A unified theory for organic matter accumulation

Emily J. Zakem,[1] B. B. Cael,[2] Naomi M. Levine[1]

[1]Department of Biological Sciences, University of Southern California, Los Angeles, USA
[2]National Oceanography Centre, Southampton, UK

Abstract

Organic matter constitutes a key reservoir in global elemental cycles and the biological sequestration of carbon. However, our understanding of the dynamics of organic matter and its accumulation remains incomplete. Seemingly disparate hypotheses have been proposed to explain organic matter accumulation: the slow degradation of intrinsically recalcitrant substrates, the depletion to concentrations that inhibit microbial consumption, and a dependency on the consumption capabilities of nearby microbial populations. Here, using a mechanistic model, we develop a theoretical framework that explains how organic matter predictably accumulates due to biochemical, ecological, and environmental factors. The new framework subsumes the previous hypotheses. Changes in the microbial community or the environment can move a class of organic matter from a state of functional recalcitrance to a state of depletion by microbial consumers. The model explains the vertical profile of dissolved organic carbon in the ocean, and it connects microbial activity at subannual timescales to organic matter turnover at millenial timescales. The threshold behavior of the model implies that organic matter accumulation may respond nonlinearly to changes in temperature and other factors, providing hypotheses for the observed correlations between organic carbon reservoirs and temperature in past earth climates.

Introduction

Heterotrophic organisms consume organic matter (OM) for both energy and biomass synthesis. Their respiration transforms much of it back into the inorganic nutrients that fuel primary production. Residual OM accumulates into large reservoirs in the ocean, sediments, and soils. Together, these pools store about five times more carbon than in the atmosphere and play a central role in global biogeochemistry[1]. The dynamics of OM cycling and accumulation are key to understanding how the carbon cycle changes with climate[1,2].

Standing stocks of OM comprise a heterogeneous mix of thousands of compounds, many of which are uncharacterized[3,4]. Compounds are often conceptually described in terms of a degree of 'lability' that correlates with consumption rates, thereby inversely related to abundance and longevity in the environment[5,6]. In most biogeochemical models, OM degradation is dictated by simple rate constants rather than explicit microbial consumption[7,8]. Though current efforts are progressing towards the integrating OM cycling and microbial ecological dynamics[6,9,10,11], we still lack a mechanistic understanding of the ecological controls on OM and its accumulation.

Three hypotheses have been invoked to explain OM accumulation: 1. *"Recalcitrance":* Compounds may accumulate because they are relatively slowly degraded or resistant to further degradation by microorganisms[5,6,12,13]. This is consistent with observations, theory, and inferences of a wide range of consumption rates and compound ages in the environment[14,7,15,8,16,17,18,19,20,21,6]. 2. *"Dilution":* The accumulation may represent the sum of low concentrations of many organic compounds, each having been "diluted" by microbial consumption to a minimum amount[22]. This is supported by evidence that concentrating apparently recalcitrant dissolved OM from the deep ocean fuels microbial growth[23]. Modeling efforts have reconciled observed carbon ages with this mechanism and have interpreted the minimum concentrations as resource subsistence concentrations, the minimum concentrations to which populations can deplete their required resources[11,24,25]. 3. *"Ecosystem dependency":* The accumulation may result from a mismatch between OM characteristics and the metabolic capability of the proximal microbial community (e.g. the substrate-specificity of enzymes)[26,27,28]. For example, the dispersal of microbial populations, which is controlled by the connectivity of the environment and which may manifest as a stochastic process[29], can allow for intermittent or sporadic OM consumption events[27].

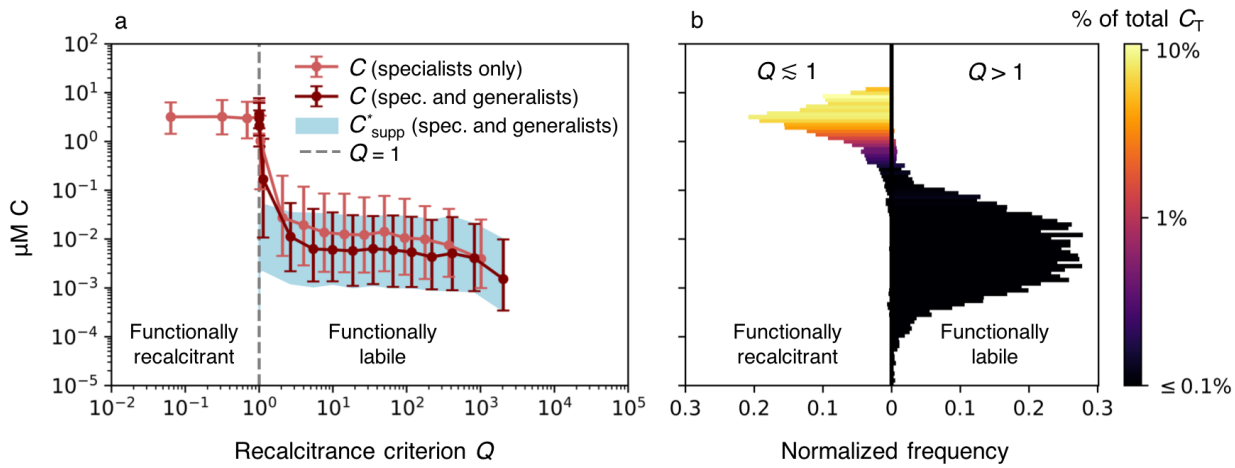

Figure 1: Resulting concentrations of OM pools in the stochastic model. (a) OM concentration C and the supply-limited subsistence concentration C^*_{supp} for ten simulations each of two model versions: with 1,000 'specialist' microbial populations, and with an additional 1,000 'generalist' populations that consume varying numbers of pools. Each of the ten simulations resolves 1,000 OM pools. Bars and shaded area indicate the 16th and 84th percentiles (equivalent to one standard deviation). (b) Normalized frequencies of C and contribution to total carbon for the version with generalists, delineated by $Q = 1$ (cutoff at 1.01).

Here, we investigate why OM accumulates using a stochastic model that simulates the complex dynamics of microbial OM consumption. We find that the mechanisms underlying each of the three hypotheses come into play simultaneously in the model. We develop a quantitative definition of functional recalcitrance that depends on the microbial community and the environmental context, in addition to substrate characteristics. The threshold behavior of the recalcitrance criterion suggests nonlinear OM responses to changes in the environment.

