Supporting Information for “Inverse carbon isotope patterns of lipids and kerogen record heterogeneous primary biomass”

Hilary G. Close1*, Roderick Bovee1, Ann Pearson1*

1Department of Earth & Planetary Sciences, Harvard University, Cambridge MA 02138

*Correspondence: hclose@fas.harvard.edu; pearson@eps.harvard.edu

Case 1: Simple uniform degradation through multiple trophic levels

We begin with the simplest case, where all organic matter degrades to a total extent (or exposure time), \( d \), with a rate constant \( k \). The amount of preserved (“sedimentary”) organic matter from a given trophic level \( N \), as a function of \( P_0 \), or the starting amount of total primary photosynthetic organic matter, is

\[
P_N(d) = e^{-Kd}P_0E^N(1 - e^{-Kd})^N
\]

(1)

where primary OM is represented as trophic level \( N = 0 \), heterotrophy proceeds from trophic levels \( N = 1 \) through infinity or some derived terminal level \( T \), \( E \) is the heterotrophic efficiency (fraction of OM incorporated into \textit{de novo} biomass rather than respired as \textit{CO}_2), \( e^{-Kd} \) is the fraction of biomass that is preserved at each trophic level, and \( 1 - e^{-Kd} \) is the fraction of biomass that is transferred to the next trophic level. The amount of biomass in the cells of a trophic level \( N \), before being degraded further by trophic level \( N + 1 \), is thus \( P_0E^N(1 - e^{-Kd})^N \).

The total preserved organic matter from all trophic levels then is

\[
P(d) = e^{-Kd}P_0 \sum \frac{P_N(d)}{P_0}E^N(1 - e^{-Kd})^N
\]

(2)

We can determine the final, net isotopic value for preserved organic matter simply by multiplying the fractional contribution of biomass from each trophic level, \( \frac{P_N(d)}{P_0} \), by its isotopic value, which is determined by compounded heterotrophic isotopic enrichment from the primary signature \( \delta_{P_0} \).
\[ \delta_P(d) = \sum \frac{P_N(d) \left( \delta_{P_0} + hN(1-E) \right)}{P(d)} \]  

where \( h \) is the maximum isotopic enrichment at each trophic level and actual enrichment per trophic level depends on the heterotrophic efficiency \( E \) as \( h(1-E) \).

In calculating the resulting isotopic content of net preserved organic matter according to these equations, over a range of exposure times \( d \) and using the parameter values in Table 1 of the main document, we find that heterotrophic biomass contribution is substrate-limited enough that the \( \delta^{13}C \) value of total preserved biomass is only enriched slightly in \( ^{13}C \) due to heterotrophic processing, in comparison to starting primary organic matter (Supporting Figure 1). Because no distinction is made between degradation rates for different biochemical classes, the amount and isotopic content of all compounds types will relate to the trophic level at which they were biosynthesized. Therefore, the \( \Delta \delta_{n-alkyl-TOC} \) value will always reflect that of biosynthetic product.

Case 2a: Heterotrophic enrichment hypothesis- Distinct degradation rates for compound classes

Starting again with total photosynthetic biomass \( P_0 \), we now account for differences in degradation rate constants for different compound classes. Photosynthetic biomass is modeled in three components: resistant biopolymer (R), lipid (L), and other, labile biomass (B), which are biosynthesized in photosynthetic cells (N=0) in the corresponding proportions \( F_R, F_L, \) and \( F_B \), such that \( F_R + F_L + F_B = 1 \). Heterotrophic cells are modeled as containing only lipid and labile biomass \( (F_R = 0) \). We will use the notation \( F_X \) to refer to generalized heterotrophic (H) components in the following formulations, and simply stipulate that for trophic levels \( N = 1 \) and above, \( F_R = 0 \), \( F_L = F_{HB} \) and \( F_B = F_{HL} \). Total biomass synthesized in a given trophic level has the isotopic signature \( \delta_{P_N} \), and isotopic values for resistant biopolymer \( (\delta_R) \) and lipid \( (\delta_L) \) are calculated by subtracting biosynthetic fractionations \( \varepsilon_R \) and \( \varepsilon_L \), respectively, from \( \delta_N \). The isotopic value of the remaining labile biomass \( (B) \) is calculated by difference from the isotopic mass balance

\[ \delta_{P_N} = F_{R} \delta_B + F_{L} \delta_L + F_{R} \delta_R \]  

