted $\delta^{13}C$ (~ -35‰) of longer-chain n-alkanes (nC_{18+}), in contrast to $\delta^{13}C$ of about -30 to -27‰ for n-alkanes in marine oils. This is consistent with previous studies (Bjorøy et al., 1991; Guthrie et al., 1996). This contrast can be explained by contribution of terrestrial and lacustrine versus marine organic matter to the source rock. Terrestrial plant synthesizes organic matter by photosynthesis using CO_2 in the air, resulting in a ~ -30‰ carbon isotope fractionation. Marine phytoplankton utilizes dissolved inorganic carbon (DIC) as inorganic carbon substrates and the equilibrium carbon isotope fractionation between DIC and CO_2 in the air ($\varepsilon_{DIC\text{-}CO2}$) is about 8‰ at 25°C. Some marine organisms also use carbon concentrating mechanisms that would reduce the isotope fractionation in carbon fixation process. These factors lead to marine source organic matter more enriched in $\delta^{13}C$ than terrestrial source. Phytoplankton in lacustrine environments also uses DIC for carbon fixation. Lacustrine DIC is more influenced by input from terrestrial sources, such as weathering of carbonate rocks and soil respiration, and is often depleted in ^{13}C compared to marine DIC. Therefore, n-alkanes from lacustrine oils typically have more depleted $\delta^{13}C$ values compared to n-alkanes in marine oils. Additionally, n-alkanes in bitumen from US shales and the North Sea oil have lower $\delta^{13}C$ values than the Brazilian marine oils, which may indicate higher proportion of terrestrial organic matter contribution to these US shales and the North Sea source rock.

There is no uniform correlation of compound-specific $\delta^{13}C$ and thermal maturity. Only bitumen samples show an increase in compound-specific $\delta^{13}C$ with thermal maturity.

3.3.2. Carbon isotope profile of abiotic *n*-alkanes

The $\delta^{13}C$ of FTT synthesis products is about -50 to -60‰ and consistent with previous studies (Table 6.4, Figure 6.4) (McCollom and Seewald, 2006; McCollom et al., 2010). The isotope fractionation between the product and the inorganic carbon substrate is about -30‰, indistinguishable from carbon fixation by Calvin cycle. Some carbon isotope profiles of Fischer–Tropsch product have a positive slope and are most likely due to a reservoir effect on the inorganic carbon substrate CO. I.e., n-alkanes with fewer carbons are earlier products in the chain elongation synthetic process and their isotope composition approximates $\delta^{13}C$ of the original CO substrate offset by the KIE of carbon-carbon bond formation. As the Rayleigh fractionation progresses, $\delta^{13}C$ of CO becomes more enriched and the $\delta^{13}C$ of later products also becomes more enriched. This explanation is consistent with the reservoir sizes of the inorganic carbon substrates (McCollom and Seewald, 2006; McCollom et al., 2010), where the concentrations of HCOOH and CO substrates are 125 and 80 mmol/kg H2O, respectively, and the carbon isotope profiles of experiments using CO have slopes more positive than HCOOH experiments. In conclusion, such carbon isotope profile with a positive slope may be a signature for kinetically-controlled elongation synthesis with a limited reservoir of substrates where the slope is determined by the reservoir size of the substrate. This could be a potential signature for higher *n*-alkanes from some abiotic synthesis because sedimentary and biological higher *n*-alkanes have not been found to have carbon isotope profiles with positive slopes.

Table 6.4 Compound-specific carbon isotope composition of *n*-alkanes from sedimentary hydrocarbons, Fischer–Tropsch type synthesis and the rose essential oil.

	$\delta^{13}C_{VPDB}$ of n-alkane (‰)																					
sample name	nC12	nC13	nC14	nC15	nC16	nC17	nC18	**nC19**	nC20	nC21	nC22	nC23	nC24	nC25	nC26	nC27	nC28	nC29	nC30	nC31	nC32	nC33
Boquillas					-31.6	-31.5	-31.3	**-31.4**	-31.2	-31.1	-30.9	-30.8	-30.8	-30.7	-31.0							
Marcellus			-30.1	-30.4	-29.9	-30.3	-29.7	**-30.1**	-30.0	-30.1	-30.0											
Niobrara		-29.5	-29.7	-29.8	-29.8	-29.8	-29.7	**-29.7**	-29.7	-29.7	-29.6	-29.9	-29.6	-29.7								
Morkel			-30.0	-30.2	-30.0	-29.8	-29.5	**-29.2**	-29.1	-29.0	-28.9	-29.0										
26406	-26.3	-26.6	-26.5	-26.4	-26.6	-26.8	-26.8	**-26.9**	-27.0	-27.0	-26.9	-26.8	-27.0	-26.9	-27.2	-26.9	-27.0	-27.1	-27.3	-27.7	-27.4	
26404				-27.0	-26.9	-27.0	-26.8	**-27.0**	-27.0	-27.0	-27.0	-27.1	-27.2	-27.3	-27.4	-26.6	-26.9	-26.6	-26.9	-27.6		
26405			-27.4	-27.0	-27.0	-27.3	-26.9	**-26.9**	-27.1	-27.2	-27.1	-27.1	-27.2	-27.0	-27.1	-26.7	-27.0	-26.7	-27.1	-27.4	-27.4	
26397			-26.8	-27.0	-27.1	-27.2	-27.3	**-27.2**	-27.1	-27.1	-27.1	-27.1										
26401	-27.1	-27.4	-27.7	-27.5	-27.5	-27.6	-27.4	**-27.2**	-27.3	-27.3	-27.2	-27.1	-27.2	-27.2	-27.1	-27.0	-27.1	-27.3	-27.5	-27.9	-27.5	
26396				-26.5	-27.4	-27.1	-27.2	**-27.1**	-27.4	-27.2	-27.2	-26.9	-26.9	-26.9	-26.9	-26.9	-26.9	-26.7	-27.2	-26.7	-27.1	-26.7
26399							-33.2	**-33.9**	-34.2	-34.8	-35.0	-35.5	-35.2	-35.5	-35.1	-35.2	-34.9	-34.8	-34.4	-34.9	-33.3	-33.5
26400			-28.6	-31.4	-32.1	-33.3	-34.0	**-34.9**	-35.5	-35.8	-35.9	-36.3	-35.4	-35.5	-35.6	-35.7	-35.0	-35.6	-35.1	-35.1	-34.9	
26403				-31.1	-33.9	-35.3	-34.9	**-35.1**	-34.9	-35.7	-35.4	-35.6	-35.4	-35.3	-35.5	-35.4	-35.2	-35.4	-35.3	-35.9	-35.9	
26402		-29.4	-29.4	-29.9	-30.2	-30.9	-30.8	**-31.2**	-31.1	-31.4	-31.4	-31.7	-31.5	-31.9	-31.8	-31.8	-31.4	-31.8	-30.9			
26398			-29.9	-30.2	-30.7	-31.6	-32.3	**-33.4**	-34.1	-34.7	-35.0	-35.5	-35.9	-36.1	-36.3	-36.5	-36.4	-36.2	-36.2	-35.8	-35.8	-35.4
FT22-1&2				-61.9	-61.5	-61.2	-60.5	**-60.4**	-60.1	-59.6												
FT22-7				-62.0	-61.9	-61.6	-61.1	**-61.0**	-60.9													
FT21			-60.1	-60.4	-60.1	-59.9	-59.9	**-59.5**	-59.4	-58.6	-58.4	-57.3										
rose								**-28.6**		-28.3												
CO #1 (68 hr)[a]				-56.7	-55.8	-54.1	-53.7	**-52.7**	-51.8	-51.0	-50.7	-49.7	-48.8	-47.7	-47.5	-46.4	-47.3	-46.6	-46.6			
CO #2 (285 hr)[a]				-54.9	-54.2	-53.3	-51.5	**-51.5**	-50.2	-50.0	-49.2	-48.3	-47.5	-47.1	-46.8	-45.8	-46	-44.6	-44.6			
CO #3 (241 hr)[a]				-56.3	-54.9	-52.7	-50.5	**-52.6**	-51.4	-51.4	-50.2	-49.3	-49.3	-48.5	-47.9	-47.5	-47.3	-46.9	-46.7	-46.2	-46.7	
HCOOH (390 hr)[a]						-59.2	-59	**-60.5**	-60.2	-60.4	-60.2	-60	-59.7	-59.7	-59.7	-59.5	-59.4	-59.9	-60.3	-60.4	-60.8	-60.2
HCOOH (86 hr)[b]	-50.5	-50.7	-50.7	-51	-50.1	-50.3	-49.8	**-50.2**	-50.1	-50.6	-50.5	-50.3	-50.6	-50.5	-51	-51.8						

[a] FTT synthesis data from McCollom et al., 2010. Experimental temperature 250°C. [b] FTT synthesis data from McCollom and Seewald, 2006. Experimental temperature 250°C.

