Resurrected Rubisco suggests uniform carbon isotope signatures over geologic time

Mateusz Kedzior[1], Amanda K. Garcia[2], Meng Li[3], Arnaud Taton[4], Zachary R. Adam[5], Jodi N. Young[3] and Betul Kacar[1,2*]

[1]Department of Molecular and Cellular Biology, University of Arizona, Tucson, AZ 85721, USA
[2]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA
[3]School of Oceanography, University of Washington, Seattle, WA 98195, USA
[4]Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093 USA
[5]Department of Geosciences, University of Wisconsin-Madison, Madison, WI 53706 USA

*Corresponding author: bkacar@wisc.edu

Lead contact: Betül Kaçar

Category: Paleobiology, Evolutionary Biology

Keywords: RuBisCO, carbon fixation, carbon isotope record, molecular paleobiology, paleogenetics, ancestral sequence reconstruction

Summary

The earliest geochemical indicators of microbes—and the enzymes that powered them—extend back almost 3.8 billion years on our planet. Paleobiologists often attempt to understand these indicators by assuming that the behaviors of modern microbes and enzymes are consistent (uniform) with those of their predecessors. A uniformitarian assumption (i.e., the idea that fundamental geobiological processes have occurred in much the same manner over Earth history) seems at odds with our understanding of the inherent variability of living systems. Here, we examine whether a uniformitarian assumption for an enzyme thought to generate carbon isotope indicators of biological activity, RuBisCO, can be corroborated by independently studying the history of changes recorded within RuBisCO's genetic sequences. Specifically, we resurrected a Precambrian-age, ancient RuBisCO by engineering its ancient DNA inside a modern cyanobacterium genome and measured the engineered organism's fitness and carbon-isotope-discrimination profile. The envelope of ancestral RuBisCO isotopic fractionation observed here indicates that uniformitarian assumptions may be warranted, but with important caveats. Our results suggest that further inquiries that link molecule-level evolutionary changes with planet-level geochemical conditions are needed to discern whether enzyme-affected isotope fractionation trends extend deeper into the early Precambrian. Experimental studies illuminating life's early molecular innovations are crucial to explore the foundations of Precambrian uniformitarian assumptions.

Introduction

The history of life on Earth may be broadly subdivided into two, mutually exclusive macroevolutionary phases, the Phanerozoic and the Precambrian (Butterfield, 2007). The Phanerozoic (~542 million years ago to present) may be characterized by physiological and anatomical innovations and their resultant effects on ecosystem expansion, trophic tiering, and sociality (Bambach, 1986; Bottjer and Ausich, 2016; Hannisdal and Peters, 2011). Hard- and soft-anatomical preservation provides a rich template for reconstructing Phanerozoic adaptive trends, correlating them with geographical and climatological changes (Sepkoski, 2015; Vermeij, 1973), and for testing observed diversity trends against possible systematic effects of preservation bias (Hendy, 2010; Peters, 2005; Peters and Foote, 2016). By contrast, the Precambrian (the ~4 billion years preceding the Phanerozoic) is primarily characterized by genetic and metabolic biomolecular innovations, traded amongst microscopic organisms of uncertain phylogenetic assignment (Knoll, 2015; Schopf, 1994). The Precambrian record of evolutionary change appears to be cryptic and may have been comparatively static. This may be attributable to macroevolutionary dynamics that were distinctly non-Phanerozoic, or it may merely indicate a lack of direct paleontological and geological evidence of the specific timing and extent of intermediate biomolecular adaptive steps (Butterfield, 2007).

Comparative analyses of extant organisms have traditionally been the most informative means of interpreting the scant direct evidence of Precambrian life, but such analyses inevitably face pitfalls. A reasonable null hypothesis is that evolution is largely a uniformitarian process, such that rates or tempos may change but the primary modes of evolutionary change (as established through observations of the more comprehensive Phanerozoic record) are stable across geologic timescales (Erwin, 2011). Similarly, observations regarding present-day or recent geobiological and molecular processes might thus be relevant for interpretation of paleobiological signatures from the Precambrian. Such uniformitarian assumptions of ancient biology inferred from extant or Phanerozoic phenotypes are often employed to make sense of the Precambrian record, including body plan function (Adam et al., 2017; Yin et al., 2016), cladistic assignment (Butterfield, 2000; Golubic and Seong-Joo, 1999; Schopf, 2011; Tang et al., 2020), and isotope biosignature traces (Anbar and Rouxel, 2007; Beard, 1999; Garcia et al., 2021a; Schidlowski, 2001; Schopf et al., 2015). However, more recent studies question these assumptions, namely indicating that modes of biological variation can actually vary through time (Erwin, 2011). A major crux of the problem is that even the simplest modern organisms, as well as the macromolecules that compose them, differ from their ancient predecessors in having been shaped by the cumulative effects of billions of years of Earth-life co-evolution and ecosystem upheaval. For this reason, Precambrian functional or sequence comparisons may be of limited utility at organismal or biomolecular levels of adaptation, undercutting interpretations made possible by uniformitarian view. Novel experimental approaches may help to distinguish inferred paleobiological phenotypes from characteristically modern adaptive overprints.

The interpretation of the Precambrian carbon isotope record, comprising the oldest signatures of life on Earth, may be aided by novel experimental constraints on ancient phenotypes. Interpretations of this record is conventionally subject to uniformitarian assumptions regarding ancient biogeochemistry (Krissansen-Totton et al., 2015; Schidlowski, 2001). The $^{13}C/^{12}C$ isotopic mean of preserved organic carbon ($\delta^{13}C_{org} \approx -25‰$) has remained notably static over geologic time (Garcia et al., 2021a; Krissansen-Totton

et al., 2015; Schidlowski, 2001), and is leveraged as a general signature of ancient biological activity (Bell et al., 2015; Schidlowski, 1988; Schopf et al., 2018). Given its role in the Calvin-Benson-Bassham cycle—likely the predominant mode of carbon fixation for much of Earth history (Schidlowski, 2001)—the majority of contextual information used to assess how carbon isotope biosignatures might have been generated over Earth's deep history comes from studies of the modern enzyme RuBisCO (Ribulose 1,5-Bisphosphate (RuBP) Carboxylase/Oxygenase, EC 4.1.1.39). RuBisCO catalyzes the uptake of inorganic CO_2 from the environment and facilitates CO_2 reduction and incorporation into organic biomass. The $^{13}C/^{12}C$ isotopic fractionation of modern Form I RuBisCO variants in photosynthetic organisms consistently measures ~-25‰ (Guy et al., 1993; Roeske and O'Leary, 1984; Scott et al., 2007; von Caemmerer et al., 2014), approximately the same isotopic difference observed between inorganic and organic carbon in the Precambrian rock record. The carbon isotope discrimination of RuBisCO has therefore been presumed to have remained constant over the history of life.

Recent data, however, demonstrate that RuBisCO can produce significantly different carbon isotope signatures within organic matter in response to external factors, such as levels of atmospheric CO_2 and/or O_2 (Eichner et al., 2015; Freeman and Hayes, 1992; Hurley et al., 2021; Wilkes et al., 2018) or cellular carbon concentrating mechanisms, which affect the catalytic efficiency of RuBisCO (Flamholz et al., 2019; Studer et al., 2014; Tcherkez et al., 2006). Internal factors may also affect fractionation, such as single-point mutations that can alter the interaction between RuBisCO and CO_2 (McNevin et al., 2007). The sensitivity of RuBisCO activity and fractionation to such factors is particularly significant given evidence for broad-scale trends in environmental conditions and biological innovations over the enzyme's evolutionary history. These include the progressive oxygenation of the Earth atmosphere (Lyons et al., 2014), the evolution of novel carbon concentrating mechanisms (Badger, 2003), and the diversification of RuBisCO-hosting taxa. Thus, it seems unlikely that ancestral forms under ancient environmental conditions generated the same isotope fractionation signal as descendent homologs in modern organisms. Though some variability might be expressed by individual carbon isotopic measurements in the geologic record within a broader range, one might expect more frequent and significant excursions in the ~-25‰ isotopic mean given the enormity of geobiological developments over the past >3 billion years (Garcia et al., 2021a). An experimental assessment that combines the phylogenetic history of RuBisCO with the study of intra- and extracellular conditions may provide a more insightful basis for comparing extant and Precambrian carbon isotope fractionation patterns.

The Form I clade of the RuBisCO phylogeny (and its macroevolutionary tractability afforded through green plant, algal and cyanobacterial fossil lineages) makes it an exemplary paleomolecular system for assessing uniformitarian assumptions applied to Precambrian biosignatures. Here, we establish an experimental system for the reconstruction of ancestral biomolecules with which to interpret evidence of Precambrian biological activity. Specifically, we report the resurrection and genetic incorporation of a phylogenetically reconstructed, ancient Form IB RuBisCO variant in a modern strain of cyanobacteria, *Synechococcus elongatus* PCC 7942 (thereafter *S. elongatus*) (Berla et al., 2013; Chen et al., 2012). We compared expression and activity levels of RuBisCO variants and the resulting changes in the growth of *S. elongatus* and isotope fractionation under ambient air as well as a CO_2 concentration that reflects Precambrian conditions.