If each class of compounds also has a distinctive degradation rate \( K_X \), the amount of preserved organic matter from primary producers alone \( (N = 0) \) is

\[ P_0(d) = P_0 \left( F_R e^{-K_R d} + F_L e^{-K_L d} + F_B e^{-K_B d} \right) \]  

and its isotopic content can be calculated as

\[ \delta_{P_0}(d) = \frac{F_{R} e^{-K_R d} \delta_R + F_{L} e^{-K_L d} \delta_L + F_{B} e^{-K_B d} \delta_B}{F_{R} e^{-K_R d} + F_{L} e^{-K_L d} + F_{B} e^{-K_B d}} \]
Accounting for fractional contribution and different degradation rates of the compound classes, as well as the presence of resistant biopolymer \( R \) only in trophic level \( N = 0 \), the analog for Equation 1 above, for trophic levels \( N = 1 \) and above only, is

\[
P_N(d) = P_0 E^N \left( F_T e^{-K_L d} + F_B e^{-K_B d} \right) \left[ 1 - F_T e^{-K_L d} - F_B e^{-K_B d} \right]^{N-1} \left( 1 - F_R e^{-K_R d} - F_L e^{-K_L d} - F_B e^{-K_B d} \right)
\]

for \( N = 1 \) to \( T \).

The total preserved organic matter, starting from \( N = 0 \) (primary producers) is then

\[
P(d) = \sum P_N(d)
\]

The isotopic signature of organic matter preserved from each heterotrophic level will depend on the organic matter consumed, the metabolic efficiency and corresponding net isotopic enrichment, isotopic fractionation during biosynthesis of different compound classes, and the differential loss of these classes to degradation by higher trophic levels. Before degradation by the trophic level \( N + 1 \), we know that isotopic value is inherited from the degraded portion of the previous trophic level and enriched such that

\[
\delta P_N = \delta P_{(N-1)Deg} + h(1 - E)
\]

where \( \delta P_{(N-1)Deg} \) can be calculated by difference from \( \delta P_{(N-1)} \) and \( \delta P_{(N-1)}(d) \), and the second term is the isotopic enrichment by a single heterotrophic level \( N \).

Lipid in trophic level \( N \) has an isotopic value of

\[
\delta_T = \delta P_{(N-1)Deg} + h(1 - E) - \varepsilon_L
\]

The isotopic value for the more labile components of biomass, by difference, is then

\[
\delta_B = \left[ \delta P_{(N-1)Deg} + h(1 - E) \right] (1 - F_T) + F_T \varepsilon_L
\]

After degradation by trophic level \( N + 1 \), the residual biomass is preserved in new proportions according to the different rates of decay; the amount of preserved organic matter from a given compound class \( X \) at trophic level \( N \) and after degradation \( d \) is

\[
X_N(d) = P_0 E^N F_T e^{-K_L d} \left[ 1 - F_T e^{-K_L d} - F_B e^{-K_B d} \right]^{N-1} \left( 1 - F_R e^{-K_R d} - F_L e^{-K_L d} - F_B e^{-K_B d} \right)
\]
The fractional contribution to total preserved organic matter from a compound class $X$ is thus $X_{P_N}(d) / P_N(d)$; therefore the net isotopic value for total preserved organic matter from trophic level $N$ after degradation $d$ is

$$\delta_{P_N}(d) = \frac{B_N(d)}{P_N(d)} \delta_B + \frac{L_N(d)}{P_N(d)} \delta_L$$

(13)

For the next trophic level, we can calculate the isotopic value of biomass that is degraded from trophic level $N$ by difference, as mentioned:

$$\delta_{P_{N+1}} = \frac{F_B(1 - e^{-K_Bd})\delta_B + F_L(1 - e^{-K_Ld})\delta_L}{1 - F_Be^{-K_Bd} - F_Le^{-K_Ld}}$$

(14)

for $N = 1$ to $T$.

