UC San Diego
UC San Diego Previously Published Works

Title

Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter

Permalink

Journal

ISSN

Authors

Arakawa, Neal
Aluwihare, Lihini I
Simpson, Andre J
et al.

Publication Date

DOI

Peer reviewed

GEOCHEMISTRY

Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter

Neal Arakawa,[1] Lihini I. Aluwihare,[1]* Andre J. Simpson,[2] Ronald Soong,[2]
Brandon M. Stephens,[1] Daniel Lane-Coplen[2]

The ocean's biota sequester atmospheric carbon dioxide (CO_2) in part by producing dissolved organic matter (DOM) that persists in the ocean for millennia. This long-term accumulation of carbon may be facilitated by abiotic and biotic production of chemical structures that resist degradation, consequently contributing disproportionately to refractory DOM. Compounds that are selectively preserved in seawater were identified in solid-phase extracted DOM (PPL-DOM) using comprehensive gas chromatography (GC) coupled to mass spectrometry (MS). These molecules contained cyclic head groups that were linked to isoprenoid tails, and their overall structures closely resembled carotenoid degradation products (CDP). The origin of these compounds in PPL-DOM was further confirmed with an in vitro β-carotene photooxidation experiment that generated water-soluble CDP with similar structural characteristics. The molecular-level identification linked at least 10% of PPL-DOM carbon, and thus 4% of total DOM carbon, to CDP. Nuclear magnetic resonance spectra of experimental CDP and environmental PPL-DOM overlapped considerably, which indicated that even a greater proportion of PPL-DOM was likely composed of CDP. The CDP-rich DOM fraction was depleted in radiocarbon (^{14}C age > 1500 years), a finding that supports the possible long-term accumulation of CDP in seawater. By linking a specific class of widespread biochemicals to refractory DOM, this work provides a foundation for future studies that aim to examine how persistent DOM forms in the ocean.

INTRODUCTION

The ocean contains approximately 700 gigatons of carbon (1) in the form of dissolved organic matter (DOM). This large reservoir size has drawn attention to the possible role of DOM in modulating Earth's past climate (2). Parts of this reservoir accumulate on long time scales, as evidenced by depleted radiocarbon (^{14}C) values for DOM throughout the ocean (3). A variety of methods can be used to extract ^{14}C-depleted (refractory) DOM from seawater (4–7), and two-dimensional (2D) nuclear magnetic resonance (NMR) experiments have provided new insights into the chemical composition of this material (8–11). Available NMR data identify specific chemical features that relate some DOM molecules to the broad chemical class of terpenoids. Specifically, the identification of alicyclic methylene groups has been used to support the presence of fused cyclic terpenoids similar to steroids and hopanoids (8, 11), whereas conjugated double bonds and branched methyl groups in close proximity to various oxygen-bearing carbon atoms have been linked to oxidized derivatives of linear terpenoids such as phytoene and zeaxanthin (9, 10). These classes are considered to be distinct and commonly referred to in the DOM literature as carboxyl-rich alicyclic material (CRAM) (8) and material derived from linear terpenoids (MDLT) (9), respectively.

The CRAM model, in particular, has been assimilated broadly into literature discussions of refractory DOM cycling. Both CRAM and MDLT are based on a "best-fit" logic in that they are structures that were generated to be consistent with some NMR and elemental ratio features in DOM. So far, the chemical complexity of DOM has obstructed molecular-level confirmation of either or both terpenoid-based models. Thus, any links between refractory DOM and biosynthetic precursors remain hypothetical. Backbone structures resembling terpenoids have been detected in terrestrial DOM by mass spectrometry (MS) following

chromatographic separation of thermally or chemically degraded DOM (12, 13). However, these studies did not demonstrate how common terpenoid biomolecules (alicyclic or linear) could be transformed into the structures currently proposed for refractory DOM. In summary, CRAM and MDLT models identify key molecular features that exist in DOM, but the manner in which these features are connected in DOM molecules remains to be confirmed. For example, methylene groups in fused ring structures have been detected based on NMR resonances, and some of these methylene groups contain carboxylic acids near them (8, 11). However, no molecules containing these features and whose backbones can be related to biosynthetic precursors have been identified in DOM. Analytical methods capable of identifying precise contiguous carbon skeletons in refractory DOM are necessary to link biosynthetic precursors to refractory DOM.

To confirm the terpenoid model, this study examined DOM isolated from a coastal site in the eastern North Pacific Ocean using solid-phase extraction (SPE) with PPL resin (PPL-DOM) (5). Molecular-level identification of terpenoids was facilitated by gas chromatography (GC), which was used to separate the complex PPL-DOM mixture. Before using GC, oxygen functional groups abundant in PPL-DOM were reduced because their presence in DOM limited the efficiency of chromatographic separation. However, it was also necessary to preserve the contiguous carbon skeletons present in PPL-DOM to ensure that the distinct isoprenoid-based carbon backbone of terpenoids could be identified. To satisfy these requirements, the current study applied a chemical reduction method that had been previously shown to effectively remove oxygen functional groups while preserving carbon backbones (12). Once PPL-DOM was reduced, it was analyzed using comprehensive GC (GC×GC) coupled to MS to identify specific terpenoid molecules. The relative abundances and chemical composition of terpenoid molecules in PPL-DOM before the reduction were further examined with 2D NMR techniques. To confirm the structural assignments made by GC-MS and NMR, a simple laboratory experiment was conducted using a model compound, β-carotene (BC), as a representative of a widespread class of terpenoids (14). Together, the results

[1]Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093–0244, USA. [2]Environmental Nuclear Magnetic Resonance Center, University of Toronto, Scarborough Campus, Toronto, Ontario M1C 1A4, Canada.
*Corresponding author. Email: laluwihare@ucsd.edu

presented here provide the first direct confirmation of the abundance of isoprenoid structures in marine DOM.

RESULTS AND DISCUSSION

The marine PPL-DOM isolated from surface waters near the Scripps Institution of Oceanography (SIO) represented 39.6% of measurable dissolved organic carbon (DOC) in the original seawater. The sample had a proton (^{1}H) NMR spectrum consistent with other PPL-isolated samples (11). It also represented a temporally consistent fraction of this pool, as evidenced by an NMR spectrum that was nearly identical to other PPL samples from a monthly time series at the SIO Pier (fig. S1 and table S1).

For initial characterization by GC, this PPL-DOM was chemically reduced with n-butylsilane (n-BS) and a tris(pentafluorophenyl)borane $[B(C_6F_5)_3]$ catalyst (12, 15). The reduction converts most oxygen-containing functional groups (for example, esters, carboxylic acids, and alcohols) to their respective hydrocarbon backbones (scheme S1) and makes PPL-DOM constituents amenable to separation by GC×GC.

Insights from chemical reduction

The chemical reduction of SIO Pier PPL-DOM yielded 10% of its original carbon as closely eluting GC-amenable products, as quantified by flame ionization detection (FID). Comprehensive GC×GC MS (Fig. 1A) revealed these products to be several related series of isoprenoid-based hydrocarbons, with alicyclic groups similar to those previously reported in terrestrial DOM (12). These alicyclic groups were identified from the observed chemical formulas of reduction products together with knowledge of the reaction mechanism revealed from reducing standard compounds. The compositional coherency of marine DOM reduction products, as revealed by GC×GC time-of-flight (TOF)–MS, resembled results from direct infusion, high-resolution MS studies that inferred the presence of homologous series in DOM based on molecular formula assignments (16–18).

Two prominent reduction products (Fig. 1A, black structures) were completely structurally characterized by matching their retention times (^{1}RT, ^{2}RT) and fragmentation patterns with those of reduced model compounds (fig. S2). The GC-MS identification of these carbon skeletons (Fig. 1) unambiguously confirmed previous NMR reports implicating terpenoids as major components of DOM (8, 9). These structures further identified carotenoids as plausible biosynthetic precursors of some fraction of PPL-DOM.