Results

Construction of a *S. elongatus* strain harboring ancestral RuBisCO. To experimentally investigate the generation of carbon isotope biosignatures in deep time, we designed a paleomolecular system to engineer computationally inferred, ancestral RuBisCO enzymes in extant cyanobacteria. We previously reconstructed a comprehensive phylogenetic history of RuBisCO and inferred maximum-likelihood ancestral RuBisCO large-subunit (RbcL) protein sequences (Kacar et al., 2017c) (Fig. 1A). For this study, we selected the ancestor of the Form IB RuBisCO clade (cyanobacteria, green algae, and land plants) for laboratory resurrection, designated "AncIB." Chlorophyte and land plant RuBisCO homologs are nested among cyanobacterial sequences within the Form IB clade. The Form IB topology therefore recapitulates a primary plastid endosymbiotic history from cyanobacterial to Chlorophyte ancestors (Maruyama and Kim, 2020; Shih et al., 2016; Yoon et al., 2004) and constrains the minimum age of ancestral Form IB to older than the Archaeplastida. As a conservative estimate, AncIB is thus likely older than ~1 Ga (the age of the oldest well-characterized, crown-group red and green algal fossils (Gibson et al., 2017; Tang et al., 2020)) and younger than maximum age estimates of cyanobacteria (~3 Ga, as constrained by oxidized sediments potentially indicating the early presence of oxygenic photosynthesis (Anbar et al., 2007; Crowe et al., 2013)).

The ancestral AncIB and *S. elongatus* native RbcL proteins differ at 37 sites and share 92% amino acid identity (Fig. 2). This site variation is evenly spread across the length of the protein, except for a highly conserved region between approximately site 170 to 285 (site numbering here and hereafter based on aligned WT *S. elongatus* RbcL; Fig. 2A) that constitutes a portion of the catalytic C-terminal domain and is proximal to the L-L interface and active site. Critical residues for carboxylase activity, including the Lys-198 site that binds CO_2, are conserved in AncIB. Homology modeling of AncIB using the *S. elongatus* RbcL template (PDB: 1RBL (Newman et al., 1993)) indicates high structural conservation without predicted disruption to secondary structure (Fig. 2B). The nucleotide sequence for the reconstructed AncIB RbcL protein was codon-optimized for *S. elongatus* and cloned within a copy of the *rbc* operon into pSyn02 (Garcia et al., 2021b) for insertion into the *S. elongatus* chromosomal neutral site 2 (NS2). The native *rbc* operon was subsequently deleted to create the AncIB strain. In addition, we generated the control strain Syn01 harboring WT *rbcL* at NS2 (see Materials and Methods for full genome strategy). Thus, the AncIB and Syn01 strains carry a single ectopic copy of the *rbc* operon at NS2 and only differ in the coding sequence of the large RuBisCO subunit (Table S1).

Ancestral RuBisCO complements photoautotrophic growth of extant *S. elongatus*. We cultured wild-type (WT) and engineered (AncIB and Syn01) strains of *S. elongatus* in both ambient air and 2% CO_2 to evaluate the physiological impact of ancestral RuBisCO under estimated Precambrian CO_2 concentrations (Catling and Zahnle, 2020) (Fig. 3A). The AncIB strain was capable of photoautotrophic growth in both ambient air and 2% CO_2. Maximum growth rates for all strains were relatively comparable under each atmospheric condition (averaging doubling times of ~15 to 20 hours), and generally increased under 2% CO_2 relative to air ($p < 0.001$; Table S2). The AncIB strain exhibited a significantly diminished maximum cell density ($OD_{750} \approx 5$) relative to both WT and Syn02 strains ($OD_{750} \approx 8$; $p < 0.001$). No significant difference between WT and Syn01 growth rate or maximum cell density was observed under ambient air or 2% CO_2. In addition, a moderate increase in the midpoint time for the AncIB strain was observed under

air, indicating lag in growth ($p < 0.01$). These growth parameters taken together suggest decreased fitness of the AncIB strain relative to WT and Syn01.

AncIB RuBisCO protein produced a more modest impact on the oxygen evolution of *S. elongatus* compared to that on growth parameters. Cell suspensions were briefly incubated in the dark and subsequently exposed to saturated light in an electrode chamber to detect evolution of molecular oxygen (normalized to chlorophyll *a* concentration (Zavřel et al., 2015b)). For cells sampled from cultures grown in air, there was no statistical difference detected between any of the WT, Syn01, or AncIB strains (Fig. 3B). We did find a modest but significant decrease in photosynthetic activity for AncIB relative to WT for cells cultured in 2% CO_2, generating ~200 and ~320 nmol $O_2 \cdot h^{-1} \cdot \mu g^{-1}$ chlorophyll *a*, respectively. However, a slower O_2 evolution rate was also observed for Syn01 at 2% CO_2.

Ancestral RuBisCO is overexpressed and less catalytically active relative to WT. We assessed the impact of ancestral RuBisCO on gene expression at both the transcript and protein levels. *S. elongatus rbcL* transcript was measured by quantitative reverse-transcription PCR (RT-qPCR) and normalized to that of the *secA* reference gene (Szekeres et al., 2014). For strains cultured in ambient air, we found that the AncIB strain produced a ~29-fold increase in *rbcL* transcript relative to WT or the control strain Syn01 ($p < 0.001$; Fig. 4A). The magnitude of AncIB *rbcL* overexpression was lower in 2% CO_2, with a ~4 and ~2-fold increase observed relative to WT and the control strain, respectively ($p < 0.001$).

RbcL protein was quantified for all *S. elongatus* strains by immunodetection using rabbit anti-RbcL antibody. We found that the amount of RbcL protein was also increased in the AncIB strain by ~3-fold ($p < 0.05$) and ~5-fold ($p < 0.001$) relative to WT or Syn01 in air and 2% CO_2, respectively (Fig. 4B). No difference in RbcL quantity was detected between WT and Syn01 strains at either atmospheric condition. Finally, we confirmed assembly of the hexadecameric L_8S_8 RuBisCO complex in the AncIB strain by native PAGE and detection by anti-RbcL antibody (Fig. S1).

The total carboxylase activity of *S. elongatus* harboring ancestral RuBisCO was measured from crude cell lysates. Activity was assessed by an *in vitro* spectrophotometric coupled-enzyme assay that measures NADH oxidation and is reported as the RuBP consumption rate normalized to total soluble protein content (Kubien et al., 2008). For two sets of assays using either 2.5 mM or 5 mM HCO_3^-, AncIB lysate generated less than half the carboxylase activity of WT lysate ($p < 0.05$) (Fig. 4C).

***S. elongatus* harboring ancestral RuBisCO produces greater carbon isotopic fractionation than wild-type.** We measured the carbon isotope discrimination of *S. elongatus* strains cultured in ambient air and 2% CO_2 to evaluate how the fractionation behavior of ancestral RuBisCO might influence the interpretation of ancient isotopic biosignatures preserved in the geologic record. The $^{13}C/^{12}C$ carbon isotope composition of *S. elongatus* was measured for biomass ($\delta^{13}C_{biomass}$) as well as dissolved inorganic carbon (DIC; $\delta^{13}C_{DIC}$) in the growth medium (Table S4). We calculated the carbon isotope fractionation associated with photosynthetic CO_2 fixation (ε_p) following Freeman and Hayes (Freeman and Hayes, 1992) after estimating $\delta^{13}C_{CO2}$ from measured $\delta^{13}C_{DIC}$ (Mook et al., 1974; Rau et al., 1996) (see Materials and Methods).

Overall, ε_p values were greater for *S. elongatus* strains cultured in 2% CO_2 compared to ambient air. At 2% CO_2, ε_p ranged between 21‰ and 26‰ compared to only 8‰ and 14‰ in air. We found that *S. elongatus* engineered with AncIB RbcL had an ε_p ~5‰ greater than both WT and Syn01 when cultured in air ($p < 0.001$) and ~2‰ to 4‰ greater than WT and Syn01 when cultured at 2% CO_2 ($p < 0.01$) (Fig. 1C), though these differences in ε_p appear driven by the inorganic carbon pool composition rather than biomass (Table S4). A substantially smaller increase in ε_p (~1‰; $p < 0.001$) was observed for Syn01 relative to WT under ambient air. Conversely, under 2% CO_2, Syn01 ε_p was decreased by ~2‰ ($p < 0.001$) relative to WT.

Discussion

Form IB ancestral RuBisCO, when engineered into an extant strain of *S. elongatus*, decreased both the organismal growth capacity and the total cell RuBisCO activity. The genetic engineering strategy for insertion of the AncIB *rbcL* sequence in the cyanobacterial genome cannot solely account for these physiological differences since differences between the WT and Syn01 control strains were insignificant for most measured properties, or not comparable in magnitude to differences between the WT and AncIB strains. Rather, the observed differences for the AncIB strain indicate that the resultant phenotype is likely attributable to the functionality of the ancestral enzyme itself. The unique phenotype of the AncIB strain could be a direct result of the ancestral RbcL subunit or due to impediments to the assembly and activation of a hexadecamer RuBisCO complex containing both the ancestral RbcL and modern RbcS subunits, the latter of which was not targeted for ancestral reconstruction in this study. Our purposes here were to clarify the downstream physiological effects of a single ancestral RuBisCO gene, without potentially complicating factors arising from the interaction of two ancestral proteins. However, future studies could additionally reconstruct the Form IB RbcS and compare with the results presented here. Another possibility is hampered integration of ancestral RbcL given a modern suite of associated proteins required for RuBisCO folding and assembly (Gunn et al., 2020). Further, while overexpression of the ancestral RbcL occurred at the level of transcription and translation, the ancestral strain appears to have comparable levels of assembled hexadecamer RuBisCO, suggesting lower rates of RuBisCO assembly (or faster degradation). Even lower rates of measured total carboxylase activity suggest that the AncIB has decreased efficacy, which could be directly representing ancestral RuBisCO kinetics as well as the challenges associated with hybrid enzyme activation and activity.