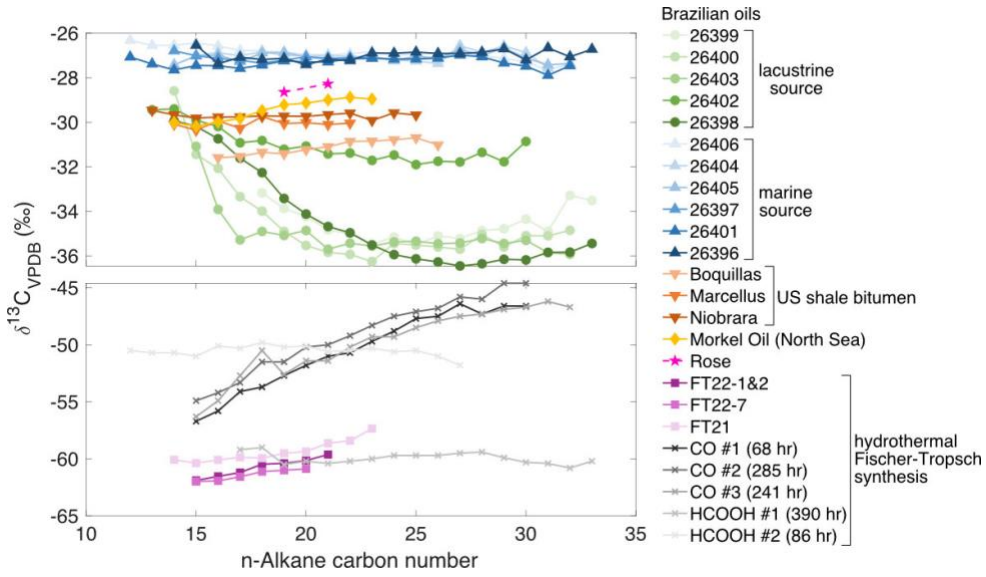


Figure 6.4 Compound-specific carbon isotope composition of *n*-alkanes from sedimentary hydrocarbons, Fischer–Tropsch type synthesis and the rose essential oil. $\delta^{13}C$ values of *n*-alkanes in oils from lacustrine origin are more depleted (~ -35‰ for nC_{18+}) compared to marine oils (-30 to -27‰), consistent with the contribution of terrestrial and lacustrine versus marine organic matter to the source rocks. The $\delta^{13}C$ values of *n*-alkanes from hydrothermal Fischer–Tropsch synthesis are about -50 to -60‰, with some profiles showing a positive slope likely due to a reservoir effect on the inorganic carbon substrate CO.

3.3.3. Hydrogen isotope profile of sedimentary *n*-alkanes

For sedimentary samples, the $\delta^{2}H$ values of *n*-alkanes in marine oils and bitumen are generally more depleted than in lacustrine oils (Table 6.5, Figure 6.5). This observation is consistent with previous studies and is likely due to the higher $\delta^{2}H$ of low-latitude evaporative lake and meteoric waters (Sessions, 2016). The bitumen n-alkanes appear to be most depleted in $\delta^{2}H$ among sedimentary samples, also consistent with previous observations and could indicate hydrogen exchange processes and kinetic fractionations to a lesser extent than expelled oils (Sessions, 2016), though the detailed mechanism is elusive. The hydrogen isotope profile of lacustrine oils appears to be "rockier". This could result from a more variable depositional setting in lacustrine environments, with fluctuations in water level, salinity, and input of terrestrial organic matter. These variations can lead to heterogeneity in

the hydrogen isotope composition of the organic matter and the resulting *n*-alkanes. There is no uniform correlation of compound-specific δ^2H and thermal maturity. Only Brazilian lacustrine-sourced oils show a general increase in compound-specific δ^2H of *n*-alkanes with thermal maturity.

3.3.4. Hydrogen isotope profile of abiotic *n*-alkanes

We report for the first time δ^2H of higher *n*-alkanes produced by FTT synthesis (Table 6.5, Figure 6.5). The hydrogen isotope composition of higher *n*-alkanes from the experiment using non-deuterated DI water is comparable to H_2, methane, and ethane products from previous experiments (McCollom et al., 2010). The hydrogen isotope fractionation factor between higher *n*-alkane products and water substrate is about -500‰ for FTT experiments at 200°C. The equilibrium hydrogen isotope fractionation factor between methyl or methylene hydrogen and water is extrapolated to be -110 to -120‰ (Wang et al., 2009). The FTT *n*-alkane products are clearly out of isotopic equilibrium with water. *n*-Alkanes in FT22-7 appears to have higher δ^2H values than FT22-1&2. This difference might be a result of a higher conversion rate of H_2 to organics for FT22-7 than FT22-1&2. The slope of hydrogen isotope profile for FTT synthesis products is comparable to sedimentary samples.

Table 6.5 Compound-specific hydrogen isotope composition of *n*-alkanes from sedimentary hydrocarbons, Fischer–Tropsch type synthesis and the rose essential oil.

	δ^2H_{VSMOW} of n-alkane (‰)																						
sample name	nC13	nC14	nC15	nC16	nC17	nC18	**nC19**	nC20	nC21	nC22	nC23	nC24	nC25	nC26	nC27	nC28	nC29	nC30	nC31	nC32	nC33	nC34	nC35
Boquillas			-123	-125	-121	-119	**-117**	-116	-114	-112	-108	-109	-104	-109	-87	-90	-88	-90	-93				
Marcellus		-161	-160	-157	-155	-156	**-151**	-148	-147	-143	-141	-138	-138										
Niobrara	-152	-150	-148	-144	-142	-141	**-138**	-137	-137	-137	-132	-130	-127	-134	-115	-113	-110						
Morkel	-123	-128	-129	-128	-126	-124	**-124**	-121	-117	-116	-116	-112	-108	-107									
26406	-112	-113	-114	-111	-112	-111	**-110**	-110	-113	-112	-108	-108	-104	-104									
26404		-109	-107	-110	-103	-106	**-109**	-103	-105	-107	-106	-105	-104	-103	-101	-100	-100	-91	-87				
26405		-96	-110	-106	-103	-106	**-110**	-102	-104	-107	-106	-105	-104	-101	-103	-97	-102	-90	-85				
26397	-113	-112	-113	-110	-111	-111	**-108**	-109	-112	-109	-106	-107	-106	-104	-88								
26401				-117	-120	-117	**-116**	-118	-120	-118	-117	-118	-118	-119	-116	-116	-122	-121	-116	-119	-118	-116	
26396		-105	-114	-113	-112	-112	**-109**	-108	-107	-107	-106	-106	-106	-106	-104	-102	-102	-100	-95	-95	-94	-99	-85
26399				-106	-104	-100	**-98**	-96	-92	-94	-91	-94	-92	-91	-96	-89	-85	-95	-89	-87	-79	-73	-81
26400				-108	-101	-100	**-99**	-95	-93	-95	-93	-93	-93	-97	-106	-93	-85	-98	-96	-78			
26403					-92	-89	**-89**	-84	-86	-89	-83	-86	-84	-87	-90	-88	-78	-97	-82				
26402		-72	-79	-77	-75	-75	**-74**	-74	-73	-73	-71	-70	-70	-70	-68	-68	-69	-71	-65	-60	-56	-58	-58
26398			-85	-84	-77	-78	**-78**	-77	-77	-76	-74	-74	-73	-70	-73	-69	-66	-75	-74	-56	-59	-59	
FT22-1&2			-176	-174	-171	-168	**-166**	-167	-167	-164	-160	-152											
FT22-7			-129	-126	-122	-121	**-124**	-122	-124	-120													
FT21		-576	-576	-572	-574	-571	**-568**	-564	-562														
rose							**-240**		-218														