Once carotenoids were implicated as possible source materials of PPL-DOM, a simple laboratory experiment was devised to test whether related compounds could be produced from a model carotenoid. Specifically, BC in filtered seawater or Milli-Q was photooxidized to mimic abiotic processes that could generate water-soluble carotenoid degradation products (CDP). As a control, an identical experiment was set up in the dark. Exposure of BC to light successfully produced measurable quantities of oxidized CDP that were water-soluble and could be extracted in a manner similar to that of PPL-DOM from both seawater and Milli-Q (fig. S3). The dark, seawater incubation also produced some water-soluble compounds (~2% conversion of the starting material), but the recoveries were too small to allow structural characterization. Once extracted, the photooxidized CDP samples were reduced and analyzed by GC×GC TOF-MS (Fig. 1B). Retention times and MS fragmentation patterns were used to confirm that structures identified in PPL-DOM were also present in the reduced CDP (Fig. 1, A and B, black structures). The CDP sample also contained additional structures not directly identified in reduced PPL-DOM (Fig. 1B, in gray). Even without authentic standards, these structures could be inferred based on consistent retention time offsets and mass fragmentation patterns.

The chromatograms of reduced marine DOM and reduced water-soluble CDP contained compounds easily identified by the 95 mass/charge ratio (m/z) ion ($C_7H_{11}^+$), which appeared in GC×GC space in contiguous patterns indicative of structurally related compounds arranged in series from C_{10} to >C_{15} (Fig. 1). Identified structures and elution patterns in Fig. 1B suggested that the mechanism of CDP formation, following exposure to seawater and sunlight, cleaved carbon-carbon bonds along the entire isoprenoid backbone of BC (Fig. 1B, inset) and not only at double-bond sites (for example, C_{10} and C_{11} compounds are detected). Particularly interesting was the prominence of $C_{12}H_{22}$ (Fig. 1B, second gray structure), which, based on the structure of BC (Fig. 1B, inset), would necessitate the breaking of two carbon-carbon bonds. Water-soluble CDP were also analyzed by GC-MS before reduction, and this analysis detected several precursor compounds that, upon reduction, would produce the alicyclic hydrocarbons observed in reduced CDP (fig. S4). These structures could be representative of some native compounds in PPL-DOM.

The isomeric diversity of reduction products could be visualized in the chromatogram as many peaks of identical mass but distinct GC×GC retention times. Related isomers were clustered within a particular GC×GC space defined by dashed ovals in Fig. 1 (A and B). Many of the unresolvable compounds in Fig. 1A could also be assigned to CDP (fig. S5). These data unequivocally established CDP hydrocarbon backbones within PPL-DOM but could not determine whether these backbones originated from identical unreduced compounds in each sample (fig. S4). The distribution of isomers differed noticeably between Fig. 1A and Fig. 1B, which suggested that starting materials in PPL-DOM were not adequately represented by BC alone. Such differences are consistent with both the diversity of terpenoids, including carotenoids, in nature and the numerous CDP formation mechanisms that are possible.

NMR confirmation of CDP in PPL-DOM

Only 10% of PPL-DOM carbon was recovered from the reduction and definitively linked to CDP, but several lines of evidence indicated that the recovery of CDP from PPL-DOM was low and underestimated the CDP contribution. For example, double bonds associated with CDP are likely to participate in cross-reactions during the reduction (12), and so, compounds containing these functional groups can be lost from the sample during cleanup. To obtain more insights into the relative abundance of CDP in PPL-DOM, 2D NMR experiments were performed on unreduced samples. Two different 2D NMR experiments were applied to CDP and a terpenoid-rich fraction of bulk PPL-DOM (Fig. 2). The latter fraction was isolated following mild acid hydrolysis of PPL-DOM, which liberated several small, polar compounds (sugars, amino acids, and small acids) that could be removed with a second PPL cleanup. The terpenoid-rich fraction was less complex, and its 2D NMR spectrum differed from PPL-DOM primarily in regions where resonances associated with carbohydrates and amino acids appeared (fig. S6). Ultimately, the hydrolysis enabled cleaner extraction of the terpenoid-specific H-C correlations by NMR spectroscopy.

Heteronuclear single-quantum coherence (HSQC) NMR experiments (Fig. 2, A to C) can detect the correlation between proton chemical shifts and the chemical shifts of their directly bonded carbons. Heteronuclear multiple-bond correlation (HMBC) experiments (Fig. 2, D to F) correlate long-range interactions between protons and carbons up to
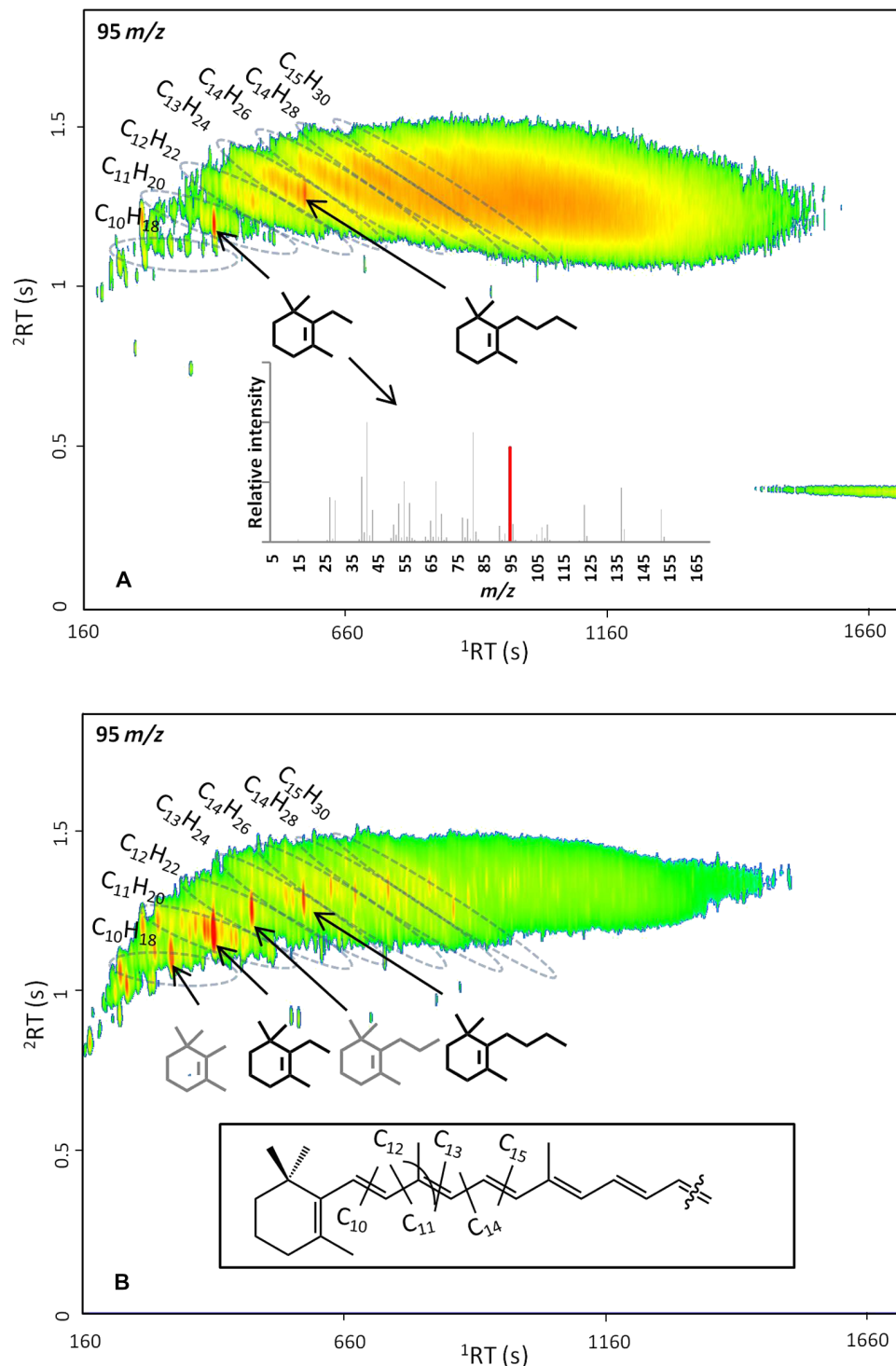
Fig. 1. GC×GC selected-ion chromatograms (95 *m/z*) of chemically reduced PPL-DOM and CDPs. (A and **B)** Ion abundances are displayed as green (low)→red (high). Individual compounds appear as sharp ellipses. Exact chemical structures are linked to two distinct peaks (black structures), and the mass spectrum is provided with 95 *m/z*, a prominent hydrocarbon fragmentation ion ($C_7H_{11}^+$), identified in red. Chemical formulas shown in the figure were validated with hydrocarbon standards. Gray structures were inferred on the basis of their similarity to confirmed products. The inset in (B) shows a generic carbon backbone of carotenes/carotenoids and identifies positions where carbon bond cleavage can form alicyclic compounds of sizes C_{10} to C_{15} denoted in the chromatogram.
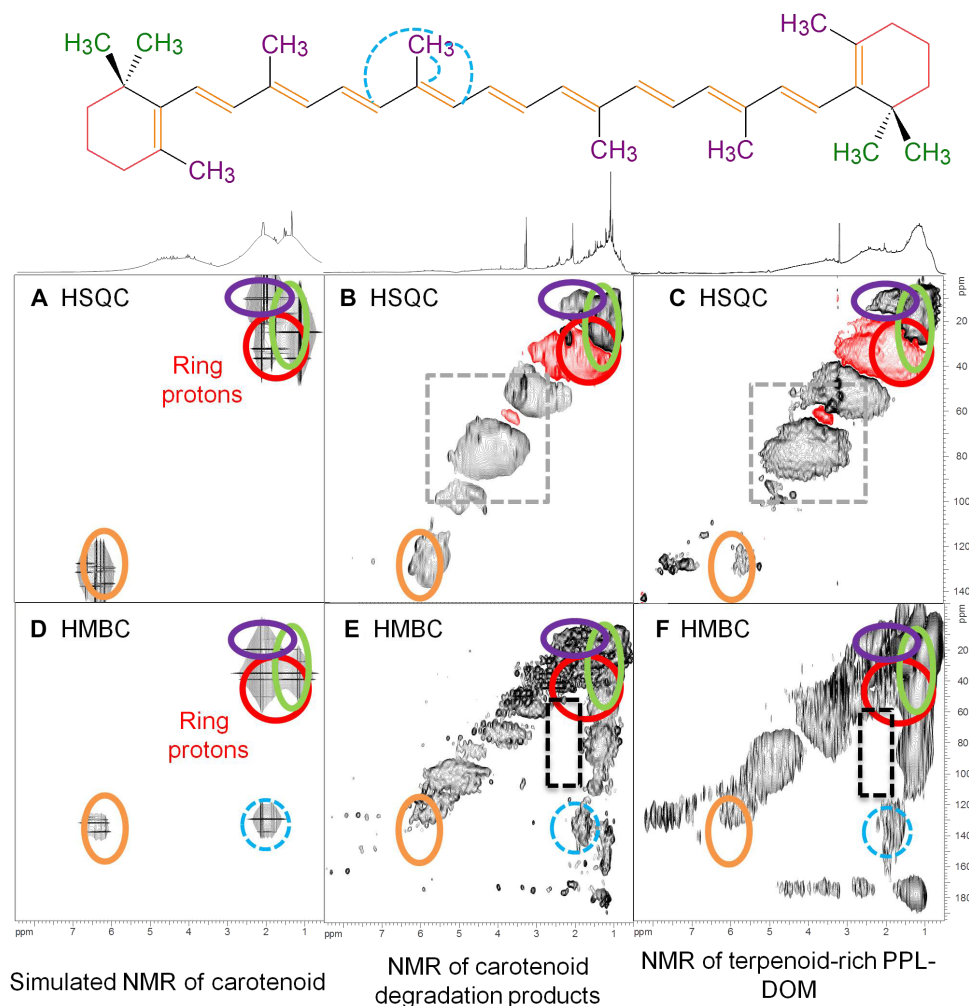
Fig. 2. Identification of CDP through ^{1}H-^{13}C NMR correlations. Simulated one-bond (**A**) and multiple-bond (**D**) modeled ^{1}H-^{13}C correlations for an intact carotene. Colors highlighted on the structure correspond to the same colored regions in each spectrum. Linewidth in the simulations (A and D) matches that of the main spectral features seen in DOM. This is to demonstrate the spectral area that may be occupied in complex mixtures of similar overlapping chemical shifts. Exact chemical shifts from the simulated structure were superimposed upon the simulated 2D NMR spectrum and appear as darker crosses (spikes in the 1D ^{1}H projection). One-bond (**B** and **C**) and multiple-bond (**E** and **F**) ^{1}H-^{13}C correlations for BC degradation products (that is, CDP) and terpenoid-rich PPL-DOM, respectively, with overlaid regions corresponding to key one-bond correlations observed in the model compound. Blue regions in HMBC spectra correspond to long-range correlations. The black dashed rectangle (E and F) identifies the region where correlations that would shift methylene protons further downfield are expected but do not appear. This is discussed in greater detail in the main text.