Overexpression of the amount of ancestral RuBisCO shown by RT-qPCR and immunodetection assays is a common physiological response to decreased enzymatic efficacy throughout the cell (e.g., (Goldsmith and Tawfik, 2009; Kacar et al., 2017a)). However, expression compensation is insufficient to fully restore the extant WT phenotype, as indicated by the reduced fitness (i.e., decreased maximum cell density, oxygen evolution, and total carboxylase activity) of the ancestral strain harboring the ancestral RuBisCO compared to WT and Syn01.

There are few *in vitro* measurements of the kinetic isotope effect of Form IB RuBisCO in modern-day organisms, but those available range from ~22‰ to 28‰ for cyanobacteria and C_3 plants, respectively (Guy et al., 1993; Roeske and O'Leary, 1984; Scott et al., 2007; von Caemmerer et al., 2014). The ~26‰ ε_p of AncIB strain biomass grown under 2% CO_2 suggests that the ancestral RuBisCO also fractionates within this range. It has been theorized that RuBisCO kinetics have adapted in response to CO_2 availability, either due to increased environmental CO_2 (Tcherkez et al., 2006) or the emergence of CCMs (e.g., C_4

258 photosynthesis in plants (Christin et al., 2008)). Considering the positive relationship between enzymatic
259 fractionation and RuBisCO's specificity to CO_2 (Tcherkez et al., 2006), reconstructing ancient RuBisCO
260 kinetic isotope effect could provide insights into the co-evolution of atmospheric concentrations of CO_2
261 and O_2 and carbon fixation strategies during the Precambrian, in particular the emergence of carbon
262 concentrating mechanisms (CCMs) (Burnap et al., 2015; Cameron et al., 2013; Hurley et al., 2021; Long et
263 al., 2021). This is relevant as precise estimates of the magnitude of atmospheric CO_2 elevation during the
264 Precambrian relative to the present, as well as the emergence and effectiveness of Precambrian CCMs,
265 are unknown.
266
267 We did observe statistically significant differences in ε_p of the ancestral strain compared to WT *S.*
268 *elongatus* under both ambient air and 2% CO_2 atmospheric conditions. However, upon inspection it
269 appears that the diminished activity of the AncIB strain is influencing the composition of the DIC pool
270 (both $\delta^{13}C$ and concentration) in our cultures, and it is in fact the differences in DIC composition driving
271 the calculated differences in ε_p. Though strains were harvested at similar cell densities, small differences
272 in cell concentrations at high densities can strongly influence carbonate chemistry of the media (Shi et al.,
273 2009). The lower CO_2 availability in the air treatment is more sensitive to cellular influence, resulting in a
274 larger difference in ε_p compared to the 2% CO_2 treatment. These differences in DIC are unlikely to be due
275 to experimental setup (e.g., variations in CO_2 bubbling), as biological replicates showed similar values.
276 Therefore, the differences in fractionation reported here are likely implicated indirectly with the less
277 efficient AncIB ancestral enzyme. Further comparative biomolecular characterization of AncIB and WT *S.*
278 *elongatus* RbcL forms is needed to determine the degree to which enzymatic inefficiencies are
279 contributing directly to the AncIB strain phenotype.
280
281 The observed carbon isotopic fractionation values corroborate a uniformitarian assumption for applying
282 the maximal range of extant organism-level isotope fractionation values to interpret deep time isotopic
283 biosignatures. The maximal observed isotopic fractionation of the AncIB strain under simulated
284 Precambrian conditions with elevated CO_2 is comparable to the ~-25‰ isotopic mean of preserved,
285 organic carbon across the plausible age range of the AncIB RuBisCO enzyme. These results suggest the
286 conservation of RuBisCO fractionation behavior over the tested span of molecular evolution. There are,
287 however, several potentially important contextual caveats for broader interpretation of these findings.
288 The Form IB ancestor represents predecessors that are at least 1 billion years old, but it is also genetically
289 and functionally still likely to be very different from the putative 'root' or common ancestor of all RuBisCO
290 variants that emerged much earlier. With the observed *in vivo* functionality of AncIB RuBisCO reported
291 here, we expect the successful reconstruction of older RuBisCO variants by similar means to be feasible.
292 Reconstruction of older ancestors may further expand this maximal envelope of RuBisCO-generated
293 carbon fractionation, or it may indicate that the extant maximal envelope is pervasive (and perhaps
294 characteristic) across all functional variants of RuBisCO.
295
296 Another important caveat lies in the observation that, whereas all strains produce increased isotopic
297 fractionation under elevated CO_2, the comparative difference between ancestral AncIB and WT RbcL
298 fractionation is relatively muted under 2% CO_2 relative to ambient air. One possibility is that elevated CO_2
299 brings the intrinsic fractionation properties of RuBisCO into relief (Bidigare et al., 1997; Hayes, 1993;
300 Schubert and Jahren, 2012; Wilkes et al., 2018), at least compared to fractionation effects deriving from
301 the overlying organismal physiology. By contrast, in present-day conditions, RuBisCO-mediated

302 fractionation processes may be more significantly overprinted by physical factors that can affect RuBisCO
303 catalytic efficiency, including cellular diffusion of O_2/CO_2 or other factors such as the presence of carbon
304 concentration mechanisms.
305
306 There are many fundamental attributes of extant and ancestral metabolism for which the systemic effects
307 on biosignature production have yet to be characterized. Disentangling these effects is critical for
308 interpretation of the oldest biogeochemical record. A host cyanobacterium *S. elongatus* engineered with
309 a Form IB RbcL ancestor confirms that organism- and enzyme-level effects on biosignature production are
310 not always synonymous but differ in nuanced ways. These differences are contingent upon changes to
311 internal (cellular, physiological) and external (environmental) conditions that have demonstrably varied
312 over Earth's long history. Cyanobacteria, with well-characterized genetic and morphological features
313 (Berla et al., 2013; Burnap et al., 2015; Cameron et al., 2013; Chen et al., 2012; Long et al., 2021) and a
314 tractable paleobiological history (Kacar et al., 2017b; Schopf, 2011), are ideal hosts for investigating a
315 range of early Precambrian metabolic processes (Garcia and Kacar, 2019; Garcia et al., 2020; Kacar et al.,
316 2017b).
317
318 Discernible trends (or steadfast consistencies) in metabolic outputs over macroevolutionary timescales
319 can lead to foundational uniformitarian approaches to deep time molecular paleobiology. The available
320 rock record becomes vanishingly sparse with greater age, but it is arguably well-sampled across key global-
321 scale and biotically relevant isotopic systems at least through the early Archaean. Greater geologic
322 sampling will therefore likely generate diminishing returns for shedding new light on deep time
323 paleobiological trends. Innovative approaches that can chart a comprehensive envelope of biomolecular
324 variability over time are a promising new means of reconciling coarse geochemical data with the nuance
325 and complexity of ancient biological activity.
326
327 The engineering of ancient-modern hybrid organisms and their characterization can be used to
328 complement the existing array of fossil remains, biogeochemical signatures, and modern organismal and
329 molecular proxies to assess and contextualize plausible ranges of Precambrian carbon isotope
330 biosignature production. Hybrid organisms may be particularly useful to disentangle the regulatory,
331 physiological, and inter- and intramolecular factors that have impacted isotope fractionation, none of
332 which are individually expressed in the geologic record. These factors must be systematically accounted
333 for when interpreting bulk fractionation signals, even if only to elucidate the evolutionary molecular
334 underpinnings of uniformitarian phenomena over geologic time.
335
336 After engineering a cyanobacterium with an ancient RuBisCO large protein subunit and cultivating it under
337 conditions that mimic those prevailing through much of the Precambrian, we found the resultant carbon
338 isotope fractionation to be within the range of organisms utilizing modern Form IB RuBisCO. The
339 underlying biomolecular and organismal adjustments made by the cell to accommodate the ancestral
340 gene were tracked, and we conclude that the small fractionation differences observed are likely
341 attributable indirectly to decreased fitness of the AncIB strain, which influenced the inorganic carbonate
342 chemistry of the media. The consistency of isotopic signatures generated by this strain indicates that
343 uniformitarian assumptions based on the range of phenotypes of modern RuBisCO variants may apply for
344 Precambrian environmental conditions, but that further study is warranted to discern organism- and
345 enzyme-level trends in carbon isotope fractionation that may extend deeper into the early Precambrian.

Limitations of the study

Our investigation characterized the *in vivo* attributes of a reconstructed, ancestral Form IB RuBisCO within a modern cyanobacterial to assess the commonly used, uniformitarian approach of interpreting ancient carbon isotope biosignatures based on comparable signatures generated by modern organisms and enzymes. Our results are supportive of the conservation of isotope fractionation behavior in cyanobacteria-hosted RuBisCO dating to the Precambrian. However, we do not provide an exhaustive assessment of all potential external factors that would together have shaped the isotope composition of preserved carbon. We target one time point in the evolution of RuBisCO using cyanobacteria as a model. Though the activity of RuBisCO enzymes have likely dominated primary production for billions of years, it has not been the only molecular determinant of the carbon isotope record (reviewed in (Garcia et al., 2021a)). A comprehensive exploration of the constancy of the carbon isotope trend across geologic timescales would require many detailed studies across environmental factors, including carbon fixation pathways catalyzed by different enzymes, and taxa that would have fluctuated in their proportional contributions to primary productivity over Earth history (e.g., prokaryotic versus eukaryotic contributions). Furthermore, our experimental system places an inferred, ancestral Form IB RuBisCO within the physiological context of a modern cyanobacterium. Such an approach is currently necessary to provide a tractable, experimental window into organism-level biogeochemical signatures stemming from the behavior of an ancestral enzyme. However, it is possible that the modern organism—including the associated modern proteins that interact with the ancestral RuBisCO protein—could influence the ancestral RuBisCO phenotype. These potential effects should in the future be investigated to provide more nuanced constraints on the range of isotopic biosignatures generated by ancient RuBisCO in the Precambrian. Our work here is a small start in the opening of an entirely new research direction that combines evolution, synthetic and molecular biology to experimentally evaluate the bases for interpreting the oldest record of life on Earth.

