

# **Modeling the Interplay between Photosynthesis, CO<sub>2</sub> Fixation, and the Quinone Pool in a Purple Non-Sulfur Bacterium**

Adil Alsiyabi<sup>1</sup>, Cheryl Immethun<sup>1</sup>, Rajib Saha\*<sup>1</sup>

<sup>1</sup>Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln

\*Corresponding author:

Rajib Saha

Assistant Professor

Chemical and Biomolecular Engineering,

University of Nebraska-Lincoln

Lincoln, NE-68588, USA

Email: [rsaha2@unl.edu](mailto:rsaha2@unl.edu)

## Abstract

*Rhodospseudomonas palustris* CGA009 is a purple non-sulfur bacterium (PNSB) that can fix CO<sub>2</sub> and nitrogen or break down organic compounds for its carbon and nitrogen requirements. Light, inorganic, and organic compounds can all be used for its source of energy. Excess electrons produced during its metabolic processes can be exploited to produce hydrogen gas or biodegradable polyesters (polyhydroxybutyrate). A genome-scale metabolic model of the bacterium was reconstructed to study the interactions between photosynthesis, carbon dioxide fixation, and the redox state of the quinone pool. A comparison of model-predicted flux values with published *in vivo* MFA fluxes resulted in predicted errors of 5-19% across four different growth substrates. The model predicted the presence of an unidentified sink responsible for the oxidation of excess quinols generated by the TCA cycle. Furthermore, light-dependent energy production was found to be highly dependent on the rate of quinol oxidation. Finally, the extent of CO<sub>2</sub> fixation was predicted to be dependent on the amount of ATP generated through the electron transport cycle, with excess ATP going toward the energy-demanding CBB pathway. Based on this analysis, it is hypothesized that the quinone redox state acts as a feed-forward controller of the CBB pathway, signaling the amount of ATP available.

## Introduction

Purple non-sulfur Bacteria (PNSB) are considered to be among the most metabolically versatile groups of bacteria<sup>1,2</sup>. Within this class, *Rhodospseudomonas palustris* CGA009 (hereafter *R. palustris*) demonstrates this elasticity through its ability to survive in a myriad of diverse environmental conditions<sup>3</sup>. It can grow either aerobically or anaerobically, utilize organic (heterotrophic) or inorganic (autotrophic) carbon sources, and exploit light to obtain energy when growing anaerobically<sup>3</sup>. Several interesting features have been observed in this bacterium, such as its consumption of fatty acids, dicarboxylic acids, and aromatic compounds including lignin monomers<sup>4-6</sup>. It is also one of two known bacteria that can express three unique nitrogenases, each with a different transition-metal cofactor<sup>7</sup>. Furthermore, this metabolically versatile strain's genome includes the aerobic and anaerobic pathways for three of the four known strategies that microbes use to break down aromatic compounds, such as lignin breakdown products (LBPs)<sup>8</sup>. Harnessing *R. palustris*' unique metabolic versatilities for the conversion of plant biomass to value-added products, such as polyhydroxybutyrate (PHB)<sup>9</sup>, n-butanol<sup>10</sup>, and hydrogen<sup>11,12</sup>, has garnered increasing interest. However, a system's level understanding of how the bacterium's complex web of metabolic modules operates in response to environmental changes is hindering the development of the PNSB as a biochemical chassis.

Several studies conducted on *R. palustris* have shown that in addition to the Calvin-Benson-Bassham (CBB) cycle's role of carbon assimilation during autotrophic growth, the pathway plays a major role in maintaining redox balance under heterotrophic growth<sup>10,12-14</sup>. It has been shown that heterotrophic growth of the PNSB on substrates that are more reduced than biomass, such as LBPs, is dependent on the availability of an electron sink<sup>13</sup>. CO<sub>2</sub>-fixation using the enzyme ribulose-1,5-biphosphate carboxylase/oxygenase (RuBisCO), nitrogen-fixation through the enzyme nitrogenase<sup>12</sup>, and the supplementation of an electron acceptor (e.g., TMAO)<sup>15</sup> all prevent the inhibitory accumulation of excess reducing agents. The use of CO<sub>2</sub> as a redox balancing strategy for the conversion of plant biomass to value-added products is an attractive

approach that could increase profitability while improving sustainability. However, the complex interplay between the electrons supplied by the catabolism of different carbon sources, CO<sub>2</sub> fixation, and the cyclic electron flow during photosynthesis is not fully understood, thus diminishing the ability to engineer this promising bacterium.

A Genome-Scale Metabolic Model (GSMM) provides a mathematical representation of an organism's metabolic functionalities<sup>16</sup>. The metabolic network represents an organism's available repertoire of biochemical transformations constructed as a stoichiometric matrix<sup>17</sup>. Due to the underdetermined nature of metabolic networks, optimization tools are used to predict reaction rates for a pre-specified objective function such as the maximization of biomass<sup>18</sup>. One of the most commonly applied optimization tools used to model metabolism is Flux Balance Analysis (FBA). FBA performs a pseudo-steady state mass balance for each metabolite in the network to predict the maximum growth rate and corresponding reaction fluxes during the cell's exponential growth phase<sup>19</sup>. Due to the high dimensionality of the network, other tools such as Flux Variability Analysis (FVA) are used to determine the sensitivity of growth rate as a function of each reaction flux<sup>20</sup>. Finally, a modified FBA formulation can be used to predict the set of essential genes under a specified growth condition<sup>21</sup>. Thus far, a limited number of small-scale metabolic reconstructions have been developed for PNSB, examining either the central carbon metabolism<sup>22</sup> or the electron transport chain<sup>23</sup>. However, these models are limited in scope, as they consider less than 4% of the organism's metabolic functionality and are therefore incapable of capturing system-wide interactions between different metabolic modules. Very recently, a GSMM of the bacterium was reconstructed and used to test an array of cellular objectives during phototrophic growth. Anaerobic growth on acetate, benzoate, and 4-hydroxybenzoate was simulated using 8 different biologically relevant objective functions. The model predicted that the organism primarily optimized for growth, ATP production, and metabolic efficiency. However, the model could further be improved by integrating recently annotated metabolic pathways of lignin monomer degradation<sup>24</sup>, and by making use of experimental data on gene essentiality<sup>25</sup> and metabolic flux analysis for growth under different carbon sources<sup>13,14</sup> to validate and refine the network.