three bonds away. Thus, whereas HSQC experiments identify H-C fragments in a structure, HMBC experiments explain how these H-C fragments link together in extended substructures. Modeled (in silico) HSQC (Fig. 2A) and HMBC (Fig. 2D) spectra of intact BC were used to assign several of the correlations observed in experimental CDP NMR spectra (Fig. 2, B and E). Key one-bond correlations from conjugated double bonds are designated as color-coded ovals in Fig. 2A (orange, olefin-H to olefin-C; purple, methyl-H to carbon α to olefins). In Fig. 2D, long-range correlations between double bonds and attached methyl groups were also identified (blue dashed circle). The presence of these correlations in 2D NMR spectra of experimental CDP (Fig. 2, B and E) showed that CDP still retained observable double-bond character, although they were oxygenated enough to be dissolved in seawater. The terpenoid-rich fraction of PPL-DOM contained several major correlations observed in the experimental CDP sample, which indicated that remnants of relatively intact CDP were more prominent in ma-

rine DOM (Fig. 2, C and F) than estimated above with molecular-level identifications.

These NMR data confirmed that terpenoid-rich PPL-DOM and experimental CDP shared many dominant molecular features. The two samples shared a particular set of correlations that were especially important: the correlation between methyl resonances in the NMR spectrum (green oval) and the varied chemical environments along the carbon scale from 30 to 180 parts per million (ppm) (Fig. 2, E and F). These HMBC correlations highlighted the proximity (two to three bonds away) of methyl groups to variously functionalized methine and methylene carbons and carboxylic acids. This set of correlations is a prominent and relatively unique feature of PPL-DOM and was previously presented as evidence of dissolved degraded linear terpenoids in aquatic environments (9).

Methylene proton resonances that were shifted downfield in the methylene region (proton chemical shifts between ~2.5 and 3.5 ppm)

were present in CDP and PPL-DOM. These resonances are unusual because they are deshielded (surrounded by electronegative functional groups) but only correlated to carboxylic acid groups as detected by HMBC. These resonances have been previously linked to carboxylated sterol–like compounds (for example, CRAM) (8, 9). The fact that experimental CDP contained similar correlations required that similarly placed proton resonances in PPL-DOM could not be attributed exclusively to sterol precursors.

1D ^{1}H NMR spectra of experimental CDP and terpenoid-rich PPL-DOM were not identical (Fig. 2, projected ^{1}H NMR spectra). For example, PPL-DOM contained a higher relative abundance of proton resonances in the chemical shift region >2.0 to 3.5 ppm (Fig. 2, B and C, proton axes). Thus, additional compounds, derived from fused alicyclic ring precursors, may be present in PPL-DOM but absent in experimental CDP. Carboxylic acid resonances were also more prominent in PPL-DOM than in CDP. Mechanisms capable of producing –COOH groups may have been absent from the BC experiment, and/or the length of the incubation could have destroyed –COOH groups. Ultimately, these samples were not expected to be identical given that only one carotenoid was tested in this experiment but numerous starting materials are likely to be important, the experiment excluded biological transformations, and modification of DOM, including CDP, happens over much longer time scales in the ocean.

The presence of downfield methine and oxomethylene correlations [gray dashed square; $O_{(1,2)}CH_{(1,2)}$] in CDP represented another remarkable finding. These correlations that were also present in PPL-DOM had been previously attributed to carbohydrates (11). However, the prominence of these resonances in experimental CDP implicated oxidized aliphatic compounds as additional sources in PPL-DOM. Carbohydrates were identified in these PPL-DOM samples using GC-MS and NMR (for example, fig. S7), but they represented only a small component of the sample. The presence of aliphatic oxygenated groups in PPL-DOM is consistent with the fact that some of these resonances remained in the terpenoid-rich PPL-DOM fraction even after hydrolysable carbohydrates had been removed.