A mechanistic model of organic matter consumption

The model resolves multiple pools of OM ($n = 1,000$) that are supplied stochastically and consumed by one or more microbial populations ($n = 1,000$ or $2,000$, Eqns. 3–5, Fig. 1). We represent the net impact of each complex OM-microbe interaction (e.g. hydrolysis, enzymatic rates, cellular allocation of enzyme, mineral protection, free energy released by OM oxidation) with a simplified set of parameters: maximum uptake rate, half-saturation concentration, and biomass yield. To include the impact of variable community composition, we modulate the consumption by each

population over time according to its stochastically assigned probability of presence. We vary the degree of 'specialist' vs. 'generalist' ability, incorporating a penalty that increases with the number of pools consumed to represent the tradeoff among the strategies. We vary both the number of pools consumed by each population and the number of consumers of each pool (Fig. S1). Population loss rates are proportional to biomass according to both quadratic and linear mortality parameters, simulating predation, viral lysis, and cell maintenance demand.

We sample all parameter values from uniform distributions over wide, plausible ranges (Table S1). We numerically integrate the equations forward in time, allowing the concentrations of OM pools to emerge from ecological interactions. The dynamics presented here are robust across the parameter space, variations in the model structure, and variations in the number of OM pools and populations (SI Text 1 and 2, Figs. S2–S7). Sequential transformation of one OM pool to another due to incomplete oxidation gives quantitatively but not qualitatively different solutions, although this may increase compound age[11]. We present results compiled from ten simulations integrated for ten years (Fig. 1).

The solutions reveal a set of qualitatively distinct controls on OM accumulation. The distinct controls result in a bimodal distribution of OM concentrations (Fig. 1, Fig. S2). Discernible bimodality depends on the parameter distributions (Fig. S12), as well as abiotic sources and sinks not included in the model. In the simple model, the majority of pools are depleted to relatively low concentrations (10^{-4} to 1 μM C), while a subset accumulates to substantially higher concentrations (0.1 to 10 μM C). The latter accumulated pools comprise the bulk of total carbon content (Fig. 1b).

Diagnosing functional recalcitrance

The pools that are depleted have reached equilibrium, and their concentrations match the minimum subsistence concentrations among the microbial populations (SI Text 3, Fig. S8). We term these pools 'functionally labile'. The populations continue to consume these pools in proportion to their supply while maintaining the subsistence concentrations[11]. The subsistence concentration C^*_{ij} for

'specialist' population j, subsisting solely on OM pool i, is:

$$C^*_{ij} = \frac{k_{ij}}{P_j \frac{\rho^{\max}_{ij}}{L_j(B)} y_{ij} - 1} \tag{1}$$

where P_j is the probability that population j encounters OM pool i, $\rho^{\max}_{ij}$ is the maximum uptake rate, k_{ij} is the half-saturation concentration, y_{ij} is the biomass yield, and L_j is the population loss rate, which varies as a function of biomass B (Eqn. 5). The lowest concentrations among the functionally labile pools equate to consumption that is limited predominantly by their supply rates (C^*_{supp} in Fig. 1a; derived with the assumption that $C \ll k$; SI Text 3).

Most of the pools that accumulate to higher concentrations do not equilibrate in the simple model. We diagnose these pools as 'functionally recalcitrant' using recalcitrance criterion Q. For a pool to equilibrate at the subsistence concentration of the specialist ($C^*_{ij} > 0$ in Eqn. 1), the maximum rate of local biomass synthesis $P_j \rho^{\max}_{ij} y_{ij}$ must exceed L_j, which is the population turnover rate at steady state. When we extend the analysis to consider the entire microbial community (SI Text 3), including 'generalists' that can consume more than one pool of OM, the recalcitrance criterion for pool i can be defined as

$$Q_i = \max_j \left[P_j \frac{\rho^{\max}_{ij}}{L_j(B)} \left(y_{ij} + \underbrace{\sum_k y_{kj} \frac{\rho_{kj}}{\rho_{ij}}}_{\text{impact of other pools}} \right) \right] \tag{2}$$

where k denotes another pool consumed by population j, and ρ_{kj}/ρ_{ij} is the relative uptake of pool k to pool i by population j. If $Q_i > 1$, pool i can be depleted to the lowest subsistence concentration of the community given sufficient time (SI Text 3), with equilibration occurring over a spectrum of timescales dictated by the parameters. If $Q_i \leq 1$, pool i is functionally recalcitrant (Fig. 1).

This demonstrates how recalcitrance is simultaneously governed by chemical, biological, ecological, and environmental characteristics. In recalcitrance criterion Q_i, both y_{ij} and $\rho^{\max}_{ij}$ reflect the enzyme-dependent substrate–microbe interaction i–j, which in turn reflects the energetic content and the accessibility of the OM[30,31,32]. The encounter probability P_j and the loss function L_j reflect the ecological context: the diversity and abundance of the local microbial populations, predators, and

viruses. This is controlled by environmental connectivity in addition to selection[29], which is shaped in part by physical conditions such as circulation and sinking particles in the ocean and porosity in soils. Diversity and connectivity also modulate the availability of other pools, which can also include inorganic nutrients and the impact of chemolithotrophy[33]. In Eqn. 2, the uptake of another pool k by population j can increase its potential to deplete pool i (i.e. Q_i may increase). This is a mechanistic explanation for the 'priming effect' in which the addition of other substrates allows for the metabolization of a given pool[34,35].

The framework posits recalcitrance as a community- and context-specific phenomenon that can change in time and space (Fig. S10). Because Q_i of each OM pool is the maximum of multiple population-specific values, one for each population consuming that pool, the 'community recalcitrance' – the fraction of the community that experiences each pool as recalcitrant – varies in the model as a function of the parameter combinations that distinguish the populations (Fig. 2). For example, for local maximum consumption rates of about 0.1 d^{-1}, about 5% to 20% of populations experience a particular pool as recalcitrant (Fig. 2). This implies that statistically, functional recalcitrance may be less prominent when OM is exposed to a higher diversity of heterotrophic microorganisms. This also implies that functional recalcitrance may arise from the substrate-specificity of the heterotrophic populations[36]. This is consistent with evidence that specific heterotrophic clades consume recalcitrant carboxyl-rich alicyclic molecules (CRAM), which comprise a significant fraction (up to 8%) of marine DOC[37].