In this work, a Genome-Scale Metabolic Model (GSMM) of *R. palustris* was constructed to model the bacterium's metabolic functionality under different environmental conditions. The model was used to simulate growth under different carbon sources and shows excellent agreement with experimentally measured fluxes<sup>13,14</sup>. Gene essentiality analysis was also performed for aerobic and anaerobic growth on acetate. The predicted essential genes were compared with available trans-mutagenesis data<sup>25</sup>, and an accuracy of 84% was achieved. After the model indicated the presence of an unidentified quinol sink, *in silico* simulations were combined with published *in vivo* flux measurements<sup>13,14</sup> to study the effect (and extent) of the quinone redox state on cellular growth, electron transport rate, and CO<sub>2</sub> fixation. These results suggest that redox state acts as a feed-forward controller of the highly energy-demanding CBB cycle by regulating the rate of light-generated ATP. Overall, an understanding of the metabolic control points of this interconnected system constitutes the first step towards engineering strains capable of more efficiently harnessing photosynthetic energy and rerouting this energy towards bio-production and lignin valorization.

## Methods

## **Model reconstruction**

A draft model was first generated in KBase<sup>26</sup> based on *R. palustris*' genome (downloaded from the NCBI database on 04/12/2018). KBase uses annotated features in the genome to construct a list of reactions associated with genes in the organism. Previously published work of the bacterium's metabolic network<sup>22</sup> was used to manually curate pathways from the central carbon metabolism and to ensure correct cofactor usage and gene association. This resulted in an expanded network of high-confidence reactions, all associated with genes in *R. palustris*. Experimentally measured concentrations of biomass components are available for *R. palustris* when grown on acetate<sup>13</sup> and were used to develop the biomass equation (see supplementary file 1). To minimize the addition of low-confidence reactions during gapfilling, the process was broken down into two steps. First, a subset of high-confidence reactions from a recently published genome-scale model of *R. palustris*<sup>27</sup> was added to the draft model. Here, high-confidence reactions are those that are associated with a published source of annotation. At this point, the majority of the reactions required to gapfill the biomass equation existed in partially incomplete linear pathways. Therefore, the ModelSEED database<sup>28</sup> was used to gapfill the model generated in KBase. In addition, annotated metabolic pathways for the breakdown of multiple aromatic compounds including lignin breakdown products were found in literature<sup>24</sup> and in organism-specific biochemical databases<sup>29,30</sup>, and were subsequently added to the model (supplementary file 2). Finally, annotated *R. palustris* genes were mined from three databases (KEGG<sup>29</sup>, BioCyc<sup>30</sup>, and UniProt<sup>31</sup>) to validate the Gene-Protein-Reaction (GPR) associations established in the model and to include GPR relationships for reactions added during the gapfilling process (supplementary file 3).

## **Model simulations**

Parsimonious Flux Balance Analysis (pFBA)<sup>32</sup> was used to simulate growth under different environmental conditions. pFBA is analogous to FBA but adds an outer objective that minimizes the sum of all reaction fluxes. Objective tilting<sup>33</sup> was used to formulate both objectives in one function as shown below.

$$\text{Maximize } v_{biomass} - 0.0001 \sum_{j \in J - v_{biomass}} v_j$$

subject to

$$\sum_{j \in J} S_{ij} \cdot v_j = 0 \quad \forall i \in I \quad (1)$$

$$LB_j \leq v_j \leq UB_j \quad \forall j \in I \quad (2)$$

Where  $I$  and  $J$  are the sets of metabolites and reactions in the model, respectively.  $S_{ij}$  is the stoichiometric coefficient of metabolite  $i$  in reaction  $j$  and  $v_j$  is the flux value of reaction  $j$ . Parameters  $LB_j$  and  $UB_j$  denote the minimum and maximum allowable fluxes for reaction  $j$ , respectively.  $v_{biomass}$  is the flux of the biomass reaction which mimics the cellular growth rate.