Of the hundreds of unique marine carotenoids, many share two defining features that are derived from isoprenoid units: cyclic head groups and branched methyl side chains, with conjugated double bonds (14). Resonances associated with alicyclic compounds in marine DOM were previously assigned to CRAM, a definition that originally referred to carboxylic acid–rich, fused alicyclic ring structures (8, 11). Linear isoprenoid NMR resonances in DOM were previously linked to MDLTs (9). However, new 2D NMR data reported here showed that cyclic terpenoid, linear terpenoid, and oxygenated aliphatic fragments, which are prominent and defining features of PPL-DOM, were also produced from the oxidation of carotenoids. By linking molecules with carotenoid features found in both PPL-DOM and experimental CDP to NMR data, a definitive connection has been made between common NMR resonances in DOM and an aliphatic, biosynthetic precursor.

## In silico transformation of BC into CDP to satisfy observed HMBC NMR correlations

In the previous section, predicted HMBC correlations for BC helped to identify double bond–containing material in CDP. Although these correlations were present in PPL-DOM and water-soluble CDP (Fig. 3, A and B, blue ovals), double bonds in BC must have been extensively oxidized to reproduce the water solubility of CDP. The conjugated double bonds are likely the most reactive subunit for oxidation to hydroxyl (OH) and carboxyl (COOH) functional groups (19).

In silico oxidation was performed on the heavily unsaturated main-chain unit in an effort to reproduce HMBC NMR correlations in CDP (Fig. 3B). Widespread oxygenation of the linear carbon backbone was necessary to approximate several of the observed NMR correlations (Fig. 3, B and C, black, green, and purple boxes). The presence of carboxylic acids is a well-known modification in DOM, as inferred from the common loss of COO$^-$ (44 Da) fragments from PPL-DOM when analyzed by electrospray ionization MS (20). Correlations of COOH carbon atoms to aliphatic protons in CDP (methyl and others; Fig. 3B, red square) were apparent from the HMBC spectrum as well (Fig. 3C). To approximate these observed NMR correlations, some methyl groups on the side chain of the model carotenoid had to be converted to carboxylic acids (fig. S8). The methyl correlations were then consistent with CDP (Fig. 3, B and C, black, green, and red boxes). It was not possible to distinguish carboxylic acid–containing products in experimental CDP by GC-MS. Therefore, mechanisms and sites of this carboxylation remain unknown. There is biological precedent for the position-specific enzymatic cleavage of the main chain to produce ketones and aldehydes that are subsequently oxidized to carboxylic acids (21), as well as the conversion of some branched methyl groups into COOH groups (22). However, the carboxylation mechanism in this experiment was most likely abiotic because the seawater in the BC experiment was prefiltered.

There were prominent features in the HMBC spectrum of CDP that were not anticipated by the in silico spectrum of oxidized BC (Fig. 3B). Most notably, the in silico model did not contain methylene proton resonances that were deshielded to the extent observed in CDP. In the predicted HSQC spectrum (Fig. 3B), these resonances appeared to be centered at 2 ppm, whereas in CDP and PPL-DOM spectra, these correlations extended past 3 ppm (Fig. 2, B and C, red shaded correlations). As discussed earlier, these resonances are typically assigned to degraded sterol-like, CRAM compounds (8). It is possible that the model HMBC spectrum did not adequately account for the deshielding effects of nearby functional groups (alcohols and olefins), especially when line broadening is taken into consideration. Alternatively, alicyclic head groups in CDP may be more oxidized than shown here for BC. Cyclization of some oxidized CDP could also lead to methylene resonances that resonate further downfield. In any case, available data are unable to identify the proximal functional groups that give rise to these methylene resonances in either CDP or PPL-DOM. However, the presence of these unusual resonances in both samples confirmed that oxidized carotenoids could give rise to resonances previously attributed exclusively to CRAM.

The oxomethylene correlation (Fig. 3C, orange box) in DOM has been largely attributed to carbohydrate-like material, yet CDP contained similar resonances. The oxidized carotenoid model structure presented here did not contain an oxomethylene group, and as a result, the model HMBC NMR (Fig. 3B, orange box) lacked the resulting correlation.

The oxidized carotenoid model was presented to facilitate discussion of CDP structures and was not intended to mimic the exact structures present in either CDP or PPL-DOM. The GC-MS data set demonstrated that the oxidation of BC resulted in a mixture of compounds in which the main chain had been cleaved at various points, whereas the chemical model for CDP (Fig. 3B) depicted one compound with an intact carotenoid backbone. The line broadening observed in the proton NMR spectrum of CDP further indicated that a complex mixture of compounds with overlapping resonances had been generated from BC. It is likely that individual CDP compounds contained highly oxidized regions and intact BC-like regions. Still, of the hundreds of
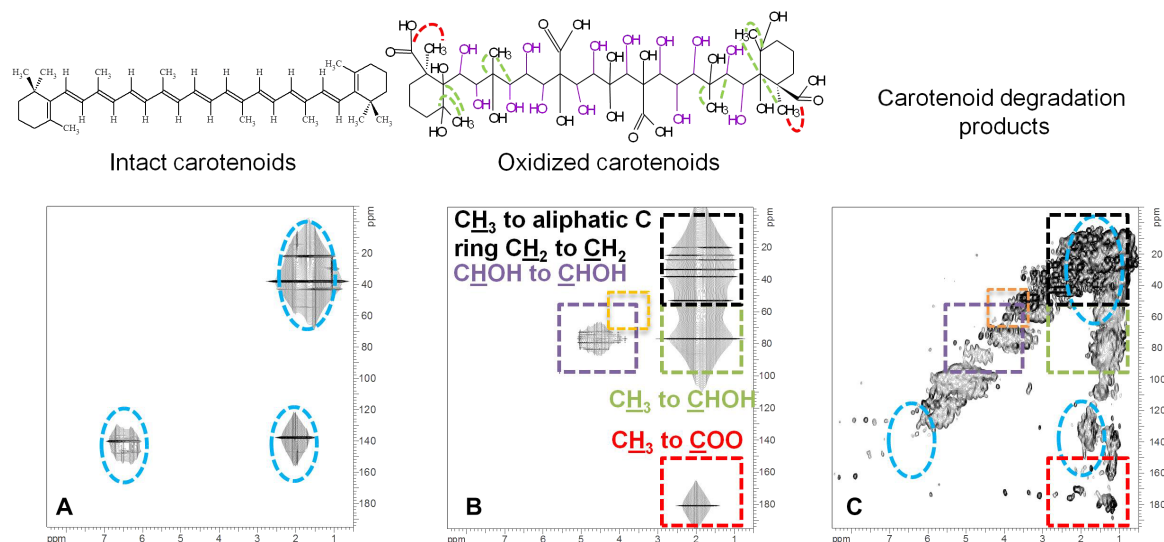
Fig. 3. Simulation of oxidized carotenoids through ^{1}H-^{13}C NMR correlations. (**A**) Simulated long-range ^{1}H-^{13}C correlations for BC. See Fig. 2 for the description of regions. (**B**) Simulated long-range ^{1}H-^{13}C correlations for oxidized carotene. (**C**) Observed long-range ^{1}H-^{13}C correlations for CDP with those correlations simulated in (A) and (B) superimposed. Linewidth in the simulations (A and B) matched that of the main spectral features seen in CDP. This was used to demonstrate the spectral area that may be occupied in complex mixtures of similar overlapping chemical shifts. The exact chemical shifts from the simulated structure were superimposed upon the simulated NMR and appear as darker crosses.

NMR model structures tested, the extent of oxidation presented in Fig. 3B most accurately represented the observed HMBC correlations in terpenoid-rich DOM and CDP. This preceding discussion highlights the need for more research into mechanisms that converted BC to CDP.

Quantification of CDP in PPL-DOM

On the basis of the GC-MS data, only 10% of PPL-DOM can be confidently attributed to CDP. Following the reduction, GC-MS directly identified molecular structures related to CDP, but 2D NMR data suggested that this identification underestimated the contribution of CDP to PPL-DOM. The reduction process has poor recoveries of model compounds with olefin groups (12), such as those on the conjugated main chain of carotenoids (Fig. 2). Olefins react with the reducing agent to form polymers that cannot be efficiently recovered. For example, a model linear diene was completely lost during the reduction, whereas reduction of a branched diene produced the corresponding alkane in only low yield (<25%) (12).

