environmental microbiology

Environmental Microbiology (2016) **18**(12), 4946–4960 doi:10.1111/1462-2920.13444

Stoichiometric modelling of assimilatory and dissimilatory biomass utilisation in a microbial community

Kristopher A. Hunt,[1,2,3] Ryan deM. Jennings,[3,4†] William P. Inskeep[3,4] and Ross P. Carlson[1,2,3*]

[1]*Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA.*

[2]*Department of Chemical and Biological Engineering, Montana State University, Bozeman, MT, USA.*

[3]*Thermal Biology Institute, Montana State University, Bozeman, MT, USA.*

[4]*Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA.*

Summary

Assimilatory and dissimilatory utilisation of autotroph biomass by heterotrophs is a fundamental mechanism for the transfer of nutrients and energy across trophic levels. Metagenome data from a tractable, thermoacidophilic microbial community in Yellowstone National Park was used to build an *in silico* model to study heterotrophic utilisation of autotroph biomass using elementary flux mode analysis and flux balance analysis. Assimilatory and dissimilatory biomass utilisation was investigated using 29 forms of biomass-derived dissolved organic carbon (DOC) including individual monomer pools, individual macromolecular pools and aggregate biomass. The simulations identified ecologically competitive strategies for utilizing DOC under conditions of varying electron donor, electron acceptor or enzyme limitation. The simulated growth environment affected which form of DOC was the most competitive use of nutrients; for instance, oxygen limitation favoured utilisation of less reduced and fermentable DOC while carbon-limited environments favoured more reduced DOC. Additionally, metabolism was studied considering two encompassing metabolic strategies: simultaneous versus sequential use of DOC. Results of this study bound the transfer of nutrients and energy through microbial food webs, providing a quantitative foundation relevant to most microbial ecosystems.

Received 3 April, 2016; accepted 30 June, 2016. *For correspondence. E-mail rossc@montana.edu. Tel: (406) 994-3631; Fax: (406) 994-5308. †Present address: Department of Biology, Mercer University, Macon, GA, USA.

Introduction

Nutrient and energy transfers across trophic levels are essential in nearly all environmental, industrial and medical microbial ecosystems. Primary producers obtain nutrients, including carbon, often via autotrophy using energy from chemolithotrophy or phototrophy. Heterotrophic utilisation (e.g., predation, decomposition and metabolite exchange) of resources from primary production requires nutrient assimilation and dissimilation, which contributes directly to biogeochemical cycling. Theoretical and experimental analysis of resource transfer between trophic levels is challenging due to the diversity of biomass constituents (i.e., macromolecules, monomers, cofactors, free metabolites and minerals), the complexity of biochemical networks and the number of intercellular metabolite exchanges that occur within microbial communities. Generalized kinetic models have been used for decades to analyze processes such as wastewater treatment (Grady *et al.*, 2011), composting (Mason, 2006) and carbon flow in sediments (Arndt *et al.*, 2013). However, these analyses typically use generic, instead of organism specific, yields or rates because direct measurements for each participating population of microorganisms and substrates are challenging to obtain.

Most naturally occurring, mesophilic ecosystems are complex assemblages of archaea, bacteria, eukaryotes and a multitude of potential nutrients and energy sources. For instance, a gram of soil has an estimated 10^3-10^6 species representing all three domains of life and numerous trophic levels from primary producers to decomposers (Dance, 2008). These ecosystems are too complex to represent using current molecular-level modelling methods without major simplifications. Relatively simple microbial mat communities, like those found in high-temperature environments of Yellowstone National Park (YNP), are tractable systems for examining assimilatory and dissimilatory biomass utilisation. High-temperature acidic, iron-oxidizing microbial mats found in YNP have been studied extensively (Kozubal *et al.*, 2012; Inskeep *et al.*, 2013;

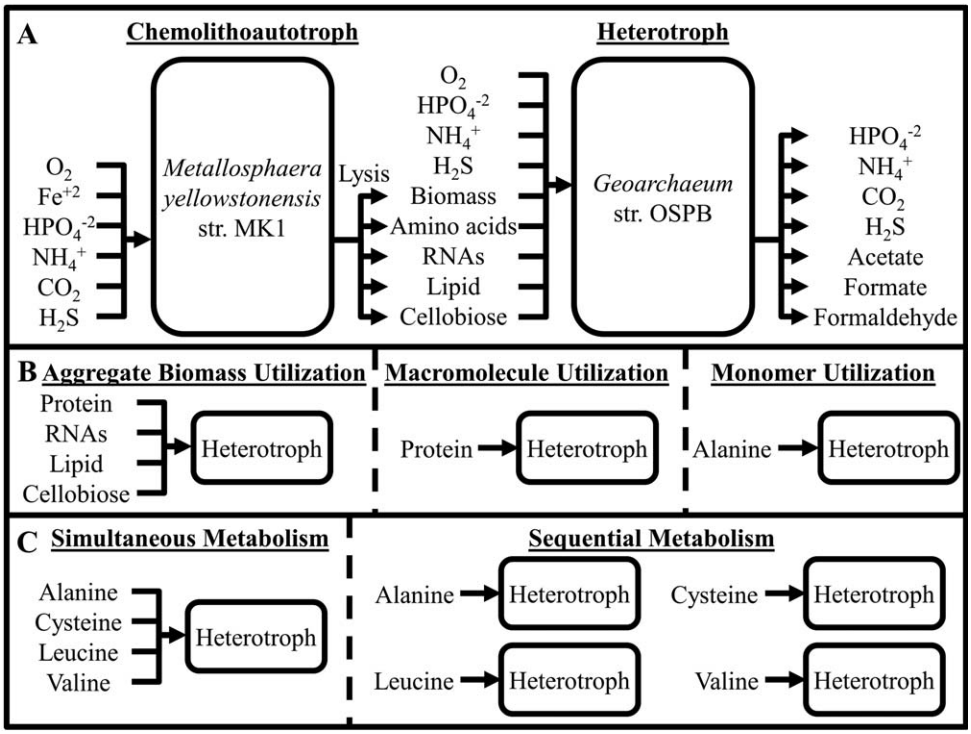

Fig. 1. Conceptual representation of high-temperature acidic Fe(III)-oxide microbial mats.
Primary productivity was represented by the chemolithoautotroph, *Metallosphaera yellowstonensis* str. MK1, which fixes carbon dioxide through iron oxidation and provides potential carbon and energy sources for community heterotrophs. Community heterotrophy was represented by *Geoarchaeum* str. OSPB, a numerically dominant heterotroph in the mats. Exchanges of nutrients and energy across trophic levels were analyzed considering 29 different biomass-derived dissolved organic carbon (DOC) (B) ranging from individual monomers to macromolecules to aggregate autotroph biomass. Simulations also explored two distinct heterotroph strategies: simultaneous versus sequential metabolism of DOC pools (C).

Jennings *et al*., 2014) and provide ideal model systems for examining nutrient and energy transfer across trophic levels (i.e., assimilatory and dissimilatory biomass utilisation Fig. 1A). High temperature (65–75°C) and low pH (3–3.5) contribute to the simplicity of these microbial communities by precluding growth of photosynthetic organisms and eukaryotes (Breitbart *et al*., 2004; Boyd *et al*., 2009; Rohwer *et al*., 2009).

A well-studied system in One Hundred Springs Plain, Norris Geyser Basin YNP, contains 5–7 dominant microbial populations, of which 2–3 are metabolically active in the oxic region based on metagenomic analyses (Kozubal *et al*., 2012; Inskeep *et al*., 2013; Jennings *et al*., 2014). *Metallosphaera yellowstonensis* str. MK1 is a dominant primary producer in these mats, and has been shown to fix inorganic carbon (DIC) during aerobic oxidation of ferrous iron (Kozubal *et al*., 2011; Jennings *et al*., 2014). The numerically dominant heterotroph (30–50% of the total aerobic community) in these mats is an archaeon from the newly described group of Geoarchaeota, referred to here as *Geoarchaeum* str. OSPB (Kozubal *et al*., 2012; 2013). This organism functions as an aerobic chemoorganoheterotroph based on metabolic analysis of metagenome assemblies (Kozubal *et al*., 2013). Viral lysis and senescence are two candidate mechanisms for the release of resources derived from chemolithoautotroph biomass utilised by heterotrophic organisms including *Geoarchaeum* str. OSPB (Breitbart *et al*., 2004; Rohwer *et al*., 2009). *M. yellowstonensis* str. MK1 and *Geoarchaeum* str. OSPB both require oxygen as a terminal electron acceptor for growth and cellular energy production, and *in situ* analysis of oxygen as a function of mat depth indicated that oxygen diffusion, not reaction rate, is rate-limiting for aerobic growth (Bernstein *et al*., 2013). Therefore, the autotroph and heterotroph may compete for this limiting resource.

Metagenome-enabled stoichiometric modelling can examine molecular-level phenomena without requiring extensive kinetic parameters. Stoichiometric modelling of metabolic systems identifies all possible genome-encoded physiologies, including ecologically relevant optimal and suboptimal phenotypes for nutrient utilisation (Klamt and Stelling, 2003; Reed and Palsson, 2003; Trinh *et al*., 2009; Orth *et al*., 2010). There are two widely applied methods of stoichiometric modelling. Flux balance analysis (FBA) is an optimisation approach that uses a metabolic model to determine possible routes of nutrient flow that, for

example, maximize biomass production assuming a substrate utilisation rate (Varma *et al.*, 1993). Elementary flux mode analysis (EFMA) is an unbiased stoichiometric modelling method that enumerates the smallest set of genotypes that describe nutrient flow through a metabolic model (elementary flux modes, EFMs) (see Supporting Information Fig. S1 for a detailed description). Nonnegative linear combinations of EFMs can represent all feasible phenotypes of the modelled organism (Schilling *et al.*, 2000; Klamt and Stelling, 2003; Llaneras and Picó, 2010). EFMA has enabled both strain development and community analysis without prior knowledge of the function of an organism (Carlson *et al.*, 2002; Taffs *et al.*, 2009).