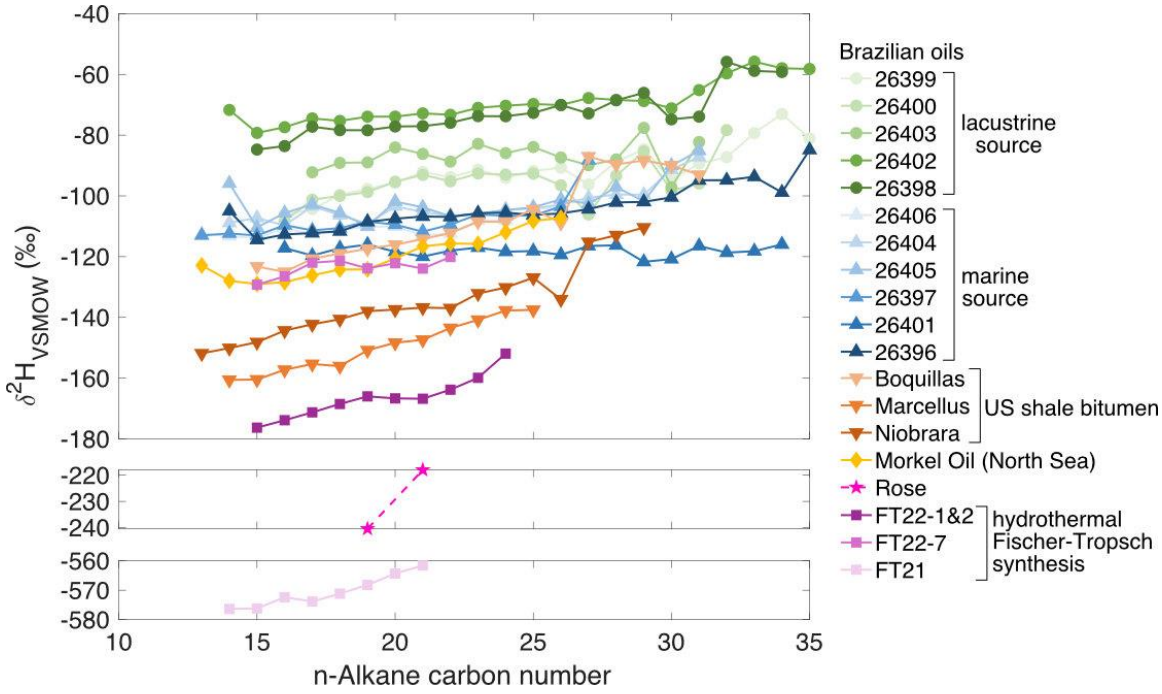


Figure 6.5 Compound-specific hydrogen isotope composition of *n*-alkanes from sedimentary hydrocarbons, Fischer–Tropsch type synthesis and the rose essential oil. The δ^2H values of *n*-alkanes in marine oils and bitumen are generally more depleted than in lacustrine oils, likely reflecting the higher δ^2H of low-latitude evaporative lake and meteoric waters. Bitumen *n*-alkanes are the most depleted in δ^2H among sedimentary samples. The hydrogen isotope profile of lacustrine oils is more variable, potentially due to fluctuations in depositional settings. The δ^2H values of *n*-alkanes from Fischer–Tropsch synthesis at 200°C indicate a fractionation factor of about -500‰ between products and water substrate, suggesting the products are out of isotopic equilibrium with water.

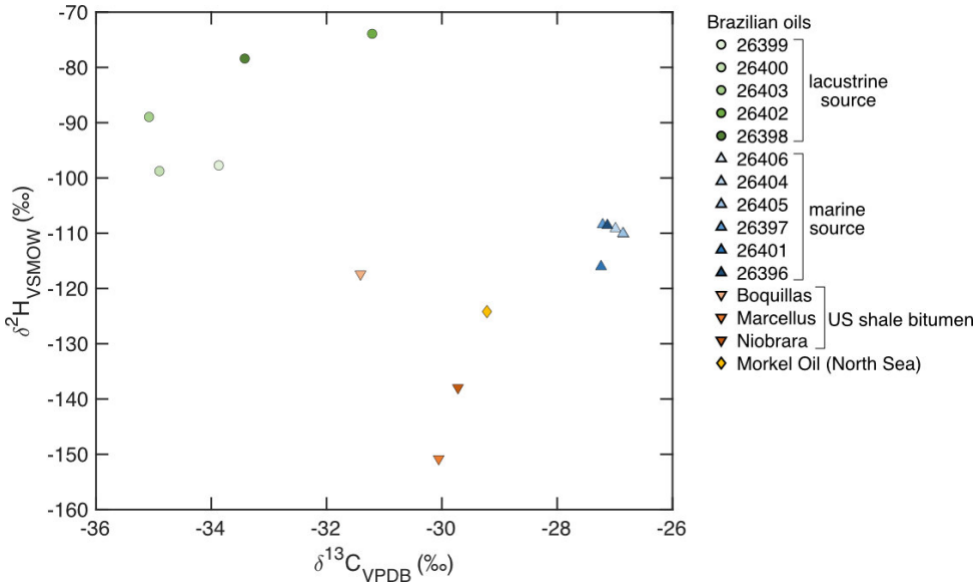


Figure 6.6 Carbon and hydrogen isotope composition of *n*-nonadecane from sedimentary samples. The long-term experimental reproducibility for measurements of standards and replicate samples on these instruments is 0.15‰ for $\delta^{13}C$ and 5‰ for $\delta^{2}H$ (1SE).

3.4. Intramolecular hydrogen and carbon isotope composition of *n*-nonadecane

Normalized carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} fragment ions from *n*-nonadecane are shown in Table 6.6 and Figure 6.7. After offsetting the molecular average isotope fractionation, the normalized fragment-specific isotope composition solely reflects the intramolecular isotope variation. The fragment-specific $\Delta^{13}C$ values of most of the sedimentary, abiotic and biological samples span a range of 2–3‰ (2–3 times long term measurement reproducibility) whereas the three lowest-maturity Brazilian lacustrine-sourced oils have $\Delta^{13}C$ values 1–2‰ more enriched than the majority. The homogeneity suggests that intramolecular carbon isotope variations of n-nonadecane from different sources are indistinguishable with the fragment-specific carbon isotope analysis.

The total range of fragment-specific $\Delta^{2}H$ values of *n*-nonadecane span 40‰ with most of the samples in the range of 0–10‰ and the Brazilian lacustrine-sourced oils solely responsible

for the total 40‰ range. The fragment-specific Δ^2H values of these lacustrine oils are in a negative correlation with their thermal maturity (Figure 6.8). Whereas marine-sourced oils and bitumen here all have average fragment-specific Δ^2H between 0 and 10‰. Fragment-specific Δ^2H of n-nonadecane from rose essential oil and FTT synthesis are mostly in the 0–10‰ range except for FT22-7, which has Δ^2H of C_4H_9 and C_6H_{13} fragments closer to -10‰, one of the lowest values in the total range.

Table 6.6 Carbon and hydrogen isotope composition of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of *n*-nonadecane in sedimentary, abiotic, and biological samples. Normalized fragment-specific carbon and hydrogen isotope composition $\Delta^{13}C$ and Δ^2H are calculated after offsetting the molecular average isotope fractionation between the sample and the CIT-C19 standard as defined in equation (6-7) and (5-6). SE is 1 standard measurement error except for the rose sample, which is standard error of 9 replicate measurements.

	$\Delta^{13}C_{CIT\text{-}C19}$ (‰)				SE (‰)			$\Delta^2H_{CIT\text{-}C19}$ (‰)				SE (‰)		
sample name	C4H9	C5H11	C6H13	average	C4H9	C5H11	C6H13	C4H9	C5H11	C6H13	average	C4H9	C5H11	C6H13
Boquillas	-0.8	0.5	-0.5	-0.3	0.5	0.7	0.8	1	5	8	5	4	4	4
Marcellus	0.6	-0.9	-1.4	-0.6	0.6	0.7	0.7	1	0	6	2	4	4	5
Niobrara	0.5	-0.7	-1.5	-0.6	0.5	0.6	0.7	4	-3	8	3	5	4	4
Morkel	-0.5	0.3	-1.4	-0.5	0.6	0.3	0.3	4	3	5	4	5	4	4
26406	-0.7	0.7	0.3	0.1	0.5	0.5	0.6	-2	5	4	2	4	4	4
26404	-0.2	1.7	-0.7	0.2	0.4	0.6	0.5	5	9	8	7	4	4	4
26405	-0.3	1.1	-0.5	0.1	0.5	0.4	0.4	6	7	5	6	5	5	4
26397	0.0	1.3	-0.2	0.3	0.7	0.5	0.5	2	1	2	2	4	5	4
26401	-0.5	2.0	0.4	0.7	0.6	0.7	0.6	6	11	10	9	5	4	4
26396	0.1	1.4	-0.4	0.3	0.5	0.6	0.6	3	4	4	4	5	4	4
26399	0.7	2.7	2.7	2.0	0.5	0.6	0.7	27	28	32	29	4	4	4
26400	2.8	3.9	1.5	2.7	0.5	0.5	0.5	14	24	21	20	3	4	4
26403	2.5	2.9	3.2	2.9	0.5	0.5	0.5	10	12	15	12	4	4	4
26402	0.0	0.2	1.9	0.7	0.5	0.6	0.4	-1	-3	3	0	4	4	4
26398	0.4	0.2	-1.1	-0.2	0.5	0.4	0.6	-13	-13	-12	-12	4	4	3
FT22-1&2	-0.4	0.3	0.1	0.0	0.7	0.6	0.6	-1	-2	4	0	6	4	5
FT22-7	0.5	1.6	0.1	0.8	0.5	0.7	0.7	-13	2	-4	-5	4	5	4
FT21	2.2	1.4	0.2	1.2	0.6	0.5	0.5							
rose*	-1.2	-0.3	-0.3	-0.6	0.4	0.3	0.4	12	5	11	9	1	2	2

* rose data are from Dong et al., chapter 4. The SEs are standard errors of 9 replicate measurements