Following acid hydrolysis, 70% of the original PPL-carbon was recovered. Reintroducing the hydrolyzed sample onto a second PPL column separated the hydrolysate into a permeate dominated by small, polar biochemicals (12% of PPL-DOM carbon; fig. S8) and a retentate [eluted with methanol (MeOH)] enriched in DOM functional groups attributed to oxygenated, aliphatic compounds (fig. S8). Assignment of 2D NMR correlations in the preceding section (Fig. 2, B and C) indicated that the retained PPL-DOM was terpenoid-rich and exhibited near overlap of resonances with experimental CDP. Although 2D NMR resonances overlap, the relative contribution of CDP resonances to PPL-DOM cannot be quantified by this method. It is also clear from projected ^{1}H NMR spectra (Fig. 2, B and C, x axis) that differences exist between CDP and PPL-DOM. Also, we cannot preclude the possibility that compounds unrelated to CDP, such as carbohydrates not released from the hydrolysis, shared NMR resonances in PPL-DOM. For these multiple reasons, GC-MS/FID was chosen to provide the only reliable quantitative data and showed that at least 10% of PPL-DOM, or 4% of total DOC at this site, was derived from CDP. On the basis of the re-

covery of only 58% of PPL-DOM as terpenoid-rich DOM (fig. S8), the absolute upper limit for CDP-related compounds was 58% of PPL-DOM carbon.

Radiocarbon distribution in PPL-DOM

PPL-DOM had a measured radiocarbon signature of −183‰ (table S1), corresponding to a ^{14}C age of 1625 years. This value was consistent with previous reports for marine PPL-DOM (6, 7). The low–molecular weight polar compounds that were liberated upon acid hydrolysis of PPL-DOM could mask the potential, true radiocarbon signature of the terpenoid fraction. As noted previously, the terpenoid fraction was purified from these polar compounds using a second PPL step, which yielded 58% of the original PPL-carbon as terpenoid-rich DOM. This terpenoid-rich fraction had a measured radiocarbon value of −240‰, or 2700 radiocarbon years (table S1, fraction 1). These radiocarbon data, together with the abundance of CDP in terpenoid-rich DOM, suggested that CDP contributed some aged carbon to PPL-DOM. Previous NMR experiments using deep ocean DOM identified methyl fragments characteristic of intact and oxidized carotenoids (for example, Fig. 2 and fig. S6) (11). The radiocarbon value of −240‰ is still relatively enriched compared to deep ocean DOM (for example, −540‰) (3). However, carotenoid and other terpenoid degradation products in the surface ocean are likely to occupy a range of radiocarbon signatures, from depleted to modern. Fresh carotenoid sources in surface waters would supply modern degradation products that mix with recycled, long-lived derivatives to yield an average radiocarbon signature for the terpenoid fraction that is not as depleted as deep ocean DOM. This is the first study to link the terpenoid structural components previously attributed to refractory DOM (8, 9) to a depleted radiocarbon signature. However, definitive values for the terpenoid fraction remain unconstrained because CDP has not been completely purified from PPL-DOM.

Concluding remarks

This study identified carotenoids as widespread biosynthetic precursors to a significant component of refractory DOM. Carotenoids are a specific

class of terpenoids and play numerous roles in marine and terrestrial environments (14), which is consistent with the similar chemical features identified in DOM from both types of aquatic environments (8, 9). Assignment of chemical structures in DOM to CDP was supported by a photooxidation experiment using a single carotenoid, BC. At a minimum, 10% of PPL-carbon, or 4% of total DOC, is derived from CDP. The analytical challenges posed by the complexity of DOM also make the molecular-level classification of 4% of total DOC (~28 Gt C) as a single class of compounds a significant achievement. On the basis of the independent evidence offered by GC/MS, 2D NMR, and the BC experiment, CDP provides a distinct (albeit mechanistically unclear) pathway to refractory DOM formation through a single class of bio-synthetic precursors. Although the evidence supporting CRAM/MDLT is robust, to the authors' knowledge, this is the first experiment to dem-onstrate direct formation of terpenoid compounds in DOM from any biosynthetic precursor. Hence, CDP are identified here as the first molecular-level evidence of terpenoids accumulating in refractory DOM.

Still, CDP is very much a nascent model for DOM formation and cannot account for all of the complexity observed in PPL-DOM. This was apparent when the overall ^{1}H NMR spectra of CDP and PPL-DOM were compared (Fig. 2, B and C, projections). The observed differences could be a result of various factors, including the relatively constrained experimental conditions that were used in this study. The experiment was limited to relatively high concentrations of only a single carotenoid, it was conducted for only 40 days compared with the millennial time scale of refractory DOM accumulation, and CDP was produced in a closed system targeting abiotic transformation pathways. For all of these reasons (and others), the diversity of CDP structures in PPL-DOM must be greater than what was observed for experimental CDP.

The remarkable diversity of compounds in CDP provides an exciting area for future research. These CDP are able to explain and partially reproduce the notable complexity in the chemical environment surround-ing prominent branched methyl groups apparent from 2D NMR spectra of DOM (8–11). The long, conjugated, side chain of common and abundant carotenoids, when cleaved at different positions, can plau-sibly give rise to the numerous, chemically related compounds between C_{10} and C_{25}, detected globally in aquatic DOM by various techniques (Fig. 4) (8–13, 16–18). How quickly this isomeric diversity is achieved during the degradation of carotenoids is not known and may be determined by whether abiotic (23) or biotic (21) pathways predomi-nate. The laboratory experiment conducted in this study demonstrated that such diversity could be produced in as little as 40 days. The abun-dance of carotenoids in nature also strongly supports the importance of CDP as a precursor for refractory DOM. An approximate, steady-state calculation based on the estimated carotenoid content of phytoplankton cells (~0.1 to 3% of cell carbon; 24–27) and an estimated primary production rate of 50 Gt C year^{-1} (28), and assuming that carotenoids are produced proportionally to their abundance in cells but that only 6% is converted to water-soluble CDP (this study), provides a CDP residence time of between 550 and 3300 years in the ocean. The range in possible residence times is constrained by the fraction of the DOC reservoir (700 Gt C) that is derived from CDP. These residence times are shorter than the average radiocarbon age of DOC in the deep North Pacific Ocean (~6000 years) (3), but the calculation is only poorly constrained and primarily used here to demonstrate that the carotenoid flux is sufficient to support the estimated size of the dissolved CDP reservoir. Without better radiocarbon data, it is not yet possible to determine the role carotenoids play in the very long-term cycling of DOC. Uncertainties surrounding this calculation demonstrate why the results
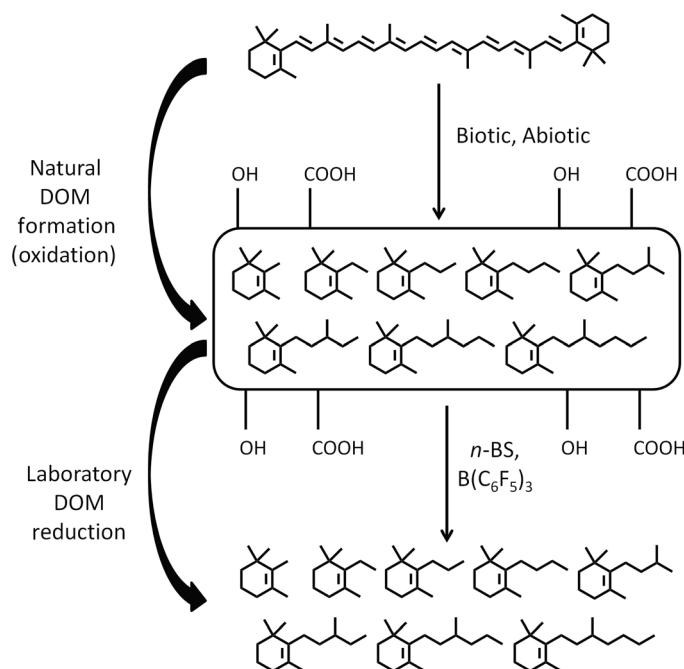
Fig. 4. Proposed transformation of carotenoids to CDP through oxidation. Biotic (bacterial oxidation/reduction) and abiotic (for example, photochemistry) processes are depicted as leading mechanisms of post-biosynthesis modification, resulting in smaller (carbon chain) compounds, such as those identified in Fig. 1. Oxygen-containing functional groups must be abundant in native CDP. Although OH and COOH are expected to be the most prominent groups, their positions along the carbon backbone are unknown. Not shown are heterocyclic compounds that may also be generated, but ring opening during reduction would produce compounds similar to those shown here. Laboratory chemical reduction of PPL-DOM produced alicyclic compounds that resembled those produced in the reduction of experimental CDP. These reduction products were used to link to the proposed biosynthetic precursor.

of this study are so valuable: By assigning carotenoids as a specific bio-synthetic precursor for some of the DOM that accumulates in the ocean, the study enables testable hypotheses to be designed to identify impor-tant processes that shape refractory DOM and control its removal from the ocean (6, 29–31).

