Dynamic allocation of carbon storage and nutrient-dependent exudation in a revised genome-scale model of *Prochlorococcus*

Shany Ofaim[1,2*], Snorre Sulheim[1,3,4*], Eivind Almaas[3,5], Daniel Sher[2], Daniel Segrè[1,6,7¥]

1. Bioinformatics Program and Biological Design Center, Boston University, Boston, MA, USA
2. Department of Marine Biology, University of Haifa, Haifa, Israel
3. Department of Biotechnology and Food Science, NTNU-Norwegian University of Science and Technology, Trondheim, Norway
4. Department of Biotechnology and Nanomedicine, SINTEF Industry, Trondheim, Norway
5. K.G. Jebsen Center for Genetic Epidemiology, NTNU-Norwegian University of Science and Technology, Trondheim, Norway
6. Department of Biomedical Engineering, Boston University, Boston, MA
7. Department of Physics and Department of Biology, Boston University, MA

* These authors contributed equally to this work

¥ Corresponding author, Email: dsegre@bu.edu

Abstract

Microbial life in the oceans impacts the entire marine ecosystem, global biogeochemistry and climate. The marine cyanobacterium *Prochlorococcus*, an abundant component of this ecosystem, releases a significant fraction of the carbon fixed through photosynthesis, but the amount, timing and molecular composition of released carbon are still poorly understood. These depend on several factors, including nutrient availability, light intensity and glycogen storage. Here we combine multiple computational approaches to provide insight into carbon storage and exudation in *Prochlorococcus*. First, with the aid of a new algorithm for recursive filling of metabolic gaps (ReFill), and through substantial manual curation, we extended an existing genome-scale metabolic model of *Prochlorococcus* MED4. In this revised model (*i*SO595), we decoupled glycogen biosynthesis/degradation from growth, thus enabling dynamic allocation of carbon storage. In contrast to standard implementations of flux balance modeling, we made use of forced influx of carbon and light into the cell, to recapitulate overflow metabolism due to the decoupling of photosynthesis and carbon fixation from growth during nutrient limitation. By using random sampling in the ensuing flux space, we found that storage of glycogen or exudation of organic acids are favored when the growth is nitrogen limited, while exudation of amino acids becomes more likely when phosphate is the limiting resource. We next used COMETS to simulate day-night cycles and found that the model displays dynamic glycogen allocation and exudation of organic acids. The switch from photosynthesis and glycogen storage to glycogen depletion is associated with a redistribution of fluxes from the Entner-Doudoroff to the Pentose Phosphate pathway. Finally, we show that specific gene knockouts in *i*SO595 exhibit dynamic anomalies compatible with experimental observations, further demonstrating the value of this model as a tool to probe the metabolic dynamic of *Prochlorococcus*.

Introduction

Marine phytoplankton perform about one-half of the photosynthesis on Earth (Field et al., 1998). *Prochlorococcus* is one of the most abundant phytoplankton clades in the world's oceans and is estimated to produce about 4 Gt of organic carbon annually (Flombaum et al., 2013). As such, these clades play a key role in a variety of ecosystems (reviewed by (Partensky and Garczarek, 2010; Biller et al., 2015)). Recent evolutionary studies suggested several evolved metabolic innovations contributing to high picocyanobacterial abundance in the harsh oligotrophic ocean waters, (Field et al., 1998) usually limited by several nutrients such as nitrogen, phosphorus, and iron. These innovations include a proteome that contains less nitrogen rich amino acids (Gilbert and Fagan, 2011), membranes that contain glyco- and sulfolipids rather than phospholipids (Van Mooy et al., 2006) and streamlining of the genome associated with outsourcing of important cellular functions to co-occurring organisms (Holtzendorff et al., 2008; Partensky and Garczarek, 2010; Morris et al., 2012; Ma et al., 2018a; Braakman, 2019).

Another innovation employed by these organisms is an increased metabolic rate that in turn manifest in the exudation of organic compounds (Braakman et al., 2017; Braakman, 2019), resulting in typically 2–50% of the carbon fixed by photosynthesis leaving the cell (Kujawinski, 2011; Thornton, 2014). Consequently, dissolved organic matter is an omnipresent component of natural waters. However, it is currently impossible to provide a universal chemical description of the dissolved organic matter (Kujawinski, 2011; Arrieta et al., 2015; Moran et al., 2016), partly because the exuded organic compounds differ between strains and environmental conditions (e.g. (Becker et al., 2014; Ma et al., 2018b)). Nevertheless, in general, phytoplankton exudate includes a small proportion of low-molecular weight compounds, such as organic acids, carbohydrates, and amino acids (Bertilsson and Berglund, 2014), as well as a larger proportion of complex, high-molecular weight compounds (Kujawinski, 2011). Another

64 strategy employed by these bacteria to manage their carbon budget is the internal storage of carbon in
65 polymeric form, specifically, glycogen (Zinser et al., 2009; Reimers et al., 2017; Luan et al., 2019; Szul et
66 al., 2019). Glycogen accumulates in the bacterial cell during the light hours and was recently suggested to
67 have two primary roles; as energy storage in preparation for darkness and as a regulation strategy to
68 manage high-light photosynthesis byproducts (Welkie et al., 2019). The allocation of glycogen is suggested
69 to be tightly associated with the overflow metabolism hypothesis and also known to be widely affected
70 by nutrient limitations (Damrow et al., 2016; Cano et al., 2018; Forchhammer and Schwarz, 2019; Szul et
71 al., 2019). Importantly, the carbon fixed and released by phytoplankton is then used by heterotrophic
72 organisms as a source of energy, whereas the heterotrophic bacteria may recycle nutrient elements and
73 support the growth of phytoplankton in other ways, as suggested by the Black Queen Hypothesis (Amin
74 et al., 2012; Morris et al., 2012; Moran et al., 2016; Cirri and Pohnert, 2019; Moran and Durham, 2019).
75 Thus, carbon fixation, storage and release are tightly intertwined with microbial interactions and microbial
76 ecosystem dynamics.
77 Quantitative models at various scales have provided critical insights into how ocean microbial ecosystems
78 function, and how they are related to broader biogeochemical cycles (Braakman et al., 2017; Coles et al.,
79 2017; Foster et al., 2018; Moradi et al., 2018). Most of these models represent organisms in terms of
80 simplified stoichiometric reactions converting elements into biomass, thus making it possible to
81 incorporate biological processes into dynamic-coupled Earth System models (Follows et al., 2007; Reid,
82 2011). The exponential increase in genomic information on marine organisms provides an opportunity to
83 seek methods to link such detailed genome-scale information to biochemical flows (Coles et al., 2017). In
84 recent years, genome-scale metabolic models (GEMs), combined with linear programming, have made it
85 possible to produce testable predictions of metabolic phenotypes of individual organisms or microbial
86 communities (Gu et al., 2019). This computational framework is based on the identification of individual
87 enzymes and transporters in an organism's genome, and on simplifying assumptions that bypass the need

88 for kinetic parameters (Maarleveld et al., 2013; O'Brien et al., 2015; Casey et al., 2016; Kim et al., 2016;
89 Reimers et al., 2017). While genome-scale modeling has proven to be a powerful approach in
90 cyanobacterial model organisms such as *Synechocystis* sp. PCC 6803, *Synechococcus* elongatus PCC 7942
91 and *Prochlorococcus marinus* MED4 (Knoop et al., 2013; Broddrick et al., 2016; Casey et al., 2016;
92 Yoshikawa et al., 2017), the exudation of organic compounds in phototrophic organisms has not been
93 studied in detail through Flux Balance Analysis (FBA) or similar methods (Varma and Palsson, 1994; Orth
94 et al., 2010). On the other hand, several examples exist of FBA-based predictions of exudation-mediated
95 interactions between different species, including using the COMETS platform (Harcombe et al., 2014). In
96 fact, FBA calculations also suggest that "costless" secretions (i.e. secretions that do not induce a fitness
97 cost) might be quite common, and can support the growth of co-occurring organisms (Pacheco et al.,
98 2018).
99 Experimental evidence and theoretical considerations indicate that *Prochlorococcus* exudes different
100 metabolites in a way that strongly depends on environmental conditions (Dubinsky and Berman-Frank,
101 2001; Szul et al., 2019) as well as on the strain's genetic makeup (Becker et al., 2014; Roth-rosenberg et
102 al., 2019). While GEMs can be used to predict these fluxes, they require modifications to deal with
103 processes not usually considered in FBA, including: (a) the special nature of photon fluxes (which, unlike
104 molecular fluxes, cannot easily be "shut off" at short time scales, (Dubinsky and Berman-Frank, 2001)); (b)
105 the buffering role of intracellular storage molecules such as glycogen. The primary focus of this study is to
106 obtain better knowledge of the potential metabolic effect of a combination of key nutrients (carbon,
107 nitrogen, phosphorous and light) and carbon fixation rate on the allocation (including storage and
108 exudation) of carbon in *Prochlorococcus* using a revised genome-scale metabolic model (Figure 1). We
109 start by describing model revisions and updates to capture the current, most complete metabolic
110 knowledge available for *Prochlorococcus*. Next, we use a variety of FBA approaches to uncover the
111 potential relationships between a set of key nutrients, carbon storage and exudates in static and dynamic

112 (time dependent) settings. The implementation and use of these approaches improve our understanding
113 of the intricate metabolic workings of *Prochloroccoccus* and provide insights on its storage and exudation
114 trends under different environmental conditions.
115

116 Methods

117 Model update and curation

118 The *i*JC568 genome-scale reconstruction of *Prochlorococcus marinus subsp. pastoris str.* CMP1986
119 (referred to throughout the manuscript as MED4) as described by Casey et al. (2016), was used as the
120 starting point for model enhancement. The update process started with an in-depth study of the
121 reconstructed network and available knowledge not previously incorporated into the model of the
122 organism. During this process, we ended up implementing the following specific steps of curation and
123 update: (i) A key modification to the model was the decoupling between the glycogen storage flux and
124 the biomass production. In standard stoichiometric reconstructions for FBA modeling (Thiele et al., 2011;
125 Nogales et al., 2012; Feist et al., 2014; Broddrick et al., 2016; Monk et al., 2017; Kavvas et al., 2018),
126 glycogen is listed as one of the biomass components, thus accounting for the carbon flux into storage.
127 However, given the fixed stoichiometry of biomass composition, this classical implementation cannot
128 account for the time-dependent storage and re-utilization of glycogen observed in picocyanobacteria. We
129 thus removed the glycogen from the biomass function and streamlined the existing glycogen granule
130 representation to a direct link between ADP-Glucose to the production of glycogen (Figure 1B). (ii) In
131 addition to targeted refinement of selected reactions, we used the KEGG database (Kanehisa and Goto,
132 2000) to perform an extensive search of previously known but missing metabolic reaction annotations.
133 Indeed, we found 354 reactions that could be potentially added to the existing network. To incorporate
134 this knowledge, we developed a semi-automated algorithm (ReFill, described below). (iii) We coupled the

135 implementation of the algorithm with several steps of manual curation. These included the addition of
136 transports, such as that of hydrogen peroxide and ethanol, known to diffuse across the cell membrane
137 (Seaver and Imlay, 2001; Noreña-Caro and Benton, 2018), and the addition of the complete Entner-
138 Doudoroff pathway, that has recently been discovered in cyanobacteria (Chen et al., 2016). Additionally,
139 we performed a BLAST search (Supplemental material 1) (Altschul et al., 1990) from which we identified
140 6PG-dehydratase (EC: 4.2.1.12) encoded by PMM0774, thereby completing this pathway in the model
141 reconstruction. (iv) The revised model was checked for redox and elemental balance. Since the biomass
142 function was based on experimental data (Casey et al., 2016), it was not updated. In line with best
143 practices, a *memote* quality assessment (Supplemental material 2) (Lieven et al., 2020), as well as model
144 files and a detailed changelog, are provided at https://github.com/segrelab/Prochlorococcus_Model. All
145 reactions added to iJC658 to form iSO595 are found in Table S1 and modified reactions are found in Table
146 S2.

147 ReFill algorithm

148 Following an extensive search of literature and the KEGG (Kanehisa and Goto, 2000), TransportDB
149 (Elbourne et al., 2017) and Metabolights (Haug et al., 2013) databases, we found a large number of new
150 or previously known but missing reaction, transporter and metabolite annotations. Adding large amounts
151 of data to an existing network might create new gaps and may give rise to new blocked reactions and
152 orphan metabolites. These, in turn, may affect optimization simulations. To add this knowledge to the
153 network in a controlled approach, we developed the semi-automated recursive algorithm ReFill
154 (Recursive Filler (of metabolic gaps)). The algorithm is based on the principle of using diverse information,
155 such as enzyme and reaction annotations, and experimental data (such as metabolomics), to selectively
156 increase the metabolic knowledge of an organism's existing curated genome-scale metabolic network.
157 ReFill makes use of a repository of reactions, in this case KEGG reaction annotations for MED4 absent

158 from the model, to construct all potential chains of reactions connecting two metabolites in the existing
159 network. It systematically tests the potential of adding each new reaction and suggests adding it only if it
160 can be a part of a chain in which all the metabolites are part of a path in the network (Figure 2). This
161 prevents the creation of new orphan metabolites and potential blocked reactions that may lead to
162 infeasible FBA solutions. The algorithm starts by selecting a reaction from the repository. It then inspects
163 each metabolite in the reaction for presence in the existing network. In case a metabolite is not present,
164 the set of available reactions is scanned for other reactions using this metabolite as a substrate or product.
165 If such a reaction is found, it is added to the chain of potential reactions. The algorithm then iteratively
166 expands the chain until either the repository is exhausted or all the metabolites in the most recent
167 reaction added are present in the network. After all the possible chains of new reactions are expanded,
168 the algorithm examines the connectivity of all the metabolites in each chain (See example in Figure 2B).
169 Following the manual addition of transporters found through TransportDB (Elbourne et al., 2017) and
170 Metabolights (Haug et al., 2013) (Study MTBLS567), using the ReFill algorithm, we updated reactions that
171 belong to several different pathways, including metabolism of cofactors and vitamins, carbohydrate
172 metabolism, amino acid metabolism and nucleotide metabolism. A complete list of added reactions can
173 be found in Table S1. ReFill was coded in python 3.7 and generates MATLAB-compatible files formatted
174 to be used with the COBRA MATLAB toolbox, including a list of suggested reactions to add and their gene-
175 reaction rules. Other outputs include the added reaction chains and possible metabolic circuits that can
176 be formed by these additions. All code (including ReFill) and model files are available on
177 https://github.com/segrelab/Prochlorococcus_Model.