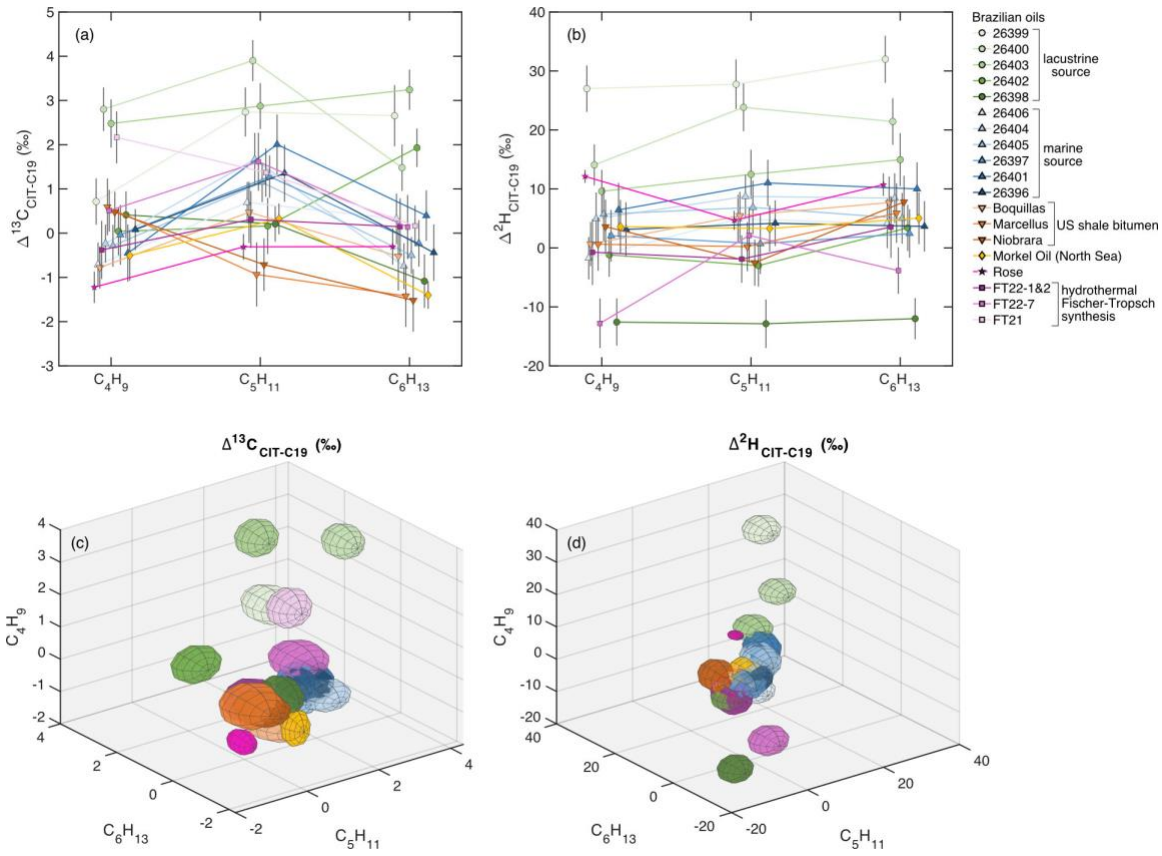


Figure 6.7 Normalized fragment-specific (a, c) $\Delta^{13}C$ and (b, d) $\Delta^{2}H$ of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of *n*-nonadecane in sedimentary, abiotic, and biological samples. After offsetting the molecular average isotope fractionation, the normalized fragment-specific isotope composition solely reflects the intramolecular isotope variation. The fragment-specific $\Delta^{13}C$ values of most of the sedimentary, abiotic, and biological samples span a range of 2–3‰ whereas the three lowest-maturity Brazilian lacustrine-sourced oils have $\Delta^{13}C$ values 1–2‰ more enriched than the majority (a,c). The total range of fragment-specific $\Delta^{2}H$ values of *n*-nonadecane span 40‰ with most of the samples fall in a range of 0-10‰ and the Brazilian lacustrine-sourced oils solely responsible for the total 40‰ range (b,d). The fragment-specific $\Delta^{2}H$ values of these lacustrine oils are in a negative correlation with their thermal maturity. We postulate that the intramolecular hydrogen isotopic equilibrium is at fragment-specific $\Delta^{2}H_{CIT\text{-}C19}$ of ~5‰. The color scheme in (c) and (d) are the same as in (a) and (b). Error bars (a, b) and elliptic radii of ellipsoids (c, d) are 1 standard measurement error. Long term measurement reproducibility is about 1‰ and 5‰ for carbon and hydrogen isotope composition, respectively.

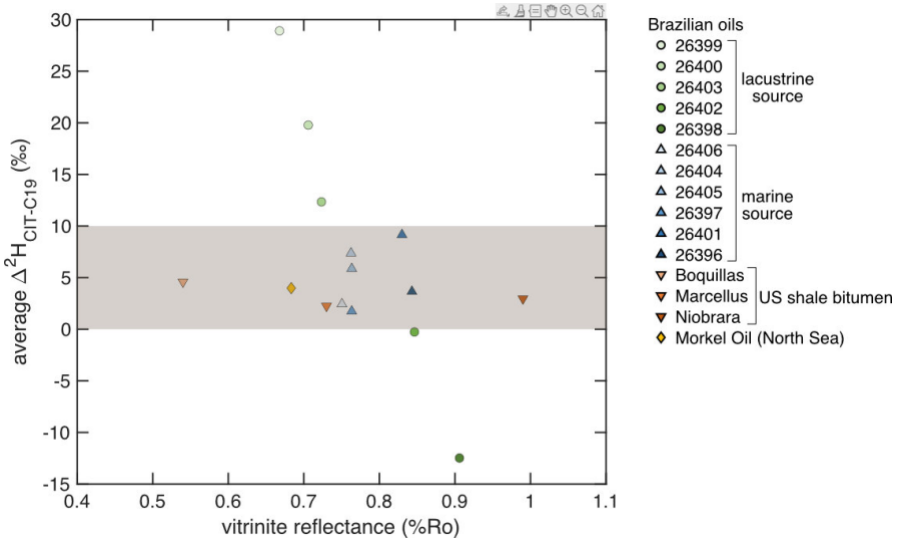


Figure 6.8 Thermal maturity (in vitrinite reflectance, %Ro) versus average normalized fragment-specific hydrogen isotope composition Δ^2H of C_4H_9, C_5H_{11}, and C_6H_{13} fragments of *n*-nonadecane in sedimentary samples. The average fragment-specific Δ^2H values of the Brazilian lacustrine-sourced oils are negatively correlated with thermal maturity whereas marine-sourced oils and bitumen here all have average fragment-specific Δ^2H between 0 and 10‰. We postulate that the intramolecular hydrogen isotopic equilibrium is at fragment-specific $\Delta^2H_{CIT-C19}$ of ~5‰.

3.4.1. Understanding the evolution of intramolecular hydrogen isotope structure of sedimentary *n*-nonadecanes

We analyzed the processes controlling the hydrogen isotope composition of sedimentary *n*-alkanes to understand the associated evolution of intramolecular hydrogen isotope structure.

3.4.1.1. Inheritance from biological lipids and polymers

The biological precursors for *n*-nonadecane are thought to be membrane components such as lipopolysaccharides, phospholipids and sphingolipids produced by bacteria and algae (Brocks and Summons, 2014), polymethylenic biopolymers such as algaenans biosynthesized by microalgae (Tegelaar et al., 1989; Blokker et al., 1998), and cuticular

waxes and aliphatic macromolecules such as cutan and suberan introduced by plant debris (Nip et al., 1986; Tegelaar et al., 1989; Mckinney et al., 1996).

There has been limited hydrogen isotope characterization of these biopolymers. But studies on biological lipids shed some lights on their hydrogen isotope properties. Biosynthesis of lipids by biota imparts a large fractionation, resulting in the ubiquitous D-depletion of organic H relative to environmental water. This molecular average hydrogen isotope fractionation for n-alkyl lipids is typically -100‰ to -150‰ in terrestrial plants and -170‰ in marine phytoplankton (Sessions, 2016). A previous study on intramolecular hydrogen isotope structure of long-chain fatty acids suggests variations as large as 1000‰ with distinct isotope composition for odd versus even numbered positions (Billault et al., 2001).

The fragment-specific Δ^2H values of *n*-nonadecane from rose essential oil indicates 10-15‰ fragment-specific fractionation from the CIT-C19 standard (Dong et al., Chapter 4). Modeling results suggest that the odd-even-position zigzag pattern in the intramolecular isotopic structure can be obscured in the analytical process, resulting in fragment-specific isotope composition indistinguishable from a flat pattern, but large isotope variations at certain molecular positions can be retained in fragment-specific isotope compositions (Dong et al., chapter 4).

3.4.1.2. Kinetic fractionation of hydrocarbon production from kerogen

Carbon-carbon bond cleavage in hydrocarbon production process is thought to impose KIEs at the cleavage and adjacent positions of a few hundred ‰ (Tang et al., 2005; Ni et al., 2011). The terminal and exterior positions in *n*-nonadecane may be products of carbon-carbon bond cleavage. Additional hydrogen needs to be added to the cleavage sites to result in the stoichiometry of final *n*-alkane products and the added capping hydrogen or H_2 for hydrogenation reaction is likely depleted in δ^2H (~ -500 to -800‰) (Dong et al., 2021; Milkov, 2022). Whereas the interior methylene groups may be the residue of preferential hydrogen abstraction from those positions due to the stability of the resulting radical or

charged species. This discrepancy may result in enriched interior and depleted exterior in the hydrogen isotope structure of *n*-nonadecane with a contrast of hundreds of ‰ under geologic conditions (Gray et al., 1971; Tang et al., 2005; Ni et al., 2011). We tested the potential impact of intramolecular isotopic contrasts and found that these upside-down V-shaped intramolecular isotope structures can result in positive fragment-specific Δ^2H values (Table 6.7, Figure 6.9). This suggests that kinetic isotope fractionations of hydrocarbon production could contribute to the positive excursion in fragment-specific Δ^2H of some of the sedimentary samples.