MATERIALS AND METHODS
Materials
All solvents (American Chemical Society grade or better), model compounds, and hydrocarbon standards were purchased commercially and used as supplied. n-BS and B(C₆F₅)₃ were purchased from Sigma-Aldrich and stored in an argon-filled desiccator when not in use. Dichlor-omethane (DCM; CH₂Cl₂) was stored over activated molecular sieves (3 Å, 10% w/v).

DOM isolation
The isolation protocol was adapted from Dittmar et al. (5). Seawater (330 liters) from the SIO Pier was collected and prefiltered through an AcroPak 0.8/0.2 μm polyethersulfone (PES) Supor membrane filter. The filtrate was acidified to pH 2 with concentrated hydrochloric acid (HCl) and passed over activated (one cartridge volume of MeOH) 1 g Agilent Bond Elut PPL styrene-divinylbenzene polymer cartridges. The flow rate for the extraction was maintained at approximately

15 ml/min. Cartridges were eluted after 20 liters of seawater had passed through them. Before elution, each cartridge was rinsed with two cartridge volumes of 0.01 M HCl followed by two cartridge volumes of Milli-Q ultrapure water. The water rinse was done to limit methylation of DOM under MeOH elution conditions. Methylation was visible in NMR spectra when the water rinse was omitted. The cartridges were then dried under N_2 gas, and DOM bound to the cartridge (PPL-DOM) was eluted with two cartridge volumes of MeOH. The combined eluents were dried under N_2 gas, resuspended in Milli-Q, frozen, and lyophilized. The total weight of recovered DOM was 235 mg, and elemental analysis showed that this fraction was composed of 48% C. Table S1 reports on some of its other element and isotope properties.

Oxidation of BC to generate CDP

BC was purchased from Sigma-Aldrich and stored at −20°C. For oxidations, 20 to 60 mg of BC were placed into combusted quartz tubes and prefiltered (AcroPak 0.8/0.2 µm PES Supor membrane filter) seawater from the SIO Pier was added to the tubes. "Dark" replicates were covered with opaque plastic bags. Tubes were placed in a seawater bath on the SIO Pier and left exposed to natural sunlight cycles for 40 days. At the end of the experiment, contents of each tube were emptied into a separation funnel, and tubes were subsequently rinsed with DCM to remove residual BC. More DCM was added to the funnel to extract unreacted BC from the aqueous layer (ratio, 1:1). The aqueous layer was extracted again twice with DCM, and organic layers were combined as "residual starting material." An emulsion layer formed during the extraction and was also collected. The aqueous layer was bubbled with N_2 gas under mild heat (50°C) for 30 min to remove any residual DCM. Once free of any organic solvent, the aqueous layer was acidified to pH 2 and CDP were extracted onto PPL resin following the DOM isolation protocol. From 200 mg of BC in 900 ml of filtered seawater, 20 mg of CDP was recovered. On a carbon basis, this represented 6% total recovery of BC carbon as water-soluble and PPL-extractable CDP. The dark incubation resulted in 2.7% recovery at 40 days. The CDP isolated by PPL contained 58.26% carbon by elemental analysis, which contrasted with the 90% carbon content of BC.

DOM fractionation/purification

PPL-DOM was hydrolyzed with 2 M HCl at 90°C overnight to further purify the terpenoid fraction. The dried hydrolysate was redissolved in acidic water (0.01 M HCl) and extracted against ethyl acetate to remove highly lipophilic material. The extraction was repeated twice with 0.01 M HCl, and this process removed approximately 30% of PPL-DOM carbon. The acidic water layers containing the remaining 70% of PPL-carbon were combined and subsequently freeze-dried. The dry, polar PPL-DOM hydrolysate was resuspended in 0.02 M HCl and extracted again using a 1 g PPL cartridge. The permeate and successive acidic and neutral Milli-Q column rinses were collected and freeze-dried. This fraction comprised 12% of the total PPL-carbon and contained primarily small polar compounds, such as sugars and amino acids, as determined by both NMR (fig. S7) and GC-MS (not presented). Finally, the PPL cartridge was dried and eluted with MeOH; this eluent was identified as the terpenoid-rich fraction in the text, table (table S1), and figures (for example, Figs. 2 and 3 and fig. S8) and comprised 58% of the total PPL-carbon. Comparing NMR spectra of terpenoid-rich DOM and bulk PPL-DOM (Fig. 2) showed that these samples were very similar. For example, the methylene resonances (Fig. 2, B and C, red shading) that occur downfield in both carbon and proton axes had been previously attributed to carbohydrates but remained in the terpenoid-

rich sample, indicating that they represented ether functionalities or terminal, aliphatic alcohols. Methine cross peaks centered at a carbon NMR resonance of ~80 ppm previously attributed primarily to carbohydrates also remained in the PPL fraction. Some fraction of these resonances must have been removed in correspondence with the confirmed removal of carbohydrates, but a significant number of these resonances remained. Residual hydrolysis-resistant carbohydrates may have been the source of some of these remaining resonances, but a fraction likely corresponded to other compounds, such a polyhydroxylated lipids (for example, Fig. 3).

Radiocarbon and elemental analysis

Radiocarbon measurements were made on PPL-DOM fractions at the W.M. Keck Carbon Cycle Accelerator Mass Spectrometry Laboratory at the University of California, Irvine (UCI). Sample graphitizing backgrounds have been subtracted based on combusted glycine and glycine. Sample processing backgrounds have been subtracted based on processed glycine and unprocessed glycine. All results have been corrected for isotopic fractionation according to the conventions of Stuiver and Polach (32), with $\delta^{13}C$ values measured on prepared graphite using the Accelerator Mass Spectrometer (AMS). These can differ from $\delta^{13}C$ of the original material, if fractionation occurred during sample graphitization or the AMS measurement, and are not shown. Independent elemental and stable isotopic ($\delta^{13}C$, $\delta^{15}N$) characterization of samples was performed using standard elemental analyzer (EA) isotope ratio MS at SIO (table S1).

Gas chromatography–flame ionization detection/mass spectrometry

An Agilent 7890A Gas Chromatograph system coupled simultaneously to an Agilent 5975C quadrupole mass spectrometer and a flame ionization detector was used for GC-MS/FID characterization of DOM samples. Splitless injection with 1 µl of the analyte was used. Separation was performed on a 5% phenyl poly(dimethylsiloxane) column [J&W 123-5731DB-ht, 30 m, 320 µm inside diameter (ID), 0.1-µm film] with a temperature program from 50°C (hold time, 0.5 min) to 140°C (at 10°C/min; hold time, 0 min) to 320°C (at 15°C/min; hold time, 10 min). Helium was the carrier gas with a constant flow of 1.8 ml/min. After separation, the effluent was split between the FID (operating at 310°C) and the mass spectrometer (70-eV ionization, scanning 50 to 750 m/z). Quantification of reduced model compounds and reduced SIO Pier DOM was accomplished using a calibration curve generated using a hydrocarbon standard (decahydronaphthalene) and assuming a constant FID response factor across unknown hydrocarbons. When available, reduced model compound structures were confirmed by comparison to the National Institute of Standards and Technology (NIST) Mass Spectral Standard Reference Database. For SIO pier DOM, manual integration was used to remove peaks associated with catalyst degradation. Unreduced PPL-DOM and standard compounds were analyzed when necessary by GC-MS following derivatization. Samples were first weighed (~1 mg, in MeOH, dried overnight at 60°C) into combusted 400-µl GC vial inserts on a balance accurate to ±0.05 mg. Labile oxygen residues were then derivatized with trimethylsilyl (TMS) groups by reacting the sample in the insert in 400 µl of N,O-bis(trimethylsilyl)trifluoroacetamide and 10% trimethylsilylchlorosilane in pyridine (3:1) at 70°C for 1 hour.

