Prebiotic Fingerprints

Thesis by

Laura Marie Chimiak

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2021

(Defended 19 Nov 2020)

Laura Marie Chimiak
ORCID: 0000-0002-5047-5421

ACKNOWLEDGEMENTS

This thesis could not have been completed without the effort and assistance of too many people to mention here. First and foremost, I would like to thank John Eiler, who has served as a mentor for the past six years. He has both supported my scientific ambition and challenged me when embarking upon those ambitions. The many hours of intellectual debate are primary drivers for the work I present here, especially the final chapter.

I must also thank the communities at Caltech, Princeton, and Amherst College. Each provided years of illuminating conversations and friendships. Firstly, thank you to Mark Marshall, Helen Leung, and Anthony Bishop, who taught me what I know about physical and organic chemistry and extended my undergrad tenure by a semester so that I add a chemistry degree to my literature one. I don't think I would be here if it weren't for that detour and the incredible warmth and love that the Amherst College chemistry department has. Secondly, I would like to acknowledge the help of Michael Bender, who taught me how to think like a scientist and took hours out of his days to give me a crash course in stable isotope biogeochemistry, to discuss interstellar chemistry, and to discuss bluegrass at length. Thank you for creating such an intellectually fulfilling environment for the years at Princeton and for keeping an interest in my work since.

For those at Caltech, I would like to thank the rest of my committee without whom I would not have made it this far (and without whom I could not have had such a wide-range of angles in my thesis). Alex Sessions has gone beyond academic advising and mentorship these past six years and taught me to always consider side reactions when looking at isotopic data and provided guidance and the lab space required to do all of the Strecker chemistry. George Rossman spent his free time to humor my interest in the intersection of chemistry and minerology and was somehow always available to discuss a stray science question or mountaineering recommendation. Geoff Blake has provided insight and clarity into potential reaction pathways in the ISM and was always a purveyor of good insight into jobs and fellowships. Finally, Don Burnett helped me understand the importance of the petrology of a meteorite sample and appreciate nuclear chemistry. I will note here that a major reason I applied to Caltech was due to reading Don's papers on the Genesis mission.

There is also no way that my thesis would be complete without the NASA Goddard Space Flight Center astrobiology group. Jamie, Jason, and Jose (I am noticing a trend here with names) were willing to meet with me when I proposed a collaboration during my second year. From that grew a great many conversations, sample exchanges, and hopefully many more years of joint pursuits.

The astrobio community, specifically the AbGradCon and Gordon Conference Origins of Life communities, have provided close friendship and scientific discussion over the years. I am grateful to them and hope to carry on the conversations as new samples and chemistries emerge. In these communities, I must thank Beck Rapf—she is a rising star

and co-chaired (by which I mean taught me how to co-chair) the Gordon Research Seminar and has been a driving force behind many of the younger astrobiology meetings. I met her at the first conference that I attended and appreciated the time that she spent both discussing interstellar reactions and exploring Paris with me.

Both the Sessions and Eiler lab groups provided constant sources of support and insight. Specifically, Brooke Dallas co-designed the Orbitrap Site-Specific Isotope Ratio measurements with me and was immensely gifted at understanding how the instrument could impact measurements. Also, the other site-specific isotope ratio group members—Amy Hofmann, Eilse Wilkes, Jaime Cesar, and Sara Zeichner shared the trials and tribulations of optimizing the electron impact source measurements and Caj Neubauer, Gabriella Weiss, Tim Csernica and Elliot Mueller continue to impress me pushing the possibilities of the electrospray source Orbitrap.

In the many hours outside the lab, I had the good fortune to be involved in campus groups. My time at Caltech would not be as profitable had I not had the ability to exercise my civic desires to make things better for others in the Title IX committee (thanks to Felicia Hunt—you deserve more than one mention but I have typed a short story as it is) and the student government. I would like to thank both the Princeton Mountaineering Club and the Caltech Alpine Club. In both groups, I was able to challenge myself physically while discussing research with some of the best, brightest, and—most importantly—kindest people I can imagine. Through these groups, I met Carl Blumenfeld, who volunteered his time during his PhD to help me perfect the Strecker set up, and Ilana Golbin, who shared coffees and conversations and has been a true friend. Thank you to Joel Scheingross, who taught me how to scramble around rocks, took me to some of the most beautiful parts of the Sierras, and taught me that geomorphologists are not the correct people to ask about minerology while you hike. Also, thank you to Jono Squire who put up with me as a housemate and climbing friend for many years. I always appreciated your enthusiasm to discuss nucleosynthesis at 8 am or to push me to get over my head game both in presentations and in sports. In this section I cannot fail to mention Sujung Lim, one of the most fearless and kindhearted people I know. Thank you for being a climbing partner, adventure buddy, and all-around wonderful human being.

Finally, I would like to thank family without whom I would literally not be here. Thanks to my mother and my grandparents for providing unending love and support and always asking questions and putting up with my attempts at explaining my work in French. And thanks to my family here in LA—Greg MacCabe, Woody, and Poochio. Greg, thank you for being a supporting partner, especially when we were both working late nights in lab—you set an impossible precedent for commitment to lab work and were still always willing to join me if I was working the graveyard shift—may we have many adventures to come. And finally thanks to Woody and Poochio, who licked my face after long days and paper rejections to remind me that who one surrounds oneself with is what is most important in life. I suppose in the end, if this were a dedication, it would have to be to all the good dogs out there who keep us grounded when a pandemic is afoot and ask for precious little in return.

ABSTRACT

Meteorites contain organic compounds that occur in all known life. These compounds, commonly referred to as prebiotic compounds, include α-amino acids and are most prevalent on carbonaceous chondrites. As carbonaceous chondrites are pristine samples from early in the solar system that have not had living organisms on them, we can study the chemistry that produced α-amino acids on them to better understand the processes by which they might have formed on early Earth or on other bodies. Multiple syntheses have been put forth as routes to form amino acids on meteorites and include ice-grain chemistry on interstellar ices and Strecker synthesis in meteorite parent bodies. Prior measurements of molecular-average carbon isotope ratios ($^{13}C/^{12}C$) have found ^{13}C enrichments of up to 53‰ in certain α-amino acids and molecular-average hydrogen isotope ratios (D/H) have found enrichments of 100s of ‰. With this data, it has been suggested that Strecker synthesis—a synthesis in which an aldehyde or ketone reacts with ammonia and cyanide to produce an α-aminonitrile that is hydrolyzed into an α-amino amide and then an α-amino acid—is the primary pathway to produce α-amino acids on aqueously altered meteorites.

Here, we develop an instrument that can measure site-specific isotope ratios (SSIR) for carbon—that is the $^{12}C/^{13}C$ at each site in a molecule—and use it to first constrain the site-specific isotope effects associated with Strecker synthesis and then the carbon SSIR of an alanine sample extracted from the Murchison meteorite. The instrument, the Q-Exactive Orbitrap, is a Fourier Transform Mass Spectrometer that has resolution of 240,000 full width-half maximum and can measure site-specific carbon isotope ratios on samples as small as 1 picomole. When we use it to measure the carbon SSIR in multiple samples of alanine produced by Strecker synthesis, we find a -20 ‰ equilibrium isotope effect between the product alanine's C-2 site (amine carbon, ^{13}C-depleted) reactant acetaldehyde's carbonyl carbon (^{13}C-enriched), a potential -15 ‰ kinetic isotope effect on the C-1 site (eventual carboxyl carbon) for the first hydrolysis of α-aminopropanenitrile (^{13}C-enriched) into alaninamide (^{13}C-depleted), and a -15.4 ‰ kinetic isotope effect on the C-1 carbon for the second hydrolysis step in which α-

alaninamide (^{13}C-enriched) becomes alanine (^{13}C-depleted). Through conventional isotope ratio mass spectrometry, we also measure a +56.4 ‰ equilibrium isotope effect between ammonia (^{15}N-depleted) and the amine site on alanine (^{15}N-enriched). When we measure the sample of alanine from the Murchison meteorite, we find site-specific carbon isotope ratios of -29 ± 10 ‰, 142 ± 20 ‰, and -36 ± 20 ‰ for the C-1, C-2, and C-3 (methyl) sites, respectively. This pattern agrees with the hypothesis that Strecker synthesis created alanine in Murchison. Combining these data with the isotope effects found for Strecker synthesis, we find initial site values of -7 ± 10 ‰, 162 ± 20 ‰, and -36 ± 20 ‰ for the C-1, C-2, and C-3 sites, respectively. With these values, we create a model of potential organic synthesis on the Murchison parent body that predicts the molecular-average δ^{13}C values of 19 other prebiotic compounds.

Finally, we create a model that uses the previously measured molecular average carbon and deuterium isotope ratios for organics on Murchison to create models that predict site-specific and molecular average isotope ratios for organic compounds. This model finds that organic compounds with have methyl sites that are enriched in deuterium by up to 3000 ‰ relative to other sites in the compound and that the degree of enrichment scales both with a compound class's solubility in water and with a sample's degree of aqueous alteration and terrestrial weathering. These patterns suggest that a primordial ISM-derived deuterium signal exchanges with water and that the methyl site hosts the highest amount of this enrichment due to its low acidity. The carbon model demonstrates that using only the aldehyde and cyanide values measured on Murchison and isotope effects inferred from other studies, we can predict 59 of 82 organic compounds on it (72%) that have δ^{13}C values spanning over 149 ‰ with an average residual of 6 ‰. To achieve this level of prediction, the model combines Strecker synthesis, reductive amination, and oxidation of aldehydes to create straight-chain α-H hydroxy and amino acids, amines, and monocarboxylic acids with subsequent formaldehyde addition to these compounds to create branches.

PUBLISHED CONTENT AND CONTRIBUTIONS

Spaulding N. E., Higgins J. A., Kurbatov A. V., Bender M. L., Arcone S. A., Campbell S., Dunbar N. W., Chimiak L. M., Introne D. S. and Mayewski P. A. (2013) Climate archives from 90 to 250ka in horizontal and vertical ice cores from the Allan Hills Blue Ice Area, Antarctica. *Quaternary Research* **80**, 562–574. doi:10.1016/j.yqres.2013.07.004

L.M.C. measured the $\delta^{18}O$, $\delta^{15}N$, $\delta^{40}Ar$ and O_2/N_2 ratios in ice cores used in this study.

Higgins J. A., Kurbatov A. V., Spaulding N. E., Brook E., Introne D. S., Chimiak L. M., Yan Y., Mayewski P. A. and Bender M. L. (2015) Atmospheric composition 1 million years ago from blue ice in the Allan Hills, Antarctica. *Proc. Natl. Acad. Sci. USA* **112**, 6887–6891. doi: 10.1073/pnas.1420232112

L.M.C. measured the $\delta^{18}O$, $\delta^{15}N$, $\delta^{40}Ar$ and O_2/N_2 ratios in ice cores used in this study.

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. *Org. Geochem.* doi:10.1016/j.orggeochem.2017.07.016

L.M.C. helped compile and synthesize $\delta^{13}C$ and δD data used in this paper.

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. *Int. J. Mass Spectrom.* **422**, 126–142. doi:10.1016/j.ijms.2017.10.002

L.M.C. aided in developing the Oribitrap Mass Spectrometry method used in this paper. Specificallty L.MC. performed offline chemical cleavage and isotopic analysis of standards for verification of the instrument's accuracy, provided site-specific measurements of alanine data, and helped analyze factors that contribute to inaccuracies.

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. *Org. Geochem.* doi:10.1016/j.orggeochem.2019.06.002

L.M.C. aided in developing the Oribitrap Mass Spectrometry method used in this paper.

Hofmann A. E., Chimiak L., Dallas B., Griep-Raming J., Juchelka D., Makarov A., Schwieters J. and Eiler J. M. (2020) Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures. *Int. J. Mass Spectrom.* **457**, 116410.
L.M.C. provided isotopic measurements of alaine and alaninamide used in this study. doi:10.1016/j.ijms.2020.116410

Chimiak L., Elsila J. E., Dallas B., Dworkin J. P., Aponte J. C., Sessions A. L. and Eiler J. M.

(2020) Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system. *Geochim. Cosmochim. Acta*. doi: 10.1016/j.gca.2020.09.026

L.M.C. helped conceive this project, helped develop the methodology used in this project, conducted all syntheses, collected site-specific isotope measurements of meteoritic alanine, interpreted results found in this project, and wrote the manuscript.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION: CHEMICAL UNDERPINNINGS OF THE ORIGINS OF LIFE

Certain compounds such as amino acids and lipids are present in all known life. It is reasonable, therefore, to hypothesize that these or similar compounds were present when life emerged, that is prebiotic compounds likely aided in the formation of life. Many if not all of these compounds can be formed through multiple abiotic syntheses. Since specific syntheses occur in specific physical and chemical conditions (aptly called 'physiochemical conditions') such as temperature, pressure, and pH and use specific reactants, by constraining the mechanism that create these compounds in abiotic processes in natural settings—outside of the lab and without biological intervention—we can better understand the conditions likely required for the emergence of life both on Earth and on other bodies.

Earth is teeming with biology, so it is difficult to find samples of these compounds that are not biologically contaminated to study. However, meteorites—pieces of asteroids and comets that have landed on earth—also contain prebiotic compounds such as amino acids, hydroxy acids, and amines and do not contain life (Glavin et al., 2018). Meteorites therefore can offer glimpses into the chemistry that formed prebiotic compounds and thereby constrain boundary conditions under which these compounds form abiotically.

Carbonaceous chondrites, a type of meteorite that comprise some of the most primitive materials in the solar system, contain 2-5 weight % carbon other biologically relevant compounds such as amino acids, lipids, monocarboxylic acids, and aldehydes (Glavin et al., 2018). Carbonaceous chondrites are grouped by class (*e.g.*, CM, CI, CR, *etc.*) that are perceived to come from the same or similar parent bodies based on mineralogical and isotopic data. Within these classes the meteorites are assigned a number based on the degree of alteration, where 3.0 refers to a sample that has not undergone any alteration,

values less than 3.0 are increasingly aqueously altered as values decrease, and values above 3.0 are increasingly thermally altered as values increase.

In aqueously altered meteorites, indigenous amino acid abundances range from 0.1 to 320 parts per million by weight (Glavin et al., 2018). The structures of meteoritic amino acids can differ from those made biologically. In the 20 biologically-produced amino acids, the carbon bound to a carboxyl group typically is bound to an amine group ('α-amino acid'), a hydrogen ('α-H'), and a functional group (the atoms in the amino acid bonded to $COOH\text{-}CH(NH_2)$ that lead to functional diversity) (Figure 1.1). The four moieties are attached such that in crystalline form, they optically rotate light only in the leftward direction and are therefore said to be have left-handed chirality at the aminated carbon. Carbonaceous chondrites contain over 100 structurally distinct amino acids, including those in which the aminated carbon is between 1 and 4 carbons away from the carboxyl group (known as α-amino, β-amino, γ-amino, and δ-amino acids from 1 to 4 carbons away, respectively). Furthermore, in amino acids from meteorite samples, the amine-bound carbons can either be bound a hydrogen ('α-H' as above) or to a methyl group ('α-CH_3') in addition to functional groups (Figure 1.1). Finally, meteoritic amino acids are typically close to a 50:50 mixture of left- to right-handed amino acids ('racemic') (Martins and Sephton, 2009). Common hypotheses of their formation include ice-grain chemistry in the interstellar medium (ISM) and Strecker synthesis, reductive amination, shock synthesis, and Fischer Tropsch type synthesis (FTT) on a parent body in the solar nebula (Kerridge, 1999).

Isotopic measurements provide a window into the formation of amino acids in meteorites. Isotopes are forms of an element with the same number of protons and electrons but different numbers of neutrons, and in stable (non-radioactive) isotopes, the heavier isotopes tend to be less abundant in nature. Isotopes of an element undergo the same chemical reactions and have the same chemical properties (*e.g.*, electronegativity, solubility) but the chemical bonds they form differ in zero-point energies with the heavier

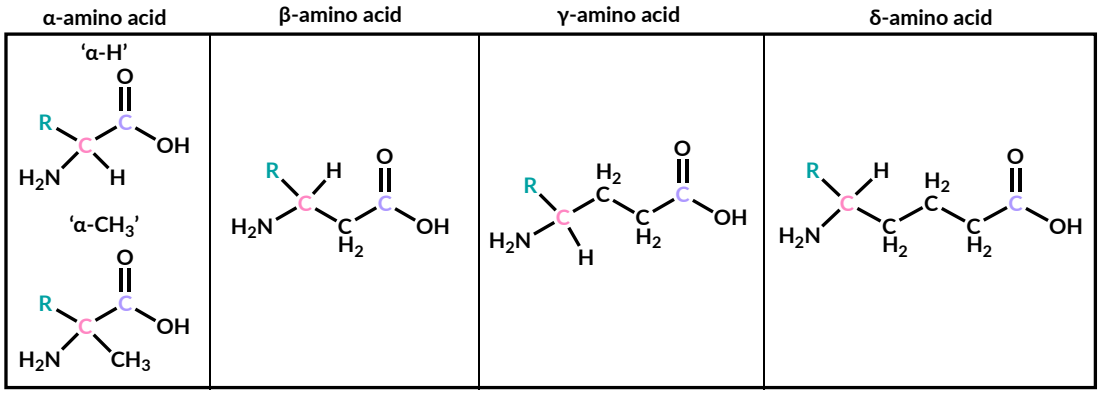


Figure 1.1: Structural forms of amino acids. Carboxyl groups are in purple, amine-bound carbons are in pink, and R groups are in teal. R can represent any group added to the amino acid including hydrogen, methyl, and carboxyl groups.

isotopic substitution resulting in a lower zero point energy. These differences lead to differences in reaction rates and the degree of difference can be influenced by a reaction's physiochemical conditions such as temperature and pressure. In this work, isotopes abundance ratios will be discussed in delta space.[1.1] While biological α-amino acids found in terrestrial systems tend to have negative $\delta^{13}C$ values that fall between -10 ‰ and -35 ‰, and most terrestrial organic compounds are 0 ‰ or lower, those on meteorites tend to fall between 10 ‰. and 60 ‰ (Glavin et al., 2018).

The high $\delta^{13}C$ values in meteorite amino acids occur in other groups of meteoritic soluble organic compounds. In meteorites, amino acids, aldehydes, amines, sulfonic acids, and dicarboxylic acids all have $\delta^{13}C$ that are enriched up to 120 ‰ (Sephton, 2002; Aponte et al., 2014). The high enrichments suggest a linked synthetic history between the

[1.1] Delta notation, $\delta^{h}X$, compares the heavy isotope of element X (^{h}X) to the light isotope (^{l}X) in a sample relative to a standard. It is computed by the equation: $\delta^{h}X = \left[\frac{\left({}^{h}X/{}_{l}X\right)_{sa}}{\left({}^{h}X/{}_{l}X\right)_{st}} - 1\right] \times 1000$ and is reported in units of per mil (‰).

compounds and are reminiscent of the high $\delta^{13}C$ values of organic materials measured in the interstellar medium (ISM).

The 10 – 40 K temperatures found in areas of the ISM lead to the preferential incorporation of heavy isotopes into materials and measurements of ISM organics suggest enrichments of 117 ‰ and higher (Simkus et al., 2019). Specifically, the reaction $^{12}CO + {}^{13}C^{+} \leftrightarrows {}^{13}CO + {}^{12}C^{+}$ favors the forward direction, which is exothermic by 35 K (Charnley and Rodgers, 2006) or 4.8 x 10^{-22} J [2.1]. This equilibrium drives ^{13}C into the CO pool where it often remains due to the 35 K energy barrier to return to the C^{+}. As CO is a dominant carbon pool in the ISM, this reaction sequesters large amounts of the available ^{13}C into the CO pool and leaves the other carbon species relatively ^{13}C-depleted. In this work, we will refer to these two pools as the CO pool and the reduced carbon pool. When carbon from these pools react, they can form compounds such as aldehydes, which would have relatively high ^{13}C abundances on their CO-sourced carbonyl site and low ^{13}C abundances on reduced carbon-sourced aliphatic carbon chains.

While measurements of each site in an organic compound would enable a direct view into which sites were bearing ^{13}C enrichments, previous isotopic analyses of organic compounds could not analyze the samples at the site level due to limitations in instrument mass resolution. Instruments lack the resolution to distinguish different isotopic substitutions if they are all due to the addition to one neutron (*e.g.*, a ^{13}C substitution and a D substitution in a molecule), so samples are typically chromatographically separated by compound, combusted, and either oxidized or reduced into a gas with only one unknown isotopic value for measurement (CO_2 and H_2 to measure the substitutions above). This technique enables molecular-average isotopic composition measurements of individual compounds, but destroys information contained in the structure such as which sites bear heavy isotope enrichments. Therefore, to gain insight into which sites or how

[2.1] To be consistent with the units used in astrochemical and cosmochemical papers, energy in reactions is described in units of Kelvin. The conversion to Joules can be found using the Arrhenius equation, $E = k_B T$.

many sites might be enriched in ^{13}C, the molecular average $\delta^{13}C$ and δD is compared to the number of carbons in structurally similar compounds. If a compound has one relatively enriched site, its measured values would decrease in $\delta^{13}C$ or δD with the number of carbons or hydrogens, respectively, in a molecule.

The Murchison meteorite—a carbonaceous chondrite with the highest number of molecular and isotopic abundance measurements of soluble organic material (Elsila et al., 2016)--is the most suitable sample for the analyses presented in this thesis. The α-amino acids and aldehydes extracted from samples of it have $\delta^{13}C$ measurements decrease with an increase in the number of carbons. This trend could result from a high ^{13}C component, potentially from the CO pool, diluted by lower ^{13}C carbons that are added as the carbon number increases. In ISM chemical models, the carbonyl site in aldehydes tends to arise from CO (Charnley and Rodgers, 2006; Aponte et al., 2017), so this site in aldehydes could have the highly enriched ^{13}C. In amino acids, this site is typically interpreted to be the amine-bound carbon, as in Strecker synthesis—a synthesis in which α-amino acids are synthesized from aldehydes or ketones, ammonia, and cyanide (Figure 1.2)— the amine-bound carbon would be derived from the carbonyl carbon in an aldehyde.

The δD measurements in α-amino acids are also consistently enriched relative to terrestrial values by hundreds to thousands per mil (*i.e.*, enrichment up to about a factor of two) (Pizzarello and Huang, 2005; Elsila et al., 2012). This enrichment suggests that the precursors formed in the ISM, where hydrogen will undergo even stronger enrichments in heavy isotopes than carbon (Rodgers and Charnley, 2002). Unlike the $\delta^{13}C$ measurements, the δD measurements either do not vary (Pizzarello and Huang, 2005) or increase with the number carbon or hydrogen atoms in a molecule (Elsila et al., 2012). The low temperatures of the ISM heavily favor deuteration over hydrogenation due to the decrease in zero-point energy for every deuterium substitution, so each hydrogen site in ISM-formed molecules have a higher chance of a deuterium-substitution than those formed in the higher temperatures in meteorite parent bodies. An increase in the fractional deuterium abundance with number of hydrogen sites is also anticipated for

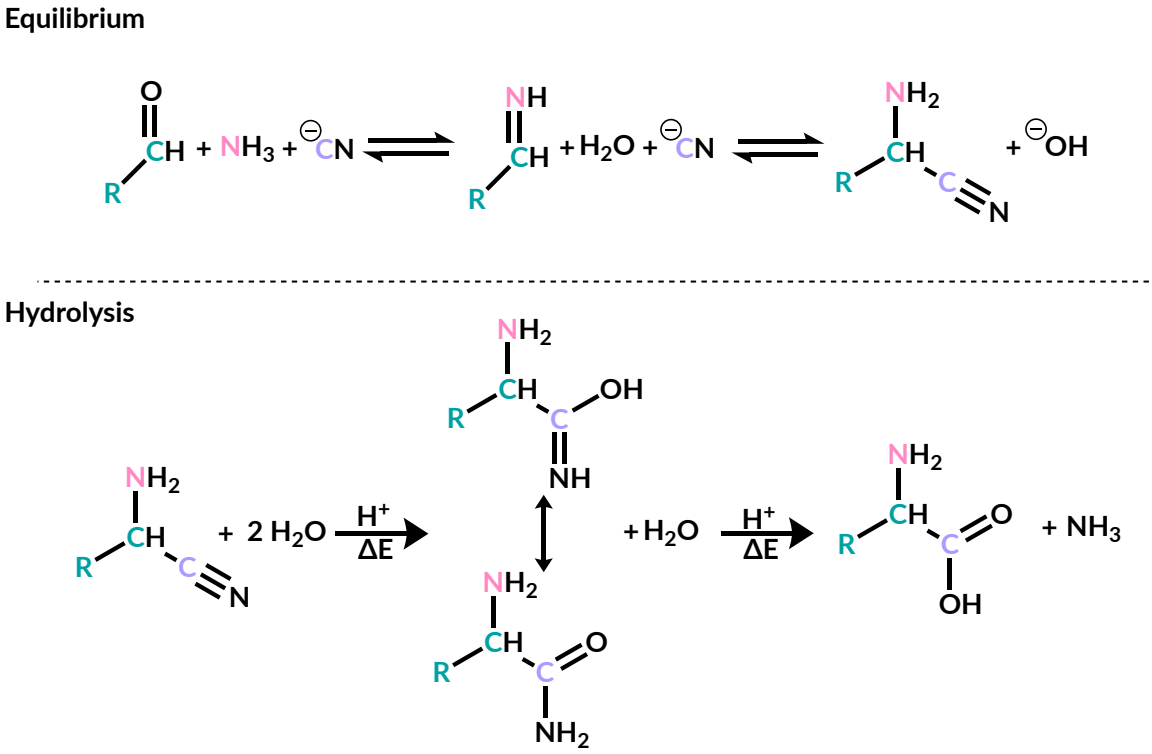


Figure 1.2: Strecker synthesis. Equilibrium and hydrolysis steps are listed separately for clarity. Teal sites represent carbons (and functional group such as CH_3) from aldehyde. Purple carbons are derived from cyanide. Pink nitrogens are derived from ammonia. All other species are depicted in black.

molecules that form in the ISM (Elsila et al., 2012), so these results and the high levels of δD enrichments both suggest a link to the ISM either in direct synthesis of the amino acids or their precursors. In the case of Strecker synthesis, the δD data suggest that either the aldehydes or the α-amino acids themselves are formed in the ISM. If so, the ^{13}C anomaly would likely derive from CO inherited at the carbonyl site in aldehyde or the amine-bound carbon in α-amino acids.

The combined carbon and hydrogen isotopic measurements of amino acids, aldehydes, and other prebiotic compounds point at an integrated network of aldehyde chemistry (‘integrated aldehyde network) (Aponte et al., 2017; Simkus et al., 2019; Aponte et al., 2019; Chimiak et al., 2020) (Figure 1.3). In this network, aldehydes are synthesized in the

ISM and are concentrated on meteorite parent bodies (*i.e.*, the asteroid or comet that is broken to form meteorites). There, the aldehydes can be oxidized into carboxylic acids or react with ammonia to form an imine. The imine can be reduced into an amine or react with CN and undergo hydrolysis to form an α-amino acid.

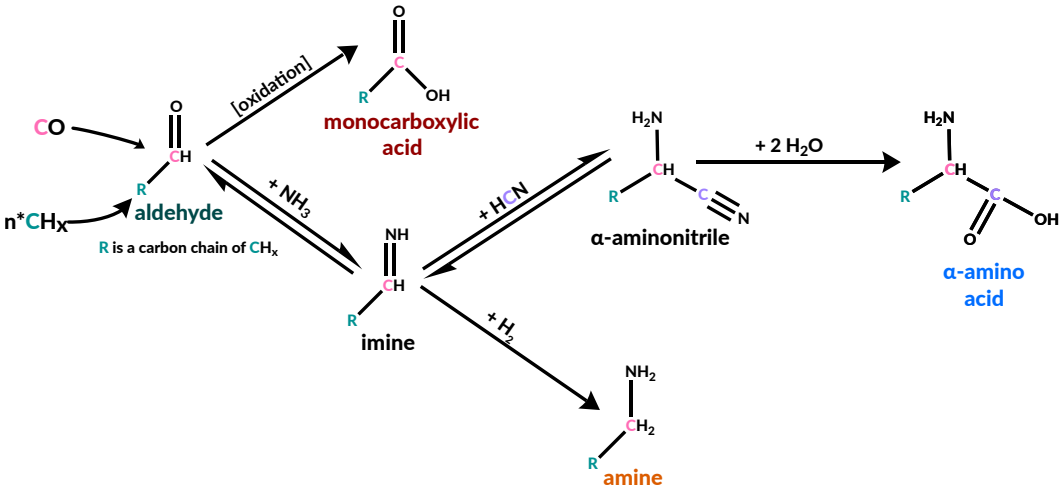


Figure 1.3: Integrated aldehyde chemistry. In this network all pink carbons are initially from the CO pool in the ISM, all teal R groups are functional groups that are derived from the reduced carbon pool in the ISM, and purple carbons are derived from CN.

To directly test this network requires site-specific measurements of the ^{13}C abundances. Molecular-average isotope ratios can provide compelling evidence for the likelihood of one enriched site that is diluted by other carbons in a compound, and they can provide evidence for likely synthetic links between amines, aldehydes, and amino acids. However, molecular-average measurements cannot rule out the possibility that all carbons in the compound are equally enriched in ^{13}C and happen to decrease with chain length or that all sites are similar in $\delta^{13}C$ but that an isotope effect from carbon addition leads to the decrease in $\delta^{13}C$ with chain length (as has been proposed for FTT synthesis (Yuen et al., 1984)). Site-specific measurements of $\delta^{13}C$ are required to understand what leads to the perceived patterns in molecular-average measurements.

In this work, we develop a method by which to measure site-specific isotope ratios (Chapter 2). This method uses Fourier Transform Isotope Ratio Mass Spectrometry on a Q-Exactive Orbitrap with a gas chromatogram and custom-built inlet attached to its front end. The method is demonstrated to be primarily shot-noise limited (*i.e.*, the precision of a measured value is limited as the square root of the number of individual measurements) and able to measure site-specific ^{13}C ratios on samples of 1 picomole or more.

We apply this technique to measure the site-specific isotope effects associated with Strecker synthesis of alanine (Chapter 3). We measure a -20 ‰ equilibrium isotope effect between the carbonyl carbon on acetaldehyde, which is relatively ^{13}C-enriched at equilibrium, and the eventual amine-bound carbon in alanine, which is relatively ^{13}C-depleted (Figure 1.2). We also measure two kinetic isotope effects on the eventual carboxyl carbon associated with the two elementary reactions in the hydrolysis of the C-1 site from nitrile to amide to carboxyl group. The first step has at least one kinetic isotope effect creates an amide carbon that is up to 15 ‰ ^{13}C-depleted relative to the starting nitrile carbon. The second step has a kinetic isotope effect measured to be -15.4 ‰ and so creates a carboxyl carbon that is up to 15.4 ‰ depleted relative to the amide carbon (Figure 1.2). The amine nitrogen is measured to have a 54 ‰ equilibrium isotope between the more ^{15}N-depleted ammonia and the more ^{15}N-enriched amine nitrogen at equilibrium.

We further apply the novel measurement method to interrogate the site-specific isotope ratios of alanine in a sample of Murchison meteorite (Chapter 4). This sample is found to have site-specific $\delta^{13}C$ values of -29 ± 10 ‰, 140 ± 20 ‰, and -36 ± 20 ‰ for the carboxyl, amine-bound, and methyl carbons, respectively (Chimiak et al., 2020). The pattern is consistent with Strecker synthesis of the alanine, and the highly enriched amine-bound carbon is equal to the lower end of ISM formaldehyde measurements (Simkus et al., 2019).

Finally, we combine these insights and molecular average data from the literature to build an integrated model of meteorite organic synthesis (Chapter 5). This model separately tests the δD and the $\delta^{13}C$ data. The δD model uses different chemical moieties (*e.g.*, methyl hydrogens or those attached to carbon in a resonant ring) to create the measured molecular average δD values. The δD model predicts high methyl hydrogens relative to others by up to 5000 ‰ depending on the meteorite sample and the compound class. The methyl hydrogen's enrichment scales with the compound class's solubility in water and with the meteorite sample's weathering grade and degree and type of alteration. Thermally altered meteorites have the δD enrichment erased and more highly-weathered or aqueously-altered meteorites have the methyl δD enrichment partially to fully erased. The $\delta^{13}C$ model tests integrated aldehyde chemistry, Formose reactions, and FTT synthesis and finds that most of the molecular-average isotope data can be best described if the integrated aldehyde chemistry creates straight-chain and α-H compounds that then have formaldehyde added to them to create branched-chain and α-CH_3 compounds. We leave with the recommendation for further site-specific isotope measurements to test the site-specific isotope ratio predictions put forth by this model.

Chapter 2

ORBITRAP ISOTOPE RATIO MASS SPECTROMETRY: A NEW METHOD TO PROBE ISOTOPIC STRUCTURES

Abstract

Site-specific isotope ratios contain information constraining a molecule's synthetic history, including the reservoirs from whence the precursors came, the physiochemical conditions during synthesis, and the synthetic mechanism. In spite of the richness of information, conventional measurement techniques are limited to molecular-average isotope-ratio measurements. Here, we discuss new methods that use a Q-Exactive Orbitrap ('Orbitrap'), a Fourier Transform Isotope Ratio Mass Spectrometer (FT-IRMS) with resolutions of up to 240,000 (FWHM at 200 u), to measure the site-specific isotope structure of alanine. The Orbitrap allows for two modes of analysis: a dual reservoir mode, which is similar to a dual inlet measurement on a conventional IRMS system in that a user dynamically switches back and forth between an unknown sample and a reference standard, and a single-reservoir mode in which peaks eluting from a chromatogram can be captured and analyzed, and standardization is achieved by comparison with a separately captured peak of a standard. Isotope ratios measured using the Orbitrap show high stability on week-long timescales and can be measured on samples as low as 1 picomole. We discuss methods by which to process data to address non-linear effects with sample size.

2.1. Introduction

Isotopes are forms of the same element having different numbers of neutrons. As such, they have distinct atomic masses and their motions in chemical bonds have distinct vibrational energies. While isotopes of an element have approximately the same chemical properties and generally undergo the same reactions, they vary in reaction rate and in the free energies of formation of molecules in which they are bound. Due to these disparities, a compound's isotopic structure (ratio of heavy to light isotopes at each

atomic site in a molecule) is impacted by and therefore contains information about its synthetic pathway(s) and physiochemical conditions of formation and storage such as pressure, temperature, and pH. As these effects occur in precursor materials too, the isotopic structure of a product potentially encodes the chemical history of several reactants that ultimate contribute to that product. It is unsurprising, therefore, that isotopic measurements are prominent in food and drug forensics studies, molecular biology, and astrochemistry amongst other fields (Martin, 1996; Scott et al., 2006; Glavin et al., 2018).

Elements that commonly participate in organic chemical reactions (*e.g.,* C, H O, N) have two orders of magnitude or higher abundances of light stable isotopes compared to heavy ones (*e.g.*: ^{1}H vs D, ^{12}C vs ^{13}C, ^{14}N vs ^{15}N, ^{18}O or ^{17}O vs ^{16}O) and those ratios typically vary by percent or fractions of a percent, relative; thus, observing variations in isotopic composition necessitates highly sensitive and precise measurements. Historically, these measurements were accomplished by measuring a sample relative to a standard with a gravimetrically determined isotope ratio via isotope ratio mass spectrometry (IRMS), which is now a conventional technique that can measure isotope ratios on micromoles to picomoles of sample. Most IRMS instruments have a mass resolution of ~200 full-width half maximum (FWHM) (Eiler et al., 2013) and therefore cannot differentiate M+1 isotopic substitutions (*i.e.* any substitution of an isotope that is approximately 1 Da above the lighter stable isotope such as a ^{13}C substitution for ^{12}C or a ^{15}N substitution for a ^{14}N) from one another, so prior to measurements, compounds are combusted and converted into a gas that contain fewer or only one element with unknown isotope ratios from a sample (CO_2 for C and N_2 for N). While this step enables measurement of a sample's molecular average isotopic ratio, it destroys information contained at the site level. Even ultra-high resolution IRMS have mass resolution of 50,000 FWHM so can differentiate multiple isotopic substitutions form one another including $^{13}CH_3D$ and $^{12}CH_2D_2$ but can only do so for low molecular mass gases (*e.g.* N_2O, NO_2, CO_2, CH_4) (Eiler et al., 2013). A non-destructive technique, cavity-ringdown spectroscopy (CRDS), leverages differences in vibrational frequencies between molecular isotopic forms to distinguish isotope ratios without chemical conversion and often without significant risk of

interferences by species other than isotopes of interest. However, CRDS can only measure gas-phase samples and the requirement for recognized, spectrally 'clean' absorption features restrict its utility to structurally simple, low-molecular-weight gases (CO_2, CH_4, H_2O, N_2O). More complex molecules must be converted to these forms for CRDS analysis, again destroying site specific isotopic information.

Several classes of instruments and techniques enable site-specific measurements, in some cases defining a full carbon or hydrogen isotopic structure. These data can improve constraints on reaction pathways because bond formation and destruction impart greater primary isotope effects (isotope effects at the site of the reaction) than secondary (adjacent to the reaction site) ones, and sites derived from different precursors can be measured separately. These techniques include site-specific natural isotopic fractionation by nuclear magnetic resonance (SNIF-NMR), chemical and pyrolytic cleavage of compounds prior to IRMS measurements, and ultra-high resolution IRMS, specifically aimed at two or more fragment ions that differ in the molecular sites they sample. While these technologies have helped elucidate chemical pathways, SNIF-NMR and chemical cleavage generally require 10's to 100's of mg-sized samples of a pure compound (Thomas et al., 2010), which is too large for many environmental and biomedical applications. An exception is the GC-combustion-GC-IRMS technique, which permits site specific analysis of carbon isotopes on a limited number of compounds for sample sizes down to nanomoles (Dennis et al., 1998; Oba and Naraoka, 2006; Li et al., 2018).

This paper presents methods developed for measuring site-specific carbon isotope composition of small organic compounds on the Q-Exactive GC Orbitrap ('Orbitrap'), a Fourier transform mass spectrometer. With a resolution of 240,000 (FWHM at 200 u; better at lower mass and poorer at higher mass), the Orbitrap mass analyzer resolves different M+1 (^{15}N, ^{13}C, D, *etc.*) substitutions in samples, and its high sensitivity enables the study of the picomole-sized samples. The Q-Exactive has options for either electron impact (EI) and electrospray ionization (ESI) sources (on the 'GC' and 'HR' models, respectively), enabling the study of gas and liquid samples and for front-end attachments such as gas and liquid chromatographs (GC and LC).

Recent papers discuss Orbitrap measurement techniques for pure samples (Eiler et al., 2017; Neubauer et al., 2018; Neubauer et al., 2020), with co-eluting peaks (Hofmann et al., 2020), and demonstrate their use on applied problems (Cesar et al., 2019; Chimiak et al., 2020). This study demonstrates a full carbon isotopic structure measurement for alanine, including in-line GC separation from complex mixtures with the ability to reach 1 ‰ precision at each site. The methods presented here can be adapted to other amino acids and organics without prior separation and can work on picomole-sized samples.

2.2. Materials and Methods

2.2.1. Alanine standards

Three alanine standards are used in this study: one from Alfa Aesar (99 % Purity, 'Alfa Aesar alanine'), one from VWR (Purity >99 %, Lot# 2795C477, 'VWR alanine'), and one produced by Strecker synthesis in 2015 (purity confirmed by NMR, 'Strecker alanine'). Strecker alanine is synthesized following methods modified from Kendall et al. (1929) to take place in a closed system (see Chapter 3 for further information). In addition to these natural abundance standards, we use three ^{13}C-labelled alanine standards purchased from Sigma Aldrich, a ^{13}C-1 standard (Lot # EB2220V), a ^{13}C-2 standard (Lot # SZ0643V), and a ^{13}C-3 standard (Lot # EB2211V).

2.2.2. Molecular-Average Measurements of Standards

The δ^{13}C measurements averaged across the whole molecular ('molecular-average δ^{13}C') of the pure alanine standards described above are made on an Elemental Analyzer attached to a Delta V IRMS. Before measurements, each alanine sample and acetanilide standard are placed in an aluminum capsule. Capsules are handled with tweezers that have been wiped down in isopropyl alcohol between each sample handling to avoid contamination. Alanine samples and acetanilide standards are measured on the same day. Sample measurements are processed on IsoDat, a proprietary software made by Thermo

that converts the measurement data (*e.g.*, current on a Faraday cup, Fourier Transform in an Orbitrap) into isotope ratio values relative to a standard (here $\delta^{13}C_{VPDB}$). Molecular average $\delta^{13}C$ values are listed as the average of at least three measurements (Table 2.1). Measurements used here were made on two separate days. EA acetanilide standards $\delta^{13}C$ values are -28.0 ± 0.1 ‰, which is within error of the -27.7 ± 1.7‰ value calibrated against a NIST sucrose standard.

Type	Reactants and Buffer (µmol)			rxn time (min)	% yield	$\delta^{13}C_{VPDB}$ (‰)	$\delta^{13}C_{VPDB}$ ± st err (‰)		
	ala	nin	citric acid				C-1	Molecular Avg	C-2 + C-3
Alfa Aesar	219	415	460	32	28%	-28.6	-28.5 ± 0.1	-19.4 ± 0.2	-14.8 ± 0.3
Alfa Aesar	155	460	478	32	39%	-28.4			
Alfa Aesar	137	237	481	32	33%	-28.7			
Alfa Aesar	182	298	480	32	44%	-28.5			
Alfa Aesar	203	310	479	7.5	37%	-30.9			
VWR	135	295	473	32	39%	-29.5	-29.6 ± 0.1	20.0 ± 0.2 ‰	-15.2 ± 0.3 ‰
VWR	139	285	469	32	45%	-29.7			
Strecker	162	288	513	32	48%	-43.4	-43.5 ± 0.1	-32.9 ± 0.1	-28.0 ± 0.2
Strecker	181	278	471	32	42%	-43.6			

Table 2.1: Ninhydrin (nin) decarboxylation results for the $\delta^{13}C_{VPDB}$ of the C-1 site on three alanine standards. All but one sample have 32 minutes for reaction time, but the 7 minute time is close to the same value and had similar yields. However, in order to compare similar materials, we only include the 32-minute reactions in the averaged values for the $\delta^{13}C_{VPDB}$ of the C-1 site standard.

2.2.3. Measurement of the C-1 site by ninhydrin chemical cleavage

Ninhydrin cleaves amino acids' carbon (the C-1 site), which becomes CO_2 (Donald D. Van Slyke et al., 1941). Combined with conventional IRMS, the ninhydrin reaction followed by capture of the produced CO_2 provides a method by which to measure an amino acid's $\delta^{13}C$ of the carboxyl site (the C-1 site) (Abelson and Hoering, 1961). Here we combine the ninhydrin decarboxylation reaction with IRMS measurements to constrain $\delta^{13}C$ of the C-1 site for the alanine standards.

The ninhydrin reaction (D D Van Slyke et al., 1941) is performed in a custom-made 50 mL borosilicate flask (Figure 2.1) with one port that attaches to a custom-built gas-line (that is used similarly to a Schlenk-line in this study) and can be sealed and one port that has a Swagelok fitting and septum. The gas-line has multiple water traps in line and two pressure gauges attached to it to enable pressure readings from 10^{-3} mTorr to 760 Torr. Ninhydrin, citric acid, and milli-Q water (18.2 MΩ cm; hereafter 'water') are used with the alanine standards in the reaction. Water for the ninhydrin reaction is bubbled with helium (UHP, >99.9 % purity) in a 40 mL vial capped with a Teflon cap by using two 26-gauge sterile needles (BD, Item# 305111), one connected to the helium line and one to allow the gas to purge. Water is added to the custom-made vial flask a 10 mL sterile Luers-lock plastic syringe (BD, multiple lots) and sterile needle (same as above). Carbon dioxide is collected in a borosilicate tube that is sealed with an acetylene torch after carbon dioxide is frozen in it.

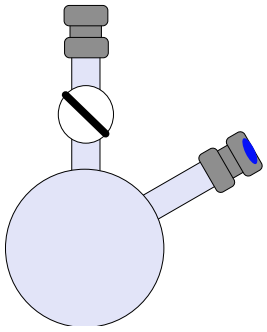


Figure 2.1: Custom made borosilicate flask used in this study for ninhydrin decarboxylation experiments.

The ninhydrin decarboxylation follows methods adapted from Van Slyke et al. (1941). In short, we add 57 mg of ninhydrin, 15 mg alanine, and 92 mg citric acid (exact quantities in Table 2.1) to a custom-made borosilicate flask. This flask is then attached to a gas-line via a Swagelok with a frit to avoid solid materials from entering the line and is then evacuated. While the flask is being evacuated, 10 mL water is added to a 40 mL vial that is subsequently capped. Helium is bubbled through the water for 10 minutes to remove any carbon dioxide. Once the flask is evacuated (0.005 mTorr or below as read by the pressure gauge), both the port to the reaction flask and to the gas-line are closed. The flask is taken off the line and place in a fume hood where 5 mL of degassed water is

injected into the vial through the septum. The flask is then placed in an ethylene glycol bath at 100 °C for 32 minutes after which it is removed from heat and connected to the gas-line. The flask is placed in a -15 °C ethylene glycol-dry ice bath for 20 minutes or until all liquid in the flask is frozen, whichever is longer. During this time, a borosilicate tube is attached to another port on the Gas-line, evacuated, and the port to it is closed. Once the contents of the reaction vessel are frozen, the ports to the pumps are closed and those to the reaction vessel are opened and the pressure is recorded. One of the water traps is then submerged in an ethanol-dry ice bath for 5 minutes or until the pressure reaches 0.005 mTorr. At this point, the port to the reaction vessel is closed, the ethanol-dry ice bath is removed, and the gas equilibrates in the line. We record the pressure, open the port to the borosilicate tube, and immerse the tube in liquid nitrogen. One minute after the pressure reaches 0.005 mTorr, the tube is sealed with a O_2 torch.

Once trapped in a borosilicate tube, CO_2 is measured by IRMS on a Thermo Fisher Delta V with two ethanol-dry ice traps and a GC on the front end for sample purification. To introduce sample into the instrument, borosilicate tubes are scored, and the scored end is sealed into a flexible metal tube via a Swagelok. The other end of the metal tube has a second Swagelok with a frit on it—this Swagelok is attached the purification line on the front end of the Delta V. After evacuating inlet to the tube, a line is open between it and the first gas trap and the tube is broken. Gas equilibrates for one minute before the inlet port is shut. Measurements of CO_2 for each sample is measured in dual-inlet mode against a CO_2 standard with a known $\delta^{13}C_{VPDB}$ composition.

2.2.4. Site-specific Isotope Ratio Measurements by Orbitrap IRMS with an EI Source

To attain site-specific isotope ratio measurements on Orbitrap IRMS a user measures the isotope ratio for multiple different fragments of a molecule for both a standard and for a sample. Fragments that sample different carbon atoms from the parent molecule can be analyzed for isotope ratio and provide separate, independent constraints on a compounds'

$^{13}C/^{12}C$ structure. One can then constrain a compound's site-specific isotope ratios by difference between the site's fractional isotopic abundance

The path of the sample from the GC to the Orbitrap and its potential impacts on the isotopic measurements is described at length in Eiler et al. (2017). In short, gas is separated on a gas chromatogram and then enters the EI Source. There, electrons from the source impact the analyte and break chemical bonds and/or remove electrons to produce charged fragments. These fragments are then collimated into a beam and filtered by their mass-to-charge ratio via the Advanced Quadrupole Selector (AQS)—a quadrupole mass selector. Typically, fragments have a charge of +1, so the mass-to-charge ratio corresponds to the fragments' mass. We will refer to the mass range that is filtered in the quadrupole as the 'mass window'. Ions in the mass window are collected in an electrostatic trap (C-trap) until a user-defined threshold for the total ion current (TIC) in the C-trap is reached. This threshold measured via the automatic gain control (AGC). Once the TIC in the C-trap reaches the threshold, ions are injected from the C-trap into the Orbitrap where they oscillate and travel back and forth in the Orbitrap at a rate proportional to mass (Makarov, 2000). The time the ions spend in the C-trap prior to injection is referred to as the injection time (IT). If the TIC in the C-trap remains below the threshold for a user-defined amount of time (the max IT), the ions are injected into the Orbitrap and measured. For these scans, the IT is said to have maxed out. We refer to each process of injecting an ion packet into the Orbitrap and measuring it in this manner as a scan and the combination of scans that forms a measurement as an acquisition. Using Excalibur, the proprietary software program by Thermo Fisher Scientific, the frequency signal produced by the back and forth travel of ions is deconvolved via a Fourier Transform function and the mass and signal intensity of the ions is reported to a user for each scan in an acquisition.

2.2.5. Derivatization

Alanine is not adequately volatile for measurement by EI Orbitrap IRMS. Prior to measurement, alanine and other amino acids are derivatized into an N-trifluoroacetyl-O-

methyl-ester. This derivative has the advantage of increasing alanine's volatility such that it can be introduced into a GC; however, it adds carbons to the compound that will be injected into the instrument. Samples and standards are derivatized in tandem using the same derivatizing agents so that the added carbons and hydrogens will have the same $\delta^{13}C$, but as fragments will also sample these carbons, the additional carbons added to the molecule will dilute the differences between the sample and standard and therefore increase the uncertainty in the calculated difference between sites in the compound. To overcome this, either more sites need to be measured or measurements must occur for longer periods of time. The N-trifluoroacetyl-O-methyl-ester derivative adds three carbons to alanine (Figure 2.2): two are added on the amine nitrogen and one is added to one of the oxygens in alanine's carboxyl group.

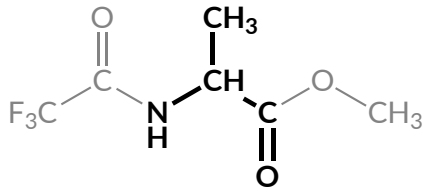


Figure 2.2: Illustration of the N- trifluoroacetyl-O-methyl-ester derivative of alanine used in this study. Carbons from alanine are bolded and in black font. Bonds and atoms from derivatizing agents are in unbolded and in gray.

2.2.5.1. Derivatization Materials

Each alanine standard is derivatized into an N-trifluoroacetyl-methyl-ester for Orbitrap analyses. Standards greater than 1 mg are derivatized in 40 mL borosilicate vials with Teflon-lined caps and those 1 mg and smaller are derivatized in 2 mL GC vials. All vials are washed and combusted at 450°C prior to use.

Alanine and amino acid mixtures are derivatized with methanol (MeOH; >99.8% purity, Macron Fine Chemicals), acetyl chloride (>99 % purity, Sigma Aldrich), hexane (>98.5 % purity, Millipore Sigma, HPLC grade), and trifluoracetic anhydride (>99 % purity, Sigma Aldrich). All standards are derivatized in tandem to avoid any potential day-to-day isotopic fractionation in the derivatizing agents.

2.2.5.2. Derivatization Methods

Derivatization uses the same procedure outlined in Chimiak et al. (2020). Derivatization was conducted on samples from 1 mg to 100 mg. Here, we provide a description of the methods for 1 mg alanine, but this method can be scaled up by increasing the reagents proportionally to the increase in alanine. The times for and temperatures of each heating step remain the same.

A vial with 100 uL of methanol is place in ice and 25 uL acetyl chloride is added dropwise as the vial is gently agitated. The acidified methanol is transferred into a vial that contains 1 mg alanine, and that vial is capped and heated to 70 °C for one hour after which it is uncapped and dried under nitrogen. Next, 120 uL hexane and 60 uL trifluoroacetic anhydride are added to the vial. This solution is capped and heated to 60 °C for 30 minutes. The vial is uncapped and dried until approximately 50 uL hexane remains for microsamples (1 mg of alanine or less) or until only derivative remains as determined by gravimetry for larger samples. The derivative is then either kept pure (for dual reservoir measurements) or diluted to 1 part in 100,000 in hexane.

2.2.6. Measurements on the Orbitrap

2.2.6.1. Orbitrap inlet systems

The Q Exactive GC Orbitrap has a gas chromatograph front-end to enable the separation of compounds. For this work, we modified a Thermo Scientific™ TRACE™ 1300 GC with a 30 m TraceGOLD TG-5SilMS column to allow for two configurations in addition to that of the GC directly connected to the Orbitrap source. In both configurations, reservoirs sit in the GC Oven, which allows a user to increase or decrease the temperature that the analyte experiences as needed.

2.2.6.1.1. Dual reservoir system

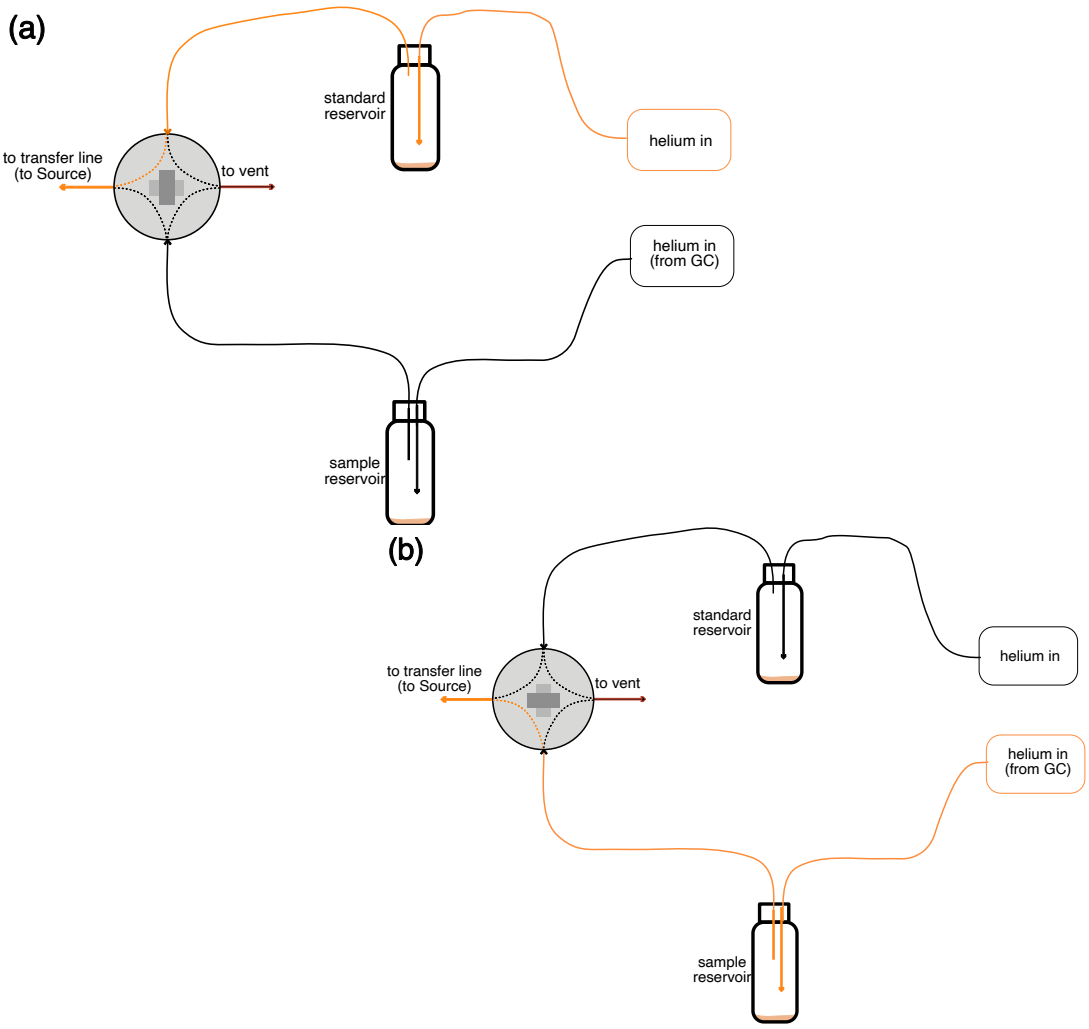


Figure 2.3: Two settings for the dual reservoir configuration. Orange trace follows the path of the analyte entering the source. In the dual reservoir system, only the valve on the left (to source/vent) is changed during the course of a run.

The dual reservoir system is analogous to a dual-inlet IRMS system: it has two reservoirs, one that contains standard and one that can contain a sample, and only analyte from one can enter the source at a time (Figure 2.3). By having them enter the source sequentially—standard, then sample, then standard, etc.—one can perform bracketed isotope ratio measurement of samples and standards with relatively constant amounts of analyte entering the source (Figure 2.4). Here, each reservoir consists of a 2 mL

borosilicate glass tube capped with a Swagelok fitting that has a two-hole ferrule through which two capillaries are threaded. One capillary delivers helium into the reservoir and one carries the gas from the reservoir to either the Orbitrap source or to a vent capillary connected to a scroll pump. The helium inlet is from the GC (though any other pressure-regulated helium supply could be used in its place) for one reservoir and from a capillary attached to an external regulator for the other, so such that both helium flow rates can be independently controlled. The outgoing capillaries are attached to ports in a two-way valve such that when one is directed to the Orbitrap source, the other is directed to vent. Samples are highly volatile liquids that are transferred directly to the sample tubes before the tubes are attached into the system via the Swagelok fitting.

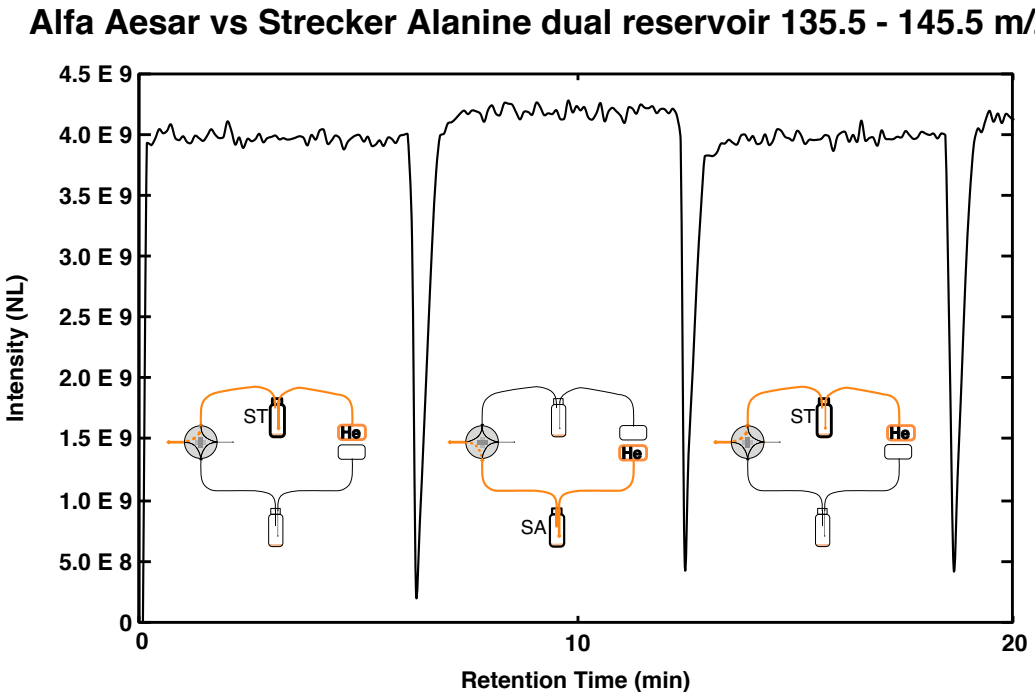


Figure 2.4: Example of part of a measurement using the dual reservoir configuration on the front end of the Orbitrap. Orange trace follows the analyte to the Orbitrap source. Valve changes are apparent by the steep decrease in beam intensity.

Dual reservoir measurements of alanine made in this study use 5 uL of derivatized alanine placed into the bottom of a reservoir using a syringe (Hamilton, Item 80330). Prior to injection and between samples, syringes are rinsed three time with 10 uL hexane,

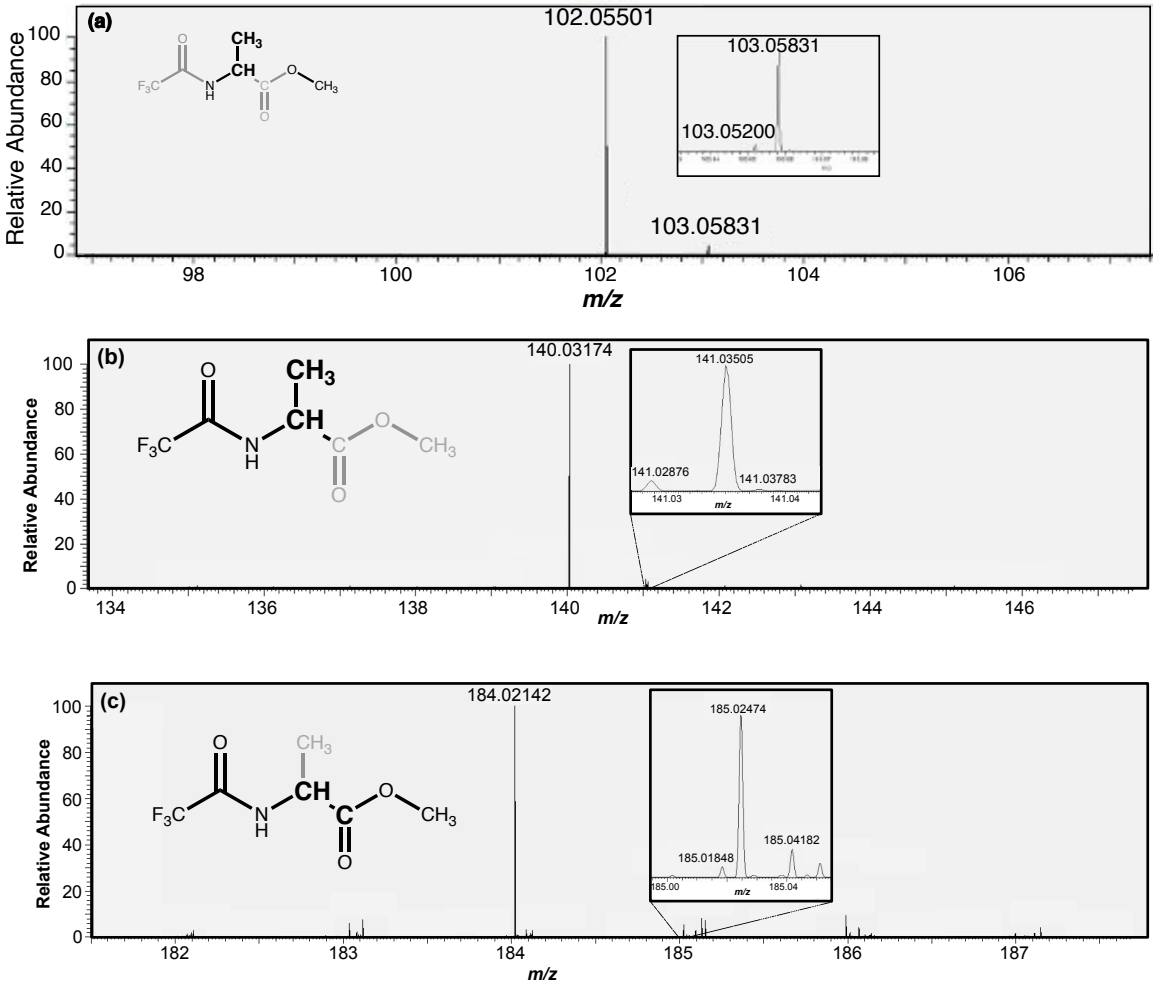
three times with 10 uL dichloromethane, and one time with 10 uL hexane. Reservoirs with samples are connected to the system and a leak check is performed. Helium flow rates are initially set to 1.2 mL/min. For measurements of alanine N-trifluoroacetyl-O-methyl ester, the GC oven temperature is held at 60 °C.

To perform measurements, the mass window centered between the two fragment peaks of interest—typically a monoisotopic peak (only containing ^{12}C, ^{16}O, ^{14}N, and ^{1}H isotopes) and a singly-substituted one—with 5 *m/z* on either side. If the ± 5 *m/z* mass window includes significant contributions from contaminant peaks, the window can be decreased, such that the center of the window is still between the two peaks of interest. However, this was unnecessary for the dual reservoir measurements conducted here. Measurement windows used here are listed in Table 2.2 and shown in Figure 2.5.

