A unified theory for organic matter accumulation

Emily J. Zakem¹,², B. B. Cael³, and Naomi M. Levine⁴

¹Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089; and ²Ocean Biogeochemistry and Ecosystems, National Oceanography Centre, Southampton SO14 3ZH, United Kingdom

Edited by David M. Karl, University of Hawaii at Manoa, Honolulu, HI, and approved December 30, 2020 (received for review August 9, 2020)

Organic matter constitutes a key reservoir in global elemental cycles. 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 in natural environments due to biochemical, ecological, and environmental factors. Our 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 connects microbial activity at subannual timescales to organic matter turnover at millennial 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.

Significance

Organic matter in the global ocean, soils, and sediments stores about five times more carbon than the atmosphere. Thus, the controls on the accumulation of organic matter are critical to global carbon cycling. However, we lack a quantitative understanding of these controls. This prevents meaningful descriptions of organic matter cycling in global climate models, which are required for understanding how changes in organic matter reservoirs provide feedbacks to past and present changes in climate. Currently, explanations for organic matter accumulation remain under debate, characterized by seemingly competing hypotheses. Here, we develop a quantitative framework for organic matter accumulation that unifies these hypotheses. The framework derives from the ecological dynamics of microorganisms, the dominant consumers of organic matter.

Author contributions: E.J.Z. and N.M.L. designed research; E.J.Z. performed research; E.J.Z., B.B.C., and N.M.L. analyzed data; and E.J.Z., B.B.C., and N.M.L. wrote the paper.

Published February 3, 2021.

See online for related content such as Commentaries.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

1To whom correspondence may be addressed. Email: zakem@usc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2016896118/-/DCSupplemental.
framework can also extend to soil and sediment environments. Finally, the threshold behavior of the recalcitrance indicator suggests nonlinear OM responses to changes in the environment.

A Mechanistic Model of OM Consumption. We develop a model of OM consumption by microbial populations using established forms of equations for microbial growth and respiration (12, 36, 37). The model resolves multiple pools of OM \((n = 1,000)\) that are supplied stochastically and consumed by one or more microbial populations \((n = 1,000 or 2,000; Eqs. 4-6 and Fig. 1)\). Stochastic supply captures the variable nature of the release of organic compounds, which is a function of complex biological dynamics (e.g., exudation, lysis, and grazing). We represent the net impact of each complex OM-microbe interaction [e.g., hydrolysis, enzymatic rates, cellular allocation of enzyme, and free energy released by OM oxidation (16, 32, 33, 38)] with a simplified set of parameters: maximum uptake rate, half-saturation concentration, and biomass yield (Materials and Methods). To include the impact of variable community composition, we modulate the OM consumption by each population over time according to its stochastically assigned probability of presence. We vary the degree of “specialists” (consuming a single OM pool) vs. “generalists” (consuming multiple OM pools), incorporating a penalty that increases with the number of pools consumed to represent a tradeoff among the strategies. We vary both the number of pools consumed by each population and the number of consumers of each pool (Fig. 1; SI Appendix, Fig. S1). Population loss rates are proportional to biomass according to both quadratic and linear mortality parameters, simulating predation, viral lysis, senescence, and maintenance demand.

Because we expect the values of these growth and mortality parameters to vary widely among organisms and substrates, we sample all parameter values from uniform distributions over wide, plausible ranges (Table 1; SI Appendix, SI Text 1). We numerically integrate the equations forward in time, allowing the concentrations of OM pools to emerge from the 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 Appendix, SI Text 2 and 3 and Figs. S2–S7). Sequential transformation of one OM pool to another due to incomplete oxidation gives qualitatively similar solutions (SI Appendix, SI Text 2), although this may increase compound age (17). We present results from simulations integrated for 10 y (Fig. 2).

The solutions reveal a bimodal distribution of OM concentrations (Fig. 2), implying a set of qualitatively distinct controls on OM accumulation. Whether or not the bimodality is discernible depends on the parameter distributions (SI Appendix, Fig. S13), as well as other sources and sinks not included in the model (e.g., photolysis). In the simple model, the majority of pools are depleted to relatively low concentrations \((10^{-4} \text{ to } 1 \text{ mM C})\), while a subset accumulates to substantially higher concentrations \((0.1 \text{ to } 10 \text{ mM C})\). The latter accumulated pools comprise the bulk of total carbon content (Fig. 2B).