Table 6.7 Position- and fragment-specific hydrogen isotope composition of sedimentary and abiotic *n*-alkanes influenced by different processes. The reference standard has homogeneous position-specific δ^2H_{VSMOW} with an arbitrary value that does not affect the intramolecular contrast of isotope composition.

			δ^2H_{arb} (‰) @ hydrogen position number*																		Δ^2H_{arb} (‰)*			
process			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	C4H9	C5H11	C6H13	average
1	KIE of hydrocarbon production or differential equilibration		-763	-520	-313	-151	-16	83	156	200	227	237	227	200	156	83	-16	-151	-313	-520	-763	24	18	49	30
2			-631	-437	-271	-141	-33	46	104	140	161	169	161	140	104	46	-33	-141	-271	-437	-631	19	14	39	24
3			-433	-312	-208	-127	-59	-10	26	49	62	67	62	49	26	-10	-59	-127	-208	-312	-433	12	9	24	15
4			-470	-50	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	-50	-470	7	5	10	7
5	equilibrium @ T (°C)	60	0	75	61	65	62	67	61	65	62	67	62	65	61	67	62	65	61	75	0	0.6	0.4	0.9	0.6
6		100	0	59	47	50	48	52	47	50	48	52	48	50	47	52	48	50	47	59	0	0.5	0.3	0.7	0.5
7		150	0	43	32	35	33	37	32	35	33	37	33	35	32	37	33	35	32	43	0	0.3	0.2	0.5	0.3
8	FT synthesis (limited substrates or 2 mechanisms)		-500	-499	-499	-497	-496	-491	-490	-484	-475	-467	-455	-437	-417	-390	-353	-307	-251	-182	-100	-6.1	-4.5	-12.4	-7.7

* The reference standard has homogeneous position-specific δ^2H_{VSMOW} with an arbitrary value that does not affect the intramolecular contrast of isotope composition.

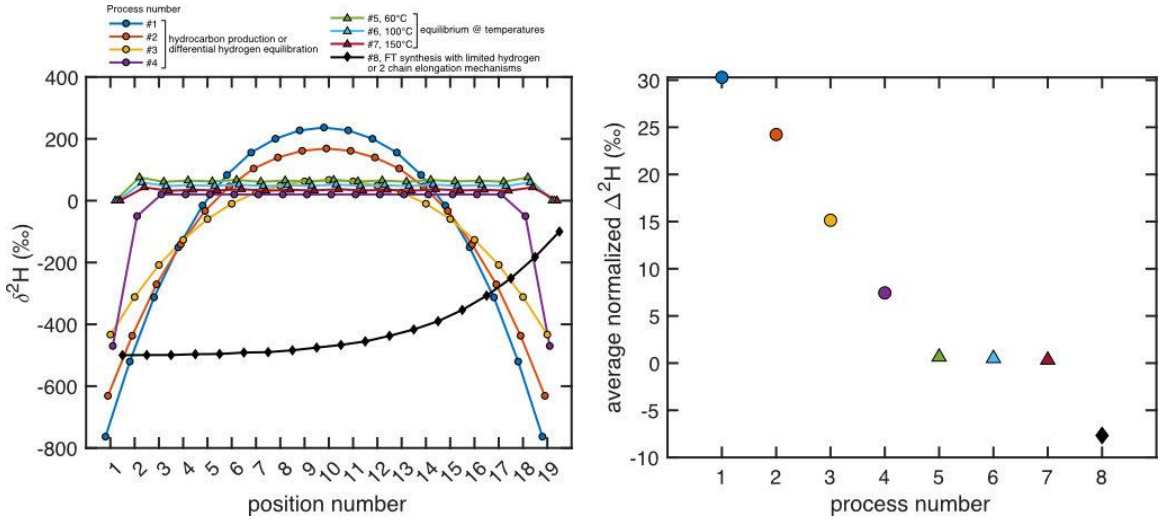


Figure 6.9 (a) Position-specific and (b) average normalized fragment-specific hydrogen isotope composition of *n*-nonadecane influenced by different processes of kinetic fractionation of hydrocarbon production, partial and full equilibration through hydrogen exchange, and Fischer–Tropsch synthesis using a limited reservoir of hydrogen. KIEs from the onset of hydrocarbon production as well as differential equilibration of hydrogen at different positions can result in positive excursion in fragment-specific Δ^2H. At the endpoint of hydrogen equilibration, fragment-specific Δ^2H are brought to thermodynamic equilibrium. Fischer–Tropsch synthesis using limited hydrogen or two chain-elongation mechanisms may lead to a Rayleigh distillation curve within the intramolecular isotope structure (before averaging by molecular symmetry) and negative excursion in fragment-specific Δ^2H. Data in Table 6.7. The reference standard has homogeneous position-specific δ^2H_{VSMOW} with an arbitrary value that does not affect the intramolecular contrast of isotope composition.

3.4.1.3. Equilibration through hydrogen exchange

Hydrogen exchange is an important process that can alter the hydrogen isotope content of organic materials post deposition (Schimmelmann et al., 2006; Sessions, 2016). The endpoint of hydrogen exchange is thermodynamic equilibrium of organic hydrogen with the exchanging reservoir, which could be ancient waters trapped in source rock pores, mineral-bound hydrogen such as that in clays, and/or organic hydrogen itself. The kinetics of hydrogen exchange are controlled by temperature and presence of catalysts such as clay minerals. For *n*-nonadecane or its precursors, the equilibrium fractionation between methyl and methylene groups and water is -70 to -100‰ under geologic conditions and therefore

the δ^2H value is expected to become more enriched during maturation. Different reaction rates of hydrogen exchange for methyl and methylene groups could result in intramolecular discrepancies in distance to isotopic equilibrium of hundreds of ‰, with methyl and exterior methylene groups slower in hydrogen exchange than interior methylene groups (Sessions, 2016). The intramolecular isotopic contrasts resulted from differential isotope equilibration are again upside-down V-shaped structures and can result in positive fragment-specific Δ^2H values (Table 6.7, Figure 6.9). This suggests that partial isotope equilibration could contribute to the positive excursion in fragment-specific Δ^2H of some of the sedimentary samples. We also tested fragment-specific Δ^2H of thermodynamic equilibrium at different temperatures (60, 100, 150°C, Table 6.7, Figure 6.9) and concluded that the whole range for equilibrium fragment-specific Δ^2H under sedimentary and geologic conditions is about 1‰.

3.4.1.4. Thermal and biological degradation

There are a few different degradation mechanisms in the evolution of sedimentary n-alkanes (1) thermal degradation in high maturity regime past-peak oil window, (2) biodegradation possible at any stage of oil generation, and (3) microbial degradation of biological precursors during diagenesis. We have studied the intramolecular isotope effect of thermal degradation of n-nonadecane (Dong et al., chapter 5). Normalized fragment-specific $\Delta^{13}C$ and Δ^2H values of the residual nonadecane show no statistically significant variations with pyrolysis extent after correcting for molecular average isotope fractionations. Although after a ~90% decomposition, fragment-specific Δ^2H values appear to become depleted by ~10‰ relative to the starting material. If such decrease in fragment-specific Δ^2H values is robust, it could explain the negative excursion in fragment-specific Δ^2H of some of the most mature sedimentary samples. The decrease in fragment-specific Δ^2H suggests the exterior part of *n*-nonadecane becomes more enriched in position-specific δ^2H and could be a result of natural gas production and cleavage of short carbon chains from *n*-nonadecane.

Previous studies on microbial degradation and diagenetic effects on n-alkanes have mixed results: either microbial rework can significantly alter the hydrogen isotope composition of

n-alkane (Zech et al., 2011), or δ^2H of *n*-alkanes can be preserved for 15–20 million years in sediments (Yang and Huang, 2003). Little is known about how microbial degradation affects the hydrogen isotope composition of biopolymers, or position-specific hydrogen isotope distribution within the molecules being biodegraded.

3.4.1.5. Summary

The normalized fragment-specific Δ^2H of sedimentary samples can be explained by the processes discussed above, especially regarding the correlation of Δ^2H with thermal maturity for lacustrine oils and the different behavior of marine-sourced oils and bitumen. Positive excursion in fragment-specific Δ^2H can result from the onset of hydrocarbon production and hydrogen exchange. Subsequent equilibration can bring fragment-specific Δ^2H to thermodynamic equilibrium. As maturation progresses, thermal degradation can cause negative excursion in fragment-specific Δ^2H. It is known that isotopic exchange can erase biosynthetic compound-specific hydrogen isotope signal by the start of oil generation in marine shales, whereas limited evidence suggests that such processes are much slower in carbonate rocks from lacustrine sources lacking the catalytic activity of clay minerals (Sessions, 2016). This can explain our observation of the maturation sequence in lacustrine oils whereas the marine oils and bitumen appear to have homogeneous fragment-specific Δ^2H regardless of thermal maturity (Figure 6.8). We postulate that the intramolecular hydrogen isotopic equilibrium is at fragment-specific $\Delta^2H_{CIT\text{-}C19}$ of ~5‰ and this hypothesis can be tested with hydrogen exchange experiments.