Mono-isotopic Mass	Date	Window Size	AGC	µ scan	Res	$^{13}R_{sa}/^{13}R_{st}$ Meas	$^{13}R_{meas}/^{13}R_{pred}$ -1	St Err norm
102	**5/21/16**	**97-108**	**5E5**	**10**	**240k**	**0.9929**	**-0.0004**	**0.0014**
	5/16/16	97-108	1E6	10	240k	0.9983	0.0050	0.0016
	Predicted Difference						0.993325	
113*	**7/16/18**	**108-118**	**5E4**	**1**	**120k**	**0.9945**	**0.0038**	**0.012**
	Predicted Difference						0.990658	
140	4/22/16	135-146	1E5	3	120k	0.9908	-0.0025	0.0021
	4/25/16	135-146	1E5	3	120k	0.9905	-0.0029	0.0036
	4/25/16	**130-190**	**1E5**	**3**	**120k**	**0.9922**	**-0.0011**	**0.0026**
	4/25/16	135-146	1E5	3	120k	0.9897	-0.0036	0.0035
	5/16/16	135-146	5E5	10	240k	0.9868	-0.0066	0.0025
	5/17/16	135-146	1E5	10	240k	0.9929	-0.0005	0.0038
	5/19/16	135-146	5E5	10	240k	0.9859	-0.0075	0.0026
	5/21/16	135-146	1E5	10	240k	0.9934	0.0001	0.0034
	Predicted Difference						0.993325	
184	**4/25/16**	**170-210**	**1E5**	**3**	**120k**	**0.9946**	**0.0002**	**0.0054**
	4/25/16	170-210	5E5	5	120k	0.9928	-0.0017	0.0004
	5/06/16	170-210	5E5	5	120k	0.9925	-0.0019	0.0009
	5/13/16	170-210	5E5	10	240k	0.9923	-0.0022	0.0006
	5/13/16	170-210	5E5	5	120k	0.9925	-0.0019	0.0009
	5/13/16	170-198	1E6	10	240k	0.9905	-0.0039	0.0002
	5/12/16	176-194	5E6	10	240k	0.9823	-0.0122	0.0018
	Predicted Difference						0.994453	

Mono-isotopic Mass	Date	Window Size	AGC	μ scan	Res	$^{13}R_{sa}/^{13}R_{st}$ Meas	$^{13}R_{meas}/^{13}R_{pred}$ -1	St Err norm
200	4/25/16	170-210	1E5	3		0.9993	0.0062	0.0118
	5/06/16	**170-210**	**1E5**	**5**		**0.9891**	**-0.0042**	**0.0033**
	5/12/16	190-207	1E6	10		0.9714	-0.0219	0.0037
	5/12/16	190-207	5E6	5		1.0034	0.0103	0.0019
	5/12/16	190-207	5E6	10		1.0014	0.0082	0.0007
	Predicted Difference						0.993188	

Table 2.2: Data for Alfa Aesar vs Strecker alanine samples for five fragments of interest. Predictions computed using the differences in $\delta^{13}C$ between the Alfa Aesar and Strecker alanine found from the ninhydrin data, full-molecular average measurements on the EA, and from prior work conducted on the Dual Focusing System, which constrained the C-2 sites. The measurements that most closely match predicted values for each fragment are in bold font. Standard error is normalized to the measurement.



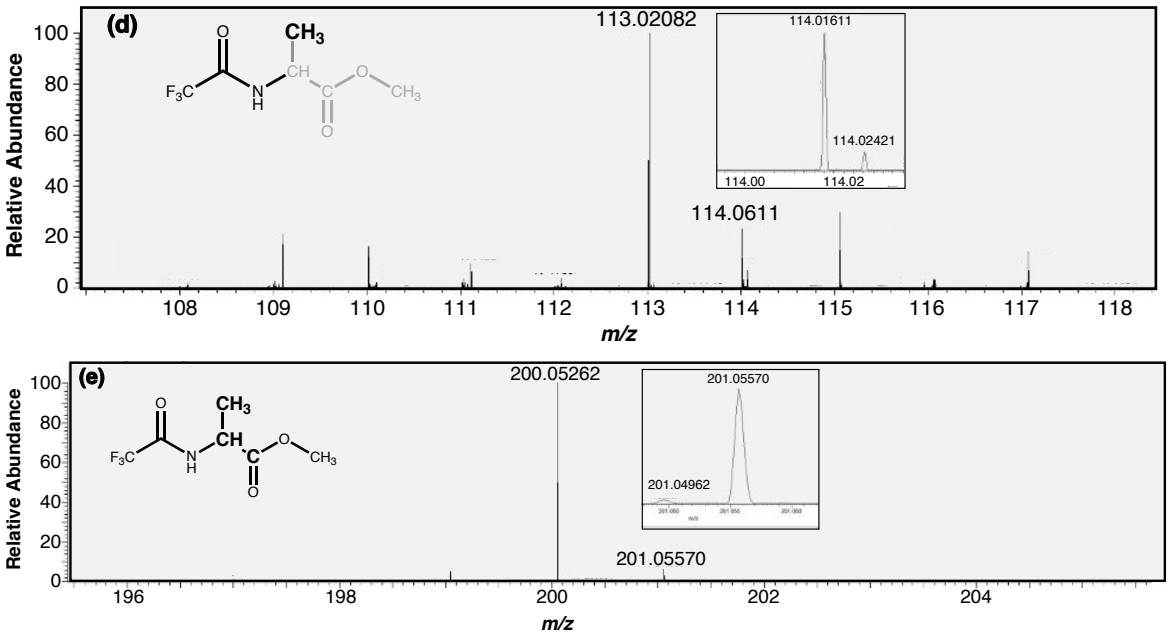


Figure 2.5: Fragments of interest and their mass spectrum. Insets show the ^{13}C-substituted peak. For each fragment, atoms in grey are not included and enlarged, bold font denotes alanine carbons that are included in a fragment. Insets in the mass spectrum show the ^{13}C-substitued peak and nearby fragments that are resolved.

Prior to starting analysis, helium flow rates are adjusted such that the amount of analyte entering the ion source from each reservoir—measured by the intensity in NL units (the nominal measure of intensity reported by the software controlling the Q Exactive GC Orbitrap mass spectrometer) of the monoisotopic fragment peak—are within 10 % of each other. For the analyses here, the standard analyte enters the source for 7 to 10 minutes after which the valve is switched to route the sample to the source and the standard to vent. After 7 to 10 minutes, the valve is switched back to route the standard to the source and the sample to vent (Figure 2.4). These operations continue for the duration of a measurement.

Measurements here explore the following parameter spaces: automatic gain control (AGC) from 5 x 10^4 ions to 5 x 10^6 ions per scan, resolution from 120,000 to 240,000 (nominal FWHF at 200 u; actually, resolutions are greater at the masses of ion species analyzed here), and microscans from 1 to 10. The AGC setting defines the amount of ions that are collected prior to injection into the Orbitrap for each scan, resolution controls the ability to differentiate peaks that fall within the analyzed mass window, and microscans

are the number of scans that are signal averaged into each reported cycle prior to fast-Fourier-transform analysis to retrieve peak masses, intensities, and noise.

2.2.6.1.2. Single reservoir system

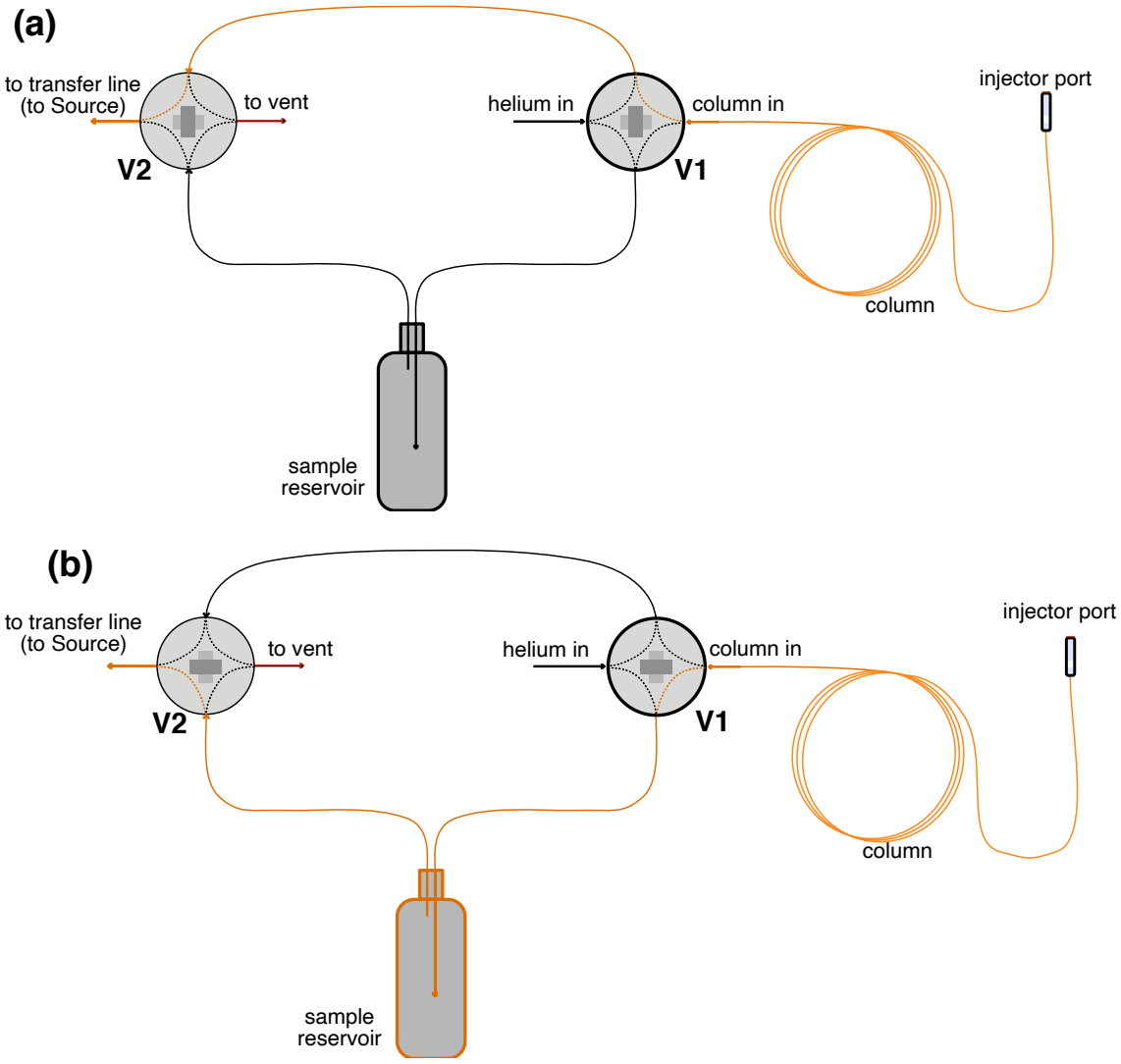


Figure 2.6: Two ways to measure the column eluent in the single reservoir configuration. In (a) the orange trace follows the direct path to from the column to the Orbitrap source and (b) the orange trace follows the path from the column to the source as a peak is captured.

The single reservoir system is depicted in Figure 2.6 and operates similarly to an exponential dilution flask. Depending on the application temperature, reservoirs consist of a borosilicate reservoir that ranges in size from 2 mL to 20 mL (custom-made from borosilicate glass) or a steel reservoir coated with deactivated silica that ranges in size from 10 mL to 30 mL. The single reservoir system has two valves. The first valve (V1 in Figure 2.6) has inputs from the GC column and from a capillary attached to the helium tank by way of a pressure regulator, and outputs are either directed to the second valve (V2) or to the reservoir. V2 is the same as the valve in the dual reservoir system; it has inputs from the V1 and the reservoir and outputs to either the Orbitrap source or to vent.

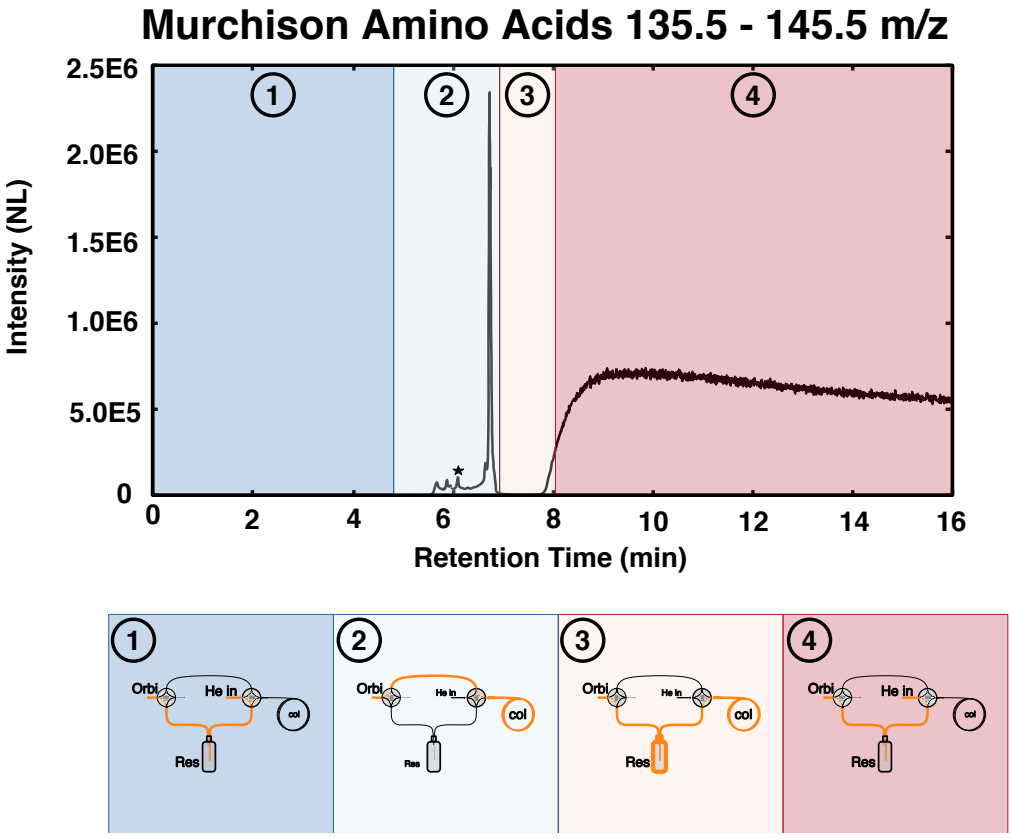


Figure 2.7: Example of the beginning of a single reservoir measurement. The bottom panel shows the different analytes flowing into the Orbitrap at different during the measurement. The numbers and shading in the bottom panel correspond to those in the top panel. Star in the top panel marks the marker peak to enable timing for the valve turns needed to capture the peak

In single reservoir analyses, samples are injected into the GC injection port and can be collected and broadened in a reservoir (Figure 2.7). Initially valves are set so helium is

directed to the reservoir and the reservoir is directed to the Orbitrap source (the GC eluent is directed to vent). The GC is set to start at 50 °C and ramp at 10 °C/minute with a helium flow rate of 1.2 mL/min. The studies here are initially performed with the injector port using a 1:6 split but are expanded to include splitless injections for smaller peaks in microsamples. The single reservoir system allows users to measure compounds directly from the GC ('GC direct mode') or as it is being broadened in a reservoir ('peak broadening mode') both of which are described below.

In GC direct mode the GC eluent directly enters to the Orbitrap. In this mode, valves are initially set to route helium to the reservoir and the Orbitrap source. With this configuration, a user injects the sample and waits for the solvent to elute (for alanine derivative samples, using the conditions outlined above, this is the first 4 minutes after injection), and after the solvent elutes, the valves are changed to allow for the GC eluent to directly enter the Orbitrap (Figure 2.6a). A failure to vent the solvent creates large backgrounds. This mode can either be conducted on large mass windows (50-500 *m/z*) to note elution times of the diverse compounds that might be contained in an injected sample or on smaller mass windows (±5 *m/z*) similar to those used for isotope ratio measurements to recognize contaminants within the mass-range of interest and to identify the masses and relative intensity of marker peaks—those that occur before a peak elutes and from which a user can count to know when the peak of interest will elute—as they will occur during a measurement. This mode is not typically used for direct measurements.

Peak broadening mode enables a user to capture and measure peaks for minute to hour-long timescales. In this mode, the valves are initially set up such that helium is routed to the reservoir and then the source (Figure 2.7). After the solvent peak elutes, helium remains routed to the reservoir, but the GC eluent (routed to V2) enters the Orbitrap source. At this time, a user can use marker peaks (described above) to indicate when valves should be turned to capture the sample peak in the reservoir. Typically, one aims to capture a peak 60 seconds prior to its elution as the GC column fractionates eluting peaks. When required, a 30 second period before the elution of a peak can be used. To

capture a peak, the capillary carrying GC eluent is routed to the reservoir, which is routed to the Orbitrap source (Figure 2.6b). After the compound elutes and before another does, as determined by number of seconds after elution in peak scanning mode, helium is routed to the reservoir and the GC eluent is routed to V2 and to vent.

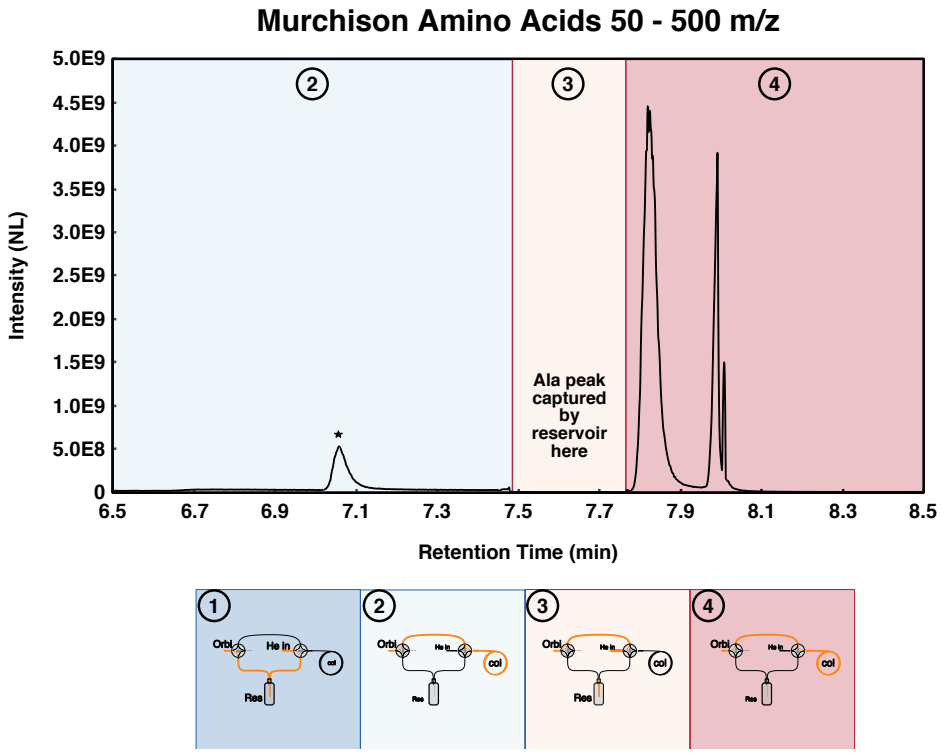


Figure 2.8: Example of peak capture timing verification. Here, the direct line from V1 is always routed to the Orbitrap. As before, the bottom image shows the different valve configures and the orange path highlights the path to the source. The first configuration on the bottom panel occurs prior to the time points selected to highlight the peak capture. The star denotes the marker peak that is used to time the third and fourth valve changes.

Contamination can lead to peak suppression (Eiler et al., 2017; Hofmann et al., 2020), so multiple methods are employed to confirm that a peak is captured without contamination. Firstly, we compare the integrated NL score of the monoisotopic peak to that captured in the reservoir and ensure that they are within a factor of 2. Secondly, we conduct the following experiment using a large mass window: following the solvent elution, we capture a peak as outlined above but we always have V1 routed directly to the source (Figure 2.8). During the peak's elution pure helium will enter the source and afterwards

the other peaks in the sample will enter the source, so a successful peak capture will only have the peak of interest missing from the chromatogram (Figure 2.8). This second experiment enables a user to determine if they are not capturing all of the peak of interest (in which case they might see some of the compound's elution curve before or after the period in which helium is routed to the source) or if they are capturing a contaminant peak (in which case they will see a peak or part of a peak missing before or after the period in which helium is routed to the source).

2.2.6.2. Fragment Selection

Fragmentation by electron impact is a high energy process, so we anticipate potential recombination reactions occur within the source that could impact the sites that are sampled by each fragment. The carbon sites that different fragments sample are identified by using a 1:9 mixture of ^{13}C-labelled alanine (at the C-1, C-2, or C-3 site) with unlabeled alanine that is derivatized into a N-trifluoro-O-methyl ester and diluted to 1:100,000 (parts per volume) in hexane and injected through the GC into the Orbitrap.

This derivative can be used either through direct injection to find potential peaks on a broad range (50-300 *m/z*) (Table 2.3) or in a more focused measurement to quantify the amount of each site that contributes to a given fragment.

2.2.6.3. Data Processing

All data files are processed through a proprietary software program provided by Thermo Fisher Scientific, 'FT statistic', which extracts information from a RAW file including intensity, peak noise and total ion counts for each scan pertaining to requested masses and converts them into a .csv file. Following extraction in FT statistic, isotopic intensities are converted to counts according to Eiler et al. (2017).

		Fractional ^{13}C abundance measured by NL for alanine with a 10% ^{13}C label at the site:				
Monoisotopic Mass(*m/z*)	**Stoichiometry**	unlabeled	unlabeled	C-1	C-2	C-3
65.01977	C2H3F2	0.0151	0.0161	0.0257	0.0270	0.0254
66.01501	CH2NF2	0.0089	0.0096	0.0073	0.0091	0.0084
68.99466	CF3	0.0112	0.0115	0.0112	0.0110	0.0108
70.02874	C3H4ON	0.0325	0.0327	0.0682	0.1193	0.0768
71.03207	C2(13C)H4ON	0.0006	0.0007	0.0035	0.0146	0.0077
71.04913	C4H7O	0.0302	0.0101	0.0315	0.0337	0.0405
71.08551	C5H11	0.0499	0.0473	0.0519	0.0507	0.0532
81.01465	C2H3OF2	0.0179	0.0148	0.0152	0.0196	0.0193
85.06478	C5H9O	0.0297	0.0120	0.0333	0.0409	0.0480
92.03062	C3H4NF2	0.0281	0.0308	0.0619	0.1118	0.0738
93.01466	C3H3OF2	0.0293	0.0297	0.0488	0.0404	0.0401
94.04132	C6H6O	0.0565	0.0373	0.0541	0.0531	0.0551
96.00554	C2HNF3	0.0203	0.0215	0.0312	0.0770	0.0445
102.05496	C4H8O2N	0.0384	0.0391	0.0740	0.1200	0.0811
113.02082	C3H4OF3	0.0285	0.0293	0.0484	0.0404	0.0409
129.01570	C3H4O2F3	0.0272	0.0284	0.0262	0.0272	0.0261
132.04546	C5H7O2NF	0.0391	0.0459	0.0759	0.1198	0.0842
138.01614	C4H3ONF3	0.0242	0.0176	0.0711	0.1082	0.0774
140.03173	C4H5ONF3	0.0417	0.0419	0.0751	0.1193	0.0813
149.02326	C8H5O3	0.0739	0.0629	0.0708	0.0696	0.0722
156.02653	C4H5O2NF3	0.0402	0.0417	0.0503	0.1079	0.0651
167.01873	C4H7O7	0.0427	0.0460	0.1263	0.1306	0.1114
167.01873	C5H4O2NF3	0.0427	0.0460	0.1263	0.1306	0.1114
184.02147	C5H5O3NF3	0.0485	0.0503	0.1142	0.1271	0.1037

Table 2.3: Peaks of interest for alanine on Orbitrap. Stoichiometry is predicted by Excalibur. The ^{13}R values are calculated directly from the chromatograms using the ion intensity. Two analyses are included for unlabeled alanine to give a reference for variability between measurements

We then convert the carbon isotope ratios to the Vienna Pee Dee Belemnite (VPDB) scale and perform a mass balance calculation that enables us account for contributions from the carbons in the derivatizing agents, which we assume are the same between derivatized samples and standards measured here (see Section 2.5). To this end we use Equation 2.1 as described in Chimiak *et al.* (2020):

$$^{13}R_{corr} = (((^{13}R_{sa,meas}/^{13}R_{st,meas} - 1) \times {}^{nC_{frag}}\!\Big/\!_{nC_{MOI}}) + 1) * {}^{13}R_{st,VPDB} \qquad (Eqn\ 2.1)$$

Where $^{13}R_{corr}$ is the ^{13}R values on the VPDB scale, $^{13}R_{sa,meas}$ and $^{13}R_{st,meas}$ are the $^{12}C/^{13}C$ of the sample and standard measured at the fragments-of-interest on the Orbitrap (which is described for the dual reservoir and single reservoir measurements below), nC_{frag} is the number of total carbons in the fragment, nC_{MOI} is the number of carbons that the molecule of interest (alanine in this study) contributes to the fragment, and $^{13}R_{st,VPDB}$ is the ^{13}R value of the standard in the VPDB scale as found by the molecular average and chemical degradation measurements described above in addition to measurements on a dual-focusing system (DFS) (Dallas et al., 2015). It is worth noting that the ratios are not the average of $^{13}C/^{12}C$ averaged across carbon sites as they are in conventional IRMS measurements. Instead, the $^{13}R_{sa,meas}$ and $^{13}R_{st,meas}$ here are the total abundances of the singly ^{13}C-substituted fragment ion to the unsubstituted fragment ion. Consequently, in a sample with natural abundance of ^{13}C at each site, the $^{13}R_{meas}$ values would be 0.011, 0.022, and 0.033, for a fragment that contains 1, 2, and 3 carbons, respectively.

For dual reservoir data, we find the $^{13}R_{sa,meas}/^{13}R_{st,meas}$ by first separating the scans according to which reservoir that was entering the source at the time of measurement. To this end, we plot scan number or retention time vs signal intensity (reported in NL score or calculated ion counts) and note the time points at which valve changes—which appear as spikes in the data—occur (Figure 2.4). All scans that occur 30 seconds prior to a valve change and 1 minute after are culled to minimize the chances of sample and standards mixing and of data being influenced by the valve turning. For the remaining data, for

each scan the ratio between the calculated ion counts of the monoisotopic and substituted peaks is recorded and these ratios are averaged for each sample or standard measurement across the time in which the reservoir is entering the source (not including the time periods that were culled). The isotope ratio of the sample to the standard is then calculated bracketed by the standard run before and after the sample as in Equation 2:

$$\frac{{}^{13}R_{sa,meas}}{{}^{13}R_{st,meas}} = \frac{\left({}^{13}C/_{12}C\right)_{avg,sa\ 2}}{0.5\left({}^{13}C/_{12}C\right)_{avg,st\ 1} + 0.5\left({}^{13}C/_{12}C\right)_{avg,st\ 3}} \qquad (Eqn.\ 2)$$

where $({}^{13}C/{}^{12}C)_{avg}$ is equal to the counts of the ^{13}C-subsitituted peak divided by that of the monoisotopic peak averaged during a measurement cycle and the numbers in the subscript refer to the measurement cycle as labelled in Figure 2.4.

For the single reservoir data, isotope ratios are calculated for each scan by dividing the counts in ^{13}C-substituted fragment peak by those in the unsubstituted fragment peak. These ratios are then multiplied by the intensity of the unsubstituted peak for the scan and divided by the sum of the intensities of the unsubstituted peak for the measurements being considered. These operations are done separately for measurements from the standard and the sample to calculate the $^{3}R_{st,meas}$ and the $^{13}R_{sa,meas}$, respectively. The $\delta^{13}C_{VPDB}$ is calculated with Equation 2.1 above. Here, parameters for data analysis of single reservoir measurements are explored in terms of the total ion count multiplied by the injection time (TICxIT), injection time (IT), and intensity of the unsubstituted peak relative to the TIC. The TICxIT can be used as an indicator of the system's stability as the IT adjusts inverse to the TIC (*e.g.,* higher TIC causes less time for C-Trap collection to meet the ion threshold and therefore lower IT). Rapid IT compensation for changes in TIC—that is when IT changes can keep up with TIC changes—will result in a stable TICxIT.

2.3. Results and Discussion

2.3.1. Ninhydrin Decarboxylation

The C-1 site of three alanine samples are constrained using ninhydrin decarboxylation. The $\delta^{13}C_{VPDB}$ of the C-1 sites for Alfa Aesar and VWR alanine standards are nearly equal at -28.5 ± 0.1 ‰ and 29.6 ± 0.1 ‰, respectively. The Strecker alanine standard's C-1 site has a $\delta^{13}C_{VPDB}$ of -43.5 ± 0.1 ‰, which approximately 15 ‰ below that of the other standards (Table 2.1). For each standard, the $\delta^{13}C$ values of the C-1 sites are invariant with respect to yield, reaction time, or the proportion of ninhydrin to alanine, so these factors did not likely impact the measured $\delta^{13}C$ of the C-1 sites. Instead, as the Alfa Aesar and VWR alanines were synthesized by microbial fermentation of acetate, differences between these alanines and Strecker alanine likely reflect a unique KIE imparted in Strecker synthesis that differs from the microbial fermentation of alanine.

Combining the $\delta^{13}C_{VPDB}$ of the C-1 sites of alanine with the molecular-average $\delta^{13}C_{VPDB}$ measurement, we calculate $\delta^{13}C_{VPDB}$ for the averaged C-2 and C-3 sites for all three standards (Table 2.1) as well.

2.3.2. Fragment Identification

Fragment identification provides information on the amount of the C-1, C-2, and C-3 sites that a given fragment contains. Here, labelling studies use 1:9 mixtures of ^{13}C-1, ^{13}C-2, or ^{13}C-3 labelled alanine with unlabeled alanine, so if a fragment contained 100% of the C-1 site, the C-1 labeled compound would have a $\delta^{13}C$ that is 11500 ‰ to 2800 ‰ enriched relative to the unlabeled alanine standard for a fragment that contains one to five carbons, respectively. With this analysis, we recognize candidate fragments listed in Table 2.3 and further explore which alanine sites contribute102.055, 113.021, 140.032, 184.021, and 200.053 fragments and their ^{13}C-substited versions. For this paper, we will

refer to pairs of fragments by the monoisotopic fragment's mass (see Figure 2.5 for images of mass spectra of monoisotopic fragments and ^{13}C-substited versions). Results from single and dual reservoir measurements of the labelled peaks are listed in Table 2.4. These results demonstrate that the 102.055 fragment contains the C-3 site and two derivative carbons, the 113.021 fragment contains the C-3 site and two derivative carbons, the 140.032 fragment contains the C-2 and C-3 sites in addition to two derivative carbons, the 184.021 fragment contains the C-1 and C-2 carbons in addition to three derivative carbons, and the 200 and 201 fragments contain all three carbons in alanine and all three derivative carbons. In each case, fragments returned either 0 % or 100 % of a given site's label, so they are each likely formed from a single mechanism.

Labelled site (label:unlabeled)	**C-1** (10.15:89.6)		**C-2** (10.369:90.9)		**C-3** (10.052:89.8)	
Mass	**% label returned**	*st err*	**% label returned**	*st err*	**% label returned**	*st err*
102	0%	*4%*	100%	*4%*	101%	*4%*
113	0%	*4%*	0%	*4%*	111%	*4%*
140	0%	*4%*	101%	*4%*	102%	*4%*
184	105%	*3%*	105%	*3%*	0%	*3%*
200	108%	*3%*	104%	*3%*	103%	*3%*

Table 2.4: Percentage of label returned at five fragments of interest. Percentages are computed as the $\delta^{13}C_{\text{Alfa Aesar}}$ of the labelled alanine at the fragment divided by the expected $\delta^{13}C_{\text{Alfa Aesar}}$ if 100% of the labelled alanine is present in the fragment.

2.3.3. Optimal Instrument Conditions During Measurements

Combining known $\delta^{13}C_{\text{VPDB}}$ values for the C-1 site, the C-2 and C-3 sites averaged, and the molecular average for the alanine standards, we optimize measurement conditions through dual-reservoir and single-reservoir studies. Dual reservoir mode studies explore AGC settings from 5 x 10^4 ions to 5 x 10^6 ions (Table 2.2). For *m/z* 135.5 to 146.5 and 170 to 210 scans, measurements with AGC values below 1 x 10^5 ions replicate the known differences between Strecker and Alfa Aesar alanine. Differences in microscans do not

impact the data beyond error as analyte enters the source as a constant stream so has no significant scan-to-scan variation.

Dual reservoir studies find the following structure of Strecker alanine relative to Alfa Aesar -17.0 ± 8.6 ‰, -9.8 ± 3.9 ‰, and -21.2 ± 1.4 for the C-1, C-2, and C-3 sites, respectively. The error follows the shot noise limit and can reach 1 ‰ precision within 2 hours for each site.

Single-reservoir studies enable three key observations: (1) the impact of maximum IT on results, (2) the impact of microscan binning on results, and (3) the impact of the intensity of the peak relative to the background. Data from scans that were conducted at maximum IT were highly fractionated. Consequently, we set maximum IT to 3000 ms and only use scans in which the IT remains below this value. To keep scans below this value, we either increase the concentration of analyte or change the reservoir size because smaller reservoirs increase the concentration of analyte entering the Orbitrap source.

When the instrument functions properly, different microscan binning does not impact $\delta^{13}C$ values (Table 2.2). However, binning microscans together can hide scans in which IT maxes out or in which there are instabilities such as electronic noise. Given that data is processed using software, we find no disadvantages to having each scan consist of a single microscan and so conduct scans using 1 microscan binning when possible.

Finally, because the sample concentration in the reservoir decreases over time, the intensity of the sample peak relative to background in the instrument (when the background fragment ions are approximately constant over the timescales of the measurements) also decreases over time. Consequently, the proportion that the sample contributes to the total ion current of a given scan decreases over the course of a measurement. When this occurs, the less abundant peak will experience high ion suppression than the more abundant one, which can lead to ^{13}R measurements that are lower than the true values. This effect is discussed at length in Hoffman et al (2020) for mixtures. For example, for alanine and alaninamide, measurements in which alanine is

less than 25 % of the mixture, the measured ^{13}R is fractionated by more than 40 ‰ relative to a pure alanine sample. We find similar effects in the tails of peaks suggesting that when background contribute to more than 70 % of the total ion counts, the ^{13}R values of the analyte peaks are suppressed.

2.3.4. Fractionation during peak capturing

Capturing peaks isolates them for measurement. The current system requires manual valve control, so we test the effect of the timing before a peak was captured on the measured ^{13}R value (Table 2.5). Peaks in which the valve is turned before 30s prior to the peak's elution have relative ^{13}R values within 2 standard error, so data in these conditions can be compared to one another. However, for this data we note a far larger scatter as the timing is closer to that of the peak's elution. This effect is potentially due to the strong ^{13}C fractionation along a GC column that might result in the closer valve turns to miss small amounts of the eluted analyte in an inconsistent fashion.

Date	Into reservoir (min from start)	Out of reservoir (min from start)	^{12}C Counts	^{13}R	st err
6/15/18	6.8	7.8	8246855	0.04259	0.00009
6/18/18	6.8	8	13850451	0.04255	0.00007
6/18/18	6.8	7.8	1439061	0.04220	0.00021
6/18/18	6.8	7.8	31183238	0.04215	0.00005
6/19/18	6.8	8	41351194	0.04241	0.00004
6/19/18	6.8	8	73704467	0.04284	0.00003
6/19/18	6.8	8	19179735	0.04192	0.00006
6/19/18	6.8	8	19865222	0.04192	0.00006
6/20/18	6.8	8	30611362	0.04198	0.00005
6/20/18	6.8	8	33731557	0.04194	0.00005
6/20/18	6.8	8	43209415	0.04201	0.00004
6/21/18	6.8	9	37255595	0.04219	0.00004
6/21/18	6.8	8.3	44976517	0.04190	0.00004
6/21/18	6.5	8.14	36913425	0.04202	0.00004
6/22/18	6.7	8	34322676	0.04231	0.00004
6/22/18	6.5	8	37735466	0.04209	0.00004
6/25/18	5.5	11	33429294	0.04200	0.00005

6/25/18	5.5	11	73130052	0.04211	0.00003
6/25/18	6.74	11	82939344	0.04235	0.00003
6/25/18	6.74	11	42122402	0.04219	0.00004
6/25/18	6.64	7.74	70095285	0.04223	0.00004
6/25/18	6.64	7.74	73703342	0.04219	0.00003
6/26/18	6.64	7.74	13625942	0.04243	0.00007
Date	Into reservoir (min from start)	Out of reservoir (min from start)	^{12}C Counts	^{13}R	st err
6/26/18	6.64	7.74	16294948	0.04205	0.00007
6/26/18	6.64	7.74	45428987	0.04270	0.00004
6/27/18	6.64	7.74	46702323	0.04243	0.00004
6/27/18	6.64	7.74	82419792	0.04253	0.00003
6/27/18	6.64	7.74	82644301	0.04249	0.00003

Table 2.5: Variations in the ^{13}R of the 140.031 peak for Alfa Aesar alanine captured with different valve timings.

2.3.5. Sample Size

Typically, we aim to have the monoisotopic peak as the most abundant peak in its measurement window. Therefore, when sample size is not a limitation, increasing the concentration of sample injection and decreasing the measurement window is recommended.

However, natural samples are often limited in size. Two methods can combat size limitations, especially when backgrounds start to dominate a spectrum: (1) decrease the mass range being measured and (2) decrease the reservoir size. While decreasing the measurement window can eliminate background peaks, it also decreases the ion fluence and therefore decreases the measurement precision per sample amount. When the reservoir size decreases, the concentration of analyte entering the source at a given time increases which also increases the number of scans in which background interference on the measured ^{13}R value is less significant. This latter solution results in less peak broadening and therefore less time during which the peak could be measured, but it can lead to more scans that can be used in a given measurement.

To constrain the lower bound of the sample size whose isotope ratio can be accurately measured in the Orbitrap, we examine how ^{13}R values change with changes in the proportion that a fragment's monoisotopic peak contributes to the TIC. Doing so allows us to understand the minimum ratio of a fragment's ion count over the TIC that is permissible for accurate measurements. Because backgrounds are relatively constant when the source and transfer lines are clean—which is the most likely condition in which to analyze microsamples—this minimum ratio can also be used to define the lowest sample concentration that one can use for each fragment analyzed.

To this end, we compare data collected from 25 picomole samples of Alfa Aesar and Strecker alanine derivatives that were analyzed in peak broadening mode using a single 30 mL reservoir and the 135.5 to 145.5 Da mass window. For both the Alfa Aesar and Strecker alanine derivative data sets, we cull scans from the data in six manners: we use the ratio of the monoisotopic peak's intensity to the TIC and cull data where the ratio is less than 0 (use all data), 0.1, 0.2, 0.3, 0.4, and 0.5 (Figure 2.9). While the $\delta^{13}C$ of the Strecker alanine relative to the Alfa Aesar alanine is within error of the measured value for all conditions, we note that the value decreases as culling threshold for the monoisotopic peak's intensity to the TIC increases from 0.1 to 0.5 (Figure 2.9). This decrease could be due, in part, to non-linearity in the measured ^{13}R with the size of sample as the experiments use manual injection, which can vary between measurements.

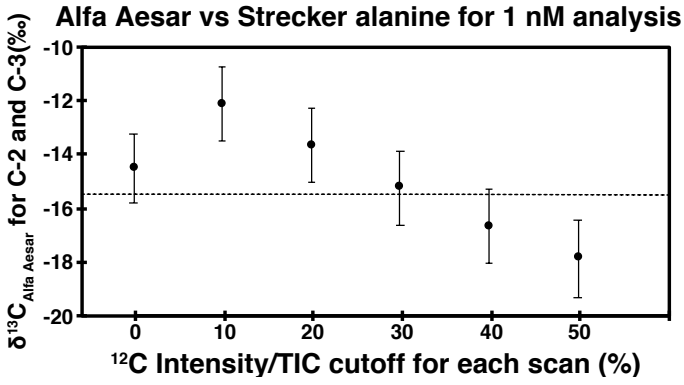


Figure 2.9: Alfa Aesar vs Strecker alanine samples relative ^{13}R values as the minimum ^{12}C/TIC minimum cutoff for a scan to be included in analysis changes. Here all alanine samples were 25 picomoles of derivative and were measured in a 30 mL reservoir. Error bars are 1 standard error for the measurement. Dotted line is the true value of $\delta^{13}C$ of Strecker relative to Alfa Aesar alanine for the averaged C-2 and C-3 sites.

To understand how the ^{13}R of a sample changes with the amount injected, we compare the ^{13}R for the 140.031 peak from two 1.25 nM samples (3 picomoles of analyte in a 2 mL reservoir) and one 0.5 nM sample (1 picomoles of analyte in a 2 mL reservoir) with the ^{13}R from a 10 nM sample (100 picomoles in a 10 mL reservoir). The data demonstrate the ^{13}R suppression anticipated with the decreased ratio of analyte to background: as the scans are culled to permit only those in which the monoisotopic peak contributes 20 %, 30 %, and 40 % of the total ion count, the ^{13}R value of the microsample approaches that of the normal sample. The smallest sample's ^{13}R value is within 2 standard error of the normal sample when its data are culled with a monoisotopic peak contribution minimum cutoff of 30 % (Figure 2.10). This trend can also be seen when the same concentration analyte differs in injection quantity, leading to a change in ^{13}R with the integrated TIC for a run that decreases as data are culled with a monoisotopic peak contribution minimum that reaches 30 % (Figure 2.11).

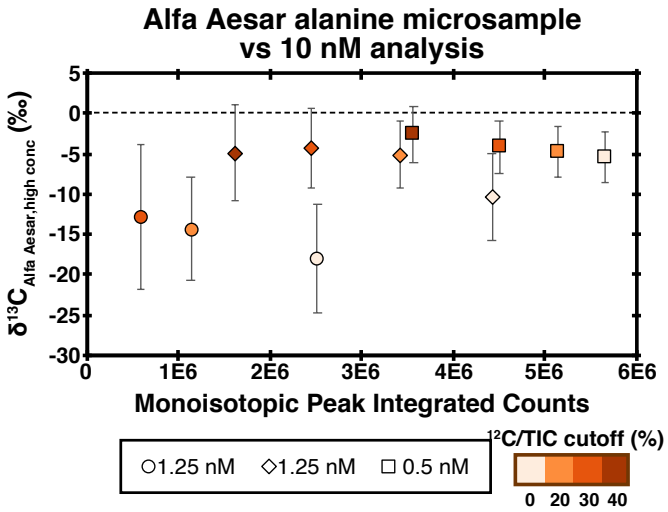


Figure 2.10: Comparison of Alfa Aesar alanine sample's ^{13}R values for the 140.024 peak. All $\delta^{13}C$ values are taken as the diluted sample relative to the normal sample (a 2uL sample dilute 1:100,000 in hexane using a 10 mL reservoir). The two 1.25 nM samples are comprised of injections of 3 picomoles of alanine derivative into a 2 mL reservoir and the 0.5 nM sample uses 1 picomole of derivative and a 2 mL reservoir. Symbols represent different analyses and shading denotes different ^{21}C intensity/TIC cutoffs during data processing. Relative ^{13}R values decrease with decreasing concentration and increase with increasing TIC cutoff threshold.

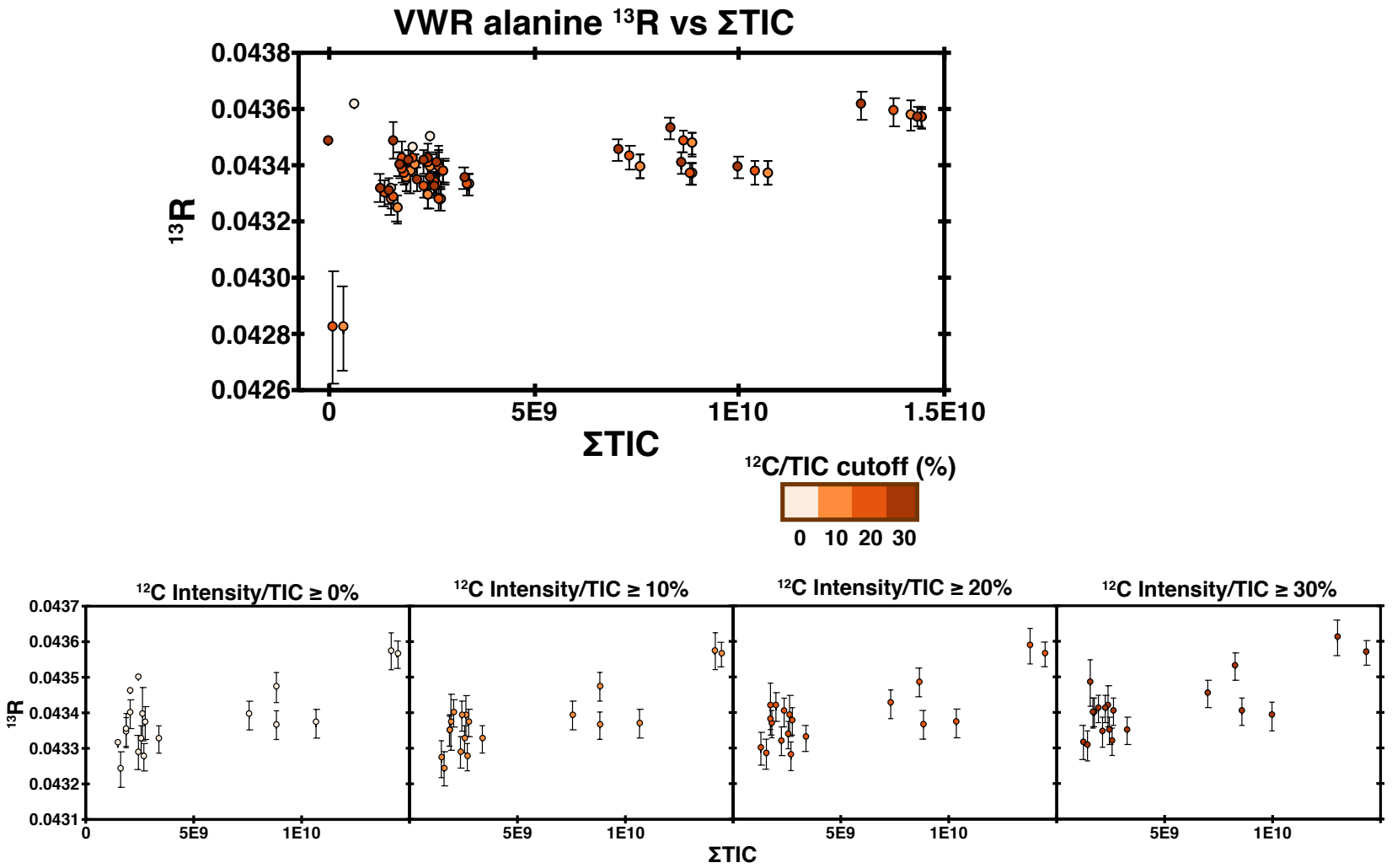


Figure 2.11: TIC cutoff effect on VWR alanine ^{13}R values. The lowest intensity sample does not have any measurements above the TIC cutoff 30 panel. Error bars are 1 standard error. Lowest ΣTIC data point for 0 and 30 % cutoffs has error bars that extend beyond the top margin of the graph so were omitted. Lower panel is a close up of the same data without the lowest ΣTIC data point with the data sets separated by the minimum ^{12}C Intensity/TIC that any scan needed to have to be included in the analysis.

2.3.6. Stability

The stability of ^{13}R values depend on the condition of the ion source. Over a sixteen-day period during which multiple analytes were measured, there is no evidence of drift within 2 standard error in measured ^{13}R values of the VWR alanine standard. Consequently, measurements over the period of weeks should be comparable to one another (Figure 2.12).

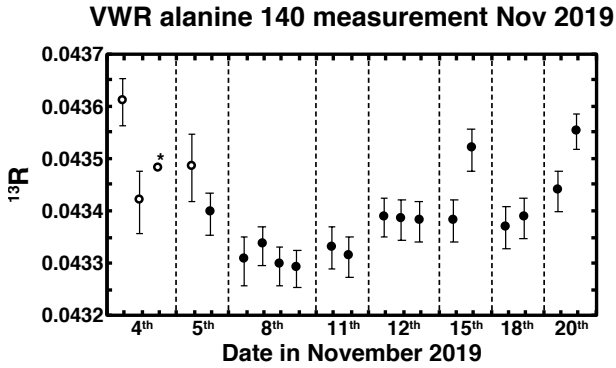


Figure 2.12: Demonstration of stability of ^{13}R values over a sixteen-day measurement period. All measurements were taken in the single-reservoir configuration from the same container of derivatized alanine. Open symbols denote measurements in which the TICxIT were oscillating between two regimes. Starred symbol has error bars beyond the y-axis. Error bars represent 1 standard error from the measurement.

2.3.7. Data Processing

Once a sample has been measured, the intensities are converted to counts through FT Statistic (See Methods for further details). The studies conducted here provide guidance on data processing methods and are outlined for normal samples and microsamples in Figure 2.13.

2.3.7.1. Non-limited samples

When measurements are not limited by sample size, only acquisitions in which the following requirements are met are considered: (1) at its highest intensity (its elution's peak from the reservoir), the unsubstituted peak is at least 50 % of the total ion intensity, (2) the product of TICxIT does not vary by more than 20 %, (3) the peak is captured at

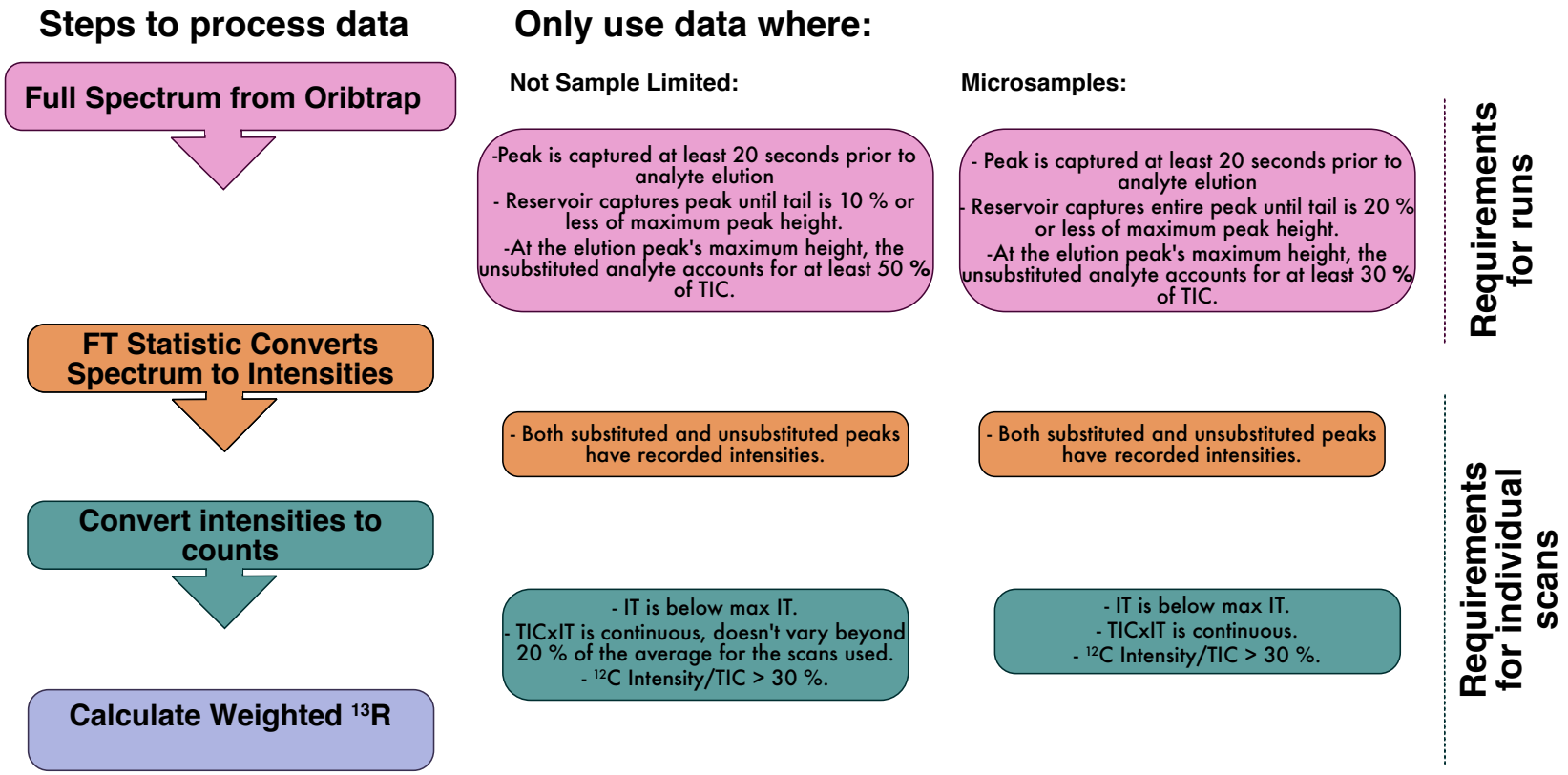


Figure 2.13: Data processing for conditions in which the sample does not limit measurements and those in which the sample needs to be conserved.

least 20 seconds prior to its elution. For data sets in which we would need to relax the third rule, we would do so and only include peaks for which the timing of the peak capture is within 5 seconds of each other and is at least 20 seconds prior to peak elution.

In these data sets, we first convert intensities to counts following equation 3 (see Methods: Data Processing for more details). In the resulting data set, we only include scans for which both fragments being considered have measured values, in which the IT remains below its maximum, and in which the TICxIT does not varies beyond 30 % of its median value. Finally, we exclude any scans for which the monoisotopic peak is less than 30 % of the total ion intensity. For many large samples, these additional constraints have no impact on the isotope ratio.

2.3.7.2. Microsamples

For measurements that are limited by sample size ('microsamples'), acquisitions must meet the following requirements to be considered: (1) the unsubstituted peak is at least 30 % of the total ion intensity at its elution peak from the reservoir, (2) the product of TICxIT is continuous during the elution of the peak from the reservoir—that is, from scan to scan it does not vary by more than 30 %—and (3) the peak is captured at least 30 seconds prior to its elution.

To process these acquisitions, we convert intensities to counts (Equation 2.3, Methods: Data Processing). In each acquisition, scans for which IT reaches the maximum value set by the user are culled. In the remaining data set scans are culled when the main monoisotopic peak is less than 30 % of the TIC and for the period before the TICxIT is continuous (defined here as exceeding 30% of a 10 scan median value) at the beginning of an acquisition and after the TICxIT is continuous at the end of an acquisition. With these parameters, the difference in ^{13}R between the averaged C-2 and C-3 sites for Strecker and Alfa Aesar alanine can be reproduced (Figure 2.9).

3. Conclusion

Isotopic structures can elucidate elementary steps in chemical reactions or constrain the identities of reactant pools in a product. Site-specific isotope ratio studies have already demonstrated their value in constraining the origins of vanillin and wine. Developing

methods that enable the measurement of site-specific isotope ratios on pico-to-micromole-sized samples will open the field of site-specific isotope studies.

We present a method by which to constrain the C-1, C-2, and C-3 sites in alanine and demonstrate the ability to reach 1 ‰ precision within 2 hours per site. While smaller samples will have lower precision, we demonstrate the ability to measure alanine samples as small as 1 picomole for site-specific isotope ratios for the C-1 site. Combined with conventional methods such as GC-C-IRMS, these analyses allow for natural samples attain site-specific isotope structures.

Chapter 3

ISOTOPE EFFECTS AT THE ORIGIN OF LIFE:

FINGERPRINTS OF THE STRECKER SYNTHESIS

Abstract

Strecker synthesis is widely hypothesized to be a key mechanism in the chemistry that led to life on Earth and potentially on other planets. It uses prebiotically plausible substrates to create α-amino acids. Because α-amino acids are widely used in life, understanding where they form abiotically can help constrain where life could form. Therefore, to aid in the search for extraterrestrial life by finding signatures of biogenic amino acids and to better constrain the synthetic environments and precursors of abiotic α-amino acids, we measured the molecular-average and site-specific isotope effects of carbon and nitrogen during the Strecker synthesis of alanine. The reaction steps of the Strecker synthesis can be divided into two groups: an initial series of reversible amination and nitrile forming reactions ('equilibration') and a second series of irreversible hydrolysis reactions. The equilibration of cyanide, acetaldehyde, and ammonia with the intermediate, α-aminopropionitrile (α-APN), has a measured +56.6 ‰ equilibrium nitrogen isotope effect between the ^{15}N-rich amine nitrogen in α-aminopropionitrile and ^{15}N-poor ammonia, and a -20.0 ‰ equilibrium carbon isotope effect between the ^{13}C-poor C-2 site in α-aminopropionitrile and the ^{13}C-rich carbonyl carbon in acetaldehyde. The first irreversible hydrolysis step is inferred to have a normal isotope effect (*i.e.,* faster for ^{12}C, slower for ^{13}C) that could be up to -10 ‰ but also has one or more side reactions that deplete the residual α-APN reservoir by 15 ‰. The second hydrolysis step has a -15.4 ‰ normal kinetic isotope effect on the amide (C-1) site of alaninamide, which becomes the carboxyl site of alanine. This step had no recognized side reactions. Other α-amino acids will likely experience similar nitrogen isotope fractionations between ammonia and their amine sites, and similar carbon isotope fractionations between the carbonyl carbon in reactant aldehydes or ketones and the intermediate α-aminonitrile, and

between cyanide and the carboxyl site. These isotope effects allow us to predict the carbon and nitrogen isotopic contents and intramolecular structures of α-amino acids based on known isotopic compositions of substrates, or to infer the isotopic compositions of substrates from which amino acids formed, for example in the case of the amino-acid-rich carbonaceous chondrites. The site-specific C and N isotope contents of amino acids formed by Strecker chemistry contrast with those typical of terrestrial biosynthetic amino acids, so these data also provide a means of forensically discriminating between biogenic and abiogenic sources.

3.1. Introduction

In Strecker synthesis, simple abiotically-available precursors react to form α-amino acids. More specifically, an aldehyde or ketone reacts with ammonia and cyanide to form an α-aminonitrile that is subsequently hydrolyzed into an α-amino acid (Van Trump, 1975) (Figure 3.1).

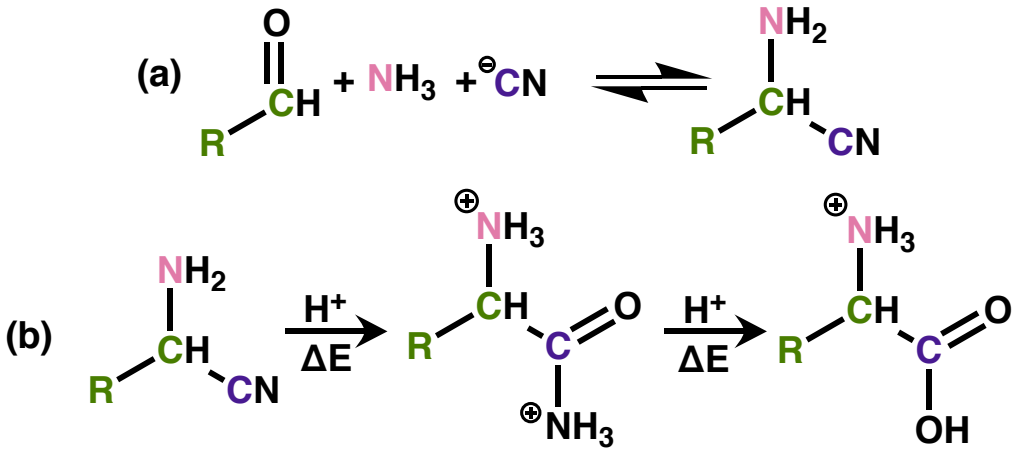


Fig 3.1: Simplified version of Strecker synthesis mechanism. Step (a) is the low temperature equilibrium step. Step (b) depicts the two hydrolysis steps (H1 and H2). Each hydrolysis step is denoted with the forward arrow.

Strecker synthesis is one of the simplest routes by which abiotically available precursors produce α-amino acids, the subset of amino acids most prevalent in known life.

Consequently, it is often hypothesized to have contributed to the prebiotic chemistry that facilitated the origins of life on Earth and potentially in extraterrestrial settings. The reactants for Strecker synthesis have been detected in the interstellar medium (ISM) and on meteorites (Ehrenfreund and Charnley, 2000; Pizzarello and Shock, 2010). These findings support the hypothesis that Strecker synthesis occurs in the ISM, early solar nebula and/or meteorite parent bodies. Delivery of these reactants to hydrous environments on planetary bodies also could lead to Strecker chemistry on those planets. Even without the delivery of reactants, planetary bodies could have Strecker synthesis occur in their atmospheres or near-surface waters, as the synthesis is a mechanism for α-amino acid production in Miller-Urey-style spark-discharge experiments, which simulate lightning in neutral or reduced atmospheres (Miller, 1957; Pizzarello and Shock, 2010).

On meteorites, the reactants of Strecker synthesis, such as aldehydes and ammonia, and potential products, α-amino acids, have stable isotope abundance ratios of H, N and C (*i.e.,* D/H, $^{13}C/^{12}C$ and $^{15}N/^{14}N$) that are higher than those observed in common terrestrial materials and in the bulk average (*i.e.*, at hand-sample scale) meteoritic organic matter. This shared heavy isotope enrichment has been interpreted as evidence that the meteoritic α-amino acids are derived from the co-occurring Strecker reagents and thus plausibly formed by the Strecker mechanism (Pizzarello et al., 1994; Elsila et al., 2012; Aponte et al., 2017; Simkus et al., 2019). Furthermore, recent site-specific isotope ratio (SSIR) measurements of an alanine sample from the Murchison meteorite identified that the α-carbon (*i.e.*, $HO_2C—H\underline{\mathbf{C}}NH_2—CH_3$) hosts a pronounced ^{13}C enrichment — a pattern consistent with derivation by Strecker chemistry from acetaldehyde with a ^{13}C-rich carbonyl carbon (Chimiak et al., 2020). This finding is also consistent with a pattern of decreasing $^{13}C/^{12}C$ with increasing carbon number among the C2+ aldehydes and α-amino acids in Murchison, and thus re-enforces the argument that alanine in that sample formed by Strecker reactions involving aldehyde precursors that are also found in that sample (Elsila et al., 2012; Aponte et al., 2017; Chimiak et al., 2020). The argument that Strecker synthesis created α-amino acids in meteorite parent bodies —rocky bodies

containing water, simple organics, and (presumably) no life—bolsters its potential as a pathway to form amino acids on early Earth and other planets.

Past experimental and theoretical work on Strecker synthesis has detailed the effect of physiochemical conditions on its reaction mechanism and kinetics. These studies show that Strecker synthesis occurs in aqueous phase, at basic or acidic pHs, and likely has a water-catalyzed hydrolysis step (Van Trump, 1975; Moutou et al., 1995). However, despite the importance of the Strecker synthesis to the study of prebiotic chemistry and the role of stable isotope data in recognizing its products in natural samples, we are not aware of any prior published studies of its isotope effects. This paper documents the carbon and nitrogen isotope effects during Strecker synthesis of alanine and includes constraints on site-specific carbon isotope effects. These findings provide a framework for recognizing and quantifying the products of Strecker synthesis in meteorites and other natural samples and environments (*i.e.,* Mars, Europa, or terrestrial hydrothermal systems).

3.2. Materials and Methods

3.2.2. Materials

3.2.2.1. Strecker synthesis

All Strecker syntheses performed in this study used sodium cyanide (Sigma Aldrich, 97 % purity, Item # 380970, Lot # MKBX0939V), acetaldehyde (>99.5 % purity, Sigma Aldrich, Item # 402788, Lot # SHBG8084V), and ammonium chloride (99.5% purity, Mallinckrodt Chemicals, Item # 3384-12, Lot # C43615). Ultrapure water was obtained from a MilliPore ultrahigh-purity water system (18.2 MΩ cm; hereafter ‘water’). Hydrochloric acid (36.5-38 wt. %, Macron Fine Chemicals, Item # 251546) was diluted to produce 6N HCl on the same day as each synthesis.

Strecker syntheses were conducted in heavy-walled 5 mL borosilicate vials (Sigma Aldrich, Item # 29361-U) with new, sterile PTFE stir bars (VWR). Between uses, vials were washed with water and combusted at 450 °C. Vials were capped by 20mm PTFE/Silicone septa (Supelco, 27237-U, lot), which were changed between syntheses. Liquid reagents were added with sterile Luers-lock plastic syringes (BD, multiple lots). Water and acetaldehyde were added with 23-gauge needles (BD, Item # 305145) and HCl was added using a 20-gauge PTFE needle (Aldrich, Z117315) threaded through a 14-gauge needle (Monoject, 2021-01). Holes in septa were sealed with Krytox LVP (Du Pont, Z273546), a non-reactive thermally stable lubricant.

3.2.2.2. Purification

The chlorine salts of alanine and alaninamide produced in Strecker syntheses were desalted on a cation-exchange resin column (AG50W-X8, 100-200 mesh, hydrogen form, Bio-Rad) using 2M ammonium hydroxide diluted from 28.0-30.0 % (J.T. Baker, multiple lots). Alanine and alanine amide were separated using methanol (MeOH; >99.8 % purity, Macron Fine Chemicals, multiple lots).

3.2.2.3. Analytical Standards

Two alanine standards were used in this study: VWR alanine (Purity >99 %, Lot # 2795C477) and one sample of alanine synthesized by us via Strecker synthesis in 2015 (Purity confirmed by NMR).

3.2.2.4. Derivatization

Alanine was derivatized prior to isotopic analysis to enhance volatility and to enable GC separation and introduction to a mass spectrometer electron impact ion source. Anhydrous methanol (MeOH; >99.8 % purity, Macron Fine Chemicals, multiple lots), n-hexane (>98.5 % purity, Millipore Sigma, HPLC grade, multiple lots), acetyl chloride (AcCl; >99.0 % purity, Sigma Aldrich, Item # 00990, Lot # BCBT8141) and

trifluoroacetic anhydride (TFAA; >99 % purity, Sigma Aldrich, Item # 106232, Lot # SHBJ0051) were used to derivatize alanine to alanine-N-trifluouroacetyl-O-methyl ester.