## **Model validation**

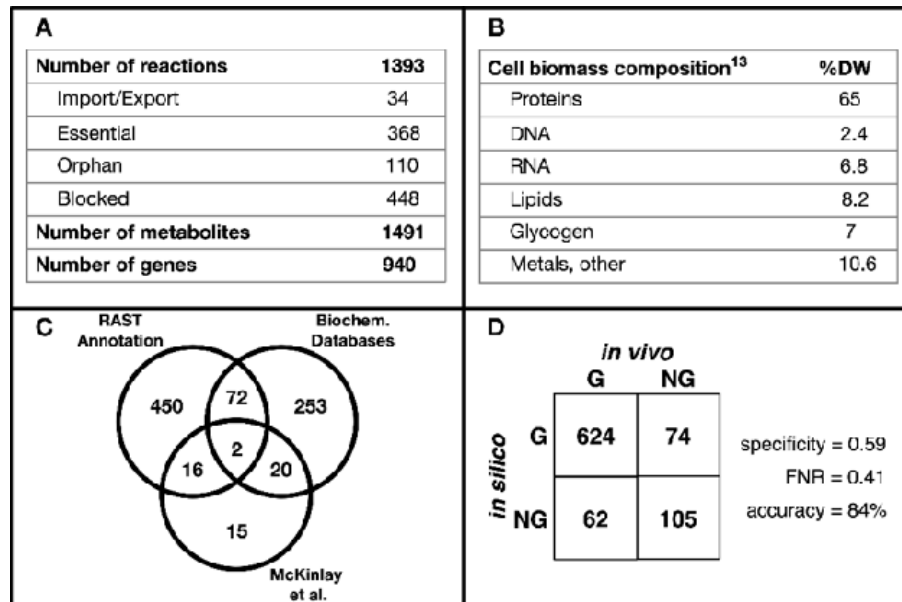
Metabolic Flux Analysis (MFA) measurements from anaerobic growth on acetate<sup>13</sup>, fumarate, succinate, and butyrate<sup>14</sup> were compared with model predicted fluxes. Model accuracy for each growth condition was calculated by taking the sum of percent errors between pFBA predicted and MFA values (see supplementary file 4 for example). In addition, *R. palustris*' essential genes, determined experimentally for aerobic growth on acetate<sup>25</sup>, was used to validate the essential genes predicted by the model. Gene essentiality was predicted in the model by sequentially knocking out each reaction and determining the resulting effect on the biomass reaction rate<sup>21</sup>. If a reaction knockout resulted in a predicted growth rate that was less than 10% of the wild type growth rate, the reaction was considered essential. Reaction GPRs were then used to map the list of essential reactions to essential genes. Finally, the list of experimentally determined essential metabolic genes<sup>25</sup> were compared with model predicted essential genes to determine the specificity and sensitivity of the predictions (see supplementary file 5).

## **Results and Discussion**

### **Model Reconstruction and validation**

A summary of the model's major statistics is shown in Figure 1A. Overall, the 940 genes associated with model reactions account for 62% of the genes involved in energy metabolism, biosynthesis, carbon & nitrogen metabolism, and cellular processes in *R. palustris*' genome<sup>3</sup>. Experimental measurements of biomass component concentrations were obtained for growth on acetate<sup>13</sup> (Figure 1B) and converted into stoichiometric coefficients for the model's biomass equation (see supplementary file 1). Thus, an initial high-confidence model containing 540 genes and 915 reactions with no orphan reactions was constructed. The gap-filling procedure was carried out next in KBase<sup>26</sup> using reactions from the ModelSEED database<sup>28</sup>. Organism-specific biochemical databases including KEGG, UniProt, and BioCyc were then used to annotate added reactions (Figure 1C). This resulted in the addition of 328 annotated and 110 unannotated (orphan) reactions. The inclusion of these reactions is necessary to ensure biomass production. pFBA was used to simulate growth on a number of different carbon sources, including carboxylic acids (acetate, fumarate, succinate and butyrate) and lignin monomers (supplementary file 2). pFBA is analogous to FBA but adds an outer objective that minimizes the sum of all reaction fluxes (see Methods). This is justified by the assumption that cells synthesize the minimum amount of cellular machinery required to maintain the maximal growth rate<sup>32</sup>. Simulating growth using pFBA has two main advantages over FBA. First, pFBA avoids unrealistic flux predictions for reactions participating in thermodynamically infeasible cycles (TICs)<sup>34</sup>. TICs are usually removed from GSMs to avoid false predictions; however, when analyzing highly connected networks like that of *R. palustris*, removing these cycles can lead to the model missing certain functionalities and metabolic modes utilized by the organism. pFBA avoids these false predictions by the additional constraint that reaction fluxes should be minimized. Second, the pFBA formulation results in a significantly reduced set of optimal solutions compared to FBA. Flux Balance Analysis usually results in a large number of alternate optimal solutions (especially in highly connected networks), most of which are not biologically relevant, and can therefore lead to false conclusions<sup>35</sup>. pFBA's additional objective greatly restricts the solution space and leads to more biologically insightful conclusions<sup>32</sup>. Essentiality analysis identified 368 essential reactions, out of which 249 were associated with gene annotations in the model. Comparison with *in vivo* gene essentiality data for aerobic growth on

acetate<sup>25</sup> was then used to check the model accuracy (Figure 1D). The calculated sensitivity and false negative rate (FNR) are consistent with recently published GSMMs<sup>36,37</sup>. Moreover, given that this is a non-model organism with no well-characterized close relatives, high-confidence annotation was not available for less-studied pathways, therefore an automated pipeline like GrowMatch<sup>38</sup> could not be implemented to further improve essentiality predictions.



**Figure 1. Summary of model statistics and validation. (A)** Overall model statistics. **(B)** Model biomass component compositions. **(C)** Sources of gene annotation. **(D)** Gene essentiality analysis results.

### The effect of the quinone pool on light uptake, carbon dioxide fixation, and growth

During initial growth simulations, growth was observed to be hindered due to the accumulation of excess quinols formed in the TCA cycle. Since no high-confidence reaction was found to consume quinols in *R. palustris*, a quinol “sink” reaction was added to the model. Sink reactions are often incorporated into metabolic models when a metabolite is known to be produced during metabolism but for which no means of consumption have been identified<sup>39</sup>, or to describe the accumulation of a storage compound<sup>39</sup> (e.g. glycogen). Furthermore, recent experimental work with *R. palustris* TIE-1 reported the presence of an unidentified quinol-oxidizing reaction that had not been accounted for previously<sup>40</sup>, giving further support to this prediction. To determine the effect of the quinone pool on growth, pFBA simulations were conducted under different quinol sink rates to qualitatively predict how changes in the quinone redox state affected the rest of the metabolic network. The quinol sink reaction was treated as a parameter in the model and pFBA simulations were conducted at varying quinol oxidation (sink) rates to determine how light uptake (i.e. Electron Transport Rate or ETR), growth, and CO<sub>2</sub> fixation are affected by changes in the quinone redox state (Figure 2). Carbon uptake was restricted to a maximum value of 100 mmol/gDW/hr for acetate and 50 mmol/gDW/hr for fumarate, succinate, and butyrate to ensure the same number of carbons were being up taken. MFA values were scaled to the same carbon uptake rates<sup>13,14</sup>. For growth on butyrate, the supplementation of CO<sub>2</sub> is required for growth, as the substrate is more reduced than biomass and requires an electron sink<sup>14</sup>. The media