Diagnosing Functional Recalcitrance. We evaluate whether each OM pool equilibrates or accumulates in the model. Equilibration indicates that the pool can sustain a microbial population in the given environment, and we classify that pool as “functionally labile.” Otherwise, the pool accumulates in the environment, and we classify that pool as “functionally recalcitrant.” For example, we describe the population dynamics of specialist population \(j\), subsisting solely on OM pool \(i\), as (Materials and Methods):

\[
\frac{\partial B_j(t)}{\partial t} = \left( P_j y_j \rho_j^{\text{max}} \frac{C_i(t)}{C_i(t) + k_{ij}} - L_j(t) \right) B_j(t),
\]

where \(B_j\) is the biomass, \(P_j\) is the probability of the presence of population \(j\), \(y_j\) is the biomass yield, \(\rho_j^{\text{max}}\) is the maximum uptake rate, \(k_{ij}\) is the half-saturation concentration for uptake, \(C_i\) is the concentration of the OM pool, and \(L_j\) is the population-loss rate, which varies as a function of the biomass (Eq. 6 and Table 1). When the system is at or close to steady state \(\left(\frac{d B_j(t)}{dt} \approx 0\right)\), the concentration of pool \(i\) can be estimated as:

\[
C_{ij}^* = \frac{k_{ij}}{\frac{\rho_j}{\rho_j^{\text{max}}} + \frac{1}{L_j}},
\]

which is the subsistence concentration of OM pool \(i\) for specialist population \(j\) (30). For a pool with multiple competing consumers, the concentration of that pool will be set by the population with the lowest subsistence concentration for that pool (30). The population can then continue to consume the pool in proportion to its supply while maintaining the subsistence concentration (17, 30).

![Fig. 1. Schematic of the OM consumption model. Multiple OM pools C and microbial populations B are resolved. The parameter values dictating the supply of each OM pool, the interaction between each pool and the microbial population (uptake kinetics and yield), and the loss of biomass (to viral lysis, grazing, senescence, and cell maintenance) are assigned stochastically. Here, we show an illustrative example where the fluxes dictated by these parameter values are represented with different widths of arrows. The supply and the presence or absence of each population vary stochastically over time in the model according to assigned probabilities.](https://doi.org/10.1073/pnas.2016896118)
For the OM pool to equilibrate ($C^*_j > 0$ in Eq. 2), the maximum rate of local biomass synthesis ($P_j y_{ij} \rho_{ij}^{\text{max}}$) must exceed the biomass loss rate at steady state ($L_j$). Using $C^*_j$ as a diagnostic, and extending the expression to generalists that can consume more than one OM pool (SI Appendix, Eqs. S17 and S18), we find that many concentrations of the modeled pools precisely match the minimum subsistence concentration among their consumers, and thus have equilibrated (Fig. 2A; SI Appendix, Figs. S8 and S9). Because these pools sustain microbial growth in this particular model environment, we consider these functionally labile. These low concentrations are consistent with the measured nanomolar or lower concentrations of known labile constituents of marine DOM, such as free amino acids and glucose (4, 39).

Most of the pools that accumulate to higher concentrations never equilibrate in the simple model. For these pools, the loss rates of all consuming populations match or exceed their maximum biomass synthesis rates. We consider these pools to be functionally recalcitrant. We can robustly define the threshold where pools transition from being functionally labile (depleted to $C^*_j$) to functionally recalcitrant (accumulating). We define a recalcitrance indicator $Q_j$ for pool $i$ as:

$$Q_j = \max_j \left[ \frac{P_j y_{ij} \rho_{ij}^{\text{max}}}{L_j} \left( y_{ij} + \sum_k y_{ik} \rho_{ik} / \rho_{kj} \right) \right]. \tag{[3]}$$

where index $k$ denotes a pool other than pool $i$ consumed by generalist population $j$, and $\rho_{ij} / \rho_{kj}$ is the relative uptake of pool $k$ to pool $i$ by population $j$ (see SI Appendix, SI Text 4 for derivation). For specialists, the term representing the impact of other pools drops out of the equation. If $Q_j > 1$, pool $i$ is functionally labile: At least one population can deplete it to its subsistence concentration given sufficient time, with the equilibration timescale dictated by the associated growth and loss parameters. If $Q_j \leq 1$, pool $i$ is functionally recalcitrant, and it accumulates over time in our model. Thus, $Q_j = 1$ serves as an emergent threshold between functional lability and functional recalcitrance (Fig. 2).