This formula holds for $N = 1$ and above. The isotopic value of primary organic matter consumed by the first heterotrophic level differs slightly, since it must account for resistant biopolymer from primary producers:

$$\delta_{P_Deg} = \frac{F_R(1 - e^{-K_Rd})\delta_R + F_L(1 - e^{-K_Ld})\delta_L + F_B(1 - e^{-K_Bd})\delta_B}{1 - F_Re^{-K_Rd} - F_Le^{-K_Ld} - F_Be^{-K_Bd}}$$

(15)

Thus we can successively derive both amounts and isotopic values for total organic matter and lipid preserved from each trophic level. After summing the total amount preserved, we then calculate the fractional contribution of each trophic level, and create a weighted average of isotopic values in order to derive the final “sedimentary” isotopic value for TOC and lipid:

$$\delta_P(d) = \sum \frac{P_N(d)\delta_{P_N}(d)}{P(d)}$$

(16)

$$\delta_L(d) = \sum \frac{L_N(d)\delta_{L}(d)}{L(d)}$$

(17)

**Case 2b: Logan hypothesis- Attenuation of degradation effects in higher trophic levels**

For calculations involving an attenuation of exposure time for increasing trophic levels, amounts of heterotrophic material can no longer be expressed as a simple sum: these must also be calculated successively and summed. Where $D(d, N)$ is some function of exposure time $d$ and trophic level $N$ that determines the attenuation of effective exposure time with trophic level,
\[ P_N(d) = EP_{N-1_{Desg}} \left( F_{HB} e^{-K_B d(N)} + F_{HL} e^{-K_L d(N)} \right) \]  

(18)

and for the next level, the degraded organic matter from trophic level N, by difference, is

\[ P_{N_{Desg}} = EP_{N-1_{Desg}} - P_N(d) \]  

(19)

The isotopic values for each trophic level can be calculated as above, but replacing \( d \) in all cases with \( D(d, N) \). For the cases presented in which effective exposure time varies with trophic level, we have formulated the function \( D \) as

\[ D(d, N) = de^{-\zeta N} \]  

(20)

where \( d \) is the “maximum exposure time” experienced by primary photosynthetic organic matter, \( \zeta \) is some attenuation scaling factor as described (cases discussed include values of 0.16 and 0.5), and \( N \) is the trophic level. The influence of \( \zeta \) on effective exposure time is shown in Supporting Figure 2.

**Case 3: Size-population-dependent signature of primary organic matter**

We now model two surface populations with different starting isotopic values and, in one case, with differing amounts of cellular lipid; the composition and isotopic signature of primary organic matter is thus dependent on the relative population of small and large cells in the surface ocean and some proportion or size scaling that determines their lipid content. For two size classes of cells, small (S) and large (G), we treat the average size of each as a ratio of small:large cell radius, or the fractional radius of the small cells \( R_z \) relative to the large cells. Similarly, the population of small:large cells is treated as a ratio, and thus this population ratio \( R_p \) can be treated as the numerical abundance of small cells for every one large cell. We then model the total amount of biomass from a given cell type as population \( \times \) radius\(^3\), assuming that the abundance of biomass in a cell is proportional to the cell volume. The amount of total biomass produced by small cells in the surface ocean is thus \( P_S = R_p R_z^3 \) and by large cells is \( P_G = 1 \times 1^3 \), and the total amount of photosynthetic biomass is

\[ P_0 = R_p R_z^3 + 1 \]  

(21)

The amount of lipid in a cell is modeled as some fraction of its biomass: the respective fractions are the values \( F_{SL} \) and \( F_{GL} \). Retaining nomenclature for other compound classes as before (the fractional contributions of which are calculated by difference from \( F_L \)), and assuming resistant biopolymer is only produced by large cells, the amount of biomass from primary producers that remains preserved after some degree of degradation \( d \) is
In calculating the isotopic value of this primary biomass, we introduce an isotopic enrichment in small cells, of magnitude $\beta_{S-G}$ (per mil). We define the isotopic composition of large cells as $\delta_{0G}$, and the starting isotopic composition of small cells is $\delta_{0S} = \delta_{0G} + \beta_{S-G}$. Assuming the biosynthetic isotopic fractionations as discussed previously, and calculating isotopic values for biosynthetic components by difference from the cellular total, the isotopic value of total preserved biomass from primary producers after some amount $d$ of degradation is

$$
\delta_{P0}(d) = R_p R^3_z \left( F_{SLE} e^{-KLd} + F_{SBE} e^{-Kld} \right) + F_{GLE} e^{-KLd} + F_{GBL} e^{-Kld} + F_{GR} e^{-Krd} (22)
$$

Heterotrophic biomass amounts, compound distribution, and isotopic content are calculated as in Case 2a, above.