was supplied with CO<sub>2</sub> at a maximum uptake rate of 32.1 mmol/gDW/hr to match MFA observations. Since steady-state GSMs cannot capture metabolite concentrations, the redox state cannot be calculated directly. Instead, the qualitative behavior of the redox state was predicted by varying the rate of the quinol sink. As the quinol oxidation rate increases, the quinone pool becomes more oxidized. Using experimental MFA data<sup>13,14</sup>, the quinol oxidation rate was predicted for each of the four substrates (Table 1). These values were calculated by minimizing the sum of errors between the *in silico* generated pFBA fluxes and the *in vivo* MFA flux values. The table also shows the quinone reduction rate through the TCA cycle for each carbon source. The percentage of CO<sub>2</sub> fixed was defined as the rate of CO<sub>2</sub> fixation divided by the total rate of CO<sub>2</sub> produced metabolically. Figure 3 shows the resulting flux predictions obtained at the predicted quinol oxidation rates for growth on acetate (figure 3A), and the calculated percent errors of these predictions for each carbon substrate (figure 3B). A comparison of flux predictions with MFA values for the other three carbon sources is provided in supplementary file 2.

For growth on acetate and butyrate, light uptake (i.e. ETR) shows two distinct regions based on the extent of quinol oxidation (Figure 2A). Under low oxidation rates, flux through the quinol-producing succinate dehydrogenase reaction is avoided by using the glyoxylate shunt and subsequently the CBB cycle. Therefore, both light uptake and CO<sub>2</sub> fixation increase rapidly in this region. In the second region, at high quinol oxidation rates, flux shifts toward the oxidative TCA cycle. Therefore, in this region, both the Electron Transport Chain (ETC) activity and the rate of CO<sub>2</sub> fixation decrease with increasing quinol oxidation. Furthermore, as can be seen from Table 1, the ratio of quinol oxidation rate to quinone reduction rate is similar for both carbon sources. Due to the supplementation of CO<sub>2</sub> during growth on butyrate, the percentage of CO<sub>2</sub> fixation could not be calculated. During growth on succinate, the production of quinols through succinate dehydrogenase cannot be avoided, therefore light uptake rate increases linearly with the rate of quinol oxidation. Moreover, the rates of quinol oxidation and quinone reduction are equivalent, indicating that the quinone pool is more reduced when compared to the redox state during growth on acetate and butyrate. This leads to a reduced electron flow through the ETC, and subsequently lower ATP generation. Finally, the model predicts that during growth on the highly oxidized (compared to cell biomass) carbon source fumarate, the rate of the quinol sink does not affect the flux distribution.

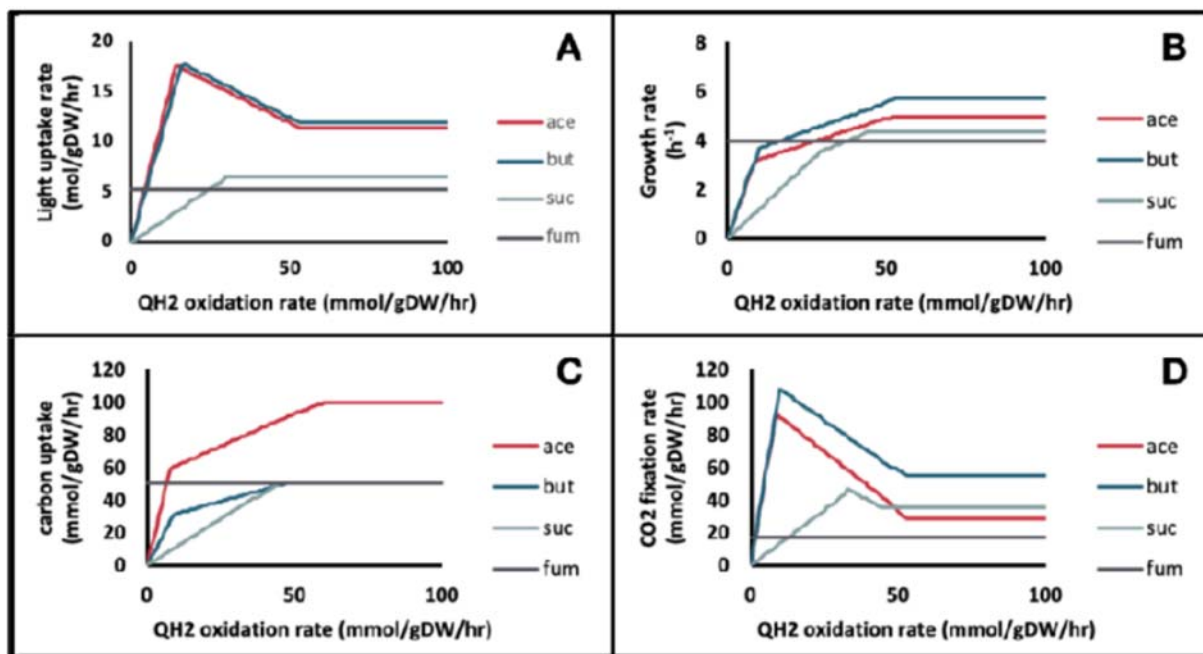


Figure 2. Effect of the Quinol sink rate on: (A) Light uptake rate, (B) Growth rate, (C) Carbon source