Stoichiometric modelling has been used to study microbial community ecology (Stolyar *et al.*, 2007; Taffs *et al.*, 2009; Zhuang *et al.*, 2011; Cerqueda-García *et al.*, 2014); however, an assessment of assimilatory and dissimilatory uses of biomass-derived, dissolved organic carbon (DOC) as mechanisms of nutrient and energy transfer across trophic levels has not been conducted. This study examined biomass and cellular energy production in a tractable microbial mat community from YNP. Study aims were to (1) identify the optimal biomass-derived DOC for biomass and cellular energy production by heterotroph *Geoarchaeum* str. OSPB as a function of different environmental scenarios (i.e., DOC-, oxygen- and enzyme-limitation), (2) develop a relationship between microbial community member abundance based on intertrophic nutrient transfer, (3) contrast simultaneous and sequential metabolism of biomass-derived DOC for heterotroph biomass and cellular energy production and (4) predict distribution of byproducts during metabolism of biomass-derived DOC.

Results

Construction of in silico *model and application of ecological theory*

An *in silico* stoichiometric model was built to dissect potential autotroph-heterotroph interactions based on utilisation of autotroph biomass components. Inorganic carbon fixed by *M. yellowstonensis* str. MK1 (hereafter termed autotroph) was modelled as the source of reduced carbon for the microbial community (Fig. 1A). Community heterotrophy was represented by *Geoarchaeum* str. OSPB (hereafter termed heterotroph) utilisation of autotroph biomass components (Jennings *et al.*, 2014). Heterotroph utilisation of autotroph biomass was simulated using 29 forms of biomass-derived DOC ranging from individual monomer pools (e.g., alanine), macromolecular pools (e.g., protein), to aggregate biomass (Fig. 1B and Table 1). The composition of monomers, macromolecules and aggregate biomass were based on genetic attributes of the autotroph and assumptions detailed in the Materials and Methods. The geothermal source water was assumed to contain nonlimiting quantities of ammonium, phosphate, sulfide and all micronutrients. All other nutrient exchanges were modelled as depicted in Fig. 1A. The metabolic modell with metagenome justification and data verifying atom and electron conservation can be found in the Supporting Information.

The metabolic model was analyzed using EFMA and FBA. Biologically relevant flux distributions, represented as EFMs or FBA optimisation solutions, were identified using ecological resource allocation theory, which assumes an organism will maximize the utility of a limiting resource (Taffs *et al.*, 2009; Beck *et al.*, 2016). The theory was applied to the complete set of EFMs by calculating the resource costs of each EFM as defined by the amount of reduced carbon source (biomass-derived DOC), electron acceptor (oxygen) or enzyme resource (e.g., nitrogen, sulfur, iron or cell volume) required to produce either biomass or cellular energy (phosphodiester bond equivalents). Figure 2 provides a graphical representation of the analysis concepts using aggregate biomass as the example DOC. The enzyme resource cost for biomass and cellular energy production (Fig. 2C and D, respectively) was modelled by summing the number of participating reactions in each EFM, which is proportional to the number of participating enzymes and therefore, the amount of enzyme resource required to synthesize these enzymes (Folsom and Carlson, 2015). This approach approximates all enzyme complexes to be the same size with the same amino acid distribution.

For every EFM, pairs of resource costs for biomass or cellular energy production were plotted to define tradeoff curves or 'pareto fronts' (Fig. 2) (Beck *et al.*, 2016), which identified the optimal phenotype for any position along the resource limitation gradient. The EFMs whose resource costs were closest to the plot axes and their non-negative linear combinations (e.g., solid black lines in Fig. 2) minimize resource costs, making them ecologically competitive (Carlson and Srienc, 2004). EFMs not on the curve represent suboptimal strategies for the simulated environment (x symbols in Fig. 2). Abundance of suboptimal EFMs illustrates the metabolic robustness of the reaction network, and EFMs near the tradeoff curve may have roles in metabolic resilience, such as responding to resource or genetic perturbations (circles in Fig. 2). Biological implications of these specific suboptimal EFMs are discussed below.

Optimal utilisation of macromolecules

Approximately 50 million EFMs were enumerated during the analysis of the 29 forms of DOC (Table 1). Every EFM was analyzed for the resource cost to produce biomass or cellular energy on gradients of DOC, oxygen or enzyme resource limitation (Figs. 3–5). Only the optimal tradeoff surfaces are presented for each form of DOC for clarity; the total number of EFMs for each form of DOC and common associated byproducts are listed in Table 1. Tradeoff

Table 1. Summary of simulations for 29 forms of biomass-derived DOC.

Carbon source	Number of EFMs	Chemical formula	Degree of reduction[a]	Nutrients consumed[b]	Byproducts
Biomass	14 994 531	$CH_{1.6}N_{0.2}O_{0.4}P_{0.02}S_{0.005}$	4.27 (4.92)		Acetate, Formate, NH_4^+, HPO_4^{-2}, H_2S, H_2O_2
Cellobiose	794 557	$C_{12}H_{22}O_{11}$	4.00 (4.00)	HPO_4^{-2}, NH_4^+, H_2S	Acetate, Formate, Formaldehyde, H_2O_2
Archaeal Lipid	216 476	$C_{86}H_{170}O_{12}P_2$	5.86 (5.86)	H_2S,NH_4^+	Acetate, Formate, Formaldehyde, HPO_4^{-2}, H_2O_2
Nucleotides					
RNA	8 332 169	$CHN_{0.4}O_{0.8}P_{0.1}$	2.91 (4.07)	H_2S	Acetate, Formate, NH_4^+, HPO_4^{-2}, H_2O_2
AMP	438 345	$C_{10}H_{12}N_5O_7P$	3.00 (4.50)		
CMP	137 746	$C_9H_{12}N_3O_8P$	3.33 (4.33)		
GMP	433 861	$C_{10}H_{12}N_5O_8P$	2.80 (4.30)		
UMP	138 091	$C_9H_{11}N_2O_9P$	3.33 (4.00)		
Amino Acids					
Protein	303 907	$CH_{1.6}N_{0.3}O_{0.3}S_{0.006}$	4.21 (5.01)	HPO_4^{-2}	Acetate, Formate, Formaldehyde, H_2O_2, NH_4^+, H_2S
Ala	85 822	$C_3H_7NO_2$	4.00 (5.00)		
Arg	8 721	$C_6H_{15}N_4O_2$	3.67 (5.67)		
Asn	116 017	$C_4H_8N_2O_3$	3.00 (4.50)		
Asp	92 836	$C_4H_6NO_4$	3.00 (3.75)		
Cys	117 743	$C_3H_7NO_2S$	3.33 (4.33)		
Glu	12 408	$C_5H_8NO_4$	3.60 (4.20)		
Gln	14 202	$C_5H_{10}N_2O_3$	3.60 (4.80)		
Gly	100 664	$C_2H_5NO_2$	3.00 (4.50)		
His	16 487	$C_9H_6N_3O_2$	3.33 (4.83)		
Ile	79 510	$C_6H_{13}NO_2$	5.00 (5.50)		
Leu	88 657	$C_6H_{13}NO_2$	5.00 (5.50)		
Lys	38 383	$C_6H_{15}N_2O_2$	4.67 (5.67)		
Met	118 907	$C_5H_{11}NO_2$	4.40 (5.00)		
Phe	55 714	$C_9H_{11}NO_2$	4.44 (4.78)		
Pro	15 419	$C_5H_9NO_2$	4.40 (5.00)		
Ser	104 376	$C_3H_7NO_3$	3.33 (4.33)		
Thr	60 396	$C_4H_9NO_3$	4.00 (4.75)		
Trp	165 116	$C_{11}H_{12}N_2O_2$	4.18 (4.73)		
Tyr	56 065	$C_9H_{11}NO_3$	4.22 (4.56)		
Val	30 653	$C_5H_{11}NO_2$	4.80 (5.40)		

Biomass production by heterotroph *Geoarchaeum* str. OSPB was evaluated using DOC ranging from monomers, to macromolecules, to aggregate biomass derived from the autotroph *M. yellowstonensis* str. MK1.
[a] Degree of reduction was calculated on $NH_3(N_2)$ bases.
[b] Nutrients consumed during biomass production.

curves generated using either EFMA or FBA were generally identical (Supporting Information).

Lipid required the fewest Cmoles of DOC to produce a Cmole of heterotroph biomass or mole of cellular energy and was therefore the best form of DOC for carbon-limited environments relative to the other 28 forms of DOC (moving to left on *x*-axis in Fig. 3A and B). The lower DOC resource cost for lipid was due primarily to the high degree of reduction (six electrons per Cmole) (Table 1). After lipid, the most competitive macromolecular DOC to produce a Cmole of heterotroph biomass was aggregate biomass followed by protein, cellobiose and lastly, RNA. The DOC produced biomass or cellular energy at varying resource costs due to (1) the degree of reduction of the DOC, which is a gross measure of energy content, (2) chemical structure of the DOC, which determines the entry point into central metabolism relative to substrate-level phosphorylation reactions, (3) the elemental stoichiometry of the DOC relative to heterotroph biomass and/or (4) the biological accessibility of energy and nutrients in the DOC, which is related to the number of enzymes required to process the DOC. The ratio of Cmole DOC utilized per Cmole biomass produced quantifies the fraction of autotroph-derived carbon assimilated into heterotroph biomass with the balance dissimilated for cellular energy. This ratio establishes an important conceptual constraint between different trophic levels; at least 2.4 Cmoles of aggregate autotroph biomass are required to produce 1 Cmole of heterotroph biomass.