When an OM pool is functionally recalcitrant ($Q_i \leq 1$), it can slowly accumulate or diminish at a rate dependent on production, consumption, and physical transport over time, or it can equilibrate due to an abiotic, concentration-dependent sink such as photolysis[5]. In model versions with many generalists (Fig. 1), $Q_i \approx 1$ (to within 1%) for the recalcitrant pools because there is always at least one consuming population sustained by other pools (though not when generalists switch their consumption over time to optimize growth; Fig. S7). When $Q_i < 1$, no population able to consume that pool is locally sustainable (Fig. S9).

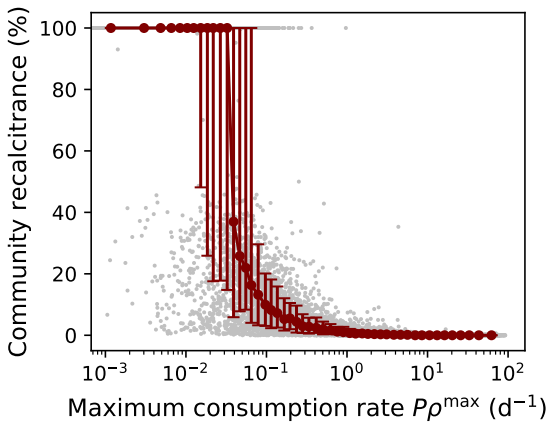

Figure 2: Community recalcitrance in the model. The percentage of consumers for which an OM pool is functionally recalcitrant is plotted against the local maximum consumption rate of each OM pool (the highest combination of probability of presence P and maximum uptake rate $\rho^{\max}$ among the populations consuming each pool). The grey dots indicate each of 10,000 pools in the ten simulation compilation, and the markers and bars indicate the binned mean, 16th, and 84th percentiles.

Unification of hypotheses

The three current hypotheses for OM accumulation – recalcitrance, dilution, and ecosystem dependency – each explain aspects of the total amount of carbon in the model. We may consider each hypothesis individually as a limit case for the formation of large organic carbon reservoirs in natural environments. Total organic carbon content (C_{T}) is the sum of all pools. A traditional view of recalcitrance, focused on intrinsic qualities of the substrate or of the microbe–substrate-specific enzymatic reaction, is represented in the model as $y_{ij}\rho_{ij}^{max}$. As the maximum uptake rate or yield becomes small (i.e. $y_{ij}\rho_{ij}^{\max} \to 0$), $C_{\mathrm{T}} \to \infty$. The number of OM pools n that are present can impact C_{T} in two opposing ways. As $n \to \infty$, $C_{\mathrm{T}} \to \infty$ even for low, equilibrated subsistence concentrations ("dilution"). However, the impact of other OM pools ("priming") means that as n increases, the likelihood of $Q_i > 1$ increases, decreasing the likelihood of functional recalcitrance and thus potentially decreasing C_{T}. Ecosystem dependency is encapsulated in the population turnover rate L_j and the probability P_j of co-occurrence with consuming populations. As $L_j \to \infty$, $C_{\mathrm{T}} \to \infty$. As the number of nearby consumers increases, the likelihood of $Q_i > 1$ increases. Thus as $P_j \to 0$, $C_{\mathrm{T}} \to \infty$.

Therefore, the parameter space setting the population and OM characteristics – the specified distributions in our stochastic model – determines the degree to which each mechanism controls

OM accumulation in different environments. Here, we estimate the distributions using plausible ranges of parameters in natural environments (SI Text 1, Table S1). The model is consistent with that of ref.[11] in that intrinsic recalcitrance is not necessary for OM accumulation, as well as with experimental evidence for dilution-limited consumption (SI Text 4, Fig. S11)[23]. Here we provide a more generalized framework that encapsulates a wider and more complete set of dynamics, and so is also consistent with evidence of recalcitrance[5,12,13] and the impact of the microbial community[26,28,36,37]. The modeled distributions of OM degradation rates are consistent with theory and observations that remineralization rates are lognormally distributed over a wide range due to multiplicative stochasticity in the underlying processes[19,38] (Fig. S12).

A key aspect of the framework is the threshold behavior of the accumulation. The threshold, $Q = 1$, reflects the sustainability of the microbial populations that consume OM. $Q = 1$ represents an ecological threshold along a continuum of OM and microbial characteristics that encompasses a variety of mechanistic factors, including those known to influence recalcitrance such as thermodynamic limitations[30], enzymatic control[28], mineral protection[31,35,39], and molecular properties[13].

Predicting accumulation at global scales

Our framework can help explain large-scale patterns in OM accumulation. Here, we use the framework to understand the vertical structure of dissolved organic carbon (DOC) in the ocean, though it can potentially also be used to understand microbial control on OM in soils and sediments. Globally, DOC concentrations peak at the sea surface and approach a minimum at depth (Fig. 3a)[5]. Since the stochastic model is not practical for multi-dimensional biogeochemical models, we utilize a reduced-complexity model analog that captures the essence of the stochastic model, resolving 25 aggregate pools. We incorporate this model analog into a fully dynamic ecosystem model of a stratified marine water column, where production and consumption of all organic and inorganic pools are resolved mechanistically as the growth, respiration, and mortality of photoautotrophic and heterotrophic microbial populations (SI Text 5, Fig. S13). The model is integrated for 6,000 years to quasi-equilibrium (SI Text 5).

Ecological interactions in the model result in characteristics typical of a marine water column (Fig. 3,

Fig. S14). Modeled DOC accumulates throughout the water column. Total DOC decreases smoothly with depth, with higher surface DOC transported to depth by vertical mixing (Fig. 3a). Most pools are depleted to subsistence concentrations throughout the water column (Fig. 3b,c). One pool remains functionally recalcitrant throughout the entire water column due to its slow consumption rate (lightest yellow line in Fig. 3b), which is consistent with the observed homogenous composition of aged marine DOC[4]. Many pools accumulate at the surface and become depleted at depths of 500 to 1,000 m. This transition is due to the increase in Q from the surface to depth (Fig. 3b). Productivity peaks in the sunlit surface, and so total biomasses, activity rates, and therefore population loss rates reach maximum values at the surface and attenuate with depth. This transition is consistent with observations that a subset of DOC was resistant to consumption by surface communities but able to be remineralized by deep communities[26]. Thus an ecologically determined transition from functional recalcitrance to functional lability, dictated by the sustainability of microbial populations represented in aggregated form, explains much of the decrease in DOC with depth. The subsistence concentrations also decrease with depth as loss rates decrease, and thus functionally labile OM may also contribute to the vertical DOC gradient.