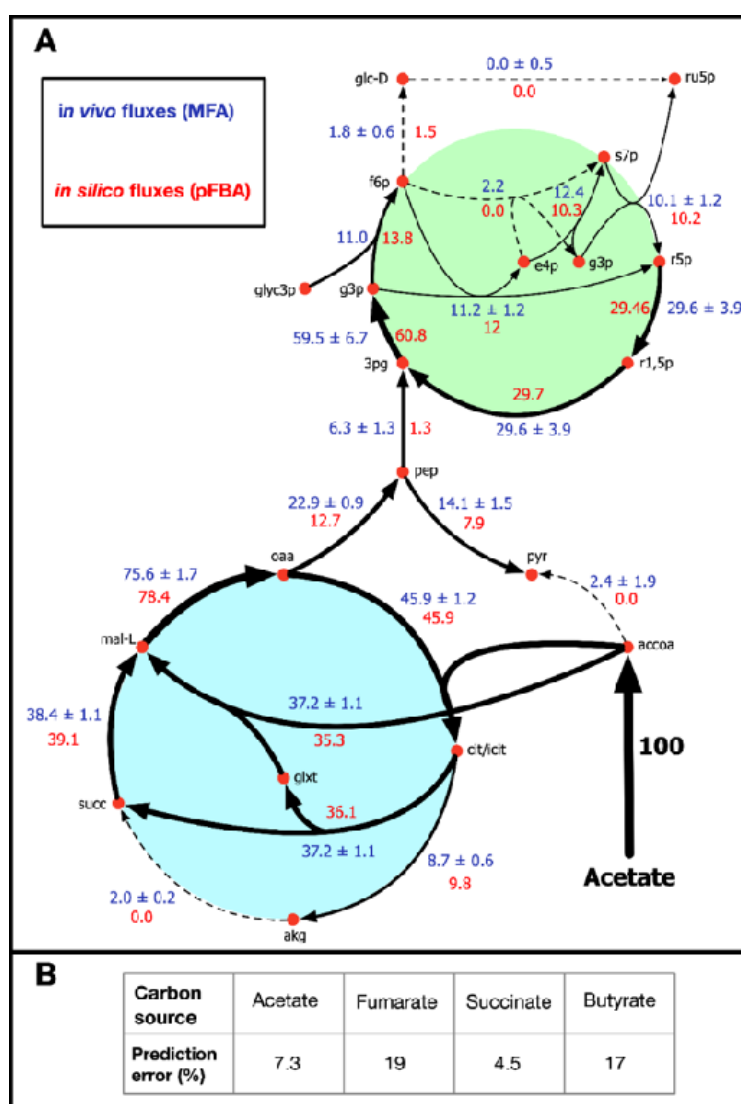
**Table 1.** Predicted reaction rates for growth on four different carbon sources

Carbon source	QH <sub>2</sub> oxidation rate (mmol/gDW/hr)	Q reduction rate <sup>a</sup> (mmol/gDW/hr)	Electron transport rate (mol/gDW/hr)	CO <sub>2</sub> fixation rate (mol/gDW/hr)	% CO <sub>2</sub> fixed <sup>b</sup>	Net CO <sub>2</sub> excretion rate (mmol/gDW/hr)
Acetate	52.5	39.1	5.3	29.7	73.2	10.9
Butyrate	54.9	37.4	5.4	57.8	-	-18.6 <sup>c</sup>
Succinate	49.8	49.2	3.0	35.6	50.6	34.8
Fumarate	0	0	2.3	17.3	25.1	51.5

<sup>a</sup>The rate of quinone reduction in the TCA cycle.

<sup>b</sup>The rate of CO<sub>2</sub> fixation divided by the rate of total CO<sub>2</sub> produced.

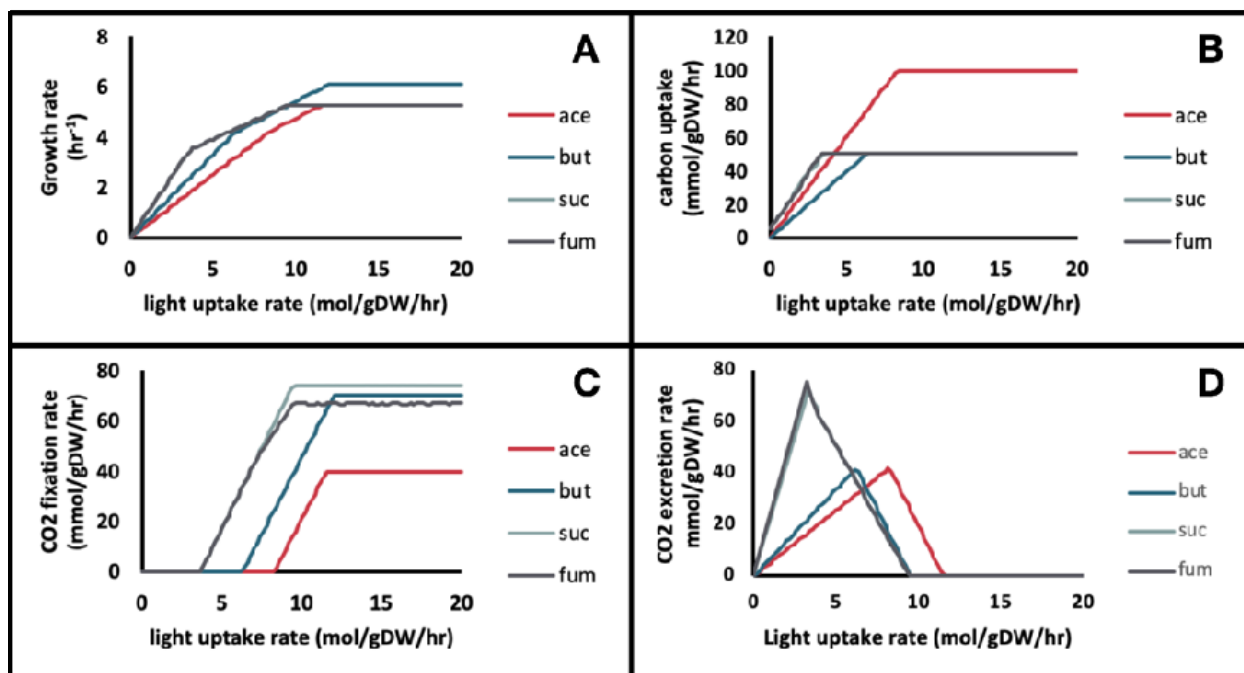
<sup>c</sup>CO<sub>2</sub> was supplied in the media during growth on butyrate.