There are a few uncertainties in the theory we proposed above.

1. The proposed theory requires intramolecular isotope variations of several hundreds of ‰. The variations are admittedly large, but not unheard of, especially for biological materials (Billault et al., 2001). Additionally, first order quantitative analysis above suggests the discussed processes can explain the magnitude of these variations. If the intramolecular hydrogen isotope variations are indeed as large as

several hundreds of ‰, they present very attractive targets to study especially in biological samples.

2. We do not have enough constraints or knowledge on the initial intramolecular isotope structure of various biological lipids and polymers. The distinct behavior of intramolecular hydrogen isotope composition of *n*-nonadecane from lacustrine versus marine oils might also be attributed to different initial intramolecular isotope structure of biological precursors in lake versus marine environments.

3. To derive fragment-specific hydrogen isotope fractionation from site-specific isotope fractionation predicted for the discussed processes, we used a transfer function based on the fragmentation scheme of *n*-nonadecane derived in Chapter 4. This fragmentation scheme was developed using constraints on fragmentation schemes of *n*-hexane, *n*-heptane, *n*-nonane, *n*-tetradecane and *n*-hexadecane from previous studies (Lavanchy et al., 1978, 1979) and site-specific labeled compounds. Different fragmentation schemes could result in 15–25‰ uncertainty in predicted fragment-specific hydrogen isotope composition (Dong et al., chapter 4).

4. It is unclear why the compound-specific hydrogen isotope composition of *n*-alkanes from solvent extracts of US marine shales is lower than the generally expected δ^2H values for thermodynamic equilibrium for sedimentary *n*-alkanes (~ -100‰, (Sessions, 2016) (Figure 6.5, Figure 6.6). Source rock extracts more depleted in δ^2H than expelled oils have also been observed in previous studies in Brazil (Sessions, 2016). It is possible that *n*-nonadecane in these oil-producing source rocks may be in isotope equilibrium with a hydrogen reservoir that is more depleted than that for general sedimentary *n*-alkanes. Source rocks contain polar compounds (such as resins and asphaltenes) and higher molecular weight hydrocarbons. These compounds are typically associated with less mature organic matter and are preferentially retained in the source rock during the early stages of oil generation. These compounds may constitute an exchanging reservoir that is depleted in δ^2H.

Overall, fragment-specific hydrogen isotope composition presents many potentials for studying paleohydrology, thermal maturation and source of organic matter in sedimentary systems. We develop a framework for its application in the final section.

3.4.2. Interpreting the intramolecular hydrogen isotope structure of *n*-nonadecanes from Fischer–Tropsch-type synthesis

N-nonadecanes from FTT synthesis have fragment-specific hydrogen isotope composition close to 2–3 standard measurement errors from the postulated thermodynamic equilibrium. This result may be expected from the thermodynamic equilibrium between alkane, alkene, alcohol, and carboxylic acid under hydrothermal conditions (Seewald, 1994, 2001). And hydrogen exchange can happen through double bond migration in alkenes.

$\Delta^2H_{CIT\text{-}C19}$ of C_4H_9 and C_6H_{13} fragments of *n*-nonadecane from FT22-7 is -13 and -4‰, a little over 2–3 standard measurement errors. If this negative excursion is robust, it may indicate *n*-nonadecane from FT22-7 has a V-shaped intramolecular hydrogen isotope structure relative to the thermodynamic equilibrium. The V-shaped structure is after consideration of the symmetry of the n-nonadecane and taking the average of two halves of the molecule. Considering the unidirectional chain elongation in FTT synthesis, the site-specific hydrogen isotope composition of n-nonadecane could be a concave up curve (Figure 6.9a). Such structure may result from a reservoir effect from Rayleigh fractionation on hydrogen gas supplying the polymerization, or two different polymerization mechanisms that prevail in short versus long chain elongation (Gaube and Klein, 2008). The more enriched compound-specific δ^2H values of n-alkanes in FT22-7 than FT22-1&2 corroborated the reservoir effect hypothesis. Further mass balance analysis could help test the reservoir effect hypothesis. It is also possible that both reservoir effect and two elongation mechanisms are operating and contributing to the intramolecular hydrogen isotope structure.

From the current survey of samples from different sources, abiotic n-nonadecane has distinguishable fragment-specific hydrogen isotope composition from biological counterpart

but is less distinct from sedimentary source. Although large intramolecular hydrogen isotope variation is expected for biological n-alkanes, the site-specific variations are probably muted due to the dilution and offset in the fragmentation and ionization of n-alkane in the analytical process. More data of abiotic and biological sourced *n*-alkanes are urgently needed to understand the internal variations of these sources.

3.5. A framework for the application of fragment-specific hydrogen isotope analysis of higher n-alkanes

We developed a roadmap for the application of fragment-specific hydrogen isotope analysis (Figure 6.10).

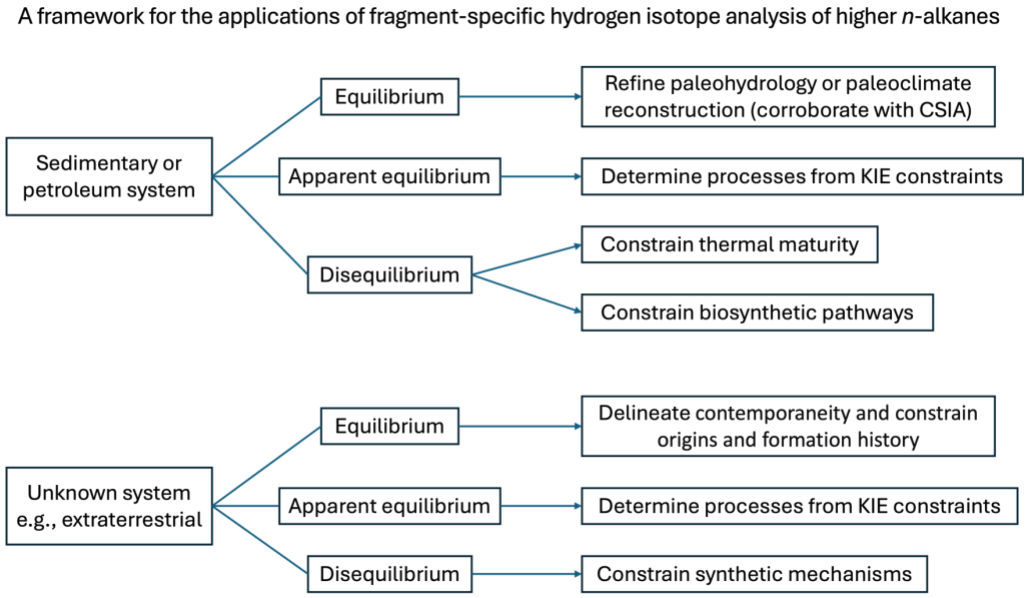


Figure 6.10 A framework for the application of fragment-specific hydrogen isotope analysis of higher n-alkanes

For sedimentary system, equilibrium in fragment-specific isotope composition of higher n-alkanes can be used to refine paleohydrology and paleoclimate reconstruction. The equilibrium range for compound-specific hydrogen isotope composition of n-alkanes can

span 40‰ (Sessions, 2016). As we have seen in the Brazilian lacustrine oils and previous studies, often times the variation of δ^2H of the whole maturation sequence is only 40‰ (Li et al., 2001). Therefore, the determination of equilibrium can be coarse, especially without the knowledge of hydrogen isotope composition of the ancient formation water or hydrogen exchanging reservoir. Intramolecular hydrogen isotope equilibrium of *n*-nonadecane is most likely attained through heterogeneous equilibration in sedimentary systems and inter- and intra-molecular isotopic equilibrium should be achieved at similar stages. There is also little temperature dependence for the equilibrium composition. Therefore, attainment of intramolecular hydrogen isotope equilibrium can facilitate the determination of the hydrogen isotope composition of the exchanging reservoir. This can potentially (1) answer the outstanding question as to which reservoir is relevant for the hydrogen exchange process in sedimentary environments (organic H, mineral H, or formation water) (Sessions, 2016), (2) provide constraints on depositional environments, such as seawater, D-enriched low-latitude strongly evaporative lakes and basins, or D-depleted high-latitude or continental interior environments.

Disequilibrium in fragment-specific isotope composition of higher n-alkanes can be used to constrain thermal maturity or biosynthetic pathways. Like clumped isotopes, this parameter has the advantage of being internal to n-alkanes and without knowledge of isotopic content of co-existing compounds and substrates. Additional study and calibration are needed to develop the fragment-specific hydrogen isotope signature into thermal maturity and/or source parameter for oil-oil and oil-source rock correlation, e.g., hydrous pyrolysis experiments on kerogen and sources rocks, measurements of petroleum samples from terrestrial sources. An area of development with large potentials may be on light hydrocarbons in light oils and condensates where biomarker-based maturity parameters become saturated or absent.