Implications

Our analysis shows that it is possible that the majority of the diverse types of OM in soils, sediments, and the ocean are present at low concentrations, while the majority of the total standing stock is functionally recalcitrant (Fig. 1b, Fig. 3b). Recalcitrant OM pools may equilibrate if subjected to abiotic concentration-dependent sinks[5], or they may change slowly with time on average[4,27]. Our framework emphasizes that apparently slow consumption rates may be controlled by the frequency of encounter of 'the right' populations and substrates in addition to cellular processing limitations. This is consistent with the understanding that localized sinks cause the 10% to 20% decrease in deep ocean DOC along the deep ocean circulation pathway that accompanies a decrease in oxygen, a signature of respiratory demand[5,27].

The water column model links the millenial timescales of OM turnover[20] to microbial consumption timescales of about 100 days (Fig. 3d). Although explicitly tracking organic transformation through a complex interaction network can also explain old carbon ages[11], slow change is consistent with

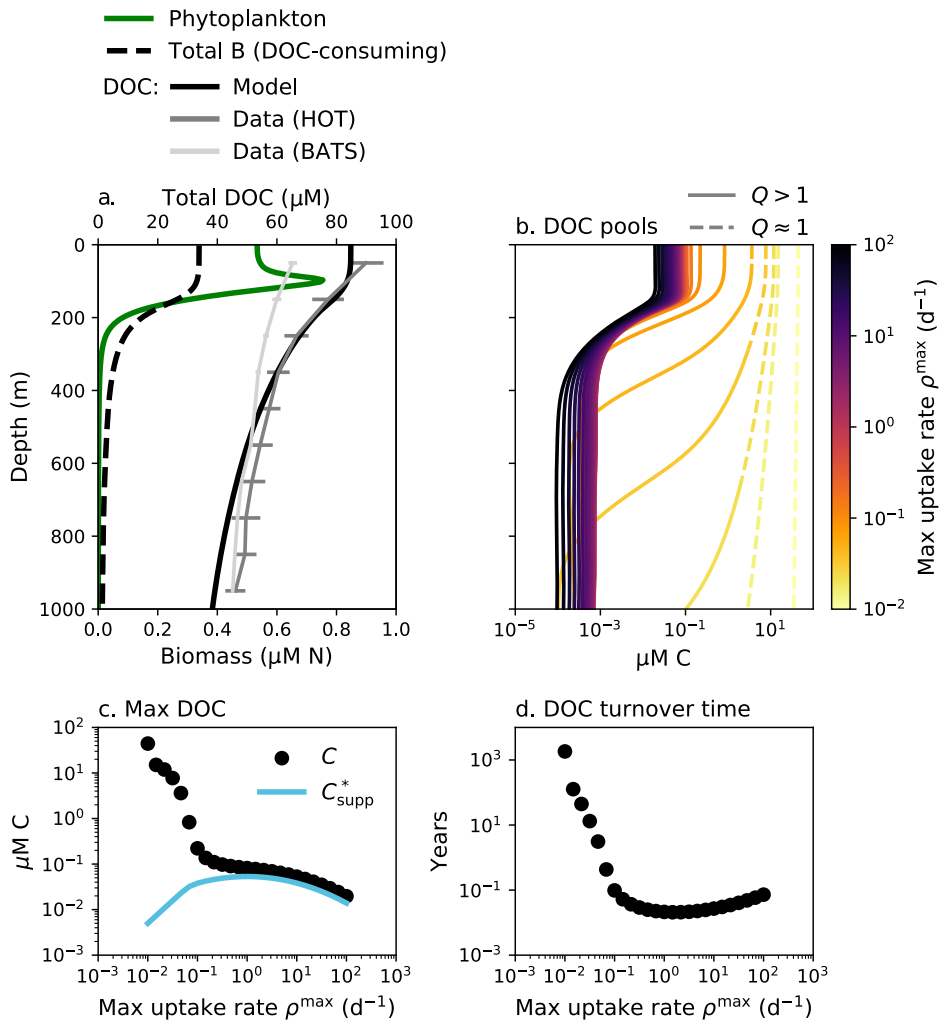

Figure 3: Marine ecosystem model solutions, simulating the accumulation of dissolved organic carbon (DOC). (a) Phytoplankton, total DOC-consuming biomass, total modeled DOC, and observed DOC, showing annual profiles from two ocean time series stations (HOT, the Hawaii Ocean Time-series, and BATS, the Bermuda Atlantic Time-series Station)[40]. The (b) concentration, (c) maximum concentration, and (d) bulk turnover timescale of each DOC pool.

inferences that nonsteady dynamics characterize the size of organic carbon reservoirs over geologic timescales[2].

Consumption of recalcitrant OM depends on the rate of microbial processing, which increases with temperature. All else the same, the model predicts that less OM accumulates at higher temperatures (SI Text 2, Fig. S10). Indeed, the loss of soil OM is a likely positive feedback to current warming[41]. The framework here additionally suggests that a decrease in OM with warming may be nonlinear due to some of the pools crossing the threshold from functionally recalcitrant to functionally labile (Fig. S10). This may help to understand the correlations between temperature and organic carbon reservoirs in past earth climates, such as increased ocean carbon burial, 'inert' soil carbon reservoirs, and perhaps marine DOC in glacial periods[42,43]. Using this framework to predict changes in organic carbon reservoirs with current increases in global temperature requires accurate estimates of microbial community loss rates, as well as how temperature will impact these rates and the diversity of the community.