178 Parameter sampling

179 To study the effects of combinations of key nutrients on glycogen production and exudation in the *i*SO595
180 model we focused on four parameters representing the uptake fluxes of light, bicarbonate, ammonium,

181 and phosphate. Light and inorganic carbon (bicarbonate) are the substrates for photosynthesis, whereas
182 nitrogen and phosphorus limit the growth of *Prochlorococcus* in large regions of the world ocean (Davey
183 et al., 2008; Moore et al., 2013; Saito et al., 2014), and nutrient limitation is likely to influence the
184 exudation of fixed carbon (Dubinsky and Berman-Frank, 2001). We sampled 10,000 different
185 environmental conditions by drawing random values from uniform distributions of these four parameters.
186 The range of each parameter was based on physiologically relevant ranges we extracted from the
187 literature and on the requirement that each range covers important phase transitions, such as nutrient
188 and light limitations (Table S3). Light flux was converted from micromole quanta m^{-2} sec^{-1} to mmol $gDW^{-1}h^{-1}$
189 similarly to (Nogales et al., 2012) using 8% photosynthesis efficiency rate (Zinser et al., 2009). All
190 parameter uptake fluxes are described in FBA compatible units (mmol $gDW^{-1}h^{-1}$), corresponding values in
191 biogeochemistry relevant units are illustrated in Table S3. As MED4 is a photoautotroph, it is exposed to
192 a constant stream of light during daylight hours. The bacterium is then forced, or 'pushed', to fix carbon
193 even when there is not enough of other elements, such as nitrogen or phosphate, to combine the fixed
194 carbon into biomass. To capture this phenomenon *in silico* we developed a 'push'-FBA framework where
195 we fixed both upper and lower uptake rates of light and bicarbonate (Figure 1A). For the other sampled
196 nutrients, ammonium and phosphate, we defined standard FBA bounds where the maximal uptake rate
197 was set to the sampled value and the lower bound was set to zero. Note that, we only considered uptake
198 of sulfur in the form of sulfate (not hydrogen sulfide), and no upper limit was set for the uptake of sulfate
199 because of its abundance in seawater. The maximum rate of RuBisCO (R00024) was fixed to 4.7 mmol
200 gDW^{-1} h^{-1}, as previously reported (Casey et al., 2016). Before sampling we blocked a set of artificial
201 exchange reactions that were added in the previous version of the model, most likely to allow export of
202 dead-end metabolites that would otherwise limit flux feasibility (Table S4). Subsequently, we removed all
203 unconditionally blocked reactions in the model to speed up computations. For each random sample, we
204 first tested the model for feasibility using FBA (Varma and Palsson, 1994). If the solver returned a solution

205 that was feasible and optimal, we further calculated optimal fluxes with parsimonious FBA (Lewis et al.,
206 2010), and determined the range of possible fluxes at optimum with Flux Variability Analysis (FVA)
207 (Gudmundsson and Thiele, 2010). Exchange fluxes from FBA, parsimonious FBA and FVA were recorded
208 and used in subsequent analyses. All environmental sampling and calculations were performed using
209 CobraPy (Ebrahim et al., 2013) and GUROBI 8.1.1 (Gurobi Optimization, Inc., Houston, TX, USA). The
210 python script used to run the environmental sampling is available in the GitHub repository
211 (https://github.com/segrelab/Prochlorococcus_Model).

212 Statistical analysis of the sampled spaces

213 We sampled 10,000 different environmental conditions based on the flux ranges described above, and
214 analyzed the results of FBA optimization, with the goal of characterizing the distribution of specific
215 import/export fluxes and the correlations between different fluxes. To that end, we calculated Pearson
216 correlations between reaction fluxes in the sampling data using the python (version 3.7) Pandas package
217 version 1.0.3 (McKinney, 2010). While negative values are normally used to define uptake in FBA, we
218 converted them to positive values for the uptake of light, bicarbonate, phosphate, ammonium, and sulfate
219 when calculating correlations to ease interpretation of the results. We also performed hierarchical
220 clustering using the Nearest Point Algorithm in SciPy (Virtanen et al., 2020) to sort the order of the
221 compounds in the correlation matrix. We performed dimensionality reduction on normalized reaction
222 fluxes using the T-distributed Stochastic Neighbor Embedding (t-SNE) method (Maaten and Hinton, 2008)
223 in Scikit-learn (Pedregosa et al., 2011) with perplexity of 50 and 3000 iterations. The reaction fluxes were
224 normalized to [-1,1] by dividing by the maximum absolute flux of each reaction. Finally, the t-SNE
225 transformed data was clustered using HDBSCAN (McInnes et al., 2017) with a minimum cluster size of 200.
226 Transport of inorganic ions, water, and protons were not considered when calculating correlations,

227 dimensionality reduction or clustering. We also discarded transport reactions with no absolute flux value
228 above 10^{-3} mmol gDW^{-1} h^{-1} in any of the environmental samples.
229

230 Dynamic modelling of light absorption during the diel cycle in COMETS

231 Cyanobacteria follow a diel cycle. To capture this dynamic behavior, we extended the Computation Of
232 Microbial Ecosystems in Time and Space (COMETS) platform (Harcombe et al., 2014), and developed a
233 module for diurnal-cycle simulations allowing oscillations of light intensity and light absorption.
234 Attenuation of light through each grid cell was modelled using the Beer-Lambert law, as described
235 previously (Yang, 2011; Gomez et al., 2014):
236

$$I(t,z) = I_0(t)e^{-(a_w + a_{dw}X)z} \quad (1)$$

237

238 Here, $I(t,z)$ is the light irradiance given in mmol photons m^{-2} s^{-1}, t is the time, z is the depth (from the
239 top of the grid cell), a_{dw} is the cell- and wavelength-specific absorption coefficient given in m^2 gDW^{-1},
240 a_w the absorption coefficient of pure water given in m^{-1}, X the biomass concentration in gDW m^{-3}, and
241 $I_0(t)$ time time-dependent incident light irradiance at the top of the grid cell. In the current version, we
242 simplified the process by assuming that the light irradiance is either monochromatic or a sum of the total
243 light bandwidth, and the absorption coefficient should match the wavelength(s) of the light source. The
244 total light attenuation (ΔI) through a grid cell of thickness Δz is then

$$\Delta I(t) = I(t,0) - I(t,\Delta z) = I_0\left(1 - e^{-(a_w + a_{dw}X)\Delta z}\right) \quad (2)$$

245

246 The light absorbed by the cells is a fraction of the total light attenuation, i.e.

$$I_{abs}(t) = \frac{\Delta I(t) \cdot a_{dw}X(t)}{a_w + a_{dw}X(t)}. \quad (3)$$

247

248 The total number of photons absorbed per dry cell weight ($\Phi(t)$) in mmol photons gDW^{-1} s^{-1} by the cells

249 within a grid cell of thickness Δz, volume *V,* and surface area *A* is then

$$\Phi(\mathrm{t}) = \frac{I_{abs}(t) \cdot A}{X \cdot V} = \frac{I_0(t) \cdot A}{X \cdot V} \frac{a_{dw} X(t)}{a_w + a_{dw} X(t)} \left(1 - e^{-(a_w + a_{dw} X)\Delta z}\right). \quad (4)$$

250

251 For all COMETS simulations presented here we have used monochromatic light at 680 nm with a

252 calculated biomass-specific absorption coefficient a_{dw} as previously described (Morel and Bricaud, 1981;

253 Bricaud et al., 2004). Briefly, the biomass-specific absorption is the weighted sum of the absorption

254 coefficients of the light-absorbing pigments divinyl-chlorophyll A and B, since none of the other pigments

255 in *Prochlorococcus* absorb light at 680 nm. Additionally, to account for the discrete distribution of

256 chlorophyll into separate cells, the absorption coefficient is scaled by the packaging factor. All coefficients

257 used to calculate light attenuation and absorption are provided in Table 1.

258

259 Table 1: Coefficients and values used to calculate light absorption in COMETS.

Symbol	Description	Value	Unit	Reference
λ	Wavelength	680	nm	
a_w	Absorption coefficient of water at 680 nm	0.465	m^{-1}	Pope and Fry, 1997
a_{dw}	Biomass-specific absorption coefficient	0.285	m^2 gDW^{-1}	
$a_{dvchl\text{-}A}$	Absorption coefficient of divinyl-chlorophyll A at 680 nm	0.0184	m^2 (mg dvchl-A)$^{-1}$	Bricaud et al., 2004
$a_{dvchl\text{-}B}$	Absorption coefficient of divinyl-chlorophyll B at 680 nm	0.0018	m^2 (mg dvchl-B)$^{-1}$	Bricaud et al., 2004

$c_{dvchl\text{-}A}$	Amount of divinyl-chlorophyll A in	0.0163	g dvchl-A gDW^{-1}	Casey et al., 2016
$c_{dvchl\text{-}B}$	Amount of divinyl-chlorophyll B	0.0013	g dvchl-B gDW^{-1}	Casey et al., 2016
d	Average diameter of MED4	0.6	µm	
n'	Imaginary part of the refractive index at 675 nm	0.01377		Stramski et al., 2001
Q*	Packaging effect at 680 nm	0.945		

The changing light conditions throughout a diel cycle was modelled as

$$I_0(t) = A \max\left(\sin(\omega t), 0\right), \tag{5}$$

where the angular frequency is $\omega = \frac{2\pi}{T}$.

Following the development of the diel cycle simulation capability in COMETS we set out to dynamically simulate the growth of MED4. Since the nutrient uptake follows Michaelis–Menten kinetics, we estimated the kinetic parameters V_{max} and K_m using a heuristic approach from experimental data (Grossowicz et al., 2017), first by finding the range of possible parameter combinations corresponding to the gross growth rate of 0.5 d^{-1} (Figure S1A), and secondly by comparing predicted growth and ammonium depletion with the experimental time-series cultivation data (Figure S1B). The estimated parameters were used in the remaining dynamic FBA simulations in COMETS. Finally, to simulate the dynamic storage and consumption of glycogen we applied a multiple objective approach consisting of the following four steps: (1) Maximization of the flux through the non-growth associated maintenance reaction. Note that, this reaction has an upper bound of 1 mmol gDW^{-1} h^{-1} (Casey et al., 2016) In contrast to standard practice, where one uses a lower bound for the non-growth associated maintenance reaction, this method provides a more realistic scenario where the organism continues to consume resources trying to keep up cellular

277 maintenance even at zero growth; (2) Maximization of growth; (3) Maximization of glycogen production
278 (storage); and (4) Parsimonious objective which minimizes the sum of absolute fluxes. To simulate
279 nitrogen-abundant and nitrogen-poor growth conditions, we used the PRO99 medium with standard (800
280 µMol) and reduced (100 µMol) ammonium concentration, as previously described (Grossowicz et al.,
281 2017). Light availability was modelled as described in Equation 5, with an amplitude of 40 µmol Q m^{-2} s^{-1}
282 and a period of 24 hours. All parameter values used in the COMETS simulations are given in Table S5. All
283 Dynamic growth simulations were performed using COMETS v.2.7.4 with the Gurobi 8.1.1 solver, invoked
284 using the associated MATLAB toolbox (https://github.com/segrelab/comets-toolbox).
285

286 # Simulating growth of knockout mutants

287
288 Simulations of the knockout mutants where performed by constraining the flux to zero for the reactions
289 catalyzed by the enzymes encoded by *glgC* (PMM0769) and *gnd* (PMM0770), respectively. For *glgC*, the
290 reaction is glucose-1-phosphate adenylyltransferase (R00948) and for *gnd* the two reactions are $NADP^+$
291 and NAD^+ associated 6-phosphogluconate dehydrogenases (R01528 and R10221). We then used dynamic
292 FBA in COMETS with PRO99 medium (Moore et al., 2007) with limited ammonium and diel light conditions
293 to simulate growth over 7 days. We included an arbitrarily derived death rate of 0.1 d^{-1} (Grossowicz et al.,
294 2017) in these simulations. The growth curves where qualitatively compared with experimental data from
295 Shinde and co-workers (Shinde et al., 2020). The model parameters are given in COMETS parameter and
296 model files in https://github.com/shashany/iSO595/tree/master/COMETS_files.
297

298 # Results and Discussion

299 ## Model curation and update

300 *Prochlorococcus* fixes carbon through photosynthesis during daytime. Fixed carbon that is neither used
301 for cell growth nor stored in the form of glycogen is exuded. Here, we set out to study dynamic changes
302 in the carbon allocation and storage mechanisms in MED4 using a genome-scale metabolic modelling
303 approach. To that end, we first re-curated and updated the available *i*JC586 model (Casey et al., 2016), as
304 described in detail in the Methods section. The update involved the development of a new semi-automatic
305 algorithm (ReFill), which can be broadly applied to other reconstructions (Methods). Concurrently, we
306 introduced a revised mechanism for carbon storage, effectively treating glycogen as an independent
307 component of biomass. This configuration makes it possible for glycogen to be accumulated and depleted
308 at variable rates (Figure 1), aligning with the overflow metabolism hypothesis (Szul et al., 2019; de Groot
309 et al., 2020). Other key modifications induced by the ReFill algorithm and subsequent manual curation
310 (see Methods) include the completion of the Entner-Doudoroff (ED) pathway, recently discovered in
311 cyanobacteria (Chen et al., 2016) and proposed as the primary *Prochlorococcus* glucose metabolism
312 pathway under mixotrophic conditions (Biller et al., 2018). Additional revisions focused on the exudation
313 of fixed carbon products from the cell and included various transports such as pyruvate, fumarate, citrate,
314 ethanol, various nucleotides and hydrogen peroxide as well as metabolites found in both the endo- and
315 exo- metabolome of *Prochlorococcus* (Metabolights study MTBLS567). The end product of our revision,
316 reconstruction iSO595, has 595 genes, 802 metabolites and 994 reactions, i.e. 27 genes, 123 metabolites
317 and 196 reactions more than the previous version, *i*JC568 (Figure 1D).
318