**Figure 3.** Comparison of model-predicted vs MFA-generated flux values for reactions involved in central carbon metabolism. **(A)** Metabolic flux map showing reaction rates for growth on acetate **(B)** Percentage error



A similar parameter sampling procedure was performed to determine the effect of light uptake on growth. Light uptake rate was set as a parameter and the quinol oxidation rate was fixed to the value predicted based on MFA fluxes (Figure 4). Again, the plots show two distinct growth regions: (i) a low-light (LL) energy-limited region, and (ii) a high-light (HL) carbon-limited region. In the LL region, growth is highly dependent on the amount of light available and the model predicts that all of the ATP produced is used to convert the carbon source into biomass precursors. Therefore, no ATP remains for the energy-intensive CBB pathway. In the HL region, the maximum substrate uptake rate is reached, and the carbon source cannot be incorporated any faster. The additional energy produced from light is then directed towards CO<sub>2</sub> fixation. Although the model predicts that the rate of CO<sub>2</sub> fixation increases linearly with light uptake rate, kinetic and thermodynamic constraints on the highly inefficient CO<sub>2</sub>-fixing RuBisCO enzyme<sup>41</sup> hinder this process at high light uptake.



**Figure 4. Effect of the light sink rate on (A) Growth rate, (B) Carbon source uptake rate, (C) Carbon fixation rate, and (D) Carbon dioxide excretion rate for growth on four carbon sources. Ace: acetate, but: butyrate, suc: succinate, fum: fumarate. In A, B, and D, the lines for succinate and fumarate lie on top of each other.**

### **Proposed mechanism for the interplay between the quinone redox state, the electron transport rate, and CO<sub>2</sub> fixation.**

Based on the effect of the quinol oxidation rate on light uptake and on the model's flux distribution, a mechanistic explanation of the system-wide metabolic interactions can be postulated. As shown in Figure 5, increased flux through the oxidative TCA cycle leads to the accumulation of reduced quinols. This in turn leads to a restriction in the flow of electrons

through the ETC and consequently in the amount of ATP produced. The CBB system thus lacks the energy required to fix CO<sub>2</sub>. Therefore, the quinone redox state is predicted to act as a feed-forward controller to the energetically expensive CBB pathway, indicating how much ATP is available at a given condition.

Comparison of pFBA generated growth simulations with MFA data leads to the hypothesis that an unidentified quinone:oxidoreductase reaction has to occur to obtain the observed flux distribution. A previous study on the PNSB *R. capsulatus* suggests that complex I, the NADH:quinone oxidoreductase enzyme, is responsible for the observed flux through reverse electron flow<sup>42</sup>. However, the model predicts that the rate of quinol oxidation required cannot be accounted for through complex I only, which showed low activity. Furthermore, based on the high thermodynamic cost of reverse electron flow, it appears unlikely that it can account for the predicted rate of quinol oxidation<sup>23</sup>.

Although the source of quinol oxidation (sink) has yet to be identified, there are a number of candidate reactions that can carry out this role. Primarily, the malate:quinone dehydrogenase (MDH) appears to be a potential reaction for oxidizing excess quinols. In the forward direction, this reaction converts malate into oxaloacetate and produces ubiquinol in the process. A second NAD-dependent malate dehydrogenase is also coded for by *R. palustris* that could perform the same function. Knocking out and over-expressing these enzymes could be employed to investigate their role in ETR, ATP production, and CO<sub>2</sub> fixation.

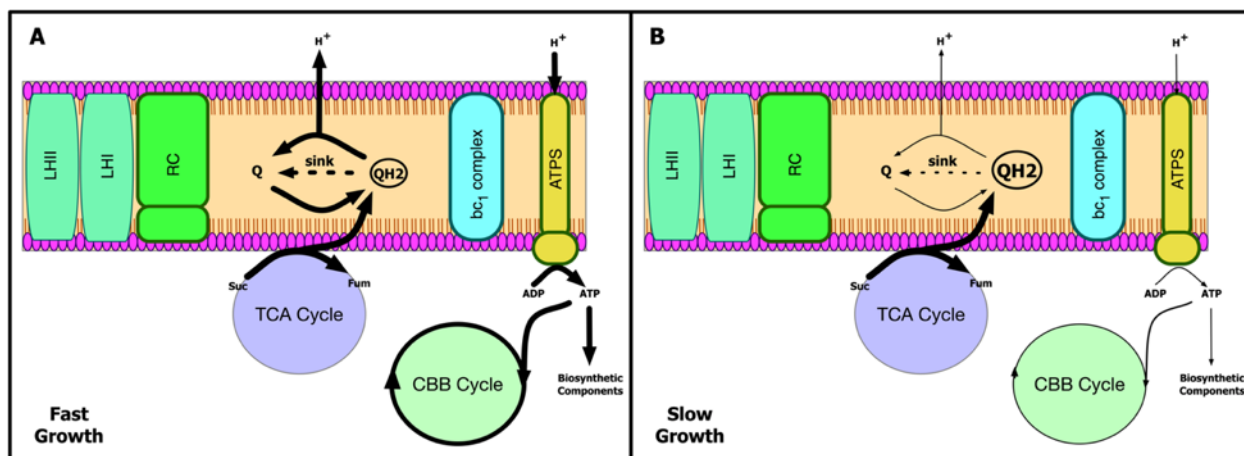


Figure 5. Schematic of a proposed mechanism for the interaction between the quinone redox state, electron transport rate, and carbon fixation. (A) High rate of quinol oxidation. (B) Low rate of quinol