Derivatization occurred in new, sterile gas-chromatography borosilicate vials (Microsolv, Item # 9502S-VWC) with new caps. All other glassware used in derivatization was cleaned in DI water and combusted at 450 °C prior to use. Chemical lab syringes (Hamilton, 250 µL) were cleaned with methanol prior to and after derivatization reactions, and instrument inlet syringes (Hamilton, 10 µL) were cleaned with 30 µL hexane between and before all analyses.

3.2.3.Methods

3.2.3.1. Syntheses

Alanine samples were synthesized following methods from Kendall *et al*. (Kendall et al., 1929) modified for the small scale and closed system used in this study. For convenience, throughout this paper we refer to the first step of the Strecker synthesis, which involves relatively low-temperature, mostly reversible reactions among aldehyde, ammonia, and cyanide precursors as the 'equilibration step', and the second step, which here involves higher temperature hydrolysis of intermediates in an acidic medium, as the 'hydrolysis step' (Figure 3.1). We note that the hydrolysis step can also occur in basic media (O'Connor, 1970; Van Trump, 1975).

Solid ammonium chloride (approximately 163 mg) and sodium cyanide (approximately 90 mg) were weighed in a reaction vial to which a PTFE stir bar was added. Vials were capped, and 150 µL acetaldehyde and 1 mL water were added to the vial via syringe. The puncture site in each reaction vial cap was sealed with Krytox, and then vials were put on a stir plate for four hours. Syntheses performed with an initial 'equilibration' step at room temperature (22.5 to 24 °C) sat directly on the stir plate; those at 0 °C were placed in an ice bath on the stir plate; and that at 54 °C was placed in a heated sand bath on the stir

plate. Table 3.1 provides exact amounts of reactants and notes of temperature conditions for each synthesis.

After four hours, samples were removed from the stir plate and hydrolyzed with 6N HCl. We needed to avoid iron contamination from the syringe needles while adding HCl, so the 1 mL of added 6N HCl was measured in a sterile plastic syringe with a PTFE needle attached via the Luers lock. This needle was threaded through a 14-gauge stainless steel needle. Both needles entered the vial's cap and headspace together, and the PTFE needle was threaded below the stainless-steel needle before HCl was added to the solution. Krytox was then spread around the exterior needle and upon removal of the needles smeared over the opening to prevent gas escape.

Sa	T(°C)		mmol reactants			mmol products	
	EQ	Hyd	CN^-	NH_4^+	ace	amide	ala
1	23	83	2.11	3.32	2.7	1.11	0.47
2	23	83	2.08	3.00	2.7	1.23	0.33
3	23	83	2.02	3.03	2.7	1.26	0.44
4	24	85	1.93	2.90	2.7	1.38	0.89
5	23	83	1.97	2.98	2.7	1.36	1.36
6	24	93	2.12	2.99	2.7	1.29	0.60
7	24	93	2.12	3.00	2.7	1.34	0.47
8	54	95	1.88	3.02	2.7	1.36	0.88
9	23	105	1.95	3.11	2.7	1.47	0.52
10	23	105	2.04	3.08	2.7	1.27	0.27
11	23	105	2.07	3.01	2.7	1.04	0.12
12	25	110	2.00	3.07	2.7	1.15	0.50
13	25	110	1.99	3.07	2.7	1.44	0.84
14	0	110	2.32	3.23	2.7	1.19	0.31
15	0	110	1.94	3.12	2.7	0.88	0.05
16	25	115	2.12	2.99	2.7	1.65	1.14
17	25	115	2.08	2.99	2.7	1.48	0.27
18	25	115	2.21	3.14	2.7	1.87	0.79
19	24	120	1.85	2.94	2.7	0.45	0.12
20	24	120	2.04	3.09	2.7	0.65	0.50

Table 3.1: Experimental conditions and moles of yields. Errors are as follows: 0.5°C for temperature, 0.02 mmol for initial mmol of cyanide and ammonium, 0.1 mmol for initial mmol acetaldehyde, and 0.01 mmol for final mmol alaninamide and alanine.

Samples were then stirred and heated for one hour in a sand bath at temperatures ranging from 80 °C to 120 °C (Table 3.1) for the hydrolysis step. After hydrolysis, samples were removed from heat and vials were attached to a bubbling apparatus by two needles. One needle carried nitrogen gas that was bubbled into samples while the other needle carried away any released gas into a 2M NaOH solution that would capture any residual HCN. Needles were attached to the gas lines and to each other with Tygon tubing. Bubbling occurred for 2 minutes after which samples were dried under nitrogen overnight.

Dried samples were desalted on a BioRad column using 2M NH_4OH as the eluent, and fractions were collected in test tubes. Each sample fraction was tested for the presence of an amine group via thin layer chromatography with a mobile phase of 1:1:1:1 butanol, ethyl-acetate, acetic acid, water solution paired with a ninhydrin reagent (1.5 g ninhydrin dissolved in 100 mL butanol and acidified with 3.0 mL glacial acetic acid) (Pirrung, 2007). For a given sample, fractions that tested positive for amines were combined and subsequently dried under nitrogen. Samples were analyzed with proton NMR and liquid chromatography mass spectrometry, which identified masses 87 and 89 for alaninamide and alanine, respectively. Methanol was added to dried samples, gently shaken, and pipetted into a new vial once particulates had settled to the bottom of the vial. Alaninamide fully dissolved in the methanol layer while alanine remained undissolved. Alanine vials were washed three times in this manner and then both vials were dried under nitrogen. Alanine samples were tested for purity via proton NMR.

3.2.3.2. Isotope Ratio Measurements

Molecular-average $\delta^{13}C$[3.1] of alanine, alaninamide, and sodium cyanide and the $\delta^{15}N$ of alanine and ammonium chloride were measured on a Thermo Fisher Scientific Flash Elemental Analyzer (EA) coupled to a Delta-V Isotope Ratio Mass Spectrometer (IRMS).

[3.1] Delta notation, $\delta^{h}X$, describes the ratio of a heavy isotope of element X (^{h}X) to the light isotope (^{l}X) relative to a standard. It is computed by : $\delta^{h}X = \left[\frac{^{h}R_{sa}}{^{h}R_{st}} - 1\right]$ where ^{h}R is equal to the ratio of the heavy-to-light isotope ($\frac{^{h}X}{^{l}X}$). The delta value is often multiplied by 1000 and reported in units of per mil (‰).

Samples were analyzed over a two-day period throughout which standards and blanks were analyzed to calibrate and detect possible drift in the instrument measurements. For $\delta^{13}C$ measurements, measurements of acetanilide standard were interspersed throughout the analysis period and a single measurement of urea and sucrose each occurred at the end of the first day of analysis. Carbon isotope standards had measured $\delta^{13}C_{VPDB}$ values and standard errors of -28.00 ± 0.42 ‰ (acetanilide), -10.98 ‰ (sucrose, 1 measurement), and -27.67 ‰ (urea, 1 measurement) and acetanilide $\delta^{13}C$ values did not drift during the measurement period during which we analyzed alanine and alaninamide samples. These values are within error of recommended values of -27.7 ‰ for acetanilide, of -27.8 ‰ for urea, and almost within error of the -10.47 ± 0.13 ‰ of sucrose. Recommended values for acetanilide and urea were found by calibration against NIST sucrose standard, which is also the sucrose standard used here.

Measurements of $\delta^{15}N$ were standardized by comparison with a potassium nitrate standard as well as the aforementioned alanine and acetanilide standards. Our measurements of $\delta^{15}N_{AIR}$ for these nitrogen standards had means and standard errors of 3.12 ± 0.13 ‰ (acetanilide) and 5.78 ± 0.16 ‰ (potassium nitrate) during the measurement period, which are 1 ‰ higher than the value given by NIST for NIST potassium nitrate reference material RM 8549 and the recommended value for acetanilide found by calibration to the potassium nitrate. Consequently, we subtracted 1 ‰ from our $\delta^{15}N$ measurements of our samples (all data reported below reflect that correction).

Molecular-average $\delta^{13}C_{VPDB}$ of acetaldehyde was measured on a Thermo gas chromatography-combustion-isotope ratio mass spectrometer (GC-C-IRMS) over the course of one day. A mixture of eight fatty acids (F8 mix) was used to check the instrument's precision prior to acetaldehyde measurements. The F8 mix consisted of a C_{14}-methyl ester, C_{14}-ethyl ester, C_{16}-methyl ester, C_{16}-ethyl ester, C_{18}-methyl ester, C_{18}-ethyl ester, C_{20}-methyl ester, C_{20}-ethyl ester with $\delta^{13}C$ values of -29.98 ‰, -29.13 ‰, -29.90 ‰, -30.92 ‰, -23.24 ‰, -28.22 ‰, -30.68 ‰, and -26.10 ‰, respectively. Measurements were made relative to a laboratory tank of CO_2 with a $\delta^{13}C$ value of -12 ‰. To make the acetaldehyde measurements, the sample was diluted in hexane 1

part in 1000 and injected into a GC held at 80°C. Due to its fast elution, the acetaldehyde mixture was injected 180 seconds after the start of the GC method and the inlet port from the GC was manually vented after the elution of the acetaldehyde peak in order to prevent the subsequent hexane peak from entering the source. We only use data in which we waited 20 minutes between each injection to avoid contamination from the hexane peak's tail.

Techniques for site-specific isotope ratio (SSIR) measurements are outlined in Chimiak et al. (2020) and Eiler et al. (2017). Briefly, alanine was derivatized as the N-trifluouroacetyl-O-methyl ester, diluted to $1{:}10^5$ concentration in hexane, and analyzed on a Thermo Fisher Q-Exactive GC-Orbitrap ('Orbitrap') with a modified inlet system that includes a reservoir to enable peak trapping and broadening. When the peak for derivatized alanine from a sample or standard elutes from the GC it enters the trapping reservoir, from which it is purged by helium over tens of minutes. Each slowly eluting peak was measured for isotope ratios of interest for 30 minutes to one hour, where the Advanced Quadrupole Selector mass filtering system—a filtering system that operates similarly to a quadrupole mass spectrometer (Eiler et al., 2017)—was set to pass to the Orbitrap only ions in the window *m/z* of 135.5 to 145.5. This window permits observation of the 140.031 *m/z* peak ($^{12}C_4H_5ONF_3$) and the 141.035 *m/z* peak ($^{13}C^{12}C_4H_5ONF_3$) which contain the C-2 + C-3 sites in alanine in addition to two carbons from the derivatizing reagent, trifluoracetic anhydride. We report these fragments as their raw masses. Measurements of the ion intensity ratios of these two peaks ($^{13}R_{sa,meas}$ and $^{13}R_{VWR,meas}$ for the sample and VWR standard, respectively) were converted to $\delta^{13}C_{VPDB}$ values for the average of the sample's C-2 and C-3 sites using Equation 3.1:

$$^{13}R_{corr} = (((^{13}R_{sa,meas}/^{13}R_{VWR,meas} - 1) \times {nC_{frag}}/{nC_{ala}}) + 1) * {}^{13}R_{VWR,C\text{-}2+C\text{-}3} \quad \textit{(Eqn.3.1)}$$

where the additional variables that appear in Eqn. 1 are: (i) nC_{frag}, the total number of carbon atoms in the measured fragment ion (4 in this case), (ii) nC_{ala} , the numbers of carbons from alanine in that fragment ion (2); and (iii) $^{13}R_{VWR,C\text{-}2+C\text{-}3}$, the $^{13}C/^{12}C$ ratio of VWR alanine averaged across the C-2 and C-3 sites, which are sampled by the

140.031fragment. The $^{13}R_{VWR,C\text{-}2+C\text{-}3}$ value was constrained independently by determining the molecular average ^{13}R value of this standard using EA-IRMS and by using the ninhydrin decarboxylation to determine the ^{13}R value of its C-1 site and then by solving for the ^{13}C content of the C-2 and C-3 sites using Equation 3.2:

$$^{13}F_{C\text{-}2\,+\,C\text{-}3} = (3 * {}^{13}F_{molec.\ avg.} - {}^{13}F_{C\text{-}1})/2 \qquad \textit{(Eqn. 3.2)}$$

where ^{13}F is the fractional abundance of ^{13}C in a sample (*i.e.*, $^{13}C/(^{12}C + {}^{13}C)$), and $^{13}F_{C\text{-}1}$, $^{13}F_{C\text{-}2\,+\,C\text{-}3}$, and $^{13}F_{molec.\ avg.}$ are the ^{13}F values for the C-1 site, the average of the C-2 and C-3 sites, and molecular average, respectively. Note here and elsewhere in this paper we perform mass balance calculations involving mixtures of isotopically distinct materials or molecular sites using $^{i}F_j$ values rather than $^{i}R_j$ or $\delta^{h}X$ values in order to avoid errors arising from non-linearities in the scales of the latter two variables. Note also that $^{i}F_j$ values can be readily interconverted with $^{i}R_j$ or $\delta^{h}X$ values using well established arithmetic relationships.[3.2]

Similarly, after establishing the $\delta^{13}C_{VPDB}$ value of the averaged C-2 + C-3 sites in each sample via Equation 3.1 and determined the molecular-average $\delta^{13}C_{VPDB}$ of each sample via EA-IRMS measurements, we use Equation 3.3 to calculate the ^{13}C content (which we express below as a $\delta^{13}C_{VPDB}$ value) of the C-1 site:

$$^{13}F_{C\text{-}1} = 3 * {}^{13}F_{molec.\ avg.} - 2 * {}^{13}F_{C\text{-}2\,+\,C\text{-}3} \qquad \textit{(Eqn. 3.3)}$$

As explained in Eiler et al. (2017) and Chimiak et al. (2020), all samples are derivatized with the same derivatizing agents, so the derivative sourced carbons are assumed to have identical $\delta^{13}C$ values between samples and standards and thus only dilute differences in ^{13}C content between a given molecular or fragment species of alanine samples and standards. This assumption was verified by analysis of the differences in isotopic contents between the VWR, Alfa Aesar, and Strecker-synthesis standards by both the

[3.2] $^{i}F_j$ is related to $^{i}R_j$ by the formula: $^{i}F_j = \frac{^{i}R_j}{1+\ ^{i}R_j}$ and $^{i}R_j$ is related to $\delta^{h}X$ as described in footnote 3.1.

mass spectrometric method described here and previously established values determined by combining EA-IRMS and ninhydrin decarboxylation data for the molecule and C-1, respectively (Eiler et al., 2017; Chimiak et al., 2020).

3.2.3.3. Calculations of reaction yield and reaction progress

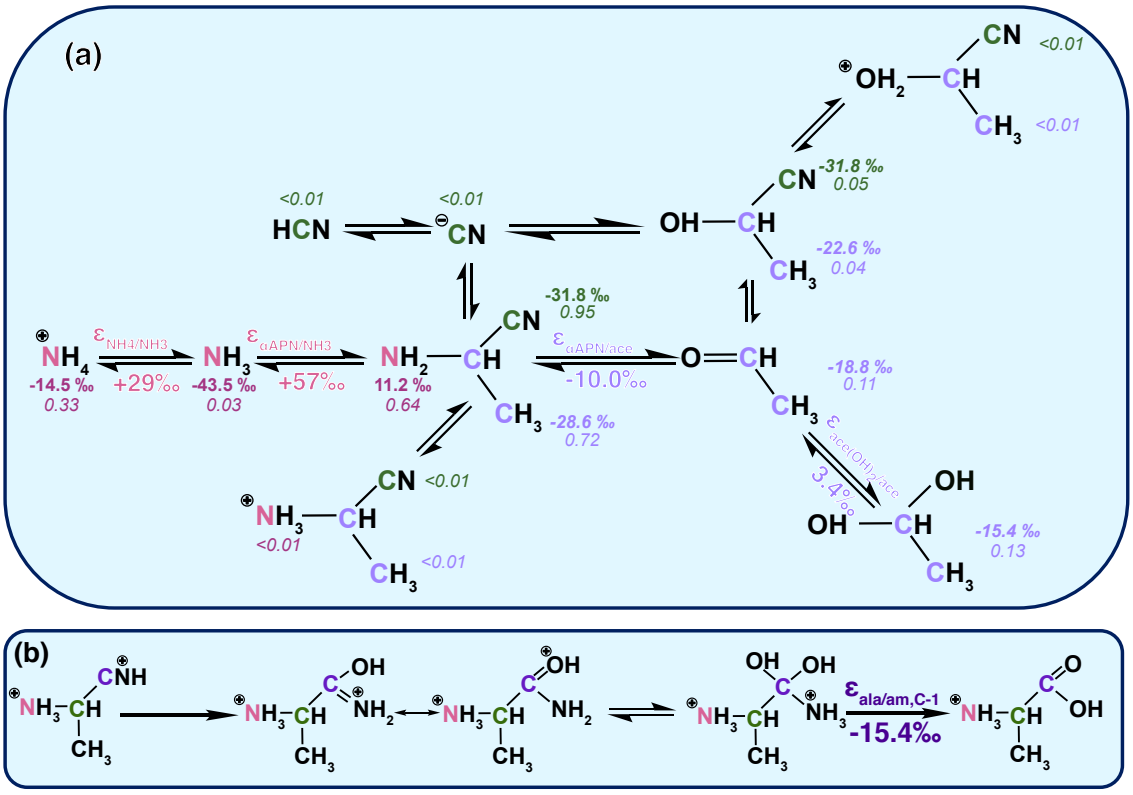


Figure 3.2: Mechanism and reactant and isotopic partitioning for (a) the equilibrium step and (b) the two hydrolysis steps in Strecker synthesis. In panel (a), the equilibrium delta values are listed under each compound and equilibrium concentrations in mMol are listed below the delta values. EIEs are listed above the arrows between the two compounds. All information is color-coded with nitrogen from the initial NH_4Cl pool represented in pink, carbon from the initial acetaldehyde in purple, and carbon from the initial pool NaCN in green for each compound in the same colors as listed above. Compounds with less than 0.01 mMol are included for completeness but do not include isotope effects or final isotope values as they are present in such minimal amounts that they do not impact the isotope ratios of the more abundant compounds. Panel (b) uses the same color scheme.

Previous studies of Strecker synthesis demonstrate that, for the conditions used in this study, the establishment of chemical equilibrium between the initial reactants (aldehyde or ketone, cyanide, and ammonia) and their immediate products in water occurs within minutes and is reversible (Van Trump, 1975; Atherton et al., 2004). Consequently, the equilibrium constants for the reactions that occur during the equilibration step provide a

means to estimate the proportions of species at intermediate steps that we have not directly analyzed. We predict the equilibrium concentration of α-aminopropanenitrile (α-APN), ammonia, acetaldehyde, and cyanide by solving the system of equilibria between reactants, intermediates, and products depicted in Figure 3.2 and listed in Table S3.1. Equilibrium concentrations of these species are solved for in Equations 3.4 to 3.17 below. These equations include equilibrium equations, which relate the relative concentrations of products and reactants to the known values of equilibrium constants (*i.e.,* the ratio of forward and reverse reactions, listed in Table 3.2) for reactions in which those products and reactants participate (Eqn. 3.4 to 3.11, an equation to calculate the water activity at the end of the equilibrium step (Eqn. 3.12), and mass balance equations, which assert that the number of moles of an element of interest contained in a specific pool of initial reactants (e.g., the N atoms in ammonium chloride) at the start of the equilibration step (here denoted with a $t = 0$ subscript) must equal the number of moles of that element contained in the residual pool of those reactants plus the number of moles of that element that were transferred to products at the end of that step (Eqn. 3.13 to 3.17; note no subscript is used for the concentrations at the end of equilibrium).

Reaction Number	Reaction	Equilibrium Constant
1	$NH_4^+ \leftrightarrow NH_3 + H^+$	5.6×10^{-10} [a]
2	$HCN \leftrightarrow CN^- + H^+$	6.03×10^{-10} [a]
3	$CH_3CH(OH)_2 \leftrightarrow CH_3CHO$	8.57×10^{-1} [b]
4	$NH_3 + CN^- + CH_3CHO \leftrightarrow CH_3CH(NH_2)CN + OH^-$	3.8×10^6 [b] (80°C: $k_{fwd} = 11.5$ [b], $k_{rev} = 36.8$ [b] 120°C: $k_{fwd} = 20.2$ [b], $k_{rev} = 1442.5$ [b])
5	$CH_3CH(NH_3^+)CN \leftrightarrow CH_3CH(NH_2)CN + H^+$	4.07×10^{-6} [b]
6	$CN^- + CH_3CHO \leftrightarrow CH_3CH(O)CN$	1.77×10^4 [c]
7	$CH_3CH(OH^+)CN \leftrightarrow CH_3CH(O^-)CN + H^+$	3.2×10^{-12} [d]

Table 3.2: Reactions and equilibrium constants used to calculate concentrations at the end of the equilibrium step. Reaction numbers correspond to K subscripts in equations 4-11. Equilibrium constants are from (a)(PubChem.), (b) (Van Trump, 1975), (c) (Moutou et al., 1995), (d) (Schlesinger and Miller, 1973).

$$K_1 = \frac{[NH_3][H^+]}{[NH_4{}^+]} \qquad \textit{(Eqn. 3.4)}$$

$$K_2 = \frac{[CN^-][H^+]}{[HCN]} \qquad \textit{(Eqn. 3.5)}$$

$$K_3 = \frac{[CH_3CHO]*a_1}{[CH_3CH(OH)_2]} \qquad \textit{(Eqn. 3.6)}$$

$$K_4 = \frac{[CH_3CH(NH_2)CN][OH^-]}{[NH_3][CN^-][CH_3CHO]} \qquad \textit{(Eqn. 3.7)}$$

$$K_5 = \frac{[CH_3CH(NH_2)CN][H^+]}{[CH_3CH(NH_3{}^+)CN]} \qquad \textit{(Eqn. 3.8)}$$

$$K_6 = \frac{[CH_3CH(O)CN]}{[CN^-][CH_3CHO]} \qquad \textit{(Eqn. 3.9)}$$

$$K_7 = \frac{[CH_3CH(O)CN][H^+]}{[CH_3CH(OH^+)CN]} \qquad \textit{(Eqn. 3.10)}$$

$$K_8 = \frac{[H^+][OH^-]}{a_1} \qquad \textit{(Eqn. 3.11)}$$

$$a_1 = \frac{n_{H_2O}}{n_{CN} + n_{HCN} + n_{CH_3CHO} + n_{CH_3CH(OH)_2} + n_{NH_3} + n_{NH_4} + n_{CH_3CH(NH_2)CN} + n_{CH_3CH(NH_3)CN} + n_{CH_3CH(O)CN} + n_{CH_3CH(OH)CN}} \qquad \textit{(Eqn. 3.12)}$$

$$[NH_4Cl]_{t=0} = [NH_4^+] + [NH_3] + [CH_3CH(NH_3^+)CN] + [CH_3CH(NH_2)CN] \qquad \textit{(Eqn. 3.13)}$$

$$NaCN_{t=0} = [HCN] + [CN^-] + [CH_3CH(NH_3^+)CN] + [CH_3CH(NH_2)CN] + [CH_3CH(OH^+)CN] + [CH_3CH(O^-)CN] \qquad \textit{(Eqn. 3.14)}$$

$$[CH_3CHO]_{t=0} = [CH_3CH(OH)_2] + [CH_3CHO] + [CH_3CH(NH_3^+)CN] + [CH_3CH(NH_2)CN] + [CH_3CH(OH^+)CN] + [CH_3CH(O^-)CN] \quad \textit{(Eqn. 3.15)}$$

$$[H^+]_{t=0} + [NH_4Cl]_{t=0} + x = [H^+] + [NH_4^+] + [CH_3CH(NH_3^+)CN] + [CH_3CH(OH^+)CN] + [HCN] \quad \textit{(Eqn. 3.16)}$$

$$[OH^-]_{t=0} + x + [CH_3CH(NH_3^+)CN] + [CH_3CH(NH_2)CN] = [OH^-] \quad \textit{(Eqn. 3.17)}$$

The variable *x* is defined below.

This system of equations accounts for all possible species in solution during the equilibrium step. Initial amounts of ammonia and cyanide were weighed and recorded, and volumes of acetaldehyde and water were measured in syringes and recorded from each synthesis (Table 3.1). For the case of Equation 3.16, we account for OH^- sources, which are initial OH^-, OH^- produced when α-APN is formed, and OH^- produced from water between the start and end of the reaction step. This final OH^- source is denoted as *x* and accounts for the degree of auto-dissociation of water: because the equilibrium constant for water dissociation at a given temperature is always equal to the product of the concentration of H^+ and OH^- divided by the water activity (a_1) (Eqn. 3.11)—the partial vapor pressure of water in a solution relative to that of pure water, which is roughly equal to 0.98 in our system as defined by Equation 3.12—*x* is the change in $[H^+]$ and $[OH^-]$ from water formation or dissociation that must occur to satisfy Equation 3.11. Gas phase concentrations of reactants were negligible (<0.1 %), so we did not include these in our model.

Initial water is assumed to have a pH of 7.0, which does not account for the uptake of CO_2 from air. However, water was added to a sealed 5 mL vial, which at most would contain 84 nmol CO_2. If all of this CO_2 were absorbed into the reaction solution and converted into $H_2CO_3^-$, a chemically improbable scenario, it would only produce 168 nmol of H^+, which is less than 0.009 % of that provided by the conversion from NH_4^+ (*aq*) to NH_3 (*aq*) during the equilibrium step. Consequently, we consider it a

reasonable approximation to not account for the chemistry associated with dissolution of atmospheric CO_2 in our experiments.

Following the equilibration step, α-APN is hydrolyzed into alaninamide and then alanine (the second reaction set in Figure 3.1), both of which are recovered and weighed. Because α-APN is not recovered, we calculate its residual concentration at the end of the hydrolysis steps by taking its calculated concentration after the equilibration step (above), subtracting the amount converted to alaninamide (equal to the sum of the moles of recovered alanine and alaninamide), and subtracting any amount that might have been created from or decomposed into the initial reactants (aldehyde, cyanide, and ammonia) during the hydrolysis steps. The α-APN formation and decomposition reactions are the forward and reverse components of Reaction 4, respectively (Table 3.2). We can solve for their separate rates using Equations 3.18 and 3.19 and rate constants from Van Trump (1975) (Table 3.2):

$$v_{fwd} = k_{f1} * \frac{[\mathrm{RCHO}][\mathrm{CN^-}][\mathrm{NH_3}]}{[\mathrm{OH^-}]} \qquad \textit{(Eqn. 3.18)}$$

$$v_{rev} = k_{r1} * [\alpha\text{-APN}] \qquad \textit{(Eqn. 3.19)}$$

In these equations, ν_{fwd} is the rate of α-APN production from reactants (acetaldehyde, cyanide, and ammonia), and ν_{rev} is the rate of α-APN decomposition into reactants. Rate constants k_{f1} and k_{r1} (Table 3.2) are for the reaction rate for the forward and reverse reactions, respectively. The forward and reverse velocities depend on the concentrations α-APN, CN^-, NH_3, and OH^-, all of which are pH dependent, so we solve Equations 3.18 and 3.19 for two pH conditions, the pH value at the end of equilibrium (pH = 8.2), which we solve for above, and the pH value during hydrolysis (pH < 1).

The quantities [NH_3], [CN^-], and [αAPN] present at the beginning of the hydrolysis step for the two pH conditions listed above are calculated using the relationship:

$$[base] = [acid + base] * \frac{K_a}{K_a + [H^+]} \quad \textit{(Eqn. 3.20)}$$

where [base] is the concentration of the basic species of a compound of interest, [acid + base] is the total concentration of that compound in solution, and K_a is the acid dissociation constant for the compound (Table 3.2). In the case of ammonia, [base] is $[NH_3]$, [acid + base] is $[NH_4^+ + NH_3]$, and K_a equals $K_{a,NH4}$, which equals 5.6×10^{-10} (Table 2).

Final masses of alaninamide (*i.e.*, the residual alaninamide present at the end of the hydrolysis steps) and alanine were gravimetrically determined. The moles of initial alaninamide are calculated as the sum of the moles of alanine and final alaninamide. The yield of alanine amide from α-APN is calculated as the moles of initial alaninamide divided by the moles of α-APN at equilibrium.

3.2.3.4. Calculated isotopic compositions of reaction intermediates

Our isotope ratio measurements of starting materials, products, and select intermediates offer direct constraints on the net isotopic carbon and nitrogen fractionations associated with the Strecker synthesis. Interpreting these net effects as constraints on elementary isotope effects (*i.e.*, isotope effects associated with individual reaction steps) requires a model that considers the isotopic compositions of the intermediates that were not recovered and analyzed for isotopic composition (*e.g.*, α-APN). In this section, we explain how those estimates were made.

To find isotope effects for each step of the Strecker synthesis of alanine, we calculate the isotope ratios of reactants and products at each step in the reaction depicted in Figure 1. As we do not have measurements of the quantities and isotopic composition of reactants, intermediates, and products for Strecker synthesis we use measurements when possible and supplement them with calculations of the quantities and isotopic composition of reactants, intermediates, and products as described in Section 3.2.2.3 and below.

We consider the isotopic composition a given compound or molecular site can have at four different periods (though most compounds are relevant only to a subset of these): 1) before any reaction ('initial'); 2) when the system ends its 4-hour equilibration ('equilibrium'); 3) at the end of the first hydrolysis step ('H-1'); and 4) at the end of the second hydrolysis step ('H-2'). For the initial compositions, the $\delta^{13}C$ of the acetaldehyde and sodium cyanide and the $\delta^{15}N$ of the ammonium chloride reactants that were added to the reaction vial are known from direct measurements. For the equilibrium step, we calculate the $\delta^{13}C$ for all species listed in show in in Figure 3.2 and $\delta^{15}N$ for ammonia and ammonium and for the amine site of α-APN and α-APN^{+}. We treat the equilibrium step as an equilibrium process because the system is closed and all of the relevant reactions are reversible on the time scales of our experiments. Consequently, the isotopic compositions of all reacting compounds (and, when relevant, their site-specific isotopic compositions) must be consistent with equilibrium isotope effects (EIEs) among those species, constraints for which are discussed below. (Note that most reversible reactions permit equilibration between only a subset of sites in each participating molecule; i.e., the equilibrium step may lead to isotope exchange equilibrium between one site of compound A and another site of compound B, but not to equilibration of the site-specific isotopic structures of either compound.), For the H-1 step, we calculate the $\delta^{13}C$ of α-APN and alaninamide, and for the H-2 step, we measured the $\delta^{13}C$ of alanine and residual alaninamide and the $\delta^{15}N$ of alanine. Both H-1 and H-2 are modeled as Rayleigh processes because they are in closed systems and irreversible reactions with associated kinetic isotope effects (KIEs). Compounds and isotope ratios considered in these model calculations are listed in Tables S3.3 to S3.5.

Changes in isotope ratios between reactants and products for each step are calculated separately for the $\delta^{15}N$, for the molecular-average $\delta^{13}C$, and for the site-specific $\delta^{13}C$ of the C-2 + C-3 sites and of the C-1 sites of α-APN, alaninamide, alanine, and the reactants that form these sites in alanine. We calculate isotope effects for all atomic sites that gain or lose covalent bonds in all elementary reaction steps. This definition includes the nitrogen, C-1, and C-2 sites for the equilibrium step, and the C-1 site for the H-1 and H-2 steps. We also estimate the molecular-average $\delta^{13}C$ isotope effect that results from an

isotope effect acting on one or more carbon sites (e.g., if a 9 ‰ effect is estimated for one carbon site of a three carbon molecule, the corresponding molecule-average effect will be 3 ‰). Our analysis also considers measured changes in isotope ratios for sites that are not directly involved in hydrolysis reactions during H-1 and H-2 steps relative to reaction progress for the two hydrolysis steps for the sites that are not directly involved in hydrolysis (N and C-2 + C-3), as these sites can potentially express secondary kinetic isotope effects. However, at the 2 standard error confidence level no consistent increase or decrease in these sites' isotopic compositions with reaction progress for both H-1 and H-2.

The following paragraphs detail how the preceding approaches were applied to specific atomic sites and reaction steps:

Isotope effects on nitrogen for the equilibrium step are solved for using Equation 3.21 (subscripts refer to the molecule of interest of f_x and $^{15}F_x$ values and refer to the product/reactant for the α values):

$$^{15}F_{tot} = \frac{f_{NH_3}}{1+\alpha_{\frac{\alpha APN}{NH_3}}\left(\frac{1}{^{15}F_{\alpha APN}}-1\right)} + \frac{f_{NH_4}\,\alpha_{\frac{NH_4}{NH_3}}}{\alpha_{\frac{NH_4}{NH_3}}+\alpha_{\frac{\alpha APN}{NH_3}}\left(\frac{1}{^{15}F_{\alpha APN}}-1\right)} + f_{\alpha APN}\,^{15}F_{\alpha APN}$$

(Eqn. 3.21)

where $^{15}F_{tot}$ is calculated from the $\delta^{15}N$ value of ammonia added to the reaction vessel and $^{15}F_{\alpha APN}$ is the ^{15}F value of the α-APN pool at equilibrium, which is assumed to be equal to the ^{15}F value of final product alanine, as ^{15}F does not change significantly and consistently with the reaction progress of steps H-1 or H-2 (see Results). The f_x values in Equation 21 are the fraction of compound *x* at equilibrium and are equal to the concentration of x calculated at equilibrium divided by the concentration of NH_4Cl added to the reaction vial. The $\alpha_{P/R}$ values for nitrogen, defined as the ratio of $^{15}R_P/^{15}R_R$, are equilibrium isotope effects between the product (subscript P) and the reactant (subscript R; ammonia in this case). We use a value of 1.029 for $\alpha_{NH4/NH3}$ (Walters et al., 2019).

Carbon isotope for the carbons that will ultimately be transferred to the C-2 and C-3 sites of alanine are calculated for the equilibrium step with Equation 3.22:

$$^{13}F_{tot} = \frac{f_{ace}}{1+\alpha_{\frac{\alpha APN}{ace}}\left(\frac{1}{^{13}F_{\alpha APN}}-1\right)} + \frac{f_{ace(OH)_2}\,\alpha_{\frac{ace(OH)_2}{ace}}}{\alpha_{\frac{ace(OH)_2}{ace}}+\alpha_{\frac{\alpha APN}{ace}}\left(\frac{1}{^{13}F_{\alpha APN}}-1\right)} + f_{\alpha APN}\,^{13}F_{\alpha APN}$$

(*Eqn. 3.22*)

Equation 3.22 is analogous to Equation 3.21 but considers the carbons transferred from acetaldehyde to αAPN. In Equation 3.22, $^{13}F_{tot}$ is based on the molecular average $\delta^{13}C$ value of acetaldehyde added to the reaction vessel and $^{13}F_{\alpha APN,eq}$ is the ^{13}F value of the carbons in α-APN that will become C-2 and C-3 in alanine. Each f_x value is the concentration of compound x at equilibrium for a specific experiment divided by the concentration of acetaldehyde initially added to the reaction vial for that experiment. Here the compounds are abbreviated as follows: acetaldehyde is 'ace', acetaldehyde hydrate is 'ace$(OH)_2$', and α-APN is αAPN. The $\alpha_{P/R}$ values are as defined in the preceding paragraph but for ^{13}R instead of ^{15}R. As there are no measurements of the equilibrium isotope effect between acetaldehyde and acetaldehyde hydrate, we adopt a value of 1.0034 for $\alpha_{ace(OH)2/ace}$, which is half the value for CO_2 hydration a it would be averaged over two carbons in acetaldehyde (Marlier and O'Leary, 1984).

The equilibrium concentrations calculated in Equations 3.4 to 3.17 find that effectively all (> 95 %) cyanide added at the start of an experiment is transferred to the C-1 site of α-APN (with the remainder being transferred to the C-1 site of α-hydroxypropionitrile). α-APN and α-hydroxypropionitrile have the same bonding environment at the C-1 site and therefore we adopt the approximation that they have the same equilibrium carbon isotope effects with respect to cyanide. Consequently, we can assert that the C-1 site of α-APN (and α-hydroxypropionitrile) has a $\delta^{13}C$ value at the end of equilibrium closely similar to that of initial HCN reagent. However, we do not have an estimate of the equilibrium fractionation factor between HCN and α-APN (and α-hydroxypropionitrile), and thus we cannot assign a value to the $\delta^{13}C$ of the trace of residual CN at the end of the

equilibrium step; however, this uncertainty has no impact on our ability to predict or interpret $\delta^{13}C$ values of other species of interest.

Our experimental results (detailed below) indicate negligible isotopic fractionation of the carbon atoms that will end up in the C-2 and C-3 sites of alanine during the H-2 reaction step because the $\delta^{13}C$ value of these sites are approximately invariant across the progress of this irreversible reaction. This implies that H-2 has negligible secondary and tertiary carbon isotope effects. Consequently, we can use the ^{13}F value for the C-2 and C-3 sites in product alanine and assume they are the same for initial alaninamide and with these values calculate the ^{13}F and therefore ^{13}R values at the C-1 site in alaninamide and alanine according to equation 3. Combining these C-1 site ^{13}R values and with that of the α-APN's C-1 site provides the ^{13}R values for all of the reactants and products in H-1 and H-2. These values enable us to calculate the isotope effect for H-1 and H-2 by using the Rayleigh equation, written below in Equation 3.23, for each hydrolysis step separately.

$$f^{\frac{\alpha_P - 1}{R}} = \frac{^{13}R_R}{^{13}R_{R,0}} \qquad \textit{(Eqn 3.23)}$$

where $^{13}R_R$ and $^{13}R_{R,0}$ are the ^{13}R values for the residual and initial reactants (α-APN for H-1 and alaninamide for H-2), f is the fraction of residual to initial reactant, and $\alpha_{P/R}$ is fractionation factor between the product and reactant.

Finally, we calculate molecular-average isotope effects for the equilibrium step, H-1, and H-2. As the CN is calculated to have nearly quantitative incorporation into the intermediate α-APN, the equilibrium isotope effect for the full molecular average can be calculated as 2/3 times that of the acetaldehyde's equilibrium isotope effect with α-APN's averaged C-2 and C-3 sites. The molecular-average isotope effects at H-1 are calculated using the Rayleigh equation (Eqn. 23) with α-APN's calculated molecular-average ^{13}R values and the initial alaninamide's calculated ^{13}R values. Finally, the molecular-average isotope effects at H-2 are similarly calculated from the Rayleigh

equation (Eqn. 3.23) and the calculated initial ^{13}R for alaninamide and measured ^{13}R of alanine.

3.3. Results

Using the methods described above, we measured the initial concentrations of reactants, the yields of alaninamide and alanine, the molecular average $\delta^{13}C$ of reactant acetaldehyde and cyanide added to the reaction vessel, the $\delta^{15}N$ of reactant ammonium chloride added to the reaction vessel, the molecular average $\delta^{13}C$ of alaninamide at the end of the reaction period, and the molecular average $\delta^{13}C$, the average $\delta^{13}C$ of the C-2 and C-3 sites, and the $\delta^{15}N$ of alanine at the end of the reaction period. With these measurements and Equations 3.4 to 3.20, we calculated the reaction rates and yields. Using Equations 3.3 and 3.21 to 3.23, we also calculated the carbon and nitrogen stable isotopic compositions of all reactants and intermediates and the equilibrium and kinetic isotope effects for the equilibration and hydrolysis steps respectively.

3.3.1. Rate of and concentrations at α-aminopropionitrile equilibrium

In the initial step of the Strecker synthesis, acetaldehyde, cyanide, and ammonia form and equilibrate with α-aminopropionitrile (α-APN). Using the concentrations of acetaldehyde, cyanide, and ammonia present in our reaction vessel in their reactive forms (*i.e.,* CH_3CHO, CN^-, NH_3; see section 2.2.3 for calculations of these concentrations and Table 1 for moles of reactant added to reaction vessel), and the kinetics equations from Van Trump (1975), we calculate that at the initial pH of 7 with initial reactant values the forward rate of α-APN production in the equilibrium step is 1.6×10^4 M/hour. However, our equilibrium reaction network summarized in equations 4-12 and Figure 2a indicates that at equilibrium the pH reached 8.2. Using the initial reactant concentrations of reactants at this higher pH, the forward reaction rate rises to 2.1×10^4 M/hour. Because the calculated equilibrium concentration of α-APN is approximately 1.7 M, these rates suggest that equilibrium was rapidly achieved (on a sub-minute timescale), and we can

assume that by the end of the 4-hour equilibration step, the reactants and intermediates were in equilibrium with one another.

The average proportions of reactants and products for the equilibrium step are depicted in Figure 3.2a. At the end of the 4-hour equilibration step, the amount of α-APN produced ranges from 1.79 ± 0.09 mmol to 2.23 ± 0.12mmol and residual cyanide, acetaldehyde, and ammonia pools are ~0, 0.38 ± 0.01 to 0.85 ± 0.02, and 0.98 ± 0.02 to 1.31 ± 0.03 mmol, respectively (Table S1). Throughout this paper, 'pools' of a compound will refer to all related and rapidly interconvertible forms of that compound; for cyanide this is CN^- and HCN combined, for acetaldehyde it is CH_3CHO and $CH_3CH(OH)_2$ combined, and for ammonia it is NH_3 and NH_4^+ combined. Relative to the cyanide, acetaldehyde, and ammonium pools initially added to the system, the yield of α-APN is 95%, 62 to 82 %, and 59 to 68 % respectively (Table S3.2). For the equilibrium step, therefore, cyanide is the limiting reagent in all syntheses.

Following the equilibrium step, 6N HCl is added to the reaction vessel and α-APN is hydrolyzed into alaninamide and then alanine during the 1-hour hydrolysis period. It is important that we evaluate the extent to which this hydrolysis could be accompanied by continued production of a-APN or the degradation of it, as any significant reactions of this kind would complicate our interpretations of the isotope effects associated with reaction steps H-1 and H-2. To this end, we use Equation 3.18 to calculate the maximum rate at which α-APN could be produced by continued reaction among residual cyanide, acetaldehyde and ammonia during hydrolysis. Rate scales with reactant concentration and temperatures, so to find the maximum rate we use the initial concentrations of acetaldehyde, cyanide, and ammonia rather than that present at equilibrium (which are lower) and the highest hydrolysis temperature (120°C). Using these parameters (which clearly overestimate rate for most or all experiments), the maximum rate of α-APN production is 2.7×10^{-3} M/hr. We calculate the maximum rate of α-APN decomposition into its reactants using Equation 3.19. As the rate of α-APN decomposition scales with α-APN concentration and reaction temperature, to find the maximum rate possible in our system, we use the highest possible α-APN concentration, the amount that would be

present if we had quantitative conversion of reactants into α-APN and the 120°C hydrolysis temperature. With these parameters, the maximum rate of α-APN decomposition (again, clearly an overestimate) is 2.0×10^{-3} M/hr. Based on these calculated rates, during the one-hour hydrolysis, less than 0.2 % of the α-APN at equilibrium would be produced or degraded. This amount is insignificant to our analysis.

Given that the α-APN production and decomposition after equilibrium are insignificant, yields and isotope effects associated with hydrolysis were calculated according to the following scheme: We assume the reactants and α-APN reach mutual equilibrium, after which the equilibrium α-APN serves as a reactant pool for the subsequent hydrolysis, which can be treated as a relatively straightforward series of two irreversible steps.

3.2.2. Yields of alaninamide and alanine.

Using the quantities of α-APN present after the equilibrium step (Table 3.1) and the measured masses of alaninamide and alanine produced in each reaction, we calculated the yields of alaninamide and of alanine. The yield of alaninamide as a fraction of pre-hydrolysis α-APN ranges from 26 % to 89 % and the yield of alanine as a fraction of alaninamide formed by reaction step H-1 ranges from 5 % to 100 % (Table S3.5). Relative to the initial cyanide pool, the yields of total alaninamide and of alanine range from 24 % to 85 % and 2 % to 69 %, respectively (Table S3.2).

3.2.3. Nitrogen isotopes

Measurements of alanine's $\delta^{15}N$ value are between 8.8 ‰ and 12.9 ‰ — ^{15}N enriched relative to the initial ammonia pool ($\delta^{15}N = 0.8 \pm 0.3$ ‰, based on measurements of the ammonium chloride reactant). The measured $\delta^{15}N$ of alanine shows no consistent trend with respect to yield for H-1 or H-2 (Table S3.2, Figure S3.1), so we assert that in each sample, α-APN's amine site's $\delta^{15}N$ is equal to that of alanine and that no KIEs on the amine nitrogen are associated with H-1 and H-2. At equilibrium, roughly two-thirds of the initial ammonia pool is in the form of α-APN's amine site (57 ± 4 % to 68 ± 4 %),

and significant portions of the initial ammonia pool remain as ammonia (~3 %) or ammonium (29 ± 1 % to 38 ± 1%). By mass balance with α-APN and the initial ammonia pool, the residual ammonia pool has a $\delta^{15}N$ of -17 ‰, which is nearly 30 ‰ below the alanine value. Given the 29 ‰ EIE between ammonia and ammonium, the $\delta^{15}N$ of ammonia at equilibrium lies between -44.69 ± 0.42 ‰ and -42.35 ± 0.37 ‰ (Table S3.2). The $\delta^{15}N$ of alanine is therefore 42.35 ‰ to 57.23 ‰ enriched relative to ammonia at equilibrium (Figure 3.3, Table S3.2). The wide range is due to the $\delta^{15}N$ of sample 14, and barring that sample all alanine values are between 53.22 ‰ and 57.23 ‰ enriched relative to ammonia at equilibrium.

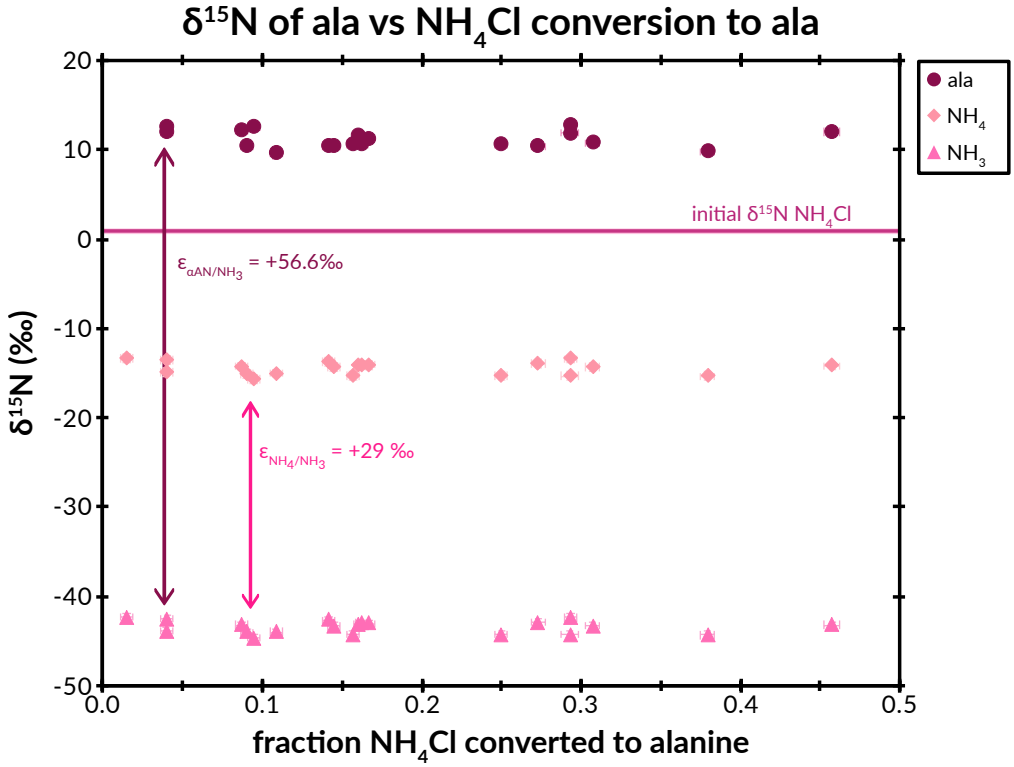


Figure 3.3: The equilibrium partitioning of nitrogen species is plotted relative to the total yield of alanine with respect to initial NH_4Cl input. The initial $\delta^{15}N$ NH_4Cl and the $\delta^{15}N$ of alanine and the fraction of NH_4Cl converted to alanine are measured. The $\delta^{15}N$ of ammonia and ammonium are calculated according to their equilibrium concentrations and that of α-APN

Our finding that nitrogen isotope KIEs are negligible (~0 ‰) for reaction steps H-1 and H-2 is expected, as these hydrolysis steps include no bond reordering at the amine nitrogen. Our finding of a large (56.6 ‰) EIE between ammonia and α-APN ($\varepsilon_{aAPN\text{-}NH3}$) during equilibrium (Table 3.3) can be attributed to the change in nitrogen bonding

Step	Isotope	Site	Value	Comments
EIE	^{15}N		56.6 ‰	Meas: Concentrations: initial NH_4Cl, final alaninamide and alanine; $\delta^{15}N$: initial NH_4Cl, final alanine Calc: Concentrations: all species at equilibrium, α-APN at the end of H-1 $\delta^{15}N$: NH_3 and NH_4^+ at equilibrium Assumptions: No chemical steps other than H-1 and H-2 occur during hydrolysis, H-1 and H-2 do not impact $\delta^{15}N^a$
	^{13}C	Molec Avg	-6.7 ‰	Meas: Concentrations: initial NaCN and CH_3CHO, final alaninamide and alanine; $\delta^{13}C$: initial NaCN and CH_3CHO, final alanine Calc: Concentrations: all species at equilibrium, α-APN at the end of H-1 $\delta^{13}C$: α-APN (weighted average of C-1 and C-2 + C-3 values), CH_3CHO, and $CH_3CH(OH)_2$ at equilibrium Assumptions: Same as those for C-2 + C-3 and C-1
		C-2+C-3	-10.0 ‰	Meas: Concentrations: initial CH_3CHO, final alaninamide and alanine; $\delta^{13}C$: initial CH_3CHO, final alanine C-2 + C-3 sites Calc: Concentrations: all species at equilibrium, α-APN at the end of H-1 $\delta^{13}C$: CH_3CHO, $CH_3CH(OH)_2$, $CH_3CH(O)CN$ pool at equilibrium Assumptions: CH_3CHO hydration's EIE = CO_2 hydration's EIE, α-APN and $CH_3CH(O)CN$ have the same EIE with respect to CH_3CHO, no chemical steps other than H-1, and H-2 occur during hydrolysis, H-1 and H-2 do not impact C-2 + C-3 sites' $\delta^{13}C^a$
		C-1	N/A	Meas: Concentrations: initial NaCN and CH_3CHO $\delta^{13}C$: initial NaCN Calc: Concentrations: all species at equilibrium Assumptions: α-APN and $CH_3CH(O)CN$ have the same EIE with respect to CN
H-2 KIE	^{15}N		0 ‰	Meas: Concentrations: initial NH_4Cl, final alaninamide and alanine; $\delta^{15}N$: initial NH_4Cl, final alanine Calc: α-APN at the end of H-1 Assumptions: No chemical steps other than H-1 and H-2 occur during hydrolysis

Step	Isotope	Site	Value	Comments
H-2	^{13}C	Molec Avg	-5.4 ‰	Meas: Concentrations: final alaninamide and alanine; $\delta^{13}C$: final alaninamide and alanine Calc: Initial alaninamide (sum of alanine + alaninamide) Assumptions: Alaninamide and alanine measured represent the H-2 step (*i.e.*, no additional loss processes for alaninamide or alanine)
		C-2+C-3	0 ‰	Meas: Concentrations: final alaninamide and alanine; $\delta^{13}C$: final alanine C-2 + C-3 sites Calc: α-APN at the end of H-1 Assumptions: no chemical steps other than H-1and H-2 occur during hydrolysis, H-1 and H-2 do not impact C-2 + C-3 sites' $\delta^{13}C^{a}$
		C-1	-15.2 ‰	Meas: Concentrations: final alaninamide and alanine; $\delta^{13}C$: final alanine C-1 site, final alaninamide molecular average Calc: α-APN at the end of H-1, $\delta^{13}C$ of initial alaninamide's C-1 site Assumptions: H-2 does not impact C-2 + C-3 sites' $\delta^{13}C^{a}$, alaninamide and alanine measured represent the H-2 step

Table 3.3: Isotope effects for Strecker synthesis. Meas lists values related to the isotope effect that are directly measured. Calc lists values that were calculated (see section 2.2.2.3) that are used to calculate the isotope effect. Assumptions lists assumptions that were made in order to calculate the related isotope effect.
[a]Whether H-1 and H-2 impact species is determined cross plots of measured isotope ratios and reaction progress for each step. If no impact is assumed, then no significant trend in isotope ratio and reaction progress was noted.

environment from sp^2 hybridization to sp^3. This value is of higher magnitude but the same sign for as the 29 ‰ EIE between NH_3 and NH_4^+, in which the sp^3 product (NH_4^+) is also heavier and is of similar sign and magnitude to that seen in the equilibrium between nitrate and nitrite in which nitrite is roughly 30 to 60 ‰ heavier (Brunner et al., 2013).

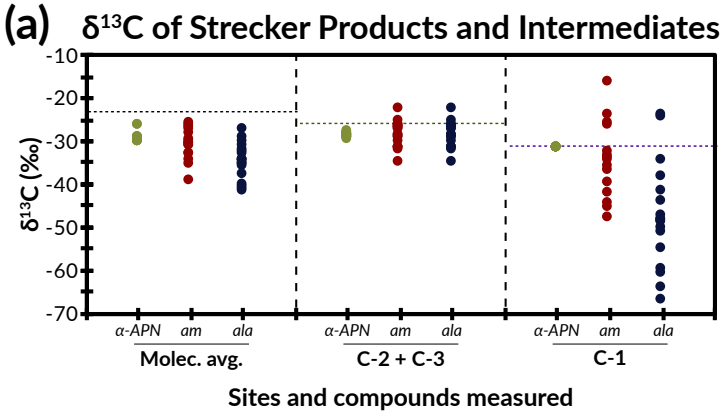
3.2.4.Carbon isotopes

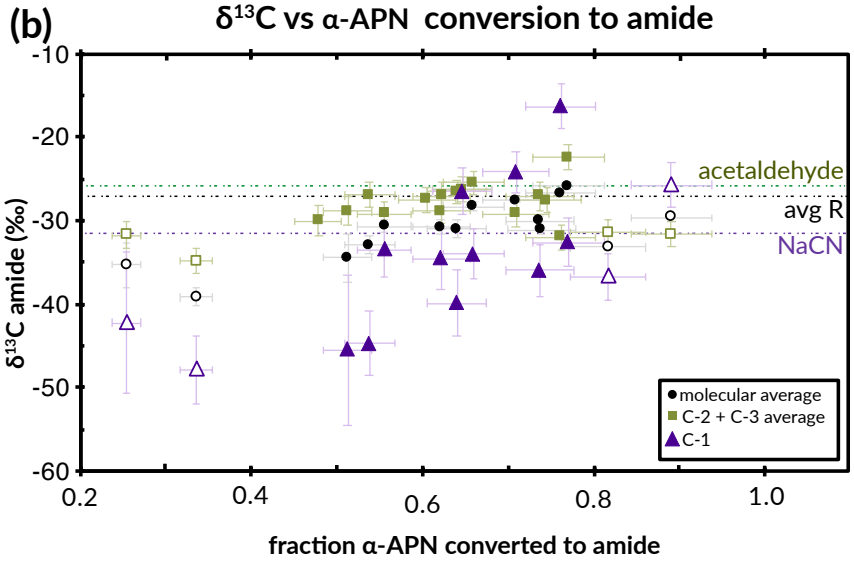
3.2.4.1.Molecular-Average $\delta^{13}C$

Measurements of the carbon isotope compositions of acetaldehyde and sodium cyanide reagents yield $\delta^{13}C$ values of -26.00 ± 0.01 ‰ and -31.8 ± 0.2 ‰, respectively. Under the simplest assumption that Alanine is produced from Strecker synthesis with no isotopic fractionation between reactants and products (or carbon-bearing side products, such as acetaldehyde hydrate), product alanine would have a molecular average $\delta^{13}C$ equal to the average of acetaldehyde and cyanide's molecular-average $\delta^{13}C$ values, weighted by their contributions of carbon to alanine (2/3 and 1/3 respectively), or -27.9 ± 0.1 ‰. Accounting for the equilibration between acetaldehyde and acetaldehyde hydrate (Figure 3.2a), for which we use an $\varepsilon_{ace\text{-}ace(OH)2}$ of 3.4 ‰ (half the EIE for the hydration of CO_2 (Marlier and O'Leary, 1984)), the partitioning of modestly ^{13}C-depleted carbon into the small amount of acetaldehyde hydrate would lead us to expect a molecular average $\delta^{13}C$ of alanine of -27.8 ‰ to -26.7 ‰ (Table S3.3).

We calculate the molecular-average $\delta^{13}C$ values of α-APN as the weighted average of the C-1 site and the C-2 and C-3 sites in α-APN for each synthesis with the same weighting as above. These site-specific values are discussed in the C-1 and the C-2 and C-3 site sections (sections 3.4.2 and 3.4.3). The molecular average value for α-APN computed this way ranges from -26.2 ± 0.5 ‰ to -30.2 ± 0.5 ‰ (Table S3.4). We compute the initial molecular average $\delta^{13}C$ of alaninamide—that is the value of the entire alaninamide pooled after H-1 but before H-2—as a weighted average the molecular-average ^{13}F (see Footnote 3.2 for definition of ^{13}F and its relation to ^{13}R) values for alanine and alaninamide. Molecular-average $\delta^{13}C$ measurements of residual alaninamide and alanine range from -22.19 ± 0.17 ‰ to -33.63 ± 0.02 ‰ and -27.54 ± 0.16 ‰ to -41.50 ± 0.03 ‰, respectively (Table S3.5). Weighted by the alaninamide and alanine abundances at the end of at the end of hydrolysis (Table S3.1), the initial molecular average $\delta^{13}C$ of alaninamide is calculated to span from -25.8 ± 2.9 ‰ to -39.2 ± 1.1 ‰ (Table S3.4).

Therefore, the predicted $\delta^{13}C$ value of alanine produced by Strecker synthesis with no fractionations is between 0.1 ‰ and 14.1 ‰ higher than was observed (Table S3.4, Figure 3.4a), which indicates that the overall carbon isotope effect of the Strecker synthesis is 'normal' (*i.e.*, faster for ^{12}C species; slower for ^{13}C species). A normal carbon isotope effect of the Strecker synthesis is also suggested by the fact that the molecular-average $\delta^{13}C$ of product alanine increases (approaching the weighted average of reactants) as alanine yield increases (Figure 3.4b and c).





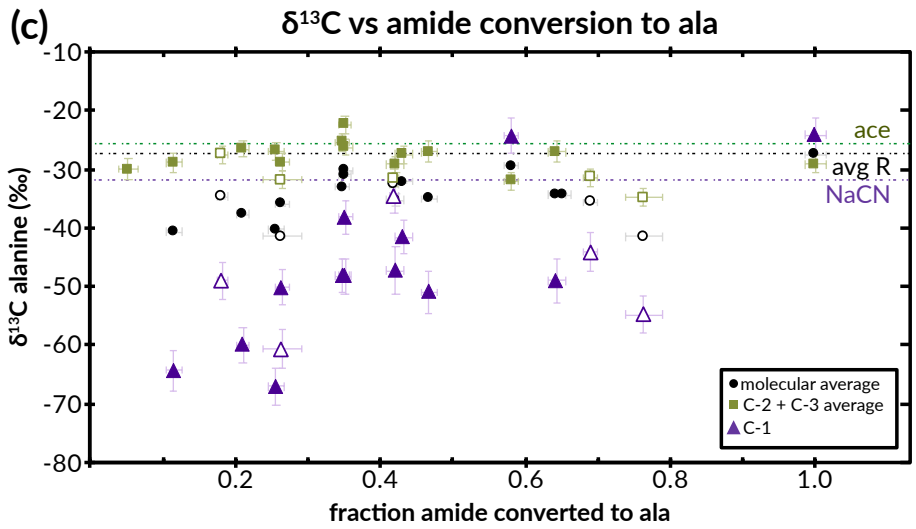


Figure 3.4: Alanine and intermediates' $\delta^{13}C$ values. Panels present carbon isotope ratios of (a) the total amide pool relative to the fractional conversion to amide and (b) alanine relative to the fractional conversion to alanine. All values for intermediates represent the intermediate pool after formation but before the next reaction (*e.g.*, α-APN $\delta^{13}C$ values represent those after equilibrium and before H-1). Both alanine (ala) values are measured, and amide (am) values are calculated as the weighted sum of the measured alanine and amide values, and α-APN values are predicted using the equilibrium concentration calculations and the calculated EIE between acetaldehyde and α-APN. Lines represent the measured $\delta^{13}C$ of the reactants added to the vessel. Molecular average and site-specific $\delta^{13}C$ values are plotted against fractional conversion for the hydrolysis steps. Filled symbols represent data from experiments conducted below 115ºC while open symbols are for data from experiments conducted at and above 115ºC.

More specifically, these relationships indicate the estimated α-APN molecular average $\delta^{13}C$ values reflects a -6.7 ‰ EIE between α-APN (lower in $\delta^{13}C$) and the weighted average of cyanide and acetaldehyde (higher in $\delta^{13}C$; Figure 3.5a, Table 3.3). The hydrolysis of α-APN to alaninamide is calculated to have no statistically significant KIE as the ratio of the calculated residual to initial α-APN ^{13}R values is relatively flat when the fraction of residual α-APN is above 40% and varies between extremely high and low values (±30 ‰) with no clear trend when the residual α-APN is below 25% (Figure S3.2a). Regardless, we do note an increase in the difference between the molecular average $\delta^{13}C$ of the amide at the end of H-1 and the α-APN at the beginning of H-1 (Figure 3.5b) with some points at high yield having values above 0 ‰. A secondary reaction that enriches the α-APN pool in ^{13}C—and causes the fraction of α-APN residual to H1 to be lower that we calculate here—could create this effect. In the highly acidic conditions used here, it is possible that the α-APN underwent side reactions that could cause the α-APN pool available for hydrolysis to increase in $\delta^{13}C$. Doing so would also

decrease the number of mols of α-APN available for hydrolysis and thereby the fraction of residual α-APN after hydrolysis, too.

The best constrained step, H2, for which $\delta^{13}C$ and abundances of the product alanine and reactant alaninamide are directly measured has a normal 5.2 ‰ KIE (Figure 3.5c, Table 3.3).

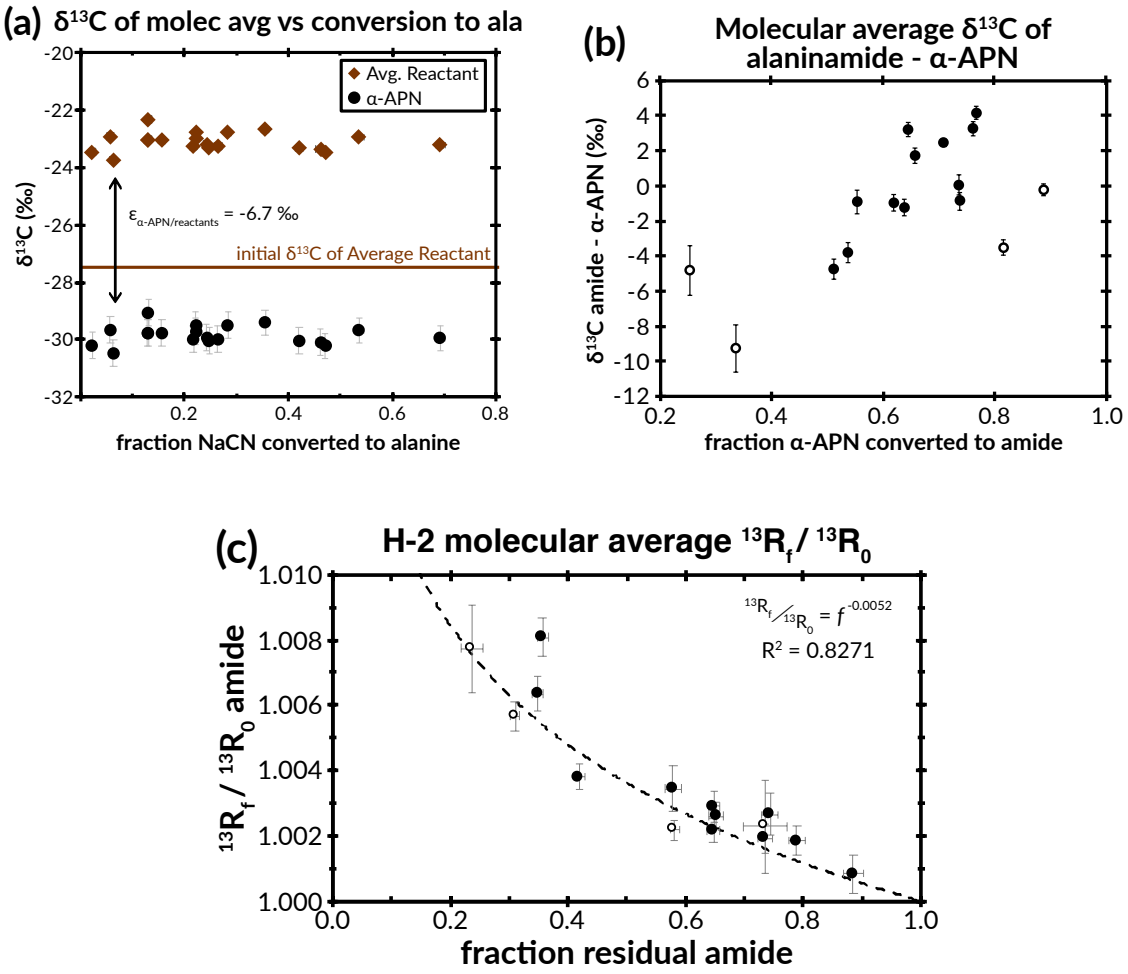


Figure 3.5: Molecular-average carbon isotope effect models and data. (a) Equilibrium $\delta^{13}C$ values predicted for α-APN and average reactants. Here the $\delta^{13}C$ value represent those at the end of equilibrium but prior to the beginning of H-1. (b) The difference between $\delta^{13}C$ values calculated for α-APN at the end of equilibrium and $\delta^{13}C$ calculated from the measured alanine and alaninamide values at the end of H-1 but before H-2. (c) Ratios of $^{13}R_{residual}$ vs. $^{13}R_{initial}$ for alaninamide versus the fraction of alaninamide remaining after H-2. Alaninamide $^{13}R_{residual}$ is measured and $^{13}R_{initial}$ is calculated as the weighed sum of ^{13}F of alanine and alaninamide collected at the end of H-2.