RNA required the fewest moles of oxygen per Cmole biomass or mole cellular energy produced among the

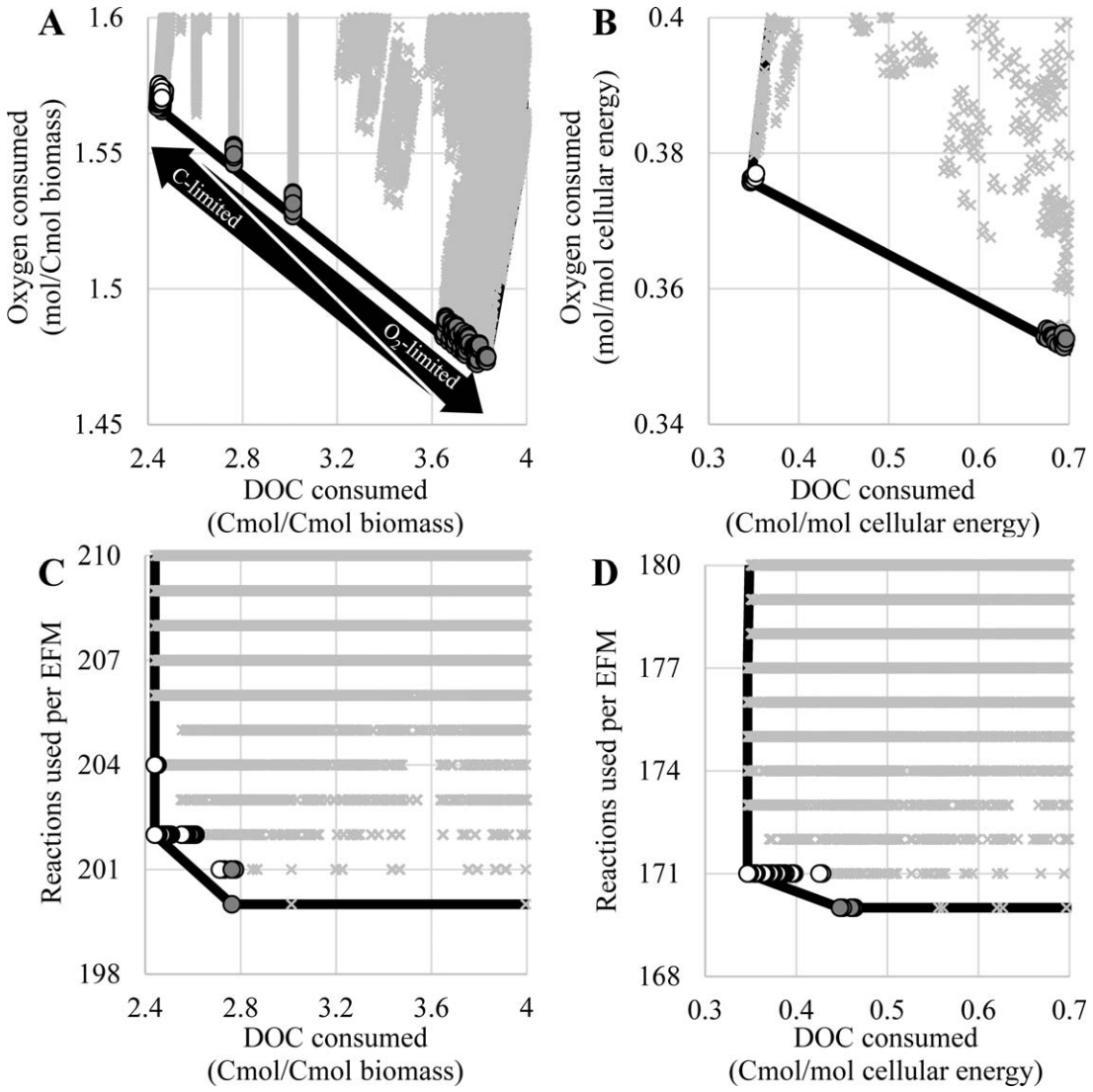

Fig. 2. Summary of ecological resource utilisation theory used for quantifying competitive physiologies during utilisation of autotroph biomass by the heterotroph, *Geoarchaeum* str. OSPB.
Resource costs (biomass-derived DOC, oxygen and enzymes) were calculated for every elementary flux mode (EFM) that synthesized biomass or cellular energy. Biomass and cellular energy production costs are plotted for the carbon- versus oxygen-limited scenarios (A) and (B), respectively. Biomass and cellular energy production costs are plotted for the carbon- versus enzyme-limited scenarios (C) and (D), respectively. Linear combinations of the most efficient EFMs define an optimal tradeoff surface (line). EFMs plotted with circles represent suboptimal ecologically competitive strategies that are phenotypically indistinguishable from the optimal strategies. These strategies ranged from complete oxidation of DOC to acetate production (white and grey circles, respectively) in (A), (B) and (D), as well as energy production (grey circles) in (C). EFMs plotted with an x are suboptimal and phenotypically distinct from the optimal strategies. Reactions used per EFM are the number of reactions with a nonzero flux in each EFM, which is a proxy for the enzyme resource cost.

macromolecular DOC for simulated oxygen-limited conditions (moving down *y*-axis in Fig. 3A and B). The low requirement for oxygen was due to the low degree of reduction of RNA and the fermentable pentose sugar backbone. Oxygen resource costs to produce biomass or cellular energy increased for the other macromolecular DOC in the following order: cellobiose, aggregate biomass, protein and finally, lipid. Reduced carbon byproducts were produced from most DOC as oxygen became more limiting with acetate and formate being common (a complete list of byproducts is available in Table 1). The metabolic model of the heterotroph did not permit biomass or cellular energy production under anoxic conditions for any considered forms of DOC, which agrees with experimental observations (Beam *et al.*, 2016).

Heterotroph biomass produced from protein required the fewest enzymatic reactions, and by extension, the lowest enzyme resource cost of all macromolecular DOC