We identify a set of controls on OM accumulation and turnover rooted in the complexity of microbial ecosystems. Previously disconnected hypotheses for OM accumulation, including the many mechanisms giving rise to functional recalcitrance, are subsumed within one framework. OM concentrations are mediated by the characteristics of substrate–microbe interactions, the heterogeneity of organic substrates, microbial community dynamics, and the ecological and biogeochemical diversity set by the connectivity of the environment (Fig. 4). The model is consistent with a comprehensive set of observations and theory of OM concentrations, turnover rates, and ages. The framework can be used to quantify the degree to which each of the subsumed hypotheses explains OM accumulation in different environments and to develop testable hypotheses for how organic reservoirs change with the biogeochemical environment.

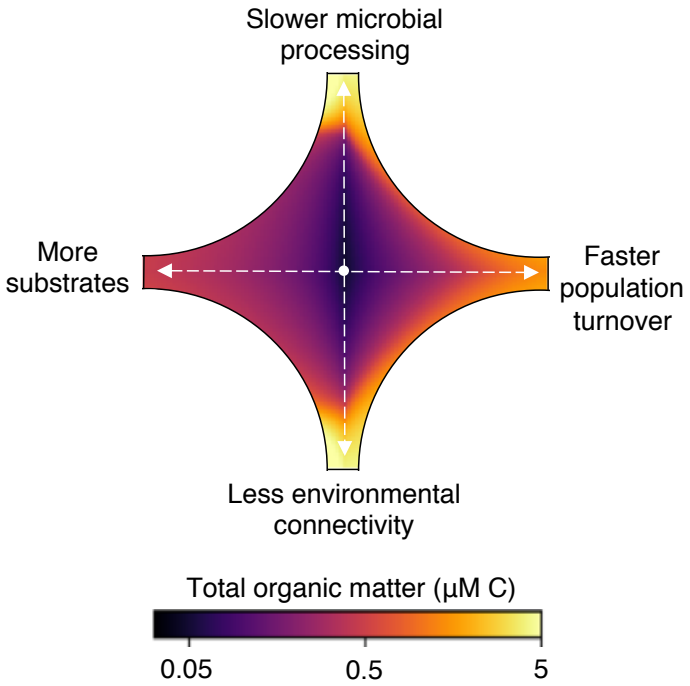

Figure 4: Summary of controls on organic matter accumulation by microbial consumption. Starting from a representative concentration in the center, change in total carbon is calculated for a ten-fold change in each of four parameters (i.e. two parameters vary in each quadrant): slower processing via reduced maximum uptake rate, faster turnover via an increased loss rate, less connectivity via a reduced population presence, and more substrates via multiplication of the center value.

Methods

OM consumption model We model microbial consumption and growth on pools of organic carbon. The model framework is sufficiently general to also account for inorganic nutrients, and may be extended to account for the cycling of other elements. We aggregate the many complex factors influencing the consumption rate of OM into a simplified set of uptake kinetics. We model the uptake ρ_{ij} of OM pool i by microbial population j using a saturating (Michaelis-Menten) form as:

$$\rho_{ij} = \rho_{ij}^{\max} \frac{C_i}{C_i + k_{ij}} \tag{3}$$

where $\rho^{\max}$ is the maximum uptake rate and k is the half-saturation constant. We consider $\rho^{\max}$ and k to reflect the net impact of enzymatic rate constants, extracellular degradation, the variable allocation of total cellular synthesis towards enzymes for substrate-specific hydrolysis and uptake, and other processes[44,10,27,28]. To account for the enormous complexity of these processes, we vary these parameters (specifically, $\rho^{\max}$ and the affinity $\rho^{\max}k^{-1}$) widely, over many orders of magnitude (Table S1).

Each population synthesizes biomass according to a growth efficiency (yield y), and loses biomass at a rate proportional to its biomass according to a quadratic mortality parameter m_q (implicitly representing higher order predators and viruses) and linear mortality parameter m_l (representing cell maintenance demand). The rates of change of the concentration C_i of pool i and the biomass B_j of population j are:

$$\frac{\partial C_i}{\partial t} = s_i - \sum_j P_{t_j} \rho_{ij} B_j \tag{4}$$

$$\frac{\partial B_j}{\partial t} = \sum_i P_{t_j} y_{ij} \rho_{ij} B_j - m_{q_j} B_j^2 - m_{l_j} B_j \tag{5}$$

where P_{t_j} is a boolean that varies with time t according to the overall probability of presence P_j of population j (see detail below), and s_i is the supply of pool i, which is a varying fraction of the total supply s_T to the box according to another assigned probability. At each timestep, a random draw determines whether pool i is supplied in accordance with P_{s_i} (Table S1). All parameters ($\rho^{\max}, k$ (via the affinity $\rho^{\max}k^{-1}$), y, m_q, m_l, P, and P_s) are set for each population by randomly sampling

from even distributions listed in Table S1.

Yield y reflects the cost of each enzyme and the free energy released by OM oxidation. $\rho^{\max}$ and y may be interdependent due to cellular optimization strategies, reflecting inherent tradeoffs between protein allocation and efficiency[45,46]. Here, different modes due to this plasticity are represented by multiple static realizations. The varying combinations of $\rho^{\max}$ and y can also represent the different modes of uptake of high molecular weight dissolved OM[47]. Analogously, the different parameter combinations can account for the additional feedback between the external concentration and the rate of cellular processing $\rho^{\max}$[48]. Real populations may change their cellular machinery due to plasticity, where in the model the many sets of parameters represent the static phenotypes among these different modes.

Community composition Observations show that community composition impacts OM consumption: community composition dictates the character of DOM remineralization in seemingly unpredictable ways[49,50], and evidence supports localized sinks of deep DOM[27]. To simulate this impact, we include a "presence-absence" dynamic in the model. Each population is assigned an overall probability of occurrence P, simulating the sporadic presence of rare functional types and the nearly guaranteed presence of ubiquitous types. At each timestep, a random draw determines whether population j consumes each pool or not in accordance with P_j. If the draw is below P_j, the population does not consume DOM or grow at that timestep, but its biomass is still subject to loss, and DOM may still be supplied. In effect, this dynamic extends the range of maximum uptake rates to lower values, demonstrating how the occasional absences of particular functional types contribute to longer effective remineralization timescales. This dynamic results in the majority of interactions at intermediate (though still widely ranging) rates, with 'very slow' and 'very fast' interactions being fairly rare (Fig. S12). This is similar to the lognormal distribution of supply with respect to the maximum uptake rate imposed in the water column ecosystem model (Fig. S14d). Over time, the consumption averages out so that incorporation of the overall probability P_j into the steady state balances closely match the dynamic model solutions.