Comprehensive 2D GC (GC×GC TOF-MS)

A LECO Pegasus 4D GC×GC TOF-MS was used for comprehensive GC×GC analysis. The term GC×GC refers to the use of two distinct

chromatography columns with different chemical selectivity, connected in series. Compounds are separated primarily by volatility in the first column and polarity in the second. The analysis is comprehensive because all of the effluent from the first column is cryofocused and transferred onto the second column. Effluent from the second column is then analyzed by the TOF-MS, which benefits from high spectral acquisition speed (50 to 500 Hz). Finally, the data are compiled into a 2D chromatogram that is visualized and processed by the ChromaTOF software. Both columns were housed within an Agilent 7890A gas chromatograph. The splitless inlet temperature was set at 300°C. The first-dimension column was a semipolar Crossbond diphenyl dimethyl polysiloxane column (Restek Rxi-17 Sil) (30 m length; 0.2550 mm ID; 0.25 μm film thickness). The column was programmed to remain isothermal at 40°C for 1 min and ramped to 315°C at 3°C/min. The modulator temperature was offset by +30°C to the primary oven. The secondary oven (within the GC) housed the second-dimension nonpolar Crossbond diphenyl dimethyl polysiloxane column (Restek Rxi-1) (1.58 m length; 0.250 mm ID; 0.25 μm film thickness). The secondary oven temperature was offset by +25°C to the primary oven. The modulation period was 2.5 s, with a hot pulse time of 1.05 s and a cool time of 0.2 s. The carrier gas was helium at a constant flow of 1.5 ml/min. The acquisition delay on the TOF-MS was set to 160 s. The acquisition rate was set to 50 Hz, with a range of 5 to 1000 m/z. Electron ionization was run at 70 eV. Again, reduced model compound structures were confirmed by comparison to the NIST Mass Spectral Standard Reference Database.

Reduction procedure

The reduction procedure directly followed Arakawa and Aluwihare (12), which was modeled after Nimmagadda and McRae (15). For model compound reductions, substrates (~2.5 mg) were transferred by syringe to a 2-ml single flame-dried vial equipped with a stir bar. Under an argon atmosphere, 10 mg of the catalyst $B(C_6F_5)_3$ was added [100 μl of 100 mg $B(C_6F_5)_3$/1 ml DCM solution]. Immediately after the catalyst was added, the reductant (100 μl of n-BS) was added, also under an argon atmosphere. The mixture was then stirred overnight. Note that samples went from being insoluble in DCM before the reduction to being completely soluble in this solvent after the reduction. Following reduction, each sample was treated to remove excess catalyst and siloxane by-products. Approximately 50 μl of the reaction mixture was transferred into 250 μl of KOH/MeOH. After 1 min, the mixture was extracted with pentane (3× 100 μl). The pentane fractions were collected, washed with H_2O, and dried over Na_2SO_4. The organic fraction was then dried under N_2 to 100 μl for GC-FID/MS and GC×GC TOF-MS analysis. This complete method was reproduced during the reduction of SIO Pier PPL-DOM with 5 mg of starting material. See Arakawa and Aluwihare (12) and scheme S1 for additional discussion of the reduction method.

NMR analysis

At SIO, ^1H NMR spectra were determined on a 500-MHz Varian NMR spectrometer. Dry PPL-DOM or hydrolysate fractions (1 to 4 mg) were dissolved in 0.5 ml of deuterated MeOH (CD_3OD) or deuterated water (D_2O) and typically acquired with 128 scans. NMR spectra were referenced to 3.31 ppm (MeOD) or 4.79 ppm (D_2O).

At the University of Toronto, samples (50 mg) were resuspended in 1 ml of deuterated MeOH (CD_3OD) and analyzed using a Bruker Avance III HD 500 MHz NMR spectrometer equipped with a ^1H-^{13}C-^{15}N 5-mm, triple-resonance inverse (TXI) cryoprobe. HSQC spectra were collected in phase-sensitive mode using echo/anti-echo gradient

selection, sensitivity enhancement, and multiplicity editing during the selection step. All 180° carbon pulses were performed using matched adiabatic pulses for inversion and refocusing. Scans (512) were collected for each of the 128 increments in the F1 dimension. Data points (1024) were collected in F2, using a $^1J^1H$-^{13}C value of 145 Hz. The F2 dimension was multiplied by an exponential function corresponding to a 15-Hz line broadening, whereas the F1 dimension was processed using a sine-squared function with a π/2 phase shift and a zero-filling factor of 2.

HMBC spectra were carried out in phase-sensitive mode using echo/anti-echo gradient selection (33) and a relaxation optimized delay of 25 ms for the evolution of long-range couplings (9, 34). Data points (2048) were collected in F2 over 1024 scans for each of the 128 increments in the F1 dimension. The F2 dimension was multiplied by an exponential function corresponding to a 15-Hz line broadening, whereas the F1 dimension was processed using a sine-squared function with a π/2 phase shift and a zero-filling factor of 2.

DEPTQ spectra, which provide data on all carbon types, including quaternary carbons, were acquired using 32,768 scans, 65,536 time domain points, and a recycle delay of 3 s, with adiabatic pulses for refocusing on the carbon channel of a 5-mm (TXI) cryoprobe. Spectra were processed using an exponential function corresponding to a line broadening of 50 Hz in a transformed spectrum and a zero-filling factor of 2.

Spectral predictions were carried out using Advanced Chemistry Development's ACD/NMR Workbook using Neural Network Prediction algorithms (version 2015.2.5). Parameters used for prediction, including spectral resolution and base frequency, were chosen to match those of the real data sets as closely as possible. Because of the relatively fast relaxation in the DOM sample, the optimal evolution delay was determined to be 25 ms in HMBC experiments. Longer delays theoretically allow longer-range couplings to evolve; however, in DOM, fast relaxation leads to the loss of signal if longer evolution delays are used. Hence, 25 ms represents the optimal compromise for the sample that provides the best balance between signal and correlations that can be observed (9). However, a 25-ms delay corresponds to a $1/2J_{lr}$ ^1H-^{13}C of 20 Hz, which will bias the stronger and shorter-range correlations in the sample; very long range couplings such as four- and five-bond correlations, which take a long time to evolve, will not be detected in DOM. To account for this and permit the simulations to reflect the real situation as closely as possible, only ^1H-^{13}C couplings greater than 6 Hz were included in calculations. Both two- and three-bond correlations were permitted, but the two-bond correlations that generally build up the quickest (that is, will be detected preferentially in DOM because of the short evolution delay that had to be used) were weighted 2:1 over the three-bond correlations. The result is that the real HMBC data for DOM will bias strong and short correlations, and this is matched, to the best possible degree, by the simulations where strong (>6 Hz) and short-range (preference of two bonds over three bonds) couplings are also given preference.

Calculating CDP residence time in the ocean

An approximate calculation was performed to examine whether marine carotenoids could plausibly be the primary source of SPE refractory DOM. Although numerous papers report on the carotenoid content of phytoplankton cells, it is not easy to scale these data to the carbon content of cells. On the basis of a few studies, it was determined that carotenoids could represent somewhere from 3% (*Prochlorococcus*) (24, 25) to 0.1% (mixed diatom assemblage) (26, 27) of the carbon in

phytoplankton cells. However, because pigment concentrations have been shown to vary with irradiation level (35), for example, these values must be viewed as crude estimates. Using a value of 50 Gt C year^{-1} for marine primary production (28), and assuming that carotenoids are produced proportionally to their contribution to cell carbon, an average carotenoid production rate of 0.78 Gt C year^{-1} was calculated. The photo-oxidation experiment found that 6% of the starting BC carbon was extractable as water-soluble CDP. Using this conversion factor, the flux of CDP carbon into the DOC reservoir was estimated at 0.05 Gt C year^{-1}. Using this flux and the conservative estimate of 4% DOC (4% of 700 Gt C) as CDP, the calculated residence time of CDP in the ocean was ~550 years. Such a short residence time does not enable these compounds to reach the deep North Pacific Ocean via advection. Using an estimate of 12% DOC as CDP, the residence time increased to ~1600 years, and if the NMR-based estimate of 58% PPL-DOC, and thus 24% DOC, was used, the residence time increased to ~3300 years.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/9/e1602976/DC1

Supplementary Text

fig. S1. ^{1}H NMR spectra of five PPL-DOM time series samples collected at the SIO Pier.

fig. S2. Detection of compounds A and B in DOM (DOM$_A$ and DOM$_B$) with structures identical to Reduced$_A$ and Reduced$_B$.

fig. S3. ^{1}H NMR spectra of PPL-extracted, water-soluble CDP produced from the photooxidation of BC in filtered seawater (blue) or Milli-Q (red).

fig. S4. Known BC degradation products identified in CDP via GC-MS.

fig. S5. Similar reduction products are observed in environmental PPL-DOM and PPL-extracted CDP.

fig. S6. Comparison of bulk PPL-DOM to terpenoid-rich PPL-DOM by 2D NMR.

fig. S7. ^{1}H NMR spectra of PPL-DOM (red), terpenoid-rich PPL-DOM (green; retained on the second PPL column following acid hydrolysis; see text), and the sugar- and amino acid–rich fraction (blue; not retained by the second PPL column following acid hydrolysis).

fig. S8. Simulated HMBC NMR correlations for oxidized carotenoids.

table S1. Isotope and elemental data for PPL-DOM samples collected from the SIO Pier (32.87° N, 117.26°W).

scheme S1. A single-step reduction of diverse oxygen-containing functional groups to their corresponding alkane.