The recalcitrance indicator $Q_j$ demonstrates how recalcitrance is simultaneously governed by chemical, biological, ecological, and environmental characteristics. In Eq. 3, an enzyme-dependent substrate–microbe interaction $i \rightarrow j$ is captured by both $y_{ij}$ and $\rho_{ij}^{\text{max}}$, which also reflect the energetic content and the accessibility of the OM (40–42). The encounter probability of the population with the OM pool ($P_j$) and the biomass loss rate ($L_j$) capture the ecological context—the diversity and abundances of the local microbial populations, predators, and viruses. Many factors control these processes, including selection and environmental connectivity (35), which is shaped in part by physical conditions such as circulation and sinking particles in the ocean and porosity and diffusion in soils. Diversity and connectivity also modulate the availability of other pools for uptake by generalists. In Eq. 3, the uptake of an additional OM pool $k$ by population $j$ can increase the population’s potential to deplete pool $i$ (i.e., $Q_j$, increases). In other words, the ability of consumers of OM pool $i$ to consume other pools increases the functional lability of pool $i$. This provides a mechanistic explanation for the observed “priming effect,” in which the addition of other substrates allows for the metabolism of a given pool (34, 43, 44).

In the environment, a functionally recalcitrant OM pool may 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 (8, 34). In the model version with many generalists, $Q_j$ reaches a minimum of one (to within 1%) (Fig. 2A). When $Q_j \approx 1$, pools are unequivocally and functionally recalcitrant, but consumption can continue by consumers whose loss rates have dynamically adjusted to approach their maximum biosynthesis rates.

Recalcitrance emerges as a community- and context-specific phenomenon that can change in time and space (SI Appendix, Fig. S10). Critically, the recalcitrance indicator for each OM pool ($Q_j$) is defined as the maximum of multiple population-specific values (Eq. 3)—one for each population $j$ that consumes pool $i$. Consequently, whether each pool is functionally labile or...
recalcitrant depends on the local microbial community. For a diverse community of consumers, we can analyze the fraction of the community that experiences each pool as recalcitrant (SI Appendix, Fig. S11). This community dependency implies that, statistically, functional recalcitrance may be more prominent when OM is exposed to a lower diversity of heterotrophic microorganisms. This also implies that functional recalcitrance may arise from the requirement for specialized enzymes or expensive consumption pathways for some types of OM (45): If specialization is required, there may be fewer possible consumers overall, and so it becomes less likely that any one consumer is present in the given environment. This is consistent with evidence that specific heterotrophic clades consume carboxyl-rich alicyclic molecules, which comprise a significant fraction (up to 8%) of marine DOC (6, 46).

**Unification of Hypotheses.** The three current hypotheses for DOM accumulation in the ocean—recalcitrance, dilution, and dependency on ecosystem properties—each explain aspects of the total amount of carbon in the model. Additional processes, such as mineral protection and diverse redox conditions of soils and sediments, can also be incorporated into the framework to modify it for these other systems. 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 is the sum of all OM pools. A traditional view of recalcitrance, focused on intrinsic qualities of the substrate or of the microbe–substrate-specific reaction, is represented in the model by the biomass yield $y_j$ and maximum uptake rate $v_j^{\text{max}}$. The quality of electron acceptor or mineral protection can also be represented by these parameters. As $y_j$ becomes small, while other parameters remain constant, OM becomes recalcitrant (Eq. 3), and the total organic carbon pool becomes large. The number of OM pools $n$ that are present can impact total carbon in two opposing ways. As $n$ increases, total carbon increases, even for low, equilibrated subsistence concentrations (the dilution hypothesis). 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 total carbon. Dependency on ecosystem properties is encapsulated in the population’s steady-state loss rate $L_j$ and probability of presence $P_j$. As $L_j$ increases, OM becomes recalcitrant, and total carbon increases. As the frequency of nearby consumers decreases, $P_j$ decreases, increasing total carbon.