Consumption matrix A consumption matrix dictates which populations consume which OM pools (Fig. S1). We vary the 'specialist' vs. 'generalist' capabilities of the populations with respect to the number of OM pools taken up by each population (n_{up}, which can vary from one to n, the number of OM pools), as well as with the widespread popularity of each pool with respect to the number of consumers of each (n_{cons}, which can vary from one to m, the number of populations). In the version with solely 'specialists' (Fig. 1a, Fig. S2), each population consumes only one unique pool. For the 'generalist populations' in Fig. 1, Fig. 2, and Fig. S2, we first randomly assign n_{up} to each population drawing from the linear range from one to n. Second, we assign a weight to each pool of its probability of being consumed n_{cons}, varying the weights linearly from one to n. Finally, we assign the specific pools taken up by each population (i.e. we fill each column of the consumption matrix) by sampling from the n possibilities with the weights. For the weighted sampling, we use the algorithms "ProbabilityWeights" and "sample" in the StatsBase package in Julia. See SI Text 1 and Figs. S3–S7 for other configurations.

Parameterization We assign parameter values in the model by drawing from wide ranges of parameter values using the random number generator "rand" in Julia (Table S1). These prior distributions are informed by estimating plausible ranges in real microbial communities as detailed below. Different prior distributions of parameter values result in different distributions of equilibrated vs. recalcitrant OM pools (Fig. S12). We emphasize that the framework provided by Q is applicable to alternative prior distributions, and thus that the model solutions remain qualitatively robust across different parameter spaces.

We vary the maximum uptake rate $\rho^{\max}$ over four orders of magnitude, drawing from 10^{-2} to 10^{2} d^{-1} (Table S1). $\rho^{\max}$ is meant to encompass not just enzymatic uptake rates, but also the net impact of other processes involved in the consumption of organic substrates, such as hydrolysis as well as internal processing constraints. $\rho^{\max}$ sets the overall growth rate ($y\rho^{\max}$ is the maximum biomass synthesis rate), and so its variation represents the wide range of observed microbial growth rates and population turnover rates[51,52]. Specifically, a recent study estimates the average growth rate of marine heterotrophic bacteria of about 0.1 ± 0.1 d^{-1} [51], while noting that growth rates of some organisms can exceed 100 d^{-1}. The lower limit of 10^{-2} d^{-1} likely does not capture the lowest

rates in sediments – evidence suggests that sediment microorganisms turnover at rates of years to hundreds of years[52] – though mortality rates there are also likely significantly lower.

We vary the uptake affinity $\rho^{\max}k^{-1}$ by two orders of magnitude (Table S1), resulting in variation of half-saturation concentration k of about six orders of magnitude, from 10^{-4} to 10^{2} μM C. This is consistent with observed ranges of the uptake affinity and half-saturation constants of phytoplankton and heterotrophic microorganisms[53,54,55].

The yield (or the growth efficiency) and the mortality constants are each sampled from linear ranges (Table S1). The upper limit to the yield is 0.5 to reflect the generally lower ranges of net heterotrophic carbon utilization efficiencies in natural environments compared to the laboratory[56,57]. The mortality constants are largely uncertain, but model results remain qualitatively robust for different values of these parameters. The quadratic mortality parameter largely influences the resulting concentrations of sustainable biomass, impacting the population loss rate. The linear mortality rate constitutes the lower limit to a population's sustainability in its environment, representing the cell's maintenance demand. Thus the maintenance demand may be lower in environments such as sediments and deep soils. A model version in which m_l was set to zero showed negligible difference in the solutions.

Simulations In the simulations illustrated in the main text, we resolve 1,000 OM classes and 1,000 or 2,000 pools of biomass: a model with 1,000 pure 'specialists', and a model with the 1,000 pure specialists and an additional 1,000 with a range of generalist ability. See Supplement for detail and experimentation with other configurations. Results with the 2,000 pools are similar to a model with only the 1,000 generalists. We illustrate ensembles of ten independent simulations, so that 10,000 datapoints comprise each of the binned results in Figs 1 and 2. For each experiment, we run the model forward in time for ten years, until the pools that have the potential to equilibrate have equilibrated. The concentrations of many of the recalcitrant pools continue to increase over time because their supply rate exceeds their consumption rate (unless a concentration-dependent sink is added to the model). Fig. S12 illustrates the resulting distributions of biomass concentrations, OM concentrations, and remineralization rates of the ensembles. These distributions are consistent with observed and inferred distributions of OM characteristics, remineralization rates, and ages in the ocean, sediments, soils, and lakes[16,14,58,7,19,59,60,17,18,20,21].

In SI Text 1 and Figs. S2-S7, we demonstrate the qualitative consistency of the solutions across variations in the model. All simulations support the conclusions presented in the main text. Solutions vary quantitatively but not qualitatively with variations in the 'specialist' vs. 'generalist' capabilities of the microbial populations, the number of OM pools resolved, the ratios of OM pools to populations resolved, the length of numerical integration, and the mode of uptake by the populations (additive consumption vs. switching over time to optimize growth).

Reduced-complexity model version For the reduced-complexity model of OM consumption used in the marine ecosystem model, we collapse the complexity onto one master 'lability' parameter – the maximum uptake rate – and we resolve fewer OM pools ($n = 25$) (Fig. S14). The values of y, m_q, m_l, and uptake affinity are kept constant, since their variation affects the solutions quantitatively but not qualitatively. In accordance with theory and our stochastic model results[38,19], we assume a lognormal distribution for the distribution of total OM production among the pools, representing the average outcome of microbial transformation over time and space (Fig. S14d).