REFERENCES AND NOTES

1. U. Siegenthaler, J. L. Sarmiento, Atmospheric carbon dioxide and the ocean. *Nature* **365**, 119–125 (1993).

2. P. F. Sexton, R. D. Norris, P. A. Wilson, H. Pälike, T. Westerhold, U. Röhl, C. T. Bolton, S. Gibbs, Eocene global warming events driven by ventilation of oceanic dissolved organic carbon. *Nature* **471**, 349–352 (2011).

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

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

5. T. Dittmar, B. Koch, N. Hertkorn, G. Kattner, A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* **6**, 230–235 (2008).

6. R. Flerus, O. J. Lechtenfeld, B. P. Koch, S. L. McCallister, P. Schmitt-Kopplin, R. Benner, K. Kaiser, G. Kattner, A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* **9**, 1935–1955 (2012).

7. O. J. Lechtenfeld, G. Kattner, R. Flerus, S. L. McCallister, P. Schmitt-Kopplin, B. P. Koch, Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* **126**, 321–337 (2014).

8. N. Hertkorn, R. Benner, M. Frommberger, P. Schmitt-Kopplin, M. Witt, K. Kaiser, A. Kettrup, J. I. Hedges, Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **70**, 2990–3010 (2006).

9. B. Lam, A. Baer, M. Alaee, B. Lefebvre, A. Moser, A. Williams, A. J. Simpson, Major structural components in freshwater dissolved organic matter. *Environ. Sci. Technol.* **41**, 8240–8247 (2007).

10. G. C. Woods, M. J. Simpson, A. J. Simpson, Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy. *Water Res.* **46**, 3398–3408 (2012).

11. N. Hertkorn, M. Harir, B. P. Koch, B. Michalke, P. Schmitt-Kopplin, High-field NMR spectroscopy and FTICR mass spectrometry: Powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **10**, 1583–1624 (2013).

12. N. Arakawa, L. Aluwihare, Direct identification of diverse alicyclic terpenoids in Suwannee River fulvic acid. *Environ. Sci. Technol.* **49**, 4097–4105 (2015).

13. J. A. Leenheer, M. A. Nanny, C. McIntyre, Terpenoids as major precursors of dissolved organic matter in landfill leachates, surface water, and groundwater. *Environ. Sci. Technol.* **37**, 2323–2331 (2003).

14. S. Liaaen-Jensen, Marine carotenoids: Recent progress. *Pure Appl. Chem.* **63**, 1–12 (1991).

15. R. D. Nimmagadda, C. McRae, Characterisation of the backbone structures of several fulvic acids using a novel selective chemical reduction method. *Org. Geochem.* **38**, 1061–1072 (2007).

16. A. C. Stenson, A. G. Marshall, W. T. Cooper, Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **75**, 1275–1284 (2003).

17. M. Witt, J. Fuchser, B. P. Koch, Fragmentation studies of fulvic acids using collision induced dissociation Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **81**, 2688–2694 (2009).

18. T. Dittmar, Reasons behind the long-term stability of dissolved organic matter, in *Biogeochemistry of Marine Dissolved Organic Matter (Second Edition)*, D. A. Hansell, C. A. Carlson, Eds. (Academic Press, 2015), pp. 369–388.

19. G. Britton, Structure and properties of carotenoids in relation to function. *FASEB J.* **9**, 1551–1558 (1995).

20. T. Reemtsma, A. These, J. Linscheid, A. Leenheer, A. Spitzy, Molecular and structural characterization of dissolved organic matter from the deep ocean by FTICR-MS, including hydrophilic nitrogenous organic molecules. *Environ. Sci. Technol.* **42**, 1430–1437 (2008).

21. S. H. Schwartz, B. C. Tan, D. A. Gage, J. A. D. Zeevaart, D. R. McCarty, Specific oxidative cleavage of carotenoids by VP14 of maize. *Science* **276**, 1872–1874 (1997).

22. N. Kitaoka, X. Lu, B. Yang, R. J. Peters, The application of synthetic biology to elucidation of plant mono-, sesqui- and diterpenoid metabolism. *Mol. Plant* **8**, 6–16 (2015).

23. M. Gonsior, N. Hertkorn, M. H. Conte, W. J. Cooper, D. Bastviken, E. Druffel, P. Schmitt-Kopplin, Photochemical production of polyols arising from significant photo-transformation of dissolved organic matter in the oligotrophic surface ocean. *Mar. Chem.* **163**, 10–18 (2014).

24. S. Bertilsson, O. Berglund, D. M. Karl, S. W. Chisholm, Elemental composition of marine *Prochlorococcus* and *Synechococcus*: Implications for the ecological stoichiometry of the sea. *Limnol. Oceanogr.* **48**, 1721–1731 (2003).

25. S. W. Chisholm, S. L. Frankel, R. Goericke, R. J. Olson, B. Palenik, J. B. Waterbury, L. West-Johnsrud, E. R. Zettler, *Prochlorococcus marinus* nov. gen. nov. sp.: An oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. *Arch. Microbiol.* **157**, 297–300 (1992).

26. D. J. S. Montagnes, J. A. Berges, P. J. Harrison, F. J. R. Taylor, Estimating carbon, nitrogen, protein and chlorophyll a from volume in marine phytoplankton. *Limnol. Oceanogr.* **39**, 1044–1060 (1994).

27. F. Pennington, R. R. L. Guillard, S. Liaaen-Jensen, Carotenoid distribution patterns in Bacillariophyceae (diatoms). *Biochem. Syst. Ecol.* **16**, 589–592 (1988).

28. A. Longhurst, S. Sathyendranath, T. Platt, C. Caverhill, An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* **17**, 1245–1271 (1995).

29. N. Jiao, G. J. Herndl, D. A. Hansell, R. Benner, G. Kattner, S. W. Wilhelm, D. L. Kirchman, M. G. Weinbauer, T. Luo, F. Chen, F. Azam, Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* **8**, 593–599 (2010).

30. S. Q. Lang, D. A. Butterfield, M. D. Lilley, H. P. Johnson, J. I. Hedges, Dissolved organic carbon in ridge-axis and ridge-flank hydrothermal systems. *Geochim. Cosmochim. Acta* **70**, 3830–3842 (2006).

31. J. A. Hawkes, P. E. Rossel, A. Stubbins, D. Butterfield, D. P. Connelly, E. P. Achterberg, A. Koschinsky, V. Chavagnac, C. T. Hansen, W. Bach, T. Dittmar, Efficient removal of recalcitrant deep-ocean dissolved organic matter during hydrothermal circulation. *Nat. Geosci.* **8**, 856–860 (2015).

32. M. Stuiver, H. A. Polach, Discussion: Reporting of ^{14}C data. *Radiocarbon* **19**, 355–363 (1977).

33. D. O. Cicero, G. Barbato, R. Bazzo, Sensitivity enhancement of a two-dimensional experiment for the measurement of heteronuclear long-range coupling constants, by a new scheme of coherence selection by gradients. *J. Magn. Reson.* **148**, 209–213 (2001).

34. A. J. Simpson, D. J. McNally, M. J. Simpson, NMR spectroscopy in environmental research: From molecular interactions to global processes. *Prog. Nucl. Magn. Reson. Spectrosc.* **58**, 97–175 (2011).
35. Z. Dubinsky, P. G. Falkowski, K. Wyman, Light harvesting and utilization by phytoplankton. *Plant Cell Physiol.* **27**, 1335–1349 (1986).

Citation: N. Arakawa, L. I. Aluwihare, A. J. Simpson, R. Soong, B. M. Stephens, D. Lane-Coplen, Carotenoids are the likely precursor of a significant fraction of marine dissolved organic matter. *Sci. Adv.* **3**, e1602976 (2017).