## Conclusion

In this study, a genome-scale metabolic network was used to propose a system-wide mechanistic model of the interactive system that includes photosynthesis, carbon dioxide fixation, and the quinone redox state. The model was validated using experimental genome essentiality data<sup>25</sup> (84% accuracy), and flux measurement data<sup>13,14</sup> on four carbon sources (5-19% prediction error). Model simulations predict the presence of an unidentified quinol sink. Predictions also indicate that the extent of CO<sub>2</sub> fixation is dependent on the amount of ATP present, with the quinone

redox state acting as a feed forward signal to the CBB system. Going forward, the proposed mechanism can be used to generate strategies for engineering strains capable of more efficiently harnessing photosynthetic energy, and that have the ability to reroute energy towards bio-production and lignin valorization. Future experimental work will be conducted to measure the electron transport rate, intracellular ATP concentration, and RuBisCO gene expression across different quinone redox states to strengthen the proposed hypothesis and further refine the model.

## References

- 1 McEwan, A. G. Photosynthetic electron transport and anaerobic metabolism in purple non-sulfur phototrophic bacteria. *Antonie Van Leeuwenhoek* **66**, 151-164 (1994).
- 2 Tabita, F. R. in *Anoxygenic Photosynthetic Bacteria* (eds Robert E. Blankenship, Michael T. Madigan, & Carl E. Bauer) 885-914 (Springer Netherlands, 1995).
- 3 Larimer, F. W. *et al.* Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodospseudomonas palustris*. *Nat Biotechnol* **22**, 55-61, doi:10.1038/nbt923 (2004).
- 4 Oda, Y. *et al.* Multiple genome sequences reveal adaptations of a phototrophic bacterium to sediment microenvironments. *Proceedings of the National Academy of Sciences* **105**, 18543-18548, doi:10.1073/pnas.0809160105 (2008).
- 5 Harwood, C. S. & Gibson, J. Uptake of benzoate by *Rhodospseudomonas palustris* grown anaerobically in light. *J Bacteriol* **165**, 504-509, doi:10.1128/jb.165.2.504-509.1986 (1986).
- 6 Harwood, C. S. & Gibson, J. Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodospseudomonas palustris*. *Appl Environ Microbiol* **54**, 712-717 (1988).
- 7 Oda, Y. *et al.* Functional Genomic Analysis of Three Nitrogenase Isozymes in the Photosynthetic Bacterium *Rhodospseudomonas palustris*. *Journal of Bacteriology* **187**, 7784-7794, doi:10.1128/jb.187.22.7784-7794.2005 (2005).
- 8 Fuchs, G., Boll, M. & Heider, J. Microbial degradation of aromatic compounds - from one strategy to four. *Nat Rev Microbiol* **9**, 803-816, doi:10.1038/nrmicro2652 (2011).
- 9 Ranaivoarisoa, T. O., Singh, R., Rengasamy, K., Guzman, M. S. & Bose, A. Towards sustainable bioplastic production using the photoautotrophic bacterium *Rhodospseudomonas palustris* TIE-1. *Journal of industrial microbiology & biotechnology*, doi:10.1007/s10295-019-02165-7 (2019).
- 10 Doud, D. F. R. *et al.* Metabolic engineering of *Rhodospseudomonas palustris* for the obligate reduction of n-butyrate to n-butanol. *Biotechnol Biofuels* **10**, 178, doi:10.1186/s13068-017-0864-3 (2017).
- 11 Scoma, A., Giannelli, L., Faraloni, C. & Torzillo, G. Outdoor H<sub>2</sub> production in a 50-L tubular photobioreactor by means of a sulfur-deprived culture of the microalga *Chlamydomonas reinhardtii*. *J Biotechnol* **157**, 620-627, doi:10.1016/j.jbiotec.2011.06.040 (2012).