The degree to which each mechanism controls OM accumulation in different environments therefore depends on the parameter space that sets the population and OM characteristics. Here, we assume uniform distributions for these parameters using plausible ranges for the ocean (Table 1; SI Appendix, SI Text 1). These ranges will vary with the environment. For example, if stochasticity in population presence does not apply to a given sediment ecosystem, then probability of presence $P$ may be set to one for analysis of that environment. The model is consistent with that of ref. 17 in that intrinsic recalcitrance is not necessary for OM accumulation in the ocean, as well as with experimental evidence for dilution-limited consumption (SI Appendix, SI Text 5 and Fig. S12) (29). Here, we provide a generalized framework that encapsulates a more complete set of dynamics than in ref. 17—one that is also consistent with evidence of recalcitrance (8, 18, and 19) and the impact of the microbial community (31, 33, 34, 45, 46). The emergent 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 (27) (SI Appendix, Fig. S13). They are also consistent with continuum intrinsic reactivity models, which assume a wide distribution of rates (10, 11) which tend toward lognormal distributions (47).

**Predicting OM Accumulation Patterns.** Our framework can help explain large-scale patterns in OM accumulation. Here, we use our model to understand the vertical structure of dissolved organic carbon (DOC) in the ocean. Globally, DOC concentrations peak at the sea surface and approach a minimum at depth (Fig. 3A) (6). Since the stochastic model is not practical for multidimensional 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 photosynthetic and heterotrophic microbial populations (SI Appendix, SI Text 6 and Fig. S14). The model is integrated for 6,000 y to quasi-equilibrium (SI Appendix, SI Text 6).

Ecological interactions in the model result in characteristics typical of a marine water column (Fig. 3; SI Appendix, Fig. S15). 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. 3 B and 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 (7, 48).

Many DOC pools in the model accumulate at the surface and become depleted at depths of 500 to 1,000 m. This transition is due to the increase in $Q$ for these pools from the surface to depth (Fig. 3B). Specifically, the loss rates of the populations are highest at the surface and attenuate with depth. This is because productivity peaks in the surface, and so total biomasses, activity rates, and, therefore, predation rates (represented implicitly in the model; Eq. 6) also peak at the surface. The subsistence concentrations for the functionally labile pools also decrease with depth as loss rates decrease, and so the total concentration of functionally labile OM also decreases with depth, contributing slightly to the vertical DOC gradient. Thus, an ecologically determined transition from functional recalcitrance to functional lability for some pools explains much of the decrease in DOC with depth. This transition is consistent with observations that a subset of DOC is recalcitrant to consumption by surface communities, but able to be remineralized by deep communities (31).

Our framework may also be employed to investigate microbial control on OM in soils and sediments. The model can be adapted to incorporate the different characteristics of these environments. For example, here, we employ a simple parameterization for the supply of each OM class, but a sediment or soil model version could include more sophisticated descriptions of how the physics and chemistry of solid particles and mineral matrices impact the supply rate. Though Michaelis–Menten uptake kinetics do not apply to the enzymatically catalyzed degradation of polymeric organic compounds to monomeric compounds, the ecological principles of our framework should still hold (SI Appendix, SI Text 7). Indeed, we find that, even in its current form, the simple model captures a key observation of sediment OM: the proportional increase in OM decomposition rates with increased OM concentration (49) (SI Appendix, Fig. S16). This further demonstrates consistency with the predictions of established first-order kinetic decomposition models (12, 49). Our framework can also be used to explore the impact of more enzymatically diverse sedimentary communities relative to pelagic communities on OM accumulation (33) by altering the community consumption matrix to include a greater degree of generalist ability. Also, varying the yields or uptake rates with electron acceptors could incorporate diverse redox conditions into the model. A decrease in yield with a lower-quality electron acceptor may suggest that some types of OM are functionally labile.
Fig. 3. Marine ecosystem water-column model results showing the accumulation of DOC. (A) Phytoplankton biomass, total DOC-consuming biomass $B$, and total DOC. Annual average profiles of total DOC from two open ocean time series stations are illustrated: HOT (the Hawaii Ocean Time-series in the Pacific Ocean) and BATS (the Bermuda Atlantic Time-series Station in the Atlantic Ocean) (50). (B) The concentration of each of the 25 resolved DOC pools, which are differentiated in the water-column model by maximum uptake rate $\rho_{\text{max}}$ (color scale). Each pool is categorized as functionally labile (solid line) or functionally recalcitrant (dashed line) as a function of depth using recalcitrance indicator $Q$ (Eq. 3). (C) The maximum (max) (surface) concentration $C$ of each DOC pool and the associated diagnostic $C^*$, the subsistence concentration of the microbial consumer population (Eq. 2), plotted against the maximum uptake rate for that pool. (D) The turnover time of each DOC pool calculated diagnostically from the integrated concentration and the integrated consumption rate, plotted against the maximum uptake rate.