Marine ecosystem model In the marine ecosystem model, production and consumption of all organic and inorganic pools are due to the growth, respiration, excretion, and mortality of microbial populations (Fig. S13). Light and vertical mixing attenuate with depth. Two populations of phytoplankton convert dissolved inorganic carbon and nitrogen into biomass using light energy. Populations of microbial heterotrophic functional biomass consume DOM and POM, oxidize a portion of the carbon for energy, and excrete DIN and DIC as waste products. For simplicity, POM is resolved as one aggregate pool sinking at a constant rate w_s. Oxygen is assumed abundant throughout the water column. The total DOM production rate is due to loss from POM degradation (due to the extracellular hydrolysis of POM) and the biomass loss rate of the microbial populations. Speciation of DIN was neglected for simplicity: solutions were visually identical in a model version that does include the DIN speciation and explicit resolution of chemoautotrophic nitrifying populations that fix carbon[61]. This suggests that the carbon fixation by chemoautotrophic populations does not substantially impact the DOC profile. The model is a modified version of a published model in which carbon and nitrogen of the organic pools and the biomasses are each resolved independently[62]. Parameter values are listed in Table S2. See SI Text 5 for model equations and

further detail.

Biomass approximations The biomass associated with recalcitrant pools reaches quasi-equilibrium ("processing"-limited B_{proc})[62], suggesting that natural assemblages may maintain steady biomass amidst varying amounts of recalcitrant OM. In addition to approximations for the supply-limited OM concentrations, we approximate the biomass associated with both the recalcitrant and supply-limited ends of the spectrum in the marine ecosystem model (Fig. 3f) as in ref.[62]. See SI Text 3 for derivations and equations (Eqns. S23 and S25).

Acknowledgments We acknowledge funding from the Simons Foundation: The Simons Collaboration on Principles of Microbial Ecology (PriME #542389 to N. M. L.) and the Simons Postdoctoral Fellowship in Marine Microbial Ecology (to E. J. Z.), the National Environmental Research Council (NE–R015953–1 to B. B. C.), and the European Union Horizon 2020 Research and Innovation Programme (grant #820989 to B. B. C.). We thank R. Letscher for providing the DOC data compilations and J. McNichol, E. McParland, and the Levine group for comments on the manuscript. The work reflects only the view of the authors; the European Commission and their executive agency are not responsible for any use that may be made of the information the work contains.

References

1. Ciais, P. *et al.* Carbon and Other Biogeochemical Cycles. Tech. Rep., Cambridge; New York (2013).

2. Rothman, D. H., Hayes, J. M. & Summons, R. E. Dynamics of the Neoproterozoic carbon cycle. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8124–8129 (2003).

3. Hedges, J. I. *et al.* The molecularly-uncharacterized component of nonliving organic matter in natural environments. *Org. Geochem.* **31**, 945–958 (2000).

4. Broek, T. A. B. *et al.* Low Molecular Weight Dissolved Organic Carbon: Aging, Compositional Changes, and Selective Utilization During Global Ocean Circulation. *Global Biogeochem. Cycles* **34**, e2020GB006547 (2020).

5. Hansell, D. A. Recalcitrant Dissolved Organic Carbon Fractions. *Ann. Rev. Mar. Sci.* **5**, 421–445 (2013).

6. Lehmann, J. & Kleber, M. The contentious nature of soil organic matter. *Nature* **528**, 60–68 (2015).

7. Boudreau, B. P. & Ruddick, B. R. On a reactive continuum representation of organic matter diagenesis. *Am. J. Sci.* **291**, 507–538 (1991).

8. Middelburg, J. J. A simple rate model for organic matter decomposition in marine sediments. *Geochim. Cosmochim. Acta* **53**, 1577–1581 (1989).

9. Schmidt, M. W. I. *et al.* Persistence of soil organic matter as an ecosystem property. *Nature* **478**, 49–56 (2011).

10. Allison, S. D. A trait-based approach for modelling microbial litter decomposition. *Ecol. Lett.* **15**, 1058–1070 (2012).

11. Mentges, A., Feenders, C., Deutsch, C., Blasius, B. & Dittmar, T. Long-term stability of marine dissolved organic carbon emerges from a neutral network of compounds and microbes. *Sci. Rep.* **9**, 17780 (2019).

12. Hedges, J. I. & Keil, R. G. Sedimentary organic matter preservation: an assessment and speculative synthesis. *Mar. Chem.* **49**, 81–115 (1995).

13. Shen, Y. & Benner, R. Molecular properties are a primary control on the microbial utilization of dissolved organic matter in the ocean. *Limnol. Oceanogr.* **65**, 1061–1071 (2020).

14. Jørgensen, B. B. A comparison of methods for the quantification of bacterial sulfate reduction in coastal marine sediments II. Calculation from mathematical models. *Geomicrobiol. J.* **1**, 29–47 (1978).

15. Williams, P. M. & Druffel, E. R. Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* **330**, 246–248 (1987).

16. Amon, R. M. & Benner, R. Rapid cycling of high-molecular-weight dissolved organic matter in the ocean. *Nature* **369**, 549–551 (1994).

17. Loh, A. N., Bauer, J. E. & Druffel, E. R. Variable ageing and storage of dissolved organic components in the open ocean. *Nature* **430**, 877–881 (2004).

18. Repeta, D. J. & Aluwihare, L. I. Radiocarbon analysis of neutral sugars in high-molecular-weight dissolved organic carbon: Implications for organic carbon cycling. *Limnol. Oceanogr.* **51**, 1045–1053 (2006).

19. Forney, D. C. & Rothman, D. H. Common structure in the heterogeneity of plant-matter decay. *J. R. Soc. Interface* **9**, 2255–2267 (2012).

20. Follett, C. L., Repeta, D. J., Rothman, D. H., Xu, L. & Santinelli, C. Hidden cycle of dissolved organic carbon in the deep ocean. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 16706–16711 (2014).

21. Lechtenfeld, O. J. *et al.* Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* **126**, 321–337 (2014).

22. Jannasch, H. W. Growth of marine bacteria at limiting concentrations of organic carbon in seawater. *Limnol. Oceanogr.* **12**, 264–271 (1967).

23. Arrieta, J. M., Mayol, E., Hansman, R. L. & Herndl, G. J. Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* **348**, 331—-333 (2015).

24. Tilman, D. *Resource Competition and Community Structure* (Princeton University Press, Princeton, N.J, 1982).

25. Wilson, J. D. & Arndt, S. Modeling radiocarbon constraints on the dilution of dissolved organic carbon in the deep ocean. *Global Biogeochem. Cycles* **31**, 775–786 (2017).

26. Carlson, C. A. *et al.* Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol. Oceanogr.* **49**, 1073–1083 (2004).

27. Hansell, D. A. & Carlson, C. A. Localized refractory dissolved organic carbon sinks in the deep ocean. *Global Biogeochem. Cycles* **27**, 705–710 (2013).