3.2.4.2.The 'amine' and 'methyl' carbon sites

The carbonyl and methyl carbons in acetaldehyde become the amine and methyl carbons (the C-2 and C-3 sites, respectively) in alanine (Figures 3.1 and 3.2). We measure the average $\delta^{13}C$ value for the C-2 and C-3 sites in alanine for each synthesis. Values range from -22.4 ± 1.5 ‰ to -34.8 ± 1.6 ‰ (Table S3.4).

As the C-2 and C-3 sites only form and break bonds during the equilibrium step (Figure 3.2) and neither hydrolysis step (H-1 and H-2) demonstrated a trend in $\delta^{13}C$ of the average of these sites with reaction progress (Figures 3.4b, 3.4c, and 3.6b), we infer that the difference in carbon isotope composition between initial acetaldehyde reagent and the average of the C2+C3 sites of alanine is predominantly due to an EIE between acetaldehyde the C-2 and C-3 sites of α-APN. During the equilibrium step, the initial acetaldehyde pool is transformed into a mixture of α-APN (50 ± 6 % to 82 ± 6 %), residual acetaldehyde (70 ± 1 % to 15 ± 1 %), acetaldehyde hydrate (7 ± 1 % to 17 ± 1 %), and α-hydroxypropionitrile (~4 %) (Table S3.1, Figure 3.6a). The bonding environments for the C-2 and C-3 carbons in α-APN and α-hydroxypropionitrile are effectively the same, so we assert that the EIE between acetaldehyde and each of these pools will be equal and combine them when calculating an EIE between acetaldehyde and α-APN. Using equation 22 and a 1.0034 value for $\alpha_{ace\text{-}ace(OH)2}$, we calculate an EIE between acetaldehyde (lower in $\delta^{13}C$) and the average of the C-2 and -C3 carbons of α-APN (higher in $\delta^{13}C$) of -10.0 ‰ (*i.e.,* this α value = 0.990; Table 3.3) (Figure 3.6a). Using this -10.0 ‰ EIE and applying the mass balance constraints from the proportions of compounds present at the end of the equilibrium step, the calculated $\delta^{13}C$ of the C-2 and C-3 sites of almost all α-APN produced at the end of equilibrium but prior to reaction H1 are within error of the corresponding measurements of alanine (Table S3.3 and S3.4); thus, within our measurement error, the $\delta^{13}C$ of the C-2 and C-3 sites is unchanged from α-APN to alaninamide to alanine (Figure 3.6b). That is to say, no resolvable secondary or tertiary isotope effects act on these two carbon sites during reaction steps H-1 and H-2.

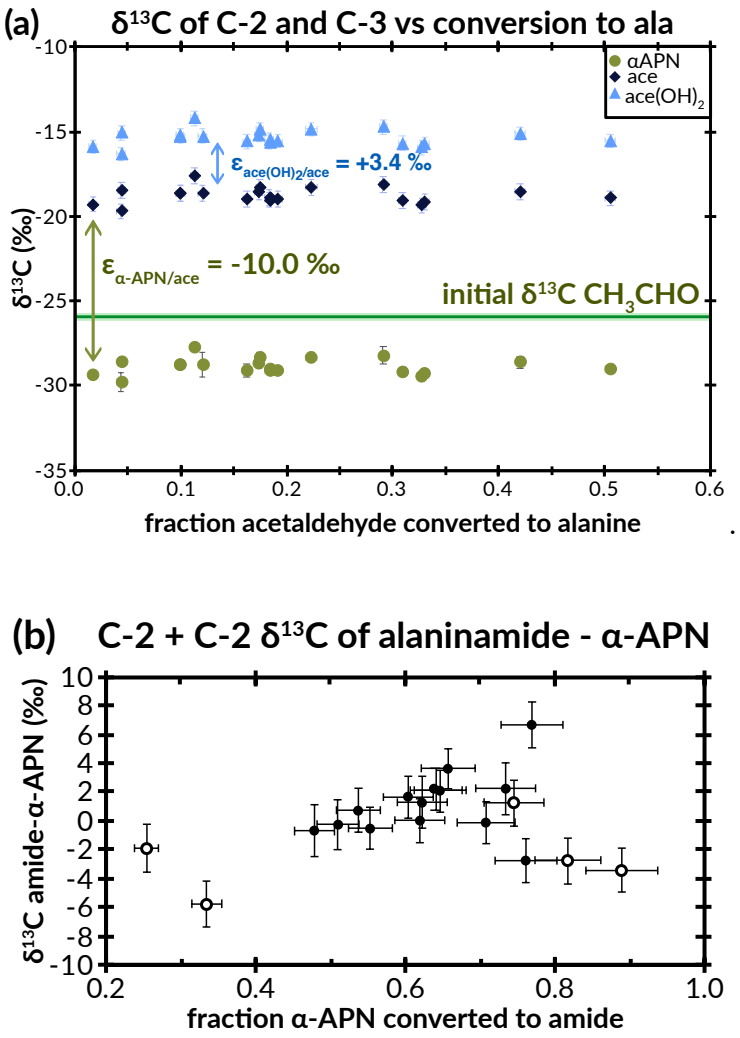


Figure 3.6: Changes in $\delta^{13}C$ for the C-2 and C-3 sites' average for each step in Strecker synthesis. (a) α-APN $\delta^{13}C$ versus that of acetaldehyde and hydrated acetaldehyde at the equilibrium of reactants and α-APN. (b) Ratio of residual to initial ^{13}R for the C-2 and C-3 site average in α-APN for H-1.

3.2.4.3.The 'carboxyl' (C-1) carbon site

In Strecker synthesis, cyanide becomes the C-1 site on α-APN, alaninamide, and alanine (Figure 3.2) as the nitrile carbon, the amide carbon, and the carboxyl carbon, respectively. We measured the $\delta^{13}C$ of cyanide (-31.8 ± 0.2 ‰) and calculate that of alanine's C-1 site using the measured molecular average $\delta^{13}C$ and average $\delta^{13}C$ of the C-2 and C-3 sites using Equation 3.3 (noting again that all such calculations transform $\delta^{13}C$ values to ^{13}F values prior to such mass balance calculations). Using this method, we find that the $\delta^{13}C$ of alanine's C-1 site ranges from -24.2 ± 3 ‰ to -67.1 ± 3 ‰ with all

but two values lying at or below the cyanide pool's -31.8 ± 0.2 ‰ value (Figure 3.4, Table S3.4).

In the experiments performed in this study, cyanide was nearly quantitatively converted (~95 %, Table S3.1 and S3.3) to the C-1 site in α-APN during the equilibrium step. Almost all of that which did not form the C-1 site on α-APN formed the C-1 site on the side product, α-hydroxypropionitrile (~5 %). The C-1 sites of α-APN and α-hydroxypropionitrile have similar bonding environments, so we infer that they have similar EIEs with respect to cyanide. This similarity and the near total conversion of cyanide into α-APN lead us to conclude that the $\delta^{13}C$ of the C-1 site in α-APN at the end of the equilibrium reaction step equals that of the starting cyanide, -31.8 ± 0.2 ‰ (Figure 3.2a, Table S3.3). Therefore, differences in isotopic composition of the C-1 site of alanine are interpreted to reflect isotope effects associated with the hydrolysis steps (H-1 and H-2).

As explained above, we infer that the average $\delta^{13}C$ of C-2 and C-3 sites of the alaninamide approximately equal that for C-2 and C-3 sites on alanine. Consequently, we can calculate the $\delta^{13}C$ of C-1 for alaninamide by subtracting the ^{13}F of C-2 and C-3 sites in alanine from the molecular average ^{13}F in alaninamide using Equation 3.3 (then converting the resulting ^{13}F value to a $\delta^{13}C$ value for the purposes of reporting; Footnotes 3.1 and 3.2). Doing so, we find that the $\delta^{13}C$ of C-1 for alaninamide ranges from -16.3 ± 3 ‰ to -48 ± 4 ‰ with all but two values within error of or below the $\delta^{13}C$ value of initial cyanide reagent (Figure 3.4, Table S3.4).

As with the molecular average $\delta^{13}C$ values, for each synthesis, the $\delta^{13}C$ of the C-1 site decrease from α-APN to alaninamide to alanine (Figure 3.4a). Additionally, for each hydrolysis step, the $\delta^{13}C$ of the C-1 site is closer to the reactant cyanide's $\delta^{13}C$ (-31.8 ± 0.2 ‰) with increasing reaction progress of that step (Figure 3.4b and 3.4c, Table S3.4). Both of these observations indicate a normal KIE acting on the C-1 site of the reactant for reactions H-1 and H-2.

A Rayleigh plot of the evolution of the C-1 $\delta^{13}C$ value with progress of the H-1 reaction (which increases from right to left in this figure) yields a trend inconsistent with a single Rayleigh distillation process, implying at least two competing processes (Figure S3.2c): one that would enrich the α-APN pool available for hydrolysis in at least some reactions and a normal KIE from hydrolysis that would further enrich the residual α-APN. It would be reasonable, therefore, to assume that given the evidence for two competing fractionations in both the C-1 and molecular-average $\delta^{13}C$ analyses, a side reaction during the first hydrolysis step has impacted the perceived KIE and yield for H-1.

As with the full-molecular average $\delta^{13}C$, the $\delta^{13}C$ of the C-1 site has a normal KIE for the hydrolysis of alaninamide to alanine (step H-2), and the data are consistent with a single Rayleigh fractionation process with a single fractionation factor ($R^2 = 0.84$) (Figure 3.7) and the H-2 step is also the best constrained step for the C-1 site as it uses direct

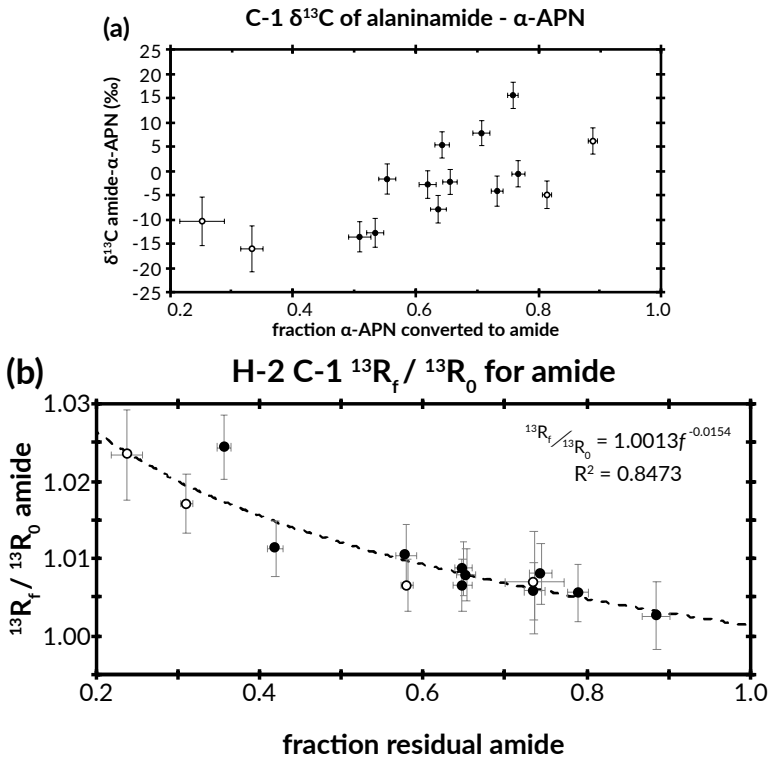


Figure 3.7: Carbon isotope concentrations for the C-1 site. (a)The difference in the α-APN reactant and amide product in H-1 with respect to the conversion of α-APN to amide. Here α-APN and the fraction of it converted are calulcted via the equilibrium values from Eqn. 4 to 17 in the text. The amide $\delta^{13}C$ is calcualted from the weighted sum of the measured $\delta^{13}C$ C-1 value of alanine and alaninamide at the end of H-2 and assume that the measured $\delta^{13}C$ of the C-2 and C-3 sites for alanine have the same values in alaninamide. (b) Ratio of residual to initial ^{13}R for the C-1 site in alaninamide for H-2.

measures of product and reactant abundances and direct measurements of the product's C-1 site $\delta^{13}C$. The normal 15.4 ‰ KIE for the C-1 site's hydrolysis from an amide to a carboxyl group (Table 3.3) is roughly three times that of the KIE found for the full-molecular average $\delta^{13}C$, further supporting the idea that the C-1 site is the only one that changes during H-2 (a corollary of our conclusion, above, that the average $\delta^{13}C$ of the C-2 + C-3 sites does not change during reaction H-2).

3.4. Discussion

3.4.1. Isotope Effects

Recall that the Strecker synthesis can be divided into the three steps depicted in Figure 3.2: (1) the equilibration between reactants and α-aminonitrile ('Equilibrium'), (2) the hydrolysis of the α-aminonitrile into an α-aminoamide ('H-1'), and (3) the subsequent hydrolysis of the α-aminoamide into an α-amino acid ('H-2') (Figures 3.1 and 3.2). Of these steps, the first occurs in equilibrium and the latter two are irreversible (Van Trump, 1975).

In circumneutral-to-basic hydrolysis conditions, the equilibrium step continues during the hydrolysis steps. The acidic hydrolysis conditions used in this study inhibit significant α-aminonitrile (α-APN) creation or decomposition following the equilibrium step. These conditions simplify the reaction pathway such that following the equilibrium step, the α-aminonitrile can be treated as an isolated reactant pool for the irreversible hydrolysis reactions. This simplification permits us to consider each of the three steps independently.

3.4.2. The Equilibrium Step

In the equilibrium step (Figure 3.2a), nitrogen from ammonia and carbon from cyanide and acetaldehyde's carbonyl carbon increase in bond order when they convert to α-APN (*i.e.,* the product α-APN has more bonds at the cyanide-, acetaldehyde- and ammonia-derived sites than do the reactants). Generally speaking, assuming a similarity of bond

types and bonding partners, when reactions are at equilibrium, we expect the compound with higher bond order to have a higher proportion of heavy-to-light isotopes. Consequently, we predict inverse equilibrium isotope effects will occur during the equilibrium step between cyanide carbon and the C-1 site in α-APN; between ammonia to the amine nitrogen site in α-APN, and between acetaldehyde's carboxyl carbon and the C-2 site in α-APN. Because the methyl carbon on acetaldehyde, which becomes the C-3 carbon in α-APN, is not involved in the reaction, we predict that it will be approximately equal in $\delta^{13}C$ in reactant and product.

Figure 3.2a displays the measured isotope effects, isotope ratios, and concentrations at equilibrium for reactants, α-APN, and side products. The $\delta^{15}N$ measurements agree with this scenario: we find a +56.6 ‰ inverse EIE between of α-APN's amine site and ammonia (Table 3.3). The calculated EIE is heavily impacted by the 29 ‰ EIE value we used for the reaction NH_3 (*aq*) ↔ NH_4^+ (*aq*) (Kirshenbaum et al., 1947; Walters et al., 2019), because at equilibrium roughly one-third of the initial ammonia pool is in the form of ammonium (29 ± 1 % to 38 ± 1 %) and only 3 % is in the form of ammonia. The large ratio of the concentration of ammonium to ammonia means that ammonium can sequester large amounts of ^{13}C and errors in this concentration or in the $\varepsilon_{NH4\text{-}NH3}$ value could result in errors in the $\varepsilon_{\alpha APN\text{-}NH3}$. We note that a recent study measured an $\varepsilon_{NH4\text{-}NH3}$ of 45 ‰ (Li et al., 2012) and past studies have found $\varepsilon_{NH4\text{-}NH3}$ values of 20 ‰ (Hogberg, 1997); however, 29 ‰ remains the $\varepsilon_{NH4\text{-}NH3}$ with the clearest consensus in both past and recent theoretical and experiments work (Walters et al., 2019). The 56.6 ‰ inverse EIE we infer agrees with our expectations and is likely due to the stronger bonding environment in α-APN (sp^3 hybridization) relative to ammonia (sp^2 hybridization).

In contrast, the estimated equilibrium isotope effect between acetaldehyde (lower in $\delta^{13}C$) and α-APN (higher in $\delta^{13}C$) did not follow our expectations, but instead exhibited a negative EIE of -10.0 ‰, averaged across the two relevant sites (Table 3.3, Figure 3.6a). As with the ammonia system, the residual acetaldehyde pool is split between two species, acetaldehyde (20 ± 1 % to 15 ± 1 %) and acetaldehyde hydrate (7 ± 1 % to 17 ± 1 %) (Figure 3.2a, Table S3.1). However, in this case, there is no experimentally verified

EIE for CH_3CHO (*aq*) ↔ $CH_3CH(OH)_2$ (*aq*). Consequently, we use values for the hydration of CO_2 (Marlier and O'Leary, 1984), to estimate an approximate $\varepsilon_{ace(OH)2\text{-}ace}$ value of 6.9 ‰ at the carbonyl site with the acetaldehyde hydrate enriched relative to acetaldehyde. This value could be incorrect especially in light of the past theoretic work that estimated $\varepsilon_{ace(OH)2\text{-}ace}$ as high as 40 ‰ at the carbonyl site (Hogg, 1980). Because the C-3 site is a spectator atom during all equilibrium steps (*i.e.*, it does not form or change bonds), we infer that the conversion of the acetaldehyde's carbonyl carbon to α-APN's C-2 site carries the entire EIE for the C-2 and C-3 sites, and thus that at equilibrium the carbonyl site of a-APN is 20 ‰ lower than that of acetaldehyde. This value does not support our hypothesis of an inverse EIE, but does agree with past work that also found the change from a double to single bond on the carbonyl carbon can lead to normal isotope effects up to -30 ‰ (O'Leary and Marlier, 1979; Marlier, 1993). Like Robins *et al.* (Robins et al., 2015), we argue the need for further SSIR work on the α-APN intermediate to elucidate the causes for the negative EIE at the C-2 site. Furthermore, future work that measures the EIE between acetaldehyde and acetaldehyde-hydrate would aid interpretation of our results.

3.4.3. The Hydrolysis Steps

Following equilibration, α-APN undergoes hydrolysis into alaninamide and then to alanine. These hydrolysis steps are irreversible and in a closed system, so associated isotope effects should be KIEs and the reactions are modeled as Rayleigh distillations. Of all sites in the final alanine, the two hydrolysis steps only directly involve the C-1 site: in H-1, the nitrile is converted into an amide and in H-2 the amide is converted into a carboxyl group. These steps are associated with significant bond reordering: during H-1, the C-1 site changes hybridization from sp to sp^2, and during H-2, the C-1 changes from sp^2 to sp^3 then back to sp^2. Theoretical and experimental work also suggests indirect involvement of the amine N during the two hydrolysis steps, by stabilizing the water lattice via hydrogen bonding; but this study predicts no involvement of the C-2 or C-3 sites (Yamabe et al., 2014). Given that the hydrolysis steps are irreversible and involve the formation of new bonds only at the C-1 site, we expect a primary normal KIE on the

C-1 site for each hydrolysis step and no isotope effects at the C-2 and C-3 sites. While we predict no primary isotope effects on the amine N, due to its potential in stabilizing intermediates we suspect an inverse secondary nitrogen isotope effect could occur, on the order of a few per mil.

The model calculations of $\delta^{13}C$ changes for the molecular average and the C-1 site of α-APN during reaction step H-1 are inconclusive, suggesting two competing fractionating processes (with one direction dominating early and the other late) and no consistent overall change; in contrast, our findings provide statistically strong and straightforward support for the expected primary KIEs during reaction step H-2 (Table 3.3). No significant variations in $\delta^{15}N$ and the average $\delta^{13}C$ of the C-2 and C-3 sites are observed during either hydrolysis step, suggesting they are spectator atoms, as expected (Figures 3.4b, 3.4c, 3.6b, and S3.1).

During H-1, the molecular-average $\delta^{13}C$ value has an overall positive trend with increasing reaction progress for both the $\delta^{13}C$ of product amide and the difference in $\delta^{13}C$ between the reactant α-APN and product amide (Figures 3.4b and 3.5b). The C-1 site appears to have these trends, too, although they are obscured by the high error bars for the C-1 measurements (Figures 3.4b and 3.7a). The difference in $\delta^{13}C$ between the reactant α-APN and product amide has a minimum value for the amide $\delta^{13}C$ subtracted from α-APN $\delta^{13}C$ at -15 ‰. However, the difference is resolvable above zero for some but not all syntheses at high (70 % - 80 %) reaction extents (Figure 3.5b), and the change in molecular average ^{13}R for α-APN versus the fraction of residual α-APN (Figure S3.2) have trends inconsistent with a single Rayleigh distillation process.

The discrepancies from a Rayleigh fractionation trend and potential for opposing isotope effects acting during reaction step H-1 lead us to conclude that H-1 included one or more side reactions and that dominated the reaction progress variable in some reactions. The reaction progress variable here is the fraction of remaining α-APN and was measured by the change in the initial moles of alaninamide (calculated from directly measured final moles of alaninamide and alanine) relative to the estimated moles of initial α-APN.

Assuming that a side reaction consumed α-APN, less alaninamide would be made from α-APN, which would appear in our model as a higher fraction of residual α-APN or a lower amide yield in that reaction. If the side reaction had a normal KIE, it would also enrich the residual α-APN reservoir and make the product alaninamide appear to be more enriched, especially when it has high yields (and thus does not express its own KIE as strongly). We suggest that a real understanding of the elementary isotope effects associated with the H-1 reaction step of the Strecker chemistry will require direct measurements of the amounts and of the isotopic compositions of both α-APN and alaninamide, including site-specific effects and will also require a wider sampling of reaction progress. For these reasons, our considerations of broader implications of this work focus on the more confident conclusions we can reach regarding the equilibrium and H-2 steps and the overall net isotope effects of the Strecker chemistry.

Data from the H-2 step affirms our expectations regarding KIE's associated with hydrolysis. Using the directly measured molecular average values for alaninamide and alanine, we find $\delta^{13}C$ normal 5.2 ‰ KIE on the molecular average $\delta^{13}C$ during H-2 (Figure 3.5c, Table 3.3) and a normal 15.4 ‰ KIE on the C-1 site (Figure 3.7b, Table 3.3) (in both cases alanine being ^{13}C-depleted relative to alaninamide). The normal 15.4 ‰ KIE on the C-1 site is roughly thrice that of the molecular average and both the models for the C-1 site, and equal to the minimum value for the $\delta^{13}C$ of amide's C-1 site subtracted from that of α-APN. We note that in Rayleigh plots (isotope ratio vs. amount of residual reactant) for both the molecular average and C-1 site carbon isotope values conform to the Rayleigh law with high correlation coefficients, suggesting a single mechanism with a constant KIE (Figures 3.5c and 3.7b). These facts support the assertion the molecular average KIE associated with converting amide to alanine is solely due the isotope effect associated with hydrolysis at the C-1 site (Figures 3.5c and 3.7b). These data also agree with past amide hydrolysis studies, which argue that the normal KIE results from the hydroxyl addition to the amide and the ensuing conversion from sp^2 to sp^3 hybridization prior to ammonia leaving (Robins et al., 2015) (Figure 3.2b).

3.5. Implications

3.5.1. Abiotic and biosynthetic isotopic fingerprints

Biological syntheses of amino acids result in carbon isotope structures that differ markedly from that produced by the Strecker (assuming precursors that are similar in $\delta^{13}C$), so a measurement of the carbon isotope structure of an amino acid of unknown origin might serve as way to distinguish biotic and abiotic processes. While the experiments here focus on alanine, the carbon sites that lead to different α-amino acids found in biological systems do not make or break bonds in Strecker synthesis. Consequently, we can assume that these sites will behave similarly to the methyl site (C-3) in this study and be spectators to the reaction. The C-1 and C-2 sites on α-amino acids found in biological systems have C-1 and C-2 sites in similar chemical environments as those in alanine, so we can assume that when these other α-amino acids are made via Strecker synthesis, they will share similar isotope effects to those measured here.

Studies by Abelson et al. (1961) find that alanine's C-1 site is between 0.2 ‰ enriched and 13.8 ‰ depleted in ^{13}C relative to starting carbon sources, that the C-2 and C-3 sites are between 11.8 ‰ and 20.5 ‰ depleted in ^{13}C relative to starting carbon sources, and that relative to the average $\delta^{13}C$ of the C-2 and C-3 sites, the C-1 site is 0.8 ‰ to 15.0 ‰ enriched in ^{13}C (Abelson and Hoering, 1961). Other biosynthetic α-amino acids in this study also had most carbon sites depleted relative to the input carbon and C-1 sites that were enriched relative to the rest of the carbons in the amino acid.

Our study demonstrates that Strecker synthesis causes a large ^{13}C depletion at the C-1 site of up to 35 ± 3 ‰ (potentially more under conditions and extends of reaction outside of the range explored here) and a moderate depletion of the averaged C-2 and C-3 sites (up to 16 ± 2 ‰ — an effect we infer results from the equilibration with acetaldehyde). Thus, the carbon isotope effect at the averaged C-2 and C-3 sites is of the same magnitude and sign as Abelson and Hoering found for biosynthetic amino acids, and the C-1 isotope

effect is of the same direction but far greater in amplitude. In amino acids other than alanine we would anticipate that this depletion of the C-2 site would remain the same but be diluted by the number of carbons in the amino acid's functional group. While in some cases, the $\delta^{13}C$ of the molecular average (or the C-1 or the averaged C-2 and C-3 sites alone) could not produce a unique isotopic signature that would readily distinguish Strecker synthesis from the biological production of amino acids, barring large variations in $\delta^{13}C$ values of initial aldehyde and cyanide, Strecker-synthesized alanine will have a C-1 site with lower ^{13}C abundances than that of the C-2 site averaged with the carbons in the functional group while the C-1 sites in biological sources will be ^{13}C-enriched relative to their C-2 site averaged with the functional group. Therefore, the intramolecular carbon isotope composition could provide a tool to in assigning biogenicity to samples of alanine. And, if the $\delta^{13}C$ values of potential substrates are known, one should be able to distinguish Strecker-synthesized from biogenic amino acids under all common circumstances.

Strecker synthesis also has a marked ^{15}N enrichment in alanine relative to initial ammonia and ammonium. As with carbon, because the ammonia is added to the carbonyl carbon on aldehydes to create any α-amino acid, we anticipate that this effect will be similar for all α-amino acids created by Strecker synthesis. Isotope effects for the microbial assimilation of nitrogen from ammonium into biomass range from -16 ‰ to +1 ‰ and -27 ‰ to 0 ‰ such that the biomass is typically depleted relative to the initial ammonium for terrestrial and aquatic ecosystems, respectively (Criss, 1999; Sigman et al., 2009). These values, especially those in aquatic systems, are both lower in magnitude and opposite in sign to the EIE between ammonium and a Strecker-synthesized α-amino acid's amine site. Starting from the same ammonium supply, therefore, we would anticipate that biological alanine would be ^{15}N-depleted relative to Strecker-synthesized α-amino acids, by up to 81 ‰. This large difference should be sufficient to distinguish biotic and Strecker-produced α-amino acids (particularly when combined with carbon isotope evidence outlined above). However, some high trophic level organisms could have amino acids with $\delta^{15}N$ similar to those synthesized by Strecker synthesis because $\delta^{15}N$ in amino acids tends to increase with trophic level by 3 to 4 ‰ per trophic level and

can increase by up to 8 ‰ per trophic level for glutamate (Bowes and Thorp, 2015) (Figure 3.8a). In these organisms, increases in $\delta^{15}N$ tend to occur with no or slight (0 to 2 ‰) positive changes in $\delta^{13}C$ per tropic level (Perkins et al., 2014). Therefore, in these cases, using the depletions in $\delta^{13}C$ combined with the enrichments in $\delta^{15}N$ could serve as a molecular-average fingerprint for abiotically synthesized α-amino acids (Figure 3.8b). And, again, inclusion of site specific carbon isotope measurements will improve this ability to discriminate biotic from abiotic nitrogen isotope signals.

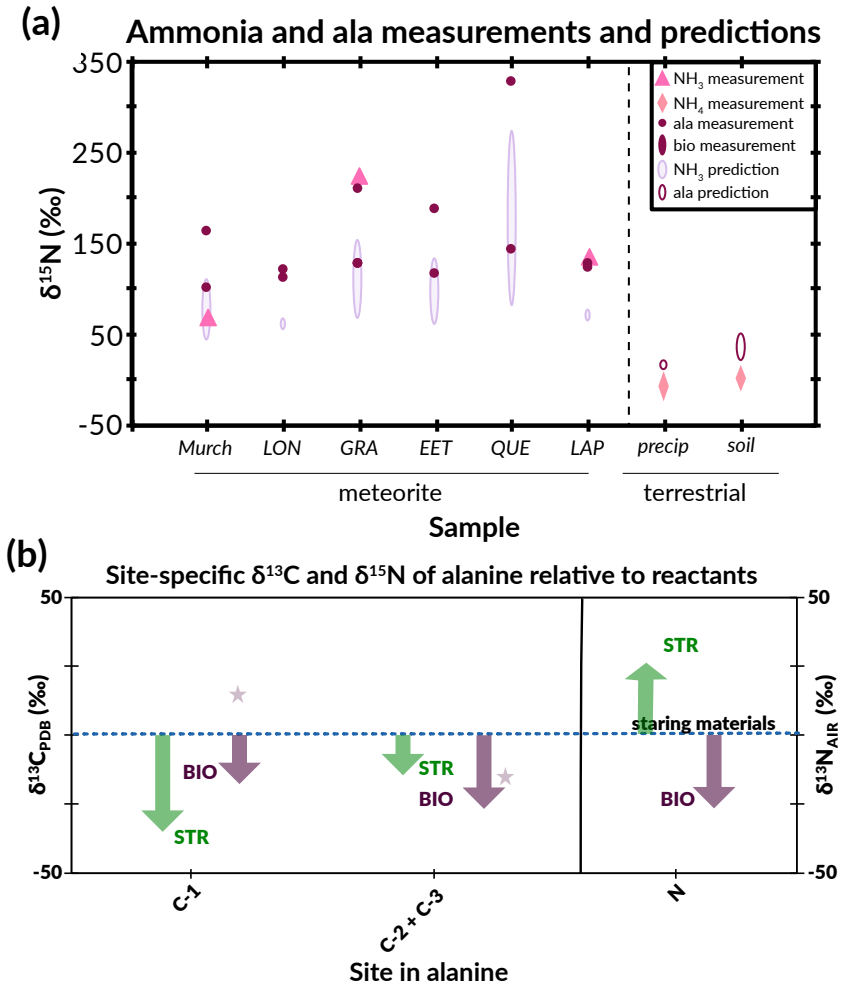


Figure 3.8: $\delta^{15}N$ predictions for ammonia and alanine. (a) $\delta^{15}N$ predictions for ammonia on meteorites based on alanine measurements assuming the alanine is produced by Strecker synthesis and predictions for terrestrial alanine produced by Strecker synthesis from precipitation or soil ammonium. On the left, $\delta^{15}N$ values of alanine from meteorites (filled burgundy circles) are used to predict possible $\delta^{15}N$ ranges ammonia used to synthesize it (open lavender ovals). Where possible, $\delta^{15}N$ ammonia measurements from meteorites (filled pink triangles) are included for comparison. On the right, measured global $\delta^{15}N$ ammonium ranges for precipitation and soil (filled pink diamonds) are used to predict $\delta^{15}N$ alanine (open burgundy ovals). (b) Site-specific $\delta^{13}C$ and $\delta^{15}N$ measurements for Strecker- and biosynthesized alanine relative to input material. The left panel shows $\delta^{13}C$ ranges relative to CN^- and acetaldehyde for C-1 and C-2 + C-3, respectively, for Strecker synthesis (green), and CO_2 for biosynthesis (burgundy) (Abelson and Hoering, 1961). Stars represent alanine biosynthesized in CO_2-limited conditions. The right panel shows $\delta^{15}N$ ranges relative to the input ammonium for Strecker synthesis and biosynthesis. Biosynthesized data is for all biomass and is from Criss *et al* (1999) and Sigman *et al* (2009).

3.5.2. Amino acids in carbonaceous chondrites

It has been previously suggested that Strecker synthesis could be an abiotic route to amino acid synthesis on parent bodies of the carbonaceous chondrites, and perhaps on other planetary bodies (Cronin et al., 1993; Pizzarello and Shock, 2010; Burton et al., 2012). This suggestion was recently supported by a study of the site-specific carbon isotope structure of alanine (Chimiak et al., 2020). If this interpretation is correct, we can combine the experimental constraints on N isotope fractionations of the Strecker synthesis presented here with prior measurements of $\delta^{15}N$ values of alanine recovered from the carbonaceous chondrites (Pizzarello et al., 1994; Pizzarello and Holmes, 2009; Pizzarello et al., 2011; Elsila et al., 2012) to estimate the $\delta^{15}N$ values for ammonia from which alanine formed (Figure 3.8a). The predictions agree with independent measurements of the $\delta^{15}N$ of ammonia recovered from Murchison, but underpredict measurements of ammonia from GRA95229 and LAP02342. The discrepancies might reflect the fact that alanine on these two samples formed by mechanisms other than (or in addition to) Strecker synthesis, differences in pH (which cause differences in proportions of NH_4 to NH_3 — a significant factor in the nitrogen isotope budget of the 'equilibrium' step of the Strecker synthesis) , or differences between these samples in post-synthetic preservation of alanine and/or ammonia.

3.6 Conclusions

Strecker synthesis has three main steps. The first step is the creation of α-APN which has a +56.6 ‰ nitrogen isotope EIE between reactant ammonia and the amine site of α-APN and a -20.0 ‰ carbon isotope EIE between the carboxyl site of reactant acetaldehyde and the C-2 site of α-APN. Isotopic values for the amine nitrogen and for the averaged C-2 and C-3 site carbons do not significantly change from their values in α-APN during subsequent hydrolysis of α-APN to alanine.

The hydrolysis of α-APN to alaninamide likely has a normal KIE, but all the carbon isotope measurements (molecular-average and site-specific) have large departures from

simple Rayleigh behavior, either because of an uncharacterized side reaction or because of uncertainties in our calculations of progress in the H-1 reaction and evolution in $\delta^{13}C$ of α-APN. To better understand the isotope chemistry of the H-1 reaction step, future experiments should focus on α-APN hydrolysis alone. However, this uncertainty does not influence our evaluation of the equilibrium isotope effects or the isotope effects associated with H2 to create alanine.

The final step in Strecker synthesis (H-2) is the hydrolysis of alaninamide to alanine. During H-2, the C-1 site has a -15.4 ‰ KIE for the conversion of alaninamide to alanine and is the only site to have a significant isotope effect during this step.

On a molecular-average level, a consistently depleted $\delta^{13}C$ in conjunction with an enriched $\delta^{15}N$ (relative to substrates) is a signal we document for the Strecker synthesis that is unlikely to be produced biologically. Using two conventional measurements—the molecular average $\delta^{15}N$ and $\delta^{13}C$ of alanine—one should have significant ability to forensically determine whether an amino acid was produced by previously studied terrestrial life or by Strecker synthesis. This ability to distinguish biotic and Strecker sources improves further with addition of site-specific carbon isotope data, as the ~15 ‰ depletion of the C-1 carbon relative to the C-2 and C-3 carbons (assuming substrates of approximately uniform $\delta^{13}C$) is a noteworthy signature of alanine produced by Strecker chemistry.

Chapter 4

CARBON ISOTOPE EVIDENCE FOR THE SUBSTRATES AND MECHANISMS OF PREBIOTIC SYNTHESIS IN THE EARLY SOLAR SYSTEM

Chimiak L., Elsila J. E., Dallas B., Dworkin J. P., Aponte J. C., Sessions A. L. and Eiler J. M. (2020) Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system. *Geochim. Cosmochim. Acta*. doi: 10.1016/j.gca.2020.09.026

Abstract

Meteorites contain prebiotic, bio-relevant organic compounds including amino acids. Their syntheses could result from diverse sources and mechanisms and provide a window on the conditions and materials present in the early solar system. Here we constrain alanine's synthetic history in the Murchison meteorite using site-specific $^{13}C/^{12}C$ measurements, reported relative to the VPDB standard. The $\delta^{13}C_{VPDB}$ values of -29 ± 10 ‰, 142 ± 20 ‰, and -36 ± 20 ‰ for the carboxyl, amine-bound, and methyl carbons, respectively, are consistent with Strecker synthesis of interstellar-medium-derived aldehydes, ammonia, and low-$\delta^{13}C$ nebular or interstellar-medium-derived CN. We report experimentally measured isotope effects associated with Strecker synthesis, and use them to constrain the $\delta^{13}C$ values of the alanine precursors, which we then use to construct a model that predicts the molecular-average $\delta^{13}C$ values of 19 other organic compounds of prebiotic significance found in Murchison if they were made by our proposed synthetic network. Most of these predictions agree with previous measurements, suggesting that interstellar-medium-derived aldehydes and nebular and/or pre-solar CN could have served as substrates for synthesis of a wide range of prebiotic compounds in the early solar system.

4.1. Introduction

Carbonaceous chondrite (CC) meteorites contain amino acids (Cronin and Moore, 1971; Engel et al., 1990; Glavin et al., 2018), the extraterrestrial origins of which are evinced by

their chemical and isotopic properties. Known life predominantly synthesizes 20 amino acids that are mostly L enantiomers and ~2 % lower in their $^{13}C/^{12}C$ ratios than the average terrestrial inorganic carbon. On the other hand, the CC meteorites contain over 90 amino acids that are nearly racemic mixtures of D and L enantiomers—likely unchanged since their arrival on Earth—and are generally ~1-3 % higher in their $^{13}C/^{12}C$ ratios than the average terrestrial inorganic carbon (Martins and Sephton, 2009; Burton et al., 2012; Elsila et al., 2016; Glavin et al., 2018).

Proposed mechanisms of meteoritic amino acid synthesis include *i)* ion/radical-molecule reactions in the interstellar medium (ISM) (*e.g.*, with the irradiation of methanol ices (Bernstein et al., 2002)), *ii)* Fischer-Tropsch type (FTT) synthesis in the protosolar nebula (Botta and Bada, 2002), and/or *iii)* aqueous chemistry (*e.g.*, Strecker synthesis or reductive amination) of ISM-derived precursor molecules that were accreted in ices by the meteorite parent bodies and reacted during aqueous alteration (Kerridge, 1999; Pizzarello et al., 2006; Glavin et al., 2018). The molecular-average $\delta^{13}C$ values[1] of individual α-amino acids from the Murchison CM2 CC decrease systematically with increasing carbon number (Pizzarello et al., 1991; Sephton, 2002; Elsila et al., 2012; Glavin et al., 2018), suggesting that they might have been assembled from smaller precursors with each newly added carbon atom being lower in ^{13}C than its source due to kinetic isotope effects (KIEs) (Yuen et al., 1984; Engel et al., 1990; Sephton, 2002). Alternatively, these trends could reflect the dilution of a high-$\delta^{13}C$ carbon atom inherited from ISM-derived CO by carbon from other, lower $\delta^{13}C$ precursors (Elsila et al., 2012). However, in the full set of prior $\delta^{13}C$ measurements of Murchison α-amino acids, $\delta^{13}C$ variations for individual amino acids compared between studies span a range equal to the extent of the proposed correlation between carbon number and $\delta^{13}C$ and so calls these explanations into question.

---

[1] $\delta^{13}C$ quantifies the ratio of $^{13}C/^{12}C$ relative to a standard, Vienna Pee Dee Belemnite (VPDB). Mathematically, $\delta^{13}C_{VPDB} = \frac{{}^{13}C/{}^{12}C_{sa}}{{}^{13}C/{}^{12}C_{st}} - 1$ where *sa* denotes the sample and *st* the VPDB standard, which has a ^{13}R value of 0.01118 (Brand et al., 2010). $\delta^{13}C$ is conventionally reported in parts per thousand (‰), *i.e.*, $= \left[\frac{{}^{13}C/{}^{12}C_{sa}}{{}^{13}C/{}^{12}C_{st}} - 1 \right] \times 1000$

These and other hypotheses regarding the origins of meteoritic amino acids can be tested through observations of their site-specific carbon isotope distributions (*i.e.*, the $\delta^{13}C$ values of individual carbon positions in each molecule). Here we present site-specific $\delta^{13}C$ measurements of the three carbon sites in alanine extracted from a sample of Murchison and measured using novel techniques conducted with an Orbitrap mass spectrometer.

4.2. Methods and Materials

4.2.1. Materials

4.2.1.1. Meteorite

We analyzed two samples of Murchison meteorite, a Methods Development sample (analyzed winter and spring 2018) and an Analytical sample (analyzed summer 2018). The Methods Development sample was a 5 g piece of Murchison from the Field Museum of Natural History via Clifford N. Matthews's research group that was known to be contaminated; although this contamination means that analytical results are of limited value, it provided a natural sample on which we could assess our novel analytical procedures. The Analytical sample was a 2.6 g sample from a different piece of Murchison and the same source; the sample has been described in Friedrich *et al.* (2018).

The D/L ratio of alanine from the methods development sample is 0.4, which is far from a pure racemic mixture's value of 1 or past measurements and therefore suggests a high proportion of terrestrial contamination. The D/L ratio of alanine from the analytical sample is 0.85, which agrees with past measurements of Murchison alanine (Cronin et al., 1995). The overall amino acid content of the Analytical sample is also similar to those measured previously in Murchison (Cronin and Moore, 1971; Martins and Sephton, 2009; Friedrich et al., 2018), which combined with the D/L ratios of amino acids in this sample suggests minimal terrestrial contamination.

4.2.1.2. Derivatization Materials

Alanine standards used in this study were Alfa Aesar L-alanine (99 % Purity) and one sample of alanine synthesized by Strecker synthesis (Purity confirmed by NMR, Figure S4.1 in Appendix C) (hereafter, 'Strecker standard'). In methods development, we also used alanine purchased from VWR (Purity >99 %, Lot # 2795C477) as a standard. The Alfa Aesar standard was synthesized via microbial aspartate fermentation. The VWR alanine standard, which has similar site-specific $\delta^{13}C$ values, was synthesized by fermentation by *Pseudomonas*. Origins of the aspartate precursor are unknown, but Hoffman and Rasmussen (Hoffman and Rasmussen, 2019), who studied supplier-bought alanine samples, demonstrate that in two different samples $\delta^{13}C$ of all sites are within 10‰ of each other. Calibration of standards is described in Appendix B.1. Ultrapure water was obtained from a MilliPore ultrahigh-purity water (18.2 MΩ cm; hereafter 'water') system at Caltech. In addition to the standards listed above, alanines with 99 % ^{13}C label at C-1, C-2, or C-3 were purchased from Sigma Aldrich (C-1: Lot # EB2220V, C-2: Lot # SZ0643V, C-3: Lot # EB2211V).

Reagents used in derivatization reactions and cleaning at Caltech included: anhydrous methanol (MeOH; >99.8 % purity, Macron Fine Chemicals, Batch# 0000042997), n-hexane (>98.5 % purity, Millipore Sigma, HPLC grade, multiple lots), acetyl chloride (AcCl; >99 % purity, Sigma Aldrich, Lot# BCBT8141), trifluoroacetic anhydride (TFAA; >99 % purity, Sigma Aldrich, Lot# SHBJ0051), and dichloromethane (DCM, Sigma Aldrich, HPLC Plus, >99.9 % purity). All derivatizing reagents were tested for amino acid contamination prior to use on samples (See Appendix B.2 for more details).

Prior to derivatization, glassware was cleaned with ultrapure water and combusted twice at 450 °C. The second combustion occurred the night before use and with no other glassware present. GC vial PTFE caps were new and handled with forceps that were pre-cleaned with methanol. Any cap whose interior was touched with forceps was discarded. Fumehoods and tubing for nitrogen gas were cleaned prior to derivatization. BioPur pipette tips were used on pipettes to prevent contamination. Chemical lab

syringes (Hamilton, 250 µL) were cleaned with methanol prior to and after derivatization reactions, and instrument inlet syringes (Hamilton, 10 µL) were cleaned with 30 µL hexane and 30 µL DCM between and before all analyses.

4.2.2. Methods

4.2.2.1 Amino Acid Extraction

Amino acids were extracted from meteorite samples at NASA Goddard Space Flight Center (GSFC) following the protocol from (Elsila et al., 2012). Briefly, each sample was ground to a homogenized powder and sealed in a glass ampoule with 1mL ultrahigh purity water (Millipore Integral 10 UV, 18.2 MΩ cm, <3 ppb total organic carbon) for 24 hours at 100 °C. The water extract was separated, dried under vacuum, and hydrolyzed in 6N HCl vapor (Sigma Aldrich, double distilled) for 3 hours at 150 °C. This hydrolyzed extract was then desalted on a cation-exchange resin column (AG50W-X8, 100-200 mesh, hydrogen form, Bio-Rad), with the amino acids recovered by elution with 2 M NH_4OH (prepared from ultrahigh purity water and NH_3 (g) *in vacuo*); this eluent was split into two fractions and dried under N_2. The Methods Development sample was processed in this way in November 2017 and the analytical sample in May of 2018.

Upon arrival at Caltech, extracts were triple bagged, boxed, and stored in a freezer. One extract from the Methods Development sample was derivatized and analyzed in December 2017; the other was derivatized and analyzed in March 2018. A portion of each derivatized extract from the Methods Development sample was sent back to GSFC along with derivatized standards for secondary analysis. The extract from the Analytical sample was split between GSFC (85 %) and Caltech (15 %). Analyses were made on the analytical sample in June 2018 at GSFC and between June to July 2018 at Caltech.

4.2.2.2. Derivatization

A flow chart for handling of samples and blanks are depicted in Figure S4.2 in Appendix C. First, 1.0 mL of water:MeOH (3:1) was added to the centrifuge vials containing meteorite extract that had been shipped from GSFC. Vials with the Methods Development samples were capped, placed in a beaker of water, and sonicated for five minutes. The analytical sample extract sat in the water-methanol mixture for 20 minutes but was not sonicated. Samples were then uncapped and transferred into 2 mL GC vials ('sample vials') via combusted glass pipettes. All original shipped extract vials were rinsed twice more with the 3:1 water-methanol mixture without sonication. The rinse liquid was again transferred into the sample vials via glass pipette. Between the second and third rinse and following the third rinse, GC samples vials were dried under slow N_2 flow.

Standards and Murchison extract samples were then derivatized as N-trifluoroacetyl-O-methyl esters. Samples were brought up in 100 µL of anhydrous MeOH and placed in an ice bath. Using a clean syringe, 25 µL of AcCl was added dropwise to the sample, which was swirled between drops to limit localized boiling (the reaction with AcCl is strongly exothermic). Forceps were used to lift vials and swirl them in order to minimize potential contamination. Samples were capped and heated to 70 °C in a heating block for 1 hour. Samples were then cooled and dried under N_2. To avoid cross-contamination, all samples, blanks, and standards were dried separately. Next, 120 µL hexane and 60 µL TFAA were added and vials were capped and heated to 60 °C in a heating block for 30 minutes. Samples were evaporated under N_2 until 50 µL remained. We ended evaporation while ~100uL of solvent still remained to avoid evaporation of the amino acid derivatives. Following this, 500 µL hexane was added for the methods-development samples and 200 µL hexane was added for the analytical sample. In initial experiments on Alfa Aesar, VWR, and Strecker alanine derivatives, evaporation was carried out until only derivative remained, as determined by gravimetry. Isotopic analysis of these samples indicated that no site-specific carbon isotope effects occurred during evaporation (within measurement errors).

A split of the analytical sample extract and Caltech alanine standards were also derivatized as N-trifluoroacetyl-O-isopropyl esters at GSFC following protocols from (Elsila et al., 2012) at GSFC. The methods are similar to those listed above but use isopropanol instead of methanol.

4.2.2.3. Amino Acid Characterization

Amino acid abundance and enantiomeric composition (*e.g.*, abundances of D- and L-alanine) of both the method development and analytical samples were measured at GSFC via liquid chromatography with fluorescence detection and time-of-flight mass spectrometry (LC-FD/ToF-MS) using methods described in (Glavin et al., 2010). For the methods development samples, 1 % of the sample was used for amino acid characterization, while for the analytical sample, 0.4 % of the initial 2.6 g sample was used for characterization (details in (Friedrich et al., 2018)).

4.2.2.4. Isotope Ratio Measurements

4.2.2.4.1. Molecule-average isotope analysis of Murchison samples

Approximately 99 % of the methods development sample was sent as two splits to Caltech where it was derivatized as N-trifluoroacetyl-O-methyl ester (See *2.2.2:* Derivatization for further details) on the two analysis dates (winter and spring 2018). One aliquot of each derivatized sample in addition to two derivatized standards (Strecker and Alfa Aesar) were sent back to GSFC where they were analyzed for molecular-average (combining both chiral forms) $\delta^{13}C$ via Gas Chromatography-Combustion-IRMS (GC-C-IRMS) with a TC-5LIMS 30 m column. For the analytical sample, the 85 % that remained at GSFC was derivatized as N-trifluoroacetyl-O-isopropyl ester and injected into a GC-MS with four 25 m Chirasil-L-Val columns (Agilent, CP7495) connected in series. This long chiral column allowed us to separate and measure the $\delta^{13}C$ values of D- and L-alanine.

4.2.2.4.2. Site-specific isotope analysis

Site-specific carbon isotope ratios of derivatized alanine samples and standards were measured at Caltech The constraints presented in this paper are based on measurements of the bulk carbon isotope ratio of the full molecule by GC-combustion IRMS (yielding the average $\delta^{13}C$ of C-1, C-2, and C-3), along with direct mass spectrometric measurements of $^{13}C/^{12}C$ ('^{13}R') of two fragment ions of the alanine derivative, one of which constrains the average ratio for C-1 and C-2 and the other of which constrains the average ratio for C-2 and C-3. For carbon number identities, see Figure 4.1. These three independent constraints permit us to calculate the $\delta^{13}C$ of each of C-1, C-2 and C-3 (see *2.3*: Data Processing).

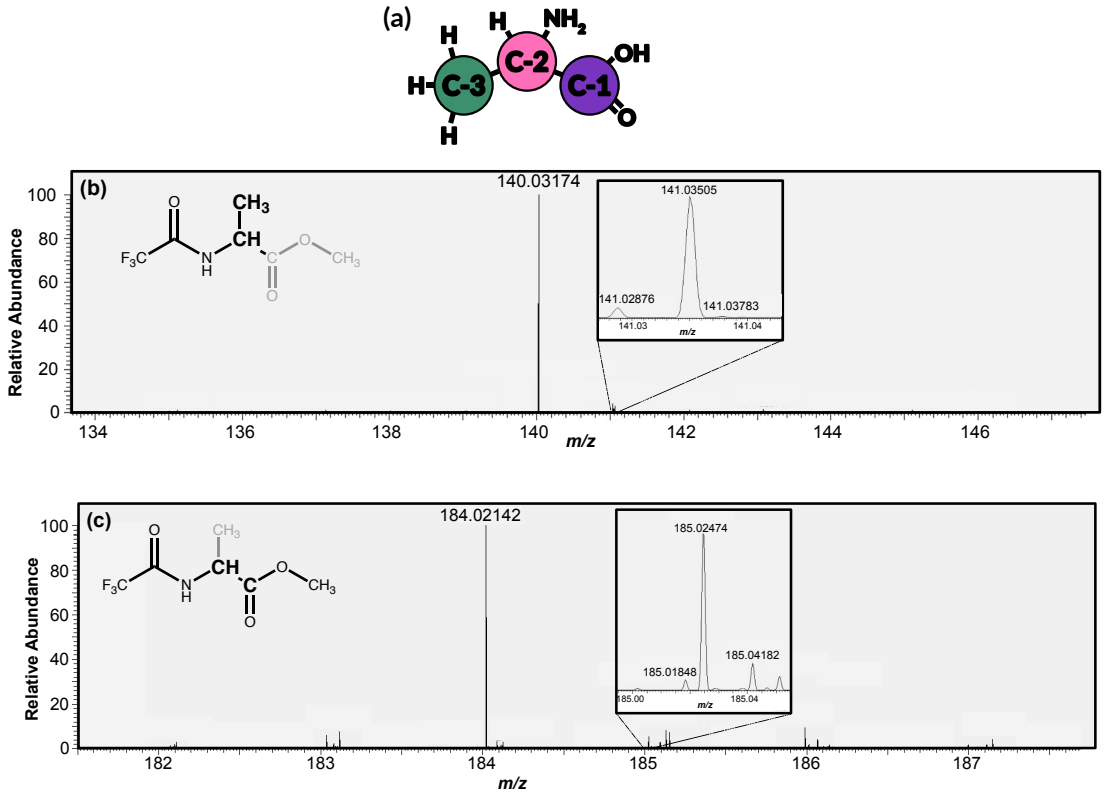


Figure 4.1: Mass spectra and fragment images for alanine from Murchison meteorite sample measured in this study. (a) Alanine with carbon sites are labelled. Mass spectra and labelled fragment images are presented for (b) *m/z* 140 and (c) *m/z* 184. In panels (b) and (c), fragments are in black with the rest of the derivative in gray, and carbon sites from alanine being measured are bolded. Insets for each mass spectra displays ^{13}C-substituted peak to demonstrate resolution from potential isobars.

The fragment ion measurements were made using the Q-Exactive GC Orbitrap mass spectrometer (hereafter 'Orbitrap'), using techniques described in (Eiler et al., 2017). The

Orbitrap can mass resolve a ^{13}C substitution from D, ^{15}N, or ^{17}O substitutions allowing a user to measure the ^{13}R of a fragment directly (*e.g.,* without combusting a molecule and then converting carbon into CO_2) (Figure 4.1c and 4.1d insets). The measured fragment ions have monoisotopic peaks (*i.e.*, the isotopologue containing only the most abundant isotope of each element, also known as the 'unsubstituted' isotopologue) of mass/charge (*m/z*) 140.0317 ($C_4H_5ONF_3$) and 184.0214 ($C_5H_5O_3NF_3$) Da (Figure 4.1). Measurements of their isotope ratios will be referred to by their monoisotopic mass (*i.e.*, 140.032 for the $^{13}C/^{12}C$ ratios derived from ions with masses 141.0350 and 140.0317 Da). The relative contributions of the carbon sites from the parent molecule to each fragment ion were determined by analyzing three mixtures, each with a 10 % ^{13}C enrichment at one carbon site (produced by mixing an appropriate site-specific, 99 % labeled alanine with the unlabeled standard). The *m/z* = 140.032 fragment contains both C-2 and C-3 from the parent alanine along with two carbons from the TFAA reagent (Figure 4.1c). The *m/z* = 184.021 fragment contains C-1 and C-2 from parent alanine along with all three carbons from the derivatizing reagents (Figure 4.1d). From labelling studies, both appear to be direct fragmentation products with no obvious evidence for recombination reactions that may add carbon atoms from one sample site into a different molecular ion.

The methods of high-precision isotope ratio analysis by Orbitrap-based mass spectrometry are described in Eiler *et al.* (Eiler et al., 2017). For the measurements presented in this paper, two configurations were used: direct analysis of analyte peaks eluting from a GC column ('Direct Elution') and analyte capturing from the GC effluent into a reservoir followed by isotopic analysis as it drained from that reservoir ('Reservoir Elution') (Figure 4.2). The Direct Elution mode was used to characterize the fragmentation pattern and retention time of derivatized alanine (Figure 4.2c). For this study, analyte eluting from the GC column was admitted to the ion source continuously following a 5.5-minute delay to avoid the solvent peak. Pre-mass selection using the AQS (quadrupole) system was set to permit all ions between *m/z* 50 and 300 Da to enter the Orbitrap mass analyzer. Reservoir Elution mode was used to measure ion-abundance ratios at a useful precision for study of natural stable isotope variations. Here Reservoir Elution mode measurements were conducted with an initial 5.5-minute solvent delay

followed peak monitoring in Direct Elution mode until 30 seconds prior to the elution of derivatized alanine, which could be timed relative to the retention times of earlier-eluting compounds (Figure 4.2d). At this point, the effluent from the GC column was rerouted into either a 5 or 20 cm^3 glass reservoir, the contents of which were flushed with He into the ion source.

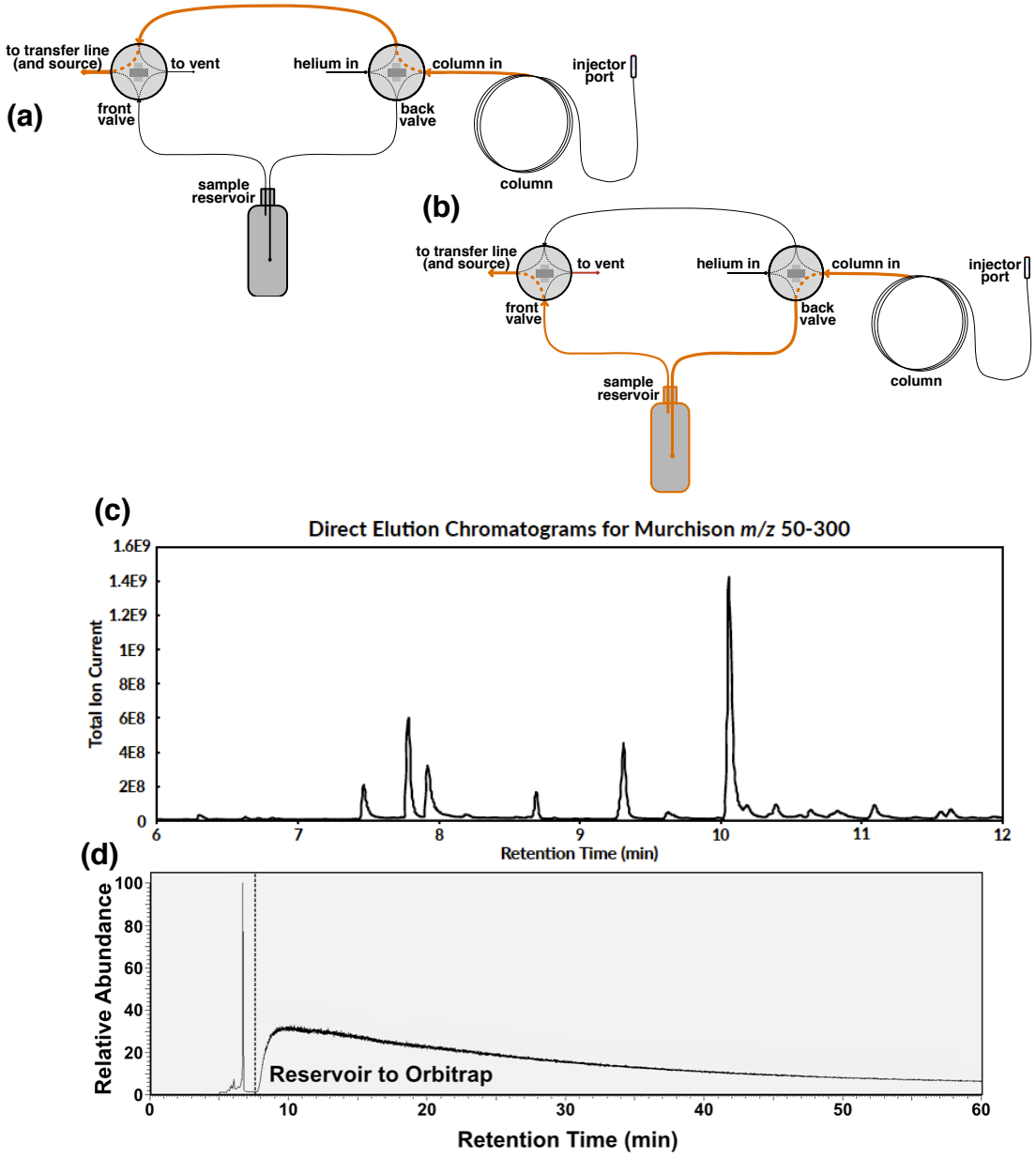


Figure 4.2: Schematic of custom inlet system for Orbitrap for (a) direct injection and (b) reservoir elution. (c) Chromatogram of 50-300 Da for direct injection that was used to find elution of alanine in Murchison sample. (d) Chromatogram of the 140.032 Da peak for reservoir elution during a typical measurement for Murchison sample.

The 20 cm^3 reservoir was used for measurements of the relatively high intensity 140.032 fragment and the 5 cm^3 reservoir was used for the weaker intensity 184.021 and 113.021 fragments, in order to increase signal-to-noise ratios (For more information on the 113.021 fragment, see Appendix B.3). Following the total collection of the derivatized alanine peak, GC column effluent was vented and clean helium was directed into the reservoir to continue purging analyte into to the ion source for the remainder of the measurement. In this fashion the glass reservoir serves as an exponential-dilution flask (Merritt and Hayes, 1994) that broadens the analyte peak from a few seconds to tens of minutes and thereby facilitates the accumulation of more ion counts—and thus greater precision for isotope ratios—by the Orbitrap. Alanine measurements were accumulated for 10- to 60-minutes depending on the reservoir size and the abundance of the fragments of interest (Figure 4.2d).

For detailed information on blanks and background analysis see Appendix B.2. In short, background scans were taken prior to each set of Murchison injections to ensure that no background alanine was present at intensities that could significantly impact the measured isotope ratios of the sample. In cases in which alanine or other contaminants were present in the reservoir or column, solvent blanks were run with only helium entering the reservoir until backgrounds subsided. Measurements of $^{13}C/^{12}C$ displayed no evidence of drift over the period of days during which the data was acquired (Electronic Annex 1) and measurements of the Strecker standard's $\delta^{13}C$ of the C-2 + C-3 sites directly following those of the Alfa Aesar standard were within 2 standard errors of independently known values (Appendix B.1). These factors suggest an absence of memory effects in the instrument.

Orbitrap measurements produce a series of 'scans', each of which reports the apparent ^{13}R of a selected fragment (*i.e.*, for the m/z = 140.032 or 184.021 fragments; see values in Electronic Annex 1). Measurements begin when the alanine peak elutes (*i.e.*, when the intensity of the monoisotopic peak is at its minimum immediately prior to alanine's elution). To minimize mass spectrometric artifacts (Eiler et al., 2017), we accept only those analyses in which both the monoisotopic and singly ^{13}C-subsituted fragments are

present, in which the monoisotopic ion makes up at least 30 % of the total ion current (TIC) in the observed mass window (Electronic Annex 1), and in which the product of the TIC and injection time (IT) varies over a narrow range (~10's of %, relative) between scans. In some cases, the trace of ion intensities provides evidence that we failed to capture all of the alanine peak in the reservoir and/or a nearly co-eluting peak has been captured with it (*e.g.*, in the case of the 113.021 peak of the Murchison sample discussed in Appendix B.3); these measurements are also discarded as procedural failures. Standard errors were calculated as the standard deviation of all accepted scans ^{13}R values for a fragment divided by the square root of the number of scans for that fragment.

The accuracy and precision of site-specific measurements was verified via a comparison of differences in $\delta^{13}C$ of C-1 measured by the Orbitrap with that measured by ninhydrin decarboxylation for the three standards (See Appendix B.1 for a detailed discussion). The average $\delta^{13}C$ values for C-2 and C-3 of Strecker alanine relative to the Alfa Aesar alanine standard measured by the 140.032 Da fragment on the Orbitrap during our experiments was -17.4 ± 1.6 ‰ (See Appendix B.4 for Error Analysis). This value is just beyond 2 standard errors from the value found from subtraction of $\delta^{13}C$ C-1 from the molecular-average $\delta^{13}C$ measured by ninhydrin decarboxylation and molecular-average EA-IRMS measurements for Strecker alanine relative to the Alfa Aesar alanine standard (-13.4 ± 0.6 ‰).