when oxygen is available, but functionally recalcitrant in anoxic environments.

**Implications.** Our model is consistent with the observations and previous sediment modeling results that the majority of the diverse types of OM are present at relatively low (<1 µM C) concentrations, while the majority of the total standing stock is functionally recalcitrant (8, 10, 51) (Figs. 2B and 3B). The recalcitrant portion may equilibrate if subjected to abiotic concentration-dependent sinks (8) or may change slowly with time (7, 8). Our framework further emphasizes that apparently slow consumption rates of recalcitrant DOC in the ocean may be controlled by the frequency of encounter of “the right” populations and substrates, in addition to biochemical and energetic 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 (32).

The water column model links the millennial timescales of OM turnover (24) to microbial consumption occurring on subannual timescales (Fig. 3D). Although OM transformation through a complex interaction network can also explain old carbon ages (17), slow turnover as an additional mechanism is consistent with inferences that the size of organic carbon reservoirs does not reach a steady state over geologic timescales (2). While our model is compatible with the dilution hypothesis, it also incorporates the other explanations for accumulation, and so it is consistent with a broader set of observations, including the compositional uniformity of ubiquitous recalcitrant classes (7, 48).

A key aspect of our framework is the threshold behavior of the accumulation. The threshold, $Q = 1$, is set by the dynamics of the microbial populations that consume the OM pools. $Q = 1$ represents an ecological threshold along a continuum of OM and microbial characteristics, including factors known to influence recalcitrance, such as thermodynamic limitations (40), enzymatic control (33), mineral protection (41, 44, 52), and molecular properties (19). The nonlinear behavior of the threshold suggests that small changes in the environment can drive large depletions or accumulations of OM.

Consumption of recalcitrant OM depends on the rate of microbial processing, which increases with temperature. If other factors remain constant, the model predicts that less OM accumulates at higher temperatures (SI Appendix, SI Text 3 and Fig. S10C). Indeed, the loss of soil OM is a likely positive feedback to current warming (53). The framework here additionally suggests that a decrease in OM with warming may be nonlinear due to some OM pools crossing the threshold from functionally
More substrates

Faster population turnover

Less environmental connectivity

Total organic matter (μM C)

0.05
0.5
5

Slower microbial processing

Fig. 4. Controls on OM accumulation by microbial consumption. Starting from a representative, arbitrary concentration in the center, the change in total OM carbon is calculated for a 10-fold change in each of four parameters (i.e., two parameters vary in each quadrant): slower microbial processing via a reduced maximum uptake rate, faster turnover via an increased population-loss rate, less connectivity via a reduced likelihood of population presence, and more substrates (chemical diversity) via a greater number of OM pools.

The framework consistent with a comprehensive set of observations and theory by the connectivity of the environment (Fig. 4). The model is dynamics, and the ecological and biogeochemical diversity set the heterogeneity of organic substrates, microbial community loss rates, as well as an understanding of microbial community turnover rooted in the complexity of microbial ecosystems. Previ-
ously 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.