Carbon fixation and storage is affected by nutrient uptake rate

Prochlorococcus thrive in oligotrophic environments (Johnson et al., 2006), where, in surface waters, its growth and carbon fixation rates are usually limited by the abundance of nitrogen, phosphate or iron (Krumhardt et al., 2013; Szul et al., 2019). Deeper in the water column *Prochlorococcus* growth becomes limited by light (Vaulot et al., 1995). We set out to explore the combined effect of different levels of light and nutrients on carbon fixation, storage and exudation. Similarly to Phenotypic Phase Plane analysis (Edwards et al., 2002), we sought a global perspective of metabolism in this multi-parameter spaces while explicitly taking into account the fact that the inflow of light and bicarbonate may not be easily controllable by the cell, and that *Prochlorococcus* may need to deal with excess amounts of fixed carbon. Thus, in contrast to normal FBA where the uptake of metabolites is constrained by an upper bound, we introduced a 'push- FBA' approach (Figure 1A), in which the influx of bicarbonate and light have a fixed imposed value (See Methods and Table S3 for specific values used). This approach attempts to mimic implications of photosynthesis, in which light is the driving force. Once photons are absorbed by the chlorophyll, the energy must be used to produce ATP and reducing power, otherwise it is dissipated in ways that may cause cell damage (Long et al., 1994). This subtle difference in applied constraints has major effects on model predictions. While flux rearrangement is usually viewed as a consequence of environmental nutrient limitations, the results of this analysis shows that a substantial rewiring of fluxes is caused by this imposed excess of fixed carbon as well.

To understand how different combinations of environmental parameters (availability of nitrogen, phosphate, light and bicarbonate) affect the way *Prochlorococcus* can manage its carbon budget, we implemented FBA under 10,000 randomly sampled growth environments. Overall, this sampling analysis demonstrated that the exudation of organic acids, amino-acids and nucleobases/nucleosides, as well as the extent of glycogen storage, are strongly modulated by environmental factors (Figure 3). To observe the full range of possible optimal solutions per sample, we implemented and compared different flux

343 balance analysis methods, including flux variability analysis (FVA) and parsimonious FBA (pFBA). These
344 two methods provide complementary insight: FVA estimates the range of possible values for the flux of
345 each reaction at the optimum, providing insight into the structure of the phenotypic space at maximal
346 growth rate. In contrast, pFBA, by minimizing the sum of fluxes at optimality, generates flux predictions
347 less likely to involve unrealistic loops, and thus potentially provides predictions closer to experimental
348 values (Lewis et al., 2010). Together, these two FBA methods help analyze the solutions of our high-
349 dimensionality dataset.

350 Our predictions simulate the metabolic effects and variability in glycogen production modulated by
351 environmental constraints (Figure 3). Glycogen production was observed only above light levels of 50
352 mmol gDW^{-1} h^{-1} (corresponding to 7.5 micromole quanta m^{-2} sec^{-1}), and decreased as ammonium and
353 phosphate concentrations increase, in agreement with previous studies (Szul et al., 2019). Interestingly,
354 FVA consistently predicted the glycogen production range minimal value to be zero across all samples.
355 This implies that glycogen storage is possible, but not necessary to achieve optimal growth in the feasible
356 solution space. This was also the case in the more stringent pFBA analysis, indicating that while
357 metabolism may be a strong modulator of glycogen metabolism, more types of regulation, not accounted
358 for in FBA, are involved. One example of such regulation may be allosteric regulation of ADP-glucose
359 pyrophosphorylase by 3-phosphoglycerate (Iglesias et al., 1991), possibly in combination with redox
360 regulation (Díaz-Troya et al., 2014). Specific regulation aimed at tuning up glycogen storage may also occur
361 at the transcriptional level, e.g. by multiple transcription factors previously suggested to be involved in
362 the regulation of glycogen metabolism in fluctuating environments (Luan et al., 2019).

363 The range of possible rates of glycogen production (through FVA) displays a bell-shaped bicarbonate
364 dependent distribution, indicating low storage of glycogen (zero flux) under both low and high uptake
365 rates of bicarbonate. When bicarbonate uptake rates are low, all available carbon is diverted into growth.
366 The reduced glycogen storage at high bicarbonate uptake, when RuBisCO is saturated, seems to be caused

367 by the increased ATP demand associated with the conversion of bicarbonate to exudation-products, since
368 the onset and rate of change of this trade-off is modulated by the ATP availability, as demonstrated by
369 phenotypic phase planes analysis (Figure S2). This agrees with recent work suggesting that
370 *Prochlorococcus* use available ATP to drive pathways to saturation by shifting reaction directions towards
371 favoring dephosphorylation of ATP to ADP, disrupting the cellular ATP/ADP ratio and increasing the
372 metabolic rate of the cell by pushing forward ATP consuming reactions, until it is restored. Together with
373 organic carbon exudation this strategy allows for growth in lower nutrient concentrations (Braakman,
374 2019).
375 We next sought to explore the effect of combinations of key nutrients on storage and exudation patterns
376 in our sampling spaces. To explore the strongest trends, we chose to employ a high stringency approach
377 and use only our set of pFBA results. Due to the nature of pFBA, any exudation observed in this analysis
378 could not be easily removed without imposing a cost on growth. To that end, we visualized the data using
379 t-SNE clustering (Figure 4A) on our pFBA set of results. We observed 6 typical phenotypes (clusters) rising
380 out of the sampling spaces (Figure 4). These 6 phenotypes are characterized by subtle differences in
381 combinations of environmental parameters, yielding significantly different exudation patterns. Generally,
382 we observed the highest biomass value in phenotype 5, and the lowest in phenotype 4. All key nutrient
383 uptake rates were highly variable (ranging from 33 to 44% variability). Phenotype 1 is characterized by
384 high light, bicarbonate, a maximum RuBisCO flux (indicating maximal photosynthesis rate) but low
385 nitrogen uptake. Additionally, we observed high exudation of pyruvate coming from the pentose
386 phosphate and Entner-Doudoroff pathways. Both are alternative routes coming out of carbon fixation
387 (Waldbauer et al., 2012; Chen et al., 2016). Together with a low biomass value, this phenotype might
388 indicate a scenario of exudation due to overflow metabolism.
389 The two largest clusters (number 2 and 3), tie together high and low light, carbon and nitrogen uptake
390 rates, and different exudation patterns. Interestingly, phenotype 3 (high light) showed exudation of

391 fumarate and malate while phenotype 2 (low light) did not. Recent work suggested that, in high light
392 conditions, fumarate is generated through oxaloacetate and malate creating a broken acyclic form of the
393 TCA cycle, while in the dark, fluxes are diverted into forming the cyclic form of it. This low light form of
394 the TCA cycle is then active and works towards energy generation (Xiong et al., 2017). Similarly, we
395 observed two forms of the TCA cycle in the high and low light phenotypes (2 and 3 respectively) with a
396 difference in the direction of one reaction (KEGG R00342, Figure S3). Phenotype 2, describing low-light
397 conditions, showed the L-Malate/oxaloacetate balance to shift in favor of oxaloacetate, completing the
398 route towards 2-Oxoglutarate, a key metabolite known to act as a starvation signal and modulator of the
399 C/N balance in cyanobacteria (Domínguez-Martín et al., 2018; Zhang et al., 2018), and subsequently into
400 energy generation. On the other hand, Phenotype 3, describing high light conditions, showed the L-
401 Malate/oxaloacetate balance to shift in favor of L-Malate and away from the formation of 2-oxoglutarate.
402 In both phenotypes fumarate is converted to L-Malate. While in Phenotype 2 it is fed into a semi-cyclic
403 form of the TCA cycle, fumarate is partly exuded and partly converted to L-malate in phenotype 3, in
404 agreement with overflow metabolism.
405 We observed a similar TCA cycle flux distribution in phenotype 4 leading to high exudation of L-Malate.
406 Interestingly, Phenotype 1 and 4 are comparable in all key nutrients except light (High in phenotype 1 and
407 low in phenotype 4). As a result of an in-depth flux distribution analysis, we observed a reaction direction
408 change in UDP-glucose:NAD+ 6-oxidoreductase [R00286, EC 1.1.1.22, PMM1261] between the two
409 phenotypes. In phenotype 4 this reaction shifted towards the creation of UDP-glucose, a precursor for the
410 production of glycogen (Due to the high stringency of this analysis we did not observe the direct formation
411 of glycogen). In phenotype 1, this reaction favored the formation of UDP-glucuronate which in turn was
412 diverted into the formation of amino sugars. These phenotypes may correlate to the 12:00 (phenotype 1)
413 and 16:00 (phenotype 4) scenarios described in (Szul et al., 2019). Finally, Phenotypes 5 and 6 may
414 represent a high-light nutrient-rich environment resulting in a high biomass value.

415

416 Nutrient uptake rates modulate exudation of organic compounds

417 The use of genome-scale metabolic models captures a comprehensive picture of the metabolic processes
418 taking place in the cell, including those that lead to metabolite exudation. From the random sampling of
419 environmental conditions, we identified conditions in which organic acids must be exuded. This was
420 noticeable by a non-zero lower bound of the FVA region (Figure 3). Interestingly, organic acids were more
421 likely to be exuded when the growth became limited by phosphate or nitrogen. Since *Prochlorococcus* is
422 known to thrive in oligotrophic ocean gyres where nitrogen or phosphate is limited (Partensky et al., 1999;
423 Flombaum et al., 2013), this represents a likely natural phenotype, and as such, supports previous findings
424 (Szul et al., 2019). Costly metabolites, essential for cell survival and growth, such as amino acids,
425 nucleobases and nucleotides, tend to be exuded in nitrogen and carbon rich conditions and might be a
426 result of overflow metabolism (Cano et al., 2018; Pacheco et al., 2018). To explore this phenomenon in
427 further detail, we looked into exudation patterns of specific metabolites as a function of key nutrient
428 limitations (Figure 5). Of the environmental factors, the uptake of nitrogen (ammonium) is a decisive
429 factor differentiating between exudation of organic acids or amino acids. While it is positively correlated
430 with the exudation of nitrogen-rich compounds such as amino acids, it is negatively correlated with
431 exudation of organic acids and glycogen. Additionally, glycogen formation is positively correlated with the
432 exudation of malate, citrate, fumarate and succinate, which are most of the TCA cycle constituents. This
433 is in line with previous findings suggesting the re-direction of carbon metabolism towards the formation
434 of macromolecules (including glycogen) in nitrogen limiting conditions (Forchhammer and Selim, 2019;
435 Szul et al., 2019). Thus, our reconstruction captured known possible aspects of the carbon/nitrogen
436 balance in *Prochlorococcus*.
437 Finally, we observed a general pattern of strong positive correlations between amino acids, nucleobases,
438 nucleosides, as well as a range of other compounds. In an interesting deviation from this general pattern,

439 L-Aspartate showed a decreased correlation with other exudates. L-Aspartate, together with its role in
440 protein nucleic acid biosynthesis, can serve as a precursor for nitrogen storage metabolites such as
441 polyamines (Szul et al., 2019). Indeed, we observed a slightly stronger correlation between L-Aspartate
442 and the uptake of nitrogen compared to other amino acids. Finally, In contrast to other amino acids, L-
443 Aspartate is negatively correlated with light uptake and hydrogen peroxide exudation. Hydrogen peroxide
444 is produced from L-Aspartate and oxygen by L-aspartate oxidase [R00481, EC 1.4.3.16, PMM0100]. L-
445 amino acid oxidases have been previously described in cyanobacteria and have been related to the use of
446 amino acids as carbon sources (Campillo-Brocal et al., 2015). The production of hydrogen peroxide is also
447 strongly correlated with light, a result consistent with the expectation that reactive oxygen species are
448 created during photosynthesis.
449

450 Dynamic allocation of carbon storage

451 Nutrient and light limitations are well-known modulators of carbon storage in *Prochlorococcus* (Zinser et
452 al., 2009; Szul et al., 2019). Recent work has suggested the storage of carbon to be one of the major
453 metabolic tasks during the day-night cycle (Cano et al., 2018; Szul et al., 2019; Shinde et al., 2020).
454 Sampling 10,000 environmental conditions showed that glycogen production decreased at lower light
455 intensities and was less saturated at higher rates (Figure 3). To explore other time-modulated trade-offs
456 and trends related to carbon storage, we performed *in silico* dynamic FBA diel-cycle simulations using the
457 Computational of Microbial Ecosystems in Time and Space (COMETS) platform (Mahadevan et al. 2002;
458 Harcombe et al. 2014). COMETS relies on kinetic information such as K_m and V_{max} to simulate the spatial
459 growth and exudation patterns of microbes in a simulated discretized time course. To improve the
460 accuracy and biological relevance of our simulations we used kinetic constants either obtained from
461 experimental measurements reported in the literature (Krumhardt et al., 2013; Hopkinson et al., 2014)
462 (Table S3) or from fitting model simulations to measured growth and depletion of ammonium rates