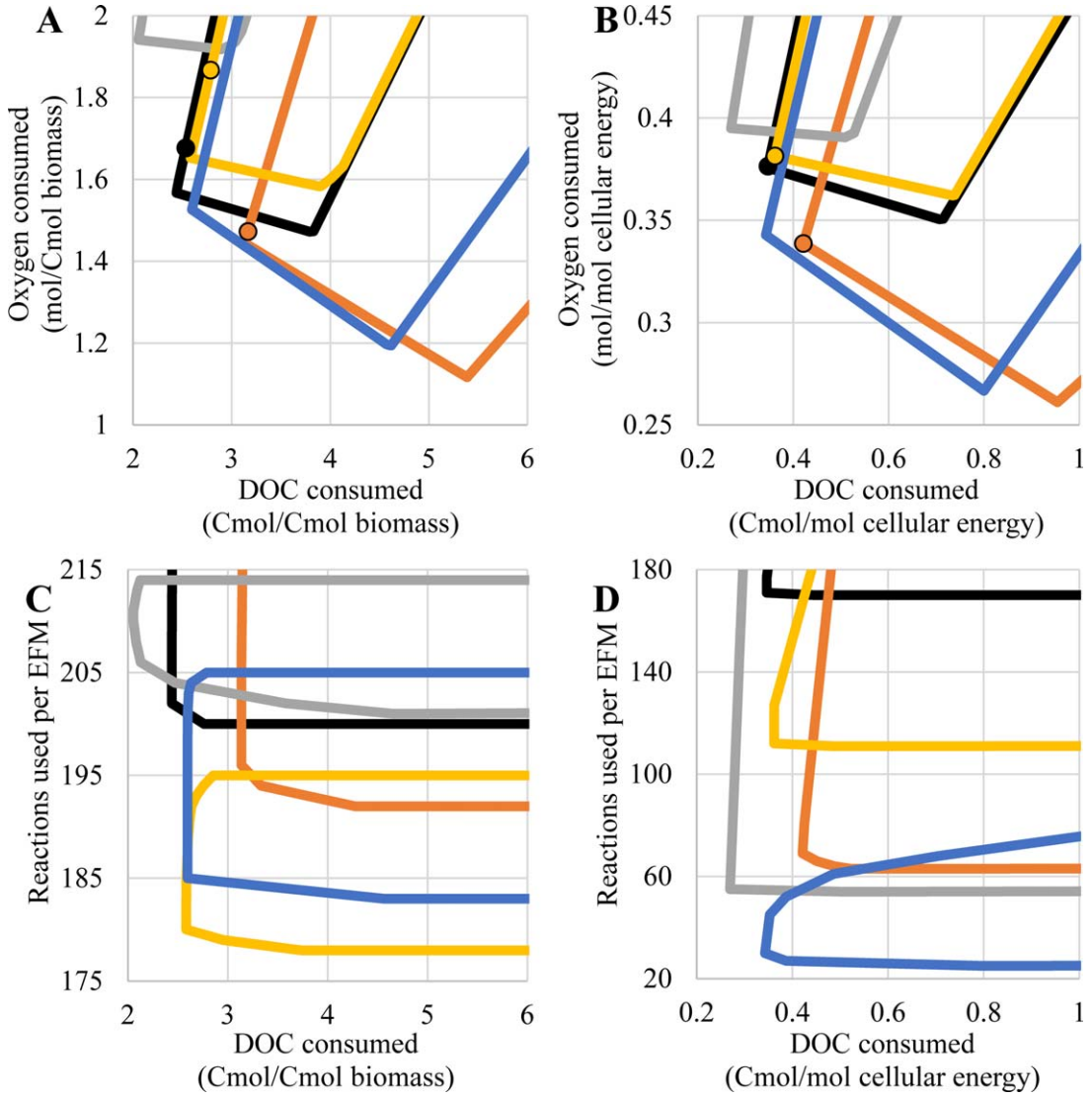

Fig. 3. Analysis of macromolecule utilisation by the heterotroph *Geoarchaeum* str. OSPB.
Biomass-derived DOC includes aggregate biomass (black), archaeal lipid (grey), cellobiose (blue), protein (yellow) and RNA (orange). Biomass and cellular energy production costs are plotted for the carbon- versus oxygen-limited scenarios (A) and (B), respectively. Biomass and cellular energy production costs are plotted for the carbon- versus enzyme-limited scenarios (C) and (D), respectively. Reactions used per elementary flux mode (EFM) are the number of reactions with a nonzero flux in each EFM, which is a proxy for the enzyme resource cost. Each tradeoff curve was calculated assuming simultaneous use of the macromolecule monomers, when applicable. Results from the sequential use of macromolecule monomers are highlighted with the circles in (A) and (B). For example, the solid yellow line in (A) outlines the tradeoff curve when the amino acids in the protein macromolecule were utilized simultaneously to produce biomass. The left most position on the yellow tradeoff curve represents the lowest DOC resource cost to produce biomass under carbon-limited conditions with its associated oxygen resource cost. The yellow circle (A) represents the DOC and oxygen resource costs to produce the same Cmole of heterotroph biomass when each amino acid is consumed sequentially; this simulation assumed all amino acids were metabolized without excretion.

considered. First, the protein macromolecule supplied all 20 common amino acids, which eliminated the need for amino acid synthesis reactions; second, the amino acid backbones were readily converted into central metabolism intermediates such as 3-phosphoglycerate and acetyl-CoA, which minimized the number of reactions involved in carbon backbone processing. After protein, the macromolecular DOC that minimized enzyme resource cost for biomass production were: cellobiose, RNA, aggregate biomass and finally, lipid (Fig. 3C). Heterotroph lipid [digeranylfarnesyl glycerol phospholipid (Kozubal *et al.*, 2013)] was modelled as chemically distinct from autotroph lipid [caldarchaeol diphosphate (Hopmans *et al.*, 2000)]; therefore, utilisation of autotroph lipid required enzymes for both degradation and reassembly to produce heterotroph biomass.

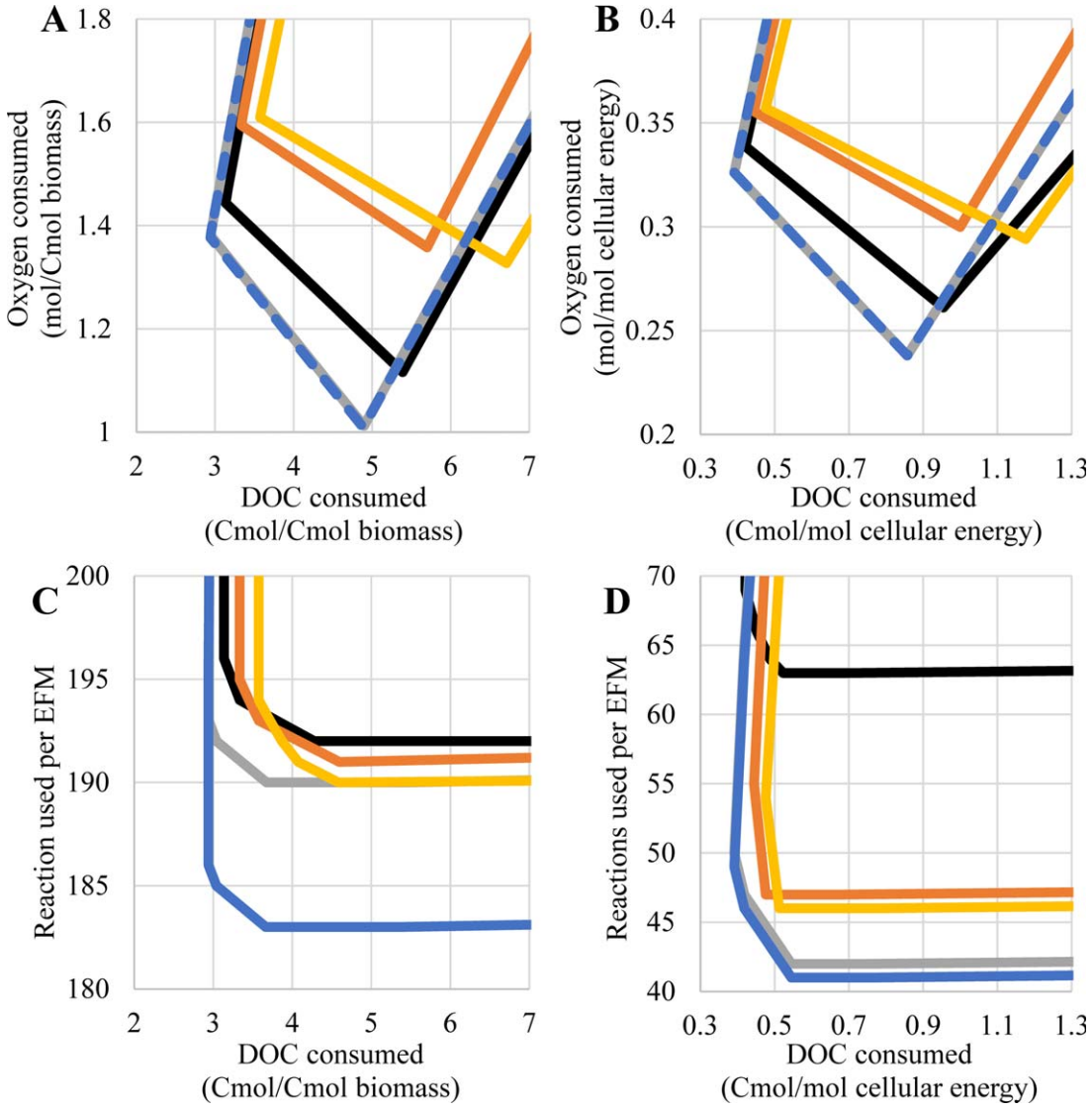

Fig. 4. Analysis of RNA macromolecule and monomer utilisation by the heterotroph *Geoarchaeum* str.OSPB.
Biomass-derived DOC includes RNA macromolecule (black), AMP (orange), CMP (grey), GMP (yellow) and UMP (blue). Biomass and cellular energy production costs are plotted for the carbon- versus oxygen-limited scenarios (A) and (B), respectively. Biomass and cellular energy production costs are plotted for the carbon- versus enzyme-limited scenarios (C) and (D), respectively. Reactions used per elementary flux mode (EFM) are the number of reactions with a nonzero flux in each EFM, which is a proxy for enzyme resource cost. UMP and CMP overlap in (A) and (B).

Cellular energy produced from cellobiose required the fewest enzymatic steps among the macromolecular DOC (Fig. 3D); all other macromolecules were degraded to multiple central metabolism precursors, requiring more enzymatic reactions to produce cellular energy. Cellobiose was followed by lipid, RNA, protein and then aggregate biomass as the most cost effective macromolecular DOC for cellular energy production with respect to enzyme resource cost. Nitrogen, phosphorous and sulfur containing DOC produced noncarbon byproducts, including ammonium, phosphate and sulfide when dissimilated to produce cellular energy (Table 1).

Selective, optimal utilisation of monomers from macromolecules

DOC simulations included both biomass-derived macromolecule and monomer pools enabling a comparison of different metabolic strategies. Biomass and cellular energy production from the RNA macromolecule and the four RNA monomers (AMP, CMP, GMP and UMP) revealed differences in DOC energy content and the biological accessibility of the carbon backbones (Fig. 4). Biomass and cellular energy production from the monomers CMP and UMP had lower DOC resource costs than either the