Kerogen versus TOC terminology

TOC and kerogen for many Proterozoic deposits are observed to have very similar $^{13}C$ contents, differing by $<1\%$ (e.g. Hieshima, 1992 - Nonesuch Formation; Summons, 1992 - Chuar Group). Additionally, it has been estimated that $\sim 90-95\%$ of sedimentary TOC is typically comprised of kerogen (Durand, 1980); therefore, the isotopic signatures of TOC and kerogen should be very similar, and deviations can sometimes be attributed to fluid migration (Strauss et al., 1992). Most Proterozoic isotopic studies cited report values for kerogen, excluding Hieshima (1992). For modeling purposes, however, we report TOC values, since the fractional partitioning of lipid products into bound kerogen rather than extractable bitumen is unclear, and could result in truncated model runs if extractable lipid concentrations are further limited. Our model-run degradation limit of 1 ppm lipid remaining in TOC would thus reflect all potentially measurable lipid, whether existing as free compounds or released from kerogen during diagenesis or analysis. Our model results are also assumed to be conservative since kerogen is usually slightly isotopically lighter than TOC, if different at all; considering kerogen rather than TOC would therefore increase the positive value of $\Delta\delta$.

Selection of Model Parameters

Degradation rate constants

Relative values for degradation rates of different compound classes are important in determining their effect on the composition and isotopic signature of both preserved primary organic matter and substrate available to heterotrophs. Absolute values are arbitrary since we do not address specific time intervals for degradation. While there are some literature examples of measured degradation rates for specific compound classes (e.g. Nguyen et al., 2003; Henrichs et al., 1992, among many), they address varying levels of chemical specificity, stages of degradation, and purity of source. The resulting rate values vary across many orders of magnitude.
We incorporate a labile component of biomass across all organisms; its degradation rate can be addressed as arbitrarily fast, since we are considering mostly scenarios in which “unprotected” labile material is fully consumed or respired. In many published results (e.g. Henrichs et al., 1992) labile biomass is degraded at rates 2-3 orders of magnitude faster than recalcitrant fractions. In others, it is degraded only a factor of 2 faster (e.g. Nguyen et al., 2003; Versteegh and Zonneveld, 2002).

We formulate degradation rates for lipid and other recalcitrant portions of organic matter so that lipid persists in concentrations of at least 1 ppm (i.e. detectable) in sedimentary scenarios where only a small portion (0.01%) of original primary organic matter is preserved. In order to ensure this, we choose rate constants such that

$$\frac{K_L}{K_R} \leq \left( \frac{\ln ([L] \cdot \frac{F_{Pres}}{F_R})}{\ln (\frac{F_{Pres}}{F_L})} \right)$$

where $F_{Pres}$ is our requirement for minimum fraction of original primary organic matter preserved (0.0001; note this does not imply sedimentary concentration, as that is dependent on corresponding inorganic sedimentation rate) and $[L]$ is the minimum required concentration of lipid in preserved OM in order to be detected (1 ppm). With these constraints, $\frac{K_L}{K_R}$ must not exceed 2.7. We use this value in our examples in order to capture some distinction between free lipid and other resistant, potentially bound/polymerized compounds. In our Monte Carlo simulations, we explore a range down to a 1:1 relationship between $K_L$ and $K_R$, in order to incorporate the possibility that lipid and the resistant fraction have similar sources and/or modes of protection from degradation.

**Chemical and isotopic distribution across cell types**

Literature values indicate that cellular lipid content can range from <5 to >25% (dry weight; Fernandez-Reiriz et al., 1989; Fleurence et al., 1992; Brown, 1991) depending on phylogeny and various environmental factors. Phospholipids, the major presumed precursors for preserved n-alkyl lipids, generally comprise some large fraction of total lipid. We therefore use an average value of 10% n-alkyl lipid for all cell types, but we consider a range of 5-25% in the Monte Carlo simulation. We use an average biosynthetic fractionation of 4.5‰ for n-alkyl lipids as compared to average biomass for a given cell, but we consider a range of 3-6‰ in the Monte Carlo simulations, since fully-expressed fractionation factors for the biosynthetic pathway of acetogenic lipids would result in larger expressed fractionation, while high cellular lipid content might result in lower expressed fractionations (Laws, 1991; Hayes, 2001).