463 (Grossowicz et al., 2017). We found K_m and V_{max} values of 0.39 mM and 0.9 mmol gDW^{-1} h^{-1} for the uptake
464 of ammonium to best fit the experimental data (Grossowicz et al., 2017) (Figure S1). Surprisingly, the
465 estimated K_m value is 3 orders of magnitude larger than previous estimates (Grossowicz et al., 2017),
466 indicating some uncertainty in this estimate that may be affected by the choice of method for calculating
467 this value.
468 Since the tight coupling between carbon and nitrogen metabolism in cyanobacteria is known to influence
469 carbon allocation and storage (Zhang et al., 2018; Szul et al., 2019), it was chosen as a case study. As such,
470 we focused in more detail on the dynamic changes in metabolism in nitrogen-abundant and nitrogen-poor
471 media, as previously defined (Grossowicz et al., 2017). Specifically, we set out to explore glycogen
472 production and consumption with COMETS in these conditions (Figure 6). We did not observe glycogen
473 storage in nitrogen-abundant simulations. One explanation for this may arise from the limitations of the
474 platform. The modelling framework is based on linear programming to optimize an ordered set of
475 assumed cellular objectives, glycogen is only stored when there is excess energy and carbon available,
476 which is the case when the growth is nitrogen limited. Another limitation that might affect glycogen
477 storage in these simulations might involve other types regulatory mechanisms not usually accounted for
478 in dynamic FBA. The addition of regulatory layers, such as proteomics data (Reimers et al., 2017), or more
479 specifically tailored objective functions could lead to smaller but nonzero generation of glycogen also
480 during nitrogen-rich conditions.
481 In agreement with previous work (Szul et al., 2019), under nitrogen-limiting conditions, glycogen
482 accumulates throughout the day and is subsequently used to support respiration and growth during the
483 night (Figure 6A). However, the predicted glycogen storage is not sufficient to support neither growth nor
484 cellular maintenance during the night and may explain the reported increased death rate during night
485 time (Zinser et al., 2009; Ribalet et al., 2015). Interestingly, the model predicts consumption of glycogen
486 during dusk to increase growth when photosynthesis is declining (Figure 6), closely resembling

487 observations in *Synechococcus,* in particular for the *ΔkaiC* mutant with a dysfunctional circadian clock
488 (Diamond et al., 2015).
489 The switch from photosynthesis at daytime to glycogen consumption at nighttime is reflected in the
490 metabolic shifts observed in key pathways (Figure 6B). Interestingly, we observed higher fluxes through
491 the Calvin cycle in nitrogen-poor conditions. This difference may be caused by the increased ATP demand
492 necessary to support higher growth rates in nitrogen-abundant conditions. Additionally, our simulations
493 predicted that the use of the Entner-Doudoroff pathway during photosynthesis creates precursor
494 metabolites for growth during light hours, and a shift to the Pentose Phosphate Pathway (PPP) during
495 nighttime. This trend might occur as an alternative for generating NADPH (Figure S4). Upregulation of the
496 PPP enzymes during dusk and the first half of the night time was also observed in the proteome of
497 *Prochlorococcus* (Waldbauer et al., 2012). Several enzymatic transformations participate in both the
498 Calvin cycle and the PPP, although in opposite directions (Waldbauer et al., 2012). These transformations
499 were captured in our simulations, specifically as demonstrated by transketolase (Figure 6B). Additionally,
500 the consumption of glycogen during nighttime leads to exudation of pyruvate and formate (Figure 6C).
501 This prediction is supported by previous observations; formate is exuded when phosphonates are
502 metabolized in the closely related strain *Prochlorococcus marinus* MIT9301 (Sosa et al., 2019) and
503 pyruvate is exuded when fixed carbon is consumed in *Synechococcus* elongatus PCC 7942 and
504 *Synechocystis* sp. PCC 6803 (Carrieri et al., 2012; Benson et al., 2016).
505 The shift from photosynthesis and carbon fixation to glycogen catabolism is also associated with a switch
506 in production and consumption of energetic cofactors (Figure S3). Generation of ATP is performed
507 concomitantly by ATP synthase in both the thylakoid membrane and the periplasmic membrane during
508 photosynthesis. The periplasmic ATP synthase is first driven by reduced cofactors (NADPH) generated by
509 the electron transport chain in the light-dependent part of photosynthesis (Figure S3). ATP is consumed
510 by two separate processes: growth- and maintenance-associated reactions reach a threshold once growth

511 is limited by the nitrogen abundance, while the recycling of precursors for the Calvin cycle follows the
512 shape of light absorption throughout the day. In agreement with previous work (Park and Choi, 2017), our
513 model predicted higher rates of NADPH production than NADH.
514 Next, we explored the ability of our model to dynamically capture biologically relevant phenotypes by
515 performing dynamic FBA simulations of knock-out mutants in *Prochlorococcus,* focusing on two gene
516 deletions disrupting different parts of glycogen metabolism. Δ*glgC* breaks synthesis of ADP-glucose and
517 thus the storage of glycogen and Δ*gnd*, knocking out 6-phosphogluconate dehydrogenase, a key reaction
518 in the Pentose Phosphate pathway found to fuel the Calvin cycle with precursor metabolites during the
519 onset of photosynthesis (Shinde et al., 2020). Our dynamic FBA simulations in COMETS (Figure S5) showed
520 similar growth between Δ*gnd* and the wild type and slightly lower growth for Δ*glgC*. We set out to
521 compare these observations with available experimental data. Since genetic tools for the modification of
522 *Prochlorococcus* are still lacking (Laurenceau et al., 2020), we chose data from the closely related
523 cyanobacteria *Synechocystis* as recent work described the impact of Δ*glgC* and Δ*gnd* on its growth during
524 diel cycles (Shinde et al., 2020). Indeed, we found very good agreement between measured and predicted
525 growth for both the wild-type and Δ*glgC* mutant where glycogen storage is disrupted (Shinde et al., 2020)
526 (Figure S5). One of the notable limitations of dynamic FBA is the ability to quantify intermediates and
527 precursor pools that might drive the initiation of a pathway. This comes mainly from the assumption of a
528 quasi steady-state of intracellular metabolite pools at each time point. Although the comparison is strictly
529 qualitative and concerns strains with known differences (Mary et al., 2004), these findings demonstrated
530 the ability of our reconstruction to capture metabolic trends in response to genetic perturbations,
531 indicating that *i*SO595 will be a valuable tool in future research of *Prochloroccus.* Overall, our dynamic
532 simulations display biological and physiological behaviors that are consistent with expectations, and at
533 the same time provide valuable insight into the putative internal metabolic processes that might modulate
534 the *Prochlorococcus* growth under environmental and genome-induced constraints.

535 # Conclusions

536 Our study provides a detailed systematic view of the underlying metabolic trends modulating carbon
537 storage and exudation in *Prochlorococcus*. *Prochlorococcus* is known to interact with other bacteria in its
538 surroundings (Sher et al., 2011; Aharonovich and Sher, 2016; Biller et al., 2018; Hennon et al., 2018). It is
539 currently impossible to predict the fluxes of organic matter (or of the myriad metabolites comprising it,
540 such as amino acids, sugars and organic acids) between phytoplankton and bacteria. Yet, quantifying such
541 fluxes and predicting them from genomic surveys, as shown here, serves a number of roles: (1) It can
542 provide experimentally testable and mechanistic hypotheses on inter-microbial exchanges and
543 competition, (2) It has the potential to increase knowledge about the specific metabolites that may
544 mediate these interactions; and (3) It would enable the construction of improved models of
545 biogeochemical cycles which consider the diverse and powerful metabolic capabilities of the ocean
546 microbiome.

547 Genome-scale metabolic-network reconstructions are a powerful tool, but not without limitations. With
548 the development of an innovative semi-automated approach (the ReFill algorithm), we were able to
549 increase existing knowledge in the reconstruction by up to 25%. In contrast to other standard gap filling
550 approaches (Henry et al., 2010; Thiele et al., 2014), or plain addition of newly found reactions, ReFill has
551 the specific capability to add individual reactions through a recursive algorithm that guarantees complete
552 connectivity to the existing network. However, this approach employs high stringency and thus adds
553 limited amounts of knowledge. Considering that *Prochlorococcus* strains have some of the smallest known
554 genomes among free-living organisms, a 25% increase in knowledge serves as a significant improvement
555 in the predictive capacity of the model. However, reconstruction of high-quality genome-scale metabolic
556 models is an iterative process, where new data, knowledge, and scope creates a demand and possibility
557 for further model improvement. One example of this possibility is the CO_2 concentrating mechanisms in
558 *Prochlorococcus.* This mechanism is known to be sustained by proton and ion gradients across the cell

559 membrane at an energetic cost (Hopkinson et al., 2014; Burnap et al., 2015). However, the comprehensive
560 knowledge and annotation of ion transporters necessary to model this mechanism are lacking, and
561 therefore, not included in iSO595. With the advancement of data collection and annotation tools, together
562 with using ReFill or similar algorithms, metabolic knowledge can be easily added to such reconstructions
563 and improve their predictive abilities and mimicry of biological and physiological processes. Our findings
564 contribute to a growing body of work on the underlying metabolic mechanisms modulating the metabolic
565 success of *Prochlorococcus*. The approaches shown here provide systematic insights corroborated in
566 recent and well-known works and provide strong foundations for future studies of *Prochlorococcus*
567 metabolism with particular interest in its interaction with other microorganisms and the effects of these
568 on community composition and larger biogeochemical cycles.
569

570 Conflict of interest statement

571 The authors declare no conflict of interest.
572

573 Author contributions

574 DSh and DSe designed the study. SO, SS and DSe developed the computational models and performed the
575 computational analyses. SO and SS wrote a first version of the manuscript. DSe, DSh and EA, oversaw the
576 project and contributed to the final version of this manuscript. All authors have read and approved the
577 final version of the manuscript.
578
579

Funding

This work was supported by the Human Frontiers Science Program (grant RGP0020/2016) and the National Science Foundation (NSFOCE-BSF 1635070) to DSe and Dsh. DSe also acknowledges support by the National Science Foundation (grants 1457695), the Directorates for Biological Sciences and Geosciences at the National Science Foundation and NASA (agreement nos. 80NSSC17K0295, 80NSSC17K0296 and 1724150) issued through the Astrobiology Program of the Science Mission Directorate, and the Boston University Interdisciplinary Biomedical Research Office. SS was funded by SINTEF, the Norwegian graduate research school in bioinformatics, biostatistics and systems biology (NORBIS) and by the INBioPharm project of the Centre for Digital Life Norway (Research Council of Norway grant no. 248885).

Supplemental material

Figure S1: Estimation of kinetic parameters for the uptake of ammonium in *Prochlorococcus.* A) All combinations of Km and Vmax along the red trajectory matches the observed gross growth rate of 0.5 d^{-1} (Grossowicz et al., 2017). However, when we compare the dynamics of cell density **(B)** and ammonium concentration **(C)** we find that the best overall prediction is achieved using Km = 0.39 mM and Vmax = 0.9 mmol gDW^{-1} h^{-1} (marked by an orange dot in **A**).

Figure S2: ATP availability influence modulates the trade-off between glycogen storage and growth. Phenotypic phase planes (Edwards et al., 2002) illustrate the combined effect of glycogen storage and bicarbonate uptake on the maximal growth rate. Compared to the base model **(A)**, we observe how that the trade-off is strongly affected by modulated ATP availability, either from an artificial reaction providing

602 extra ATP **(B)** or by increasing **(C)** or decreasing **(D)** the amount of available light. Increasing ATP allows
603 more glycogen storage without reducing the growth rate.
604
605 **Figure S3: TCA cycle flux diagram differences between the most common phenotypes.** Flux diagrams of
606 the TCA cycle in the most common phenotypes 2 (colored orange) and 3 (colored blue). Reactions are
607 denoted by KEGG reaction ids. Reaction colors correspond to cluster colors presented in figure 4.
608
609 **Figure S4: The transition from daytime to nighttime is associated with a drastic change in the production**
610 **and consumption of the energy-carrying cofactors. The figure panels show the major sources (left) and**
611 **drains (right) of the cofactors ATP, NAPDH and NADH.** The legend shows the reaction IDs used in iSO595.
612 **A)** ATP is produced by both the thylakoid (R00086th) and the periplasmic (R00086p) ATP synthase during
613 daytime, but mostly by the periplasmic ATP synthase (respiration) during nighttime. ATP is consumed by
614 reactions associated with growth (BIOMASS and BProtein), cellular maintenance (Maintenance) and
615 storage of glycogen (R00948) in addition to reactions recycling precursors for the Calvin cycle (R01512 and
616 R01523) and acetyl-CoA carboxylase (R00742). **B)** NADPH is produced by ferredoxin reductase (fdr) during
617 daytime and by the pentose phosphate pathway (R02736 and R01528) during nighttime. The NADPH is
618 either used to drive the proton gradient across the periplasmic membrane (NADPHDHp) or in
619 Gluconeogenesis (R01063) to refuel the Calvin cycle during photosynthesis. **C)** NADH production is
620 correlated with the growth rate and dominated by pyruvate dehydrogenase (R00209) during daytime and
621 the glycine cleavage system (R01221) during nighttime. NADH is consumed by 6-phosphogluconate
622 dehydrogenase (reverse, R10221) during daytime, NADH transhydrogenase (R00112) solely during dusk
623 and concomitant with methylenetetrahydrofolate reductase (R07168) during nighttime.
624

625 **Figure S5: Predicted growth curves show good agreement in a qualitative comparison with**
626 **experimental growth of *Synechocystis.*** To compare growth data we have overlaid growth curves
627 predicted for the wild-type, the *ΔglgC*-mutant and the *Δgnd*-mutant of *Prochlorococcus* with experimental
628 OD measurements of *Synechocystis.* We find a very good agreement for the wild-type and *ΔglgC*-mutant,
629 but not for the *Δgnd*-mutant. The lower panel shows how the model predicts the allocation and
630 consumption of each of the three strains.
631 **Table S1:** List of reactions added to iJC568 to form iSO595.
632 **Table S2:** List of reactions from iJC568 that are modified in iSO595.
633 **Table S3:** Parameter ranges used in the sampling of nutrient environments.
634 **Table S4:** List of blocked exchange reactions prior to sampling of nutrient environments.
635 **Table S5:** Parameter values used to run dynamic FBA in COMETS.
636 **Supplemental material 1:** Results from the BLAST-search used to identify 6PG-dehydratase (EC:
637 4.2.1.12) encoded by PMM0774 in *P. marinus* MED4.
638 **Supplemental material 2:** memote snapshot report of *i*SO595.
639
640

641 # Figure legends

642
643 **Figure 1:** iSO595 is an updated reconstruction of *P. marinus* MED4, featuring a complete Entner-Doudoroff
644 pathway, rewired glycogen metabolism and increased coverage of the genome. The central panel is a
645 simplified illustration covering the most relevant metabolic features. **A)** To simulate the natural
646 environmental constraints experience by Prochlorococcus we use "Push-FBA": Light and bicarbonate
647 uptake is given a fixed flux independent of growth rate. This contrasts standard FBA where light and
648 bicarbonate would be "pulled" in by the demand needed to support maximal growth rate up to a given

649 bound. **B)** We rewired the glycogen metabolism in iSO595 to study the dynamic allocation the dynamic
650 allocation of glycogen. **C)** Additionally, we implemented dynamic light conditions and light absorption in
651 COMETS to simulate the growth of *P. marinus* during the diel cycle. **D)** iSO595 has increased coverage
652 both in terms of genes, reactions and metabolites compared to its ancestor iJC568.