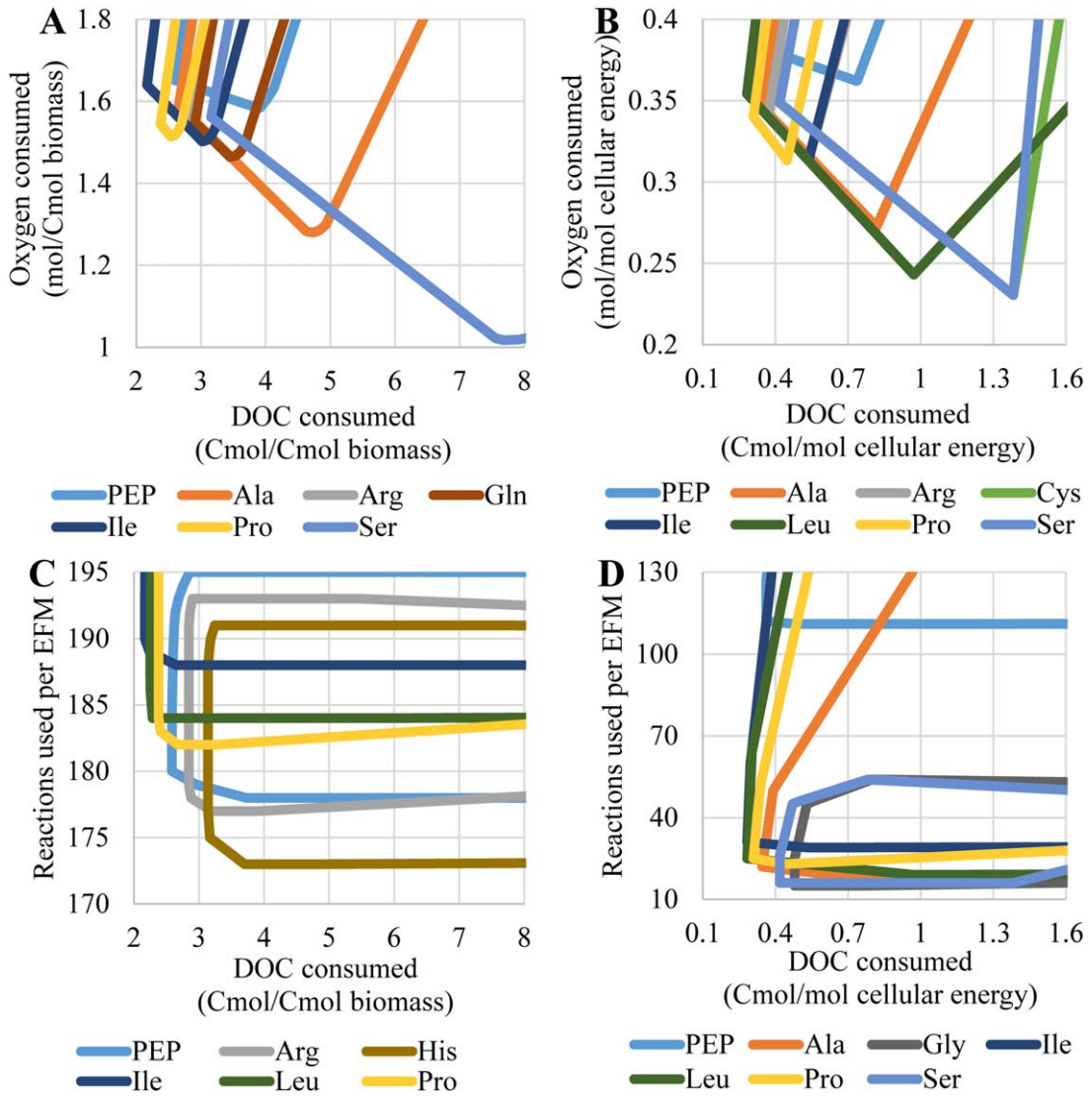

Fig. 5. Analysis of protein macromolecule and amino acid utilisation by the heterotroph *Geoarchaeum* str. OSPB. Biomass-derived DOC includes protein macromolecule and individual amino acids. Biomass and cellular energy production costs are plotted for the carbon- versus oxygen-limited scenarios (A) and (B), respectively. Biomass and cellular energy production costs are plotted for the carbon- versus enzyme-limited scenarios (C) and (D), respectively. Reactions used per elementary flux mode (EFM) are the number of reactions with a nonzero flux in each EFM, which is a proxy for enzyme resource cost. Only amino acids with optimal properties are plotted in each subfigure; Supporting Information data contains the complete amino acid data set.

aggregate RNA macromolecule, AMP or GMP (Fig. 4A and B); this was due to the higher degree of reduction and the connectivity of the carbon backbones to central metabolism. The enzyme resource cost was higher for the RNA macromolecule than any of the individual nucleotides because four separate monomer degradation pathways were necessary to generate central metabolism intermediates.

The 20 amino acids have a variety of metabolic properties including different elemental compositions and structures, which affected their potential use as substrates (Table 1). Isoleucine and leucine had the lowest DOC resource costs to produce both biomass and cellular energy compared to the protein macromolecule and other amino acids (Fig. 5). This was due largely to the high degree of reduction of these amino acids and the entry points of the carbon backbones into central metabolism. Serine had the lowest oxygen resource cost to produce biomass of the amino acids, making it more competitive than the protein macromolecule or any other amino acid under oxygen-limited conditions. Serine and cysteine had the lowest oxygen resource requirement for cellular energy

Table 2. Resource costs of heterotroph *Geoarchaeum* str. OSPB biomass production from the 29 analyzed biomass-derived DOC pools.

Carbon-limited costs			Oxygen-limited costs			Enzyme-limited costs		
	DOC per biomass	Oxygen per biomass		Oxygen per biomass	DOC per biomass		Fewest reactions used per EFM	DOC per biomass
LIP	2.06	1.94	UMP	1.01	4.89	His	173	3.70
Ile	2.17	1.64	CMP	1.01	4.90	Arg	177	3.20
Leu	2.25	1.74	Ser	1.02	7.68	PEP	178	3.73
Val	2.30	1.69	Cys	1.03	8.01	Gly	180	6.65
Pro	2.38	1.54	RNA	1.12	5.39	Ser	180	7.68
BIO	2.44	1.57	CEL	1.19	4.62	Cys	181	8.25
PEP	2.58	1.65	Asn	1.27	6.30	Pro	182	2.67
CEL	2.60	1.53	Asp	1.28	6.33	Gln	182	3.64
Ala	2.66	1.59	Ala	1.28	4.75	Glu	182	3.66
Met	2.75	2.23	GMP	1.33	6.71	Ala	182	4.88
Lys	2.84	2.24	Leu	1.34	4.86	Asn	182	6.48
Arg	2.84	1.54	AMP	1.36	5.70	Asp	182	6.51
Thr	2.86	1.79	Gly	1.40	6.54	UMP	183	3.67
Gln	2.91	1.55	Gln	1.46	3.45	CEL	183	4.57
Glu	2.92	1.56	His	1.47	4.16	Leu	184	2.29
UMP	2.94	1.38	Glu	1.47	3.47	Thr	184	4.17
CMP	2.94	1.38	BIO	1.47	3.79	Met	186	3.26
RNA	3.14	1.44	Ile	1.50	3.02	Lys	187	2.89
His	3.14	1.54	Arg	1.51	3.03	Ile	188	2.65
Ser	3.16	1.56	Pro	1.51	2.54	Trp	188	4.12
Cys	3.21	1.61	PEP	1.58	3.89	Val	190	2.35
AMP	3.33	1.59	Val	1.67	2.46	CMP	190	3.68
Trp	3.35	2.43	Thr	1.72	3.96	GMP	190	4.58
Asn	3.53	1.58	LIP	1.92	2.89	Tyr	190	6.51
Asp	3.55	1.59	Met	2.23	2.75	Phe	190	7.85
GMP	3.57	1.61	Lys	2.24	2.84	AMP	191	4.60
Gly	3.62	1.64	Trp	2.43	3.35	RNA	192	4.28
Tyr	3.76	2.90	Tyr	2.90	3.76	BIO	200	2.76
Phe	4.04	3.42	Phe	3.42	4.04	LIP	201	4.66

DOC is ordered from low to high biomass production costs for the three examined nutrient limitations.
Biomass production costs utilizing biomass (BIO), archaeal lipid (LIP), cellobiose (CEL), a representative protein (PEP) and a representative distribution of nucleotides (RNA).
Nucleotides and amino acids are labeled using standard biochemical conventions.
Units are Cmoles DOC consumed, Cmoles biomass produced and moles oxygen consumed.

production of the amino acids. The serine backbone enters glycolysis at pyruvate, which can be converted to acetate, carbon dioxide and cellular energy using only a few enzymatic reactions including substrate-level phosphorylation. Moreover, acetate can be secreted, which removes reducing equivalents that would otherwise require oxygen. The enzyme resource costs to produce cellular energy from amino acids were a function of the carbon backbone (Fig. 5C and D). Structurally simpler amino acids, such as glutamate and aspartate, could be integrated into central metabolism with fewer enzymatic reactions than cyclic or aromatic amino acids, such as phenylalanine, which required longer and more specialized enzymatic pathways. A similar trend was observed for biomass production, with the exception of histidine and arginine. Degradation of histidine and arginine required fewer enzymatic reactions than their synthesis (Supporting Information); this differential made the DOC ideal for minimizing enzyme resource costs to produce biomass.

Simultaneous versus sequential metabolism of monomer pools

Biomass utilisation strategies were examined as two metabolic scenarios: simultaneous and sequential DOC metabolism (Fig. 1C). Differences in the two metabolic strategies were exemplified by comparing the metabolism of the protein macromolecule to the sequential utilisation of each amino acid (Fig. 3). The sequential metabolism was quantified via nonnegative linear combinations of optimal EFMs for each amino acid (Tables 2 and 3). Sequential metabolism of individual amino acids had higher DOC and oxygen resource costs to produce biomass than the simultaneous metabolism of the amino acids (Fig. 3A). Sequential consumption necessitated more metabolic processing; each amino acid had to be converted into the other 19 amino acids as well as the other biomass components. The added metabolic processing consumed more cellular energy and reducing equivalents and produced

Table 3. Resource costs of heterotroph *Geoarchaeum* str. OSPB cellular energy production from the 29 analyzed biomass-derived DOC pools.