- 12 Muzziotti, D., Adessi, A., Faraloni, C., Torzillo, G. & De Philippis, R. H<sub>2</sub> production in *Rhodospseudomonas palustris* as a way to cope with high light intensities. *Res Microbiol* **167**, 350-356, doi:10.1016/j.resmic.2016.02.003 (2016).
- 13 McKinlay, J. B. & Harwood, C. S. Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. *Proc Natl Acad Sci U S A* **107**, 11669-11675, doi:10.1073/pnas.1006175107 (2010).
- 14 McKinlay, J. B. & Harwood, C. S. Calvin cycle flux, pathway constraints, and substrate oxidation state together determine the H<sub>2</sub> biofuel yield in photoheterotrophic bacteria. *MBio* **2**, doi:10.1128/mBio.00323-10 (2011).
- 15 Schultz, J. E. & Weaver, P. F. Fermentation and anaerobic respiration by *Rhodospirillum rubrum* and *Rhodospseudomonas capsulata*. *J Bacteriol* **149**, 181-190 (1982).
- 16 Fondi, M. & Lio, P. Genome-scale metabolic network reconstruction. *Methods Mol Biol* **1231**, 233-256, doi:10.1007/978-1-4939-1720-4\_15 (2015).
- 17 Saha, R., Chowdhury, A. & Maranas, C. D. Recent advances in the reconstruction of metabolic models and integration of omics data. *Curr Opin Biotechnol* **29**, 39-45, doi:10.1016/j.copbio.2014.02.011 (2014).
- 18 Lachance, J. C. *et al.* BOFdat: Generating biomass objective functions for genome-scale metabolic models from experimental data. *PLoS Comput Biol* **15**, e1006971, doi:10.1371/journal.pcbi.1006971 (2019).
- 19 Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nat Biotechnol* **28**, 245-248, doi:10.1038/nbt.1614 (2010).
- 20 Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489, doi:10.1186/1471-2105-11-489 (2010).
- 21 Joyce, A. R. & Palsson, B. O. Predicting gene essentiality using genome-scale in silico models. *Methods Mol Biol* **416**, 433-457, doi:10.1007/978-1-59745-321-9\_30 (2008).
- 22 Hadicke, O., Grammel, H. & Klamt, S. Metabolic network modeling of redox balancing and biohydrogen production in purple nonsulfur bacteria. *BMC Syst Biol* **5**, 150, doi:10.1186/1752-0509-5-150 (2011).
- 23 Klamt, S., Grammel, H., Straube, R., Ghosh, R. & Gilles, E. D. Modeling the electron transport chain of purple non-sulfur bacteria. *Mol Syst Biol* **4**, 156, doi:10.1038/msb4100191 (2008).
- 24 Austin, S. *et al.* Metabolism of Multiple Aromatic Compounds in Corn Stover Hydrolysate by *Rhodospseudomonas palustris*. *Environ Sci Technol* **49**, 8914-8922, doi:10.1021/acs.est.5b02062 (2015).
- 25 Pechter, K. B., Gallagher, L., Pyles, H., Manoil, C. S. & Harwood, C. S. Essential Genome of the Metabolically Versatile Alphaproteobacterium *Rhodospseudomonas palustris*. *Journal of bacteriology* **198**, 867-876, doi:10.1128/jb.00771-15 (2015).
- 26 Arkin, A. P. *et al.* KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* **36**, 566-569, doi:10.1038/nbt.4163 (2018).
- 27 Navid, A., Jiao, Y., Wong, S. E. & Pett-Ridge, J. System-level analysis of metabolic trade-offs during anaerobic photoheterotrophic growth in *Rhodospseudomonas palustris*. *BMC Bioinformatics* **20**, 233, doi:10.1186/s12859-019-2844-z (2019).
- 28 Henry, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology* **28**, 977, doi:10.1038/nbt.1672

<https://www.nature.com/articles/nbt.1672#supplementary-information> (2010).

- 29 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
- 30 Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **44**, D471-480, doi:10.1093/nar/gkv1164 (2016).
- 31 Wu, C. H. *et al.* The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res* **34**, D187-191, doi:10.1093/nar/gkj161 (2006).
- 32 Lewis, N. E. *et al.* 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 (2010).
- 33 Feist, A. M. *et al.* Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. *Metab Eng* **12**, 173-186, doi:10.1016/j.ymben.2009.10.003 (2010).
- 34 Schellenberger, J., Lewis, N. E. & Palsson, B. O. Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophys J* **100**, 544-553, doi:10.1016/j.bpj.2010.12.3707 (2011).
- 35 Segre, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* **99**, 15112-15117, doi:10.1073/pnas.232349399 (2002).
- 36 Norsigian, C. J., Kavvas, E., Seif, Y., Palsson, B. O. & Monk, J. M. iCN718, an Updated and Improved Genome-Scale Metabolic Network Reconstruction of Acinetobacter baumannii AYE. *Front Genet* **9**, 121, doi:10.3389/fgene.2018.00121 (2018).
- 37 Seif, Y. *et al.* A computational knowledge-base elucidates the response of Staphylococcus aureus to different media types. *PLoS Comput Biol* **15**, e1006644, doi:10.1371/journal.pcbi.1006644 (2019).
- 38 Kumar, V. S. & Maranas, C. D. GrowMatch: an automated method for reconciling in silico/in vivo growth predictions. *PLoS Comput Biol* **5**, e1000308, doi:10.1371/journal.pcbi.1000308 (2009).
- 39 Thiele, I. & Palsson, B. O. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* **5**, 93-121, doi:10.1038/nprot.2009.203 (2010).
- 40 Guzman, M. S. *et al.* Phototrophic extracellular electron uptake is linked to carbon dioxide fixation in the bacterium Rhodospseudomonas palustris. *Nat Commun* **10**, 1355, doi:10.1038/s41467-019-09377-6 (2019).
- 41 Badger, M. R. & Bek, E. J. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO<sub>2</sub> acquisition by the CBB cycle. *J Exp Bot* **59**, 1525-1541, doi:10.1093/jxb/erm297 (2008).
- 42 Tichi, M. A., Meijer, W. G. & Tabita, F. R. Complex I and its involvement in redox homeostasis and carbon and nitrogen metabolism in Rhodobacter capsulatus. *J Bacteriol* **183**, 7285-7294, doi:10.1128/JB.183.24.7285-7294.2001 (2001).

## Acknowledgement

Funding to support this work was provided by University of Nebraska-Lincoln Faculty Startup Grant to Rajib Saha. We thank all members of the Systems and Synthetic Biology Laboratory for collegial discussions.

### **Author contributions**

A.A., C.I., and R.S. conceived the idea of the project. A.A. reconstructed the model and conducted computational analysis. All authors designed research, wrote and edited the manuscript.

### **Competing Interests**

The authors declare no competing interests.

### **Data Availability**

All data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

### **Supplementary information**

1. Supplementary file 1: Biomass equation components
2. Supplementary file 2: Lignin monomer breakdown pathways and flux predictions for growth on butyrate, succinate, and fumarate.
3. Supplementary file 3: Model files for *R. palustris*
4. Supplementary file 4: Model flux predictions accuracy calculations
5. Supplementary file 5: Gene essentiality analysis results