Differences in the isotopic composition between the Alfa Aesar and Strecker standards' fragments were constant within the nominal external errors of each measurement over the course of our analysis (Electronic Annex 1) and between analysis sets (Table 4.1). Each standard had stable measurements of each fragment's ratios of the ^{13}R over the course of our measurements: when source backgrounds are low, the standard deviation for Alfa Aesar's ^{13}R between different injections normalized to the measurements' averages are 4.6 ‰ and 14.0 ‰ for the *m/z* 140.032 and 184.021 fragments, respectively, for quantities of analyte similar to those of Murchison extracts. Furthermore, the variation decreases with increasing quantity of analyte (*i.e.*, the samples that most vary from the mean tend to be of lower intensity fragments and/or measurements) because ^{13}C counts

Analysis Set	Sample	Molecular Average $\delta^{13}C_{VPDB}$ (‰)	Fragment ^{13}R (‰)		Fragment $\delta^{13}C_{AlfaAesar}$ (‰)		Fragment $\delta^{13}C_{VPDB}$ (‰)	
			140	**184**	**140**	**184**	**140**	**184**
Winter 2018	Alfa Aesar	-19.4 *(0.2)*	0.04313 *(0.00004)*	0.05392 *(0.00018)*	x	x	x	x
	Strecker	-32.9 *(0.1)*	0.04297 *(0.00004)*	0.05381 *(0.00012)*	-7.4 *(2.8)*	-5.5 *(9.9)*	**-22.1** *(2.8)*	**-27.0** *(9.9)*
	Methods Development Murchison	17 *(4)*	0.04379 *(0.00004)*	0.05458 *(0.00020)*	30.5 *(2.8)*	30.3 *(12.2)*	**15.3** *(2.8)*	**8.0** *(12.2)*
Spring 2018	Alfa Aesar	-19.4 *(0.2)*	0.04323 *(0.00003)*	0.05389 *(0.00028)*	x	x	x	x
	Strecker	-32.9 *(0.1)*	0.04289 *(0.00002)*	x	-15.5 *(2.0)*	x	**-30.1** *(2.0)*	x
	Methods Development Murchison	17 *(4)*	0.04381 *(0.00003)*	0.05503 *(0.00024)*	27.3 *(2.1)*	53.0 *(18.9)*	**12.1** *(2.1)*	**30.3** *(18.9)*
Summer 2018	Alfa Aesar	-19.4 *(0.2)*	0.04237 *(0.00002)*	0.05538 *(0.00010)*	x	x	x	x
	Strecker	-32.9 *(0.1)*	0.04203 *(0.00003)*	0.05493 *(0.00018)*	-16.0 *(1.8)*	-20.2 *(9.3)*	**-30.5** *(1.8)*	**-41.4** *(9.2)*
	Analytical Murchison	25.5 *(3)*	0.04382 *(0.00003)*	0.05714 *(0.00017)*	68.4 *(1.5)*	79.8 *(8.9)*	**52.6** *(1.5)*	**56.4** *(8.9)*

Table 4.1: Fragment ^{13}R values and $\delta^{13}C$ values (AA and VPDB scales) for samples, standards, blanks. All delta values are dilution corrected. Standard error values are listed in parentheses. The first two columns of data (Molecular-average $\delta^{13}C$ and Fragment ^{13}R) were directly measured while Fragment $\delta^{13}C$ values relative to Alfa Aesar and VPDB were calculate using equation S1. The $\delta^{13}C$ values used in the Monte Carlo simulation are in the last columns (Fragment $\delta^{13}C_{VPDB}$).

increase with analyte quantity, and the instrument's shot noise limit is inversely proportional to the square root of the number of ^{13}C counts.

We tested our ability to trap and analyze derivatized alanine in amino acid mixtures by measuring alanine in a standard mixture comprising the 20 most abundant amino acids in Murchison in relative abundances that match those in Martins and Sephton (Martins and Sephton, 2009), as well by measuring alanine in the methods development sample in two analytical periods. We used the Alfa Aesar alanine standard in the standard mixture and compared it to measurements of pure Alfa Aesar alanine (*i.e.*, not in a mixture) to ensure that the methodology used to measure mixtures would not fractionate alanine. The standard mixture was handled in a manner similar to that of the meteorite samples including being transferred in a water methanol mixture and dried down prior to derivatization. Relative to the Alfa Aesar standard, the standard mixture had a $\delta^{13}C$ of 2 ‰, which was within error of its measurement. Furthermore, excepting one methods development sample that was contaminated during derivatization (November 2017), the averaged C-2 + C-3 $\delta^{13}C$ (*i.e.*, that of the 140.032 Da fragment) and the averaged C-1 + C-2 $\delta^{13}C$ (184.021 Da fragment) values for two aliquots of methods development measured via GC-C-IRMS at GSFC and on the Orbitrap at Caltech in January and March of 2019 were within error of one another (Table 4.1, Appendix B.1). The summer 2018 analysis of the Strecker alanine is also within one standard error of the spring and winter 2018 C-2 + C-3 averaged $\delta^{13}C$ value and C-1 + C-2 averaged $\delta^{13}C$ value.

4.2.3. Data processing

4.2.3.1. Calculations of Site-Specific Isotope Ratios

Several arithmetic operations were required to calculate the site-specific $\delta^{13}C$ values for C-1, C-2, and C-3 in alanine. First, all accepted analyses for each fragment were combined (see Section 2.2.4.1 for criteria of accepted scans and Table S4.3 in Appendix C for analyses used) and the ^{13}R of each fragment was calculated as a weighted average of all counts (monoisotopic and singly ^{13}C-substituted) for the fragment (Eqn. 4.1)

$$^{13}R_{frag} = \sum_{i=1}^{n} \left[{}^{13}R_{scan} \times \frac{{}^{12}C_{cts,scan} + {}^{13}C_{cts,scan}}{\Sigma {}^{12}C_{cts,scan} + \Sigma {}^{13}C_{cts,scan}} \right]_i \qquad \textit{(Eqn. 4.1)}$$

where $^{13}R_{frag}$ is the ^{13}R value used for a fragment measurement, $^{13}R_{scan}$ is the ^{13}C counts/^{12}C counts for a single scan as defined in Eiler *et al* (2017) with a C_N (the charge conversion constant) of 3.6, $^{x}C_{cts,scan}$ is the and number of counts of isotope, *x,* for a single scan. The measurement is summed over all included scans.

This $^{13}R_{frag}$ value was then converted into $\delta^{13}C_{VPDB}$. To this end, the measured $^{13}R_{frag}$ of each fragment was standardized to Alfa Aesar by dividing the sample's $^{13}R_{frag}$ by that of Alfa Aesar alanine for the same fragment ion measured under the same analytical conditions (*i.e.*, same elution times into reservoir, same AGC conditions, similar TICxIT ranges) and temporally close *(i.e*., same measurement period). The standardized $^{13}R_{frag}$ for the alanine carbon site(s) in a fragment were then corrected for the dilution by carbon(s) from derivatizing reagents present in the fragment of interest (as these carbons have the same source in the sample and standard; see Table 4.1). This correction is found in Equation 4.2:

$$^{13}R_{corr} = (((^{13}R_{sa,meas}/^{13}R_{AA,meas} - 1) \times {}^{nC_{frag}}\!/_{nC_{ala}}) + 1) * {}^{13}R_{AA,fVPDB} \qquad \textit{(Eqn. 4.2)}$$

where $^{13}R_{corr}$ is the standardized ^{13}R value for a given fragment, $^{13}R_{sa,meas}$ is the $^{13}R_{frag}$ value for a sample directly measured for a fragment on the Orbitrap, $^{13}R_{AA,meas}$ is the $^{13}R_{frag}$ value for the Alfa Aesar standard directly measured for the same fragment on the Orbitrap, nC_{frag} is the total numbers of carbons in a fragment (*e.g.* 4 carbons for the 140.035 fragment), nC_{ala} is the numbers of carbons from alanine in that fragment (*e.g.* 2 carbon for the 140.035 fragment), and $^{13}R_{AA,fVPDB}$ Alfa Aesar's ^{13}R value for the alanine carbons in the fragment on the VPDB scale (for more information on these values see Eiler *et al.*, (Eiler et al., 2017). Finally, the standardized and corrected ^{13}R values were transcribed into $\delta^{13}C_{VPDB}$ values (Table 4.1). The corrected values assume that the derivative carbons between samples and standards have the identical ^{13}R values at each

site between sample and standard (*i.e.*, such that ratios may be treated as conservatively mixed properties) and that they have the same $\delta^{13}C$ values as the Alfa Aesar standards. We examined this assumption and found that variations less than ~50 ‰ would result in errors below the analytical uncertainty. Using corrected ^{13}R values for each fragment ion, we found the $\delta^{13}C_{VPDB}$ value (See Footnote 4.1 for formula and description).

Once each fragment was assigned a $\delta^{13}C_{VPDB}$ value, we calculated the site-specific $\delta^{13}C_{VPDB}$ of each of the three alanine sites. Our measurements of the analytical sample extracts provided three independent constraints on the site-specific $\delta^{13}C$ values of alanine: the molecule-average isotope ratio measured by compound-specific GC-C-IRMS and the two ratios measured by the Orbitrap (for the 140.032 and 184.032 Da fragment ions). The assumption that derivative carbons are the same in the sample as the alanine standard provided a fourth constraint. Each constraint is associated with its own uncertainty and weighted effect on the $\delta^{13}C$ of each alanine carbon site.

For the site-specific isotope calculation, the GC-C-MS measurement of molecular-average $\delta^{13}C$, and the Orbitrap measurements of the averaged C-1 + C-2 and the averaged C-2 + C-3 $\delta^{13}C$ were converted to fractional abundances ($^{13}F_{avg}$, $^{13}F_{C1+C2}$, $^{13}F_{C2+C3}$ respectively) using the relation $^{13}F = {}^{13}R/(1 + {}^{13}R)$. The ^{13}F values were then used to solve the following set of mass balance expressions (Eqn. 4.3a-4.3c):

$$^{13}F_{C\text{-}1} = 3 \times {}^{13}F_{molec\ avg} - 2 \times {}^{13}F_{C\text{-}2+C\text{-}3} \qquad \textit{(Eqn. 4.3a)}$$

$$^{13}F_{C\text{-}2} = 2 \times {}^{13}F_{C\text{-}1+C\text{-}2} - {}^{13}F_{C\text{-}1} \qquad \textit{(Eqn. 4.3b)}$$

$$^{13}F_{C\text{-}3} = 2 \times {}^{13}F_{C\text{-}2+C\text{-}3} - {}^{13}F_{C\text{-}2} \qquad \textit{(Eqn. 4.3c)}$$

Once fractional abundances of ^{13}C in each site were calculated, they were converted $\delta^{13}C$ values. Error analysis is discussed in Appendix B.4.

4.3. Results

The 267 ng sample of alanine recovered from an acid-hydrolyzed hot water extract of the Murchison sample studied here comprises 0.665 ppm by weight of the bulk meteorite, is a nearly racemic mixture of D and L enantiomers, and has molecular-average $\delta^{13}C_{VPDB}$ values of 25 ± 3 ‰ and 26 ± 3 ‰ for the D and L enantiomers, respectively, which is consistent with prior measurements of alanine from samples of Murchison (Engel et al., 1990; Martins and Sephton, 2009; Elsila et al., 2012). Acid hydrolysis increases yield in our samples from 2.37 ± 0.23 and 2.30 ± 0.16 nmol/g (water-extractable, or 'free' alanine) to 5.30 ± 0.88 and 5.98 ± 1.03 nmol/g (total alanine) for D- and L-alanine respectively (Friedrich et al., 2018). Past studies have demonstrated that 'free' and total alanine are indistinguishable in $\delta^{13}C$ ((Burton et al., 2013); similar results are found for other water-soluble organics as with (Aponte et al., 2014)). Procedural blanks typically yielded alanine abundances that were less than 1 % of the recovered meteoritic material (see Appendix B.2 and see Appendix C Figure S4.3 and Tables S4.1 and S4.2). Although the enantiomeric proportions of amino acids cannot conclusively establish their biogenicity, the weight of the preceding observations lead us to conclude our sample contains no detectable terrestrial contamination and closely approaches the properties of indigenous alanine found in Murchison. The site-specific $\delta^{13}C$ values for alanine are -29 ± 10 ‰, 142 ± 20 ‰, and -36 ± 20 ‰ for the C-1, C-2, and C-3 sites, respectively (Table 4.2, see Figure 4.1a for carbon site identities). Errors for each site-specific value are highly correlated due to the more precisely known molecular-average value (25.5 ± 3 ‰) and even more precisely known average of the C-2 and C-3 sites (52.6 ± 1.5 ‰) one or both of which are used to calculate the site-specific value; see the Appendix B.1 for details. The carbon isotope structure we observe for Murchison alanine, particularly the marked ^{13}C enrichment of the C-2 site, provides new constraints on the mechanism, precursors, and setting of its synthesis.

	Carbon Site(s)	$\delta^{13}C_{VPDB}$ (‰)	*st err (‰)*	$\delta^{13}C_{VPDB}$ (‰)	*st err (‰)*
Direct Measurement	L-Alanine Molecular Avg.	26	*3*	*-32.1*	*0.1*
	D- Alanine Molecular Avg.	25	*3*		
	Average of C-1 and C-2	56.4	*8.9*	-30.5	*1.8*
	Average of C-2 and C-3	52.6	*1.5*	-41.4	*9.2*
Site Specific Calculations	C-1	-28.7	*9.5*	-37.7	*3.6*
	C-2	141.5	*20.1*	-45.1	*13.6*
	C-3	-36.3	*20.4*	-15.9	*13.9*

Table 4.2: Fragment and site-specific $\delta^{13}C_{VPDB}$ values for hydrolyzed alanine from a Murchison meteorite hot water extract and the Strecker standard. The full molecular-average direct measurement $\delta^{13}C$ value was measured via GC-C-IRMS, and the fragments' $\delta^{13}C$ values were measured on the Q-Exactive GC Orbitrap mass analyzer. The site-specific $\delta^{13}C$ values were calculated using the average of the D- and L-alanine molecular averages and the fragment $\delta^{13}C$ values.

4.4. Discussion

4.4.1. Alanine Formation Mechanism

Isotopic measurements of compounds in the ISM, including the local ISM (*i.e.*, within 800 parsecs of the Sun (Lallement et al., 2018), which is at 7.94 ± 0.42 kiloparsecs from the galactic center (Eisenhauer et al., 2003)), exhibit large gradients in the $^{13}C/^{12}C$ of gas-phase carbon pools depending, in part, on their location in a molecular cloud. Furthermore, due to the high (hundreds of per mil) error associated with the measurements, the $^{13}C/^{12}C$ of most carbon pools (CO, CH_x, CN, CS) overlap. Consequently, for our analysis, we rely on the models of Charnley *et al* (2011), which combine recent theories of ISM carbon chemistry and related isotope fractionations. Future measurements from higher precision instruments will allow for more refined isotopic models of ISM chemistry and thus will test our hypothesized reaction network and/or lead to refinements of our interpretation. Charnley *et al*'s model and others

predict CO will be highly ^{13}C enriched due to its high prevalence and the 35 K lower zero-point energy of ^{13}CO relative to ^{12}CO, as shown in Equation 4.4.

$$^{12}CO + {}^{13}C^{+} \leftrightarrows {}^{13}CO + {}^{12}C^{+} + \Delta E\ (35K) \qquad \textit{(Eqn. 4.4)}$$

In the cold ISM, where temperatures are 10 to 40 K, this energetic difference drives ^{13}C into the CO pool and depletes the C^{+} and CH_x pools in ^{13}C. One possible exception is CN, which has been modeled to have a $\delta^{13}C$ value that is either similar to the ^{13}C-enriched CO or to the ^{13}C-depleted C^{+}-derived pools of carbon-bearing molecules (Langer et al., 1984; Langer and Penzias, 1990; Milam et al., 2005). We note that, the only measurements of CN are: 1) KCN from the Murchison meteorite, which has a $\delta^{13}C$ value of 5 ± 3 ‰ (Pizzarello, 2014) — ^{13}C-depleted relative to the CO-bearing molecule formaldehyde; and 2) cometary HCN that was measured to have a $H^{12}CN/ H^{13}CN$ ratio of 88 ± 18 (*i.e.*, a $\delta^{13}C$ of $\sim 16^{+262}_{-172}$ ‰), which is within error of the solar system value of 89 (Cordiner et al., 2019). In both cases, the CN reservoir in the solar system does not bear enrichments predicted in Milam *et al* (2005), supporting the hypothesis that it is isotopically light — like the C^{+} and CH_x pools. However, we note the possibility that the CN measured was not made in the ISM and that the error associated with the cometary HCN could place it in either the isotopically enriched or light pools.

When describing sources of precursor compounds in the ISM in our hypothesized reaction network, we consider two main pools: ^{13}C-enriched (CO and possibly CN) and ^{13}C-poor (CH_x and possibly CN). The first pool includes carbonyl carbons such as those in aldehydes and the second includes reduced carbon such as hydrocarbons and aliphatic carbon chains. The 10-40K temperatures in the ISM enhance the ^{13}C fractionation caused by the exchange of CO and C^{+} in Equation 4.4, and the ^{13}C-enriched and ^{13}C-poor pools inherit the CO and C^{+} carbon isotopic compositions, respectively. We consider a $\delta^{13}C_{VPDB}$ value that is above 50 ‰—the predicted $\delta^{13}C_{VPDB}$ reservoir for planets of 0 ‰ plus a potential 50 ‰ ^{13}C enrichment from isotope effects associated with synthesis (Lyons et al., 2018)—to likely include carbon derived from material that is either sourced from CO and/or CN in the ISM or in a cold nebular environment, which experience

similarly cold temperatures. Our finding of a δ^{13}CVPDB value that exceeds 100 ‰ at the C-2 site in alanine provides a strong indication that this site is derived from a precursor that was itself synthesized in the ISM or a cold nebular environment from CO and its products and/or CN.

Our finding of a relatively low δ^{13}C value of the C-1 site, however, is inconsistent with current experimental constraints on amino acid synthesis via the irradiation of methanol ices and ammonia in the ISM. An experimental irradiation of isotopically labelled methanol ices at 40 K (Elsila et al., 2007) produced adequate amounts of serine for site-specific analyses and found that both the C-1 and C-2 sites are inherited from HCN, implying that this mechanism should not lead to marked differences between the carbon isotopic compositions of the C-1 and C-2 sites. Assuming that alanine follows a similar formation pathway to serine, we conclude that alanine from Murchison inherited the C-2 carbon from a precursor that was itself formed from the CO, HCO^+, and/or CN pools in the ISM and that its ^{13}C depletion in the C-1 carbon was contributed from another, lower δ^{13}C precursor through reactions that likely occurred either in the early solar nebula or in Murchison's parent body. However, we note that further experiments should explore the potential to form alanine through alternate pathways in the ISM, such as by gas-phase reactions and gas-grain reactions, and that experiments should sample carbon sources in a manner that reflects the diversity found in in interstellar ices. Specifically, we note that ice-grain experiments presented in Elsila *et al.* (2007) produced glycine that formed by multiple formation pathways, including a minor pathway in which C-1 was derived from HCN and C-2 from methanol — a pattern of transfer from substrate to amino acid sites that could be mistaken for the Strecker synthesis. Further experiments of this kind are required to determine the factors controlling the relative rates of the pathways of amino acid synthesis that can occur in ice grain chemistry, especially those that lead to synthesis of aliphatic α-H and α-CH_3 amino acids such as alanine or isovaline.

Our findings are also inconsistent with the hypothesis that this nebular or parent body chemistry followed a predominantly FTT mechanism. FTT-synthesized alanine inherits all its carbons from the source CO, with each added carbon being only subtly lower in

$\delta^{13}C$ than the CO pool due to a KIE of approximately 0-10 ‰ (Mccollom and Seewald, 2006; Taran et al., 2007). Although this reaction mechanism is incapable of directly generating the ~170 ‰ contrast we observe between the $\delta^{13}C$ values of the C-2 site compared to the C-1 and C-3 sites, it is possible that alanine could form by an FTT-like process if the carbon in the C-2 site were derived from a secondary product of small molecules other than CO. In some conditions, FTT chemistry can create CO_2 and CH_4 that differ from one another by up to ~50 ‰ (Taran et al., 2007)— a contrast approaching that required by our data. In this case, alanine synthesis by FTT would require that the C-2 carbon is a secondary product of the ^{13}C-enriched CO_2 produced by FTT synthesis, whereas the C-3 and – most problematically – C-1 carbon are secondary products of low ^{13}C FTT-derived CH_4. We can think of no plausible chemical reaction sequences in which this would happen.

For these reasons, given the current understanding of isotope effects and mechanisms of interstellar and nebular chemistry, we conclude that alanine in Murchison likely formed via Strecker synthesis or reductive amination, that it was synthesized in the solar nebula, possibly in the meteorite's parent body, and that it had likely at least one reactant that itself was derived from CO or CN in the ISM or outer solar system. Drawing on past models and measurements, Elsila *et al.*(2012) and Aponte *et al.* (2017) argued that meteoritic alanine formed by Strecker synthesis from ISM-derived acetaldehyde with a ^{13}C-enriched carbonyl carbon inherited from CO and ^{13}C-depleted methyl carbon inherited from the CH_x pool, in addition to NH_3, and ^{13}C-depleted HCN such as that measured on Murchison. These reactants would lead to alanine with a high $\delta^{13}C$ value at the C-2 site and lower $\delta^{13}C$ at the C-1 and C-3 sites (Elsila et al., 2012) (Figure 4.3). The results presented here are consistent with this argument. If instead alanine formed by reductive amination, one of its precursors would have been pyruvic acid. If the precursor were pyruvic acid formed solely by CO grain chemistry (Elsila et al., 2012), then all of its carbon sites and those on the subsequently produced alanine will be equally ^{13}C-enriched, in disagreement with our findings. If, however, pyruvic acid formed via a ketene or aldehyde reacting with HCN and water in the ISM (Cooper et al., 2011) or cyanohydrin in the parent body, it could result in a carbon isotope structure broadly resembling that

produced by Strecker synthesis (See Appendix B.5). We consider these two mechanisms equally plausible based on the constraints of our alanine's C isotope structure. Non-α-amino acids (*e.g.*, β-, γ-) cannot be produced via the Strecker pathway and require other mechanisms of production.

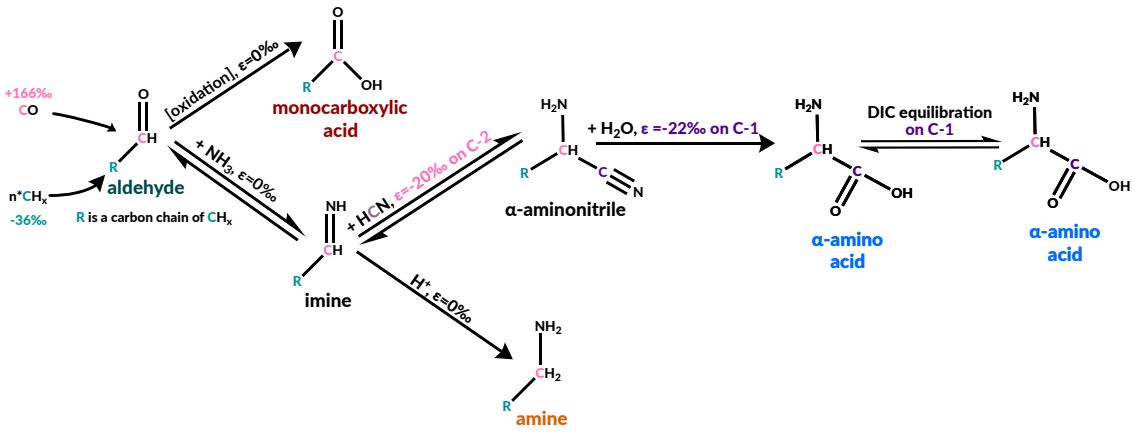


Figure 4.3: Proposed mechanisms for syntheses of organic compounds related to alanine (with R of CH_3) on the Murchison parent body, with associated carbon isotope effects. In this scheme, CO and CH_x are derived from the ISM. Reaction steps between the aldehyde and imine and between the imine and aminonitrile are reversible (Van Trump, 1975). Isotopic values for the initial CO and *n*CH_x are back-calculated using our measured alanine value and the isotopic effects shown.

Mechanism Step ***(isotope effect type)***	**$\varepsilon_{C\text{-}1}$ (‰)**	**$\varepsilon_{C\text{-}2 + C\text{-}3}$ (‰)**	**$\varepsilon_{C\text{-}2}$ (‰)**
Aminonitrile formation *(EIE)*	N/A	-10	-20
Aminonitrile hydrolysis to amide *(KIE)*	-8.5	0.7	1.4
Amide hydrolysis to amino acid *(KIE)*	-15	0	0

Table 4.3: Site-specific isotope effects measured for the Strecker synthesis of alanine. The $\varepsilon_{C\text{-}1}$ likely has a non-zero value as exists for the equilibrium between CN and HCN, but in the experiments all CN was converted into 2-propionitrile so no isotope effect could be measured for the aminonitrile formation.

4.4.2. Precursor Reservoir Values

To help us predict the isotopic contents and structures for the precursors to alanine in Murchison, we synthesized alanine via Strecker synthesis and measured its site-specific carbon isotope effects relative to the starting acetaldehyde and NaCN (see Table 4.3 and

Appendix B.6). Experiments indicate that production of the α-aminopropanenitrile intermediate has a $\delta^{13}C$ that is 10 ‰ below its acetaldehyde precursor at moderate (~60-70 %) yields. Because the C-3 carbon does not gain or lose covalent bonds in the Strecker reaction, and thus will not experience large isotope effects from the synthesis, the 10 ‰ shift in the average C-2 and C-3 $\delta^{13}C$ value suggests a -20 ‰ isotope effect on the C-2 carbon (see Table 4.3, Figure 4.3), which is consistent with the KIE on a carbonyl carbon from the addition of CN (Lynn and Yankwich, 1961). If we assume a large initial acetaldehyde reservoir such that its isotopic value is effectively constant during alanine production, and account for the reactant aldehyde's fractionation by adding 10 ‰ to the C-2 and C-3 site's average $\delta^{13}C$, we predict that the initial acetaldehyde reservoir parental to Murchison alanine had a molecular-average $\delta^{13}C$ of 62.6 ± 1.5 ‰. This value is within error of 64 ± 1 ‰, a molecular-average value for acetaldehyde recently measured in Murchison (Figure 4.4, Electronic Annex 2, and Aponte et al. (2019). This agreement is consistent with our suggestion that alanine had an acetaldehyde precursor and thus reinforces the possibility that alanine was synthesized by Strecker reaction rather than reductive amination; it also suggests that the initial aldehyde pool was not fractionated during the synthesis of alanine and was therefore either large in amount relative to the alanine produced or that aldehyde was continuously produced (*e.g.*, from the hydrolysis of other acetaldehyde-derived compounds) as it was consumed in alanine syntheses. We note that other measurements of the molecular-average $\delta^{13}C$ of acetaldehyde have found values of 25-27 ‰ (Simkus et al., 2019) possibly due to sample heterogeneity or fractionation of volatile molecules during laboratory extraction (Aponte et al., 2019). Future site-specific isotope ratio studies of Strecker synthesis reactants (*e.g.*, aldehyde, CN) and products from the same sample could resolve the reason for this discrepancy and further test our hypothesis.

The Strecker experiments also indicate that the acid hydrolysis of α-aminonitrile to an amino acid has carbon isotopic fractionation on the C-1 site of up to -50 ‰ for a 13 % conversion of cyanide to alanine and a mean value of -22 ‰ for a 20 % to 55 % conversion of α-aminonitrile to alanine (Table 4.3, Figure 4.4). Therefore, if alanine in

Murchison formed by Strecker synthesis with moderate yield in its second step (20 - 50 %), it should have inherited its C-1 carbon from reactant HCN that had $\delta^{13}C$

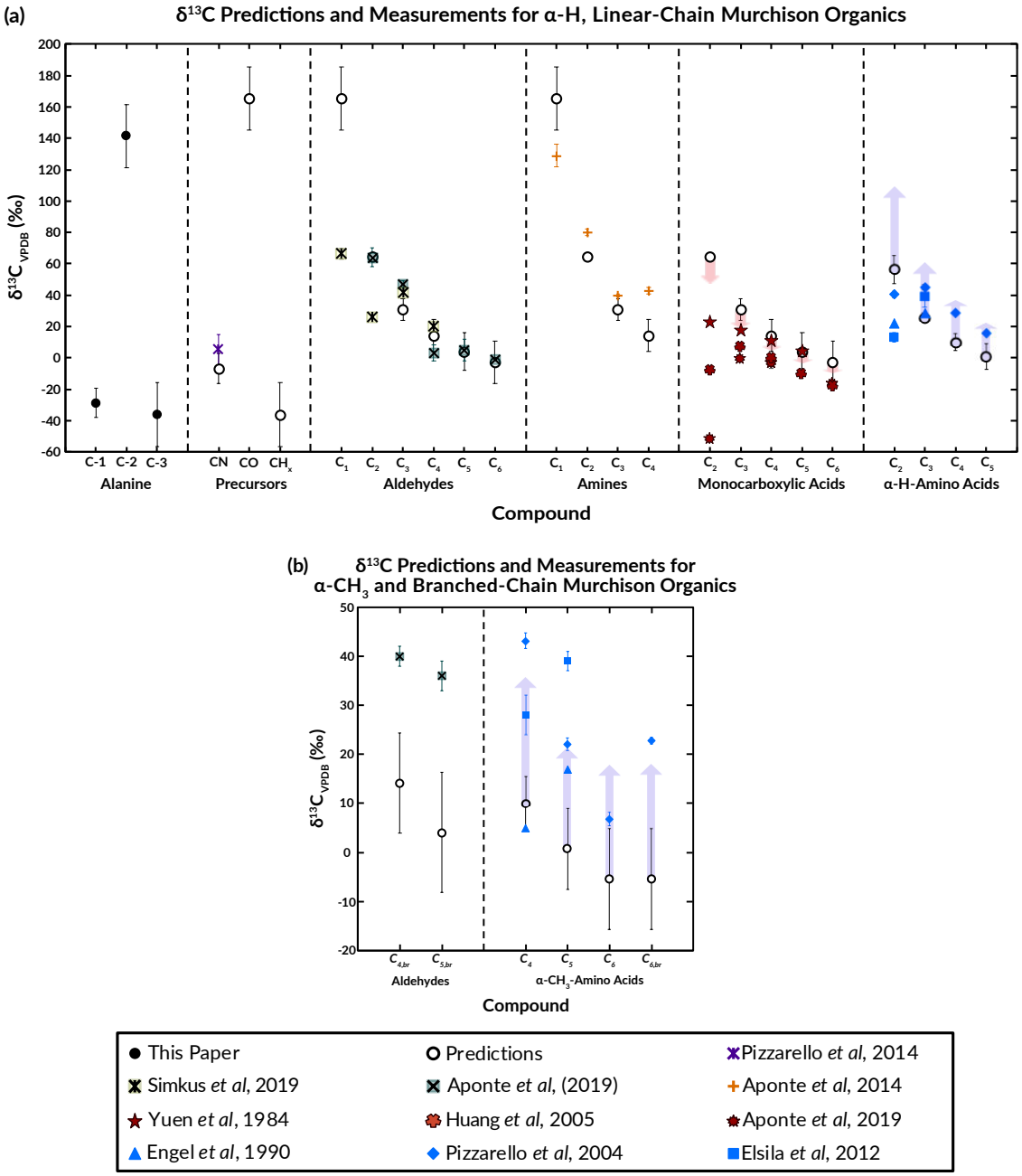


Figure 4.4: (A) Comparison of $\delta^{13}C$ measurements from this study, model predictions, and literature values for each carbon site in alanine and for molecular-averages of precursors, product aldehydes, amines, monocarboxylic acids, and α-H-amino acids with linear carbon sidechains. (B) Comparison of $\delta^{13}C$ measurements from this study, model predictions, and literature values for α-CH_3-amino acids with linear carbon sidechains (denoted with no subscript) and for aldehydes and α-CH_3-amino acids with branched carbon sidechains (denoted with *br* subscript). We only include compounds with isotopic values recorded in the literature and with alpha chiral sites as well as glycine (and possible compounds made from its proposed precursor, formaldehyde) due to its biological importance. The pink arrows display the range of values predicted based on the range of KIEs for aldehyde oxidation on the reactant CO site (the terminal COOH site in monocarboxylic acids). The purple arrows highlight the expected range of values for Strecker synthesis followed by carbon isotope exchange between DIC and the C-1 sites of α-amino acids. The subscripts denote the total

number of carbons in the molecule. The error bars for carboxylic acids are smaller than symbols and are not included in the data for the Engel *et al.* (1990) measurements as they are not provided in the 1990 paper.

of -7 ± 10 ‰ (For error analysis, see Appendix B.4). This value is within error of the previously reported 5 ± 3 ‰ $\delta^{13}C$ of HCN in Murchison (Pizzarello, 2014). It is noteworthy that the HCN extracted from Murchison has a lower $\delta^{13}C$ than do formaldehyde and acetaldehyde from that sample: this difference is consistent with the idea that the HCN reservoir available on Murchison was ^{13}C-depleted relative to the reservoir that created the alpha site on α-amino acids. Other combinations of substrate $\delta^{13}C$ values and reaction yields are also possible, but the agreement of this scenario with independent constraints for acetaldehyde and HCN support its plausibility.

4.4.3. Predictions for the $\delta^{13}C$ of other small, water soluble organics on Murchison

The preceding findings enable us to create a testable hypothesis in the form of a chemical network connecting the synthesis of alanine in Murchison and the formation of other organic compounds, including C_1 to C_6 aldehydes, amines, carboxylic acids, and other α-amino acids in the Murchison parent body **(**See Appendix B.6, Electronic Annex 2, and references (Aponte et al., 2017; Simkus et al., 2019)). Our model above predicts an acetaldehyde precursor of alanine having carbonyl and methyl groups with $\delta^{13}C$ values of 162 ± 20 ‰ and -36 ± 20 ‰, respectively (noting that the average of these two sites is predicted with a much narrower error of ±1.5 ‰). The model we present presumes that alanine and the other soluble organics we consider were synthesized from a pool of precursors (H_2O, aldehydes, HCN, NH_3) that was not significantly depleted by their growth (excepting CN, which we assume underwent 10's of % consumption by the Strecker chemistry, as in our experiments), that all reactions occurred at the same temperature, and that none of the studied compounds are residual to losses by a fractionating side-reaction. These assumptions are clearly simplifications, but generally similar models that relax these constraints (*i.e.*, allowing for variable temperature, reaction progress or side reactions) do not strongly impact our predictions (Appendix B.6). If formaldehyde and acetaldehyde have the same carbonyl source as expected for

ISM-derived aldehyde, then the $\delta^{13}C$ of formaldehyde should be 162 ± 20 ‰. Likewise, larger aldehyde precursors would be predicted to have molecular $\delta^{13}C$ values equal to the weighted average of their one ^{13}C-rich carbonyl carbon and some additional number of ^{13}C-poor R-group carbons similar in ^{13}C isotopic composition to acetaldehyde's methyl group. These predictions agree with some molecular-average measurements of individual linear aldehydes having 2-5 carbon atoms from Murchison (Figure 4.4a), but they over-predict the $\delta^{13}C$ measured for formaldehyde (Simkus et al., 2019) and under-predict the measured differences between branched and linear compounds (Figure 4.4b, *refs 23* and *24*). Data from (Simkus et al., 2019) disagree with our predicted acetaldehyde value but agree with our predictions for C_4 and C_5 linear aldehydes and still display modest ^{13}C-enrichments for C_2 and C_3 linear aldehydes.

We hypothesize that other molecules with amine functional groups in Murchison were formed by reductive amination of the same aldehyde precursors that formed alanine through Strecker synthesis. Past measurements of reductive amination have demonstrated negligible KIEs of less than 1 ‰ (Billault et al., 2007), so the carbon backbones of other organic amines should resemble the parent aldehyde in our proposed mechanism. This hypothesis leads to $\delta^{13}C$ predictions of 162 ± 20 ‰, 62.6 ± 1.5 ‰, 30 ± 7 ‰, and 13 ± 10 ‰ for methyl-, ethyl-, propyl-, and butylamine, respectively, which resemble previous measurements from Murchison (Figure 4.4 and Aponte et al., 2016). However, the predictions cannot account for the lack of measured difference in $\delta^{13}C$ between the C_3 and C_4 amines.

Similarly, we hypothesize that aldehyde precursors in Murchison can be oxidized into monocarboxylic acids via hydration and hydrogen abstraction at the carbonyl carbon. In the presence of water and metal oxides, aldehydes can be oxidized (Rajesh and Ozkan, 1993) to form carboxylic acids. Metal oxide catalysts are present in Murchison and other CM2 meteorites (Bunch and Olsen, 1975; Hanowski and Brearley, 2000), supporting the plausibility of this scenario. Accounting for previously measured KIEs associated with addition reactions to aldehydes (a 0 to -19 ‰ KIE for carbonyl carbons; (Yamataka et al., 1997; Yamataka et al., 2001)) and the likely upper limit of a ~-30 ‰ KIE for the

oxidation of a carbonyl carbon, the $\delta^{13}C$ values of the C_1-C_5 monocarboxylic acids can be calculated as a mixture of a ^{13}C-enriched carbonyl carbon and ^{13}C-depleted methyl carbons. The final predicted monocarboxylic acid molecular-average $\delta^{13}C$ values vary little between the 0 ‰ and -30 ‰ isotope effects on the carbonyl carbon, so we will consider the -30 ‰ predictions that closely agree with previous measurements for the C_3-C_5 species from Yuen *et al* (1984) and with the trends presented in more recent studies by (Huang et al., 2005) and (Aponte et al., 2019b) (Figure 4a). The overprediction of acetic acid's $\delta^{13}C$ relative to data from all studies (Yuen et al., 1984; Huang et al., 2005; Aponte et al., 2019b), in conjunction with the larger range of measured $\delta^{13}C$ values for acetic acid (~75 ‰) versus those for other monocarboxylic acids (0-20 ‰) (Figure 4.4a) supports the argument that the acetic acid measured on Murchison is a mixture of two or more sources (Huang et al., 2005). Furthermore, the differences in past monocarboxylic acid $\delta^{13}C$ measurements from both our predictions and from each other (Yuen et al., 1984; Huang et al., 2005; Aponte et al., 2019b) could reflect spatial $\delta^{13}C$ heterogeneity of these components that our model does not capture as it bases its predictions on $\delta^{13}C$ values from one compound from one meteorite sample (see Appendix B.6). Studies of site-specific isotope ratios of monocarboxylic acids and aldehydes could provide a means of further testing and refining our understanding of the relationships amongst these compounds in Murchison as they play a critical role in the network of reactions in which amino acid synthesis occurs. Despite these complexities in the prior carbon isotope data, the relatively straightforward, unified chemical reaction network we propose provides a coherent and accurate explanation for the measured $\delta^{13}C$ values of alanine, reactant HCN, and most aldehydes, amines and monocarboxylic acids in Murchison, based only on two assumed $\delta^{13}C$ values (that for CO and CH_x precursors in the ISM; see Figure 3 and Appendix B.6). The most noteworthy disagreements between our model and prior data for Murchison extracts are for formaldehyde and glycine. These are among the most volatile and easily contaminated compounds that we considered, and we suggest their high variability among prior studies and lower-than-predicted average values may reflect particularly poor preservation.

4.4.4. Model Shortcomings

Four complicating factors prevent us from extending our model to all amino acids in Murchison that have $\delta^{13}C$ measurements: 1) prior studies yield ranges of up to 30 ‰ in $\delta^{13}C$ for individual amino acids (Engel et al., 1990; Pizzarello et al., 2004; Elsila et al., 2012), possibly reflecting spatial variation in precursors, reaction progress, and/or terrestrial contamination between sub-samples; 2) amino acids as a whole are structurally diverse and draw on a variety of precursors that may not have been uniform in their ^{13}C contents; 3) amino acids can be subject to side reactions not considered in the simple reaction network outlined above; and 4) Strecker synthesis can only produce α-amino acids, so all others (*e.g.*, β, γ, δ) require other synthetic routes. Nevertheless, it is straightforward to extend our hypothesis to an approximate prediction of the molecular-average $\delta^{13}C$ values of the α-amino acids. If we assume all the C-1 and C-2 sites in α-amino acids have $\delta^{13}C$ values that are identical to those observed in alanine and that all other carbon atoms have $\delta^{13}C$ values equal to that of the C-3 site in alanine (as would occur if all form by Strecker synthesis from a closely related pool of aldehyde precursors and HCN as outlined above and in Figure 4.3), then we can calculate the molecular average $\delta^{13}C$ values of the other individual α-amino acids. The results are similar to most prior measurements of the C_2-C_5 α-H-amino acids, except a subset of glycine measurements; there are several possible explanations for this discrepancy, but we note that glycine is unusual in being achiral and is suspected to have been synthesized by multiple mechanisms (Figure 4.4) (Engel et al., 1990; Pizzarello et al., 2004; Elsila et al., 2012).

The model presented here consistently under-predicts $\delta^{13}C$ values of both branched aldehydes and α-CH_3-amino acids (Engel et al., 1990; Pizzarello et al., 2004; Elsila et al., 2012) (Figure 4.4b). One possible cause of higher measured $\delta^{13}C$ values in the amino acids, particularly the α-CH_3 amino acids, in Murchison samples is isotopic exchange between carboxyl sites and dissolved inorganic carbon (DIC) present during parent-body

aqueous alteration. Theoretical calculations demonstrate that this exchange can occur for amino acids (Rustad, 2009; Pietrucci et al., 2018), has lower energy barriers for α-CH_3 species than for α-H species (Pietrucci et al., 2018), and high-$\delta^{13}C$ carbonate minerals in Murchison attest to the presence of a ^{13}C-rich DIC pool (est. with the highest measured value of +80 ‰ (Sephton, 2002) to present the full possible range of $\delta^{13}C$ values, see SI) during aqueous alteration. The measured molecular-average $\delta^{13}C$ values of the α-CH_3 amino acids are similar to those predicted by our model of Strecker synthesis if it is followed by equilibration of carboxyl sites with the DIC pool (purple arrows in Figures 4.3 and 4.4). Our predictions represent a maximum $\delta^{13}C_{VPDB}$ change in the amino acids (top of the purple arrows in Figure 4.4). The isotopic composition of DIC varies on Murchison samples; therefore, a lower amount of exchange and/or exchange with a less enriched $\delta^{13}C_{VPDB}$ DIC pool would result in $\delta^{13}C_{VPDB}$ values that span the length of the purple arrows in Figure 4.4. It may be that partial exchange and/or lower $\delta^{13}C$ DIC pools explain why some of the α-H amino acids have $\delta^{13}C_{VPDB}$ values lower than predicted by our model. This mechanism would not function on the monocarboxylic acids without moieties on the C-2 site with a lone pair (*e.g.*, NH_2 or OH) as the C-2 site could not switch between sp^3 and sp^2 as easily. This explanation for the $\delta^{13}C$ values of the amino acids is not unique; however, it captures the full range of observations with a single plausible addition to an already parsimonious model. Branched-aldehydes and branched-sidechain amino acids require different explanations as both would require a less favorable exchange of C in saturated hydrocarbon chains. The differences in isotopic content between linear and branched compounds is another attractive target for further studies of site-specific isotopic contents of meteoritic organics.

4.5. Conclusions

The arguments and data presented here suggest that Strecker synthesis is likely the origin of alanine in the Murchison meteorite and that aldehydes formed from CO and CH_x in the ISM are essential precursors to a wide range of the prebiotic organic compounds observed in Murchison. These organic compounds include amino acids, amines, and carboxylic acids that formed when the ISM-sourced aldehydes reacted with HCN, NH_3,

and water. Following the production of amino acids, isotopic exchange between the carboxyl group and ^{13}C-rich DIC pool might have occurred in at least some α-amino acids, approaching equilibrium for the relatively exchangeable α-CH_3 amino acids. The success of a simple reaction model (Figure 4.3) in explaining most of the $\delta^{13}C$ values previously measured for these diverse compounds supports the idea that the various chemical reactions called on occurred concurrently, in a single environment, and drawing on a common pool of precursors, some of which likely originated in the ISM(de Marcellus et al., 2015). Aqueous alteration in the Murchison parent body is a plausible setting where this could have transpired.

In the future, more precise measurements of $\delta^{13}C$ values of organic compounds in the ISM will help test and refine this model by constraining initial isotope values in ISM as well as locations where these enrichments occur for different carbon pools.

Acknowledgments: We would like to acknowledge Dr. Max Klatte and Dr. Carl Blumenfeld for aiding with the Strecker synthesis of alanine. **Funding:** This project was supported by NASA through LARS grant number NNX17AE52G and through part of the Planetary Science Division Internal Scientist Funding Program through the Fundamental Laboratory Research (FLaRe) work package of the Planetary Science Division Internal Scientist Funding Program, by DOE grant DE-SC0016561, by Thermo Fisher, and by the Simons Foundations; **Author contributions:** LC, BD, ALS, and JME designed methods for site-specific carbon isotope measurements. JEE, JPD, and JA provided Murchison sample. JEE extracted amino acids and measured molecular-average isotope ratios of alanine. LC and JME created the Monte Carlo simulation to calculate isotope ratios. LC measured alanine on Murchison meteorite, processed data, and calculated site-specific isotope ratios. LC, JEE, JPD, JA, ALS, and JME contributed ideas to form the parent-body organic synthesis model.; **Competing interests:** Authors declare no competing interests.; and **Data and materials availability:** All data is available in the main text or the Supplementary Information

Chapter 5

AVERAGES TO INSIGHTS:

USING MOLECULAR-AVERAGE ISOTOPE VALUES TO MODEL ISOTOPIC STRUCTURES AND METEORITE ORGANIC SYNTHESIS

Abstract

Carbonaceous chondrites contain a variety of prebiotic compounds. A variety of mechanisms have been put forth as main synthetic pathways that create these compounds. These mechanisms range from ice-grain reactions in the interstellar medium (ISM) to aqueous chemistry in mild temperature and pH conditions in meteorite parent bodies. The ratio of heavy to light isotopes that a compound inherits depends on the chemistry that produces it. Compounds produced in the ISM are anticipated to have large enrichments in heavy isotopes that can be in the 100s of per mil for carbon and the 1000s for hydrogen. To date, studies have focused on molecular average isotope measurements and qualitative models. Here we create quantitative models for hydrogen and carbon isotope compositions and use these models to interpret likely chemistries that impacted each.

For hydrogen, we build a model that fits the measured molecular-average deuterium concentrations for compound classes—molecules that share functional groups such as monocarboxylic acids and amines—as linear combinations of hydrogen moieties, hydrogens that share the same chemical environment. We assume that hydrogens that share similar chemical environments will have similar isotopic ratios for two reasons: (1) chemical environments determine the exchangeability of a hydrogen site and (2) our model considers that compounds within each class are made from a similar process and therefore each of its moieties should be sourced from the same or similar precursors. We use this model first on compounds measured from the Murchison meteorite and then expand it to include other meteorites. In doing so, we find that methyl hydrogens are amongst the most deuterium enriched moiety by up to 3000 ‰ and hydrogens attached to α-carbons are the least. We also notice a trend in the deuterium enrichment with a given

compound class's solubility in water and with a meteorite sample's degree of aqueous alteration and terrestrial weathering. Combined, these fits put forward a scenario in which molecular inherit their hydrogen from compounds in the ISM and this primordial signal is washed out through exchange with deuterium on the parent body during aqueous alteration.

We model the carbon isotope composition of products from different reaction networks including Fischer Tropsch type synthesis, Formose chemistry, and the integrated aldehyde network (oxidation, reductive amination, and Strecker synthesis on a pool of aldehydes and ketones to create carboxylic acids, α-amino acids, and amines, respectively). The model with the highest percentage of accurate predictions and the lowest average residuals for carbon isotope ratio predictions of compounds from the Murchison meteorite was the integrated aldehyde network followed by formaldehyde addition to α-H and straight-chain compounds to create α-CH_3 and branched-chain compounds. This model predicts 61 % of compounds within two standard deviations of their analytical error and 72 % of compounds within 5 ‰ of that. The monocarboxylic and α-hydroxy acids performed worst in these predictions; however, when we model formaldehyde addition to their α-H and straight-chain compounds using these compounds' measured isotope ratios, the α-CH_3 and branched-chain compounds' predicted values agree with measurements for 50 % of the monocarboxylic acid measurements with a 5 ‰ average error amongst the remining compounds and 100 % of the α-hydroxy acid measurements. Formaldehyde addition has not been previously tested or noticed in chemical reaction networks, but the best fit network's ability to predict compounds that span over 100 ‰ in carbon isotope abundances makes it an attractive chemistry to explore.

5.1. Introduction

Carbonaceous chondrites (CCs)—a class of undifferentiated meteorites that have 2-5 wt % C (Glavin et al., 2018)—are amongst the most primitive samples of solar system material. Insights into their chemical evolution elucidate the organic chemistry in the early solar system and that which could occur in others. The organic inventory of the

CC's includes soluble and insoluble organic matter (SOM and IOM, respectively), includes over 80 amino acids and 100 polyaromatic hydrocarbons (PAHs), and is dominated by its IOM component, which comprises ~70% of the organic matter (OM) (Glavin et al., 2018).

Concentrations of OM vary both between meteorite classes—groups of meteorites thought to be from the same or similar parent bodies, as evinced by their isotopic and mineralogic compositions—and between petrologic types—categories that reflect a meteorite's degree of aqueous or thermal alteration from parent-body processes (Elsila et al., 2016; Glavin et al., 2018). Typically, CR, CM, and CI chondrites—the less thermally altered CC meteorite classes—have higher SOM concentration than the CH and CV chondrites—more thermally altered CC classes (Elsila et al., 2016). These differences could be attributed to different chemistries prior to parent body heating or varying degrees of degradation during parent body heating (Elsila et al., 2016). Meteorites with a more pristine petrologic type (closer to type 3.0) also have higher concentration of SOM as compared to more aqueously altered types (lower than 3.0) or thermally altered types (higher than 3.0) (Elsila et al., 2016).

Potential environments for synthesis of OM found in the carbonaceous chondrites include the interstellar medium (ISM), the pre solar nebula and early solar disk, and meteorite parent bodies. In the ISM, grain-surface chemistry is proposed to produce aldehydes, ketones, amino acids, amines, and hydroxy acids (Sandford et al., 2001; Elsila et al., 2007; Glavin et al., 2018). In the solar nebula and disk, it has been suggested that gas-grain reactions can enable Fischer-Tropsch type (FTT) synthesis; additionally, radiation-driven reactions, synthesis from meteorite impact ('shock synthesis'), and aqueous syntheses, including hydrothermal FTT, Formose, Strecker, reductive amination, and aldehyde oxidation, could produce the SOM we observe on meteorites (Cronin et al., 1995; Kerridge, 1999; Cooper et al., 2011; Elsila et al., 2012; Glavin et al., 2018).

These scenarios of organic synthesis have been proposed and evaluated based on concentrations of key compounds, stable isotope abundance ratios (D/H, $^{13}C/^{12}C$, *etc.*),

and relationships of these two variables to the degree and type of alteration and class of the meteorite sample from which the compound was extracted. Extraordinary enrichments in abundances of rare, heavy isotopes (D, ^{13}C, ^{15}N), beyond the range observed in most terrestrial materials (100s of ‰ for δ^{13}C and d15N and 1000s of ‰for δD)[5.1] are observed in some meteoritic OM, and generally have been interpreted as evidence for formation in the ISM or inheritance from precursors from the ISM, because at the low relevant temperatures (10's of K), zero-point energy effects of heavy isotope substitution in chemical bonds substantially increase the thermodynamic stabilities of substituted molecules (Pizzarello and Shock, 2010; Glavin et al., 2018). It is less clear whether this process is responsible for the subtle heavy isotope enrichments (10's of ‰ for δ^{13}C and d^{15}N; 100's of ‰ for δD) that characterize many meteoritic organic compounds.

Previous attempts to understand these more subtle heavy isotope enrichments of meteoritic organics have examined H, C, and N isotope differences between compound classes (*e.g.,* amino acids, MCAs, DCAs), molecular structural motifs (*e.g.,* straight- or branched-chain), molecular size (e.g., carbon chain length), and host meteorite sample types. These efforts are complicated by the fact that isotopic composition of a given organic compound can differ between portions of a given sample and between studies. Nevertheless, certain observations are consistent: the non-exchangeable hydrogens in most meteoritic OM is enriched in deuterium relative to terrestrial OM (Huang et al., 2007; Alexander et al., 2007; Glavin et al., 2018); most SOM—including α-amino acids, α-hydroxy acids, amines, sulfonic acids, dicarboxylic acids (DCAs), and aldehydes—are also enriched in ^{13}C and ^{15}N relative to terrestrial sources; and ^{13}C enrichments in SOM

[5.1] $^{h}\delta X$ is equal to the ratio of the heavy isotope, h, in atom X relative to the light isotope, l, in a sample relative to that ratio in a standard. It can be represented as: $^{h}\delta X = \frac{\left(\frac{^{h}X}{^{l}X}\right)_{sa}}{\left(\frac{^{h}X}{^{l}X}\right)_{sa}} - 1$ and is typically multiplied by 1000 to be reported in until of per mil (‰). We will use ^{h}R to represent the ratio of heavy to light isotopes ($\frac{^{h}X}{^{l}X}$) and ^{h}F to represent the fractional abundance of the heavy isotope, $\frac{^{h}X}{^{l}X + ^{h}X}$. We note that ^{h}F and ^{h}R are related such that $^{h}F = \frac{^{h}R}{1 + ^{h}R}$.

compounds generally decrease with the carbon chain length (Sephton, 2002; Glavin et al., 2018). These observations are suggestive evidence that many meteoritic compounds of prebiotic relevance formed in the ISM or from ISM-sourced precursors. However, the isotope enrichments in question are generally within a range that can be achieved by chemical isotope effects under a wide range of conditions, so this evidence alone is insufficient for a definitive association of these compounds with ISM chemistry. Conversely, IOM and monocarboxylic acids (MCAs) tend to have ^{13}C abundances that are closer to terrestrial composition (Huang et al., 2005; Huang et al., 2007; Aponte et al., 2011). Finally, heavily altered meteorites (above petrologic type 3.0 or below type 2.5) have generally lower abundances of SOM and less pronounced heavy isotope enrichments, suggesting parent-body processing and/or terrestrial contamination may influence the abundances and isotopic compositions of meteoritic OM (Elsila et al., 2016; Glavin et al., 2018).

A commonly postulated scenario for meteoritic SOM synthesis is as follows: (1) ISM-sourced aldehydes, ketones, ammonia, and possibly CN are concentrated on meteorite parent bodies; (2) the aldehyde and ketones are aminated and either reduced into amines or react with CN to form α-aminonitriles and α-amino acids upon hydrolysis; (3) some aldehydes are oxidized into MCAs (Aponte et al., 2017). Alternatively, MCAs have been hypothesized to form by a kinetically controlled synthesis such as FTT or as a combination of the two pathways (Yuen et al., 1984; Huang et al., 2005; Aponte et al., 2011; Glavin et al., 2018).

The scenario above provides a useful hypothetical framework for understanding the prebiotic synthesis of meteoritic SOM but has several shortcomings. Firstly, it requires that ketones are more ^{13}C-enriched than aldehydes to create the consistent ^{13}C enrichments amongst α-CH_3 amino and hydroxy acids and secondary amines and carboxylic acids relative to their α-H and primary analogs (Pizzarello et al., 2004; Elsila et al., 2012; Simkus et al., 2019). However, measurements of ketones are significantly lower in ^{13}C than aldehydes of equivalent size (Simkus et al., 2019; Aponte et al., 2019)

so the ketones used to create the measured products are required to have formed from a separate pool as those in measurements.

Secondly, the scenario outlined above cannot explain the high abundances of α-CH_3 compounds relative to their α-H analogs (Glavin et al., 2010; Elsila et al., 2016). For instance, the concentration of α-aminoisobutyric acid are equal to or higher than those of alanine. However, ISM measurements find that acetone (the ketone precursor to α-aminoisobutyric acid) is only 1/15 as abundant as acetaldehyde (the aldehyde precursor to alanine) (Combes et al., 1987), and ketones are less reactive than aldehydes in both Strecker synthesis and in reductive amination (Van Trump, 1975; Gomez et al., 2002; Lamm et al., 2013). Consequently, in addition to requiring another pool of ketones to explain the ^{13}C abundance data, the scenario requires that α-CH_3 and secondary compounds have a second synthetic pathway for formation. We note that while the abundance data could be explained by a large-scale degradation of α-H and primary compounds, this would result in the residual compounds being more ^{13}C-enriched, which disagrees with the ^{13}C abundance data.

A more general shortcoming of the scenario outlined above (and prior interpretations of the isotopic compositions of meteoritic organics generally) is that it is qualitative and has limited predictive capabilities. We note in particular that prior explanations of the stable isotope compositions of SOM compounds is that they do not predict isotopic structures (site-specific isotope ratios) in molecules, or quantitatively consider isotope effects in the chemistry called on to create specific compounds. For example, although dicarboxylic acids (DCAs) from the Murchison meteorite are higher in $\delta^{13}C$ than alkanes of generally similar size, it is not obvious whether we should expect the carboxyl groups are host to exceptional ^{13}C enrichments or, if so, whether the two carboxyl 'ends' of each compound are expected to be the same or different in $\delta^{13}C$. Several technologies are available for site-specific isotope ratio (SSIR) measurements of organic compounds, and Chimiak et al. (2020) has demonstrated that at least one of these methods (Orbitrap-based Fourier-transform isotope ratio mass spectrometry) is suitable for study of picomolar quantities of compounds extracted from a carbonaceous chondrite. This capability provides an

opportunity to test and substantially extend hypotheses regarding the origins of meteoritic organic matter, but such studies must be framed by quantitative predictions at the compound-specific and site-specific isotope ratios implied by a hypothesized organic synthesis scenario.

This present study compiles previous measurements of compound-specific and site-specific isotopic compositions of meteoritic organics from several carbonaceous chondrites, and uses those data to derive and demonstrate the self-consistency of a quantitative model of prebiotic organic synthesis, including a specified and ordered network of reactions, consideration of isotope effects associated with those reactions, patterns of inheritance from precursors to products, and simultaneous consideration of H and C isotope compositions. In addition to creating a more comprehensive and specific hypothesized scenario of meteoritic prebiotic chemistry, the output of this model provides falsifiable predictions of site-specific and molecular-average isotopic compositions of both previously studied and previously unstudied chemical species; thus, this model can be disproven, or revised and refined, in response to future measurements.

5.2. Methods

5.2.1. Data Compilation

We use previous measurements of organic compounds' isotope ratios for the compound classes and meteorites listed in Tables 5.1 and 5.2. These tables also include the references from which data was collected. Most studies include uncertainties of concentrations and isotope ratio measurements, and these are important to our effort because errors were considered when we examined the self-consistency of our hypothesized organic synthesis reaction network. In cases where a study omitted an error for the $\delta^{13}C$ of compound of interest to this present study, we assign an error equal the highest $\delta^{13}C$ error reported in that study. When a study provides no error analysis for $\delta^{13}C$, we use another study that measures the same property in the same compound class in the same meteorite and assign the highest error for the property of interest from that study to all the data in the study that has no error analysis.

Meteorite	**Murchison**	**Murray**	**LON 94101**	**QUE 99177**	**GRA 95229**	**ALH 83100**	**EET 92042**	**ALH 85013**	**Orgueil**	**EET 87770**	**ALH 84034**	**ALH 84033**	**WIS 91600**	**MET 00430**
Pet Type (Weather/Frac)	CM 1.6	CM 1.5	CM 1.8 (Ce,C)	CR 2.4 (Be,B)	CR 2.5 (A, A/B)	CM 1.1 (Be,B/C)	CR 2.5 (B,B)	CM 1.4 (A,A/B)	CI 1.1	CR 2 (B,A)		CM2 (Ae,B)	C2 (A/Be,A)*	CV3 (B,B)
monocarboxylic a-amino acids	10[1,2]	9[1]		6[2]										
dicarboxylic a-amino acids	5[1]	6[1]												
non-a-amino acids	9[1,2]	14[1]												
all amino acids	24[1,2]	29[1]	7[2]	8[2]	13[3]	3[2]	8[2]							
hydroxy acids	9[4]													
sulfonic acids	4[5]													
DCA	4[6]													
MCA	31[7,8]													
MCA (Aponte only)	16[8]									16[8]	11[8]	13[8]	3[8]	11[8]
IOM Aliphatic	16[8]									19[8]	16[8]	10[8]	12[8]	12[8]
IOM Aromatic	9[9]					9[9]	9[9]	9[9]	14[10]					
notes												*"heated?" in Alexander 2013, but T<20 in Aponte et al 2011	*highly altered in oxidative media	

Table 5.1: Total number of compounds in different classes that are analyzed in the hydrogen moiety model for all meteorites used in this study. Superscripts refer to references for the measurements used in this study and data are listed in Tables S5.10 and S5.11

[1]Elsila et al. (2012), [2]Elsila et al. (2012), [3]Pizzarello et al. (2008), [4]Pizzarello et al. (2010), [5]Cooper et al. (1997), [6]Pizzarello et al. (2002), [7]Huang et al. (2005), [8]Aponte et al. (2011), [9]Wang et al. (2005), [10]Remusat et al. (2006)

	Meteorite Petrologic class (Weathering/Fracturing), Alter. T (ºC)	**Murchison** CM 1.6, 20-33	**ALH 83100** CM 1.1 (Be, B/C), 50-250	**LEW 90500** CM 1.6 (B, A),0-130	**LON 94101** CM 1.8 (Ce,C), 25-35	**MIL 090001** CR 2.2 (B, B/C), 250-300	**LAP 02342** CR 2.5 (A/B, A/B), 0-88	**GRA 95229** CR 2.5 (A, A/B), 230
aldehydes/ ketones	aldehydes (straight chain)	9[1,2]	4[2]	3[2]	4[2]	2[2]	0	3[2]
	aldehydes (branched chain)	2[1,2]	2[2]	2[2]	2[2]	0	0	0
	ketone	4[1,2]	2[2]	2[2]	2[2]	2[2]	0	2[2]
	total aldehydes and ketones	**15**	**6**	**7**	**8**	**4**	**0**	**5**
CN	**CN**	1[3]	0	0	0	0	0	0
amines	amines (straight chain)	4[4]	4[5]	6[5]	4[5]	3[6]	3[5]	3[5]
	amines (branched chain and non 1°)	9[4]	6[5]	8[5]	7[5]	1[6]	3[5]	3[5]
	total amines	**13**	**10**	**14**	**11**	**4**	**6**	**6**
MCA	monocarboxylic acids (straight chain)	13[7,8,9]	3[9]	5[9]	5[9]	2[10]	4[9]	3[9]
	monocarboxylic acids (branched chain)	8[7,8,9]	0	4[9]	0	0	2[9]	1[9]
	total monocarboxylic acids	21	3	9	5	2	6	4
DCA	dicarboxylic acids (straight chain)	3[11]	0	0	0	0	0	0
	dicarboxylic acids (branched chain)	3[11]	0	0	0	0	0	0
	total dicarboxylic acids	**6**	**0**	**0**	**0**	**0**	**0**	**0**
α-amino acids	monocarboxylic α-H straight chain amino acids	12[12,13,14,15]	1[14]	2[14]	3[14]	3[10]	3[16]	5[14,17]
	monocarboxylic α-H branched-chain and α-CH3 amino acids	11[12,13,14,15]	0	0	3[14]	1[10]	2[16]	5[14,17]
	dicarboxylic α-H straight chain amino acids	3[12]	0	0	0	0	1[16]	2[17]
	dicarboxylic α-H branched-chain and α-CH3 amino acids	3[12]	0	0	0	0	0	0
	total α-amino acids	**29**	**1**	**2**	**6**	**4**	**6**	**12**

	Meteorite Petrologic class (Weathering/Fracturing), Alter. T (ºC)	**Murchison** CM 1.6, 20-33	**ALH 83100** CM 1.1 (Be, B/C), 50-250	**LEW 90500** CM 1.6 (B, A), 0-130	**LON 94101** CM 1.8 (Ce,C), 25-35	**MIL 090001** CR 2.2 (B, B/C), 250-300	**LAP 02342** CR 2.5 (A/B, A/B), 0-88	**GRA 95229** CR 2.5 (A, A/B), 230
non-α-amino acids	non-α-amino acids (straight chain)	5[13,14,15]	0	0	0	0	0	0
	non-α-amino acids (branched chain)	3[13]	0	0	0	0	0	0
	total non-α-amino acids	**8**	**0**	**0**	**0**	**0**	**0**	**0**
α-hydroxy acids	α-hydroxy acids (straight chain)	4[18]	0	0	0	0	0	0
	α-hydroxy acids (branched chain)	3[18]	0	0	0	0	0	0
	total α-hydroxy acids	**7**	**0**	**0**	**0**	**0**	**0**	**0**

Table 5.2: Total number of compounds in different classes and subclasses that are analyzed or input into the carbon reaction model for all meteorites used in this study. Superscripts refer to references for the measurements used in this study and listed in Tables S5.2 and S5.4 to S5.9. Superscripts are as follows: [1]Simkus et al. (2019), [2] Aponte et al. (2019), [3]Pizzarello (2014), [4]Aponte et al. (2016), [5]Aponte et al. (2016), [6]Aponte et al. (2018), [7]Yuen et al. (1984), [8]Huang et al. (2005), [9]Aponte et al. (2019b), [10] Burton et al. (2015,[11]Pizzarello et al. (2002), [12]Engel et al. (1990), [13]Pizzarello et al. (2004) [14]Elsila et al. (2012), [15]Glavin et al. (2020), [16]Pizzarello et al. (2009), [17]Martins et al. (2007), [18]Pizzarello et al. (2010).