Carbon-limited costs			Oxygen-limited costs			Enzyme-limited costs		
	DOC per energy	Oxygen per energy		Oxygen per energy	DOC per energy		Fewest reactions used per EFM	DOC per energy
LIP	0.270	0.395	Cys	0.230	1.382	Gly	15	0.480
Ile	0.283	0.354	Ser	0.230	1.382	Ser	16	0.418
Leu	0.283	0.354	CMP	0.238	0.857	Ala	17	0.817
Val	0.300	0.360	UMP	0.238	0.857	Cys	17	1.382
Pro	0.309	0.340	Leu	0.243	0.972	Leu	19	0.972
CEL	0.343	0.343	RNA	0.261	0.954	Asp	20	1.090
Ala	0.346	0.346	CEL	0.267	0.800	Glu	21	0.612
BIO	0.347	0.376	Ala	0.272	0.817	Gln	21	0.612
Lys	0.360	0.420	Asn	0.272	1.090	Pro	23	0.448
PEP	0.362	0.381	Asp	0.272	1.090	His	23	0.620
Met	0.366	0.439	GMP	0.294	1.176	Asn	23	1.090
Arg	0.371	0.340	Gly	0.299	1.198	Thr	24	0.705
Thr	0.375	0.375	AMP	0.300	1.000	Lys	24	0.900
Glu	0.380	0.342	Gln	0.306	0.612	CEL	25	0.800
Gln	0.380	0.342	His	0.306	0.734	Arg	26	0.537
UMP	0.391	0.326	Glu	0.306	0.612	Val	29	0.319
CMP	0.391	0.326	Arg	0.313	0.537	Ile	29	0.537
His	0.409	0.341	Pro	0.313	0.448	Met	30	0.882
Ser	0.418	0.349	Ile	0.313	0.537	Trp	31	1.082
Cys	0.418	0.349	Val	0.343	0.428	Tyr	31	1.102
RNA	0.421	0.339	BIO	0.351	0.707	Phe	34	1.349
Trp	0.437	0.457	Thr	0.353	0.705	UMP	41	0.545
AMP	0.444	0.356	PEP	0.362	0.734	CMP	42	0.545
Asp	0.461	0.346	LIP	0.391	0.509	GMP	46	0.513
Asn	0.461	0.346	Lys	0.420	0.360	AMP	47	0.476
GMP	0.476	0.357	Met	0.439	0.366	LIP	54	0.509
Gly	0.480	0.360	Trp	0.457	0.437	RNA	63	0.523
Tyr	0.495	0.523	Tyr	0.523	0.495	PEP	111	0.483
Phe	0.540	0.600	Phe	0.600	0.540	BIO	170	0.448

DOC is ordered from low to high cellular energy production costs for the three examined nutrient limitations.
Energy production costs utilizing biomass (BIO), archaeal lipid (LIP), cellobiose (CEL), a representative protein (PEP), and a representative distribution of nucleotides (RNA).
Nucleotides and amino acids are labeled using standard biochemical conventions.
Units are Cmoles DOC consumed, moles energy produced and moles oxygen consumed.

more carbon byproducts, which increased the DOC resource cost to produce biomass.

Simultaneous versus sequential DOC consumption had no effect on cellular energy production (Fig. 3B). Cellular energy production exclusively harvests energy from the DOC, while biomass production uses DOC to produce both biomass components and cellular energy. This property resulted in nonlinear relationships between different forms of DOC for biomass but not cellular energy production.

Identification of suboptimal metabolic strategies

Resource allocation theory identified optimal strategies for biomass and cellular energy production along continuous gradients of resource limitations. However, this does not capture the total robustness of the metabolic network. EFMA enumerated all mathematically possible biochemical pathways including optimal and suboptimal strategies. The suboptimal strategies can be biologically significant because they create metabolic redundancy and buffer against perturbations (Mahadevan and Lovley, 2008). The analysis of suboptimal strategies considered EFMs with biomass yields within 0.5% of the optimal tradeoff surface as phenotypically indistinguishable and were therefore part of the tradeoff surface. The 0.5% threshold value was set based on experimental observations (Dykhuizen and Hartl, 1980). Analysis of suboptimal biomass production strategies expanded the relevant number of EFMs along the DOC versus oxygen tradeoff curve from 3 to 2073 for aggregate biomass utilisation (Fig. 2) while it had no effect on the four EFMs considered on the cellobiose utilisation tradeoff curve.

The serine–glycine pathway is an example of an identified suboptimal metabolic strategy with competitive properties (Fig. 2). The serine–glycine pathway uses

serine and glycine synthesis reactions along with the glycine cleavage system to oxidize DOC (Supporting Information Fig. S2). This pathway permits complete oxidation of DOC, but requires fewer enzymatic steps than the TCA cycle, potentially making it competitive in enzyme-limited environments. However, fewer substrate-level phosphorylation reactions occurred in this pathway, which increased DOC resource cost. In the case of aggregate biomass utilisation to produce biomass, the serine–glycine pathway is suboptimal on the enzyme resource plots, though it is optimal for some forms of DOC, such as cellobiose. Additionally, the serine–glycine pathway provides flexibility in response to oxygen limitation. In the presence of sufficient oxygen, the pathway can completely oxidize DOC to maximize cellular energy production. Alternatively, under oxygen limitation, formate produced when regenerating tetrahydrofolate can be excreted to lower the requirement for oxygen. Consequently, results from EFMA analysis provide a theoretical foundation for understanding and/or predicting effects of different electron acceptor costs on microbial community response.

Discussion

Nutrient and energy transfer across trophic levels during assimilatory and dissimilatory biomass utilisation is fundamental to most ecosystems. Heterotrophic utilisation of reduced carbon constituents produced by autotrophs can occur as a result of metabolite exchange, predation (e.g., viral lysis) and/or decomposition. The presented work focused on decomposition of autotroph biomass and established quantitative relationships between community members. For example, a minimum of 2.4 Cmoles of autotroph biomass is required to synthesize 1 Cmole of heterotroph biomass. The exchange of metabolites is also common both within and across trophic levels (Walker *et al.*, 2012; Morris *et al.*, 2013). Rational analysis of these synergistic exchanges requires a numerical basis. The value of exchanged DOC can be quantified in terms of the amount of biomass or cellular energy that can be produced from it (e.g., Tables 2 and 3). For instance, the amino acid isoleucine can be oxidized to produce 3.5 moles of cellular energy per Cmole, while phenylalanine can be oxidized to produce 1.9 moles of cellular energy per Cmole (Table 3). The presented study provides a quantitative foundation for analyzing metabolite exchange within microbial communities.

Stoichiometric modelling is well suited for analyzing complex systems such as metabolism and can quantify the extremes of cellular physiology without knowledge of substrate concentrations or specific enzyme parameters; this is a major strength of the modelling technique. Cellular phenotypes over the continuum of specific growth rates can be modelled using linear combinations of the EFMs at fast and slow growth rates, namely the EFMs that produce only biomass and only cellular energy, respectively. In this study, the modelled biomass production assumes a relatively high specific growth rate (0.1 h^{-1}) and therefore, the majority of substrate is directed towards synthesizing biomass components and not towards cell maintenance (3.5 moles cellular energy per Cmole biomass). Conversely, at slow specific growth rates, most of the cellular activity is directed towards cell maintenance; these phenotypes are represented in the cellular energy plots (e.g., Fig. 2B and C). Combinations of these two scenarios can be used to approximate any phenotype between the two extremes. Stoichiometric modelling does not automatically account for enzyme affinities and therefore accounting for low substrate concentrations and the expression of high affinity transporters would need to be implemented manually; the additional energetic requirements for a high affinity transporter versus a low affinity transporter has been detailed previously (Carlson, 2007). In theory, strategies that only produce cellular energy could describe the ultimate low nutrient environments found in the deep subsurface, where biomass doubling times on the order of 10^2-10^4 years have been estimated (Hoehler and Jørgensen, 2013).

DOC utilisation when multiple forms are present was bounded by two scenarios, simultaneous and sequential metabolism. These results can be used to study microbial growth at either high or low DOC concentrations. Many experimental studies have examined sequential metabolism of simple carbohydrates and organic acids, often termed catabolite repression or diauxic growth (Wolfe, 2005; Deutscher, 2008; Görke and Stülke, 2008). Sequential metabolism is competitive at high DOC concentrations where use of a single form of DOC can saturate the activity of a minimal enzyme investment and drive high growth rates. At low DOC concentrations, such as those found in most natural environments including the studied mat system, simultaneous metabolism of multiple forms of DOC is observed. This strategy requires more investment into degradation pathways, but is necessary to maximize growth rate (Kovarova-Kovar and Egli, 1998).

The presented study establishes a foundation for the evaluation of biomass utilisation by heterotrophs; additional complexity could be integrated during future work. Utilisation of complex DOC mixtures is common in natural environments and is constrained by bioavailability (e.g., solubility and polymer hydrolysis), DOC transport into a cell and the necessary degradation pathways (Kovarova-Kovar and Egli, 1998). The solubility of DOC was not explicitly considered here, and could confound the predicted optimality of some forms of DOC, such as lipids which have lower solubilities ($\sim$5 $\mu g\ l^{-1}$) than most other DOC constituents. A review of relevant DOC solubility can be found in the Supporting Information (see Table S1). Currently, all DOC is transported and hydrolyzed via a

single macromolecule-specific mechanism (see Materials and Methods). While this permits equitable comparisons across simulations for simultaneous and sequential metabolism, it also highlights scientific knowledge gaps associated with enzyme specificity and enzyme resource costs of transport and hydrolysis reactions. Many microorganisms have multiple transporters for the same molecule (e.g., low affinity versus high affinity transporters) or same class of molecules (e.g., glucose vs. glucose oligomer transporters), which can influence phenotype (Zhang and Lynd, 2005). Another simplification made in the current study was the complete utilisation of a metabolite pool (e.g., all amino acids in the protein macromolecule) as compared to the selective use of optimal monomers (e.g., metabolism of alanine combined with the secretion of phenylalanine) or the selective use of components from a complex molecule (e.g., metabolism of the ribose unit in an RNA nucleotide with the secretion of the recalcitrant aromatic base). Wastewater and sediment studies have commonly shown that recalcitrant molecules like aromatic amino acids are often utilized last (Crawford *et al*., 1974; Tegelaar *et al*., 1989; Arndt *et al*., 2013). Finally, the broad applicability of the presented approach was demonstrated by applying the theory to an ecosystem with extensive DOC utilisation data. The analysis compares theoretical and experimentally measured utilisation of amino acids within a eutrophic marine ecosystem (presented in the Supporting Information). The experimental measurements were predicted with remarkable accuracy using the oxygen limited acclimation strategies presented in this study highlighting how basic physiological concepts can be applied to many different ecosystems.