The resistant biological component that we account for could encompass a wide range of reported preservation effects in eukaryotes. Most directly, and in keeping with a Logan-type explanation, this could represent a cell-wall biopolymer of eukaryotic algae that is resistant to degradation (e.g. algaenan; Tegelaar et al., 1989, and others). Literature values vary widely as to the commonness, cellular content, and isotopic fractionation of these types of polymers (e.g. Kodner et al., 2009). Reports suggest the cellular content is likely to range from <10% to >30% (de Leeuw and Largeau, 1993), and since these polymers are usually described as strongly aliphatic in origin, some isotopic fractionation is suspected to occur during biosynthesis (e.g. Eglington, 1994; Hõld et al., 1998). Other modes of resistance for a eukaryotic component could stem from preferential physical protection of biomass through cell wall matrix effects (Nguyen et al., 2003), or from reduced exposure time via faster sinking through the water column (e.g. Butterfield, 2009; Burd and Jackson, 2009). In keeping with a biopolymer origin for the resistant fraction, we present cases in which the resistant fraction is isotopically depleted by 1.5‰ (the algaenan model) and comprises 30% of cellular material. However, in our Monte Carlo simulation, we address a 15-35% range of cellular composition and we explore isotopic signatures ranging from 4.5‰ depleted to 1.5‰ enriched in $^{13}$C compared to average
biomass. The wide range of isotopic values accounts for the possibility that the resistant fraction is instead an early diagenetic polymerization product that could derive from a wide range of biochemical sources (e.g. Eglinton, 1994).

**Heterotrophic efficiency and isotopic enrichment**

The efficiency of conversion of organic substrate to de novo biosynthetic product by heterotrophic bacteria is commonly referred to as bacterial growth efficiency (BGE; e.g. del Giorgio and Cole, 1998), and is largely responsible for the limitations on trophic isotopic enrichment by the relationship $H = h_{max}(1 - E)$ as described above (where $H =$ actual isotopic enrichment, $h_{max} =$ maximum isotopic enrichment, and $E =$ BGE). A wide range of values from $<0.05$ to $>0.5$ for $E$ has been reported from natural environments. Del Giorgio and Cole (1998) related relative values of BGE to environmental/ecologic factors such as oligotrophy and eutrophy. We utilize an intermediate value of 0.3 in our examples, but we consider a wide range (0.05-0.6) in our Monte Carlo simulations, given the uncertainties on the ecological controls likely to be present in various Proterozoic environments.

The maximum isotopic enrichment per trophic level, $h_{max}$, is modulated by BGE as described. We utilize a maximum value of $1.5\%$, a commonly cited value originally deriving from DeNiro and Epstein (1978), but incorporate a small range of uncertainty in this number by using values from 1-2\% in our Monte Carlo simulations.
Supporting Figure 1. Final $^{13}$C composition of preserved TOC, total n-alkyl lipid, and net heterotrophic contribution to TOC under Case 1. No distinction is made between compound classes; all are assumed to degrade at a constant rate. As $d$ increases, the $^{13}$C content of heterotrophic biomass increases because a larger supply of primary material is available to extend the food chain and increase the relative amount of biomass preserved from later trophic levels. The maximum enrichment is limited, however, by the decreasing amount of substrate available to late trophic levels. Total preserved n-alkyl lipid is preserved with an isotopic offset from TOC that reflects biosynthetic fractionation from the net organic pool. Although heterotrophic organic matter can become significantly positive in its $^{13}$C value, it cannot become positive enough to drive its associated n-alkyl lipids to values $> 0\%$ in relation to preserved TOC.
Supporting Figure 2. Effective attenuation of degradative exposure (D) experienced by trophic level (N), at different values of \( \zeta \). Primary producers are denoted by trophic level 0 and always experience the highest degradative exposure (d). Heterotrophic processing begins at trophic level 1; degradative exposure experienced by heterotrophic biomass is less than that experienced by primary producers by a factor of \( e^{-\zeta N} \), such that \( D = de^{-\zeta N} \).
Supporting References