Acknowledgements

We sincerely thank Emily Peñaherrera, Ryan Ward, and the University of California-Davis Stable Isotope Facility for assistance, and Jenan Kharbush for helpful discussions. We also thank two anonymous reviewers for their constructive comments and suggestions. This work was supported by the National Aeronautics and Space Administration Early Career Faculty (ECF) Award No. 80NSSC19K1617 (BK) and Interdisciplinary Consortium for Astrobiology Research No. 19-ICAR19_2-0007 (BK), the National Science Foundation Emerging Frontiers Program Award No. 1724090 (BK), National Aeronautics and Space Administration Postdoctoral Fellowship (AKG), Simons Foundation Early Career Award No. 561645 (JNY and ML), and National Institute of General Medical Sciences of the National Institutes of Health Award No. R01GM118815 (AT, to James W. Golden at the University of California-San Diego).

Author contributions

Study design, and supervision, B.K.; methodology, M.K., A.K.G., M.L., A.T., J.Y., and B.K.; formal analysis, M.K. A.K.G. and B.K.; investigation, M.K., A.K.G., M.L., and B.K.; resources, A.T., J.Y., and B.K.; writing –

original draft, M.K., A.K.G., Z.A., and B.K.; writing – review & editing, M.K., A.K.G., M.L., A.T., J.Y., Z.A., and B.K.; visualization, A.K.G. and B.K. All authors read, edited, and approved the final manuscript.

Declaration of interests

The authors declare no competing interests.

Figures

Figure 1.

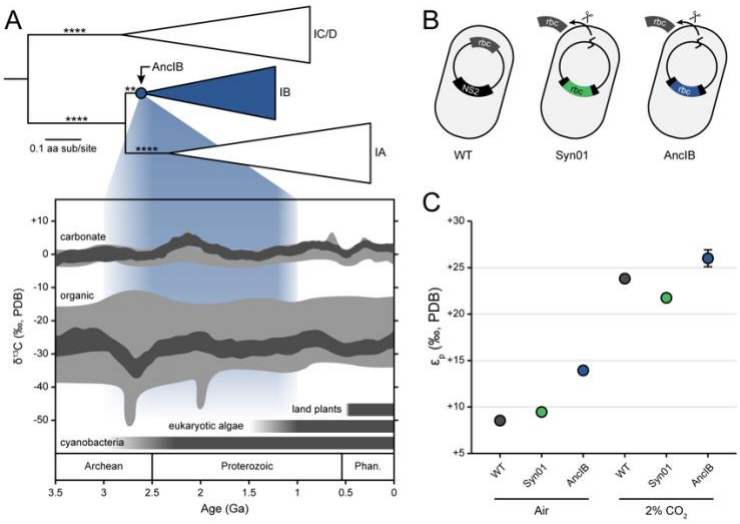
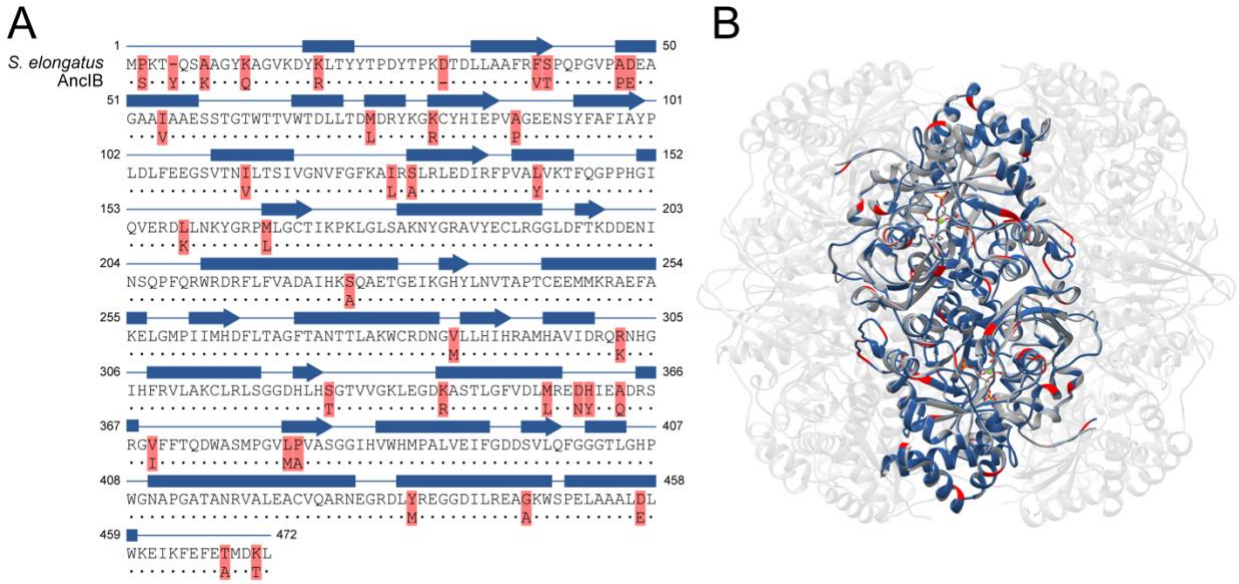
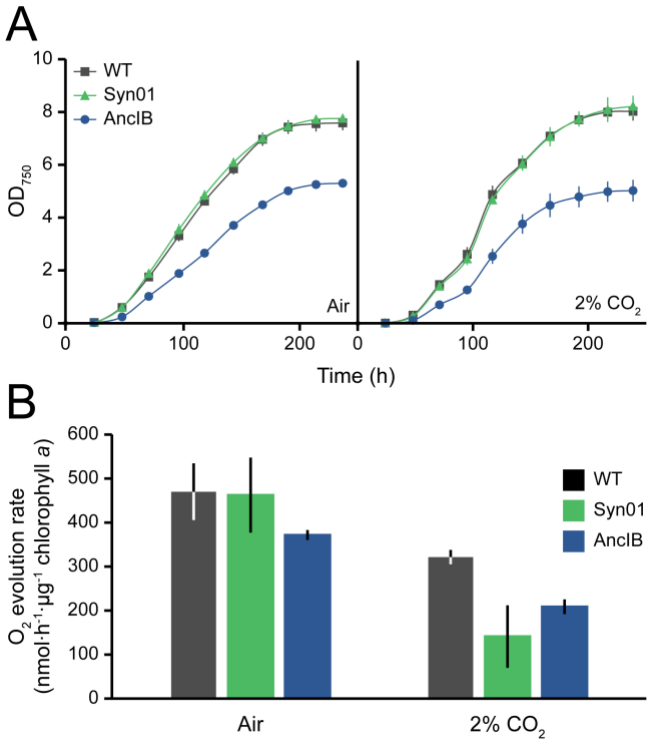


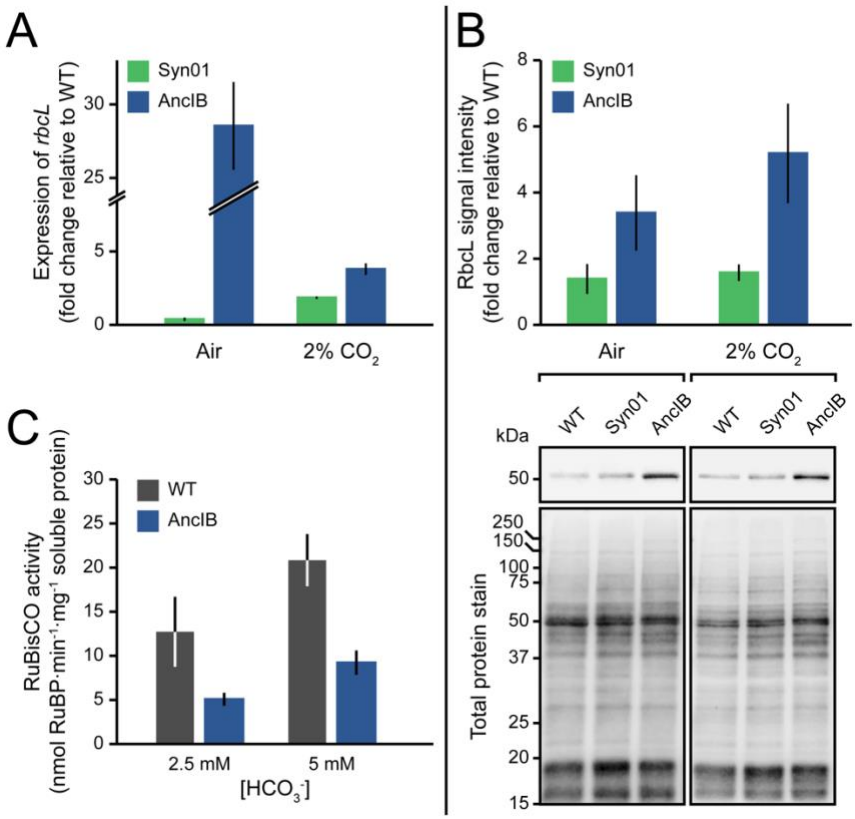
Figure 2.