The alteration temperature, weathering grade, fracturing grade, petrologic type, and classes of the meteorites surveyed and their respective references are described in Table 5.3.

	Meteorite	Petrologic Type[1]	Peak alt T (°C)[2]	Weathering[1]	Fracturing[1]
D analysis	Orgueil	CI 1.1	<220		
	ALH 84034	CM 1.1	50–250	A	A
	ALH 85013	CM 1.4	0-130	A	A/B
	Murray	CM 1.5	30		
	LON 94101	CM 1.8	25 to 35	Ce	C
	ALH 84033	CM2 (possibly heated)	<20	Ae	B
	QUE 99177	CR 2.4		Be*	B
	EET 92042	CR 2.5	220	B	B
	EET 87770	CR 2 (min alt)	<20	B	A
	MET 00430	CV3	260 to 400	B	B
	WIS 91600	C2 (extensive alt in ox media)	50–100	A/Be	A
D and ^{13}C analysis	Murchison	CM 1.6	20-33		
	ALH83100	CM 1.1	125 ± 65	Be	B/C
	LON 94101	CM 1.8	25 to 35	Ce	C
^{13}C analysis	LEW 90500	CM 1.6	0-130	B	A
	MIL 090001	CR 2.2	250-300	B	B/C
	LAP 02342	CR 2.5	≤0-88	A/B	A/B
	GRA 95229	CR 2.5	230	A	A/B

Table 5.3: Meteorites from which data in this study was compiled. Petrologic type, peak alteration temperature, weathering grade, and fracturing grade are compiled from multiple studies labelled in subscripts: [1]Alexander et al. (2013), Antarctic Meteorite Classification Database(Anon, n.d.) , Aponte et al. (2019b), Aponte et al. (2011), [2]Aponte et al. (2011), Alexander et al. (2015), Guo et al. (2007), Telus et al. (2019), Lee et al. (2013), Keller (2011), Nittler *et al* (2019).

5.2.2. Methods of model calculation

The principal product of this study is a model chemical reaction network that tracks both amounts and isotopic compositions (δD and $\delta^{13}C$ values, including site-specific differences) of organic molecules that are hypothesized to derive from an initial set of substrates.

The variations in compound-specific D/H ratios among meteoritic organics are large multiples of the variations in those ratios that generally arise from common chemical reactions (100's to 1000's of ‰ vs. the 10's to 100's of ‰ for reactions pertinent to the processes considered in our model), whereas variations in $^{13}C/^{12}C$ are broadly similar to amplitudes of common equilibrium and kinetic isotope effects. For that reason, our model of site- and compound-specific D/H ratios is simpler than our treatment of carbon isotope compositions; specifically, we focus on identifying chemical moieties that may be responsible for differences in D/H ratio between compounds, and discuss the patterns of inheritance and isotope exchange that may explain those findings, whereas the more subtle and complex variations in carbon isotope composition are described through a full treatment of inheritance and isotopic fractionation associated with each reaction in the network.

5.2.2.1. Deuterium

Our model is inspired by the trend between the number of methyl groups a compound has and its δD that suggests that the methyl group controls the degree of enrichment that a compound exhibits (Figure 5.1). To this end, we evaluate the molecular-average

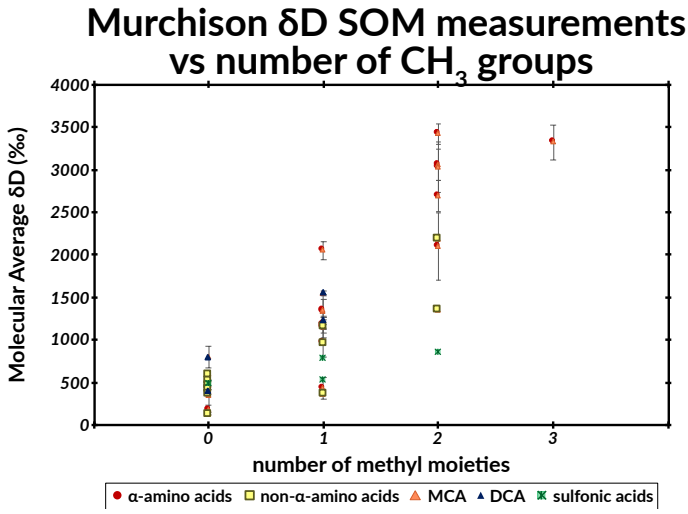


Figure 5.1: The measured molecular-average δD values of SOM plotted arroding to its number of methyl moieties

fractional abundance of deuterium as a weighted average of its abundance in different moieties—groups of deuterium that shared similar chemical environments within a molecule—for a variety of SOM and IOM components on multiple meteorites (Table 5.1, S5.1, and S5.2). These moieties consist of hydrogens bonded to five types of carbon: methyl C, non-methyl alkyl C, resonant ring C, and hydrogens attached to substituted carbons (*i.e.,* H-CNH_2, H-COH) at either the α- or non-α-sites (Figure 5.2). Throughout this paper, we will refer to these sites as methyl-H, chain-H, ring-H, α-H, and non-α-H, respectively. Exchangeable hydrogens are neither included in our analyses nor in the compound-specific molecular-average δD measurements we use. Figure 5.2 depicts the moieties used in α-aminobutyric acid and phenylmethanamine. In α-aminobutyric acid, the model includes the three methyl-Hs, two chain-Hs, and one α-H but not the two hydrogens attached to the nitrogen in NH_2 nor the hydrogen on the carboxyl group (COOH) as they are exchangeable during extraction. In phenylmethanamine, the model includes the ring-Hs and the two non-α-H but again excludes the two hydrogens attached to nitrogen in NH_2.

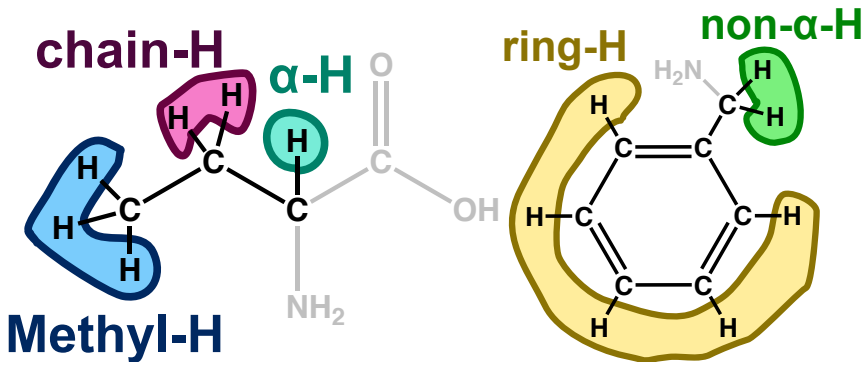


Figure 5.2: Examples of hydrogen moieties used in model.

For each compound, we assign a moiety type to each hydrogen and sum them to calculate how many hydrogens each moiety contributes to the compound's chemical structure. For example, α-aminobutyric acid has six non-exchangeable hydrogens: three are methyl hydrogens, two are on the aliphatic chain, and one is bound to the α-C (Figure 2). Compounds are then grouped by their chemical class: α-amino acid, non-α-amino acid, amine, monocarboxylic acid (MCA), dicarboxylic acid (DCA), sulfonic acid (SA),

aromatic polyaromatic hydrocarbons (PAHs), and aliphatic moieties of IOM. For each class, we combine the measured molecular-average isotope ratios and the number of each type of hydrogens for molecules in that class to find the least-square solution to Equation 5.1 for each ^2F_s (the fractional abundance of D for hydrogens of type, s, in the compound class; see Footnote 5.1 for an explanation of fractional abundance notation):

$$^2F_{\text{molec avg}} = \frac{n_{H_a}\,{}^2F_a + n_{H_b}\,{}^2F_b + \cdots + n_{H_z}\,{}^2F_z}{n_{H_a} + n_{H_b} + \cdots + n_{H_z}}\ ,\ {}^2F_s \geq 0 \qquad (\textit{Eqn. 5.1})$$

where $^2F_{\text{molec avg}}$ is the average fractional abundance of D (*i.e.,* $\frac{D}{D+H}$) for the compound, n_{H_s} is the number of hydrogens in of type, s, in the compound, and ^2F_s is described above. All fits for Equation 5.1 are over constrained as the number of compounds in each class is greater than the number of types of moieties (Table 5.1); however, the degree to which the model is over constrained varies between compound classes. The discussion section focuses on the compound classes that are more strongly over constrained.

5.2.2.2. Carbon Isotopes

Our model of the carbon isotope contents of individual compounds begins with the carbon isotope composition for the reactants that create the carbon backbone in each synthesis. These reactant compounds are listed in Table 5.4 and their values are included in Tables S5.4 to S5.10. For the integrated aldehyde network, reactants are aldehydes or ketones and cyanide. For the Formose-like reaction, the only reactant is formaldehyde. For FTT reactions, we test both CO_2 and carbonate as the reactant because the two pools are known to exchange. We assume that the $\delta^{13}C$ values of the reactants for each synthesis are equal to those reactants measured on the meteorite whose chemistry is being modeled (*i.e.*, we assume that reactants on meteorites that are measured are from pools that were either isolated from the synthesis or were part of a reservoir that was large enough relative to the products that the isotopic composition did not change during synthesis). Therefore, when modeling a Formose-like reaction on Murchison, the only

	integrated aldehyde network	integrated aldehyde network + formaldehyde addition		integrated aldehyde network + COOH equilibration	FTT	Formose type
		aldehyde as precursors	analogs as precursors			
precursors	Aldehydes Ketones Cyanide	Aldehydes Cyanide	Analog α-H and straight-chain compounds Formaldehyde	Aldehydes Ketones Carboxyl	Carbon dioxide Carbonate	Formaldehyde
Reaction (KIE)	CN to COOH (-30‰ on COOH)[1]	CN to COOH (-30‰ on COOH)[1] Aldol condensation (-30‰ on C being added, -30 ‰ on C=O added to)[2]		-	C addition (-10 ‰ per C)[3]	Aldol condensation (-30‰ on C being added, -30 ‰ on C=O added to)[2]
Reaction (EIE)	Aldehyde + CN (-20‰ on C=O)[4]				-	-

Table 5.4: Precursors and isotope effects used in each version of the model. Superscripts refer to papers in which the isotope effects are discussed. Superscripts are as follows: [1](Chimiak et al., 2020), [2](Zhu et al., 2009), [3](Taran et al., 2007)

precursor is formaldehyde, and we use the $\delta^{13}C$ measurements of formaldehyde extracted from Murchison meteorite samples.

We then constructed a reaction network that combined these precursors in various orders and combinations dictated by several different reaction mechanisms: Fischer Tropsch-type synthesis (FTT synthesis), Formose chemistry, Strecker synthesis, reductive amination, and carbon-oxidation reactions. We also consider the possibility that α-CH_3 amino acids (formed by Strecker synthesis in our model) might be able to undergo carbon isotope exchange between their carboxyl sites and dissolved inorganic carbonate (which we assume has a carbon isotope composition equal to that of solid carbonates in each sample)

(Pietrucci et al., 2018). In our model, Strecker synthesis, reductive amination, and carbon-oxidation reactions each create one type of product (α-amino acids, amines, and carboxylic acids, respectively), by reactions that draw on the same aldehyde pool. This portion of our reaction network resembles that put forth in other studies (Aponte et al., 2017; Aponte et al., 2019; Chimiak et al., 2020); here we will refer to it as 'integrated aldehyde chemistry'. In addition to modeling the chemistry in which ketones and branched-chain aldehydes are the reactants for branched-chain and α-CH_3 compounds, we also test the possibility of the integrated aldehyde chemistry forming straight-chain compounds and these compounds undergoing formaldehyde addition to form branched and α-CH_3 compounds. While we note the possibility that multiple reaction pathways might lead to the formation of a single product measured here, for simplicity, the reaction network we present assumes that each compound class is predominantly produced by one set of chemical steps.

For each elementary step in each reaction mechanism, we apply carbon isotope effects between reactants and products based on previously experimentally measured isotope effects for that reaction mechanism or for similar reactions (Table 5.4). These isotope effects are combined with assigned $\delta^{13}C$ values of precursors for each reaction step to predict the $\delta^{13}C$ value of the product. These calculations track the $\delta^{13}C$ values of each carbon transferred from precursor to product and apply relevant isotope effects to the carbon atom or atoms to which they are known to apply. The molecular average $\delta^{13}C$ of the product was then calculated as the weighted average of the site-specific carbon isotope contents of its carbon atoms.

These calculations were performed under two conditions that result in minimum and maximum molecular-average $\delta^{13}C$ values of the product. In the first condition (which yields the highest modeled site-specific and molecular-average $\delta^{13}C$ values), we consider the case where carbons from precursors are quantitatively transferred to products, in which case chemical isotope effects associated with the reaction are not expressed and the product simply inherits the carbon isotope signatures of the reactants, appropriately weighted for the sites that contribute to product and their proportions in the product.

Given that the precursors in question for most reactions of interest (e.g., aldehydes, HCN, etc.) are still present in the samples we studied, this model case would correspond to a physical situation where the reactions occur with near quantitative yield on surfaces or restricted volumes, but overall rate is limited by some intervening process that is involves minimal isotopic fractionation (such as advection or diffusion through a condensed phase; *i.e.*, much as C_4 photosynthesis leads to modest carbon isotope fractionation despite the fact that the elementary reaction of CO_2 with the RuBiSCO enzyme in those plants is highly fractionating). In this case we can use Equation 5.2 to calculate the molecular average $\delta^{13}C$ of the product:

$$\delta^{13}C_{product} = \frac{n_{C_a}\delta^{13}C_a + n_b\delta^{13}C_b + \cdots + n_z\delta^{13}C_z}{n_{C_{product}}} \qquad (Eqn.\ 5.2)$$

where $\delta^{13}C_{product}$ is the molecular-average $\delta^{13}C$ of the product, n_{C_x} is the number of carbon atoms in precursor x, $\delta^{13}C_x$ is the fractional abundance of ^{13}C in precursor x, and $n_{C_{product}}$ is the total number of carbons in the product. We use $\delta^{13}C$ values in these calculations because the uncertainties in the isotope effects and those in the measurements are larger than the imprecision that arises from non-linearities in delta space for the values used here—when tested on a representative subset of data (*i.e.*, including lower and higher enriched compounds), it results in predictions that differed from those in delta space by less than 2 ‰ and typically less than 1 ‰. These errors are lower than many of the analytical errors associated with aldehyde $\delta^{13}C$ measurements used in the model.

The second case we consider involves maximal isotopic fractionation at each reaction step, leading to the lowest possible $d^{13}C$ values of products in the context of our model reaction network (*i.e.*, because all of the carbon isotope fractionations we consider are 'normal' – favoring ^{12}C in products). To this end, modify equation two by subtracting the isotope effect from each reactant's $\delta^{13}C$ (evaluated site by site for both reactants and products) prior to calculating the weighted average of the reactant's $\delta^{13}C$:

$$\delta^{13}C_{product,low} = \frac{n_{C_a}(\delta^{13}C_a - \varepsilon_{product-a}) + n_b(\delta^{13}C_b - \varepsilon_{product-b}) + \cdots + n_z(\delta^{13}C_z - \varepsilon_{product-z})}{n_{C_{product}}}$$

(*Eqn. 5.3*)

Here $\delta^{13}C_{product,low}$ is the lower bound for the product's predicted $\delta^{13}C$ and $\varepsilon_{product\text{-}x}$ is the carbon isotope effect between the product's carbon from reactant x and the carbon in reactant x in units of per mil. (Again, we note that this expression is an approximation because it treats $d^{13}C$ values as conservative properties that can be manipulated with simple arithmetic operations, and it is known that this practice leads to subtle (~0.1 ‰) non-linear artifacts.)

Finally, we note the following exceptions to the methods outlined above, where we introduced an additional assumption or approximation to address a gap in existing analytical data or experimental constraints on our model:

In the FTT model we aim to test both CO_2 and carbonate as precursors. As there is only one CO_2 $\delta^{13}C$ measurement in the literature and that is for Murchison meteorite, most models only use $\delta^{13}C$ measurements of carbonate.

The integrated aldehyde network model has multiple exceptions that we will note here. Firstly, we use the $\delta^{13}C$ measurement of CN from Murchison as the CN reactant value in all meteorites' α-amino acid synthesis because that is the only $\delta^{13}C$ measurement of CN in the literature. Secondly, we use aldehyde and ketone $\delta^{13}C$ measurements from GRA 95229 (CR 2.5) as precursor $\delta^{13}C$ values for aldehydes and ketones in LON 94201 (CR 2.5) as these two meteorites are of the same class and petrologic type. Finally, as dicarboxylic compounds (*e.g.*, dicarboxylic acids and dicarboxylic amino acids) synthetic origins are debated and they have no measured precursors, we test two scenarios for their precursors, that of an oxonitrile precursor and that of a dialdehyde precursor. Formyl cyanide, an oxonitrile, has been detected in the ISM (McGuire, 2018), and dialdehydes such as glyoxylate are often invoked in prebiotic synthesis (de Marcellus et al., 2015) and are close relatives of molecules such as glycolaldehyde, which has been detected in the

ISM (McGuire, 2018). In each case, we predict the precursor's initial $\delta^{13}C$ value Equation 5.2 with the $\delta^{13}C$ of the corresponding straight-chain aldehyde and that of either cyanide or formaldehyde.

The formaldehyde addition chemistry also has one exception to note. Formaldehyde addition is used to create branched structures that are one carbon in length including α-CH_3 amino acids or secondary amines; however, in the case of the tertiary amine, 3-pentylamine, we model the chemistry with acetaldehyde addition in place of formaldehyde addition.

5.3. Results

The isotopic models of hydrogen and carbon provide distinct information about each system. The hydrogen data provides evidence of ISM-sourced moieties and of subsequent exchange, primarily due to interactions with local water on the meteorite parent body. The carbon model demonstrates the unique role that carbon plays in organic synthesis on meteorites. The integrated aldehyde network with subsequent formaldehyde addition has the highest accuracy and the lowest average residuals for its $\delta^{13}C$ predictions of compounds on Murchison. Combined the models demonstrate the likelihood that amines, amino acids, and carboxylic acids are derived from ISM-sourced aldehydes.

Our choices of fitted moieties for the hydrogen isotope model and our formulation of the carbon isotope reaction network (*i.e.*, selection of reaction types and organization into the presented order of reactions) were both decided on based on the systematics of isotopic data for SOM compounds from the Murchison meteorite, for the simple reason that it is far more extensively characterized than any other sample considered in this study (Table 5.1, S5.1). We then apply those models to data for other meteorites (Table 5.1, S5.2); in the case of the moiety-specific deuterium model this application entailed fitting a different set of moiety D/H ratios to data for each sample, whereas in the case of the carbon isotope reaction network we compare the predictions of the network to

the data for these other samples, with the only changes being the assignment of substrate $\delta^{13}C$ values based on measurements of aldehydes from each sample.

5.3.1. Deuterium

We fit the measured δD values of members of each of several groups of SOM compounds (MCAs, DCAs, sulfonic acids, hydroxy acids, α- and non-α-amino acids, amines, PAHs, and both aliphatic IOM linkages and aromatic IOM) as functions of the δD values for methyl-H, chain-H, ring-H, α-H, and non-α-H in each compound group, and performed these least squares fits separately for each carbonaceous chondrite studied here (Table 5.5, S5.1, S5.2). The resulting best-fit moiety-specific δD values for the Murchison meteorite follow similar trends amongst all classes of compounds, including both SOM and included components of IOM: methyl-H and non-α-H (when present) always have the highest δD values (typically by 100s of ‰ on the VPDB scale), the α-H (when present) has the lowest fitted δD (Table 5.5 and Figure 5.3b). The range of this model spans ~3000 ‰ for moieties from compound classes in the Murchison meteorite and a larger range for those in CR2 meteorites. This range is an order of magnitude greater than the typical analytical reproducibility of the data (on the order of 100-300 ‰), so we omit error bars for clarity. The fitted δD value of methyl-H for each compound class is inversely correlated with median solubility in water for compounds of that class used in this study (Table S5.3 in Appendix D, Figure 5.3c) as found by recording the water solubility for all compounds used in this study, arranging the data by compound class, and finding the median value for the water solubility. Finally, the differences in fitted D/H between methyl-H, chain-H, and ring-H are broadly similar for all compound classes and samples (Figure 5.3b).

When the moiety-specific D/H ratio model is fit to compound-specific hydrogen isotope data for other carbonaceous chondrites (*i.e.*, other than Murchison), the trends observed for moiety-specific fitted δD values of the compounds studied here—both the SOM and IOM components—are the same as those in Murchison (Figures 5.4a and 5.5). Chiefly,

		Moiety δD Predictions (‰)					Methyl-H-normalized δD Predictions (‰)			
Meteorite	**Compound class**	**Methyl-H**	**non-α-H**	**Ring -H**	**Chain-H**	**α-H**	**non-α**	**Ring-H**	**Chain-H**	**α-H**
Murchison	monocarboxylic α-amino acids	2834			835	-130			-521	-773
	dicarboxylic α-amino acids	2765			617	0			-570	-734
	non-α-amino acids	1279	794	129	327		-213	-505	-418	
	all amino acids	2788	77	129	465	-135	-716	-702	-613	-772
	hydroxy acids	1492	1889		125	1556	159		-548	26
	sulfonic acids	936	492		-19		-229		-493	
	DCA	149			2346				1912	
	MCA	2322		400	129			-579	-660	
	MCA (Aponte only)	2147			-307				-780	
	IOM Aliphatic	1401			-236	-11			-682	-558
	IOM Aromatic	1262	5216	486	372		1748	-343	-393	-558
Murray	monocarboxylic α-amino acids	2798			1484	-191			-346	-787
	dicarboxylic α-amino acids	2769			690	-1000			-552	-1000
	non- α-amino acids	2138	899		586		-395		-495	
	all amino acids	2772	272		752	-440	-663		-536	-852
ALH 83100	all amino acids	410	-399		286	355	-574		-88	-39
EET 92042	all amino acids	5493	6429		177	191	144		-819	-817
	α-amino acids	5392			843	181			-712	-815
LON 94101	all amino acids	1825	512		1318	133	-465		-179	-599
QUE 99177	all amino acids	3225	1711		1596	966	-358		-385	-535
	monocarboxylic α-amino acids	2851			4079	929			319	-499
GRA 95229	all amino acids	4121	3199		1015	344	-180		-606	-738
	α-amino acids	4139			886	333			-393	-741
	Hydroxy acids	2992	-31		2380	2439	-757		-153	-138
EET 87770	IOM Aliphatic	1871			-107	-65			-689	-674
	MCA	1334			508				-354	
ALH 84034	IOM Aliphatic	1546			-90	-36			-643	-621
	MCA	327			-172				-376	
ALH 84033	IOM Aliphatic	352			-275	260			-463	
	MCA	370			360				-7	
WIS 91600	IOM Aliphatic	1912			-342	-25			-774	-665
	MCA	458			-251				-487	

		Moiety δD Predictions (‰)					Methyl-H-normalized δD Predictions (‰)			
Meteorite	**Compound class**	**Methyl-H**	**non-α-H**	**Meteorite**	**Compound class**	**Methyl-H**	**non-α-H**	**Meteorite**	**Compound class**	**Methyl-H**
MET 00430	IOM Aliphatic MCA	709 172			-643 462	25			-791 248	
EET 92042	IOM Aromatic	4388	5430	2976	2010		194	-262	-441	
ALH 83100	IOM Aromatic	1252	2506	707	607		557	-242	-287	
ALH 85013	IOM Aromatic	1841	1152	586	-936		-243	-442	-978	
Orgueil	IOM Aromatic	2063		1839	216			-73	-603	

Table 5.5: Predictions of hydrogen moiety δD values in ‰ units.

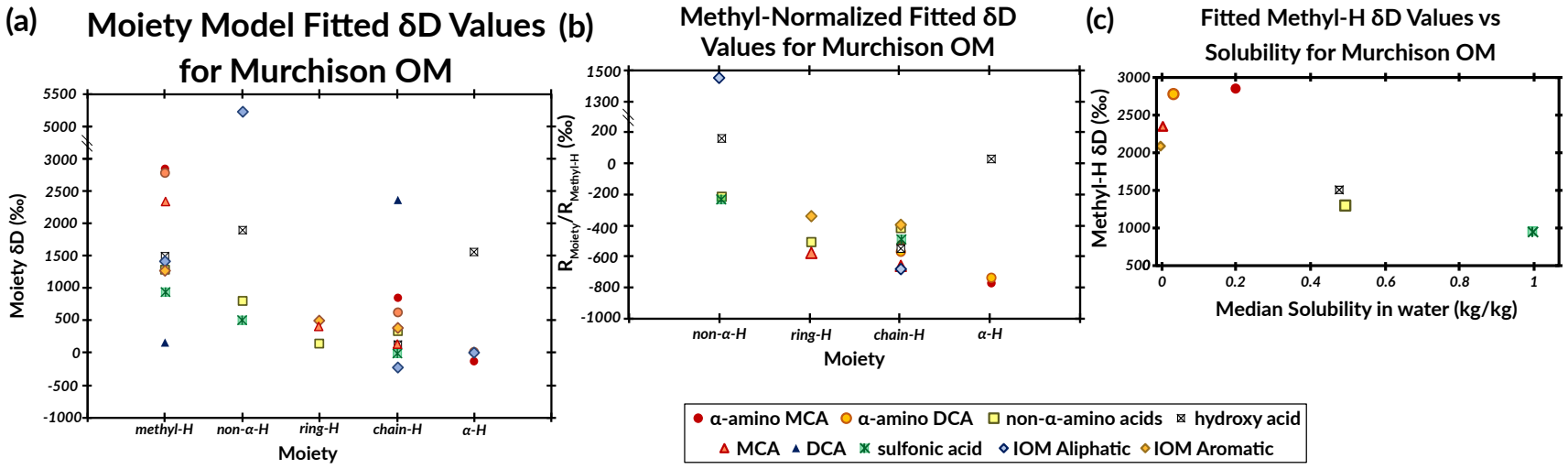


Figure 5.3: δD predictions for hydrogen moieties in Murchison meteorite. Predictions are graphed as (a) absolute values for the average δD of the moiety, (b) average δD of the moiety normalized to that of the methyl moiety, and (c) relative to the median solubility of the compound class. On plot (c) only values of the methyl moiety δD are included. Scale on each has a split in order to plot the IOM aliphatic chain on the same plot without losing vertical resolution.

the methyl-H and in some cases the non-α-H moieties are the most deuterium-enriched components and ring-H and chain-H tend to have the similar δD values.

In terms of absolute values (rather than trends), we notice two patterns: one as a result meteorite classes and one as a result of alteration. For all compounds studied, the fitted δD values for methyl-H and some non-α-H moieties are higher in CR chondrites than in CM chondrites (including Murchison) (Figures 5.4 and 5.5, Table 5.5). Orgueil, the one CI chondrite studied here overlaps with the CM chondrites for the fit δD values for methyl-H. In terms of alteration, the one thermally altered sample (MET 00430, CV 3) and the one heated sample (ALH 84033, CM 2), the methyl-H fitted δD values were amongst the lowest for their categories, suggesting a thermally influenced path to hydrogen exchange (Figure 5.5a and 5.5b, Table 5.5). Due to the small number of meteorite samples and compound classes, we cannot provide a more detailed analysis of this process. For SOM studied here, aqueous alteration and terrestrial weathering (as measured by petrologic grade and weathering index) lead to a decrease in the methyl-H fitted δD values (Figure 5.4b, Table 5.5). We find it particularly noteworthy that Murchison and Murray are the two most similar samples in terms of class (CM), petrologic grade (1.6 and 1.5 for Murchison and Murray, respectively), and weathering (A); and, they have closely similar fitted values of moiety-specific D/H ratios of all compound groups, particularly when expressed as differences of each moiety from the methyl-H fitted value (Figure 5.4a, Table 5.5).

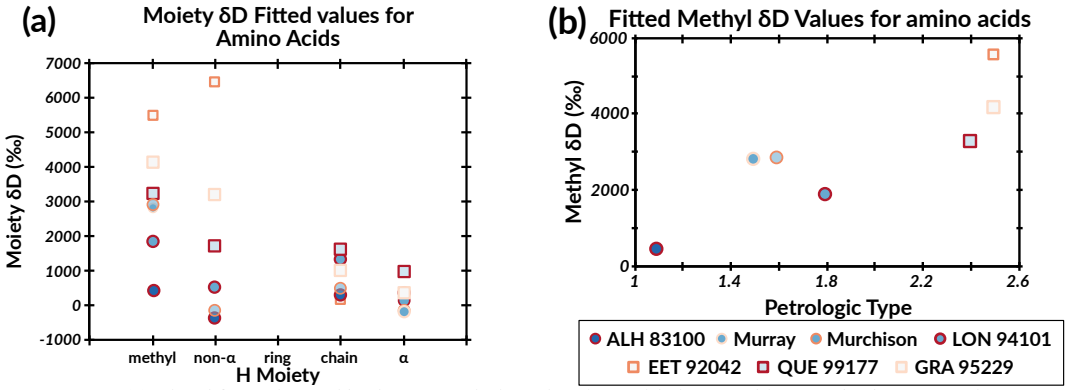


Figure 5.4: (a) Fitted δD values of hydrogen moieties of amino acids in Murchison and other meteorites. For meteorite IDs, see Table 2. Shapes of symbols refer to the meteorite class (circle = CM, square = CR). Outside of symbol is shaded from white to red accodring to the highest value for the weathering and fracturing grade with white representing a grade of A for both and red representing a grade of Ce for at least one. The interior is shaded by petrologic class with dark blue being 1.1, white being 3.0, and red being a heated sample. Sources of data are listed in Table 1. (b) Same data but only the fitted δD values of methyl-H versus the petrologic type

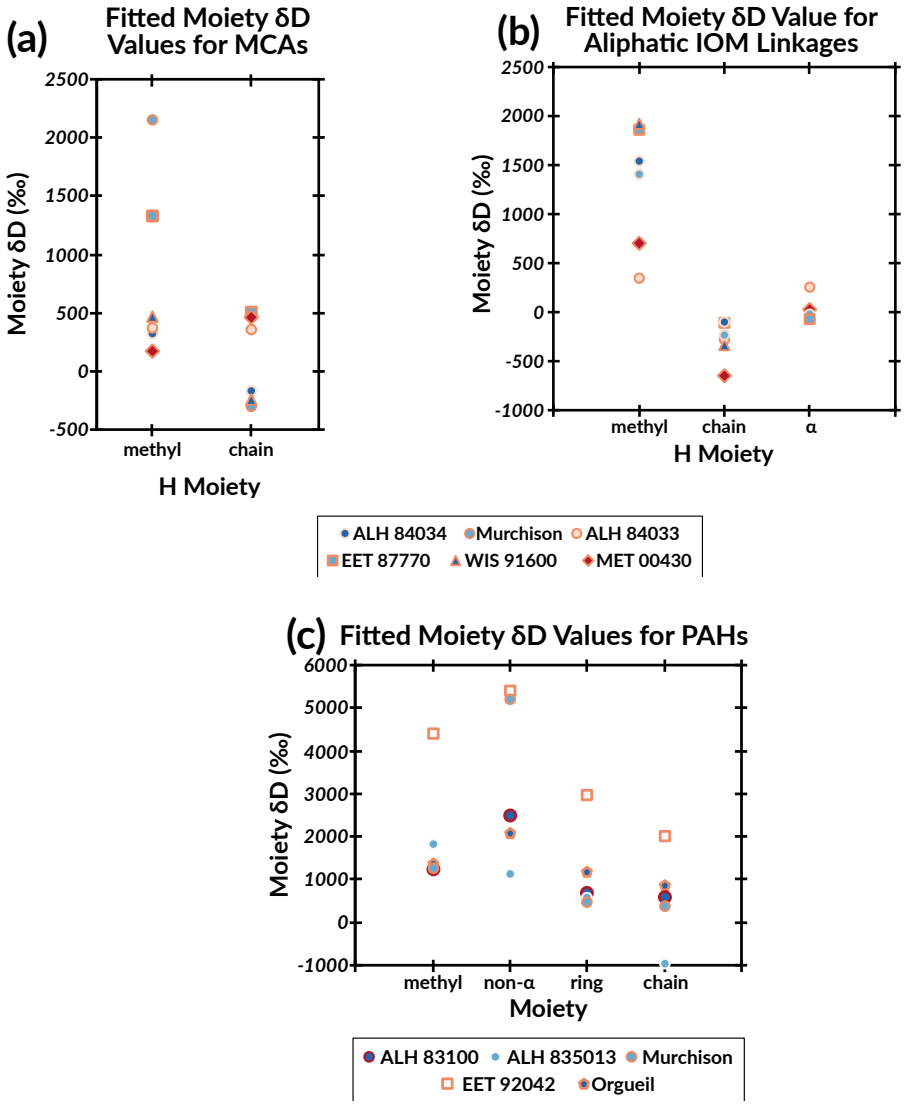


Figure 5.5: Fitted δD values for hydrogen moieties of compounds Murchison and other meteorites. For meteorite IDs, see Table 2. Shapes of symbols refer to the meteorite class (circle = CM, square = CR, diamond = CV, Pentagon = CI). Outside of symbol is shaded from white to red accodring to the weathering and fracturing grade with red being a grade of Ce. Interior is shaded by petrologic class with dark blue being 1.1, white being 3.0, and red being a heated sample. Sources of data are listed in Table 1. (a) Fitted δD values for hydrogen moieties in MCAs, (b) Fitted δD values for hydrogen moieties in aliphatic linkages in IOM, (c) Fitted δD values for PAHs.

5.3.2. Carbon

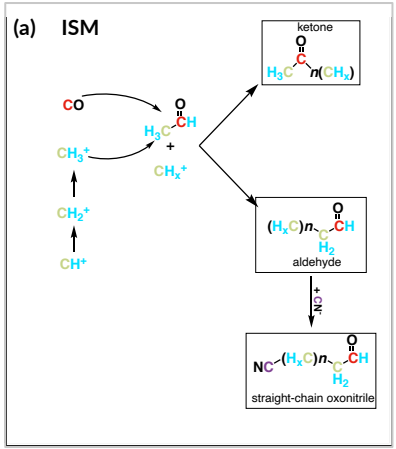
5.3.2.1. Reaction Network Comparison

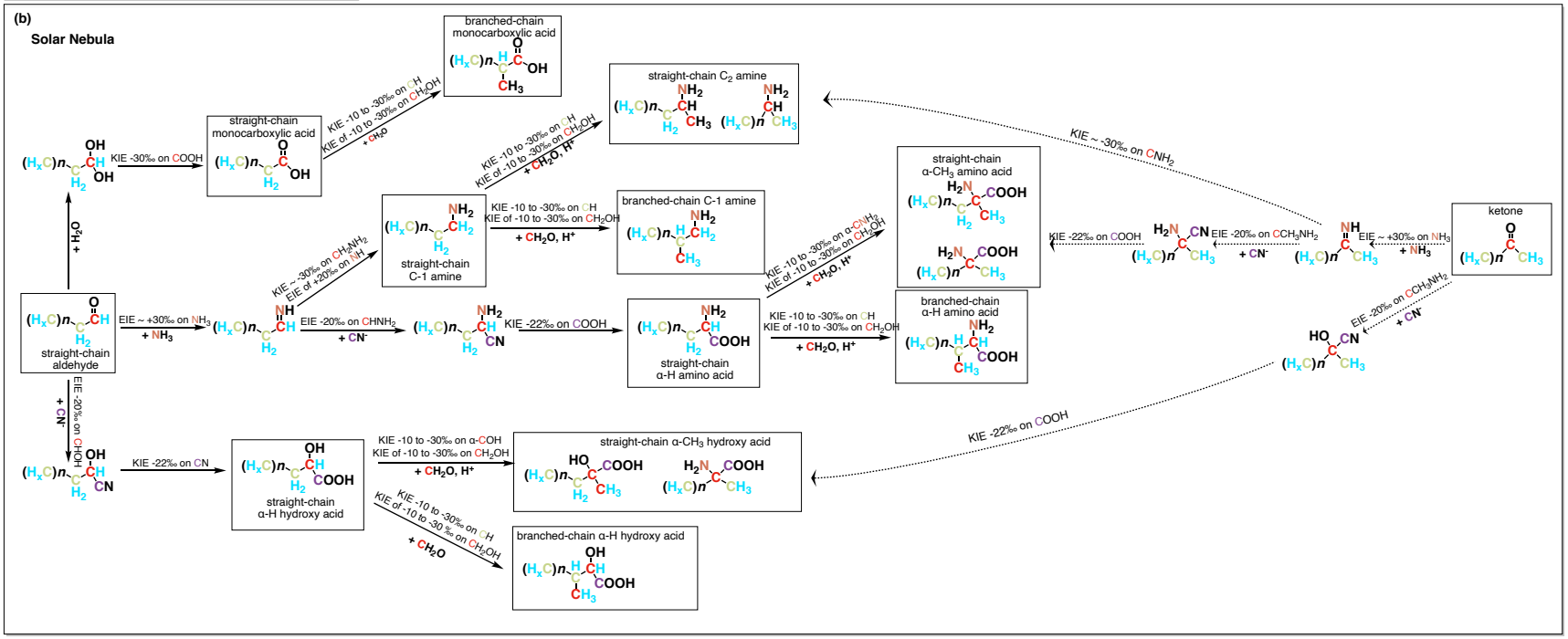
In this section, we compare the measured carbon isotope compositions of meteoritic SOM compounds to those expected based on our proposed reaction network (Figure 5.6;

we do not attempt to predict the carbon isotope contents of IOM moieties discussed in the preceding section because they are effectively insoluble in water and so would be able to form their own reservoir separate from the SOM and would likely not undergo the same aqueous chemistry proposed here). This model reaction network was not fit to these data in any formal sense; instead, we examined several chemically plausible scenarios, present the one that best matches observations for Murchison SOM $\delta^{13}C$ values (Table 5.6), and discuss how that model does (or does not) match similar observations in other, less extensively studied samples.

The central features of our proposed reaction model that are most essential to observed describing patterns of SOM $\delta^{13}C$ values are: (1) the integrated aldehyde chemistry (see 5.2.2.2 for description), which creates straight-chain and α-H compounds; (2) aldehyde addition to those straight-chain compounds (typically drawing on formaldehyde as a precursor but using acetaldehyde in the case of 3-pentylamine) to form branched-chain and α-CH_3 species; and (3) formation dicarboxylic species through chemistry resembling that which forms monocarboxylic species, but starting with oxonitrile precursors rather than aldehyde precursors (and accompanied by hydrolysis of the nitrile moiety into a second carboxyl group on those oxonitriles).

Focusing first on the data for the Murchison meteorite, if we 'seed' the proposed reaction network with the known $\delta^{13}C$ values of straight-chain aldehydes, we can predict the $\delta^{13}C$ values of the straight-chain amines, a-H amino acids, and monocarboxylic acids arising from amination, Strecker synthesis, and oxidation, respectively (Table S5.4). Furthermore, the $\delta^{13}C$ values of branched, secondary, and α-CH_3 relatives of these straight-chain, primary, and α-H_3 compounds that our proposed network forms by aldehyde addition (where the model predictions for each compound span a range that considers the range of reaction yields spanned by our 'first and second cases' described above). This exercise leads to prediction of 74 independently measured compound specific $\delta^{13}C$ values; 55 % of those predictions agree with measurements within 2 standard deviation (based on the reported analytical errors alone), 66 % agree within


(a) ISM
ketone
aldehyde
straight-chain oxonitrile


(b)
Solar Nebula
branched-chain monocarboxylic acid
straight-chain monocarboxylic acid
straight-chain C2 amine
straight-chain C-1 amine
branched-chain C-1 amine
straight-chain α-CH3 amino acid
ketone
straight-chain aldehyde
branched-chain α-H amino acid
straight-chain α-H amino acid
straight-chain α-H hydroxy acid
straight-chain α-CH3 hydroxy acid
branched-chain α-H hydroxy acid
KIE -30‰ on COOH
KIE -10 to -30‰ on CH
KIE of -10 to -30‰ on CH2OH
+ CH2O, H+
KIE ~ -30‰ on CNH2
KIE -22‰ on COOH
EIE ~ +30‰ on NH3
+ NH3
EIE -20‰ on CHNH2
+ CN-
KIE ~ -30‰ on CH3NH2
EIE of +20‰ on NH
EIE -20‰ on CH3NH2
KIE -22‰ on CN
KIE -10 to -30‰ on α-COH
KIE of -10 to -30‰ on CH2OH
KIE -10 to -30‰ on α-CNH2
KIE of -10 to -30‰ on CH2OH
+ H2O
+ CN-
EIE -20‰ on CCH3NH2

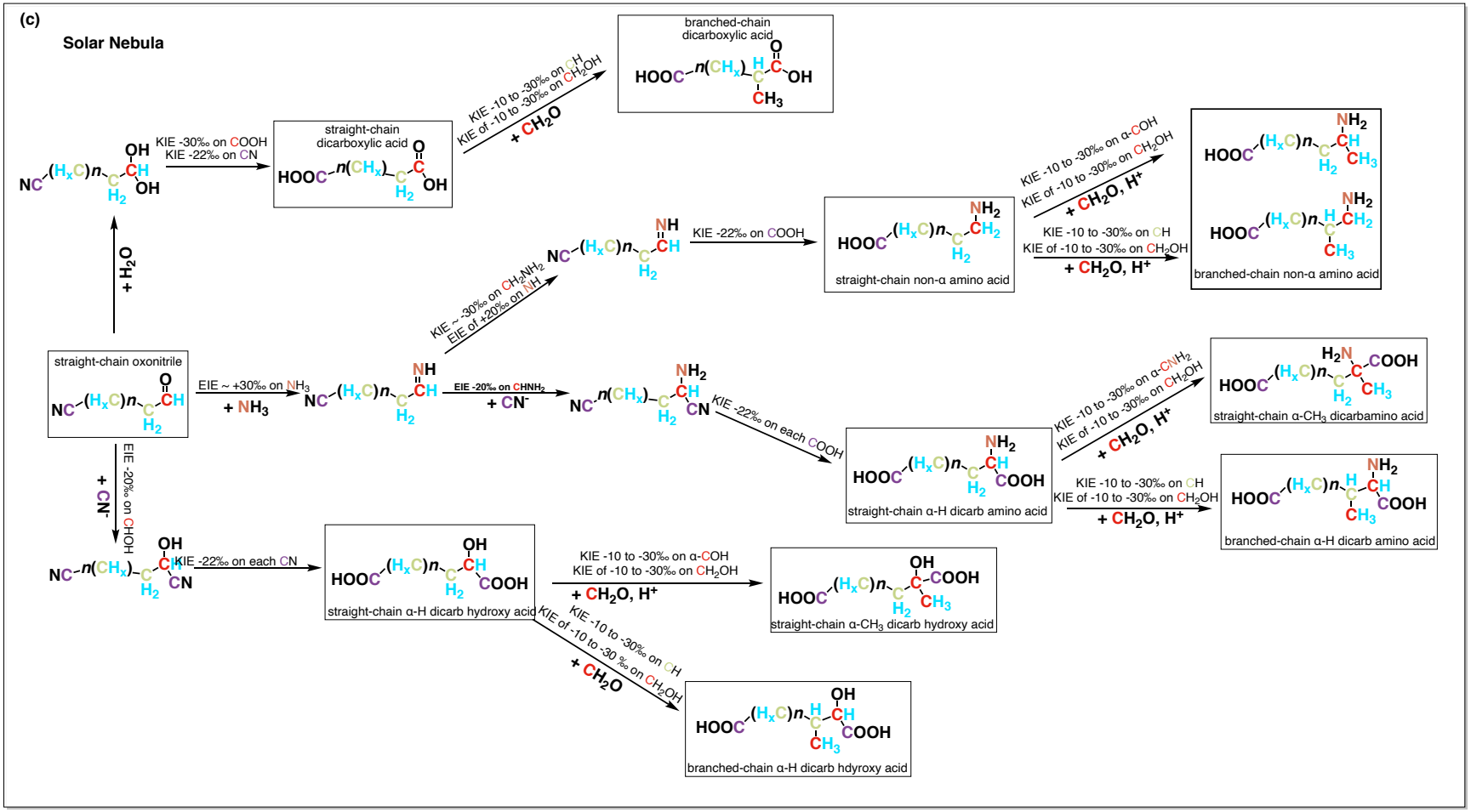


Figure 5.6: Model of carbon chemistry. Carbons and hydrogens are color-coded according to their inheritance: red carbons are from the CO pool in the ISM and are predicted to be ^{13}C-enriched, green carbons are from the CH_x pool in the ISM and are predicted to be ^{13}C-deplete, purple carbons are from CN and are predicted to be ^{13}C-deplete, blue hydrogens are from the ISM and are predicted to be D-enriched, and black hydrogens are from the meteorite parent body or are exchangable and predicted to be D-deplete. Ketones are included for completeness but they are assumed to have insignificant yields with respect to the pathway with solid arrows. Panel (a) depicts the chemistry that produces precursor compounds in the ISM, (b) depicts the chemistry in the meteorite parent body that produces monocarboxylic compounds and amines, and (c) depicts the chemsitrty in the meteorite parent body that produces dicarboxylic compounds.

	integrated aldehyde network	integrated aldehyde network + formaldehyde addition with aldehydes as precursors	integrated aldehyde network + COOH equilibration	FTT	Formose type
avg resid straight-chain set	6	6	9	11	23
avg resid branched-chain set	16	3	14	10	10
% accurate set	49%	66%	40%	30%	17%

Table 5.6: Residuals and accuracy for the different networks modeled for Murchison meteorite SOM for the 47 compounds that have $\delta^{13}C$ values predicted by all networks. A compound is considered accurate if a measurement ± 2 standard deviation was within the model's predicted range. We note that we do not include the standard deviations in the model predictions; however, we include the standard deviation propagated into the calculated predictions in Tables S5.2 and S5.4 to S5.9.

5 ‰, and the average residual (absolute value of prediction minus observation) is 6 ‰ (Table 5.7, Figure 5.7). The model reaction network is least successful at predicting the measured carbon isotope contents of MCAs and hydroxy acids (25% and 0 %, respectively, predicted within 2 standard deviations based on measurement precision alone; average residuals of (12 ‰ and 11 ‰, respectively; Table S5.4). However, elements of our proposed reaction network are more successful at describing variations among these compounds; specifically, differences in $\delta^{13}C$ between straight-chain and branched versions of otherwise similar MCAs and hydroxy acids are well described by our proposal of aldehyde addition. When the measured primary, straight-chain MCA analogs and α-H, straight-chain hydroxy acids and formaldehyde are used as precursors, the model predictions that fall within 2 standard deviations of measurement error increase from 14 % to 50 % for MCAs and 0 % to 100 % for hydroxy acids and the average residual falls from 18 ‰ to 3 ‰ and from 11 ‰ to 0 ‰ for branched MCAs and branched or α-CH_3 hydroxyl acids, respectively (Table 5.7, S5.4). Consequently, it is possible that straight-chain and α-H MCAs and hydroxy acids are synthesized entirely or in part through another mechanism and/or with an associated KIE (and/or experience a side reaction we have not considered) that results in several-per-mil differences from those predicted by our reaction network, but participate in the same aldehyde addition

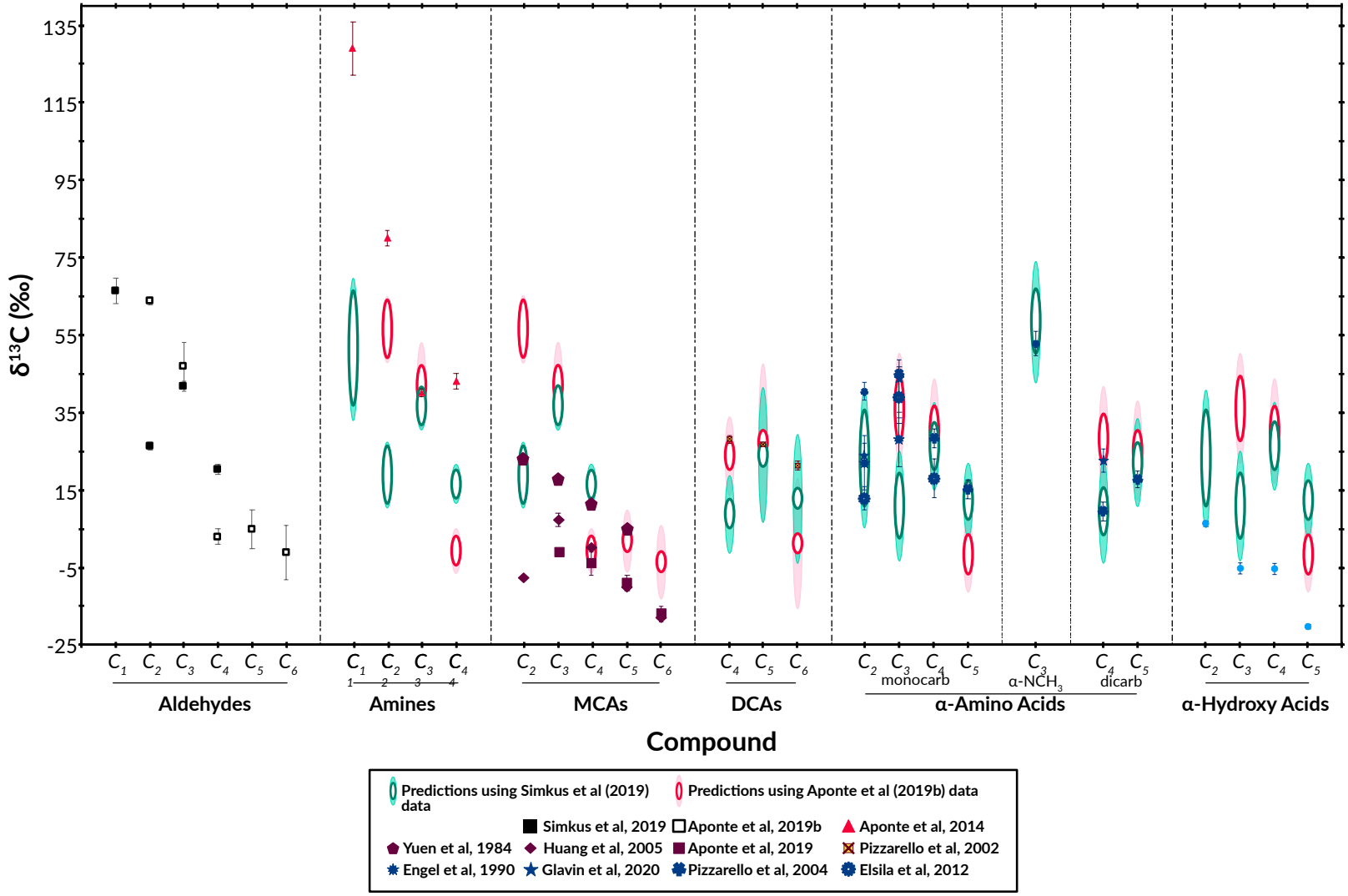
	integrated aldehyde network	integrated aldehyde network + formaldehyde addition with aldehydes as precursors	
		ald as precursors	analogs as precursors
compounds analyzed	60	73	69
% accurate	43%	58%	55%
% <5	53%	68%	74%
average resid	9	25	28

Table 5.7: Residuals and accuracy for the integrated aldehyde models. A compound is considered accurate if a measurement ± 2 standard deviations was within its predicted range. The standard deviation of the predicted values is calculated but not included in the model as in many cases it would increase the range of predicted values by 10 or 20 ‰. For formaldehyde addition, we consider two scenarios: (1) using aldehydes and applying isotope effects to predict the α-H and straight-chain precursor values and add formaldehyde to those ('ald as precursors') and (2) using the analogs α-H and straight-chain compound's $\delta^{13}C$ value as the $\delta^{13}C$ value of the backbone and adding a formaldehyde to that ('analogs as precursors').

chemistry that also successfully describes differences in $d^{13}C$ between the other straight-chain, primary, and α-H compounds (amino acids and amines) and their branched, secondary, and a-CH_3 relatives.

Data from other meteorites support the general usefulness of the proposed reaction network model as a framework for understanding the compound-specific carbon isotope geochemistry of SOM in the carbonaceous chondrites (Table 5.8). Specifically, when we seed the proposed reaction network with measured or estimated $\delta^{13}C$ values of aldehydes from each sample and then predict the $\delta^{13}C$ values of the other compounds that result from the various components of the integrated aldehyde chemistry and aldehyde addition, we predict $\delta^{13}C$ values of SOM compounds in CR and CM chondrites, with petrologic grades ranging from 1.1 to 2.5, that are, with one exception, as successful as our model predictions for the Murchison meteorite (Table 5.8 and Tables S5.5 – S5.10). The general

(a) Measurements and Predictions for α-H Straight-Chain Compounds from Murchison Meteorite



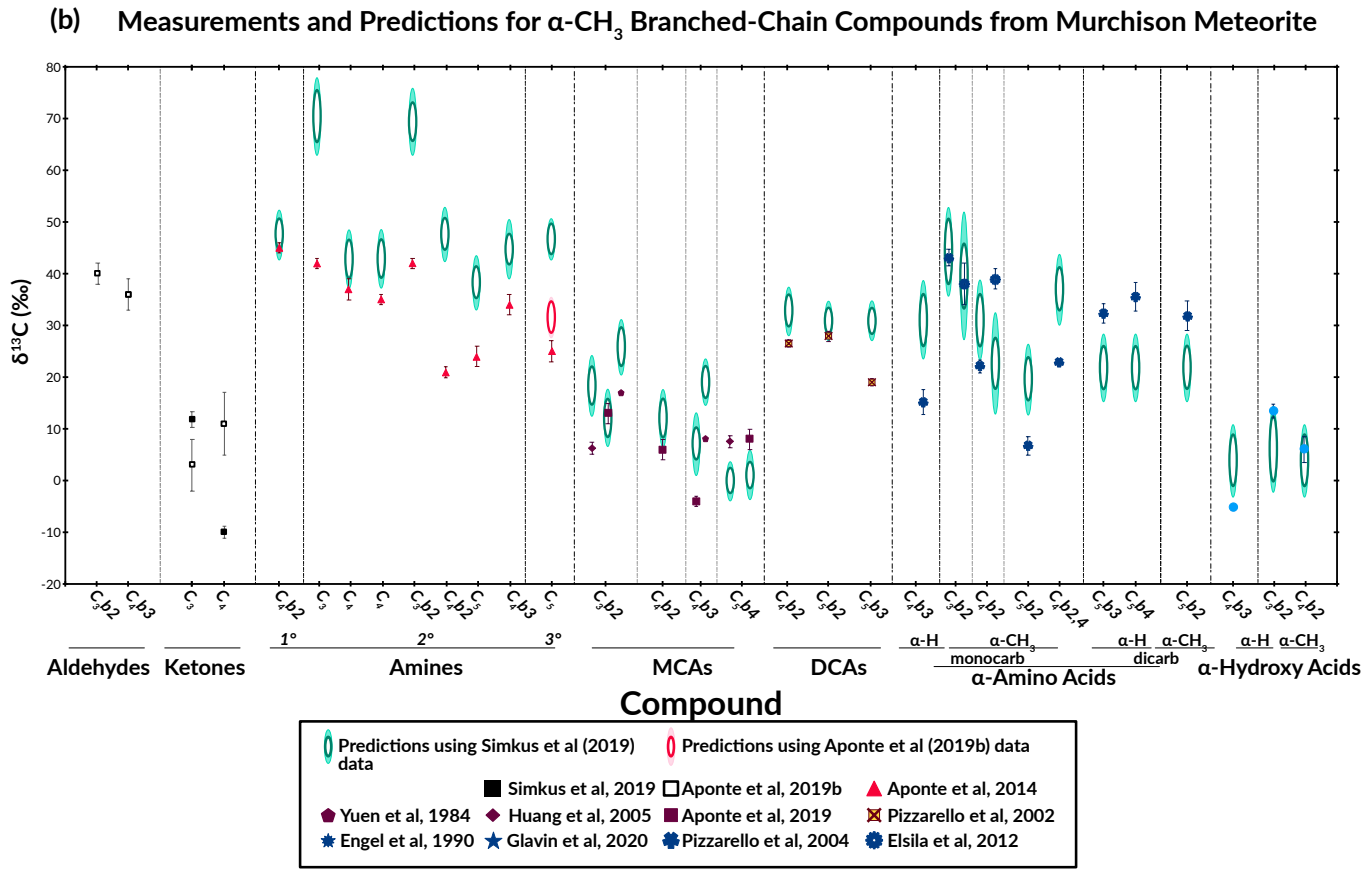


Figure 5.7: Measured molecularr-average $\delta^{13}C$ values and those predicted by the integrated aldehyde chemistry model followed by formaldehyde addition. The lighter colored oval behind the prediction is the error envelope for 1 standard deviation. Panel (a) depicts measurements and predictions for straight-chain, α-H compounds and C_n refers to the number of carbons in the compound. Panel (b) depicts measurements and prodictions for branched-chain and α-CH_3 compounds. The C_nbx notation refers to the number of carbons in the backbone (n) and the point at which a branch occurs (*x*). Amine are further separated by primary (1°), secondary (2°), and tertiary (3°). Labels are as follows: aldehydes: C_3b2 = isobutanal, C_4b3 = isopentanal; ketones: C_3 = acetone, C_4 = 2-butanone; amines: 1°: C_4b2 =2-methylbutylamine, 2°: C_3 = isopropylamine, C_4 = sec-butylamine (two isomers), C_3b2 = tert-butylamine, C_4b2 = tert-pentylamine, C_5 =sec-pentylamine, C_4b3 = 3-methylbutylamine, 3°: C_5 = 3-pentylamine; MCAs: C_3b2 = isobutanoic acid, C_4b2 = 2-methylbutanoic acid, C_4b3 = isopentanoic acid, C_5b4 =; DCAs: C_4b2 = 2-methyl succinic acid, C_5b2 = 2-methylglutaric acid, C_5b3 = 3-methylglutaric acid; amino acids: monocarboxylic: C_4b3 = valine, C_3b2 = 2-aminoisobutyric acid, C_4b2 = isovaline, C_5b2 = 2-methylnorvaline, $C_4b2,4$ = 2-amino-2,3-dimethylbutyric acid, dicarboxylic: C_5b3 = 3-methylglutamic acid, C_5b4 = 4-methylglutamic acid, C_5b2 =2-methylglutamic acid; hydroxy acids: monocarboxylic: C_4b3 = 2-hydroxy-3-methylbutyric acid, C_3b2 =2-hydroxy isobutyric acid, C_4b2 = 2-hydroxy-2-methylbutyric acid.

	Murchison (CM 1.6)	ALH 83100 (CM 1.1)	LEW 90500 (CM 1.6)	LON 94101 (CM 1.8)	MIL 090001 (CR 2.2)	LAP 02342 (CR 2.5)	GRA 95229 (CR 2.5)
compounds analyzed	82	13	17	17	7	15	21
% accurate	61%	31%	24%	29%	0%	47%	48%
% <5	72%	62%	24%	35%	0%	47%	48%
Avg resid	6	8	20	15	44	5	5

Table 5.8: Comparison of accuracy for integrated aldehyde network followed by aldehyde addition for seven meteorites. The models' fits improve with the integrated aldehyde network used for straight-chain and α-H compounds and use the measured straight-chain and α-H compounds and formaldehyde to predict values for the branched-chain and α-CH_3 compounds.

sense of disagreements between model predictions and data for this exception, sample MIL 090001, arise from the $\delta^{13}C$ values of the aldehyde and ketones measured on MIL 09000—predominantly below -50 ‰—which are both low relative to other meteorites and to the other SOM in MIL 090001 for which there only 10 measurements (Table S5.8). This implies that exceptionally intense aqueous alteration after organic synthesis leads to either the breakdown of material into straight-chain aldehydes and ketones or to isotopic exchange with a low $\delta^{13}C$ carbon pool.

One question not addressed by our evaluation of the model so far is the factors controlling the $\delta^{13}C$ values of aldehydes that are treated as precursors of most other compounds. Based on site-specific $\delta^{13}C$ measurements of an alanine sample from Murchison meteorite, Chimiak et al. (2020) propose that aldehydes are formed from ISM chemistry in which ^{13}C-enriched CO ($\delta^{13}C$ values of 166 ‰) react with ^{13}C-poor ($\delta^{13}C$ values of -36 ‰) reduced carbon. The model predicts all but two linear aldehyde values within one measurement error—these two values are that for formaldehyde and for one of two acetaldehyde $\delta^{13}C$ measurements. Therefore, we suggest that a chemistry similar to that in Chimiak et al. (2020) creates the aldehydes used here for aldehydes of acetaldehyde chain-length and above and that a parent-body altered the formaldehyde reservoir that is used in formaldehyde addition.

Taken as a whole, the reaction network we propose is capable of describing the $\delta^{13}C$ values of 95 SOM compounds in 7 carbonaceous chondrite samples with an average

residual between model and data of 10 ‰ (Tables 5.2, 5.8, and S5.4 – S5.10). This average residual is several times the reported analytical errors typical of the studies on which we have drawn but is a small fraction of the observed range in $\delta^{13}C$ we are trying to explain (162 ‰, Table S5.4), and is similar to the variations that might be expected to arise from modest changes in temperature or other reaction conditions or differences between real and assumed KIE's. The residual also is similar to the standard deviation found in aldehyde and ketone $\delta^{13}C$ measurements. For this reason, we consider the model presented to be successful, with the possible exceptions that it systematically over-predicts the $\delta^{13}C$ values of the linear MCA's and hydroxy acids and does not adequately describe data for the most extensively aqueously altered samples.

Alternative models that did not predict the data as well include the FTT and Formose reaction networks. Both networks overpredicted almost all compounds and had residuals of 10 ‰ or more (Table 5.6, S5.4). The aldehyde reaction network in which ketones and branched aldehydes were used also tended to underestimate the $\delta^{13}C$ values of the α-CH_3 and branched-chain compounds and thee underpredictions remained for amines and carboxylic acids when the carboxyl group of amino and hydroxy acids were modeled to equilibrate with the carbonate pool.

5.4. Discussion

5.4.1. The chemistry of the model meteoritic SOM reaction network

The model reaction network we propose can be viewed as simply a pattern of inheritance, modified by reaction-specific isotope effects, that provides a parsimonious explanation of the $\delta^{13}C$ values of meteoritic organic compounds. The moiety-specific hydrogen isotope model lacks the same implied family of connections between compounds of different chemical types, but potentially contributes to our understanding of the origins of recurring chemical 'units' that appear in the compounds of the reaction network. In this section, we examine the relationships between the findings of these models and our

understanding of the possible ranges of environments, precursors and reactions that might have led to synthesis of meteoritic organics.

In the ISM, surface- and gas-phase chemistry cause concentrate deuterium into molecules by an extent that scales with the number of hydrogen sites on a molecule (Sandford et al., 2001; Rodgers and Charnley, 2002; Charnley et al., 2004). This phenomenon is due to the 10-40K temperatures of the ISM and the lower zero-point-energy from deuterium bonds relative to hydrogen. Models of gas phase chemistry predict deuterium sequestration in heavily substituted species such as D_2CO, D_3^+, and CD_3^+ (Osamura et al., 2005; Millar, 2005; Roueff et al., 2005). Grain surface chemistry can also impart large δD enrichments and multiple deuterium-substitutions as explained in Rodgers et al. (2002). They explain that when deuterium absorbs on ice it is less likely than hydrogen to be desorbed: hydrogen can tunnel through the ice surface increasing its probability of bonding with another hydrogen, an exothermic reaction that releases the H_2 molecule. With this elevated deuteration, molecules formed on grains will have deuterium enrichments proportional to its number of hydrogen sites. Furthermore, when molecules such as formaldehyde form on ice grains, hydrogen is preferentially removed and thus enriches the compounds' singly to fully deuterated forms (HDCO and D_2CO, respectively, for formaldehyde). This theory leads to two predictions: (i) sites that have more hydrogens (*e.g.*, methyl) will have higher δD values than those with fewer hydrogens (*e.g.*, those on a resonant ring), and (ii) sites derived from formaldehyde or similarly formed compounds will have higher molecular average δD values (and enrichments in multiply-deuterated isotopologues) relative to compounds such as PAHs.

Existing ISM chemical models also predict ^{13}C isotopic enrichments at specific sites in organic compounds (Charnley et al., 2004). In the ISM, ^{13}C preferentially enters the CO pool due to the reaction $^{12}CO + {}^{13}C^+ \rightleftharpoons {}^{13}CO + {}^{12}C^+$. This reaction is exothermic in the forward direction by 35 K, so ^{13}C cannot leave the CO pool without an energy input greater than the ISM's typical environmental temperature. Because CO is a major species in the ISM, its probability of interacting with C^+ is higher than that of most other carbon species. Therefore, this reaction can sequester a large proportion of available ^{13}C atoms in

the CO reservoir. Sites that form from CO such as carbonyl sites are predicted to conserve this ^{13}C-enhancement while those formed from other carbon pools such as aliphatic chains predicted to be relatively ^{13}C-depleted.