Materials and Methods

Model Equations. We describe 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 model the uptake \( \rho_i(t) \) of each OM pool \( i \), according to its concentration \( C_i \), by microbial population \( j \) as a function of time \( t \) using a saturating (Michaelis–Menten) form as

\[
\rho_i(t) = \frac{\rho_{i\text{max}}}{1 + \frac{C_i(t)}{k_i}},
\]

where \( \rho_{i\text{max}} \) is the maximum uptake rate and \( k_i \) is the half-saturation constant (Table 1).

Each population synthesizes biomass according to a growth efficiency for each pool (yield \( y_{ij} \)) and loses biomass at a rate proportional to its biomass according to a quadratic mortality parameter \( m^j \) (implicitly representing predators and viruses) and linear mortality parameter \( m^j \) (representing cell maintenance and senescence). The rates of change of the concentration \( C_i \) of pool \( i \) and the biomass \( B_j \) of population \( j \) are

\[
\frac{dC_i(t)}{dt} = s_i(t) - \sum_j l_{ij}(t) y_{ij} B_j(t),
\]

\[
\frac{dB_j(t)}{dt} = \sum_i l_{ij}(t) y_{ij} B_j(t) - m_i^j B_j(t)^2 - m_j^i B_j(t).
\]

where \( s_i(t) \) is the supply rate of pool \( i \), which is governed by the probability of the supply of each pool \( q_i \) as

\[
s_i(t) = \begin{cases} c_i & \text{with probability } q_i, \\ 0 & \text{with probability } 1 - q_i. \end{cases}
\]

Because the presence of population \( j \) averages to \( P_j \) over time, we include \( P_j \) in Eqs. 1–3 for conciseness. All parameter values (\( \rho_{i\text{max}}, k, m^j, m^j \)) are set by randomly sampling from uniform distributions (Table 1; SI Appendix, SI Text 1).

Yield \( y_{ij} \) reflects the cost of enzymes and the free energy released by OM oxidation. \( \rho_{i\text{max}} \) and \( y_{ij} \) may be interdependent due to cellular optimization strategies, reflecting inherent tradeoffs between protein allocation and efficiency (57, 58). The varying combinations of \( \rho_{i\text{max}} \) and \( y_{ij} \) can also represent the different modes of uptake of high-molecular-weight DOM (59). Analogously, the different parameter combinations can account for the additional feedback between the external concentration and the rate of cellular processing \( y_{ij} \). Real populations may change their cellular machinery due to plasticity, where, in the model, the many sets of parameters represent static phenotypes among these different modes.

Probability of Presence. Observations show that community composition dictates the character of DOM remineralization in seemingly unpredictable ways (61, 62), and evidence supports localized sinks of deep DOM (32). To simulate this impact, we include a population presence–absence dynamic in the model. Each population is assigned an overall probability of presence \( P_j \) (see detail below) as

\[
l_{ij}(t) = \begin{cases} 1 & \text{with probability } P_j, \\ 0 & \text{with probability } 1 - P_j. \end{cases}
\]

Because the population of presence \( j \) averages to \( P_j \) over time, we include \( P_j \) in Eqs. 1–3 for conciseness. All parameter values (\( \rho_{i\text{max}}, k, m^j, m^j \)) are set by randomly sampling from uniform distributions (Table 1; SI Appendix, SI Text 1).

Consumption Matrix. A consumption matrix dictates which populations consume which OM pools (SI Appendix, 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 (\( n_{\text{max}} \), 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 (\( n_{\text{cons}} \), which can vary from one to \( m \), the number of populations) (Fig. 1). In the model...
version with solely specialists (Fig. 2; SI Appendix, Fig. S2), each population consumes only one unique pool. For the generalist populations (Fig. 2; SI Appendix, Fig. S2), we first randomly assign n_p to each population drawing from the linear range from one to n. Second, we assign a weight to each OM pool of its probability of being consumed \( \rho \_\text{pool} \). Varying the weights linearly, 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.

Simulations. In the simulations illustrated in Fig. 2, we resolve 1,000 OM classes, and 1,000 or 2,000 pools of biomass: a model version with 1,000 specialists, and a model version with the 1,000 specialists and an additional 1,000 with a range of generalist ability. Results with the latter 2,000 pools are similar to a model with only the 1,000 generalists. For each experiment, we integrate the model forward in time for 10 \( y \), until the pools that have the potential to equilibrate have equilibrated. The concentrations of many of the recalcitrant pools continue to increase over time (unless a concentration-dependent sink is added to the model). SI Appendix, Fig. S13 illustrates the resulting distributions of biomass concentrations, OM concentrations, and remineralization rates of the ensembles, which are consistent with observed and inferred distributions of OM characteristics, remineralization rates, and ages in the ocean, sediments, soils, and lakes (10, 21–27, 63–65).