404 **Figure 3.**



405
406
407 **Figure 4.**



408

Figure titles and legends

Figure 1. Reconstruction of ancestral RuBisCO biogeochemical signatures. (*A*) Maximum likelihood Form I RuBisCO RbcL phylogeny (derived from full RuBisCO phylogeny described in Kacar et al. (2017c)). Ancestral AncIB node and descendent Form IB clade highlighted in blue. Approximate likelihood ratio (aLR) branch support indicated by asterisks (**: >10, ****: >1000). Carbon isotope record figure adapted from Garcia et al. (Garcia et al., 2021a), with data from Schidlowski et al. (Schidlowski, 1988) (grey) and Krissansen-Totton et al. (Krissansen-Totton et al., 2015) (dark grey). Approximate age range of AncIB indicated by blue field (see text for discussion). (*B*) Genetic engineering of *S. elongatus* strains. Strain Syn01 was constructed by inserting a second copy of the *rbc* operon in the chromosomal neutral site 2 (NS2). Strain AncIB was constructed by inserting the genetic sequence encoding for the ancestral AncIB *rbcL* within the NS2 *rbc* operon. The native *rbc* operon was removed in both strains Syn01 and AncIB. (*C*) Photosynthetic carbon isotope fractionation (ε_p) of *S. elongatus* strains in this study, cultured in ambient air or 2% CO_2. Each data point represents the mean of three biological replicates and error bars indicate 1σ (error bars smaller than some datapoints).

Figure 2. Structure and sequence features of ancestral RuBisCO. (*A*) Amino acid sequence alignment between ancestral AncIB and extant *S. elongatus* RbcL. Ancestral site variation relative to the *S. elongatus* template is highlighted in red. (*B*) Modeled structure of the ancestral AncIB L_2 dimer (blue), aligned to the active conformation of the extant *S. elongatus* L_8S_8 hexadecamer (grey; PDB: 1RBL (Newman et al., 1993)). Highlighted residues in (*A*) are also highlighted in (*B*). Site numbering from extant *S. elongatus*. Conserved residues are indicated by dots and secondary structure is indicated above the sequences (blue rectangle: α-helix; blue arrow: β-sheet).

Figure 3. (*A*) Growth curves and (*B*) photosynthetic oxygen evolution of *S. elongatus* strains cultured in ambient air or 2% CO_2. (*A, B*) Each data point or bar represents the mean of three biological replicates and error bars indicate 1σ.

Figure 4. Expression and total cell activity of RuBisCO in *S. elongatus* strains. (*A*) Expression of *rbcL* in Syn01 and AncIB detected by RT-qPCR (*secA* reference gene), relative to WT. (*B*) Immunodetection of RbcL protein. *Top*, RbcL signal intensities normalized to those for total soluble protein load. *Bottom*, Western blot showing RbcL protein detected by anti-RuBisCO antibody and total protein stain from crude cell lysates (*C*) Total cell RuBisCO activity, measured with 2.5 mM and 5 mM HCO_3^- concentrations. (*A*-*C*) Each bar represents the mean of three biological replicates (except the bar for 2% CO_2 RbcL signal intensity data in (*B*), which represents the mean of six biological replicates) and error bars indicate 1σ.

Star Methods

Resource availability

Lead contact

- Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Betül Kaçar (bkacar@wisc.edu).

Materials availability

- Strains and plasmids generated in this study are available from the lead contact, Betül Kaçar (bkacar@wisc.edu), upon request.

Data and code availability

- All phylogenetic data has been deposited at GitHub (https://github.com/kacarlab/rubisco) and are publicly available as of the date of publication.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental model and subject details

Bacterial strains

S. elongatus PCC 7942 strains were cultured in BG-11 medium (Rippka et al., 1979) as liquid cultures or on agar plates (1.5% (w/v) agar and 1 mM $Na_2S_2O_3 \cdot 5H_2O$). Liquid cultures were grown at 30°C, continuous shaking at 120 rpm, sparged with ambient air or 2% CO_2, and with 115 µmol photon·m^{-2}·s^{-1} (except for cultures used to prepare samples for O_2 evolution, which were grown with 80 µmol photon·m^{-2}·s^{-1}). The 2% CO_2 gas mix was controlled by an environment chamber (Percival, Cat. No. I36LLVLC8) with a CO_2 tank input. For recombinant strains, liquid and solid media were supplemented with appropriate antibiotics: 2 µg·ml^{-1} Spectinomycin (Sp) plus 2 µg·ml^{-1} Streptomycin (Sm), 5 µg·ml^{-1} Kanamycin (Km). Cyanobacterial growth was measured at an optical density of 750 nm (OD_{750}) and growth parameters were estimated using the Growthcurver package for R (Sprouffske and Wagner, 2016). Cultures were sampled at the middle exponential growth phase, i.e., at an OD_{750} of ~2.5 (AncIB) or ~4.5 (WT and Syn01) for all subsequent experiments.

492 ## Method details

493

494 **Inference of ancestral AncIB RbcL protein.** A RuBisCO RbcL phylogeny was reconstructed as previously
495 described (Kacar et al., 2017c). Briefly, RbcL orthologs were identified from the NCBI protein database by
496 BLAST (sequence dataset and the tree can be found at https://github.com/kacarlab/rubisco). Phylogenetic
497 analysis was performed by Phylobot (Hanson-Smith and Johnson, 2016), a web portal that integrates
498 alignment, phylogenetic reconstruction by RAxML (Stamatakis, 2014), and ancestral sequence inference
499 by PAML (Yang, 2007). A maximum likelihood phylogeny was built using a MSAProbs alignment (Liu et al.,
500 2010) and the best-fit PROTCATWAG model (Lartillot and Philippe, 2004; Whelan and Goldman, 2001),
501 determined by the Akaike information criterion (Abascal et al., 2005). Ancestral states were reconstructed
502 at each amino acid site for all phylogenetic nodes, and gap characters were inferred according to Fitch's
503 parsimony (Fitch, 1971).

504

505 **Cyanobacteria genetic engineering.** Recombinant strains of *S. elongatus* were constructed by natural
506 transformation using standard protocols (Clerico et al., 2007) with minor modifications (Garcia et al.,
507 2021b). The plasmids and strains used in this study are listed in Table S1. The construction of *S. elongatus*
508 Syn01, carrying a single ectopic copy of the *rbc* operon at NS2, as well as the plasmids pSyn01 and pSyn02
509 used to construct strain Syn01, were described previously (Garcia et al., 2021b). The construction of strain
510 AncIB was performed similarly to Syn01. Briefly, pSyn03, which carries the AncIB nucleotide sequence
511 within the entire *rbc* operon (including flanking sequences and homologous regions for recombination at
512 neutral site 2 (NS2) of the *S. elongatus* chromosome), was transformed in *S. elongatus*. Transformation of
513 WT *S. elongatus* with pSyn03 generated strain Syn03 carrying a second copy of the *rbc* operon at NS2.
514 Strain Syn03 was subsequently transformed with pSyn01 to replace the native *rbc* operon with a
515 spectinomycin/streptomycin resistance gene as described previously (Garcia et al., 2021b), producing
516 strain AncIB. Transformants of Syn03 and AncIB were screened for complete segregation by colony PCR
517 using primers F06, R06, F07, and R07 (Fig. S2; Table S3) and the strain sequences at the deletion and
518 insertion sites were further verified by Sanger sequencing using the primers R07, F08, R08, F15, and F16
519 (Table S3). To construct plasmid pSyn03, pSyn02 excluding the *rbcL* coding sequence was PCR-amplified
520 and linearized using primers F13/R13 and assembled with the AncIB *RbcL* coding sequence codon-
521 optimized for *S. elongatus* and synthesized by Twist Bioscience. Both DNA fragments were assembled
522 using the GeneArt™ Seamless Cloning and Assembly Kit (Invitrogen, Cat. No. A13288). (Table S3).

523

524 **RT-qPCR analysis of *rbcL* expression.** Cells were pelleted by centrifugation and resuspended in TE buffer
525 (10 mM Tris, pH 8.0, 1 mM EDTA). Total RNA was extracted using the RNeasy® Protect Bacteria Mini Kit
526 (QIAGEN, Cat. No. 74524). DNase I-treated RNA was then used in reverse transcription (RT) performed
527 with the SuperScript™ IV First-Strand Synthesis System (Invitrogen, Cat. No. 18091050). F09/R09,
528 F11/R11, F14/R14 pairs of qPCR primers (Table S3) were designed with Primer3Plus (Untergasser et al.,
529 2007) (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The quality of cDNA and
530 primer specificity was assessed by PCR using cDNA templates (RT positive reactions) and RT negative
531 controls. qPCR was performed by the real-time thermal cycler qTOWER^3 G (Analytik Jena AG) using
532 qPCRsoft software. The relative expression of native and AncIB *rbcL* was calculated as the average fold
533 change normalized to the *secA* reference gene (Szekeres et al., 2014) using the delta-delta Ct method. The
534 experiment was carried out using three biological replicates and three technical replicates.