Meteorite parent bodies can concentrate and further react compounds formed in the ISM. One common premise is that aldehyde chemical networks create amines, MCAs, hydroxy acids, and α-amino acids (Aponte et al., 2017). Another suggests that MCAs are partially formed from the breakdown of PAHs and/or from a kinetically-controlled carbon-addition process (Yuen et al., 1984; Huang et al., 2005; Aponte et al., 2011).

On meteorite parent bodies, the abundances and isotopic compositions of products and reactants are impacted by physiochemical conditions and reaction rates of the various chemical processes (Glavin et al., 2010; Elsila et al., 2016; Aponte et al., 2016; Aponte et al., 2019). For instance, the concentration of ammonia and therefore the pH determines the proportions of α-amino and α-hydroxy acids produced by Strecker synthesis (Miller and Van Trump, 1981; Pizzarello et al., 2010) and the proportions of amines and alcohols produced by reductive amination (Gomez et al., 2002). Higher pHs lead to higher ammonia concentrations and favor the aminated species. Furthermore, the faster reaction rates for aldehydes relative to ketones in reductive amination (Gomez et al., 2002; Lamm et al., 2013) and in Strecker synthesis (Lerner et al., 1991) could lead to increased C-1 amines and α-H amino acids.

Once synthesized, compounds can undergo further alteration. Here we propose that an important process is the reaction of formaldehyde with double bonds and positively charged carbon radicals. Subsequent reduction of the formaldehyde would result in methylated compounds. This type of chemistry occurs in Formose-type reactions, which has been hypothesized for create sugars on meteorites (Cooper et al., 2011; Kebukawa et al., 2017).

Aldehyde addition is arguably the most problematic conjectured chemical reaction contained in our proposed model reaction network: while it is known that such reactions

can occur with alkenes and radicals as substrates, we are not aware of any evidence that it can lead to methylation of amino acids, amines, and other species we have suggested reacting with aldehydes. So, although the suggestion of aldehyde addition is a parsimonious and highly successful explanation for the wide range of isotopic observations that motivate our proposed model, it obviously cannot be correct unless the required chemistry (or something leading to the same net effect) occurs. We suggest two potential avenues that could enable the aldehyde additions called on by our model. The first potential route is through mineral-mediated carbon addition similar to that in FTT synthesis, or radical formation due to hydrogen loss during aqueous alteration (perhaps also surface catalyzed). Classic FTT synthesis involves CO chain elongation as C-H has a far higher bond energy; however some catalysts enable C-H bond activation by the creation of C-M organometallic bonds, and these enable additions to alkyl carbons at temperatures down to 25 °C (Vidal et al., 1997; Yeung and Dong, 2011). If a similar process occurs on meteorite parent bodies, it could enable the addition of formaldehyde on the alkyl backbones of SOM compounds. Secondly, Lerner (1995) observes that in the presence of crushed Murchison meteorite, deuterium exchange increases on α-H amino acids, including enablement of the exchange of alkyl hydrogens. If a similar mechanism enables the addition of formaldehyde rather than hydrogen on the weakened bond, it could result in formaldehyde addition. Given that the on α-H bond is weaker than the alkyl-H bond, we would anticipate that this process could generate a higher proportion of α-CH_3 compounds than branched compounds. In either case, an implication of our model is that the chemistry of aldehyde – SOM mixtures in the presence of surface catalysts should be explored.

Additional alternatives that might result in a successful (though different) model of carbon isotopic variations is to call for methylation of the various straight-chain precursors through another mechanism and calling on another substrate, which is itself also the dominant source of formaldehyde in the reaction network (*i.e.*, so that formaldehyde is a side product that coincidentally matches the carbon isotope composition required to explain the straight-chain/branched differences in $\delta^{13}C$). The principal disadvantage of this scheme (and the reason it was not presented here as our

proposed model) is that it is difficult to envision how the resulting methyl groups would retain the extreme deuteration required in our interpretation of the D/H variations of SOM compounds and IOM moieties. Finally, we consider the potential that all chemistry to create compounds occurs in the ISM as this would enable carbon-activation through photolytic cleavage of the C-H bond. As the only site-specific $\delta^{13}C$ measurements suggests that alanine—and from the model other SOM—forms from aqueous processes (*i.e.,* the integrated aldehyde network) (Chimiak et al., 2020), we do not further explore this option. As noted in Chimiak et al. (2020), experimentally verified mechanistic models of ISM chemistry that would lead to site-specific isotopic predictions are limited, so future experiments could find a pathway in which ISM chemistry could produce the results found in Chimiak et al. (2020) and therefore be more pertinent to the model explored here.

5.4.2. Hydrogen

Our δD moiety model indicates two distinct pools of hydrogen in OM: (1) hydrogen attached to methyl carbons ('methyl-H') with high δD enrichments of up to 3000 ‰ in Murchison (and up to 5000 ‰ in CR 2 chondrites) and (2) other hydrogens on chains, resonant rings, and attached to α-CNH_2 ('chain-H', 'ring-H' and 'α-H') with typically about half the D/H ratio of the methyl-H pool (Table 5.5, Figure 5.3). (We note that non-α-H enrichments vary between compound classes potentially due to having mixed origins). As can be anticipated from this result, a compound's molecular-average δD scales with its ratio of methyl-H to total hydrogen, and likely results in the number of methyl moieties strongly influences a compound's δD (Figure 5.1).

The degree of fitted D enrichment at the methyl-H moiety differs between compound classes and decreases with increasing solubility in water (Figure 5.3c). These findings suggest that methyl-H precursors' hydrogen is predominantly inherited from the ISM and after compounds are synthesized, the methyl-H exchanges with the surrounding water. This exchangeability proposition is bolstered by the fact that the difference in fitted D/H ratio between methyl-H moieties and other moieties (the 'methyl normalized δD value;

Figure 5.3b, Table 5.5) is greatest for the a-H moieties (*i.e.*, these are lowest in δD, when compared to methyl-H in the compound class in Murchison), which are the most acidic and so the most exchangeable. Finally, the high relative fitted δD of the methyl-H relative to that of the ring-H supports exchange in an aqueous environment. When considering homolytic cleavages (*i.e.*, breaking a bond in which each atom will acquire one of the shared electrons and lead to radicals), hydrogen bonded to resonant rings has a higher bonding energy than hydrogen bonded to a methyl carbon, so we would anticipate that the hydrogen on a ring would be less prone to exchange. In the case of heterolytic cleavages, such as we would expect in an aqueous environment, the hydrogen attached to methyl is less acidic and therefore less likely to exchange than hydrogen attached to a resonant ring. In this case, the methyl hydrogen would be les exchangeable, resulting in the relatively higher δD values that the model fitting produces.

On Murchison, therefore we propose that SOM and IOM are initially imprinted with ISM-sourced hydrogen with high D enrichments (equal to or greater than the 5430 ‰ δD maximum found by our moiety-specific δD model fits), and this high deuteration is diluted via exchange with low δD water on the parent body during subsequent alteration and metamorphism (and perhaps also terrestrial weathering). This analysis agrees with that of Alexander et al. (2012) with respect to the deuterium content of primordial IOM and to meteoritic water values. The highest measured values for methyl-H and non-α-H in IOM for the meteorites studied here are 4388 ‰ and 5430 ‰, respectively (EET 92042, Table 5.5). These values are above the estimated 3500 ‰ from Alexander et al. (2012); however meteorite EET 92042 is excluded from the analyses that estimate this number (Alexander et al., 2017). Barring this sample, the highest fitted δD values for methyl-H and non-α-H in IOM values are 2063 ‰ (Orgueil) and 5216 ‰ (Murchison), respectively (Table 5). Interestingly, the higher value is similar to those fitted and measured for EET 92042 (Alexander et al., 2017), which is amongst the least altered CR samples (CR 2.5, Table 5.2). When using the same value that Alexander et al. (2017)These values are above Our lowest physically plausible (*i.e.*, greater than -1000 ‰) moiety-fit δD values are -440 ‰ for CM chondrites (Murray) and -107 ‰ for CR chondrites (EET 87770). The first value nearly matches the predicted 440 ± 23 ‰

δD value of water in CM chondrites and is close to the predicted 96^{+110}_{-65} ‰ δD value for water on CR chondrites (Alexander et al., 2012). Site-specific and clumped isotope ratio measurements could confirm this scenario if it demonstrates an inverse trend between a moieties' acidity and its δD value.

Differences between the fitted δD values of the moiety-specific model vary between meteorite classes, likely reflecting differences in parent body composition, processing, and alteration temperature. For amino acids, all CR chondrites studied have higher methyl-H and non-α-H δD values than those fitted for CM chondrites (Figure 5.4 and 5.5). This difference could point to differences in the CR and CM chondrite parent body: CR 2.4 to 2.7 chondrites studied have 2-3 orders of magnitude higher amino acid concentrations than do the CM1.1 to 1.6 chondrites (Pizzarello et al., 2010; Elsila et al., 2012; Elsila et al., 2016), so a lower amount of processing or degradation of the initial inventory of amino acids on the CR parent body could explain their relatively high methyl-H δD enrichments (*i.e.*, lesser aqueous alteration leads to less exchange with parent body water). Furthermore, we can compare moiety δD predictions for amino acids in Murray and Murchison meteorites, which are CM 1.5 and CM 1.6 types, respectively, with those in a more pristine sample, GRA 95229 (CR 2.5). Our analyses predict roughly the same absolute δD moiety values for Murchison and Murray samples and higher absolute δD values for each moiety in GRA 95229, and it predicts similar methyl-normalized δD values for the chain-H and α-H moieties in all samples (Figure 5.4a). This comparison supports the idea that differences in absolute δD values are controlled by parent body composition and degree of alteration/weathering, which agrees with past analyses (Elsila et al., 2012).

5.4.3. Carbon Data

Our proposed reaction network model of SOM synthesis on the carbonaceous chondrites—crafted to provide the most parsimonious explanation of measured compound-specific $\delta^{13}C$ values—contains the following essential features: (1) straight-chain aldehydes and oxonitriles undergo oxidation, reductive amination, and Strecker

synthesis to form α-H and straight-chain MCAs, amines, and α-amino acids, respectively, (2) these compounds then undergo aldehyde addition, possibly during hydrothermal alteration or earlier in mineral-mediated reactions, and (3) bonded formaldehyde is reduced to form methyl groups (Figure 5.6). We think it most chemically plausible that the first series of reactions (the 'integrated aldehyde chemistry') likely occurs as two steps: first, an oxidative step in an environment with a low ammonia-to-aldehyde ratio, and second a reductive amination and Strecker synthesis step in an environment with a high ammonia-to-aldehyde ratio. The different environments could be achieved by a rise in pH over time: a higher pH will release ammonia from its more unreactive form ammonium. Other reactions such as deamination or decarboxylation likely occur after the listed steps, but these reactions are beyond the scope of this work as we have relatively little information about the abundances and isotopic properties of their products.

The network we propose was designed through a process of iterative hypotheses and comparisons of predicted and measured compound-specific $\delta^{13}C$ values, where we settled on a preferred version based on its relative low average residuals and high percentage of accurate predictions for data from 3 CM and 2 CR chondrites (some alternate variants we considered are summarized in the Supplementary Information). Data used in this study were collected over thirty years using diverse methods and from various different samples of the meteorites, and it seems likely to us that SOM synthesis actually occurred over a range of conditions that might have caused KIE's to vary or differ from those we assumed. Therefore, even an accurate reaction network of the kind we propose might fail to reproduce measured compound-specific $\delta^{13}C$ values to within their nominal analytical errors. For these reasons, we consider the ability of this model to match 55 % of the measured $\delta^{13}C$ values for SOM compounds from five different carbonaceous chondrites to within 2 standard deviation of measurement (and not including the expanded range in predictions that results from including the error of the predicted values, which vastly increases its percentage of fit data) and achieves an average residual of only 6 ‰ to be demonstrations of its successfulness as a comprehensive hypothesis (Tables 5.6, 5.8, and S5.4).

The network proposed here potentially answers questions raised in prior studies. Firstly, the network explains why branched-chain and α-CH_3 compounds are observed to be systematically higher in $\delta^{13}C$ than straight-chain and α-H analogs. Previously, it has been suggested that ^{13}C-enriched ketone precursors (Pizzarello et al., 2004; Simkus et al., 2019) or that the post-synthetic equilibration of the carboxyl group with a ^{13}C-enriched carbonate could lead to the enrichments observed in α-CH_3 compounds (Chimiak et al., 2020). However, carbon isotope measurements of meteoritic ketones reveal them to be lower in $\delta^{13}C$ than aldehydes (Simkus et al., 2019; José C. Aponte et al., 2019), and we do not consider it chemically plausible that carbonate equilibration could explain the relative $\delta^{13}C$ enrichments on branched-chain compounds or on compounds other than amino or hydroxy acids. In the proposed network, formaldehyde addition creates both branched and α-CH_3 compounds. Formaldehyde is amongst the most ^{13}C-enriched compounds measured in meteorites (Simkus et al., 2019), so formaldehyde-derived methyl groups will increase in molecular-average $\delta^{13}C$ from the straight and α-H compounds. The proposed scenario does not limit formaldehyde addition to a compound class and successfully explains differences in $\delta^{13}C$ between straight- and branched-chain compounds of all classes, even in cases where the linear-compound $\delta^{13}C$ values are anomalous compared to compounds of similar size in different classes; specifically, the aldehyde addition hypothesis describes the $\delta^{13}C$ differences between almost 100% of pairs of straight- and branched-chain MCA and hydroxy acids, despite the fact that the straight-chain species in question are not well matched by the 'integrated aldehyde chemistry' part of our network. And, although methyl groups forming (directly or indirectly) from CO_2 or carbonate might result in similar final $\delta^{13}C$ values (because carbonate in the carbonaceous chondrites is relatively ^{13}C rich), the addition of carbon from an oxidized pool and subsequent reduction would be expected to result in low D/H ratios of the resulting methyl-H moieties rather than the high ratios fit by the moiety-specific D/H model. This doesn't constitute a disproof of the carbonate or CO_2 addition hypothesis because the moiety-specific D/H model is a generally successful hypothesis rather than an independently known fact. Nevertheless, a parsimonious interpretation of combined H and C isotope geochemistry of meteoritic SOM compounds along the lines

developed in this paper favors formaldehyde as the ultimate source of the carbons in question.

In addition to isotopic considerations, creating branched and α-CH_3 compounds via formaldehyde addition potentially helps understand abundance patterns with respect to α-H straight- and branched-chain compounds. In spite of the slower reaction rates for ketones in Strecker synthesis (Lerner et al., 1991) and of the low ketone-to-aldehyde ratio in Murchison meteorite (Simkus et al., 2019; Aponte et al., 2019), α-CH_3 amino acids are typically more or close to as abundant as α-H amino acids of the same chain length (Kerridge, 1999; Glavin et al., 2010; Elsila et al., 2016). If formaldehyde addition is the predominant route to α-CH_3 and branched compound formation, we would anticipate that α-CH_3 compounds would be produced at a faster rate and therefore be more abundant than branched-chain ones.

The reaction network we put forth also predicts that oxonitriles are the precursors for dicarboxylic acids and dicarboxylic amino acids. Using oxonitriles as precursors, we are able to explain the measured ^{13}C values of 10 of the 11 dicarboxylic within 2 standard errors and the remaining compound had an error of 4 ‰, a modest discrepancy (Table 5.9). We can test this proposal through future site-specific $\delta^{13}C$ measurements. In the outlined mechanism, both carboxyl carbons in α-amino acids will have cyanide precursors and undergo hydrolysis so both resulting carboxyl sites should also have the same $\delta^{13}C$ values. Based on measurements in Chimiak *et al* (2020), these $\delta^{13}C$ values will be within error the carbon chain and ~150 ‰ lower than the C-2 carbon. Conversely, in DCAs, we would anticipate one carboxyl site to be formed form CN and one to be formed from the carbonyl group in an oxonitrile originally from the CO pool (Figure 5.6c), so a DCA would have two distinct carboxyl sites that differ from each other by over 100 ‰ and differ on average from chain carbon's in the same compound by 50 ‰ or more.

The oxonitrile pathway for dicarboxylic species also provides insight into a potential route to synthesize some non-α-amino acids via amine addition and subsequent reductive

						Oxonitrile							Dialdehyde						
						Simkus et al 2019 data			*Aponte et al 2019b data*			**Resid**	*Simkus et al 2019 data*			*Aponte et al 2019b data*			**Resid**
			C #	**$\delta^{13}C$**	*st err*	high	low	*st err*	high	low	*st err*		high	low	*st err*	high	low	*st err*	
DCA		*ref*	*Pizzarello et al, 2002*																
	straight	Succinic	4	**28.1**	*1*	16	8	*5*	35	27	*5*	**0**	31	24	*3*	50	42	*3*	**0**
		Glutaric	5	**26.8**	*0*	27	21	*4*	30	24	*5*	**0**	39	33	*2*	42	36	*4*	**6**
		Adipic	6	**21.4**	*1*	15	10	*3*	4	-1	*4*	**4**	25	20	*2*	14	9	*2*	**0**
	branch	2-methylsuccinic	5	**26.5**	*1*	26	20	*4*	41	35	*4*	**0**	38	32	*2*	53	47	*2*	**4**
		2-methylglutaric	6	**27.9**	*1*	27	22	*3*	36	31	*4*	**0**	44	39	*2*	46	41	*3*	**9**
		3-methylglutaric	6	**19.1**	*1*	27	22	*3*	36	31	*4*	**1**	44	39	*2*	46	41	*3*	**18**
AA	**Dicarb α-H**	*ref*	*Pizzarello et al, 2004*																
		Aspartic acid	4	**10**	*2*	16	3	*5*	32	20	*5*	**0**	31	19	*3*	49	36	*3*	**4**
		Glutamic acid	5	**18**	*2*	27	17	*4*	28	18	*5*	**0**	39	29	*2*	41	31	*4*	**7**
	Dicarb α-CH_3	3-methylglutaminic	6	**32**	*2*	34	25	*3*	36	28	*3*	**0**	44	36	*2*	46	38	*3*	**-1**
		4-methylglutaminic	6	**36**	*3*	34	25	*3*	36	28	*3*	**0**	44	36	*2*	46	38	*3*	**0**
		2-Methylglutamic	6	**32**	*3*	34	25	*3*	36	28	*3*	**0**	44	36	*2*	46	38	*3*	**-2**
analysis			total compounds									**11**	total compounds						**11**
			% accurate									**82%**	% accurate						**27%**
			average residual									**0.50**	average residual						**4.22**

Table 5.9: Comparison of oxonitrile and dialdehyde precursor predictions for dicarboxylic compounds in the Murchison meteorite.

amination of the iminonitrile species (Figure 6c). Applying the same network to predict these non-α-amino acids, we find that 4 of 7 compounds (57%) are predicted to within 2 standard errors based on analytical uncertainties alone and the average residual for the other three compounds is only 4 ‰ (Table S4).

5.4.4. Possible inclusion of nitrogen isotope data in the hypothesized reaction network

There are insufficient nitrogen isotopic data for SOM compounds to include nitrogen isotopes in our proposed reaction network. We are only aware of usefully precise compound specific $\delta^{15}N$ measurements of meteoritic SOM compounds for ammonia and amino acids. Amino acids in those studies are ^{15}N-enriched relative to the ammonia, as would be predicted by Strecker synthesis (Pizzarello et al., 1994; Pizzarello and Holmes, 2009; Elsila et al., 2012); however, these ^{15}N enrichments span a 100 ‰ range. The equilibrium between amine and imine in Strecker synthesis has an equilibrium isotope effect that enrich the imine (*i.e.*, the nitrogen that is inherited by amino acids) by 56 ‰ relative to the ammonia (See Chapter 3 for a detailed analysis), so secondary chemistries likely impact the final $\delta^{15}N$ of amino acids. One can imagine that changes in the pH would change the $\delta^{15}N$ of ammonia by up to 30‰ (Walters et al., 2019), the approximate value of the EIE between NH_4^+ (*aq*) and NH_3 (*aq*) at relevant temperatures. Furthermore, as the EIE between the α-aminonitrile's amine nitrogen and ammonia is on the order of 50 ‰, amino acids that react first will deplete the remaining ammonia's $\delta^{15}N$ by tens of ‰, potentially leading to large variations in $\delta^{15}N$ of products over the course of reaction. In principle, a detailed understanding of these variations might lead to a more nuanced understanding of the overall SOM reaction network, as it could allow to constrain the reaction progress responsible for creating some key product compounds. However, further study of the N isotope system will be required. Specifically, understanding the $\delta^{15}N$ values of amines and comparing them with the amino acids will further constrain their common precursors, their reaction order, and the pH on the meteorite parent body.

5.5. Summary and Conclusions

We have compiled and re-interpreted hydrogen and carbon data from 30 years of meteorite analysis and 4 meteorites classes (CM, CR, CI, and CV) that span petrologic

types 1.1 to 3.0. Based on these data, we created a moiety-specific model of H isotope variations and an integrated synthesis network designed to explain C isotope variations. For deuterium, we notice two main pools: (1) the methyl-H and non-α-H moieties, which are the enriched by up to 5000‰, and (2) the chain-H, ring-H, and α-H moieties, which are systematically lower in D/H by about a factor of 2 as compared with methyl-H moieties in the same compounds. The fitted δD values of methyl-H moieties decrease with increasing water solubility of each compound class and with increasing thermal and aqueous alteration of each meteorite. These predictions are consistent with the interpretation that all nominally non-exchangeable hydrogens in meteoritic SOM and select IOM compounds were D-enriched immediately after synthesis, perhaps due to inheritance from D-rich precursors synthesized in the ISM, and later were reduced in D/H ratio during aqueous alteration, with the extent of D depletion rising with the acidity (and thus exchangeability) of each H site; the solubility in water of the parent compound, and the extent of aqueous and thermal alteration of the parent body. Because D/H ratios of these compounds largely reflect the abundance of relatively refractory methyl-H moieties and the extent of post-formation D/H exchange with water, they likely cannot be interpreted in a straightforward way as constraints on the specific chemistry responsible for their synthesis. For that reason, we turn to the compound specific ^{13}C contents of SOM compounds to reconstruct their network of formation reactions.

The model reaction network we put forth to explain carbon isotope variations affirms and builds on a previously proposed model in which aldehydes react to form carboxylic acids, amines, and α-amino acids (the 'integrated aldehyde chemistry'). We go on to propose that branched and α-CH_3 species are predominantly synthesized through the addition of formaldehyde to straight-chain α-H compounds. The result is a model that can describes the $\delta^{13}C$ enrichments and the relative abundances between branched and/or α-CH_3 compounds and straight-chain α-H ones.

The moiety-based deuterium model provides a series of testable predictions both at the site-specific and clumped levels. Firstly, assuming this model is correct, we would

anticipate large (3000 to 5000 ‰) site-specific δD values at the methyl-H and the aromatic IOM's non-α-H in sites Murchison relative to the other sites. Secondly, we would anticipate that for a compound, this value would increase in δD between meteorites as they decreased in degree of aqueous alteration until reaching a petrologic type of 3.0. Thirdly, if the methyl moieties are synthesized on ice grains in the ISM, we would anticipate a high degree of multiple deuterations at a site (Rodgers and Charnley, 2002).

The formaldehyde addition chemistry proposed in the carbon model also provides specific δD predictions for the methyl sites that arise from formaldehyde addition that would be distinct from methyl sites on α-H and straight-chain compounds. In addition to high $\delta^{13}C$ values (~60 ‰), methyl groups produced from the formaldehyde chemistry proposed above will have extremely proportions of double-deuteration relative to a stochastic distribution but low proportions of triple-deuteration because two hydrogens would be inherited from the ISM and the final hydrogen would be from water on the parent body. Conversely, methyl groups that are not produced from formaldehyde chemistry (*i.e.*, the methyl groups on straight-chain, α-H compounds) would have relatively low $\delta^{13}C$ values and high proportions of doubly- and triply-deuterated methyl groups relative to stochastic distributions as the entire group would form in the ISM (Rodgers and Charnley, 2002; Charnley et al., 2004).

The reaction network proposed here also provides falsifiable predictions at both a molecular-average and site-specific level for $\delta^{13}C$. Key predictions include: (1) α-C sites are heavily ^{13}C-enriched relative to other non-branched sites in α-amino and α-hydroxy acids, (2) substituted carbons in amines and monocarboxylic acids are heavily ^{13}C-enriched relative to carbons in their aliphatic chain, (3) dicarboxylic acids have one ^{13}C-enriched and one ^{13}C-depleted carboxyl group, and (4) α-CH_3 and single-carbon branches are ^{13}C-enriched relative to carbons in the main linear carbon chain.

BIBLIOGRAPHY

Abelson P. H. and Hoering T. C. (1961) Carbon isotope fractionation in formation of amino acids by photosynthetic organisms. *Proc. Natl. Acad. Sci. USA* **47**, 623–632.

Alexander C. M. O., Bowden R., Fogel M. L. and Howard K. T. (2015) Carbonate abundances and isotopic compositions in chondrites. *Meteorit. Planet. Sci.* **50**, 810–833.

Alexander C. M. O., Bowden R., Fogel M. L., Howard K. T., Herd C. D. K. and Nittler L. R. (2012) The provenances of asteroids, and their contributions to the volatile inventories of the terrestrial planets. *Science* **337**, 721–723.

Alexander C. M. O., Cody G. D., De Gregorio B. T., Nittler L. R. and Stroud R. M. (2017) The nature, origin and modification of insoluble organic matter in chondrites, the possibly interstellar source of Earth's C and N. *Chem. Erde* **77**, 227–256.

Alexander C. M. O., Fogel M., Yabuta H. and Cody G. D. (2007) The origin and evolution of chondrites recorded in the elemental and isotopic compositions of their macromolecular organic matter. *Geochim. Cosmochim. Acta* **71**, 4380–4403.

Alexander C. M. O., Howard K. T., Bowden R. and Fogel M. L. (2013) The classification of CM and CR chondrites using bulk H, C and N abundances and isotopic compositions. *Geochim. Cosmochim. Acta* **123**, 244–260.

Anon Antarctic Meteorite Collection. Available at: https://curator.jsc.nasa.gov/antmet/ [Accessed October 20, 2020].

Aponte J. C., Abreu N., Keller L., Elsila J. E. and Dworkin J. P. (2018) Soluble Amines in Anomalous CR Chondrite Miller Range (MIL) 090001.

Aponte J. C., Alexandre M. R., Wang Y., Brearley A. J., Alexander C. M. O. and Huang Y. (2011) Effects of secondary alteration on the composition of free and IOM-derived monocarboxylic acids in carbonaceous chondrites. *Geochim. Cosmochim. Acta* **75**, 2309–2323.

Aponte J. C., Dworkin J. P. and Elsila J. E. (2014) Assessing the origins of aliphatic amines in the Murchison meteorite from their compound-specific carbon isotopic ratios and enantiomeric composition. *Geochim. Cosmochim. Acta* **141**, 331–345.

Aponte J. C., Elsila J. E., Glavin D. P., Milam S. N., Charnley S. B. and Dworkin J. P. (2017) Pathways to meteoritic glycine and methylamine. *ACS Earth Space Chem.* **1**, 3–13.

Aponte J. C., McLain H. L., Dworkin J. P. and Elsila J. E. (2016) Aliphatic amines in Antarctic CR2, CM2, and CM1/2 carbonaceous chondrites. *Geochim. Cosmochim. Acta* **189**, 296–311.

Aponte José C., Whitaker D., Powner M. W., Elsila J. E. and Dworkin J. P. (2019b) Analyses of aliphatic aldehydes and ketones in carbonaceous chondrites. *ACS Earth Space Chem.* **3**, 463–472.

Aponte José C, Woodward H. K., Abreu N. M., Elsila J. E. and Dworkin J. P. (2019) Molecular distribution, ^{13}C-isotope, and enantiomeric compositions of carbonaceous chondrite monocarboxylic acids. *Meteorit. Planet. Sci.* **54**, 415–430.

Atherton J. H., Blacker J., Crampton M. R. and Grosjean C. (2004) The Strecker reaction:

kinetic and equilibrium studies of cyanide addition to iminium ions. *Org. Biomol. Chem.* **2**, 2567–2571.
Bernstein M. P., Dworkin J. P., Sandford S. A., Cooper G. W. and Allamandola L. J. (2002) Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues. *Nature* **416**, 401–403.
Billault I., Courant F., Pasquereau L., Derrien S., Robins R. J. and Naulet N. (2007) Correlation between the synthetic origin of methamphetamine samples and their ^{15}N and ^{13}C stable isotope ratios. *Anal. Chim. Acta* **593**, 20–29.
Biver N., Bockelée-Morvan D., Hofstadter M., Lellouch E., Choukroun M., Gulkis S., Crovisier J., Schloerb F. P., Rezac L., von Allmen P., Lee S., Leyrat C., Ip W. H., Hartogh P., Encrenaz P., Beaudin G. and the MIRO team (2019) Long-term monitoring of the outgassing and composition of comet 67P/Churyumov-Gerasimenko with the Rosetta/MIRO instrument. *A&A* **630**, A19.
Botta O. and Bada J. L. (2002) Extraterrestrial Organic Compounds in Meteorites. *Surveys in Geophysics* **23**, 411–467.
Bowes R. E. and Thorp J. H. (2015) Consequences of employing amino acid vs. bulk-tissue, stable isotope analysis: a laboratory trophic position experiment. *Ecosphere* **6**, art14.
Brand W. A., Assonov S. S. and Coplen T. B. (2010) Correction for the ^{17}O interference in $\delta^{13}C$ measurements when analyzing CO_2 with stable isotope mass spectrometry (IUPAC Technical Report). *Pure Appl. Chem.* **82**, 1719–1733.
Brunner B., Contreras S., Lehmann M. F., Matantseva O., Rollog M., Kalvelage T., Klockgether G., Lavik G., Jetten M. S. M., Kartal B. and Kuypers M. M. M. (2013) Nitrogen isotope effects induced by anammox bacteria. *Proc. Natl. Acad. Sci. USA* **110**, 18994–18999.
Bunch T. E. and Olsen E. (1975) Distribution and significance of chromium in meteorites. *Geochim. Cosmochim. Acta* **39**, 911–927.
Burton A. S., Elsila J. E., Hein J. E., Glavin D. P. and Dworkin J. P. (2013) Extraterrestrial amino acids identified in metal-rich CH and CB carbonaceous chondrites from Antarctica. *Meteorit. Planet. Sci.* **48**, 390–402.
Burton A. S., Stern J. C., Elsila J. E., Glavin D. P. and Dworkin J. P. (2012) Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites. *Chem. Soc. Rev.* **41**, 5459–5472.
Canellas P. F. and Cleland W. W. (1991) ^{13}C and deuterium isotope effects on the reaction catalyzed by glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry* **30**, 8871–8876.
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. *Org. Geochem.*
Charnley S. B., Ehrenfreund P., Millar T. J., Boogert A. C. A., Markwick A. J., Butner H. M., Ruiterkamp R. and Rodgers S. D. (2004) Observational tests for grain chemistry: posterior isotopic labelling. *Mon. Not. R. Astron. Soc.* **347**, 157–162.
Charnley S. B. and Rodgers S. D. (2006) Pathways to molecular complexity. *IAU* **1**, 237.
Chimiak L., Elsila J. E., Dallas B., Dworkin J. P., Aponte J. C., Sessions A. L. and Eiler

J. M. (2020) Carbon isotope evidence for the substrates and mechanisms of prebiotic synthesis in the early solar system. *Geochim. Cosmochim. Acta*.
Cleland W. W. (2005) The use of isotope effects to determine enzyme mechanisms. *Arch. Biochem. Biophys.* **433**, 2–12.
Combes F., Gerin M., Wootten A., Wlodarczak G., Clausset F. and Encrenaz P. J. (1987) Acetone in interstellar space. *Astronomy and Astrophysics* **180**.
Cooper G., Reed C., Nguyen D., Carter M. and Wang Y. (2011) Detection and formation scenario of citric acid, pyruvic acid, and other possible metabolism precursors in carbonaceous meteorites. *Proc. Natl. Acad. Sci. USA* **108**, 14015–14020.
Cooper G. W., Thiemens M. H., Jackson T. L. and Chang S. (1997) Sulfur and hydrogen isotope anomalies in meteorite sulfonic acids. *Science* **277**, 1072–1074.
Cordiner M. A., Palmer M. Y., Val-Borro M. de, Charnley S. B., Paganini L., Villanueva G., Bockelée-Morvan D., Biver N., Remijan A. J., Kuan Y. J., Milam S. N., Crovisier J., Lis D. C. and Mumma M. J. (2019) ALMA autocorrelation spectroscopy of comets: the HCN/H^{13}CN ratio in C/2012 S1 (ISON). *Astrophys. J.* **870**, L26.
Criss R. E. (1999) *Principles of stable isotope distribution*., Oxford University Press on Demand.
Cronin, Chang J. R. and S. (1993) Organic matter in meteorites: molecular and isotopic analyses of the Murchison meteorite.
Cronin J. R., Cooper G. W. and Pizzarello S. (1995) Characteristics and formation of amino acids and hydroxy acids of the Murchison meteorite. *Adv. Space Res.* **15**, 91–97.
Cronin J. R. and Moore C. B. (1971) Amino acid analyses of the Murchison, Murray, and Allende carbonaceous chondrites. *Science* **172**, 1327–1329.
Dallas B., Eiler J. M. and Sessions A. L. (2015) Position-Specific Carbon Isotope Analysis of Amino Acids by High-Resolution Gas Source IRMS. *Goldschmidt Abstracts*, 645.
Dennis M. J., Wilson P., Kelly S. and Parker I. (1998) The use of pyrolytic techniques to estimate site specific isotope data of vanillin. *J. Anal. Appl. Pyrolysis* **47**, 95–103.
Ehrenfreund P. and Charnley S. B. (2000) Organic Molecules in the Interstellar Medium, Comets, and Meteorites: A Voyage from Dark Clouds to the Early Earth. *Annu. Rev. Astron. Astrophys.* **38**, 427–483.
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. *Int. J. Mass Spectrom.* **422**, 126–142.
Eiler J. M., Clog M., Magyar P., Piasecki A., Sessions A., Stolper D., Deerberg M., Schlueter H.-J. and Schwieters J. (2013) A high-resolution gas-source isotope ratio mass spectrometer. *Int. J. Mass Spectrom.* **335**, 45–56.
Eisenhauer F., Schdel R., Genzel R., Ott T., Tecza M., Abuter R., Eckart A. and Alexander T. (2003) A geometric determination of the distance to the galactic center. *Astrophys. J.* **597**, L121–L124.
Elsila J. E., Aponte J. C., Blackmond D. G., Burton A. S., Dworkin J. P. and Glavin D. P. (2016) Meteoritic amino acids: diversity in compositions reflects parent body histories. *ACS Cent. Sci.* **2**, 370–379.

Elsila J. E., Charnley S. B., Burton A. S., Glavin D. P. and Dworkin J. P. (2012) Compound-specific carbon, nitrogen, and hydrogen isotopic ratios for amino acids in CM and CR chondrites and their use in evaluating potential formation pathways. *Meteorit Planet Sci* **47**, 1517–1536.

Elsila J. E., Dworkin J. P., Bernstein M. P., Martin M. P. and Sandford S. A. (2007) Mechanisms of amino acid formation in interstellar ice analogs. *Astrophys. J.* **660**, 911–918.

Engel M. H., Macko S. A. and Silfer J. A. (1990) Carbon isotope composition of individual amino acids in the Murchison meteorite. *Nature* **348**, 47–49.

Friedrich J. M., McLain H. L., Dworkin J. P., Glavin D. P., Towbin W. H., Hill M. and Ebel D. S. (2018) Effect of polychromatic X-ray microtomography imaging on the amino acid content of the Murchison CM chondrite. *Meteorit. Planet. Sci*.

Gilbert A., Silvestre V., Segebarth N., Tcherkez G., Guillou C., Robins R. J., Akoka S. and Remaud G. S. (2011) The intramolecular ^{13}C-distribution in ethanol reveals the influence of the CO_2 -fixation pathway and environmental conditions on the site-specific ^{13}C variation in glucose. *Plant Cell Environ.* **34**, 1104–1112.

Glavin D. P., Alexander C. M. O., Aponte J. C., Dworkin J. P., Elsila J. E. and Yabuta H. (2018) The origin and evolution of organic matter in carbonaceous chondrites and links to their parent bodies. In *Primitive meteorites and asteroids* Elsevier. pp. 205–271.

Glavin D. P., Callahan M. P., Dworkin J. P. and Elsila J. E. (2010) The effects of parent body processes on amino acids in carbonaceous chondrites. *Meteorit Planet Sci* **45**, 1948–1972.

Glavin D. P., Elsila J. E., McLain H. L., Aponte J. C., Parker E. T., Dworkin J. P., Hill D. H., Connolly H. C. and Lauretta D. S. (2020) Extraterrestrial amino acids and L-enantiomeric excesses in the CM 2 carbonaceous chondrites Aguas Zarcas and Murchison. *Meteorit. Planet. Sci*.

Gomez S., Peters J. A. and Maschmeyer T. (2002) The reductive amination of aldehydes and ketones and the hydrogenation of nitriles: mechanistic aspects and selectivity control. *Adv. Synth. Catal.* **344**, 1037–1057.

Guo W. and Eiler J. M. (2007) Temperatures of aqueous alteration and evidence for methane generation on the parent bodies of the CM chondrites. *Geochim. Cosmochim. Acta* **71**, 5565–5575.

Hanowski N. P. and Brearley A. J. (2000) Iron-rich aureoles in the CM carbonaceous chondrites Murray, Murchison, and Allan Hills 81002: Evidence for*in situ* aqueous alteration. *Meteorit. Planet. Sci.* **35**, 1291–1308.

Hoffman D. W. and Rasmussen C. (2019) Position-Specific Carbon Stable Isotope Ratios by Proton NMR Spectroscopy. *Anal. Chem.* **91**, 15661–15669.

Hofmann A. E., Chimiak L., Dallas B., Griep-Raming J., Juchelka D., Makarov A., Schwieters J. and Eiler J. M. (2020) Using Orbitrap mass spectrometry to assess the isotopic compositions of individual compounds in mixtures. *Int. J. Mass Spectrom.* **457**, 116410.

Hogberg P. (1997) Tansley Review No. 95. ^{15}N natural abundance in soil-plant systems. *New Phytol.* **137**, 179–203.

Huang Y., Alexandre M. R. and Wang Y. (2007) Structure and isotopic ratios of aliphatic

side chains in the insoluble organic matter of the Murchison carbonaceous chondrite. *Earth Planet. Sci. Lett.* **259**, 517–525.
Huang Y., Wang Y., Alexandre M. R., Lee T., Rose-Petruck C., Fuller M. and Pizzarello S. (2005) Molecular and compound-specific isotopic characterization of monocarboxylic acids in carbonaceous meteorites. *Geochim. Cosmochim. Acta* **69**, 1073–1084.
Kebukawa Y., Chan Q. H. S., Tachibana S., Kobayashi K. and Zolensky M. E. (2017) One-pot synthesis of amino acid precursors with insoluble organic matter in planetesimals with aqueous activity. *Sci. Adv.* **3**, e1602093.
Keller L. P. (2011) Mineralogy and petrography of MIL 090001, a highly altered CV chondrite from the reduced sub-group.
Kendall E. C., McKenzie B. F., Tobie W. C. and Ayres G. B. (1929) dl-Alanine. *Organic Syntheses*, 4–4.
Kerridge J. F. (1999) Formation and processing of organics in the early solar system. *Space Sci. Rev.* **90**, 275–288.
Kirshenbaum I., Smith J. S., Crowell T., Graff J. and McKee R. (1947) Separation of the Nitrogen Isotopes by the Exchange Reaction between Ammonia and Solutions of Ammonium Nitrate. *J. Chem. Phys.* **15**, 440–446.
Lallement R., Capitanio L., Ruiz-Dern L., Danielski C., Babusiaux C., Vergely L., Elyajouri M., Arenou F. and Leclerc N. (2018) Three-dimensional maps of interstellar dust in the Local Arm: using *Gaia*, 2MASS, and APOGEE-DR14. *A&A* **616**, A132.
Lamm V., Pan X., Taniguchi T. and Curran D. P. (2013) Reductions of aldehydes and ketones with a readily available N-heterocyclic carbene borane and acetic acid. *Beilstein J. Org. Chem.* **9**, 675–680.
Langer W. D., Graedel T. E., Frerking M. A. and Armentrout P. B. (1984) Carbon and oxygen isotope fractionation in dense interstellar clouds. *Astrophys. J.* **277**, 581–590.
Langer W. D. and Penzias A. A. (1990) C-12/C-13 isotope ratio across the Galaxy from observations of $^{13}C/^{18}O$ in molecular clouds. *Astrophys. J.* **357**, 477–492.
Lee M. R., Sofe M. R., Lindgren P., Starkey N. A. and Franchi I. A. (2013) The oxygen isotope evolution of parent body aqueous solutions as recorded by multiple carbonate generations in the Lonewolf Nunataks 94101 CM2 carbonaceous chondrite. *Geochim. Cosmochim. Acta* **121**, 452–466.
Lerner N. R. (1995) Influence of Murchison or Allende minerals on hydrogen-deuterium exchange of amino acids. *Geochim. Cosmochim. Acta* **59**, 1623–1631.
Lerner N. R., Peterson E. and Chang S. (1991) Evaluation of the Strecker synthesis as a source of amino acids on carbonaceous chondrites. *Meteoritics* **26**.
Li L., Lollar B. S., Li H., Wortmann U. G. and Lacrampe-Couloume G. (2012) Ammonium stability and nitrogen isotope fractionations for $–NH_3(aq)–NH_3(gas)$ systems at 20–70°C and pH of 2–13: Applications to habitability and nitrogen cycling in low-temperature hydrothermal systems. *Geochim. Cosmochim. Acta* **84**, 280–296.
Li Y., Zhang L., Xiong Y., Gao S., Yu Z. and Peng P. (2018) Determination of position-specific carbon isotope ratios of propane from natural gas. *Org. Geochem.* **119**, 11–21.
Lynn K. R. and Yankwich P. E. (1961) Cyanide carbon isotope fractionation in the reaction of cyanide ion and methyl iodide. Carbon isotope effect in the hydrolysis of

methyl iodide. *J. Am. Chem. Soc.* **83**, 53–57.
Lyons J. R., Gharib-Nezhad E. and Ayres T. R. (2018) A light carbon isotope composition for the Sun. *Nat. Commun.* **9**, 908.
Makarov A. (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal. Chem.* **72**, 1156–1162.
de Marcellus P., Meinert C., Myrgorodska I., Nahon L., Buhse T., d'Hendecourt L. L. S. and Meierhenrich U. J. (2015) Aldehydes and sugars from evolved precometary ice analogs: importance of ices in astrochemical and prebiotic evolution. *Proc. Natl. Acad. Sci. USA* **112**, 965–970.
Marlier J. F. (1993) Heavy-atom isotope effects on the alkaline hydrolysis of methyl formate: the role of hydroxide ion in ester hydrolysis. *J. Am. Chem. Soc.* **115**, 5953–5956.
Marlier J. F. and O'Leary M. H. (1984) Carbon kinetic isotope effects on the hydration of carbon dioxide and the dehydration of bicarbonate ion. *J. Am. Chem. Soc.* **106**, 5054–5057.
Martin G. J. (1996) Tracing back the origin of vanillin by SNIF-NMR. In *The roots of organic development* Industrial Chemistry Library. Elsevier. pp. 506–527.
Martins Z., Alexander C. M. O., Orzechowska G. E., Fogel M. L. and Ehrenfreund P. (2007) Indigenous amino acids in primitive CR meteorites. *Meteorit Planet Sci* **42**, 2125–2136.
Martins Z. and Sephton M. A. (2009) Extraterrestrial Amino Acids. In *Amino acids, peptides and proteins in organic chemistry: origins and synthesis of amino acids* (ed. A. B. Hughes). Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany. pp. 1–42.
Mccollom T. and Seewald J. (2006) Carbon isotope composition of organic compounds produced by abiotic synthesis under hydrothermal conditions. *Earth Planet. Sci. Lett.* **243**, 74–84.
McGuire B. A. (2018) 2018 census of interstellar, circumstellar, extragalactic, protoplanetary disk, and exoplanetary molecules. *Astrophys. J. Suppl. S.* **239**, 17.
Merritt D. A. and Hayes J. M. (1994) Factors Controlling Precision and Accuracy in Isotope-Ratio-Monitoring Mass Spectrometry. *Anal. Chem.* **66**, 2336–2347.
Milam S. N., Savage C., Brewster M. A., Ziurys L. M. and Wyckoff S. (2005) The $^{12}C/^{13}C$ isotope gradient derived from millimeter transitions of CN: The case for galactic chemical evolution. *Astrophys. J.* **634**, 1126.
Millar T. J. (2005) Deuterium in interstellar clouds. *Astron. Geophys.* **46**, 2.29-2.32.
Miller S. L. (1957) The mechanism of synthesis of amino acids by electric discharges. *Biochim. Biophys. Acta* **23**, 480–489.
Miller S. L. and Van Trump J. E. (1981) The Strecker synthesis in the primitive ocean. In *Origin of Life* (ed. Y. Wolman). Springer Netherlands, Dordrecht. pp. 135–141.
Moutou G., Taillades J., Bénefice-Malouet S., Commeyras A., Messina G. and Mansani R. (1995) Equilibrium of α-aminoacetonitrile formation from formaldehyde, hydrogen cyanide and ammonia in aqueous solution: Industrial and prebiotic significance. *J Phys Org. Chem.* **8**, 721–730.
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. *Anal. Chem.* **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. *Int. J. Mass Spectrom.* **434**, 276–286.
Oba Y. and Naraoka H. (2006) Site-specific carbon isotope analysis of aromatic carboxylic acids by elemental analysis/pyrolysis/isotope ratio mass spectrometry. *Rapid Commun. Mass Spectrom.* **20**, 3649–3653.
O'Connor C. (1970) Acidic and basic amide hydrolysis. *Q. Rev. Chem. Soc.* **24**, 553.
O'Leary M. H. and Marlier J. F. (1979) Heavy-atom isotope effects on the alkaline hydrolysis and hydrazinolysis of methyl benzoate. *J. Am. Chem. Soc.* **101**, 3300–3306.
Osamura Y., Roberts H. and Herbst E. (2005) The gas-phase deuterium fractionation of formaldehyde. *Astrophys. J.* **621**, 348–358.
Perkins M. J., McDonald R. A., van Veen F. J. F., Kelly S. D., Rees G. and Bearhop S. (2014) Application of nitrogen and carbon stable isotopes $\delta^{15}N$ and $\delta^{13}C$ to quantify food chain length and trophic structure. *PLoS One* **9**, e93281.
Pietrucci F., Aponte J. C., Starr R. D., Pérez-Villa A., Elsila J. E., Dworkin J. P. and Saitta A. M. (2018) Hydrothermal decomposition of amino acids and origins of prebiotic meteoritic organic compounds. *ACS Earth Space Chem.* **2**, 588–598.
Pirrung M. C. (2007) *The synthetic organic chemist's companion*., John Wiley & Sons, Inc., Hoboken, NJ, USA.
Pizzarello S. (2014) The Nitrogen Isotopic Composition of Meteoritic HCN. *Astrophys. J.* **796**, L25.
Pizzarello S., Cooper G. W. and Flynn G. J. (2006) The nature and distribution of the organic material in carbonaceous chondrites and interplanetary dust particles. *Meteorites and the early solar system II* **1**, 625–651.
Pizzarello S., Feng X., Epstein S. and Cronin J. R. (1994) Isotopic analyses of nitrogenous compounds from the Murchison meteorite: ammonia, amines, amino acids, and polar hydrocarbons. *Geochim. Cosmochim. Acta* **58**, 5579–5587.
Pizzarello S. and Holmes W. (2009) Nitrogen-containing compounds in two CR2 meteorites: ^{15}N composition, molecular distribution and precursor molecules. *Geochim. Cosmochim. Acta* **73**, 2150–2162.
Pizzarello S. and Huang Y. (2002) Molecular and isotopic analyses of Tagish Lake alkyl dicarboxylic acids. *Meteorit Planet Sci* **37**, 687–696.
Pizzarello S. and Huang Y. (2005) The deuterium enrichment of individual amino acids in carbonaceous meteorites: A case for the presolar distribution of biomolecule precursors. *Geochim. Cosmochim. Acta* **69**, 599–605.
Pizzarello S., Huang Y. and Alexandre M. R. (2008) Molecular asymmetry in extraterrestrial chemistry: Insights from a pristine meteorite. *Proc. Natl. Acad. Sci. USA* **105**, 3700–3704.
Pizzarello S., Huang Y. and Fuller M. (2004) The carbon isotopic distribution of Murchison amino acids. *Geochim. Cosmochim. Acta* **68**, 4963–4969.
Pizzarello S., Krishnamurthy R. V., Epstein S. and Cronin J. R. (1991) Isotopic analyses of amino acids from the Murchison meteorite. *Geochim. Cosmochim. Acta* **55**, 905–910.

Pizzarello S. and Shock E. (2010) The organic composition of carbonaceous meteorites: the evolutionary story ahead of biochemistry. *Cold Spring Harb. Perspect. Biol.* **2**, a002105.
Pizzarello S., Wang Y. and Chaban G. M. (2010) A comparative study of the hydroxy acids from the Murchison, GRA 95229 and LAP 02342 meteorites. *Geochim. Cosmochim. Acta* **74**, 6206–6217.
Pizzarello S., Williams L. B., Lehman J., Holland G. P. and Yarger J. L. (2011) Abundant ammonia in primitive asteroids and the case for a possible exobiology. *Proc. Natl. Acad. Sci. USA* **108**, 4303–4306.
PubChem PubChem.
Rajesh H. and Ozkan U. S. (1993) Complete oxidation of ethanol, acetaldehyde and ethanol/methanol mixtures over copper oxide and copper-chromium oxide catalysts. *Ind Eng Chem Res* **32**, 1622–1630.
Remusat L., Palhol F., Robert F., Derenne S. and Francelanord C. (2006) Enrichment of deuterium in insoluble organic matter from primitive meteorites: A solar system origin? *Earth Planet. Sci. Lett.* **243**, 15–25.
Robins L. I., Fogle E. J. and Marlier J. F. (2015) Mechanistic investigations of the hydrolysis of amides, oxoesters and thioesters via kinetic isotope effects and positional isotope exchange. *BBA-Proteins and Proteomics* **1854**, 1756–1767.
Rodgers S. D. and Charnley S. B. (2002) Multiply deuterated molecules and constraints on interstellar chemistry. *Planet. Space Sci.* **50**, 1125–1132.
Roueff E., Lis D. C., van der Tak F. F. S., Gerin M. and Goldsmith P. F. (2005) Interstellar deuterated ammonia: from NH_3 to ND_3. *A&A* **438**, 585–598.
Rustad J. R. (2009) Ab initio calculation of the carbon isotope signatures of amino acids. *Org. Geochem.* **40**, 720–723.
Rustad J. R., Nelmes S. L., Jackson V. E. and Dixon D. A. (2008) Quantum-chemical calculations of carbon-isotope fractionation in CO_2(g), aqueous carbonate species, and carbonate minerals. *J. Phys. Chem. A* **112**, 542–555.
Sandford S. A., Bernstein M. P. and Dworkin J. P. (2001) Assessment of the interstellar processes leading to deuterium enrichment in meteoritic organics. *Met. Planet. Sci.* **36**, 1117–1133.
Schlesinger G. and Miller S. L. (1973) Equilibrium and kinetics of glyconitrile formation in aqueous solution. *J. Am. Chem. Soc.* **95**, 3729–3735.
Scott J. H., O'Brien D. M., Emerson D., Sun H., McDonald G. D., Salgado A. and Fogel M. L. (2006) An examination of the carbon isotope effects associated with amino acid biosynthesis. *Astrobiology* **6**, 867–880.
Sephton M. A. (2002) Organic compounds in carbonaceous meteorites. *Nat Prod Rep* **19**, 292–311.
Sigman D. M., Karsh K. L. and Casciotti K. L. (2009) Ocean process tracers: nitrogen isotopes in the ocean.
Simkus D. N., Aponte J. C., Hilts R. W., Elsila J. E. and Herd C. D. K. (2019) Compound-specific carbon isotope compositions of aldehydes and ketones in the Murchison meteorite. *Meteorit. Planet. Sci.* **54**, 142–156.
Taran Y. A., Kliger G. A. and Sevastianov V. S. (2007) Carbon isotope effects in the

open-system Fischer–Tropsch synthesis. *Geochim. Cosmochim. Acta* **71**, 4474–4487.
Telus M., Alexander C. M. O., Hauri E. H. and Wang J. (2019) Calcite and dolomite formation in the CM parent body: Insight from in situ C and O isotope analyses. *Geochim. Cosmochim. Acta* **260**, 275–291.
Thomas F., Randet C., Gilbert A., Silvestre V., Jamin E., Akoka S., Remaud G., Segebarth N. and Guillou C. (2010) Improved characterization of the botanical origin of sugar by carbon-13 SNIF-NMR applied to ethanol. *J. Agric. Food Chem.* **58**, 11580–11585.
Van Slyke D D, DA MacFadyen and Hamilton P. (1941) Determination of free amino acids by titration of the carbon dioxide formed in the reaction with ninhydrin. *J. Bio. Chem.* **141**, 671–680.
Van Slyke Donald D., Dillon R. T., MacFadyen D. A. and Hamilton P. (1941) Gasometric Determination of carboxyl groups in free amino acids. *J. Bio. Chem.* **141**, 627–669.
Van Trump J. E. (1975) The Strecker synthesis and its prebiological importance. Doctoral dissertation, University of California: San Diego.
Vidal V., Theolier A., Thivolle-Cazat J. and Basset J. M. (1997) Metathesis of Alkanes Catalyzed by Silica-Supported Transition Metal Hydrides. *Science* **276**, 99–102.
Walters W. W., Chai J. and Hastings M. G. (2019) Theoretical phase resolved ammonia–ammonium nitrogen equilibrium isotope exchange fractionations: applications for tracking atmospheric ammonia gas-to-particle conversion. *ACS Earth Space Chem.* **3**, 79–89.
Wang Y., Huang Y., Alexander C. M. O. 'D., Fogel M. and Cody G. (2005) Molecular and compound-specific hydrogen isotope analyses of insoluble organic matter from different carbonaceous chondrite groups. *Geochim. Cosmochim. Acta* **69**, 3711–3721.
Wiberg K. B. (1954) The deuterium isotope effect of some ionic reactions of benzaldehyde. *J. Am. Chem. Soc.* **76**, 5371–5375.
Yamabe S., Zeng G., Guan W. and Sakaki S. (2014) Proton transfers in the Strecker reaction revealed by DFT calculations. *Beilstein J Org Chem* **10**, 1765–1774.
Yamataka H., Sasaki D., Kuwatani Y., Mishima M., Shimizu M. and Tsuno Y. (2001) Reactions of $PhSCH_2Li$ and $NCCH_2Li$ with benzaldehyde and benzophenone: when does the mechanism change from ET to polar? *J. Org. Chem.* **66**, 2131–2135.
Yamataka H., Sasaki D., Kuwatani Y., Mishima M. and Tsuno Y. (1997) On the mechanism of addition of lithium pinacolone enolate to benzaldehyde: polar or electron transfer? *J. Am. Chem. Soc.* **119**, 9975–9979.
Yeung C. S. and Dong V. M. (2011) Catalytic dehydrogenative cross-coupling: forming carbon-carbon bonds by oxidizing two carbon-hydrogen bonds. *Chem. Rev.* **111**, 1215–1292.
Yuen G., Blair N., Des Marais D. J. and Chang S. (1984) Carbon isotope composition of low molecular weight hydrocarbons and monocarboxylic acids from Murchison meteorite. *Nature* **307**, 252–254.
Zhu H., Clemente F. R., Houk K. N. and Meyer M. P. (2009) Rate limiting step precedes C-C bond formation in the archetypical proline-catalyzed intramolecular aldol reaction. *J. Am. Chem. Soc.* **131**, 1632–1633.

Appendix A: Supplemental Tables and Figures for Chapter 3

mmol species at equilibrium															
α-APN	*st err*	α-(OH)PN	*st err*	HCN	*st err*	CN-	*st err*	NH4+	*st err*	NH3	*st err*	Ace	*st err*	Ace(OH2)	*st err*
2.01	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.20	*0.03*	0.10	*0.00*	0.28	*0.01*	0.32	*0.01*
1.99	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.92	*0.02*	0.10	*0.00*	0.29	*0.01*	0.33	*0.01*
1.93	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.01	*0.02*	0.10	*0.00*	0.32	*0.01*	0.37	*0.01*
1.90	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.00	*0.02*	0.10	*0.00*	0.34	*0.01*	0.39	*0.02*
1.94	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.04	*0.02*	0.10	*0.00*	0.32	*0.01*	0.36	*0.01*
2.08	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.91	*0.02*	0.11	*0.00*	0.25	*0.01*	0.28	*0.01*
2.08	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.91	*0.02*	0.11	*0.00*	0.25	*0.01*	0.28	*0.01*
1.86	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.16	*0.03*	0.10	*0.00*	0.36	*0.01*	0.41	*0.02*
1.93	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.07	*0.02*	0.10	*0.00*	0.32	*0.01*	0.37	*0.01*
2.00	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.08	*0.03*	0.10	*0.00*	0.28	*0.01*	0.33	*0.01*
2.04	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.96	*0.02*	0.10	*0.00*	0.27	*0.01*	0.31	*0.01*
1.92	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.07	*0.03*	0.10	*0.00*	0.33	*0.01*	0.38	*0.02*
1.91	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.08	*0.03*	0.10	*0.00*	0.33	*0.01*	0.38	*0.02*
2.23	*0.12*	0.11	*0.01*	0.00	0.00	0.00	0.00	0.91	*0.02*	0.11	*0.00*	0.18	*0.01*	0.20	*0.01*
1.87	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.18	*0.03*	0.10	*0.00*	0.35	*0.01*	0.41	*0.02*
2.02	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.87	*0.02*	0.10	*0.00*	0.27	*0.01*	0.31	*0.01*
1.99	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.91	*0.02*	0.10	*0.00*	0.29	*0.01*	0.33	*0.01*
2.11	*0.11*	0.10	*0.00*	0.00	0.00	0.00	0.00	0.93	*0.02*	0.11	*0.00*	0.23	*0.01*	0.26	*0.01*
1.79	*0.09*	0.09	*0.00*	0.00	0.00	0.00	0.00	1.08	*0.03*	0.10	*0.00*	0.40	*0.02*	0.45	*0.02*
1.96	*0.10*	0.10	*0.00*	0.00	0.00	0.00	0.00	1.06	*0.02*	0.10	*0.00*	0.31	*0.01*	0.36	*0.01*

Table S3.1: Predicted mmol of species at equilibrium from Eqn 4-16 in the main text.

Sa	% yield α-APN vs NH_3	Equilibrium		
		NH_4^+	NH_3	α-APN
1	60 *(4)*	-15.0 *(0.3)*	-35.0 *(0.4)*	12.0 *(0.3)*
2	66 *(4)*	-16.4 *(0.3)*	-36.4 *(0.4)*	10.6 *(0.3)*
3	63 *(4)*	-15.7 *(0.3)*	-35.7 *(0.4)*	11.3 *(0.3)*
4	63 *(4)*	-15.6 *(0.3)*	-35.6 *(0.4)*	11.4 *(0.3)*
5	63 *(4)*	-15.5 *(0.3)*	-35.5 *(0.4)*	11.5 *(0.3)*
6	67 *(4)*	-16.7 *(0.3)*	-36.7 *(0.4)*	10.3 *(0.3)*
7	67 *(4)*	-16.7 *(0.3)*	-36.7 *(0.4)*	10.3 *(0.3)*
8	59 *(3)*	-14.7 *(0.3)*	-34.7 *(0.4)*	12.3 *(0.3)*
9	60 *(4)*	-15.4 *(0.3)*	-35.4 *(0.4)*	11.6 *(0.3)*
10	63 *(4)*	-15.6 *(0.3)*	-35.6 *(0.4)*	11.4 *(0.3)*
11	65 *(4)*	-16.3 *(0.3)*	-36.3 *(0.4)*	10.7 *(0.3)*
12	62 *(4)*	-15.3 *(0.3)*	-35.3 *(0.4)*	11.7 *(0.3)*
13	62 *(4)*	-15.2 *(0.3)*	-35.2 *(0.4)*	11.8 *(0.3)*
14	69 *(4)*	-17.1 *(0.4)*	-37.1 *(0.4)*	9.9 *(0.3)*
15	59 *(3)*	-14.6 *(0.3)*	-34.6 *(0.4)*	---
16	67 *(4)*	-16.8 *(0.3)*	-36.8 *(0.4)*	10.2 *(0.3)*
17	66 *(4)*	-16.5 *(0.3)*	-36.5 *(0.4)*	10.5 *(0.3)*
18	67 *(4)*	-16.7 *(0.3)*	-36.7 *(0.4)*	10.3 *(0.3)*
19	60 *(4)*	-14.8 *(0.3)*	-34.8 *(0.4)*	12.2 *(0.3)*
20	63 *(4)*	-15.5 *(0.3)*	-35.5 *(0.4)*	11.5 *(0.3)*

Table S3.2: $\delta^{15}N$ values for equilibrium step.

Sa	% yield α-APN at equilibrium vs		δ^{13}C reactants (‰)		δ^{13}C α-APN (‰)		
	ace	NaCN	ace	ace (OH2)	Molec Avg	C-2 + C-3	C-1
1	74 *(6)*	95 *(6)*	-18.6 *(0.2)*	-15.2 *(0.2)*	-29.7 *(0.1)*	-28.7 *(0.1)*	-31.8 *(0.2)*
2	73 *(6)*	95 *(6)*	-18.7 *(0.2)*	-15.3 *(0.2)*	-29.8 *(0.1)*	-28.8 *(0.1)*	-31.8 *(0.2)*
3	71 *(6)*	95 *(6)*	-19.0 *(0.2)*	-15.6 *(0.2)*	-30.0 *(0.1)*	-29.1 *(0.1)*	-31.8 *(0.2)*
4	70 *(6)*	95 *(6)*	-19.1 *(0.2)*	-15.7 *(0.2)*	-30.1 *(0.1)*	-29.2 *(0.1)*	-31.8 *(0.2)*
5	71 *(6)*	95 *(6)*	-18.9 *(0.2)*	-15.5 *(0.2)*	-30.0 *(0.1)*	-29.0 *(0.1)*	-31.8 *(0.2)*
6	77 *(6)*	95 *(6)*	-18.3 *(0.2)*	-14.9 *(0.2)*	-29.5 *(0.1)*	-28.4 *(0.1)*	-31.8 *(0.2)*
7	77 *(6)*	95 *(6)*	-18.3 *(0.2)*	-14.9 *(0.2)*	-29.5 *(0.1)*	-28.4 *(0.1)*	-31.8 *(0.2)*
8	68 *(6)*	95 *(6)*	-19.3 *(0.2)*	-15.9 *(0.2)*	-30.2 *(0.1)*	-29.4 *(0.1)*	-31.8 *(0.2)*
9	71 *(6)*	95 *(6)*	-19.0 *(0.2)*	-15.6 *(0.2)*	-30.0 *(0.1)*	-29.1 *(0.1)*	-31.8 *(0.2)*
10	74 *(6)*	95 *(6)*	-18.6 *(0.2)*	-15.2 *(0.2)*	-29.7 *(0.1)*	-28.7 *(0.1)*	-31.8 *(0.2)*
11	75 *(6)*	95 *(6)*	-18.5 *(0.2)*	-15.1 *(0.2)*	-29.6 *(0.1)*	-28.6 *(0.1)*	-31.8 *(0.2)*
12	70 *(6)*	95 *(6)*	-19.0 *(0.2)*	-15.6 *(0.2)*	-30.0 *(0.1)*	-29.1 *(0.1)*	-31.8 *(0.2)*
13	70 *(6)*	95 *(6)*	-19.1 *(0.2)*	-15.7 *(0.2)*	-30.1 *(0.1)*	-29.2 *(0.1)*	-31.8 *(0.2)*
14	82 *(6)*	96 *(6)*	-17.6 *(0.2)*	-14.2 *(0.2)*	-29.1 *(0.1)*	-27.7 *(0.1)*	-31.8 *(0.2)*
15	68 *(6)*	95 *(6)*	-19.3 *(0.2)*	-15.9 *(0.2)*	-30.2 *(0.1)*	-29.4 *(0.1)*	-31.8 *(0.2)*
16	75 *(6)*	95 *(6)*	-18.5 *(0.2)*	-15.1 *(0.2)*	-29.7 *(0.1)*	-28.6 *(0.1)*	-31.8 *(0.2)*
17	73 *(6)*	95 *(6)*	-18.7 *(0.2)*	-15.3 *(0.2)*	-29.8 *(0.1)*	-28.8 *(0.1)*	-31.8 *(0.2)*
18	78 *(6)*	95 *(6)*	-18.1 *(0.2)*	-14.7 *(0.2)*	-29.4 *(0.1)*	-28.2 *(0.1)*	-31.8 *(0.2)*
19	65 *(6)*	95 *(6)*	-19.7 *(0.2)*	-16.3 *(0.2)*	-30.5 *(0.1)*	-29.8 *(0.1)*	-31.8 *(0.2)*
20	72 *(6)*	95 *(6)*	-18.9 *(0.2)*	-15.5 *(0.2)*	-29.9 *(0.1)*	-29.0 *(0.1)*	-31.8 *(0.2)*

Table S3.3: δ^{13}C values for equilibrium step.