In SI Appendix, SI Text 2 and Figs. S2–S7, we demonstrate the qualitative consistency of the solutions across variations of the model. All simulations support the conclusions presented. Solutions vary quantitatively, but not qualitatively, with variations in the 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 updating the pools. In the present paper (additive consumption vs. switching over time to optimize growth). In the model version, where generalists switch their consumption over time (SI Appendix, Fig. S7), values of \( Q < 1 \) result for some pools as generalists cease to functionally recalcitrant pools, despite their capability to do so.

Reduced-Complexity Model Version. For the reduced-complexity model version 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 (\( n = 25 \)) (SI Appendix, Fig. S15). The values of \( x, m, m', m'', m''' \), and uptake affinity are kept constant, since their variation affects the solutions quantitatively, but not qualitatively. A specialist population, which represents multiple clades in aggregate, consumes each pool. Since we don't include stochastic processes, the probability of presence \( P \sim 1 \) for all populations. In accordance with theory and our stochastic model results (27, 66), we assume a lognormal distribution for the partitioning of total OM production into the 25 pools (SI Appendix, Fig. S15D), which represents the average outcome of microbial transformation over time and space.

Marine Ecosystem Model. The reduced-complexity model version is incorporated into a dynamic marine ecosystem model of a stratified vertical water column, where the production and consumption of all organic and inorganic pools are due to the growth, respiration, excretion, and mortality of microbial populations (SI Appendix, Fig. S14). Light and vertical mixing attenuate with depth. Two populations of phytoplankton convert dissolved inorganic carbon and nitrogen into biomass using light energy. Populations of microbes, with solely specialists (Fig. 2; SI Appendix, Fig. S15), each population consumes only one unique pool. For the generalist populations (Fig. 2; SI Appendix, Fig. S2), we first randomly assign n_p to each population drawing from the linear range from one to n. Second, we assign a weight to each OM pool of its probability of being consumed \( \rho \_\text{pool} \). Varying the weights linearly, 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.
32. D. A. Hansell, C. A. Carlson, Localized refractory dissolved organic carbon sinks in the deep ocean. Global Biogeochem. Cycles 27, 705–710 (2013).
33. C. Arnott, Microbial extracellular enzymes and the marine carbon cycle. Ann. Rev. Mar. Sci. 3, 401–425 (2011).
34. Y. Shen, R. Benner, Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon. Sci. Rep. 8, 2542 (2018).
35. S. Evans, J. B. Martiny, S. D. Allison, Effects of dispersal and selection on stochastic assembly in microbial communities. ISME J. 11, 176–185 (2017).
36. S. J. Pirt, The maintenance energy of bacteria in growing cultures. Proc. R. Soc. Lond. Ser. B Biol. Sci. 183, 224–231 (1965).
37. M. J. R. Fasham, H. W. Ducklow, S. M. McKelvie, A nitrogen-based model of plankton dynamics in the ocean mixed layer. J. Mar. Res. 48, 591–639 (1990).
38. R. L. Sinsabaugh et al., Stoichiometry of soil enzyme activity at global scale. Ecol. Lett. 11, 1252–1264 (2008).
39. D. L. Kirchman, Mixing it up in the ocean carbon cycle and the removal of refractory dissolved organic carbon: A thermodynamic analysis. Geochem. Cosmochim. Acta 75, 2030–2042 (2011).
40. D. E. LaRowe, P. Van Cappellen, Degradation of natural organic matter: A thermodynamic analysis. Global Biogeochem. Cycles 29, 1389–1407 (2015).
41. J. D. Hemingway et al., Mineral protection regulates long-term global preservation of natural organic carbon. Nature 570, 228–231 (2019).