535

536 **RbcL protein immunodetection.** Cells were pelleted by centrifugation and resuspended in 95°C TE buffer
537 supplemented with 1% (w/v) SDS and incubated at 95°C for 10 min. The mixture was sonicated and
538 centrifuged to remove cell debris. Total protein concentration in the crude cell lysates was measured using

539 the Pierce™ BCA Protein Assay Kit (Thermo Scientific, Cat. No. 23225). Lysates containing 5 µg of total
540 protein in Laemmli sample buffer were loaded onto a 6% (v/v) polyacrylamide stacking gel. Proteins were
541 electrophoresed in a 12% polyacrylamide resolving gel and blotted onto a nitrocellulose membrane.
542 Detection of RbcL and total protein load was performed as previously described (Garcia et al., 2021b;
543 Kędzior and Kacar, 2021). Detection of RbcL was performed with rabbit anti-RbcL antibody (Agrisera, Cat
544 No. AS03 037) and rabbit IgG secondary antibody (LI-COR Biosciences, Cat No. 926-32211). The
545 densitometric analysis of RbcL signal intensity, normalized to total protein load, was performed with
546 Quantity One® software (Bio-Rad) for three to six biological replicates.
547
548 **RuBisCO assembly confirmation.** Assembly of the RuBisCO large and small subunits into a hexadecameric
549 complex in each strain was evaluated by native gel electrophoresis and immunodetection, as previously
550 described (Garcia et al., 2021b; Kędzior and Kacar, 2021). Immunodetection of the RuBisCO complex was
551 performed for three biological replicates with the same primary and secondary antibodies that were used
552 to detect RbcL, as described above.
553
554 **RuBisCO catalytic activity.** The activity of RuBisCO in cyanobacterial lysates was measured using a
555 spectrophotometric coupled-enzyme assay that links this activity with the rate of NADH oxidation (Kubien
556 et al., 2011). Cell lysis and the activity assay were carried out as previously described (Garcia et al., 2021b)
557 with either 2.5 mM or 5 mM $NaHCO_3$. After 20 min at 25 °C for activation of Rubisco, the reaction was
558 initialized with the addition of ribulose 1,5-bisphosphate (RuBP) (0.5 mM) and the absorbance at 340 nm
559 was recorded using a Synergy H1 plate reader (BioTek). RuBisCO activity was reported as the RuBP
560 consumption rate normalized to total soluble protein content. The assay was performed for three
561 biological replicates.
562
563 **Photosynthetic oxygen evolution rate.** *S. elongatus* strain photosynthetic activity was assayed using a
564 Clark-type oxygen electrode chamber to measure the level of molecular oxygen produced in
565 cyanobacterial cultures. Cells were pelleted and resuspended in fresh BG-11 to an OD_{750} of ~1 following
566 De Porcelinis (De Porcellinis et al., 2018). Concentration of chlorophyll *a* (for normalization) was measured
567 following the protocol by Zavrel et al. (Zavřel et al., 2015a). The remaining suspension was incubated in
568 the dark for 20 min with gentle agitation. Samples from each suspension were analyzed in an oxygen
569 electrode chamber under saturated light, using the Oxygraph+ System (Hansatech Instruments) equipped
570 with the OxyTrace+ software. Oxygen evolution rate was monitored for 10 min and expressed as
571 nanomoles of molecular oxygen evolved per hour per microgram of chlorophyll *a*. The assay was
572 performed for three biological replicates.
573
574 **Carbon isotope fractionation.** Cells were pelleted by centrifugation and washed in 10 mL of 10 mM NaCl
575 (OD_{750} for Syn-1 ~ 4.5, OD_{750} for AncIB ~2.5). Pellets were then dried at 50°C. In parallel, the supernatant
576 from centrifuged culture samples was sterilized through 0.2 µm filtration for DIC isotopic analysis of
577 growth media. Sterilized media was transferred to Exetainer vials leaving no headspace and stored at 4°C
578 until analysis. Isotopic analysis was performed for three biological replicates.
579
580 The carbon isotope composition of bulk biomass ($\delta^{13}C_{biomass}$) and DIC ($\delta^{13}C_{DIC}$) was determined at the UC
581 Davis Stable Isotope Facility. $\delta^{13}C_{biomass}$ was analyzed using a PDZ Europa ANCA-GSL elemental analyzer
582 interfaced to a PDZ Europa 20-20 isotope ratio mass spectrometer (Sercon Ltd.). DIC samples were
583 analyzed by gas evolution and composition was measured by a GasBench II system interfaced to a Delta
584 V Plus IRMS (Thermo Scientific). The carbon isotopic composition values were reported relative to the
585 Vienna PeeDee Belemnite standard (V-PDB):
586

$$\delta^{13}C_{sample} = \left(\frac{^{13}C / ^{12}C_{sample}}{^{13}C / ^{12}C_{V-PDB}} - 1 \right) \times 1000$$

The isotopic composition of dissolved molecular CO_2 ($\delta^{13}C_{CO2}$) was estimated from $\delta^{13}C_{DIC}$ following Rau et al. (Rau et al., 1996) and Mook et al. (Mook et al., 1974):

$$\delta^{13}C_{CO_2} = \delta^{13}C_{DIC} + 23.644 - \frac{9701.5}{T_K}$$

The carbon isotope fractionation associated with photosynthetic CO_2 fixation (ε_p) was calculated relative to $\delta^{13}C_{CO2}$ in the post-culture medium according to Freeman and Hayes (1992):

$$\varepsilon_p = \frac{\delta^{13}C_{CO_2} - \delta^{13}C_{biomass}}{1 + \delta^{13}C_{biomass} / 1000}$$

Quantification and statistical analysis

Results for experimental analyses were presented as the mean and the sample standard deviation (1σ) values of at least three biological replicates. For comparisons of two groups, statistical significance was analyzed by an unpaired, two-tailed *t*-test assuming equal variance. For comparisons of three or more groups, significance was analyzed by one-way ANOVA and a post-hoc Tukey HSD test.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RbcL, RuBisCO large subunit, form I (rabbit)	Agrisera	Cat# AS03 037
IRDye 800CW Goat anti-Rabbit IgG Secondary Antibody	LI-COR	Cat# 926-32211
Experimental models: Organisms/strains		
Synechococcus elongatus PCC 7942 (wild-type, "WT")	Susan S. Golden (UC San Diego)	N/A
Syn01 (*S. elongatus* strain Syn02 with the native *rbc* operon removed: Syn02 and Δ(*rbcL-rbcS-purK)*::*aadA*.)	(Garcia et al., 2021b)	N/A
Syn02 (*S. elongatus* PCC 7942 carrying a second copy of the *rbc* operon and flanking sequences at NS2: NS2::*aphI-rbcL-rbcS-purK*- Synpcc7942_1429-Synpcc7942_1430)	(Garcia et al., 2021b)	N/A

Syn03 (*S. elongatus* PCC 7942 carrying a modified copy of the *rbc* operon (with the ancestral *rbcL*) and flanking sequences at NS2: NS2::*aphI-ancIB-rbcS-purK*-Synpcc7942_1429- Synpcc7942_1430)	This study	N/A
Oligonucleotides		
Primer for genotyping (forward), F06: 5'-GACAATCCTGTTCTCCGGCA-3'	This study	N/A
Primer for genotyping (reverse), R06: 5'-ATCAACGCCGTACCCGTATC -3'	This study	N/A
Primer for genotyping (forward), F07: 5'-GGAGTCAATTCTGCAAGAGC -3'	This study	N/A
Primer for genotyping (reverse), R07: 5'-TCAAGCTCGGTCTACTGC -3'	This study	N/A
Primer for sequencing (forward), F08: 5'-GAATGCTCCGCTGGACTTGC -3'	This study	N/A
Primer for sequencing (reverse), R08: 5'-TGTACTCGATTTGTGCAGCG -3'	This study	N/A
Primer for RT-qPCR (forward), F09: 5'-ACCACCTTGGCAAAATGGTG -3'	This study	N/A
Primer for RT-qPCR (reverse), R09: 5'-TTTGTCGCCTTCCAGTTTGC -3'	This study	N/A
Primer for RT-qPCR (forward), F11: 5'-ATTACCTGCGCGACAACATG -3'	This study	N/A
Primer for RT-qPCR (reverse), R11: 5'-TGCCCGCATGTATTTTTCGC -3'	This study	N/A
Primer for plasmid construction (forward), F13: 5'-GGAGCCTCTGACTATCGCTGGGGGAG -3'	This study	N/A
Primer for plasmid construction (reverse), R13: 5'-GTCGTCTCTCCCTAGAGATATG -3'	This study	N/A
Primer for RT-qPCR (forward), F14: 5'-AAACCGGCGAAATCAAAGGC -3'	This study	N/A
Primer for RT-qPCR (reverse), R14: 5'-CGTGCATGATGATGGGCATG -3'	This study	N/A

Primer for sequencing (forward), F15: 5'-TGTTTGAAGAAGGCAGCGTG -3'	This study	N/A
Primer for sequencing (forward), F16: 5'-CCAAATGGTGCCGCGATAAC -3'	This study	N/A
Recombinant DNA		
pSyn01	This study	N/A
pSyn02	This study	N/A
pSyn03	This study	N/A
Software and algorithms		
Growthcurver	(Sprouffske and Wagner, 2016)	N/A
MSAProbs	(Liu et al., 2010)	N/A
OxyTrace+ software	Hansatech Instruments	https://www.hansatech-instruments.com/product/oxygraph-system/
PAML	(Yang, 2007)	N/A
Phylobot	(Hanson-Smith and Johnson, 2016)	N/A
Primer3Plus	(Untergasser et al., 2007)	(http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi
qPCRsoft software	Analytic Jena AG	https://www.analytik-jena.us/products/life-science/pcr-and-qpcr/real-time-pcr/
Quantity One software	Bio-Rad	SOFT-LIT-70-9600-Q1-468PC
RAxML	(Stamatakis, 2014)	N/A
Other		

Phylogenetic data (protein sequences and phylogenetic trees)	This study	https://github.com/kacarlab/rubisco

References

Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics *21*, 2104-2105.