653
654 **Figure 2:** Schematic description of the ReFill algorithm. New reaction annotations are added only if all the
655 metabolites are connected to the existing network. A connected metabolite is denoted in a green circle.
656 An orphan (unconnected metabolite is denoted in a red circle) (A) A simple case in which all the
657 metabolites comply with the selection rule. (B) In case one or more metabolites are not connected to the
658 existing network, other new reactions may be added to complete the missing connections. For example,
659 consider reaction a to be composed of two substrates and two products, one of which, x (a metabolite
660 that did not exist in the initial network), is not connected to the existing network and is currently an
661 orphan. After the expansion step, the algorithm identified one reaction, b, in which x is used as a substrate.
662 All other metabolites in b exist in the network thus creating a path from reaction a through reaction b to
663 the network.

664
665 **Figure 3:** Histograms of environmental sampling results provide insight into how the fixed uptake rates of
666 light and bicarbonate and the upper bounds on ammonium and phosphate affects exudation of organic
667 acids, nucleobases and nucleotides, amino acids as well as glycogen storage. The y-axis represents the
668 fraction of all sampled environments yielding feasible models. Although we initially sample each
669 parameter uniformly the final sample distribution is not uniform because some combination of
670 parameters represents infeasible phenotypes (no solution can satisfy the constraints). The final sample
671 distribution for each parameter is therefore shown in the figure panels as a blue line. The black and red
672 line represents the histograms of samples where exudation is predicted by FBA and pFBA, respectively.

673 The shaded green region represents the span between histograms of samples as predicted by FVA: the
674 lower and upper bound represents the number of samples where exudation is predicted using the
675 minimum and maximum value of FVA, respectively. This FVA region covers the range of possible
676 phenotypes. The lower bound of the FVA region display the number of samples where a certain outcome
677 is obligatory to maximize growth, while the upper bound of the FVA region display the number of samples
678 where the outcome is possibly without reducing growth.

679 **Figure 4:** T-SNE clustering identifies typical phenotypes from the pFBA results from the random samples.
680 **A)** The random samples are reduced into two dimensions with t-SNE. We have subsequently used
681 HDBSCAN to identify 6 disjoint clusters which represents different phenotypes. **B)** For each of the 6
682 clusters the mean uptake or exudation across all samples within the respective cluster is shown. Only
683 exchange reactions with an absolute flux above 1e-3 mmol gDW^{-1} h^{-1} in any of the random samples are
684 included.

685
686 **Figure 5: Correlation between maximal FVA values** demonstrate which compounds that can be secreted
687 in similar environmental conditions. There are two clearly correlated groups of compounds: The first
688 group comprises nucleobases, nucleotides and amino acids while the second contains organic acids,
689 glycogen and glycerol 3-phosphate. From the correlation of each of these groups with the uptake bounds
690 on ammonium and phosphate we observe that these factors determine which of the groups that can be
691 secreted. Note that the environmental constraints have been converted to positive values prior to
692 calculating the correlation.

693
694 **Figure 6: Insight into metabolic rearrangements during the diel cycle**. **A)** Light irradiance, biomass and
695 glycogen storage throughout the diel cycle. We observe that the largest accumulated growth is found in
696 the nitrogen-abundant condition (green), but glycogen is only predicted to be stored in the nitrogen-poor

condition (orange). **B)** The flux distributions shifts when the metabolism switches from photosynthesis to glycogen catabolism, displayed by five reactions representing the Calvin Cycle (CBB), Glycogen metabolism, lower part of glycolysis, The Entner-Doudoroff (ED) and the Penthose Phosphate Pathway (PPP). **C)** iSO595 predicts that the depletion of glycogen is accompanied by exudation of fumarate and pyruvate.

References

Aharonovich, D., and Sher, D. (2016). Transcriptional response of Prochlorococcus to co-culture with a marine Alteromonas: Differences between strains and the involvement of putative infochemicals. *ISME J.* 10, 2892–2906. doi:10.1038/ismej.2016.70.

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). Basic local alignment search tool. *J Mol Biol* 215. doi:10.1016/S0022-2836(05)80360-2.

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between Diatoms and Bacteria. 76, 667–684. doi:10.1128/MMBR.00007-12.

Arrieta, J. M., Mayol, E., Hansman, R. L., and Herndl, G. J. (2015). Dilution limits dissolved organic carbon utilization in the deep ocean. 348, 331–334.

Becker, J. W., Berube, P. M., Follett, C. L., Waterbury, J. B., Chisholm, S. W., Delong, E. F., et al. (2014). Closely related phytoplankton species produce similar suites of dissolved organic matter. 5, 1–14. doi:10.3389/fmicb.2014.00111.

Benson, P. J., Purcell-Meyerink, D., Hocart, C. H., Truong, T. T., James, G. O., Rourke, L., et al. (2016). Factors Altering Pyruvate Excretion in a Glycogen Storage Mutant of the Cyanobacterium,

720 Synechococcus PCC7942. *Front. Microbiol.* 7, 475. doi:10.3389/fmicb.2016.00475.

721 Bertilsson, S., and Berglund, O. (2014). Release of dissolved organic matter by Prochlorococcus RELEASE
722 OF DISSOLVED ORGANIC MATTER BY.

723 Biller, S. J., Berube, P. M., Lindell, D., and Chisholm, S. W. (2015). Prochlorococcus: The structure and
724 function of collective diversity. *Nat. Rev. Microbiol.* 13, 13–27. doi:10.1038/nrmicro3378.

725 Biller, S. J., Coe, A., Roggensack, S. E., and Chisholm, W. (2018). Heterotroph Interactions Alter
726 Prochlorococcus Transcriptome. 3, 1–18.

727 Braakman, R. (2019). Evolution of cellular metabolism and the rise of a globally productive biosphere.
728 *Free Radic. Biol. Med.* 140, 172–187. doi:10.1016/j.freeradbiomed.2019.05.004.

729 Braakman, R., Follows, M. J., and Chisholm, S. W. (2017). Metabolic evolution and the self-organization
730 of ecosystems. *Proc. Natl. Acad. Sci.* 114, E3091–E3100. doi:10.1073/pnas.1619573114.

731 Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic
732 absorption in oceanic waters: Influence of the size structure of algal populations. *J. Geophys. Res.*
733 *Ocean.* 109, 1–12. doi:10.1029/2004JC002419.

734 Broddrick, J. T., Rubin, B. E., Welkie, D. G., Du, N., Mih, N., Diamond, S., et al. (2016). Unique attributes
735 of cyanobacterial metabolism revealed by improved genome-scale metabolic modeling and
736 essential gene analysis. *Proc. Natl. Acad. Sci.* 113, E8344–E8353. doi:10.1073/pnas.1613446113.

737 Burnap, R., Hagemann, M., and Kaplan, A. (2015). Regulation of CO2 Concentrating Mechanism in
738 Cyanobacteria. *Life* 5, 348–371. doi:10.3390/life5010348.

739 Campillo-Brocal, J. C., Lucas-Elió, P., and Sanchez-Amat, A. (2015). Distribution in different organisms of
740 amino acid oxidases with fad or a quinone as cofactor and their role as antimicrobial proteins in
741 marine bacteria. *Mar. Drugs* 13, 7403–7418. doi:10.3390/md13127073.

742 Cano, M., Holland, S. C., Artier, J., Burnap, R. L., Ghirardi, M., Morgan, J. A., et al. (2018). Glycogen
743 Synthesis and Metabolite Overflow Contribute to Energy Balancing in Cyanobacteria. *Cell Rep.* 23,

667–672. doi:10.1016/j.celrep.2018.03.083.

Carrieri, D., Paddock, T., Maness, P.-C., Seibert, M., and Yu, J. (2012). Photo-catalytic conversion of carbon dioxide to organic acids by a recombinant cyanobacterium incapable of glycogen storage. *Energy Environ. Sci.* 5, 9457. doi:10.1039/c2ee23181f.

Casey, J. R., Mardinoglu, A., Nielsen, J., and Karl, D. M. (2016). Adaptive Evolution of Phosphorus Metabolism in Prochlorococcus. *mSystems* 1, 1–15. doi:10.1128/mSystems.00065-16.Editor.

Chen, X., Schreiber, K., Appel, J., Makowka, A., Fähnrich, B., Roettger, M., et al. (2016). The Entner-Doudoroff pathway is an overlooked glycolytic route in cyanobacteria and plants. *Proc. Natl. Acad. Sci. U. S. A.* 113, 5441–5446. doi:10.1073/pnas.1521916113.

Cirri, E., and Pohnert, G. (2019). Algae−bacteria interactions that balance the planktonic microbiome. *New Phytol.* 223, 100–106. doi:10.1111/nph.15765.

Coles, V. J., Stukel, M. R., Brooks, M. T., Burd, A., Crump, B. C., Moran, M. A., et al. (2017). Ocean biogeochemistry modeled with emergent trait-based genomics. *Science (80-.).* 358, 1149–1154. doi:10.1126/science.aan5712.

Damrow, R., Maldener, I., and Zilliges, Y. (2016). The multiple functions of common microbial carbon polymers, glycogen and PHB, during stress responses in the non-diazotrophic cyanobacterium synechocystis sp. PCC 6803. *Front. Microbiol.* 7, 1–10. doi:10.3389/fmicb.2016.00966.

Davey, M., Tarran, G. A., Mills, M. M., Ridame, C., Geider, R. J., and LaRoche, J. (2008). Nutrient limitation of picophytoplankton photosynthesis and growth in the tropical North Atlantic. *Limnol. Oceanogr.* 53, 1722–1733. doi:10.4319/lo.2008.53.5.1722.

de Groot, D. H., Lischke, J., Muolo, R., Planqué, R., Bruggeman, F. J., and Teusink, B. (2020). The common message of constraint-based optimization approaches: overflow metabolism is caused by two growth-limiting constraints. *Cell. Mol. Life Sci.* 77, 441–453. doi:10.1007/s00018-019-03380-2.

Diamond, S., Jun, D., Rubin, B. E., and Golden, S. S. (2015). The circadian oscillator in Synechococcus

elongatus controls metabolite partitioning during diurnal growth. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1916–E1925. doi:10.1073/pnas.1504576112.

Díaz-Troya, S., López-Maury, L., Sánchez-Riego, A. M., Roldán, M., and Florencio, F. J. (2014). Redox regulation of glycogen biosynthesis in the cyanobacterium synechocystis sp. PCC 6803: Analysis of the AGP and glycogen synthases. *Mol. Plant* 7, 87–100. doi:10.1093/mp/sst137.

Domínguez-Martín, M. A., López-Lozano, A., Clavería-Gimeno, R., Velázquez-Campoy, A., Seidel, G., Burkovski, A., et al. (2018). Differential NtcA responsiveness to 2-oxoglutarate underlies the diversity of C/N balance regulation in Prochlorococcus. *Front. Microbiol.* 8, 1–16. doi:10.3389/fmicb.2017.02641.

Dubinsky, Z., and Berman-Frank, I. (2001). Uncoupling primary production from population growth in photosynthesizing organisms in aquatic ecosystems. in *Aquatic Sciences* (Birkhauser Verlag AG), 4–17. doi:10.1007/PL00001343.

Ebrahim, A., Lerman, J. A., Palsson, B. O., and Hyduke, D. R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* 7, 74. doi:10.1186/1752-0509-7-74.

Edwards, J. S., Ramakrishna, R., and Palsson, B. O. (2002). Characterizing the metabolic phenotype: A phenotype phase plane analysis. *Biotechnol. Bioeng.* 77, 27–36. doi:10.1002/bit.10047.

Elbourne, L. D. H., Tetu, S. G., Hassan, K. A., and Paulsen, I. T. (2017). TransportDB 2.0: A database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* 45, D320–D324. doi:10.1093/nar/gkw1068.

Feist, A. M., Nagarajan, H., Rotaru, A. E., Tremblay, P. L., Zhang, T., Nevin, K. P., et al. (2014). Constraint-Based Modeling of Carbon Fixation and the Energetics of Electron Transfer in Geobacter metallireducens. *PLoS Comput. Biol.* 10. doi:10.1371/journal.pcbi.1003575.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science (80-.).* 281, 237–240.

doi:10.1126/science.281.5374.237.

Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., et al. (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proc. Natl. Acad. Sci. U. S. A.* 110, 9824–9829. doi:10.1073/pnas.1307701110.

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W. (2007). Emergent Biogeography of Microbial. *Science (80-.).*, 1843–1847. doi:10.1126/science.1138544.

Forchhammer, K., and Schwarz, R. (2019). Nitrogen chlorosis in unicellular cyanobacteria – a developmental program for surviving nitrogen deprivation. *Environ. Microbiol.* 21, 1173–1184. doi:10.1111/1462-2920.14447.

Forchhammer, K., and Selim, K. A. (2019). Carbon/nitrogen homeostasis control in cyanobacteria. *FEMS Microbiol. Rev.* 44, 33–53. doi:10.1093/femsre/fuz025.

Foster, S. Q., Al-haj, A., Church, M. J., Dijken, G. L. Van, Dutkiewicz, S., Fulweiler, R. W., et al. (2018). Ecological control of nitrite in the upper ocean. *Nat. Commun.* doi:10.1038/s41467-018-03553-w.

Gilbert, J. D. J., and Fagan, W. F. (2011). Contrasting mechanisms of proteomic nitrogen thrift in Prochlorococcus. *Mol. Ecol.* 20, 92–104. doi:10.1111/j.1365-294X.2010.04914.x.

Gomez, J. A., Höffner, K., and Barton, P. I. (2014). DFBAlab: A fast and reliable MATLAB code for dynamic flux balance analysis. *BMC Bioinformatics* 15, 409. doi:10.1186/s12859-014-0409-8.

Grossowicz, M., Roth-Rosenberg, D., Aharonovich, D., Silverman, J., Follows, M. J., and Sher, D. (2017). *Prochlorococcus* in the lab and in silico: The importance of representing exudation. *Limnol. Oceanogr.* 62, 818–835. doi:10.1002/lno.10463.

Gu, C., Kim, G. B., Kim, W. J., Kim, H. U., and Lee, S. Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biol.* 20, 1–18. doi:10.1186/s13059-019-1730-3.

Gudmundsson, S., and Thiele, I. (2010). Computationally efficient flux variability analysis. *BMC Bioinformatics* 11, 489. doi:10.1186/1471-2105-11-489.

Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., et al. (2014). Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep.* 7, 1104–1115. doi:10.1016/j.celrep.2014.03.070.

Haug, K., Salek, R. M., Conesa, P., Hastings, J., De Matos, P., Rijnbeek, M., et al. (2013). MetaboLights - An open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.* 41, 781–786. doi:10.1093/nar/gks1004.

Hennon, G. M., Morris, J. J., Haley, S. T., Zinser, E. R., Durrant, A. R., Entwistle, E., et al. (2018). The impact of elevated CO 2 on Prochlorococcus and microbial interactions with â € helper' bacterium Alteromonas. *ISME J.* 12, 520–531. doi:10.1038/ismej.2017.189.

Henry, C. S., Dejongh, M., Best, A. A., Frybarger, P. M., Linsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi:10.1038/nbt.1672.

Holtzendorff, J., Partensky, F., Mella, D., Lennon, J.-F., Hess, W. R., and Garczarek, L. (2008). Genome Streamlining Results in Loss of Robustness of the Circadian Clock in the Marine Cyanobacterium *Prochlorococcus marinus* PCC 9511. *J. Biol. Rhythms* 23, 187–199. doi:10.1177/0748730408316040.

Hopkinson, B. M., Young, J. N., Tansik, A. L., and Binder, B. J. (2014). The minimal CO2-concentrating mechanism of prochlorococcus spp. MED4 is effective and efficient. *Plant Physiol.* 166, 2205–2217. doi:10.1104/pp.114.247049.

Iglesias, A. A., Kakefuda, G., and Preiss, J. (1991). Regulatory and structural properties of the cyanobacterial ADPglucose pyrophosphorylases. *Plant Physiol.* 97, 1187–1195. doi:10.1104/pp.97.3.1187.

Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., and Chisholm, S. W. (2006). Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. *Science (80-.).* 311, 1737–1740. doi:10.1126/science.1118052.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10592173 [Accessed October 23, 2017].

Kavvas, E. S., Seif, Y., Yurkovich, J. T., Norsigian, C., Poudel, S., Greenwald, W. W., et al. (2018). Updated and standardized genome-scale reconstruction of Mycobacterium tuberculosis H37Rv, iEK1011, simulates flux states indicative of physiological conditions. *BMC Syst. Biol.* 12. doi:10.1186/s12918-018-0557-y.

Kim, J., Fabris, M., Baart, G., Kim, M. K., Goossens, A., Vyverman, W., et al. (2016). Flux balance analysis of primary metabolism in the diatom *Phaeodactylum tricornutum*. *Plant J.* 85, 161–176. doi:10.1111/tpj.13081.

Knoop, H., Gründel, M., Zilliges, Y., Lehmann, R., Hoffmann, S., Lockau, W., et al. (2013). Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of Synechocystis sp. PCC 6803. *PLoS Comput. Biol.* 9. doi:10.1371/journal.pcbi.1003081.

Krumhardt, K. M., Callnan, K., Roache-Johnson, K., Swett, T., Robinson, D., Reistetter, E. N., et al. (2013). Effects of phosphorus starvation versus limitation on the marine cyanobacterium ProchlorococcusMED4 I: Uptake physiology. *Environ. Microbiol.* 15, 2114–2128. doi:10.1111/1462-2920.12079.

Kujawinski, E. B. (2011). The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. doi:10.1146/annurev-marine-120308-081003.

Laurenceau, R., Bliem, C., Osburne, M. S., Becker, J. W., Biller, S. J., Cubillos-Ruiz, A., et al. (2020). Toward a genetic system in the marine cyanobacterium Prochlorococcus. *Access Microbiol.* 2, e000107. doi:10.1099/acmi.0.000107.

Lewis, N. E., Hixson, K. K., Conrad, T. M., Lerman, J. A., Charusanti, P., Polpitiya, A. D., et al. (2010). Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale

models. *Mol. Syst. Biol.* 6, 390. doi:10.1038/msb.2010.47.

Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* 38, 272–276. doi:10.1038/s41587-020-0446-y.

Long, S. P., Humphries, S., and Falkowski, P. G. (1994). Photoinhibition of Photosynthesis in Nature. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 45, 633–662. doi:10.1146/annurev.pp.45.060194.003221.

Luan, G., Zhang, S., Wang, M., and Lu, X. (2019). Progress and perspective on cyanobacterial glycogen metabolism engineering. *Biotechnol. Adv.* 37, 771–786. doi:10.1016/j.biotechadv.2019.04.005.

Ma, L., Calfee, B. C., Morris, J. J., Johnson, Z. I., and Zinser, E. R. (2018a). Degradation of hydrogen peroxide at the ocean's surface: The influence of the microbial community on the realized thermal niche of Prochlorococcus. *ISME J.* 12, 473–484. doi:10.1038/ismej.2017.182.

Ma, X., Coleman, M. L., and Waldbauer, J. R. (2018b). Distinct molecular signatures in dissolved organic matter produced by viral lysis of marine cyanobacteria. *Environ. Microbiol.* 20, 3001–3011. doi:10.1111/1462-2920.14338.

Maarleveld, T. R., Khandelwal, R. A., Olivier, B. G., Teusink, B., and Bruggeman, F. J. (2013). Basic concepts and principles of stoichiometric modeling of metabolic networks. *Biotechnol. J.* 8, 997–1008. doi:10.1002/biot.201200291.

Maaten, L. van der, and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.

Mahadevan, R., Edwards, J. S., and Doyle, F. J. (2002). Dynamic Flux Balance Analysis of diauxic growth in Escherichia coli. *Biophys. J.* 83, 1331–1340. doi:10.1016/S0006-3495(02)73903-9.

Mary, I., Tu, C.-J., Grossman, A., and Vaulot, D. (2004). Effects of high light on transcripts of stress-associated genes for the cyanobacteria Synechocystis sp. PCC 6803 and Prochlorococcus MED4 and MIT9313. *Microbiology* 150, 1271–1281. doi:10.1099/mic.0.27014-0.

McInnes, L., Healy, J., and Astels, S. (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* 2, 205. doi:10.21105/joss.00205.

McKinney, W. (2010). pandas: a Foundational Python Library for Data Analysis and Statistics | R (Programming Language) | Database Index. Available at: https://www.scribd.com/document/71048089/pandas-a-Foundational-Python-Library-for-Data-Analysis-and-Statistics [Accessed April 29, 2020].

Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., et al. (2017). iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.* 35, 904–908. doi:10.1038/nbt.3956.

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., et al. (2013). Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* 6, 701–710. doi:10.1038/ngeo1765.

Moore, L. R., Coe, A., Zinser, E. R., Saito, M. A., Sullivan, M. B., Lindell, D., et al. (2007). Culturing the marine cyanobacterium Prochlorococcus. *Limnol. Oceanogr. Methods* 5, 353–362. doi:10.4319/lom.2007.5.353.

Moradi, N., Liu, B., Iversen, M., and Kuypers, M. M. (2018). A new mathematical model to explore microbial processes and their constraints in phytoplankton colonies and sinking marine aggregates. 1–10.

Moran, M. A., and Durham, B. P. (2019). Sulfur metabolites in the pelagic ocean. *Nat. Rev. Microbiol.* 17, 665–678. doi:10.1038/s41579-019-0250-1.

Moran, M. A., Kujawinski, E. B., Stubbins, A., Fatland, R., Aluwihare, L. I., Buchan, A., et al. (2016). Deciphering ocean carbon in a changing world. *Proc. Natl. Acad. Sci.* 113, 3143–3151. doi:10.1073/pnas.1514645113.

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Res. Part A, Oceanogr. Res. Pap.* 28, 1375–1393. doi:10.1016/0198-0149(81)90039-X.

Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The black queen hypothesis: Evolution of dependencies through adaptive gene loss. *MBio* 3. doi:10.1128/mBio.00036-12.

Nogales, J., Gudmundsson, S., Knight, E. M., Palsson, B. O., and Thiele, I. (2012). Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 2678–2683. doi:10.1073/pnas.1117907109.

Noreña-Caro, D., and Benton, M. G. (2018). Cyanobacteria as photoautotrophic biofactories of high-value chemicals. *J. CO2 Util.* 28, 335–366. doi:10.1016/j.jcou.2018.10.008.

O'Brien, E. J., Monk, J. M., and Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell* 161, 971–987. doi:10.1016/j.cell.2015.05.019.

Orth, J. D., Thiele, I., and Palsson, B. O. (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi:10.1038/nbt.1614.

Pacheco, A. R., Moel, M., and Segrè, D. (2018). Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. 1–42.

Park, J., and Choi, Y. (2017). Cofactor engineering in cyanobacteria to overcome imbalance between NADPH and NADH: A mini review. *Front. Chem. Sci. Eng.* 11, 66–71. doi:10.1007/s11705-016-1591-1.

Partensky, F., and Garczarek, L. (2010). *Prochlorococcus* : Advantages and Limits of Minimalism. *Ann. Rev. Mar. Sci.* 2, 305–331. doi:10.1146/annurev-marine-120308-081034.

Partensky, F., Hess, W. R., and Vaulot, D. (1999). Prochlorococcus, a Marine Photosynthetic Prokaryote of Global Significance. *Microbiol. Mol. Biol. Rev.* 63, 106–127. doi:10.1128/mmbr.63.1.106-127.1999.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. Available at: http://jmlr.org/papers/v12/pedregosa11a.html.

Reid, A. (2011). Incorporating Microbial Processes into Climate Change Models. 2–32.

Reimers, A.-M., Knoop, H., Bockmayr, A., and Steuer, R. (2017). Cellular trade-offs and optimal resource allocation during cyanobacterial diurnal growth. *Proc. Natl. Acad. Sci.*, 201617508. doi:10.1073/pnas.1617508114.

Ribalet, F., Swalwell, J., Clayton, S., Jiménez, V., Sudek, S., Lin, Y., et al. (2015). Light-driven synchrony of Prochlorococcus growth and mortality in the subtropical Pacific gyre. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8008–12. doi:10.1073/pnas.1424279112.

Roth-rosenberg, D., Aharonovich, D., Omta, A.-W., Follows, M. J., and Sciences, P. (2019). Dynamic macromolecular composition and high exudation rates in Prochlorococcus. *bioRxiv*, 828897. doi:10.1101/828897.

Saito, M. A., McIlvin, M. R., Moran, D. M., Goepfert, T. J., DiTullio, G. R., Post, A. F., et al. (2014). Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science (80-.).* 345, 1173–1177. doi:10.1126/science.1256450.

Seaver, L. C., and Imlay, J. A. (2001). Hydrogen peroxide fluxes and compartmentalization inside growing Escherichia coli. *J. Bacteriol.* 183, 7182–7189. doi:10.1128/JB.183.24.7182-7189.2001.

Sher, D., Thompson, J. W., Kashtan, N., Croal, L., and Chisholm, S. W. (2011). Response of Prochlorococcus ecotypes to co-culture with diverse marine bacteria. *ISME J.* 5, 1125–1132. doi:10.1038/ismej.2011.1.

Shinde, S., Zhang, X., Singapuri, S. P., Kalra, I., Liu, X., Morgan-Kiss, R. M., et al. (2020). Glycogen Metabolism Supports Photosynthesis Start through the Oxidative Pentose Phosphate Pathway in Cyanobacteria. *Plant Physiol.* 182, 507–517. doi:10.1104/pp.19.01184.

Sosa, O. A., Casey, J. R., and Karl, D. M. (2019). Methylphosphonate Oxidation in Prochlorococcus Strain MIT9301 Supports Phosphate Acquisition, Formate Excretion, and Carbon Assimilation into Purines. *Appl. Environ. Microbiol.* 85. doi:10.1128/AEM.00289-19.