This *in silico* systems biology study quantified assimilatory and dissimilatory biomass utilisation by an aerobic heterotrophic archaeon, *Geoarchaeum* str. OSPB. The study provided a metabolic basis for comparing optimal DOC utilisation patterns under different nutrient-limiting conditions, evaluating simultaneous versus sequential DOC metabolism and investigating byproducts secreted during simulated growth and cellular energy production. These relationships describe limits for community interactions between an autotroph and a heterotroph, such as the amount and type of autotrophic biomass needed to sustain the community. The novel approaches and findings of this study are applicable to a wide range of microbial ecosystems.

Materials and methods

Metabolic model construction

The *Geoarchaeum* str. OSPB (NCBI taxon ID 1448933) genome sequence (Markowitz *et al*., 2012), which has been assembled from metagenomes of iron oxide mats sampled from One Hundred Springs Plain (Kozubal *et al*., 2013), was used to build the *in silico* heterotroph metabolic network. Incorporation of enzyme-catalyzed reactions into the model was based on protein homology (greater than 30% identity) and gene annotations. Due to the incomplete nature of metagenome assemblies, missing genes were modelled as present, when necessary, using genes found in other archaea (Huson *et al*., 2007). The metabolic model of *Geoarchaeum* str. OSPB exhibits typical heterotrophic archaeal central carbon metabolism (e.g., TCA cycle, archaeal glycolysis/gluconeogenesis and electron transport) and common biosynthetic and degradation pathways. The model did not predict auxotrophy for any of the major biomass components (i.e., amino acids or nucleotides). Acetate, formate and formaldehyde were the only reduced carbon byproducts considered in the model, based on gene annotation. No genes were identified to support the inclusion of other reduced carbon byproducts. Enzyme resource costs accounted for the number of enzymes in the manually compressed reactions, with the exception of transporters and degradation enzymes (Supporting Information).

Macromolecules and monomers were not allowed to be excreted by the heterotroph due to EFMA computational explosion. Peptides were assumed to be transported actively using an ABC transporter for trimeric peptides and depolymerized independent of cellular energy. RNA was transported and depolymerized in a cellular energy neutral reaction. Cellobiose was assumed to be transported actively using an ABC transporter and depolymerized independent of cellular energy, but required phosphorylation to enter central metabolism. Archaeal lipid was assumed to be transported actively using an ABC transporter and required complete degradation to monomers before polymerisation into heterotroph biomass. Monomers were transported and degraded as homopolymers to avoid biases associated with differences in transport costs of polymers and oligomers.

Heterotroph biomass production

The heterotroph biomass production reactions were developed using macromolecular ratios as previously described (Neidhardt *et al*., 1990), tailoring amino acid and nucleotide distributions to organism-specific genome data (Supporting Information). The amino acid distribution was determined using the average distribution of all open reading frames in the *de novo* assembly of *Geoarchaeum* str. OSPB from the metagenome. The nucleotide distribution of RNA was represented by the weighted average nucleotide distribution of major rRNA subunits (i.e., 16S and 23S); rRNA pools are the largest fraction of RNA in the cell [e.g., approximately 80% in *E. coli* (Neidhardt *et al*., 1990)]. Autotroph lipid was represented by ether-linked digeranylfarnesyl glycerol phospholipid (Hopmans *et al*., 2000). Polysaccharide was represented by cellobiose. The maintenance energy was set based on glucose utilisation by *Alicyclobacillus acidocaldarius* DSM 446, a thermoacidophilic gram-positive bacterium, grown at a pH and temperature of 4.3°C and 51°C, respectively (Farrand *et al*., 1983). The maintenance energy requirements were determined assuming the use of a cytochrome bd oxidase, which resulted in 1 proton translocated per electron, because a cytochrome c oxidase system, which resulted in 2 protons translocated per electron, resulted in uncharacteristically high

maintenance energies (i.e., 3.5 vs. 5 moles cellular energy per Cmole biomass or 150 vs. 215 mmoles cellular energy per g cell dry weight for a cytochrome bd oxidase versus cytochrome c oxidase, respectively, at a growth rate of 0.1 h^{-1}). One mole of cellular energy was modelled as one mole of phosphodiester bonds broken. EFMs using cytochrome c oxidase were subsequently separated from EFMs using cytochrome bd oxidase in the primary analysis for clarity as all trends are similar, but scaled by the difference in protons transferred per oxygen respired (Supporting Information).

Stoichiometric analyses

Model reactions and metabolites were assembled using Microsoft Excel, transferred to CellNetAnalyzer version 2014.1 (Klamt *et al.*, 2007; Klamt and von Kamp, 2011) for formatting, and exported to RegEFMTool version 2.0 (Jungreuthmayer *et al.*, 2013) for EFM enumeration. Gene regulatory rules mutually excluding cyclic reactions were incorporated with RegEFMTool to minimize futile cycles and validated against network subsets analyzed with EFMTool version 4.7.1 (Terzer, 2006; Terzer and Stelling, 2008) without regulatory rules (Supporting Information). 'Metabolic check valves' (i.e., pseudo-metabolites and irreversible reactions) were used in addition to gene regulatory rules to minimize futile cycling of nucleotides, which are often involved in reactions as coenzymes (Supporting Information). FBA-based resource costs were calculated from the *doubleRobustnessAnalysis* function of COBRA Toolbox (Schellenberger *et al.*, 2011) on the model exported using the *CNA2Cobra* function of CellNetAnalyzer. The resulting output was converted to resource costs by normalizing to either the specific growth rate or the specific cellular energy production rate, and compared to resource cost tradeoff curves produced from EFMs (Supporting Information). An example of model construction, EFMA and resource costs is available in the Supporting Information (Fig. S1). A general maintenance energy of 3.5 moles cellular energy per Cmole of biomass was held constant for all simulations, which included contributions from growth and nongrowth associated maintenance energy at a growth rate of 0.1 h^{-1}. Computations were performed on a machine with two X5690 Intel Xeon processors and 120 GB of RAM.

Acknowledgements

This study was supported by the National Science Foundation Integrative Graduate Education and Research Training (IGERT) Program (DGE 0654336) and subcontracts (112443, 254840) from Pacific Northwest National Laboratory (Richland, WA) to MSU as part of the DOE-BER Foundational Science Focus Area in Microbial Community Design Principles. K.A.H. was also supported by NIH (1U01EB019416) and the Department of Chemical and Biological Engineering at Montana State University.

Conflict of Interest

The authors declare no conflicts of interest in the preparation of this paper.

References

Arndt, S., Jørgensen, B.B., LaRowe, D.E., Middelburg, J.J., Pancost, R.D., and Regnier, P. (2013) Quantifying the degradation of organic matter in marine sediments: a review and synthesis. *Earth Sci Rev* **123:** 53–86.

Beam, J.P., Bernstein, H.C., Jay, Z.J., Kozubal, M.A., Jennings, R. deM., Tringe, S.G., and Inskeep, W.P. (2016) Assembly and succession of iron oxide microbial mat communities in acidic geothermal Springs. *Front Microbiol* **7:** 25.

Beck, A.E., Hunt, K.A., Bernstein, H.C., and Carlson, R.P. (2016) Interpreting and Designing Microbial Communities for Bioprocess Applications, from Components to Interactions to Emergent Properties. In, Eckert, C., and Trinh, C.T. (eds), *Biotechnology for Biofuel Production and Optimization*. Elsevier, Amsterdam, Netherlands, pp. 407–432.

Bernstein, H.C., Beam, J.P., Kozubal, M.A., Carlson, R.P., and Inskeep, W.P. (2013) *In situ* analysis of oxygen consumption and diffusive transport in high-temperature acidic iron-oxide microbial mats. *Environ Microbiol* **15:** 2360–2370.

Boyd, E.S., Leavitt, W.D., and Geesey, G.G. (2009) CO_2 uptake and fixation by a thermoacidophilic microbial community attached to precipitated sulfur in a geothermal spring. *Appl Environ Microbiol* **75:** 4289–4296.

Breitbart, M., Wegley, L., Leeds, S., Schoenfeld, T., and Rohwer, F. (2004) Phage community dynamics in hot springs. *Appl Environ Microbiol* **70:** 1633–1640.

Carlson, R., and Srienc, F. (2004) Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: creation of overall flux states. *Biotechnol Bioeng* **86:** 149–162.

Carlson, R., Fell, D., and Srienc, F. (2002) Metabolic pathway analysis of a recombinant yeast for rational strain development. *Biotechnol Bioeng* **79:** 121–134.

Carlson, R.P. (2007) Metabolic systems cost-benefit analysis for interpreting network structure and regulation. *Bioinformatics* **23:** 1258–1264.

Cerqueda-García, D., Martínez-Castilla, L.P., Falcón, L.I., and Delaye, L. (2014) Metabolic analysis of *Chlorobium chlorochromatii* CaD3 reveals clues of the symbiosis in "chlorochromatium aggregatum." *ISME J* **8:** 991–998.

Crawford, C.C., Hobbie, J.E., and Webb, K.L. (1974) The utilization of dissolved free amino acids by estuarine microorganisms. *Ecology* **55:** 551–563.

Dance, A. (2008) Soil ecology: what lies beneath. *Nature* **455:** 724–725.

Deutscher, J. (2008) The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* **11:** 87–93.