Adam, Z.R., Skidmore, M.L., Mogk, D.W., and Butterfield, N.J. (2017). A Laurentian record of the earliest fossil eukaryotes. Geology *45*, 387-390.

Anbar, A.D., Duan, Y., Lyons, T.W., Arnold, G.L., Kendall, B., Creaser, R.A., Kaufman, A.J., Gordon, G.W., Scott, C., Garvin, J.*, et al.* (2007). A whiff of oxygen before the great oxidation event? Science *317*, 1903-1906.

Anbar, A.D., and Rouxel, O. (2007). Metal Stable Isotopes in Paleoceanography. Annu Rev Earth Pl Sc *35*, 717-746.

Badger, M.R. (2003). CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot *54*, 609-622.

Bambach, R.K. (1986). Classes and Adaptive Variety: The Ecology of Diversification in Marine Faunas Through the Phanerozoic. In Phanerozoic Diversity Patterns, J.W. Valentine, ed. (Princeton University Press), pp. 191-254.

Beard, B.L. (1999). Iron Isotope Biosignatures. Science *285*, 1889-1892.

Bell, E.A., Boehnke, P., Harrison, T.M., and Mao, W.L. (2015). Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. Proc Natl Acad Sci USA *112*, 14518-14521.

Berla, B.M., Saha, R., Immethun, C.M., Maranas, C.D., Moon, T.S., and Pakrasi, H.B. (2013). Synthetic biology of cyanobacteria: unique challenges and opportunities. Front Microbiol *4*.

Bidigare, R.R., Fluegge, A., Freeman, K.H., Hanson, K.L., Hayes, J.M., Hollander, D., Jasper, J.P., King, L.L., Laws, E.A., Milder, J.*, et al.* (1997). Consistent fractionation of ^{13}C in nature and in the laboratory: growth-rate effects in some haptophyte algae. Global Biogeochem Cycles *11*, 279-292.

Bottjer, D.J., and Ausich, W.I. (2016). Phanerozoic development of tiering in soft substrata suspension-feeding communities. Paleobiology *12*, 400-420.

Burnap, R., Hagemann, M., and Kaplan, A. (2015). Regulation of CO_2 Concentrating Mechanism in Cyanobacteria. Life *5*, 348-371.

Butterfield, N.J. (2000). *Bangiomorpha pubescens* n. gen., n. sp.: implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. Paleobiology *26*, 386-404.

641 Butterfield, N.J. (2007). Macroevolution and Macroecology through Deep Time. Palaeontology *50*, 41-55.

642 Cameron, Jeffrey C., Wilson, Steven C., Bernstein, Susan L., and Kerfeld, Cheryl A. (2013). Biogenesis of a
643 Bacterial Organelle: The Carboxysome Assembly Pathway. Cell *155*, 1131-1140.

644 Catling, D.C., and Zahnle, K.J. (2020). The Archean atmosphere. Sci Adv *6*, eaax1420.

645 Chen, Y., Holtman, C.K., Taton, A., and Golden, S.S. (2012). Functional Analysis of the Synechococcus
646 elongatus PCC 7942 Genome. In Functional Genomics and Evolution of Photosynthetic Systems, R.
647 Burnap, and W. Vermaas, eds. (Dordrecht: Springer), pp. 119-137.

648 Christin, P.-A., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F., and Besnard, G. (2008). Evolutionary
649 Switch and Genetic Convergence on rbcL following the Evolution of C4 Photosynthesis. Mol Biol Evol *25*,
650 2361-2368.

651 Clerico, E.M., Ditty, J.L., and Golden, S.S. (2007). Specialized techniques for site-directed mutagenesis in
652 cyanobacteria. Methods Mol Biol *362*, 155-171.

653 Crowe, S.A., Dossing, L.N., Beukes, N.J., Bau, M., Kruger, S.J., Frei, R., and Canfield, D.E. (2013).
654 Atmospheric oxygenation three billion years ago. Nature *501*, 535-538.

655 De Porcellinis, A.J., Nørgaard, H., Brey, L.M.F., Erstad, S.M., Jones, P.R., Heazlewood, J.L., and Sakuragi, Y.
656 (2018). Overexpression of bifunctional fructose-1,6-bisphosphatase/sedoheptulose-1,7-bisphosphatase
657 leads to enhanced photosynthesis and global reprogramming of carbon metabolism in *Synechococcus*
658 sp. PCC 7002. Metabolic Engineering *47*, 170-183.

659 Eichner, M., Thoms, S., Kranz, S.A., and Rost, B. (2015). Cellular inorganic carbon fluxes in
660 *Trichodesmium*: a combined approach using measurements and modelling. J Exp Bot *66*, 749-759.

661 Erwin, D.H. (2011). Evolutionary uniformitarianism. Dev Biol *357*, 27-34.

662 Fitch, W.M. (1971). Toward Defining the Course of Evolution: Minimum Change for a Specific Tree
663 Topology. Syst Zool *20*.

664 Flamholz, A.I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y.M., Oltrogge, L.M., Alves, R., Savage, D., and
665 Milo, R. (2019). Revisiting trade-offs between Rubisco kinetic parameters. Biochemistry *58*, 3365-3376.

666 Freeman, K.H., and Hayes, J.M. (1992). Fractionation of carbon isotopes by phytoplankton and estimates
667 of ancient CO_2 levels. Global Biogeochem Cycles *6*, 185-198.

668 Garcia, A.K., Cavanaugh, C.M., and Kacar, B. (2021a). The curious consistency of carbon biosignatures
669 over billions of years of Earth-life coevolution. ISME J.

670 Garcia, A.K., and Kacar, B. (2019). How to resurrect ancestral proteins as proxies for ancient
671 biogeochemistry. Free Radic Biol Med *140*, 260-269.

672 Garcia, A.K., Kedzior, M., Taton, A., Li, M., Young, J.N., and Kaçar, B. (2021b). System-level effects of CO_2
673 and RuBisCO concentration on carbon isotope fractionation. In bioRxiv.

Garcia, A.K., McShea, H., Kolaczkowski, B., and Kacar, B. (2020). Reconstructing the evolutionary history of nitrogenases: Evidence for ancestral molybdenum-cofactor utilization. Geobiology *18*, 394-411.

Gibson, T.M., Shih, P.M., Cumming, V.M., Fischer, W.W., Crockford, P.W., Hodgskiss, M.S.W., Wörndle, S., Creaser, R.A., Rainbird, R.H., Skulski, T.M.*, et al.* (2017). Precise age of *Bangiomorpha pubescens* dates the origin of eukaryotic photosynthesis. Geology *46*, 135-138.

Goldsmith, M., and Tawfik, D.S. (2009). Potential role of phenotypic mutations in the evolution of protein expression and stability. Proc Natl Acad Sci USA *106*, 6197-6202.

Golubic, S., and Seong-Joo, L. (1999). Early cyanobacterial fossil record: preservation, palaeoenvironments and identification. Eur J Phycol *34*, 339-348.

Gunn, L.H., Martin Avila, E., Birch, R., and Whitney, S.M. (2020). The dependency of red Rubisco on its cognate activase for enhancing plant photosynthesis and growth. Proc Natl Acad Sci USA *117*, 25890-25896.

Guy, R.D., Fogel, M.L., and Berry, J.A. (1993). Photosynthetic fractionation of the stable isotopes of oxygen and carbon. Plant Physiol *101*, 37-47.

Hannisdal, B., and Peters, S.E. (2011). Phanerozoic Earth System Evolution and Marine Biodiversity. Science *334*, 1121-1124.

Hanson-Smith, V., and Johnson, A. (2016). PhyloBot: A Web Portal for Automated Phylogenetics, Ancestral Sequence Reconstruction, and Exploration of Mutational Trajectories. PLoS Comput Biol *12*, e1004976.

Hayes, J.M. (1993). Factors controlling 13C contents of sedimentary organic compounds: Principles and evidence. Mar Geol *113*, 111-125.

Hendy, A.J.W. (2010). Taphonomic Overprints on Phanerozoic Trends in Biodiversity: Lithification and Other Secular Megabiases. In Taphonomy, P.A. Allison, and D.J. Bottjer, eds. (Dordrecht: Springer), pp. 19-77.

Hurley, S.J., Wing, B.A., Jasper, C.E., Hill, N.C., and Cameron, J.C. (2021). Carbon isotope evidence for the global physiology of Proterozoic cyanobacteria. Sci Adv *7*.

Kacar, B., Ge, X., Sanyal, S., and Gaucher, E.A. (2017a). Experimental Evolution of Escherichia coli Harboring an Ancient Translation Protein. J Mol Evol *84*, 69-84.

Kacar, B., Guy, L., Smith, E., and Baross, J. (2017b). Resurrecting ancestral genes in bacteria to interpret ancient biosignatures. Philos Trans R Soc A *375*.