Szul, M. J., Dearth, S. P., Campagna, S. R., and Zinser, E. R. (2019). Carbon Fate and Flux in Prochlorococcus under Nitrogen Limitation . *mSystems* 4. doi:10.1128/msystems.00254-18.

Thiele, I., Hyduke, D. R., Steeb, B., Fankam, G., Allen, D. K., Bazzani, S., et al. (2011). A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2. *BMC Syst. Biol.* 5. doi:10.1186/1752-0509-5-8.

Thiele, I., Vlassis, N., and Fleming, R. M. T. (2014). FASTGAPFILL: Efficient gap filling in metabolic networks. *Bioinformatics* 30, 2529–2531. doi:10.1093/bioinformatics/btu321.

Thornton, D. C. O. (2014). Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean. *Eur. J. Phycol.* 49, 20–46. doi:10.1080/09670262.2013.875596.

Van Mooy, B. A. S., Rocap, G., Fredricks, H. F., Evans, C. T., and Devol, A. H. (2006). Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8607–8612. doi:10.1073/pnas.0600540103.

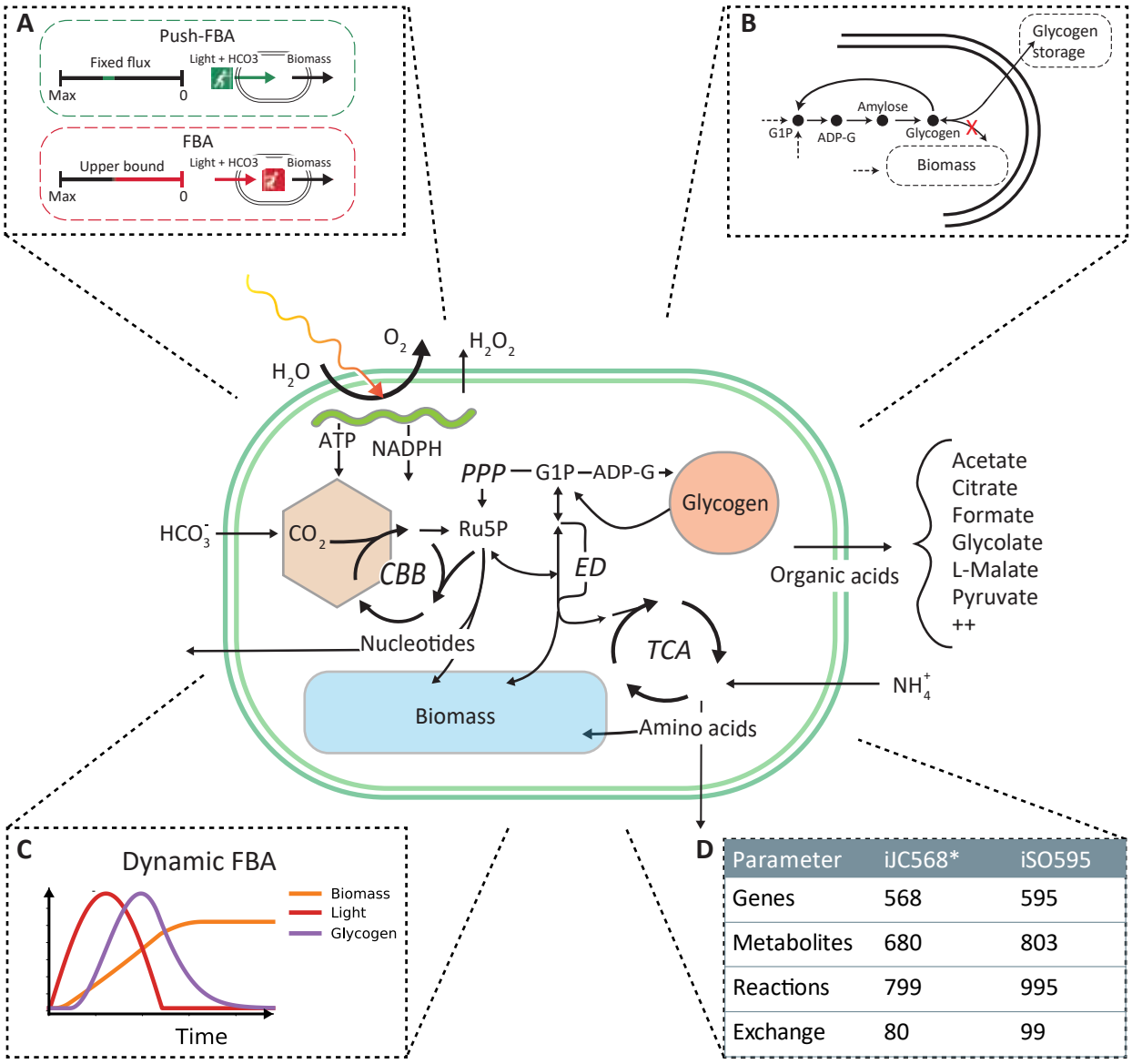
Varma, A., and Palsson, B. O. (1994). Metabolic flux balancing: Basic concepts, scientific and practical use. *Bio/Technology* 12, 994–998. doi:10.1038/nbt1094-994.

Vaulot, D., Marie, D., Olson, R. J., and Chisholm, S. W. (1995). Growth of Prochlorococcus, a photosynthetic prokaryote, in the equatorial Pacific Ocean. *Science (80-.).* 268, 1480–1482. doi:10.1126/science.268.5216.1480.

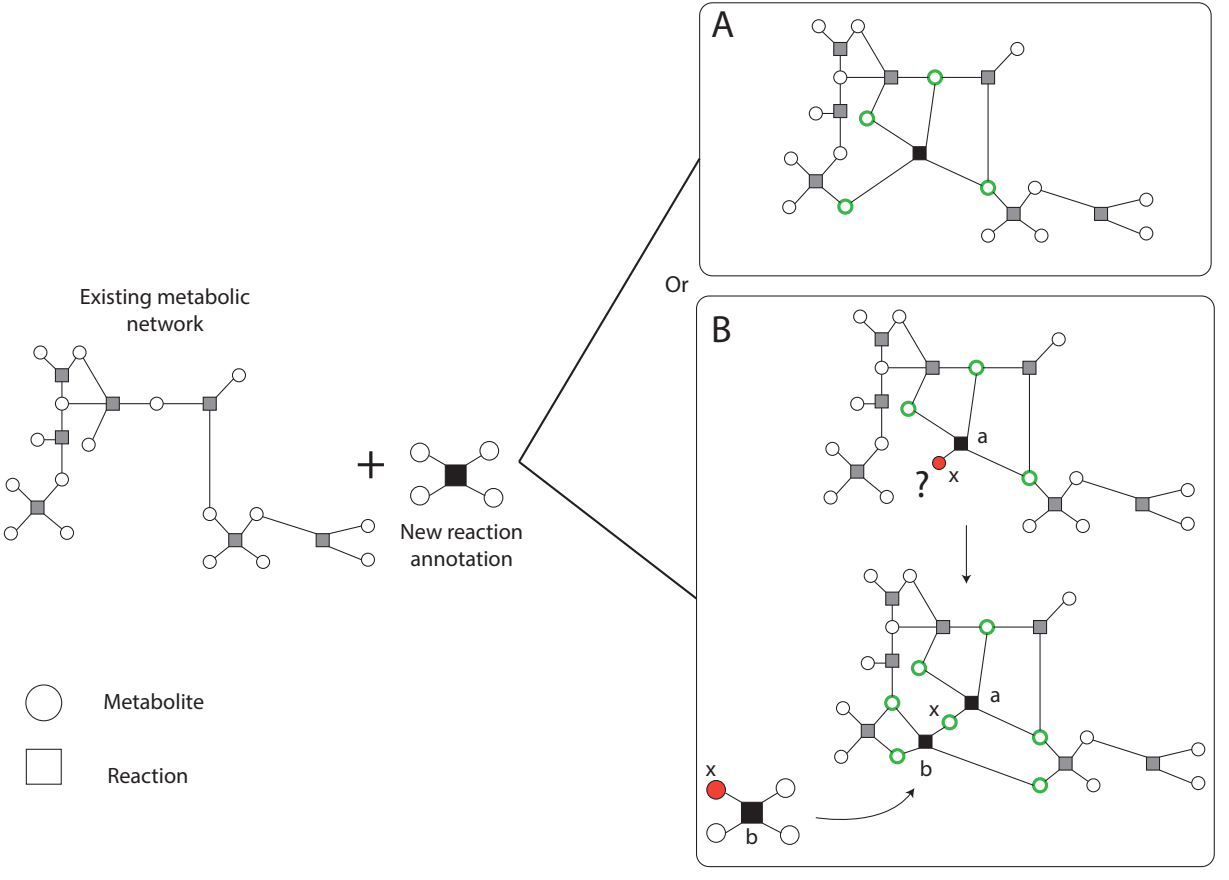
Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17, 261–272. doi:10.1038/s41592-019-0686-2.

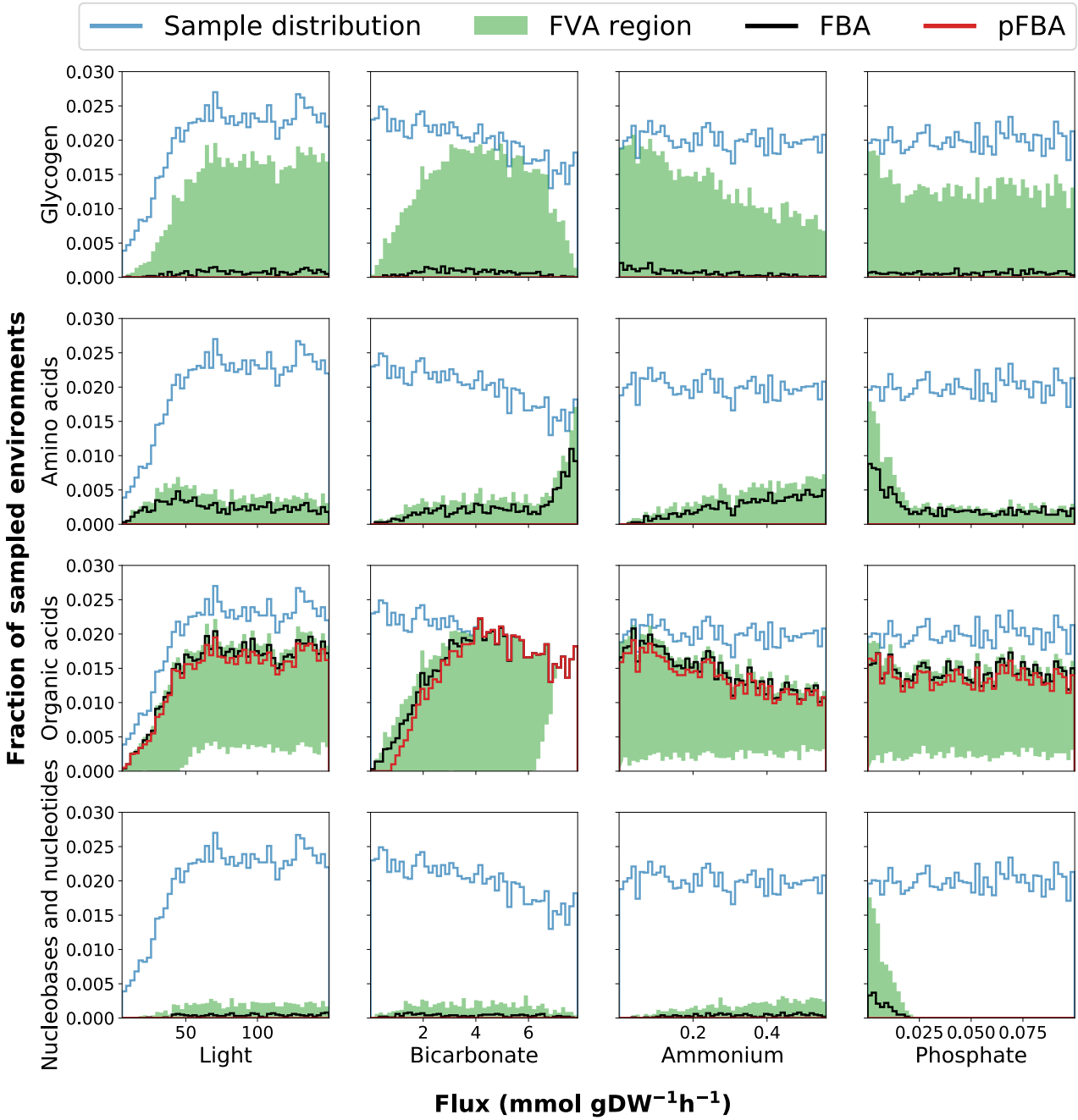
Waldbauer, J. R., Rodrigue, S., Coleman, M. L., and Chisholm, S. W. (2012). Transcriptome and Proteome Dynamics of a Light-Dark Synchronized Bacterial Cell Cycle. *PLoS One* 7, e43432. doi:10.1371/journal.pone.0043432.