Also similar to clumped isotopes, fragment-specific hydrogen isotope composition can appear to be in equilibrium but as a result of kinetic processes, as for the case of pyrolysis of

n-nonadecane and potentially also *n*-nonadecane from rose essential oils. In that case, we can derive constraints on the associated KIEs to elucidate the kinetic processes.

For *n*-alkanes from unknown systems, e.g., extraterrestrial environments, equilibrium intramolecular hydrogen isotope composition can help identify coeval compounds, substrates, and hydrogen reservoirs (water, mineral, insoluble organic matter, etc.) and constrain their origins and formation history (same generation or late addition, hydrological history on extraterrestrial bodies). Intramolecular hydrogen isotope structure can be derived from disequilibrium in fragment-specific hydrogen isotope composition and used to constrain synthetic mechanisms. Whereas apparent equilibrium can again help constrain the kinetic processes involved.

4. Conclusion and future implications

In this study, we conducted a survey of intramolecular carbon and hydrogen isotope compositions of *n*-nonadecane from a range of sedimentary samples, abiotic synthesis, and a biological source. We found that Fragment-specific hydrogen isotope compositions, reflecting intramolecular isotopic structures, vary widely in lacustrine oils in correlation with thermal maturity, but are homogeneous in marine samples. We propose a framework where positive excursions in fragment-specific Δ^2H arise from KIEs during oil generation and partial hydrogen exchange, negative excursions indicate high maturity thermal degradation, and consistent equilibrium values around +5‰ are reached with clay-catalyzed hydrogen exchange. *n*-Alkanes from Fischer–Tropsch synthesis have intramolecular hydrogen isotope compositions close to inferred equilibrium values and are currently distinguishable from biological counterparts in their fragment-specific hydrogen isotope compositions but overlap with sedimentary *n*-alkanes.

These findings demonstrate the potential of position-specific isotope compositions for tracing the origins and history of sedimentary organic matter. Further promising applications of this novel analytical approach include:

1. Analysis of n-alkanes and other aliphatic compounds in carbonaceous chondrite meteorites to evaluate their origin, formation mechanisms and potential for recording presolar or parent body processes. The high sensitivity of the Orbitrap mass spectrometer enables analysis of small extraterrestrial samples (Zeichner et al., 2023).

2. Investigation of the intramolecular isotopic structures of aliphatic alcohols and acids from hydrothermal Fischer–Tropsch synthesis, which have been proposed to play key roles as primordial cell membrane constituents. These compounds are also abundant in some meteorites and may retain isotopic signatures of their formation and early molecular evolution.

3. Employing high-field NMR and industrial scale Fischer–Tropsch synthesis to directly map the site-specific carbon and hydrogen isotope distributions within *n*-alkanes and/or fatty acids from biological and abiotic sources. This will build a more complete understanding of the underlying isotope fractionation mechanisms and may also have implications for forensics of biofuel, clean diesel and aviation fuel.

The continued development and application of intramolecular isotopic analysis of hydrocarbons promises to open many new research directions in organic geochemistry, paleoclimatology, astrobiology, and biogeochemistry.

Acknowledgements

This work was supported by the Caltech Joint Industry Partnership for Petroleum Geochemistry (Petrobras, Shell, ExxonMobil), the Simons Foundation, the Department of Energy BES program, and AkerBP. We thank Nami Kitchen, Michael Mathuri, and Fenfang Wu for laboratory operation support, Shae Silverman and Elliott Mueller for assistance with GC-IRMS analysis, and Hannah Dion-Kirschner, Gabriella Weiss, and Alex Phillips for SPE procedures. We are grateful for helpful discussions with Kate Freeman, Yanhua Shuai, Nithya Thiagarajan, Julia van Winden, and Johan Weijers. We thank Sarah Zeichner, Tim

Csernica, Elise Wilkes, Peter Martin, and Max Lloyd for their contributions to the communal Orbitrap data analysis code. During the preparation of this work, the author used Claude 3 Opus (large language model developed by Anthropic) in order to improve readability and language. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Bibliography

Billault I., Guiet S., Mabon F. and Robins R. (2001) Natural Deuterium Distribution in Long-Chain Fatty Acids Is Nonstatistical: A Site-Specific Study by Quantitative2H NMR Spectroscopy. *ChemBioChem* **2**, 425–431.

Birdwell J. E., Jubb A. M., Hackley P. C. and Hatcherian J. J. (2021) Compositional evolution of organic matter in Boquillas Shale across a thermal gradient at the single particle level. *International Journal of Coal Geology* **248**, 103859.

Birdwell J. E. and Wilson S. A. (2019) Variability in Results from Mineralogical and Organic Geochemical Interlaboratory Testing of U. S. Geological Survey Shale Reference Materials. In *Proceedings of the 7th Unconventional Resources Technology Conference* Unconventional Resources Technology Conference. American Association of Petroleum Geologists, Denver, Colorado, USA.

Bjorøy M., Hall K., Gillyon P. and Jumeau J. (1991) Carbon isotope variations in n-alkanes and isoprenoids of whole oils. *Chemical Geology* **93**, 13–20.

Blokker P., Schouten S., Van Den Ende H., De Leeuw J. W., Hatcher P. G. and Sinninghe Damsté J. S. (1998) Chemical structure of algaenans from the fresh water algae Tetraedron minimum, Scenedesmus communis and Pediastrum boryanum. *Organic Geochemistry* **29**, 1453–1468.

Brocks J. J. and Summons R. E. (2014) Sedimentary Hydrocarbons, Biomarkers for Early Life. In *Treatise on Geochemistry* Elsevier. pp. 61–103.

Collister J. W., Rieley G., Stern B., Eglinton G. and Fry B. (1994) Compound-specific δ 13C analyses of leaf lipids from plants with differing carbon dioxide metabolisms. *Organic Geochemistry* **21**, 619–627.

Coplen T. B. (1988) Normalization of oxygen and hydrogen isotope data. *Chemical Geology: Isotope Geoscience section* **72**, 293–297.

Dion-Kirschner H., McFarlin J. M., Masterson A. L., Axford Y. and Osburn M. R. (2020) Modern constraints on the sources and climate signals recorded by sedimentary plant waxes in west Greenland. *Geochimica et Cosmochimica Acta* **286**, 336–354.

Dong G., Xie H., Formolo M., Lawson M., Sessions A. and Eiler J. (2021) Clumped isotope effects of thermogenic methane formation: Insights from pyrolysis of hydrocarbons. *Geochimica et Cosmochimica Acta* **303**, 159–183.

Donovan A. D. (2016) Making outcrops relevant for an unconventional source rock play: an example from the Eagle Ford Group of Texas. *SP* **436**, 193–206.

El Attar A. and Pranter M. J. (2016) Regional stratigraphy, elemental chemostratigraphy, and organic richness of the Niobrara Member of the Mancos Shale, Piceance Basin, Colorado. *Bulletin* **100**, 345–377.

Gaube J. and Klein H.-F. (2008) Studies on the reaction mechanism of the Fischer–Tropsch synthesis on iron and cobalt. *Journal of Molecular Catalysis A: Chemical* **283**, 60–68.

Gilbert A., Yamada K. and Yoshida N. (2013) Exploration of intramolecular 13C isotope distribution in long chain n-alkanes (C11–C31) using isotopic 13C NMR. *Organic Geochemistry* **62**, 56–61.

Gray P., Herod A. A. and Jones A. (1971) Kinetic data for hydrogen and deuterium atom abstraction by methyl and trifluoromethyl radicals in the gaseous phase. *Chemical Reviews* **71**, 247–294.

Guthrie J. M., Trindade L. A. F., Eckardt C. B. and Takaki T. (1996) Molecular and carbon isotopic analysis of specific biological markers: Evidence for distinguishing between marine and lacustrine depositional environments in sedimentary basins of Brazil. In *AAPG Bulletin* Annual convention of the American Association of Petroleum Geologists, Inc. and the Society for Sedimentary Geology: global exploration and geotechnology. San Diego, CA (United States).

Horstad I., Larter S. R. and Mills N. (1995) Migration of hydrocarbons in the Tampen Spur area, Norwegian North Sea: a reservoir geochemical evaluation. *SP* **86**, 159–183.

Julien M., Nun P., Höhener P., Parinet J., Robins R. J. and Remaud G. S. (2016) Enhanced forensic discrimination of pollutants by position-specific isotope analysis using isotope ratio monitoring by 13C nuclear magnetic resonance spectrometry. *Talanta* **147**, 383–389.

Julien M., Zhao Y., Ma R., Zhou Y., Nakagawa M., Yamada K., Yoshida N., Remaud G. S. and Gilbert A. (2022) Re-evaluation of the 13C isotope fractionation associated

with lipids biosynthesis by position-specific isotope analysis of plant fatty acids. *Organic Geochemistry* **174**, 104516.