Sa	% Conversion to:		δ13C amide initial (‰)			δ13C amide resid (‰)	δ13C and δ15N alanine (‰)			
	amide	alanine	Molec Avg	C-2 + C-3	C-1	Molec Avg	Molec Avg	C-2 + C-3	C-1	N
1	56 *(3)*	42 *(1)*	-30.7 *(0.7)*	-29.2 *(1.5)*	-34 *(3)*	-27.32 *(0.22)*	-35.25 *(0.96)*	-29.2 *(1.5)*	-47 *(4)*	10.5 *(0.3)*
2	62 *(3)*	26 *(1)*	-30.8 *(0.9)*	-28.8 *(1.5)*	-35 *(4)*	-28.91 *(0.11)*	-35.93 *(0.04)*	-28.8 *(1.5)*	-50 *(3)*	9.6 *(0.7)*
3	66 *(4)*	35 *(1)*	-28.3 *(0.6)*	-25.4 *(1.5)*	-34 *(3)*	-25.76 *(0.06)*	-33.03 *(0.01)*	-25.4 *(1.4)*	-48 *(3)*	10.4 *(0.4)*
4	74 *(4)*	64 *(1)*	-30.0 *(0.5)*	-27.1 *(1.8)*	-36 *(3)*	-22.19 *(0.17)*	-34.38 *(0.44)*	-27.1 *(1.8)*	-49 *(4)*	10.8 *(0.1)*
5	71 *(4)*	100 *(1)*	-27.5 *(0.2)*	-29.2 *(1.5)*	-24 *(2)*	---	-27.54 *(0.16)*	-29.2 *(1.5)*	-24 *(3)*	12.0 *(0.2)*
6	62 *(3)*	47 *(1)*	---	-27.1 *(1.8)*	---	---	-35.06 *(0.05)*	-27.1 *(1.8)*	-51 *(4)*	11.9 *(0.3)*
7	65 *(4)*	35 *(1)*	-26.3 *(0.5)*	-26.3 *(1.4)*	-26 *(3)*	-24.22 *(0.09)*	-30.23 *(0.05)*	-26.3 *(1.4)*	-38 *(3)*	10.6 *(0.3)*
8	74 *(4)*	65 *(1)*	-31.1 *(0.4)*	---		-24.94 *(0.22)*	-34.38 *(0.15)*	---	---	12.8 *(0.1)*
9	77 *(4)*	35 *(1)*	-25.8 *(0.5)*	-22.4 *(1.5)*	-33 *(3)*	-23.01 *(0.27)*	-31.04 *(0.09)*	-22.4 *(1.5)*	-48 *(3)*	11.3 *(0.1)*
10	64 *(3)*	21 *(1)*	-31.0 *(1.1)*	-26.6 *(1.5)*	-40 *(4)*	-29.22 *(0.01)*	-37.71 *(0.12)*	-26.6 *(1.5)*	-60 *(3)*	12.1 *(0.1)*
11	51 *(3)*	11 *(1)*	-34.4 *(2.9)*	-28.9 *(1.7)*	-46 *(9)*	-33.63 *(0.02)*	-40.69 *(0.09)*	-28.9 *(1.7)*	-64 *(3)*	12.6 *(0.1)*
12	60 *(3)*	43 *(1)*	---	-27.5 *(1.5)*	---	---	-32.17 *(0.02)*	-27.5 *(1.5)*	-42 *(3)*	10.6 *(0.1)*
13	76 *(4)*	58 *(1)*	-26.8 *(0.3)*	-32.0 *(1.6)*	-16 *(3)*	-23.08 *(0.10)*	-29.45 *(0.10)*	-32.0 *(1.6)*	-24 *(3)*	10.4 *(0.1)*
14	54 *(3)*	26 *(1)*	-32.9 *(1.0)*	-27.0 *(1.5)*	-45 *(4)*	-30.34 *(0.31)*	-40.34 *(0.02)*	-27.0 *(1.5)*	-67 *(3)*	12.5 *(0.2)*
15	48 *(3)*	5 *(1)*		-30.1 *(1.8)*	---	---	---	-30.1 *(1.8)*	---	---
16	82 *(4)*	69 *(1)*	-33.2 *(0.4)*	-31.4 *(1.6)*	-37 *(3)*	-27.72 *(0.04)*	-35.67 *(0.02)*	-31.4 *(1.6)*	-44 *(3)*	9.8 *(0.3)*
17	75 *(4)*	18 *(1)*	---	-27.6 *(1.5)*	---	---	-34.70 *(0.04)*	-27.6 *(1.5)*	-49 *(3)*	10.4 *(0.1)*
18	89 *(5)*	42 *(1)*	-29.7 *(0.4)*	-31.6 *(1.5)*	-26 *(3)*	-27.55 *(0.10)*	-32.59 *(0.02)*	-31.6 *(1.5)*	-35 *(3)*	10.6 *(0.5)*
19	25 *(2)*	26 *(3)*	-35.3 *(2.7)*	-31.8 *(1.6)*	-42 *(9)*	-33.08 *(0.08)*	-41.43 *(0.06)*	-31.8 *(1.6)*	-61 *(3)*	12.0 *(0.5)*
20	34 *(2)*	76 *(3)*	-39.2 *(1.1)*	-34.8 *(1.6)*	-48 *(4)*	-31.75 *(0.03)*	-41.50 *(0.03)*	-34.8 *(1.6)*	-55 *(3)*	11.6 *(0.1)*

Table S3.4: Carbon and nitrogen isotope values during hydrolysis steps. Alanine and residual amide values are found from direct measurements. Initial amide values are calculated with the assumption that C-2 and C-3 values do not change during the conversion from alaninamide to alanine.

	D ala (‰)	st err	L ala (‰)	st err	high (‰)	low (‰)
Murchison[a]	99	*10*	162	*15*		
LON 94101[a]	120	*5*	111	*7*		
GRA 95220[a]	208	*10*	127	*3*		
GRA 95220[b]	125.7	*3*	126.2	*3.5*		
EET 92042 [a]	115	*5*	187	*11*		
QUE 99177[a]	326	*9*	141	*21*		
LAP 02342[b]	126.5	*0.25*	121.9	*1.75*		
Precipitation[c]					-18	8
Soil[d]					-8	18

Table S3.5: $\delta^{15}N$ values used to calculate ammonia values for meteorites and to calculate alanine amine $\delta^{15}N$ values for potential terrestrial samples made form Strecker synthesis. Literature sources are (a) Elsila *et al*., 2012, (b) Pizzarello *et al.*, 2009, (c) Peterson and Fry, 1987, and (d) Criss, 1999.

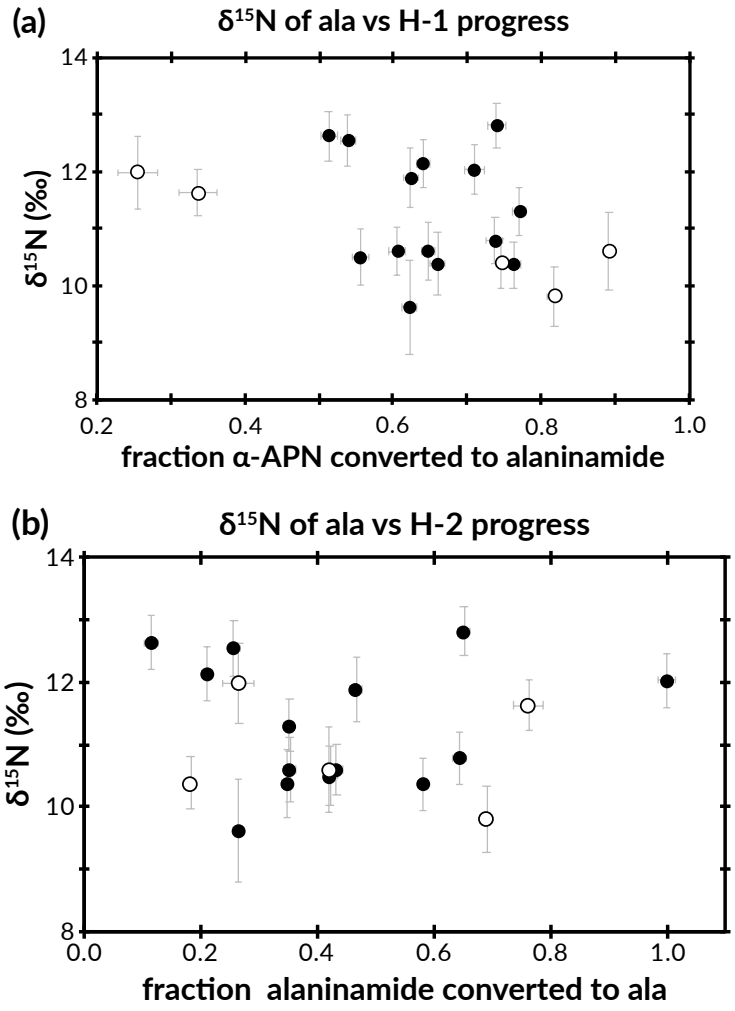


Figure S3.1: The measured $\delta^{15}N$ versus the reaction progress for (a) H-1 and (b) H-2. Reaction progress here is the fraction of initial reactant converted into product. Open symbols represent samples where the hydrolysis temperature was above 115 °C.

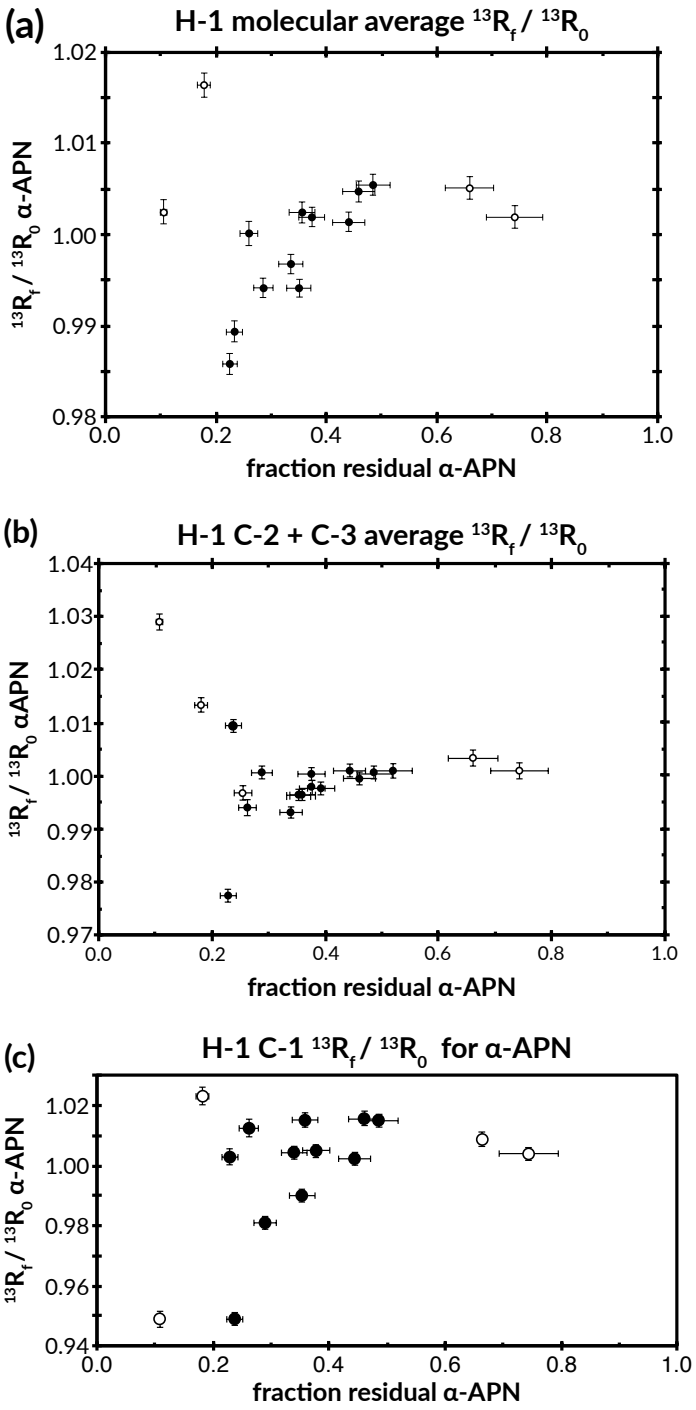


Figure S3.2: H-1 Rayleigh plots for (a) molecular-average ^{13}R, (b) the averaged C-2 and C-3site ^{13}R, (c) the C-1 site ^{13}R. Open symbols represent samples where the hydrolysis temperature was above 115 °C.

Appendix B: Supplemental Information for Chapter 4

B1. Calibration of alanine standards for Site-Specific Isotope Ratio (SSIR) Measurements

The molecular-average $\delta^{13}C$ values of pure alanine standards were measured on a Thermo Fisher Scientific Flash Elemental Analyzer (EA) coupled to a Delta-V isotope ratio mass spectrometer (IRMS) at Caltech. Alanine standards are described above (See Materials: Derivatization). A lab acetanilide standard served as check on accuracy of $\delta^{13}C$ measurements. The $\delta^{13}C$ values and associated uncertainties for the alanine standards are -19.4 ± 0.1 ‰, -20.0 ± 0.2 ‰, and -32.9 ± 0.2 ‰ for Alfa Aesar, VWR, and Strecker alanine respectively (Eiler et al., 2017) (Table 1); acetanilide was measured to have a $\delta^{13}C$ value of -27.6 ± 0.1 ‰ in good agreement with its prior measured value of -27.7 ± 1.7‰.

The Alfa Aesar and Strecker alanine standards were also analyzed at GSFC following protocols from (Elsila et al., 2012) using coupled GC-combustion-IRMS (GC-C-IRMS), which enables isotopic analysis of individual amino acids in mixtures such as those from the Murchison extracts. After accounting for dilution effects from the derivative methyl and isopropyl groups (See Data Processing and (Elsila et al., 2012) for details on dilution effects), the standards' $\delta^{13}C$ values were -19.4 ± 0.2 ‰ and -33.3 ± 0.1 ‰ for Alfa Aesar and Strecker respectively, which is within two standard errors of those measured at Caltech (Table 1).

We also measured the $\delta^{13}C$ values of C-1 in all 3 alanine standards via ninhydrin decarboxylation, following methods from (Donald D. Van Slyke et al., 1941) and (Abelson and Hoering, 1961). Resulting $\delta^{13}C_{VPDB}$ values for C-1 were -28.5 ± 0.1 ‰, -29.5 ± 0.3 ‰, -43.5 ± 0.1 ‰, for the Alfa Aesar, VWR, and Strecker standards, respectively (Eiler et al., 2017). Combining these data with the molecular-average $\delta^{13}C$ values from above allowed us to calculate the average $\delta^{13}C$ of their combined C-2 and C-3 sites (See Section 2.3: Data Processing for calculations and Figure 1a in main text for

alanine with labelled carbon sites) as -14.8 ± 0.6 ‰, -27.6 ± 0.3 ‰, and -15.6 ± 0.3 ‰. At the time of this publication, we have no independent evidence regarding the individual isotopic compositions of the C-2 and C-3 sites in these standards; however, NMR studies of site-specific carbon isotope ratios of amino acids (R. Robins pers. com.) indicate that all common terrestrial forms of these amino acids, including standards purchased from Sigma Aldrich (BioUltra, >99% Purity, Lot# BCBM6312V), have $\delta^{13}C_{VPDB}$ fractionations between C-2 and C-3 in each molecule that are 10 ‰ or less, which is in the upper range of differences between methyl and adjacent cites for other small organics (Gilbert et al., 2011). The differences we observe in the Murchison sample relative to the Alfa Aesar standard C-2 and C-3 are on the order of 170 ‰, and the error of the C-3 calculation (10 ‰) is within error of the 10 ‰ difference found between C-2 and C-3 in other alanine samples. Consequently, the potential 10 ‰ difference is negligible in our study, and for this study we assume our standards have C-2 and C-3 sites that are identical in $\delta^{13}C$. Future measurements of one or more of the standards used in this study could be used to refine the data presented here in order to account for the likely small differences between C-2 and C-3 in our alanine standards, but we think it implausible that our conclusions could be influenced by the small isotopic differences between these sites likely present in our terrestrial standards.

Site-specific $\delta^{13}C$ values for the Methods Development samples measured in December and March are within error of one another (Table 3.1). We interpret the differences in site-specific isotope ratios between methods development and analytical samples as being due to terrestrial contamination (though it is also possible that they partially reflect differences in isotopic composition between the alanine native to these two Murchison samples or fractionations arising from chemical reactions of sample alanine during storage). Regardless, we base our discussion of the Murchison sample only on the analytical sample. We present the data for the methods development sample only in order to document the development of the methods used in this study.

B2. Blanks

Multiple procedural blanks were carried through the workup and analyzed alongside the Methods Development and Analytical samples. All blanks start at their listed step (*e.g.*, extraction, transfer, derivatization; see Table S3.1) and follow all subsequent steps through derivatization as outlined in Figure S3.2. As an example, blanks designed to test the extraction of amino acids had water added to an empty ampoule after which all subsequent extraction, transfer, and derivatization steps were followed. Thus, all blanks should only contain derivatizing reagents, the products of their reactions with one another, and hexane if sample processing produced no contamination. Procedural Blanks are summarized in Table S3.1 and consisted of the following: (1a and 1b) blanks that tested reagents used in the derivatization of alanine (our ultimate analytical target), (2) a blank that starts with water leaching at GSFC and continues through chemical derivatization at Caltech, (3) a blank that starts with the water:methanol transfer of the meteorite extract into a GC vial at Caltech, and (4) a blank that starts with analyte derivatization at Caltech (See Figure S3.2). Procedural Blanks 1a and 1b occurred prior to the day of meteorite extract derivatization while Procedural Blanks 2-4 occurred on the same day as the corresponding meteorite extract derivatization. Additional solvent blanks (injections of hexane into the Orbitrap) and instrument blanks (temperature ramps with no injection) were run prior to each meteorite analysis to test the instrument background.

Each procedural blank was analyzed in Direct Injection mode on the Orbitrap, and signals were integrated between 6.5 and 8.5 minutes after injection for ^{12}C and ^{13}C counts from *m/z* 140.032 and 141.035 fragment peaks (for conversion from signal intensity to counts see Eiler et al. (2017). Alanine elutes at ~7.5 minutes and is typically transferred into the reservoir from approximately 7–8 minutes retention time, so counting the background over 2 minutes overestimates possible contamination. As with the sample data (see Site-Specific Isotope Analysis and Data Processing) data used to calculate ^{13}R was culled only to include scans that contained both the monoisotopic and singly ^{13}C-substituted fragment and was computed using a counts-weighted average of all ^{13}R values in the blank. Reported sums of ^{12}C and ^{13}C counts (Dataset S1) use all scans including those which

have only the monoisotopic or the singly ^{13}C substituted fragment without the other in order demonstrate the maximum possible error in our measurements. When compared to samples measured with the Reservoir Elution mode, the overestimation is even greater because in Reservoir Elution mode measurements are broadened over many tens of minutes, giving them a lower signal-to-noise ratio (which is inversely proportional to counts reported). The procedural blank for analytical Murchison that had the highest amount of contamination in all metrics was Procedural Blank 2 (Table S3.1), which started with the meteorite extraction at GSFC. However, compared to the 15 pmol/μL alanine in the analytical sample, Procedural Blank 2 contained 0.15 pmol/μL and could account for only 1.9 % of the integrated ^{12}C counts, 0.7% of the integrated ^{13}C counts, and 0.3 % of the integrated ^{12}C signal intensity relative to the directly injected Murchison sample. The 140.032 and 141.035 *m/z* fragments are the most abundant ones in the mass spectrum of alanine. Maximum abundances of *m/z* 140.032 and 141.035 ions in blanks were low (see Dataset S1) and did not appear during the 7.41-7.73 window during which alanine elutes, so these background signals likely either represent other compounds derived from column bleed, reagents, etc., and/or part of the instrument background. For chromatograms and spectra of blanks and Murchison, see Figure S3.3.

Solvent blanks and instrument blanks were run prior to meteorite sample analyses and also processed for integrated ^{12}C and ^{13}C counts from 6.5 to 8.5 minutes elution time (Table S3.2). These measurements find background ^{12}C and ^{13}C counts arising from the injector, column, transfer lines, etc. to typically account for less than 0.5% of the measured ^{12}C and ^{13}C counts in Murchison samples and a <0.05 ‰ change in ^{13}R values. Of the fragments used to calculate the site-specific isotope ratios of alanine, the highest background signals were observed for the *m/z* 184.021 fragment. In this case, the background counts account for approximately 0.5 % of the measured signal but change the ^{13}R value by only ~0.03 ‰, which is well within the ~10 ‰ standard error of the measurements at the 184.021 fragment. The low procedural blanks and instrument background demonstrate that our ^{13}R values reflect alanine from the meteorite rather than background or contamination.

B3. Potential additional constraints for alanine SSIR measurement

We attempted to add a fourth constraint to our characterization of the carbon isotope structure of alanine by measuring the ^{13}R of a fragment ion having a monoisotopic mass of 113.0208 ($C_3H_4OF_3$). The straightforward fragment suggested by this mass would be $CF_3CH(O)CH_3$ using C-2 from the parent alanine. However, our studies of labeled alanines suggest that this fragment only receives sample carbon atoms from C-3 of the parent alanine along with two carbons from the TFAA derivatizing reagents and none from C-2 of the parent alanine. The stoichiometry of this ion suggests it is a recombination product (*i.e.*, because direct fragmentation of the parent molecule cannot create a single piece containing these sites). We infer C-3 of alanine recombines with COH and CF_3 from the TFAA derivatizing reagent either as a two-body reaction or as two stepwise reactions. This complexity calls into question whether such a measurement could yield a consistent constraint on the ^{13}R of C-3 because the yields of recombination reactions generally depend on source pressure and other analytical variables (*i.e.*, we can imagine the same ion might be produced through other pathways when analytical conditions are varied). In any case, when we attempted to apply this method to the derivatized Murchison extract our peak captures of alanine were contaminated by at least one subsequent peak of a different compound. We recognize one such candidate contaminant peak also produces a 113.0208 Da fragment ion. Thus, we consider these measurements to have failed for reasons having to do with our chromatographic separations and peak trapping. We report these results in the for completeness, but we do not use these data as constraints on the Murchison sample carbon isotope structure.

B4. Error Analysis

Errors for the Total Orbitrap and the Combined Orbitrap/GC-C-MS calculations were weighted according to the proportion effect of their value on the final calculation and then added in quadrature (Eqn. S3.1a-S3.1c):

Combined Orbitrap(140,184)/GC-C-IRMS Calculation Error

$$^{13}\sigma_{C\text{-}1} = \{(3 \times {}^{13}\sigma_{molec\ avg})^2 + (2 \times {}^{13}\sigma_{C\text{-}2+C\text{-}3})^2\}^{0.5} \quad \textit{(Eqn. S3.1a)}$$

$$^{13}\sigma_{C\text{-}2} = \{(2 \times {}^{13}\sigma_{C\text{-}1+C\text{-}2})^2 + {}^{13}\sigma_{C\text{-}1}{}^2\}^{0.5} \quad \textit{(Eqn. S3.1b)}$$

$$^{13}\sigma_{C\text{-}3} = \{(2 \times {}^{13}\sigma_{C\text{-}2+C\text{-}3})^2 + {}^{13}\sigma_{C\text{-}2}{}^2\}^{0.5} \quad \textit{(Eqn. S3.1c)}$$

It is important to note that the resulting computed errors for the three alanine sites are highly correlated with one another due to interdependencies among the functions that relate them to the various measured ratios. In particular, the $\delta^{13}C$ of C-2 and C-3 are associated with large errors, yet their average is known to within 1.5 ‰ (1SE). The primary control on the error is the experimental uncertainty in the average C-1 + C-2 $\delta^{13}C$, which is doubled in computing the site-specific uncertainty of the C-2 site (See Eqn. S3.1b) and then propagated into the calculated $\delta^{13}C$ of the C-3 site. If future studies improve in the precision of the results presented here, it will be productive to focus on these dependencies; in particular, a highly precise molecular-average measurement that includes the derivative carbons, a high precision analysis of the *m/z* = 184.021, and a high precision analysis of the fragment *m/z* = 113.032 fragment with peak capturing that excludes subsequent peaks. These improvements were not possible during this study due to limited sample sizes, but a more ambitious effort to extract and purify alanine from Murchison might achieve errors on the order of ~1 ‰ for all sites (see (Neubauer et al., 2018) for an example of high precision amino acid C isotope structures measured using our techniques).

B5. Alternative Pathways for Alanine Synthesis

In addition to acetaldehyde and cyanide reacting via Strecker synthesis, the alanine carbon isotope structure could be explained by the reductive amination of pyruvic acid(Rustad, 2009; Robins et al., 2015). In this case, the pyruvic acid would form from a ketene (ethenone) which sources its alkyl group (C-2) from the same ^{13}C-depleted CH_x pool and its CO (C-1) from the same ^{13}C-enriched CO pool described in the main text (See Figure S3.4). The ethenone would then react with CN and water to form pyruvic

acid that could react with NH_3 on later to form alanine. Consequently, assuming a low ^{13}C ISM CN pool, this reaction network could explain our results. Furthermore, as the reaction network (Figure S3.4) still involves the addition of CN to an sp^2-hybridized carbon and the oxidation of a nitrile to a carboxyl group (Rustad, 2009), the isotope effect and thus predicted initial carbon values should not greatly change between the scenarios (excepting possible changes in isotope effect due to physiochemical conditions).

Unlike the Strecker model, the pyruvate model would not provide clear pathways to amines, aldehydes, or monocarboxylic acids. Furthermore, measured values of keto acids are, as of yet, unavailable such that we could not compare predictions of this model to our data. For this reason, we chose to focus on the Strecker synthesis possibility. The agreement between our predictions and measured values across a wide range of compound classes supports the possibility that Strecker synthesis of aldehydes and cyanohydrins produced alanine and other organic compounds.

We also considered whether Murchison alanine could be the product of a reaction network in which alanine carboxyl is derived from high $\delta^{13}C$ HCN, through Strecker chemistry. This hypothesis could be indirectly supported by the observation that monocarboxylic acids in Murchison have high molecular average $\delta^{13}C$ values (Yuen et al., 1984). If these carboxylic acids formed by hydrolysis of nitriles, then those nitriles presumably could have been high in $\delta^{13}C$. And if that ^{13}C enrichment were hosted by the terminal CN group, we should expect co-existing HCN would be ^{13}C enriched. We are not aware of measurements of $\delta^{13}C$ of Murchison nitriles (and their terminal CN groups are certainly not known). But if their terminal CN groups were enriched enough to account for the 10's of per mil enrichment of carboxylic acids, it would imply a $\delta^{13}C$ value for that group of +100 ‰ or more. This hypothesis is speculative but based on sound chemical principles and so worth considering. Nevertheless, it is strongly contradicted by data (both from previous studies and our study), so we think it must be rejected. Most simply, HCN from Murchison is relatively low in $\delta^{13}C$ (Pizzarello, 2014), and our measurement of alanine carboxyl indicates it is consistent with derivation by Strecker reaction from that measured HCN. We conclude the most parsimonious

interpretation is that alanine in fact did form from the HCN present in Murchison, and that this HCN was not derived from a strongly ^{13}C-enriched pre-solar pool.

Finally, we consider the IOM as source of organics. (Huang et al., 2007) argue that monocarboxylic acids and other small organics could be produced by the hydrothermal processing of IOM. Observations that might be taken as evidence of this idea include: 1) Correlations of the $\delta^{13}C$ values of monocarboxylic acids with their carbon numbers are similar to those for moieties from the IOM; and 2) our measurements demonstrate that the IOM has an isotopic composition similar to the ^{13}C pool that was the source of the C-1 and C-3 sites of alanine, perhaps suggesting alanine is also formed by hydrolysis of IOM. This second observation could be understood in the context of the model we present if the IOM and alanine's C-1 and C-3 sites both derive from a primordial low ^{13}C pool (*i.e.*, hydrocarbons and HCN). If, instead, alanine was made from hydrolysis of the IOM, it is not obvious how it would have acquired such an extraordinarily high $\delta^{13}C$ value in its C2 carbon site without evidence of enrichment in the C1 and C3 sites. We are aware of no high ^{13}C chemical moieties of the IOM that could readily explain this finding, and so we believe this idea could not be developed to provide a satisfactory explanation of this study's results.

Nevertheless, future compound- and site-specific measures may be able to identify IOM processing as a source of soluble organics in Murchison (and perhaps other carbonaceous chondrites). The site-specific $\delta^{13}C$ isotope ratio for compounds produced by IOM processing should mirror those found in the IOM aliphatic side chains (which have compound specific molecular average $\delta^{13}C$ values of 57.9 ‰ to 0.4 ‰). In contrast, the reaction network we propose predicts that the terminal carboxyl (C-1) sites of the carboxylic acids will be highly ^{13}C enriched compared to all other CH_x sites.

B6. Parent-Body Organic Reaction Model

Constraints on the Isotope Effects Associated with Syntheses

To calculate the $\delta^{13}C$ values of alanine precursors and organic synthesis products other than alanine, isotope effects of different synthetic steps were collated from literature review and those for Strecker synthesis were measured via experimental work conducted as part of this study. Isotope effects for Strecker synthesis were further validated by comparison to literature values for isotope effects from similar reaction mechanisms.

The reduction of aldehydes into imines via reductive amination has a maximum measured isotope effect of 0.6 ‰ (Billault et al., 2007), which is lower than our measurement errors so was treated as a 0 ‰ fractionation in the model. Studies for carbon isotope effects during the oxidation of aldehydes have observed a range of effects from negligible (aldehyde to thiohemiacetal conversion) (Canellas and Cleland, 1991) to large deuterium isotope effects that suggest possible concurrent carbon isotope effects (Wiberg, 1954); although these have not been measured. To consider both possibilities, we consider two endmember cases of 1) no isotope effect and 2) a 30‰ normal kinetic isotope effect, similar to intrinsic KIE's associated with other carbon oxidation reactions (Cleland, 2005). Mechanisms and associated isotope effects are portrayed in Figure 3.3 in the main text. Differences in our solution between the 0 ‰ carbonyl oxidation KIE and the 30 ‰ normal KIE case are depicted in Figure 3.4 in the main text.

Experimental work was conducted to constrain the isotope effects in Strecker synthesized alanine from ammonium chloride, acetaldehyde, sodium cyanide, and water at temperatures ranging from 20°C to 25°C for the creation of the aminonitrile and 80°C to 120°C for its acid hydrolysis. We measured the average isotopic composition of solid reagents and products via EA-IRMS, of acetaldehyde via combustion over CuO into CO_2 which was measured on a dual-inlet IRMS, and the site-specific isotopic composition of alanine produced by the synthesis was measured for $\delta^{13}C$ of the C-2 + C-3 (140.032 fragment) on the Orbitrap as described above. Our measurements indicated that the

average $\delta^{13}C$ of C-2 and C-3 of alanine produced by Strecker synthesis (-30.6 ± 0.9 ‰) is approximately 12‰ depleted in ^{13}C relative to the reactant acetaldehyde ($\delta^{13}C$ = -19.1 ‰) regardless of yield. Because C-3 does not participate in the Strecker reaction, we assumed the difference in the average $\delta^{13}C$ for C-2 and C-3 is due to a -24 ‰ isotope effect on C-2, which is consistent with other CN addition reactions (Lynn and Yankwich, 1961). C-1 (found by a subtraction of C-2 and C-3 from the molecular average) exhibited a normal KIE that had an average value of 22 ‰ for alanine produced between a 10 % and 55 % yield (-54.1 ± 3.2 ‰ relative to a starting CN $\delta^{13}C$ of -31.8 ± 0.2 ‰). This KIE also agrees with literature values for amide oxidation (Robins et al., 2015).

Our reaction network model assumes a low yield of products and unlimited supply of reactants relative to the products such that isotope effects would be apparent in products and but would not significantly alter the $\delta^{13}C$ of the reactants (and, consequently, other compounds produced from them). The agreement between our predicted isotope ratios and measurements in literature, particularly for acetaldehyde and HCN, is consistent with this assumption. However, below we analyze the possibility that variations in certain factors would impact our results:

Temperature: The isotope effects associated with reactions in our hypothesized reaction network range up to 30 ‰. Given that the temperatures of aqueous alteration of the CM chondrites have been demonstrated to have varied between 20 and 71 ˚C (293.15 – 344.15 K (Guo and Eiler, 2007)) through clumped isotope thermometry, and given that chemical isotope effects commonly exhibit approximately linear variations in amplitude with $1/T^2$, we estimate that these model estimates could have varied by several per mil. For moderate variations in reaction progress (below), these should lead to variations of just a few per mil in predicted $\delta^{13}C$ values of products. This is comparable to full procedural analytical precision and less than otherwise unexplained variability in the data, and so we consider it insignificant (in the context of the constraints and goals of our model).

Reaction progress: Our model presumes that essential reactants (water, aldehydes, ammonia and HCN) are more abundant than products that are created in our reaction network. If the proportions of these compounds in the Murchison parent body initially resembled those in comets (*e.g.*, (Biver et al., 2019)), this assumption would be well justified. However, if organic synthesis reactions such as the Strecker chemistry locally went to near completion (consuming most of reactants), isotope effects associated with synthesis reactions would be mitigated, as isotopic proportions in products would approach those of reactants. The largest kinetic isotope effects associated with our reaction network model (30 ‰) could be diminished in this way—and in the extreme limit of quantitative yield, reduced to nothing.

The limits one should place on this argument are difficult to evaluate because all of the reactants are more volatile than the products (*e.g.*, alanine is essentially involatile whereas its proposed substrates, acetaldehyde and HCN have boiling points of 20 and 26˚C, respectively). Thus, the abundance ratios of aldehydes to amino acids in the Murchison meteorite are likely a poor guide to their proportions early in the history of the Murchison parent body. If the synthesis chemistry had yields comparable to laboratory Strecker synthesis (10's of %), then the effective KIE's would be approximately halved, or reduced by approximately 10 ‰. That would be degrade the level of agreement between our model prediction and the measured $\delta^{13}C$ of some compounds in our model (and improve the level of agreement for others), but by amounts that are a small fraction of the isotopic variations (*i.e.*, site-specific and intermolecular differences) that motivate our model. We therefore consider it implausible that this factor significantly impacts the overall reasonableness of our model.

Alanine destruction: Free and total alanine in Murchison are about one-third as abundant as in the most alanine-rich CM chondrite (~0.20 and ~0.65 ppm, respectively), implying that it could be residual to 10's of % destruction. If this destruction was accompanied by a ^{13}C kinetic isotope effect in the range typical of irreversible organic reactions (~10-30 ‰) and operated on one or two atomic sites, then the residual alanine could have been enriched in $\delta^{13}C$ by several per mil up to perhaps 10 ‰. The most likely mechanisms for

alanine destruction (NH_2 replacement with OH, or decarboxylation) should either enrich the C-2 site or enrich both the C-1 and C-2 sites equally in the residue. These effects are less than or just at the margin of the level of significance addressed by our model and are a small fraction of the 150 ‰ site-specific effect our model was tailored to describe. Moreover, the $\delta^{13}C$ values of alanine from CM chondrites do not exhibit an inverse concentration with their concentration in the samples, so there is no empirical evidence to suggest such a fractionating loss mechanism. We conclude loss of alanine through these side reactions is unlikely to significantly impact our conclusions.

Calculation of reactant $\delta^{13}C$ values

To estimate the site-specific $\delta^{13}C$ values of reactants in our network model, we subtracted site-specific isotope effects constrained by our Strecker synthesis experiments from the measured $\delta^{13}C$ values for alanine in the analytical Murchison sample. Based on these results, the reactant CN is estimated to have a $\delta^{13}C_{VPDB}$ value of -7 ‰ and the initial acetaldehyde is estimated to have $\delta^{13}C_{VPDB}$ values of 166 ± 10‰ and -36 ± 10 ‰ for the carbonyl ($C\text{-}1_{acetaldehyde}$) and methyl ($C\text{-}2_{acetaldehyde}$) carbons, respectively. Combining our results with the ISM chemical networks described in (Elsila et al., 2012) and references therein, we predict that the carbonyl carbon in all aldehyde functional groups are from the ^{13}C-enriched CO pool in the ISM and that all alkyl carbons are from another, ^{13}C-depleted pool (that include C_xH_y compounds). Thus, in our model we assigned $\delta^{13}C$ values of 166 ± 10 ‰ to all carbonyl carbons and -36 ± 10 ‰ to all alkyl carbons. Equivalently, we calculated the molecular-average $\delta^{13}C$ values of aliphatic aldehydes with two or more carbons by calculating the carbon-weighted average values of acetaldehyde (64.6 ± 1.5 ‰) and additional aliphatic carbons (-36 ± 10 ‰) (Eqn. S3.2; See Appendix B.3).

$$^{13}F_{Cx\text{-}aldehyde} = \left(\frac{2}{x}\right)^{13}F_{molec\ avg,\ acetaldehyde} + \left(\frac{x-2}{x}\right)^{13}F_{C\text{-}2,\ acetaldehyde} \quad \textit{(Eqn. S3.2)}$$

where x is the carbon chain length and C_x-aldehyde is a molecule with one aldehyde carbon and x methylene carbons. All such calculations are made using ^{13}C mole fraction

("fractional abundance") rather than $\delta^{13}C$ values to avoid systematic errors arising from non-linearities of the δ scale.

In our model, amines form from a reactant aldehyde's reductive amination (Figure 3.3, main text), which is proposed to have an insignificant KIE, so we estimated that the $\delta^{13}C$ value of the amine molecule is equal to that of an aldehyde molecule with the same carbon backbone (See Dataset S2). Monocarboxylic acids formed from the oxidation of aldehyde precursors were assigned to have isotope effects that range from 0 ‰ to -30 ‰. In the first case, the product carboxylic acids have $\delta^{13}C$ values equal to their aldehyde precursors (See Dataset S2). For the alternate case of a fully expressed -30 ‰ KIE during oxidation of the aldehyde's carbonyl site, the isotope effect is assumed to only occur on the C-1 carbon, so the molecular-average $\delta^{13}C$ for acetic acid was calculated accounting for the isotope effect only occurring on this site (Eqn. A3). Higher carbon chain carboxylic acids (C_2 and above) were calculated as the carbon-weighted average values of acetaldehyde (64.6 ± 1.5 ‰) and additional CH_x groups (-36 ± 10 ‰) to decrease error (Eqn. *S3*.4).

$$^{13}R_{\text{molec avg, acetic acid}} = (1\text{-}0.050/2)\ ^{13}R_{\text{molec avg, acetaldehyde}} \qquad \textit{(Eqn. S3.3)}$$

$$^{13}F_{\text{C-x-carboxylic acid}} = \left(\frac{2}{x}\right)\ ^{13}F_{\text{molec avg, acetic acid}} + \left(\frac{x-2}{x}\right)\ ^{13}F_{\text{C-2, acetaldehyde}} \qquad \textit{(Eqn. S3.4)}$$

All α-amino acids (*i.e.* not only alanine) were assumed to undergo fractionation in Strecker synthesis as described above. Because our analytical Murchison alanine measurements include $\delta^{13}C$ for sites that have undergone the same fractionations associated with their synthesis (*e.g.*, the C-1 and C-2 carbons of all alpha amino acids formed by Strecker synthesis are predicted to be fractionated in the same way we predict for our model of alanine formation), we used alanine's site-specific isotopic composition as our building blocks for other amino acids. Glycine's $\delta^{13}C$ was predicted based on the 184.021 *m/z* fragment measurement (corrected for dilution with carbons from derivatizing agents) and alanine was assigned to have the $\delta^{13}C$ value directly measured in this study, 25.5 ‰ (*e.g.*, it is not predicted but serves as the basis for predicting other species, particularly acetaldehyde and HCN). All amino acids with longer alkyl chains

than alanine were assumed to have additional alkyl carbons (*i.e.,* with a $\delta^{13}C$ equal to that of C-3 in alanine) comprising the balance of the molecular carbon inventory (Eqn. S3.5).

$$^{13}F_{Cx\text{-amino acid}} = \left(\frac{3}{x}\right)^{13}F_{molec\ avg,\ alanine} + \left(\frac{x-3}{x}\right)\ ^{13}F_{C3,\ alanine} \textit{(Eqn. S3.5)}$$

In addition to Strecker synthesis, we also considered the possibility that C-1 in amino acids could equilibrate with the dissolved inorganic carbon (DIC) pool on meteorites (*e.g.*, the carbonate pool). The DIC pool is 3000 times more abundant than all amino acids on Murchison combined (Sephton, 2002). Consequently, in the case of equilibration between the two reservoirs, the $\delta^{13}C$ value of DIC would control that of C-1 in amino acids. We assumed a DIC reservoir with a $\delta^{13}C$ of 80 ‰, equal to the highest measured literature value for CM chondrites (Sephton, 2002) (and thus the maximum effect on amino acids with which it equilibrates). Using ε values for CO_3^{2-}-CO_2 and CO_2-amino acid carboxyl group equilibration from (Rustad et al., 2008) and (Rustad, 2009) respectively, we predicted the $\delta^{13}C$ of different amino acids on Murchison that had equilibrated with its carbonate pool (Dataset S2). Of amino acids with molecular-average $\delta^{13}C$ values measured on Murchison, only glycine and alanine also have ε values for CO_2 and amino acid carboxyl group in (Rustad, 2009). These values are 4.4 ‰ and 4.9 ‰, so we adopted an average value of 4.65 ‰ for $\varepsilon_{CO2\text{-amino acid C-1 site}}$ in our calculations for all amino acids.

Appendix C: Supplemental Tables and Figures for Chapter 4

Sample (p*mol ala/µL*)	Procedural Blank #	First Step	% of signal relative to Murchison			*Max pmol ala/µL*
			$\Sigma^{12}C$ Counts	$\Sigma^{13}C$ Counts	$\Sigma^{12}C$ Intensity	
Methods Development Murchison, January 2018 (29)	Hexane	Pure Hexane	0.1%	0.0%	0.0%	0.11
	1	Transfer of Meteorite to GC vial Run Prior to Derivatization	0.5%	0.1%	0.2%	0.45
	2	Extraction of meteorite at NASA	0.8%	0.2%	0.2%	2.4
	3	Transfer of Meteorite to GC vial	0.3%	0.0%	0.1%	1.0
Methods Development Murchison, March 2018 (29)	Hexane	Pure Hexane	0.0%	0.1%	0.0%	3.1×10^{-3}
	1	Transfer of Meteorite to GC vial Run Prior to Derivatization	0.2%	0.1%	0.1%	5.4×10^{-2}
	2	Extraction of meteorite at NASA	1.0%	0.0%	0.3%	0.29
	3	Transfer of Meteorite to GC vial	5.6%	3.4%	3.7%	1.6
Analytical Murchison, Summer 2018 (15)	Hexane	Pure Hexane	0.0%	0.0%	0.0%	1.3×10^{-3}
	1	Transfer of Meteorite to GC vial Run Prior to Derivatization	0.0%	0.0%	0.1%	1.8×10^{-3}
	2	Extraction of meteorite at NASA	1.0%	0.0%	0.2%	0.16
	3	Transfer of Meteorite to GC vial	0.7%	0.0%	0.1%	9.8×10^{-2}
	4	Chemical Derivatization	0.5%	0.0%	0.0%	6.9×10^{-2}

Table S4.1. Blank IDs and possible contamination to Murchison.

Analysis Set	Date	Scan Range	% of signal relative to alanine		Maximum ‰ change in ^{13}R
			$\Sigma^{12}C$ Counts	$\Sigma^{13}C$ Counts	
Methods Development Murchison, January 2018	1/3	135-146 *RE*	0.08%	0.05%	-0.061
	1/3	135-146 *RE*	0.31%	0.29%	-0.037
	1/3	135-146 *RE*	0.16%	0.14%	-0.027
	1/4	50-300 *DE*	0.00%	0.00%	0.000
	1/6	50-300 *DE*	0.04%	0.00%	-0.091
	1/7	50-300 *DE*	0.09%	0.00%	-0.086
	1/7	50-300 *DE*	0.31%	0.10%	-0.210
	1/7	181-187 *RE*	0.00%	0.05%	0.048
Methods Development Murchison, March 2018	3/26	50-500 *DE*	0.0%	0.0%	-0.002
	3/26	50-500 *DE*	0.0%	0.0%	-0.007
	3/27	135-146 *RE*	0.0%	0.0%	0.000
	3/29	181-187 *RE*	0.0%	0.2%	0.199
	3/30	181-187 *RE*	0.0%	0.3%	0.282
	3/31	181-187 *RE*	0.0%	2.8%	2.756
Analytical Murchison, Summer 2018	6/27	135-146 *RE*	0.12%	0.12%	-0.002
	6/27	135-146 *RE*	0.12%	0.11%	-0.010
	6/27	135-146 *RE*	0.16%	0.16%	-0.002
	7/12	181-187 *RE*	0.52%	0.49%	-0.033
	7/12	181-187 *RE*	0.00%	0.07%	0.045
	7/12	181-187 *RE*	0.04%	0.01%	-0.033
	7/13	181-187 *RE*	0.05%	0.04%	-0.013
	7/16	108-118 *RE*	5.52%	6.20%	0.644

Table S4.2. Background from blanks during measurement. RE denotes samples analyzed using Reservoir Elution mode while DE denotes samples analyzed using Direct Elution mode.

Analysis Set	Sample	Analysis Number used	
		140	**184**
Winter 2018	Alfa Aesar	28, 29	47-49, 51-53, 61-62
	Strecker	30	64-67
	Methods Development Murchison	24-25	45-46, 55-59
Spring 2018	Alfa Aesar	75, 83	94, 97-99, 104
	Strecker	74, 80	x
	Methods Development Murchison	78, 84-86	101-103, 105-106
Summer 2018	Alfa Aesar	128, 131, 134-137	162, 166-167
	Strecker	139-140	159, 161
	Analytical Murchison	142, 144, 147	164-165

Table S4.3. Analysis numbers of measurements used to calculate ^{13}R values from samples and standards. Analysis numbers correspond to values found in Dataset S1 and were used to calculate ratios found in Table S2.

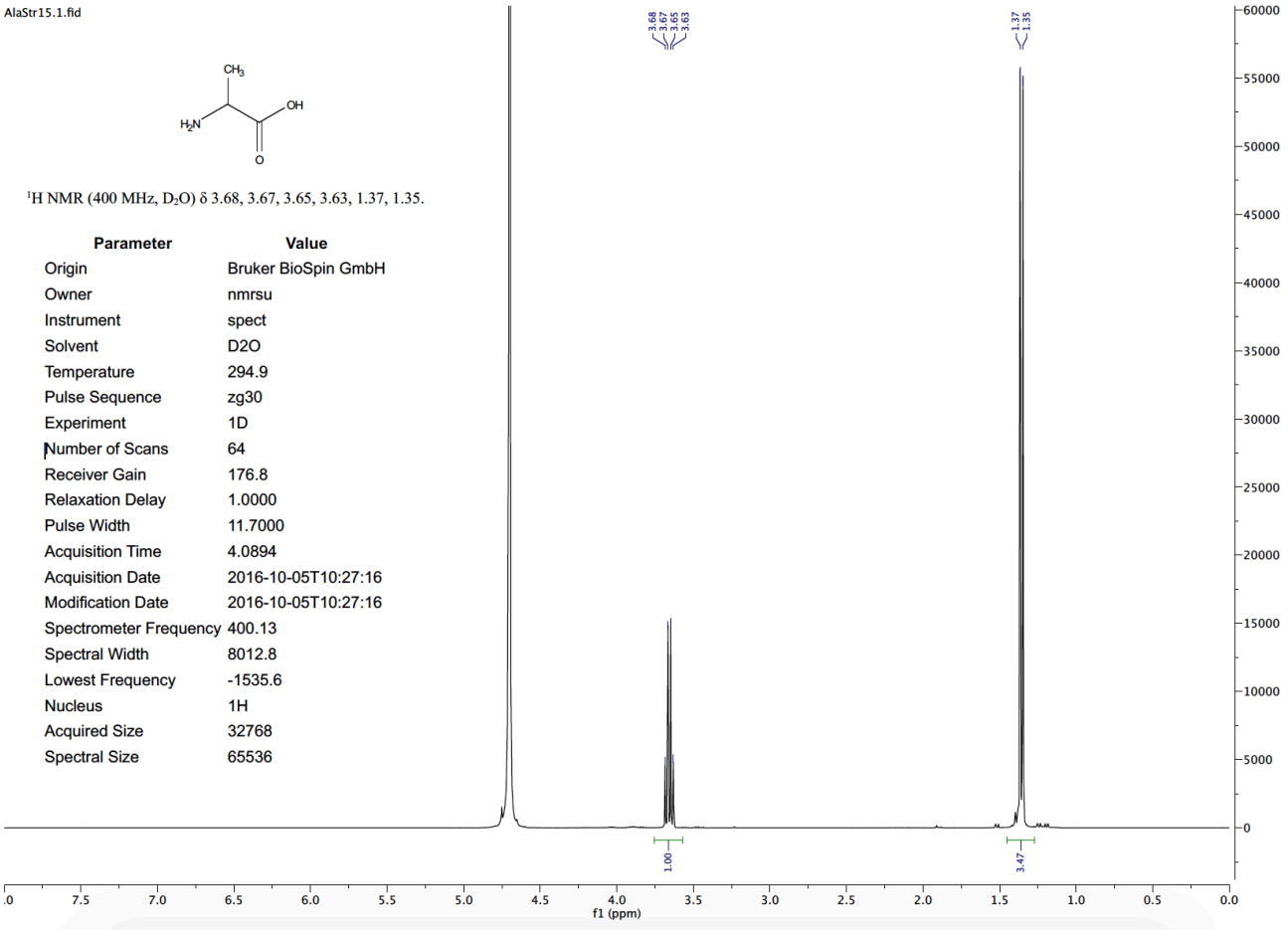


Figure S4.1: NMR of Strecker Alanine standard.

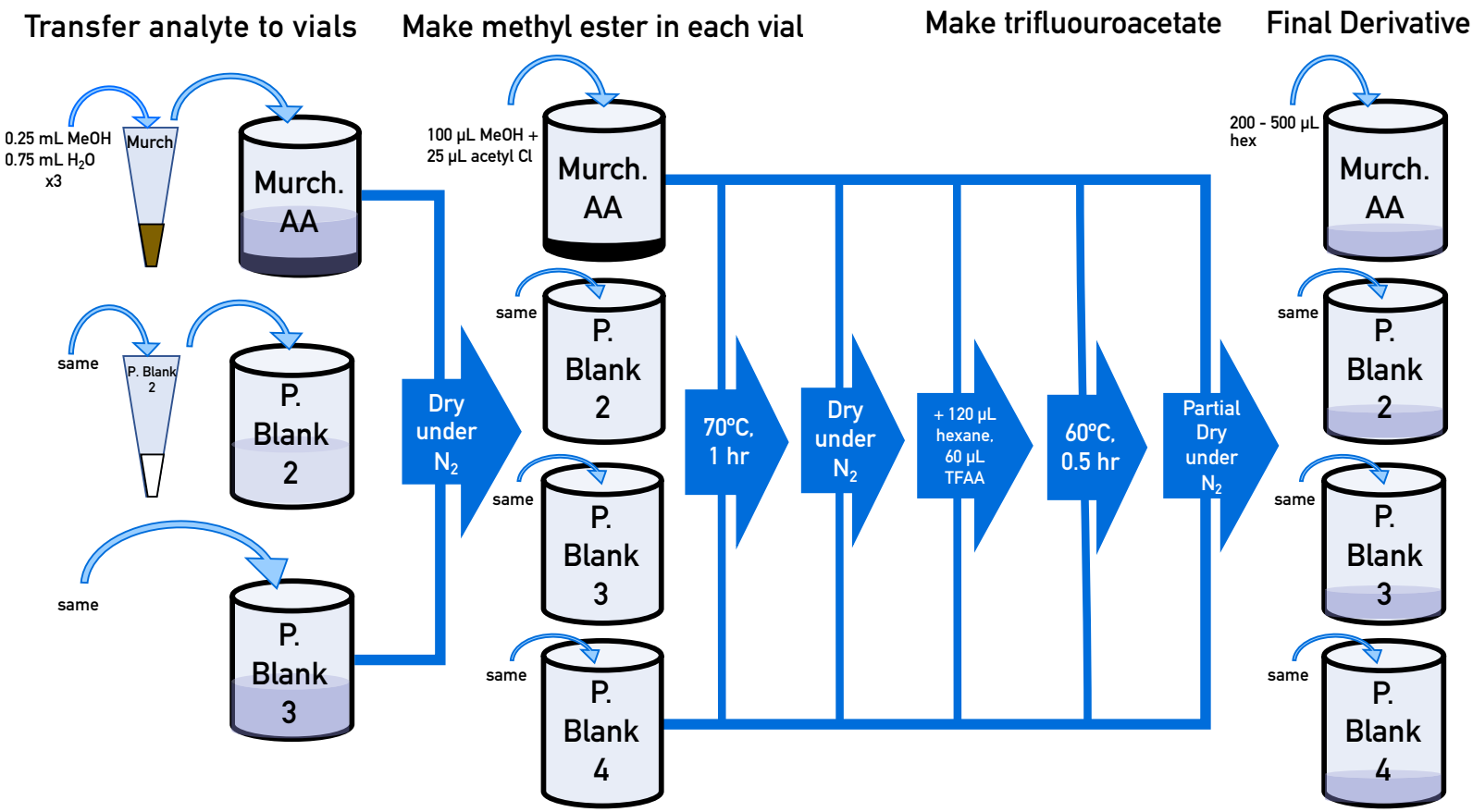
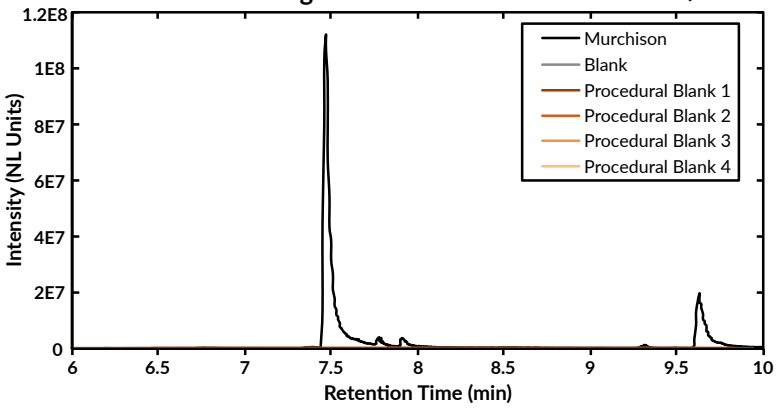
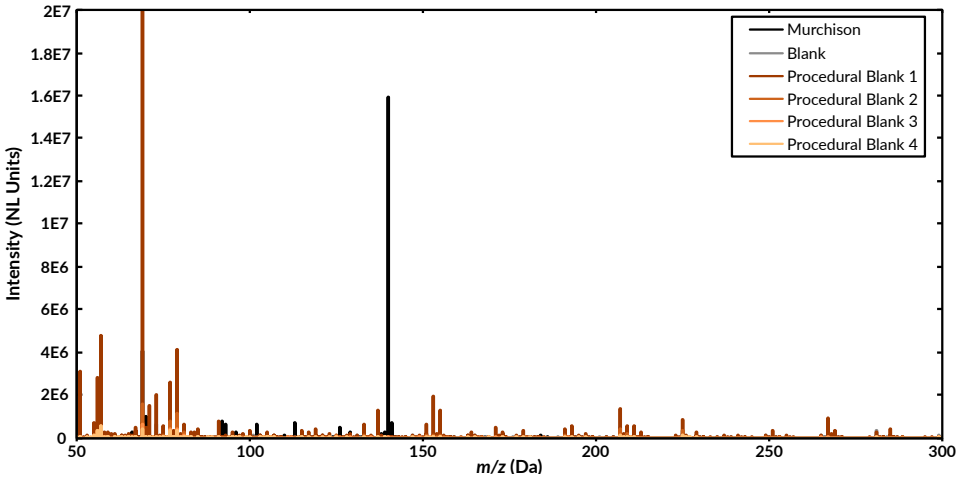


Figure S4.2: Steps in derivatization of samples, standards, and blanks.

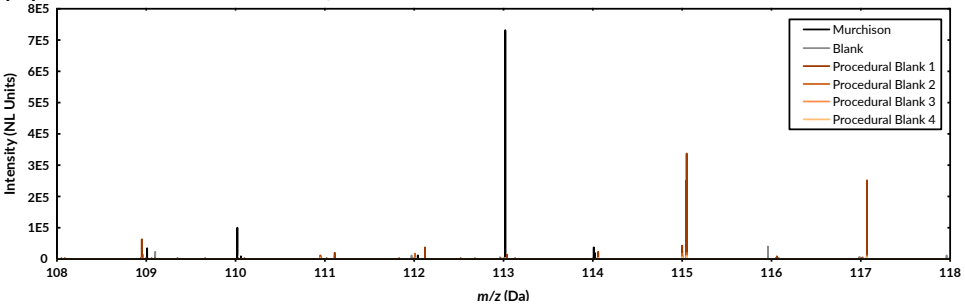
(a) Direct Elution Chromatograms for Murchison and Blanks *m/z* 140.032



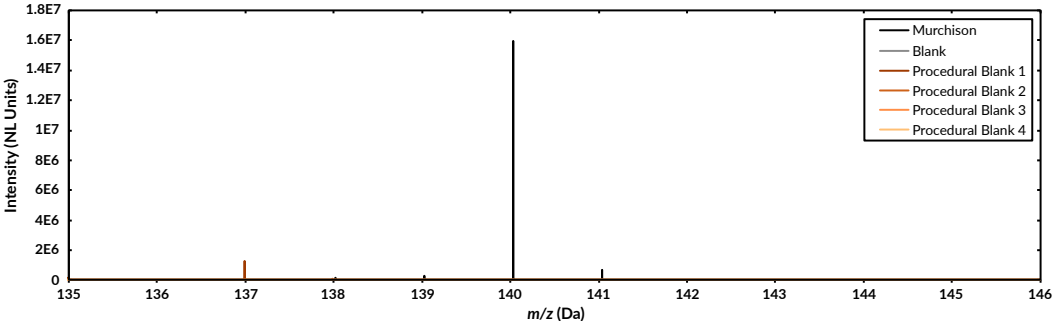
(b) Direct Elution Mass Spectra for Murchison and Blanks (Retention time 7.40 to 7.70 min)



(c) Direct Elution Mass Spectra for Murchison and Blanks (Retention time 7.40 to 7.70 min)



(d) Direct Elution Mass Spectra for Murchison and Blanks (Retention time 7.40 to 7.70 min)



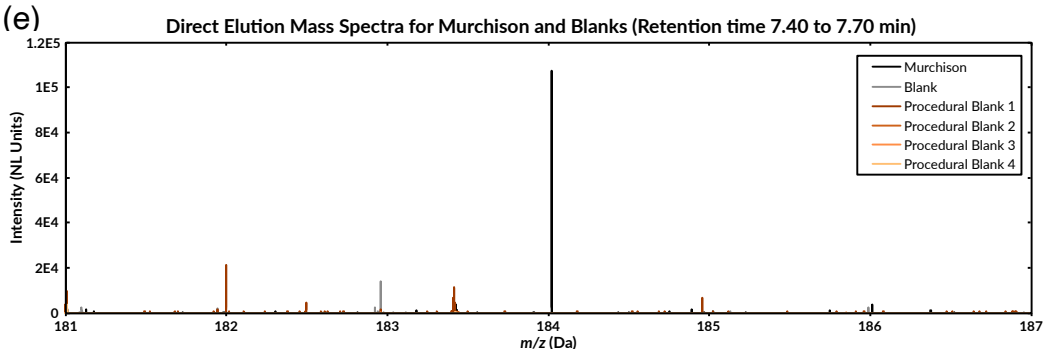


Figure S4.3: Gas chromatograms and mass spectra for the analytical Murchison sample. (a) Mass spectra from *m/z* 50-300 for Murchison sample and blanks. (b) Chromatogram for the *m/z* 140.032 trace (the main fragment for alanine) for Murchison sample and blanks from 5 to 10 minutes. (c) Mass spectra from *m/z* 108-118 (113 fragment measurement window) for Murchison sample and blanks. (d) Mass spectra from *m/z* 135-146 (140 fragment measurement window) for Murchison sample and blanks. (e) Mass spectra from *m/z* 181-187 (184 fragment measurement window) for Murchison sample and blanks. Retention time for the mass spectra capture the alanine derivative's elution time.

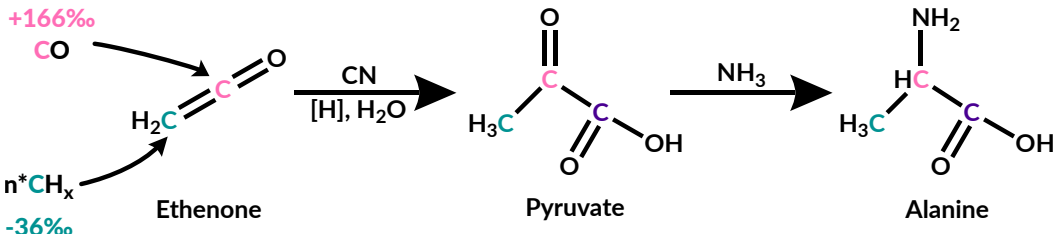


Figure S4.4: Ketene reaction mechanism to create alanine with a similar carbon isotope structure to the one measured in this paper. Pyruvate is created in the ISM and can react with NH_3 either in the ISM or on a parent body in the solar nebula.

Dataset 1: Values for ^{13}R, ^{12}C counts, and ^{13}C counts of each analysis. ^{13}R values are a weighted average for the entire run and only include scans for which the monoisotopic and singly ^{13}C-subsituted peaks are present. The ^{12}C counts and ^{13}C counts are summed from all scans including those that only contained monoisotopic or singly ^{13}C-subsituted peaks. Further discussion of culling procedures for the ^{13}R can be found in the Site-Specific Isotope Analysis and Blanks sections. Available for download at: https://ars.els-cdn.com/content/image/1-s2.0-S0016703720305901-mmc4.xlsx

Dataset 2: Predictions of molecular-average $\delta^{13}C$ values for prebiotic compounds. Literature values are from (*a*) (Pizzarello, 2014), (*b*) (Simkus et al., 2019), (*c*) (Aponte et al., 2019), (*d*) (Aponte et al., 2016), (*e*) (Yuen et al., 1984), (f) (Huang et al., 2005), (*g*) (Aponte et al., 2019b), *(h*) (Engel et al., 1990), (*i*) (Pizzarello et al., 2004), (*j*) (Elsila et al., 2012). Alternate predictions refer to a fully expressed 19 ‰ normal KIE for carboxylic acids and DIC equilibration with the carboxyl site for amino acids. Data in which cis-trans isomerism was measured separately, data from cis-isomers are denoted with (*Z*) and trans are denoted with (*E*). Available for download at: https://ars.els-cdn.com/content/image/1-s2.0-S0016703720305901-mmc5.xlsx

Appendix D: Supplemental Tables for Chapter 5

Table S5.1: The Murchison δD moiety model compounds, moieties assigned, measurements used, and predictions is at doi: 10.22002/D1.1662

Table S5.2: The δD moiety model compounds, moieties assigned, measurements used and predictions for Murchison compared to other meteorites is at doi: 10.22002/D1.1662

	Median values for compounds in Murchison				
Compound Class	Boiling Point (deg C)	Melting Point (deg C)	Kow	Solubility (Kg/kg)	Topological Polar SA (A^2)
Monocarboxylic α-Amino Acid	214	277	-2.65	0.204	63.3
Dicarboxylic α-Amino Acid	329	226	-3.1	0.0381	101
Monocarboxylic non-α-Amino Acid	241	203	-2.7	0.5	63.3
Hydroxy Acid	245	57	-0.4	0.484	57.5
Sulfonic Acid	184	7.5	-0.25	1	62.8
Monocarboxylic Acid	219	-9.25	2.2	0.00416	37.3
Polyaromatic hydrocarbon	357.5	130.5	4.65	7.45E-07	0

Table S5.3: Median values of key physical properties for the compounds studied from the Murchison meteorite. To arrive at these values, we take values for the physical property of interest for each compound in a class (see table S1 for compound classes) and then calculate the median of each set to get the median value for the compound class. Values are taken from PubChem (https://pubchem.ncbi.nih.gov) when possible and from predictions in ChemSpider (https://www.chemspider.com) when no data are available on PubChem.

Tables S5.4 -S5.10: The measured $\delta^{13}C$ values and predicted $\delta^{13}C$ ranges for all compounds in the modeled syntheses. The bottom of each table has an analysis of the model's performance including the number of compounds predicted, % of compounds predicted within 2 standard deviation of measurement and 5 ‰ of that, the average residual for each model, and similar data about the compounds that were predicted by all models for a meteorite. Tables S4 to S10 are for Murchison, ALH 83100, LEW 90500, LON 94101, MIL 090001, LAP 02342, and GRA 95229. These tables are available for download at doi: 10.22002/D1.1661.