Dykhuizen, D., and Hartl, D. (1980) Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background. *Genetics* **96:** 801–817.

Farrand, S.G., Linton, J.D., Stephenson, R.J., McCarthy, W.V., and Jones, C.W. (1983) The effect of temperature and pH on the growth efficiency of the thermoacidophilic bacterium *Bacillus acidocaldarius* in continuous culture. *Arch Microbiol* **135:** 276–283.

Folsom, J.P., and Carlson, R.P. (2015) Physiological, biomass elemental composition and proteomic analyses of escherichia coli ammonium-limited chemostat growth, and comparison with iron- and glucose-limited chemostat growth. *Microbiology* **161:** 1659–1670.

Görke, B., and Stülke, J. (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* **6:** 613–624.

Grady, C.P.L.J., Daigger, G.T., Love, N.G., and Filipe, C.D.M. (2011) *Biological Wastewater Treatment (*3rd ed.). Taylor & Francis.

Hoehler, T.M., and Jørgensen, B.B. (2013) Microbial life under extreme energy limitation. *Nat Rev Microbiol* **11:** 83–94.

Hopmans, E.C., Schouten, S., Pancost, R.D., Van Der Meer, M.T.J., and Sinninghe Damsté, J.S. (2000) Analysis of intact tetraether lipids in archaeal cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. *Rapid Commun Mass Spectrom* **14:** 585–589.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17:** 377–386.

Inskeep, W.P., Jay, Z.J., Herrgard, M.J., Kozubal, M.A., Rusch, D.B., Tringe, S.G., *et al.* (2013) Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. *Front Microbiol* **4:** 95.

Jennings, R.M., Whitmore, L.M., Moran, J.J., Kreuzer, H.W., and Inskeep, W.P. (2014) Carbon dioxide fixation by *Metallosphaera yellowstonensis* and acidothermophilic iron-oxidizing microbial communities from Yellowstone National Park. *Appl Environ Microbiol* **80:** 2665–2671.

Jungreuthmayer, C., Ruckerbauer, D.E., and Zanghellini, J. (2013) regEfmtool: speeding up elementary flux mode calculation using transcriptional regulatory rules in the form of three-state logic. *Biosystems* **113:** 37–39.

Klamt, S., and von Kamp, A. (2011) An application programming interface for cellnetanalyzer. *Biosystems* **105:** 162–168.

Klamt, S., and Stelling, J. (2003) Two approaches for metabolic pathway analysis? *Trends Biotechnol* **21:** 64–69.

Klamt, S., Saez-Rodriguez, J., and Gilles, E.D. (2007) Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Syst Biol* **129:** 329–351.

Kovarova-Kovar, K., and Egli, T. (1998) Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. *Microbiol Mol Biol Rev* **62:** 646–666.

Kozubal, M.A., Dlakic, M., Macur, R.E., and Inskeep, W.P. (2011) Terminal oxidase diversity and function in "*Metallosphaera yellowstonensis*": gene expression and protein modeling suggest mechanisms of Fe(II) oxidation in the Sulfolobales. *Appl Environ Microbiol* **77:** 1844–1853.

Kozubal, M.A., Macur, R.E., Jay, Z.J., Beam, J.P., Malfatti, S.A., Tringe, S.G., *et al.* (2012) Microbial iron cycling in acidic geothermal springs of yellowstone national park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Front. Microbiol* **3:** 109.

Kozubal, M.A., Romine, M., Jennings, R., deM Jay, Z.J., Tringe, S.G., Rusch, D.B., *et al.* (2013) Geoarchaeota: a new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. *ISME J* **7:** 622–634.

Llaneras, F., and Picó, J. (2010) Which metabolic pathways generate and characterize the flux space? a comparison among elementary modes, extreme pathways and minimal generators. *J Biomed Biotechnol* 753904.

Mahadevan, R., and Lovley, D.R. (2008) The degree of redundancy in metabolic genes is linked to mode of metabolism. *Biophys J* **94:** 1216–1220.

Markowitz, V.M., Chen, I.M.A., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., *et al.* (2012) IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res* **40:** D115–D122.

Mason, I.G. (2006) Mathematical modelling of the composting process: a review. *Waste Manag* **26:** 3–21.

Morris, B.E.L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013) Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* **37:** 384–406.

Neidhardt, F., Ingraham, J., and Schaechter, S. (1990) Physiology of the bacterial cell: a molecular approach 1st ed. Simpson, J. (ed) Sinauer Associates, Inc., Sunderland, Massachusetts.

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

Reed, J.L., and Palsson, B.O. (2003) Thirteen years of building constraint-based in silico models of *Escherichia coli*. *J. Bacteriol* **185:** 2692.

Rohwer, F., Prangishvili, D., and Lindell, D. (2009) Roles of viruses in the environment. *Environ Microbiol* **11:** 2771–2774.

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., *et al.* (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox v2.0. *Nat Protoc* **6:** 1290–1307.

Schilling, C.H., Letscher, D., and Palsson, B. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J Theor Biol* **203:** 229–248.

Stolyar, S., Van Dien, S., Hillesland, K.L., Pinel, N., Lie, T.J., Leigh, J. A., and Stahl, D.A. (2007) Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* **3:** 92.

Taffs, R., Aston, J.E., Brileya, K., Jay, Z., Klatt, C.G., McGlynn, S., *et al.* (2009) *In silico* approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst Biol* **3:** 114.

Tegelaar, E.W., de Leeuw, J.W., Derenne, S., and Largeau, C. (1989) A reappraisal of kerogen formation. *Geochim Cosmochim Acta* **53:** 3103–3106.

Terzer, M. (2006) Accelerating the computation of elementary modes using pattern trees. *Algorithms Bioinforma* **2006:** 333–343.

Terzer, M., and Stelling, J. (2008) Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* **24:** 2229–2235.

Trinh, C.T., Wlaschin, A., and Srienc, F. (2009) Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl Microbiol Biotechnol* **81:** 813–826.

Varma, A., Boesch, B.W., and Palsson, B. (1993) Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl Environ Microbiol* **59:** 2465–2473.

Walker, C.B., Redding-Johanson, A.M., Baidoo, E.E., Rajeev, L., He, Z., Hendrickson, E.L., *et al*. (2012) Functional responses of methanogenic archaea to syntrophic growth. *ISME J* **6:** 2045–2055.

Wolfe, A.J.,(2005) The acetate switch. *Microbiol Mol Biol Rev* **69:** 12–50.

Zhang, Y.H.P., and Lynd, L.R. (2005) Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. *Proc Natl Acad Sci USA* **102:** 7321–7325.

Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., and Lovley, D.R. (2011) Genome-scale dynamic modeling of the competition between rhodoferax and geobacter in anoxic subsurface environments. *ISME J* **5:** 305–316.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Summary of metabolic model building, stoichiometric analysis and application of ecological theory. (A) The metabolic model is a compilation of genomic annotations from databases, experimental yields and efficiencies reported in the literature, and hypothesized reactions. Hypothesized reactions include those missing from annotations, but expected due to literature data (i.e., an organism that grows on a mineral medium has complete biosynthetic pathways). (B) The model is represented as a series of reaction equations correlating the consumed metabolites and the produced metabolites (i.e., Ri, where i corresponds to the reaction number and is written in terms of metabolites A, B and C; W, X, Y and Z are sources and sinks). The reaction equations can be written concisely as a mathematical matrix, positive coefficients are reaction products and negative coefficients are reaction substrates. (C) These reaction equations are analyzed to identify metabolic routes through the network that are stoichiometrically balanced (i.e., all metabolites produced must be consumed and vice versa) and all reactions are limited to directions permitted by chemical thermodynamics. In flux balance analysis (FBA), these routes must also satisfy an objective function (e.g., maximize flux through R6 while constraining the magnitude of flux through reactions R1 and R2). Elementary flux mode (EFM) analysis finds all of the simplest, genetically distinct routes through the network (i.e., the five possible EFMs (color coded) for the sample metabolic model). (D) Application of ecological theory to the analysis of these routes predicts competitive metabolic behaviour. The theory assumes the reactions used to produce the desired product (e.g., biomass or cellular energy) will minimize the resource cost for the limiting nutrient (e.g., carbon, oxygen or cellular volume). In the context of the depicted example, if W is limiting and a population must produce Z, theory states the optimal strategy would be along the *x*-axis (plot to the left). Conversely, limitation of X to produce Z would be along the *y*-axis (plot to the left). In addition, alternative products may favour alternative metabolic strategies (e.g., production of Y in the plot to the right). Additional review articles on stoichiometric modelling: (Price et al., 2004; Feist et al., 2009; Trinh et al., 2009; Beck et al., 2016).

Fig. S2. Graphical representation of the serine–glycine pathway and tetrahydrolate regeneration. The net pathway of serine–glycine and tetrahydrofolate regeneration results in formate and reducing equivalents. If the formate produced is then oxidized, the net result is electronically identical to complete oxidation through the tricarboxylic acid cycle but uses fewer enzymes.

Fig. S3. Correlation of theoretical amino acid carbon resource cost (Cmol amino acid consumed/Cmol biomass or cellular energy produced) and experimental volumetric amino acid consumption rate. Carbon resource costs to produce cellular energy (Cmol amino acid/mol cellular energy) (left) and biomass (Cmol amino acid/Cmol biomass) (right) under carbon- (A), oxygen- (B) or enzyme-limited conditions (C) are shown for all amino acids except asparagine, cysteine, glutamine, histidine and tryptophan, which were not measured. Volumetric consumption rates are the average consumption rates calculated from (Crawford *et al*., 1974). Error bars are standard error of measurements for up to 12 measurements.

Table S1. Solubilities of common forms of dissolved organic carbon at 25°C.