Lavanchy A., Houriet R. and Gäumann T. (1978) The mass spectrometric fragmentation of *n* -heptane. *Org. Mass Spectrom.* **13**, 410–416.

Lavanchy A., Houriet R. and Gäumann T. (1979) The mass spectrometric fragmentation ofn-alkanes. *Org. Mass Spectrom.* **14**, 79–85.

Li M., Huang Y., Obermajer M., Jiang C., Snowdon L. R. and Fowler M. G. (2001) Hydrogen isotopic compositions of individual alkanes as a new approach to petroleum correlation: case studies from the Western Canada Sedimentary Basin. *Organic Geochemistry* **32**, 1387–1399.

Lock B. E. and Peschier L. (2006) Boquillas (Eagle Ford) upper slope sediments, west Texas: Outcrop analogs for potential shale reservoirs. *Gulf Coast Assoc. Geol. Soc. Trans.* **56**, 491–508.

Lock B. E., Peschier L. and Whitcomb N. (2010) The Eagle Ford (Boquillas Formation) of Val Verde County, Texas— A Window on the South Texas Play. *Gulf Coast Assoc. Geol. Soc. Trans.* **60**, 419–434.

McCollom T. M., Lollar B. S., Lacrampe-Couloume G. and Seewald J. S. (2010) The influence of carbon source on abiotic organic synthesis and carbon isotope fractionation under hydrothermal conditions. *Geochimica et Cosmochimica Acta* **74**, 2717–2740.

Mccollom T. M., Ritter G. and Simoneit B. R. T. (1999) Lipid Synthesis Under Hydrothermal Conditions by Fischer- Tropsch-Type Reactions. *Origins of Life and Evolution of the Biosphere* **29**, 153–166.

McCollom T. and Seewald J. (2006) Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions. *Earth and Planetary Science Letters* **243**, 74–84.

Mckinney D. E., Bortiatynski J. M., Carson D. M., Clifford D. J., De Leeuw J. W. and Hatcher P. G. (1996) Tetramethylammonium hydroxide (TMAH) thermochemolysis of the aliphatic biopolymer cutan: insights into the chemical structure. *Organic Geochemistry* **24**, 641–650.

Mead R., Xu Y., Chong J. and Jaffé R. (2005) Sediment and soil organic matter source assessment as revealed by the molecular distribution and carbon isotopic composition of n-alkanes. *Organic Geochemistry* **36**, 363–370.

Milkov A. V. (2022) Molecular hydrogen in surface and subsurface natural gases: Abundance, origins and ideas for deliberate exploration. *Earth-Science Reviews* **230**, 104063.

Monson K. D. and 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. *J Biol Chem* **255**, 11435–11441.

Monson K. D. and Hayes J. M. (1982) Biosynthetic control of the natural abundance of carbon 13 at specific positions within fatty acids in Saccharomyces cerevisiae. Isotopic fractionation in lipid synthesis as evidence for peroxisomal regulation. *J Biol Chem* **257**, 5568–5575.

Ni Y., Ma Q., Ellis G. S., Dai J., Katz B., Zhang S. and Tang Y. (2011) Fundamental studies on kinetic isotope effect (KIE) of hydrogen isotope fractionation in natural gas systems. *Geochimica et Cosmochimica Acta* **75**, 2696–2707.

Nip M., Tegelaar E. W., Brinkhuis H., De Leeuw J. W., Schenck P. A. and Holloway P. J. (1986) Analysis of modern and fossil plant cuticles by Curie point Py-GC and Curie point Py-GC-MS: Recognition of a new, highly aliphatic and resistant biopolymer. *Organic Geochemistry* **10**, 769–778.

Norwegian Offshore Directorate (2024a) 33/2-2 S.

Norwegian Offshore Directorate (2024b) DRAUPNE FM.

Norwegian Offshore Directorate (2024c) HEATHER FM.

Norwegian Offshore Directorate (2024d) LUNDE FM.

Pedentchouk N. and Turich C. (2018) Carbon and hydrogen isotopic compositions of *n* -alkanes as a tool in petroleum exploration. *SP* **468**, 105–125.

Peters K. E., Walters C. C. and Moldowan J. M. (2004) *The Biomarker Guide*. 2nd ed., Cambridge University Press.

Sachse D., Billault I., Bowen G. J., Chikaraishi Y., Dawson T. E., Feakins S. J., Freeman K. H., Magill C. R., McInerney F. A., Van Der Meer M. T. J., Polissar P., Robins R. J., Sachs J. P., Schmidt H.-L., Sessions A. L., White J. W. C., West J. B. and Kahmen A. (2012) Molecular Paleohydrology: Interpreting the Hydrogen-Isotopic Composition of Lipid Biomarkers from Photosynthesizing Organisms. *Annu. Rev. Earth Planet. Sci.* **40**, 221–249.

Schimmelmann A., Sessions A. L. and Mastalerz M. (2006) HYDROGEN ISOTOPIC (D/H) COMPOSITION OF ORGANIC MATTER DURING DIAGENESIS AND THERMAL MATURATION. *Annual Review of Earth and Planetary Sciences* **34**, 501–533.

Seewald J. S. (2001) Aqueous geochemistry of low molecular weight hydrocarbons at elevated temperatures and pressures: constraints from mineral buffered laboratory experiments. *Geochimica et Cosmochimica Acta* **65**, 1641–1664.

Seewald J. S. (1994) Evidence for metastable equilibrium between hydrocarbons under hydrothermal conditions. *Nature* **370**, 285–287.

Seki O., Nakatsuka T., Shibata H. and Kawamura K. (2010) A compound-specific n-alkane δ13C and δD approach for assessing source and delivery processes of terrestrial organic matter within a forested watershed in northern Japan. *Geochimica et Cosmochimica Acta* **74**, 599–613.

Sessions A. L. (2016) Factors controlling the deuterium contents of sedimentary hydrocarbons. *Organic Geochemistry* **96**, 43–64.

Sessions A. L. (2006) Seasonal changes in D/H fractionation accompanying lipid biosynthesis in Spartina alterniflora. *Geochimica et Cosmochimica Acta* **70**, 2153–2162.

Soeder D. J., Enomoto C. B. and Chermak J. A. (2014) The Devonian Marcellus Shale and Millboro Shale. In *Elevating Geoscience in the Southeastern United States: New Ideas about Old Terranes—Field Guides for the GSA Southeastern Section Meeting, Blacksburg, Virginia, 2014* Geological Society of America. pp. 129–160.

Tang Y., Huang Y., Ellis G. S., Wang Y., Kralert P. G., Gillaizeau B., Ma Q. and Hwang R. (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.

Tegelaar E. W., De Leeuw J. W., Derenne S. and Largeau C. (1989) A reappraisal of kerogen formation. *Geochimica et Cosmochimica Acta* **53**, 3103–3106.

Wang Y., Sessions A. L., Nielsen R. J. and Goddard W. A. (2009) Equilibrium 2H/1H fractionations in organic molecules. II: Linear alkanes, alkenes, ketones, carboxylic acids, esters, alcohols and ethers. *Geochimica et Cosmochimica Acta* **73**, 7076–7086.

Yang H. and Huang Y. (2003) Preservation of lipid hydrogen isotope ratios in Miocene lacustrine sediments and plant fossils at Clarkia, northern Idaho, USA. *Organic Geochemistry* **34**, 413–423.

Zech M., Pedentchouk N., Buggle B., Leiber K., Kalbitz K., Marković S. B. and Glaser B. (2011) Effect of leaf litter degradation and seasonality on D/H isotope ratios of n-alkane biomarkers. *Geochimica et Cosmochimica Acta* **75**, 4917–4928.

Zeichner S. S., Aponte J. C., Bhattacharjee S., Dong G., Hofmann A. E., Dworkin J. P., Glavin D. P., Elsila J. E., Graham H. V., Naraoka H., Takano Y., Tachibana S., Karp A. T., Grice K., Holman A. I., Freeman K. H., Yurimoto H., Nakamura T., Noguchi T., Okazaki R., Yabuta H., Sakamoto K., Yada T., Nishimura M., Nakato A., Miyazaki A., Yogata K., Abe M., Okada T., Usui T., Yoshikawa M., Saiki T., Tanaka Satoshi, Terui F., Nakazawa S., Watanabe S., Tsuda Y., Hamase K., Fukushima K., Aoki D., Hashiguchi M., Mita H., Chikaraishi Y., Ohkouchi N., Ogawa N. O., Sakai S., Parker E. T., McLain H. L., Orthous-Daunay F.-R., Vuitton V., Wolters C., Schmitt-Kopplin P., Hertkorn N., Thissen R., Ruf A., Isa J., Oba Y., Koga T., Yoshimura T., Araoka D., Sugahara H., Furusho A., Furukawa Y., Aoki J., Kano K., Nomura S. M., Sasaki K., Sato H., Yoshikawa T., Tanaka Satoru, Morita M., Onose M., Kabashima F., Fujishima K., Yamazaki T., Kimura Y. and Eiler J. M. (2023) Polycyclic aromatic hydrocarbons in samples of Ryugu formed in the interstellar medium. *Science* **382**, 1411–1416.