28. Arnosti, C. Microbial Extracellular Enzymes and the Marine Carbon Cycle. *Ann. Rev. Mar. Sci.* **3**, 401–425 (2011).

29. Evans, S., Martiny, J. B. & Allison, S. D. Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J.* **11**, 176–185 (2017).

30. LaRowe, D. E. & Van Cappellen, P. Degradation of natural organic matter: A thermodynamic analysis. *Geochim. Cosmochim. Acta* **75**, 2030–2042 (2011).

31. Hemingway, J. D. *et al.* Mineral protection regulates long-term global preservation of natural organic carbon. *Nature* **570**, 228–231 (2019).

32. Grabowski, E., Letelier, R. M., Laws, E. A. & Karl, D. M. Coupling carbon and energy fluxes in the North Pacific Subtropical Gyre. *Nat. Commun.* **10** (2019).

33. Swan, B. K. *et al.* Potential for Chemolithoautotrophy. *Science* **333**, 1296–9 (2011).

34. Guenet, B. *et al.* Impact of priming on global soil carbon stocks. *Glob. Chang. Biol.* **24**, 1873–1883 (2018).

35. Bailey, V. L., Pries, C. H. & Lajtha, K. What do we know about soil carbon destabilization? *Environ. Res. Lett.* **14**, 083004 (2019).

36. Sichert, A. *et al.* Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide fucoidan. *Nat. Microbiol.* s41564–020–0720–2 (2020).

37. Liu, S. *et al.* Different carboxyl-rich alicyclic molecules proxy compounds select distinct bacterioplankton for oxidation of dissolved organic matter in the mesopelagic Sargasso Sea. *Limnol. Oceanogr.* 11405 (2020).

38. Rothman, D. H. & Forney, D. C. Physical Model for the Decay and Preservation of Marine Organic Carbon. *Science* **316**, 1325 (2008).

39. Blattmann, T. M. *et al.* Mineralogical control on the fate of continentally derived organic matter in the ocean. *Science* **366**, 742–745 (2019).

40. Letscher, R. T. & Moore, J. K. Preferential remineralization of dissolved organic phosphorus and non-Redfield DOM dynamics in the global ocean: Impacts on marine productivity, nitrogen fixation, and carbon export. *Global Biogeochem. Cycles* **29**, 325–340 (2015).

41. Crowther, T. W. *et al.* Quantifying global soil carbon losses in response to warming. *Nature* **540**, 104–108 (2016).

42. Ciais, P. *et al.* Large inert carbon pool in the terrestrial biosphere during the Last Glacial Maximum. *Nat. Geosci.* **5**, 74–79 (2012).

43. Cartapanis, O., Bianchi, D., Jaccard, S. L. & Galbraith, E. D. Global pulses of organic carbon burial in deep-sea sediments during glacial maxima. *Nat. Commun.* **7**, 10796 (2016).

44. Sinsabaugh, R. L. *et al.* Stoichiometry of soil enzyme activity at global scale. *Ecol. Lett.* **11**, 1252–1264 (2008).

45. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: Origins and consequences. *Science* **330**, 1099–1102 (2010).

46. Basan, M. *et al.* Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).

47. Reintjes, G., Arnosti, C., Fuchs, B. & Amann, R. Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. *ISME J.* **13**, 1119–1132 (2019).

48. Allison, S. D., Chacon, S. S. & German, D. P. Substrate concentration constraints on microbial decomposition. *Soil Biol. Biochem.* **79**, 43–49 (2014).

49. Logue, J. B. *et al.* Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. *ISME J.* **10**, 533–545 (2016).

50. Balmonte, J. P. *et al.* Community structural differences shape microbial responses to high molecular weight organic matter. *Environ. Microbiol.* **21**, 557–571 (2019).

51. Kirchman, D. L. Growth Rates of Microbes in the Oceans. *Ann. Rev. Mar. Sci.* **8**, 285–309 (2016).

52. LaRowe, D. E. *et al.* The fate of organic carbon in marine sediments - New insights from recent data and analysis. *Earth-Science Rev.* **204**, 103146 (2020).

53. Litchman, E., Klausmeier, C. A., Schofield, O. M. & Falkowski, P. G. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecol. Lett.* **10**, 1170–81 (2007).

54. Lindemann, C., Fiksen, Ø., Andersen, K. H. & Aksnes, D. L. Scaling laws in phytoplankton nutrient uptake affinity. *Front. Mar. Sci.* **3**, 26 (2016).

55. Noell, S. E. & Giovannoni, S. J. SAR11 Bacteria Have a High Affinity and Multifunctional Glycine Betaine Transporter. *Environ. Microbiol.* **21**, 2559–2575 (2019).

56. Robinson, C. & Williams, P. J. l. B. Respiration and its measurements in surface marine waters. In del Giorgio, P. A. & Williams, P. (eds.) *Respir. Aquat. Ecosyst.*, chap. 9, 147–180 (Oxford, 2005).

57. Sinsabaugh, R. L., Manzoni, S., Moorhead, D. L. & Richter, A. Carbon use efficiency of microbial communities: stoichiometry, methodology and modelling. *Ecol. Lett.* **16**, 930–939 (2013).

58. Koch, B. P., Witt, M., Engbrodt, R., Dittmar, T. & Kattner, G. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* **69**, 3299–3308 (2005).

59. Koehler, B., Von Wachenfeldt, E., Kothawala, D. & Tranvik, L. J. Reactivity continuum of dissolved organic carbon decomposition in lake water. *J. Geophys. Res.* **117**, G01024 (2012).

60. Mostovaya, A., Hawkes, J. A., Koehler, B., Dittmar, T. & Tranvik, L. J. Emergence of the Reactivity Continuum of Organic Matter from Kinetics of a Multitude of Individual Molecular Constituents. *Environ. Sci. Technol.* **51**, 11571–11579 (2017).

61. Zakem, E. J. *et al.* Ecological control of nitrite in the upper ocean. *Nat. Commun.* **9**, 1206 (2018).

62. Zakem, E. J. & Levine, N. M. Systematic variation in marine dissolved organic matter stoichiometry and remineralization ratios as a function of lability. *Global Biogeochem. Cycles* **33**, 1389–1407 (2019).