42. E. Grabowski, R. M. Letelier, E. A. Laws, D. M. Karl, Coupling carbon and energy fluxes in the North Pacific Subtropical Gyre. Nat. Commun. 10, 1895 (2019).
43. B. Guenet et al., Impact of priming on global soil carbon stocks. Global Change Biol. 24, 1873–1883 (2018).
44. V. L. Bailey, C. H. Pries, K. Lajtha, What do we know about soil carbon destabilization? Environ. Res. Lett. 14, 083004 (2019).
45. A. Sichert et al., Verrucomicrobia use hundreds of enzymes to digest the algal polysaccharide fucoidan. Nat. Microbiol. 5, 1026–1039 (2020).
46. S. Liu et al., Different carboxyl-rich aliphatic molecules proxy compounds select distinct bacterioplankton for oxidation of dissolved organic matter in the mesopelagic Sargasso Sea. Limnol. Oceanogr. 65, 1532–1553 (2020).
47. E. W. Montroll, M. F. Shlesinger, On 1/f noise and other distributions with long tails. Proc. Natl. Acad. Sci. U.S.A. 79, 3380–3383 (1982).
48. L. I. Aluiahe, D. J. Repeta, R. F. Chen, A major biopolymeric component to dissolved organic carbon in surface sea water. Nature 387, 166–169 (1997).
49. J. T. Westrich, R. A. Berner, The role of sedimentary organic matter in bacterial sulfate reduction: The G model tested. Limnol. Oceanogr. 29, 236–249 (1984).
50. R. T. Letscher, J. K. Moore, 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).
51. G. I. Agren, E. Bosatta, Theoretical Ecosystem Ecology: Understanding Element Cycles (Cambridge University Press, Cambridge, UK, 1998).
52. T. M. Blattmann et al., Mineralogical control on the fate of continentally derived organic matter in the ocean. Science 366, 742–745 (2019).
53. T. W. Crowther et al., Quantifying global soil carbon losses in response to warming. Nature 540, 104–108 (2016).
54. P. Gais et al., Large inert carbon pool in the terrestrial biosphere during the Last Glacial Maximum. Nat. Geosci. 5, 74–79 (2012).
55. O. Cartapanis, D. Bianchi, S. L. Jaccard, E. D. Galbraith, Global pulses of organic carbon burial in deep-sea sediments during glacial maxima. Nat. Commun. 7, 10796 (2016).
56. C. Santinelli, L. Nannicini, A. Seritti, DOC dynamics in the meso and bathypelagic layers of the Mediterranean Sea. Deep. Res. Part II Top. Stud. Oceanogr. 57, 1446–1459 (2010).
57. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of cell growth and gene expression: Origins and consequences. Science 330, 1099–1102 (2010).
58. M. Basan et al., Overflow metabolism in Escherichia coli results from efficient proteome allocation. Nature 528, 99–104 (2015).
59. G. Reinjers, C. Arnosti, B. Fuchs, R. Amann, Selfish, sharing and scavenging bacteria in the Atlantic Ocean: A biogeographical study of bacterial substrate utilisation. ISME J. 13, 1119–1132 (2019).
60. S. D. Allison, S. S. Chacon, D. P. German, Substrate concentration constraints on microbial decomposition. Soil Biol. Biochem. 79, 43–49 (2016).
61. J. B. Logue 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).
62. J. P. Balmonte et al., Community structural differences shape microbial responses to high molecular weight organic matter. Environ. Microbiol. 21, 557–571 (2019).
63. B. P. Koch, M. Witt, R. Engbrodt, T. Dittmar, G. Kattner, Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Geochem. Cosmochim. Acta 69, 3299–3308 (2005).
64. B. Koehler, E. Von Wachenfeldt, D. Rothman, D. C. Forney, Physical model for the decay and preservation of marine organic matter in the ocean. Deep. Res. Part II Top. Stud. Oceanogr. 65, 1532–1553 (2020).
65. E. J. Zakem, N. M. Levine, Systematic variation in marine dissolved organic matter stoichiometry and remineralization ratios as a function of lability. Global Biogeochem. Cycles 33, 1389–1407 (2019).
66. E. Zakem, OMconsumption. GitHub. https://github.com/emilyzakem/OMconsumption. Deposited 19 November 2020.