Kacar, B., Hanson-Smith, V., Adam, Z.R., and Boekelheide, N. (2017c). Constraining the timing of the Great Oxidation Event within the Rubisco phylogenetic tree. Geobiology *15*, 628-640.

Kędzior, M., and Kacar, B. (2021). Quantification of RuBisCO Expression and Photosynthetic Oxygen Evolution in Cyanobacteria. Bio-Protocol *11*.

708 Knoll, A.H. (2015). Paleobiological Perspectives on Early Microbial Evolution. CSH Perspect Biol *7*,
709 a018093.

710 Krissansen-Totton, J., Buick, R., and Catling, D.C. (2015). A statistical analysis of the carbon isotope
711 record from the Archean to Phanerozoic and implications for the rise of oxygen. Am J Sci *315*, 275-316.

712 Kubien, D.S., Brown, C.M., and Kane, H.J. (2011). Quantifying the amount and activity of Rubisco in
713 leaves. Methods Mol Biol *684*, 349-362.

714 Kubien, D.S., Whitney, S.M., Moore, P.V., and Jesson, L.K. (2008). The biochemistry of Rubisco in
715 Flaveria. J Exp Bot *59*, 1767-1777.

716 Lartillot, N., and Philippe, H. (2004). A Bayesian mixture model for across-site heterogeneities in the
717 amino-acid replacement process. Mol Biol Evol *21*, 1095-1109.

718 Liu, Y., Schmidt, B., and Maskell, D.L. (2010). MSAProbs: multiple sequence alignment based on pair
719 hidden Markov models and partition function posterior probabilities. Bioinformatics *26*, 1958-1964.

720 Long, B.M., Förster, B., Pulsford, S.B., Price, G.D., and Badger, M.R. (2021). Rubisco proton production
721 can drive the elevation of CO_2 within condensates and carboxysomes. Proc Natl Acad Sci USA *118*.

722 Lyons, T.W., Reinhard, C.T., and Planavsky, N.J. (2014). The rise of oxygen in Earth's early ocean and
723 atmosphere. Nature *506*, 307-315.

724 Maruyama, S., and Kim, E. (2020). Evolution of Photosynthetic Eukaryotes; Current Opinion, Perplexity,
725 and a New Perspective. In Symbiosis: Cellular, Molecular, Medical and Evolutionary Aspects, M. Kloc, ed.
726 (Cham: Springer), pp. 337-351.

727 McNevin, D.B., Badger, M.R., Whitney, S.M., von Caemmerer, S., Tcherkez, G.G., and Farquhar, G.D.
728 (2007). Differences in carbon isotope discrimination of three variants of D-ribulose-1,5-bisphosphate
729 carboxylase/oxygenase reflect differences in their catalytic mechanisms. J Biol Chem *282*, 36068-36076.

730 Mook, W.G., Bommerson, J.C., and Staverman, W.H. (1974). Carbon isotope fractionation between
731 dissolved bicarbonate and gaseous carbon dioxide. Earth Planet Sc Lett *22*, 169-176.

732 Newman, J., Brändén, C.I., and Jones, T.A. (1993). Structure determination and refinement of ribulose
733 1,5-bisphosphate carboxylase/oxygenase from *Synechococcus* PCC6301. Acta Crystallogr D *49*, 548-560.

734 Peters, S.E. (2005). Geologic constraints on the macroevolutionary history of marine animals. Proc Natl
735 Acad Sci USA *102*, 12326-12331.

736 Peters, S.E., and Foote, M. (2016). Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology *27*,
737 583-601.

738 Rau, G.H., Riebesell, U., and Wolf-Gladrow, D. (1996). A model of photosynthetic ^{13}C fractionation by
739 marine phytoplankton based on diffusive molecular CO_2 uptake. Mar Ecol Prog Ser *133*, 275-285.

740 Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979). Generic Assignments,
741 Strain Histories and Properties of Pure Cultures of Cyanobacteria. Microbiology *111*, 1-61.

Roeske, C.A., and O'Leary, M.H. (1984). Carbon isotope effects on the enzyme-catalysed carboxylation of ribulose bisphosphate. Biochemistry *23*, 6275-6284.

Schidlowski, M. (1988). A 3,800-Million-Year Isotopic Record of Life from Carbon in Sedimentary-Rocks. Nature *333*, 313-318.

Schidlowski, M. (2001). Carbon isotopes as biogeochemical recorders of life over 3.8 Ga of Earth history: evolution of a concept. Precambrian Res *106*, 117-134.

Schopf, J.W. (1994). Disparate rates, differing fates: tempo and mode of evolution changed from the Precambrian to the Phanerozoic. Proc Natl Acad Sci USA *91*, 6735-6742.

Schopf, J.W. (2011). The paleobiological record of photosynthesis. Photosynth Res *107*, 87-101.

Schopf, J.W., Kitajima, K., Spicuzza, M.J., Kudryavtsev, A.B., and Valley, J.W. (2018). SIMS analyses of the oldest known assemblage of microfossils document their taxon-correlated carbon isotope compositions. Proc Natl Acad Sci USA *115*, 53-58.

Schopf, J.W., Kudryavtsev, A.B., Walter, M.R., Van Kranendonk, M.J., Williford, K.H., Kozdon, R., Valley, J.W., Gallardo, V.A., Espinoza, C., and Flannery, D.T. (2015). Sulfur-cycling fossil bacteria from the 1.8-Ga Duck Creek Formation provide promising evidence of evolution's null hypothesis. Proc Natl Acad Sci USA *112*, 2087-2092.

Schubert, B.A., and Jahren, A.H. (2012). The effect of atmospheric CO_2 concentration on carbon isotope fractionation in C_3 land plants. Geochim Cosmochim Acta *96*, 29-43.

Scott, K.M., Henn-Sax, M., Harmer, T.L., Longo, D.L., Frame, C.H., and Cavanaugh, C.M. (2007). Kinetic isotope effect and biochemical characterization of form IA RubisCO from the marine cyanobacterium *Prochlorococcus marinus* MIT9313. Limnol Oceanogr *52*, 2199-2204.

Sepkoski, J.J. (2015). Biodiversity: Past, Present, and Future. J Paleontol *71*, 533-539.

Shi, D., Xu, Y., and Morel, F.M.M. (2009). Effects of the pH/pCO_2 control method on medium chemistry and phytoplankton growth. Biogeosciences *6*, 1199-1207.

Shih, P.M., Occhialini, A., Cameron, J.C., Andralojc, P.J., Parry, M.A., and Kerfeld, C.A. (2016). Biochemical characterization of predicted Precambrian RuBisCO. Nat Commun *7*, 10382.

Sprouffske, K., and Wagner, A. (2016). Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC Bioinformatics *17*, 172.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.

Studer, R.A., Christin, P.A., Williams, M.A., and Orengo, C.A. (2014). Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. Proc Natl Acad Sci USA *111*, 2223-2228.

774 Szekeres, E., Sicora, C., Dragos, N., and Druga, B. (2014). Selection of proper reference genes for the
775 cyanobacterium Synechococcus PCC 7002 using real-time quantitative PCR. FEMS Microbiol Lett *359*,
776 102-109.

777 Tang, Q., Pang, K., Yuan, X., and Xiao, S. (2020). A one-billion-year-old multicellular chlorophyte. Nat Ecol
778 Evol *4*, 543-549.

779 Tcherkez, G.G., Farquhar, G.D., and Andrews, T.J. (2006). Despite slow catalysis and confused substrate
780 specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proc Natl Acad Sci
781 USA *103*, 7246-7251.

782 Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J.A.M. (2007). Primer3Plus,
783 an enhanced web interface to Primer3. Nucleic Acids Research *35*, W71-W74.

784 Vermeij, G.J. (1973). Adaptation, Versatility, and Evolution. Syst Zool *22*.

785 von Caemmerer, S., Tazoe, Y., Evans, J.R., and Whitney, S.M. (2014). Exploiting transplastomically
786 modified Rubisco to rapidly measure natural diversity in its carbon isotope discrimination using tuneable
787 diode laser spectroscopy. J Exp Bot *65*, 3759-3767.

788 Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from
789 multiple protein families using a maximum-likelihood approach. Mol Biol Evol *18*, 691-699.

790 Wilkes, E.B., Lee, R.B.Y., McClelland, H.L.O., Rickaby, R.E.M., and Pearson, A. (2018). Carbon isotope
791 ratios of coccolith–associated polysaccharides of *Emiliania huxleyi* as a function of growth rate and CO_2
792 concentration. Org Geochem *119*, 1-10.

793 Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol *24*, 1586-1591.

794 Yin, Z., Zhu, M., Bottjer, D.J., Zhao, F., and Tafforeau, P. (2016). Meroblastic cleavage identifies some
795 Ediacaran Doushantuo (China) embryo-like fossils as metazoans. Geology *44*, 735-738.

796 Yoon, H.S., Hackett, J.D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A Molecular Timeline for the
797 Origin of Photosynthetic Eukaryotes. Mol Biol Evol *21*, 809-818.

798 Zavřel, T., Sinetova, M., and Červený, J. (2015a). Measurement of Chlorophyll *a* and Carotenoids
799 Concentration in Cyanobacteria. Bio-Protocol *5*.

800 Zavřel, T., Sinetova, M.A., Búzová, D., Literáková, P., and Červený, J. (2015b). Characterization of a model
801 cyanobacterium *Synechocystis* sp. PCC 6803 autotrophic growth in a flat-panel photobioreactor. Eng Life
802 Sci *15*, 122-132.

803