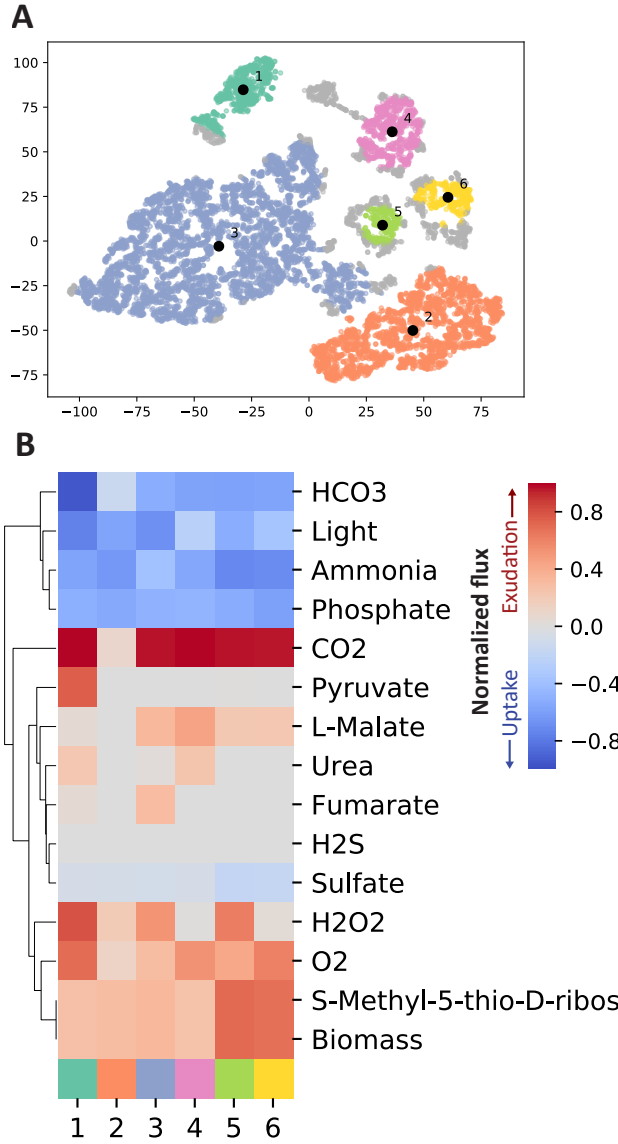
984 Welkie, D. G., Rubin, B. E., Diamond, S., Hood, R. D., Savage, D. F., and Golden, S. S. (2019). A Hard Day ’
985 s Night : Cyanobacteria in Diel Cycles. 27, 231–242.
986 Xiong, W., Cano, M., Wang, B., Douchi, D., and Yu, J. (2017). The plasticity of cyanobacterial carbon
987 metabolism. *Curr. Opin. Chem. Biol.* 41, 12–19. doi:10.1016/j.cbpa.2017.09.004.
988 Yang, A. (2011). Modeling and evaluation of CO 2 supply and utilization in algal ponds. *Ind. Eng. Chem.*
989 *Res.* 50, 11181–11192. doi:10.1021/ie200723w.
990 Yoshikawa, K., Toya, Y., and Shimizu, H. (2017). Metabolic engineering of Synechocystis sp. PCC 6803 for
991 enhanced ethanol production based on flux balance analysis. *Bioprocess Biosyst. Eng.* 40, 791–796.
992 doi:10.1007/s00449-017-1744-8.
993 Zhang, C. C., Zhou, C. Z., Burnap, R. L., and Peng, L. (2018). Carbon/Nitrogen Metabolic Balance: Lessons
994 from Cyanobacteria. *Trends Plant Sci.* 23, 1116–1130. doi:10.1016/j.tplants.2018.09.008.
995 Zinser, E. R., Lindell, D., Johnson, Z. I., Futschik, M. E., Steglich, C., Coleman, M. L., et al. (2009).
996 Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph,
997 Prochlorococcus. *PLoS One* 4. doi:10.1371/journal.pone.0005135.
998
999

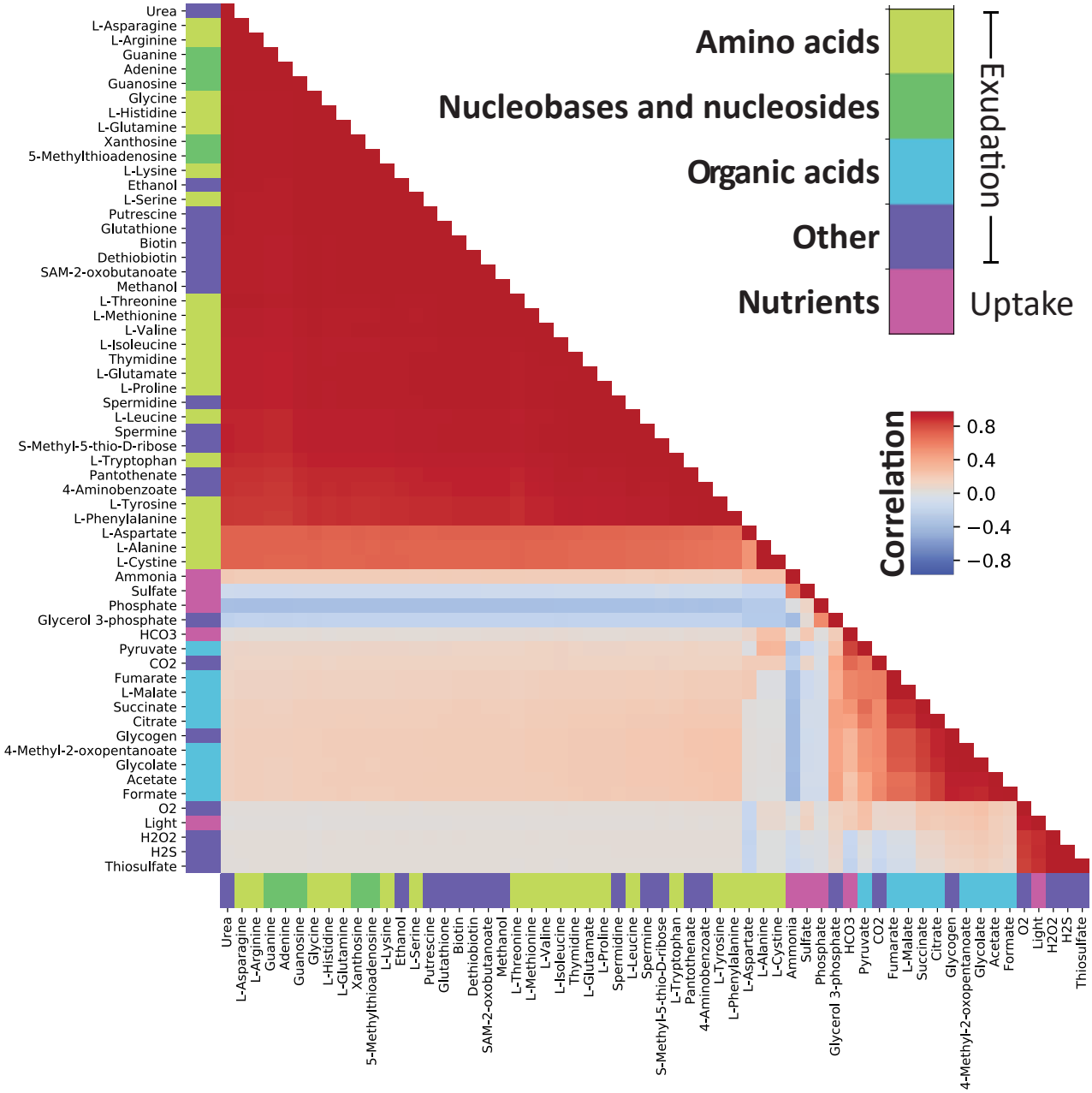


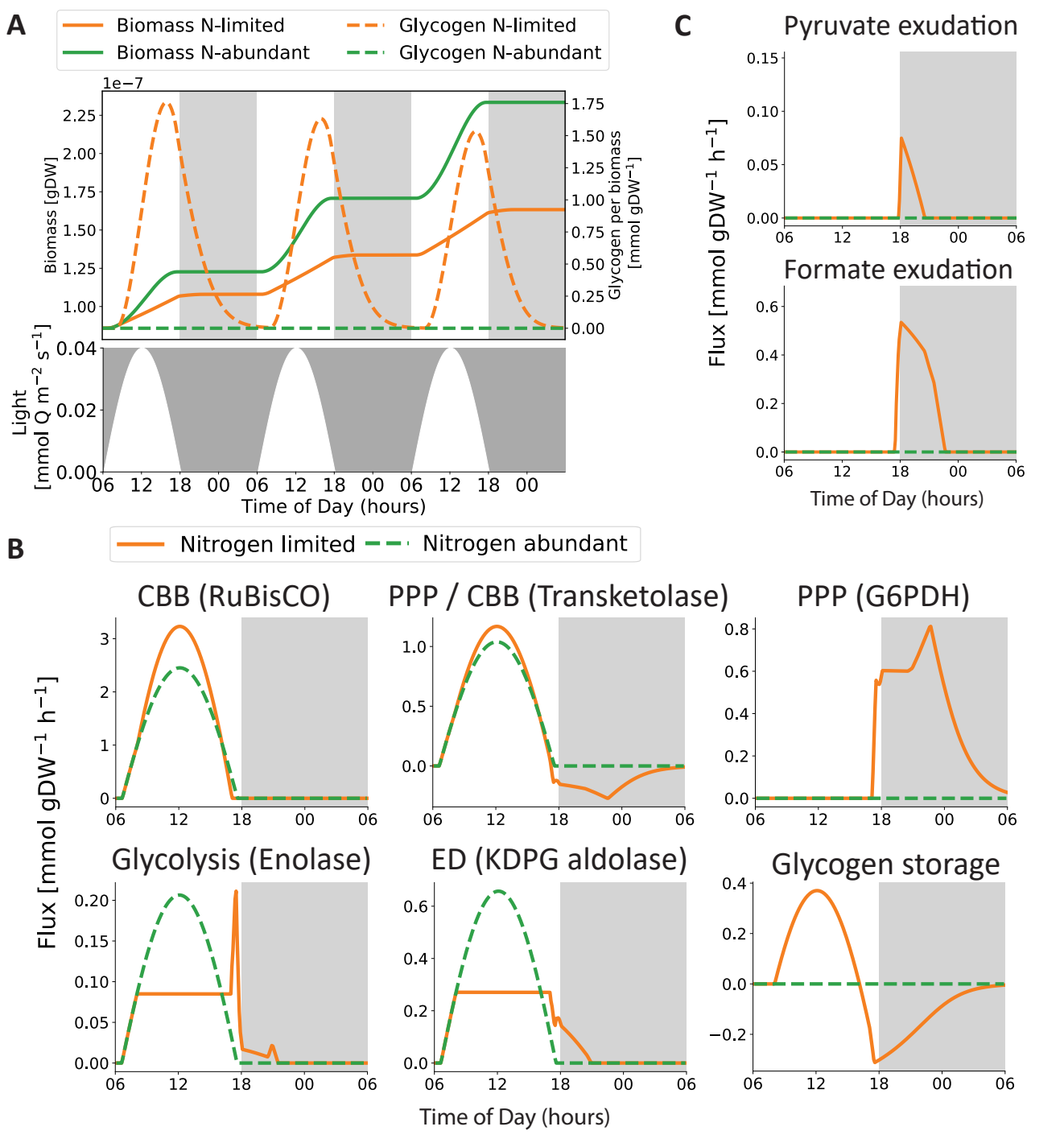
Parameter	iJC568*	iSO595
Genes	568	595
Metabolites	680	803
Reactions	799	995
Exchange	80	99


Existing metabolic network
+
New reaction annotation
Metabolite
Reaction
Or
A
B
a
x
?
x
a
b
x
b


Sample distribution
FVA region
FBA
pFBA
Fraction of sampled environments
Glycogen
Amino acids
Organic acids
Nucleobases and nucleotides
0.030
0.025
0.020
0.015
0.010
0.005
0.000
50
100
Light
2
4
6
Bicarbonate
0.2
0.4
Ammonium
0.025
0.050
0.075
Phosphate
Flux (mmol gDW^{-1}h^{-1})


A
100
75
50
25
0
−25
−50
−75
−100
−75
−50
−25
0
25
50
75
1
2
3
4
5
6
B
HCO3
Light
Ammonia
Phosphate
CO2
Pyruvate
L-Malate
Urea
Fumarate
H2S
Sulfate
H2O2
O2
S-Methyl-5-thio-D-ribose
Biomass
1
2
3
4
5
6
Normalized flux
Exudation
Uptake
0.8
0.4
0.0
−0.4
−0.8


Amino acids
Nucleobases and nucleosides
Organic acids
Other
Nutrients
Exudation
Uptake
Correlation
0.8
0.4
0.0
−0.4
−0.8
Urea
L-Asparagine
L-Arginine
Guanine
Adenine
Guanosine
Glycine
L-Histidine
L-Glutamine
Xanthosine
5-Methylthioadenosine
L-Lysine
Ethanol
L-Serine
Putrescine
Glutathione
Biotin
Dethiobiotin
SAM-2-oxobutanoate
Methanol
L-Threonine
L-Methionine
L-Valine
L-Isoleucine
Thymidine
L-Glutamate
L-Proline
Spermidine
L-Leucine
Spermine
S-Methyl-5-thio-D-ribose
L-Tryptophan
Pantothenate
4-Aminobenzoate
L-Tyrosine
L-Phenylalanine
L-Aspartate
L-Alanine
L-Cystine
Ammonia
Sulfate
Phosphate
Glycerol 3-phosphate
HCO3
Pyruvate
CO2
Fumarate
L-Malate
Succinate
Citrate
Glycogen
4-Methyl-2-oxopentanoate
Glycolate
Acetate
Formate
O2
Light
H2O2
H2S
Thiosulfate


A
Biomass N-limited
Biomass N-abundant
Glycogen N-limited
Glycogen N-abundant
1e−7
Biomass [gDW]
Glycogen per biomass [mmol gDW^{-1}]
Light [mmol Q m^{-2} s^{-1}]
Time of Day (hours)
C
Pyruvate exudation
Formate exudation
Flux [mmol gDW^{-1} h^{-1}]
Time of Day (hours)
B
Nitrogen limited
Nitrogen abundant
CBB (RuBisCO)
PPP / CBB (Transketolase)
PPP (G6PDH)
Glycolysis (Enolase)
ED (KDPG aldolase)
Glycogen storage
Flux [mmol gDW^{-1} h^{-1}]
Time of Day (hours)