

A High-Quality Genome-Scale Model for *Rhodococcus opacus* Metabolism

Garrett W. Roell, Christina Schenk, Winston E. Anthony, Rhiannon R. Carr, Aditya Ponukumati,
Joonhoon Kim, Elena Akhmatskaya, Marcus Foston, Gautam Dantas, Tae Seok Moon,* Yinjie J. Tang,*
and Hector García Martín*



Cite This: <https://doi.org/10.1021/acssynbio.2c00618>



Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: *Rhodococcus opacus* is a bacterium that has a high tolerance to aromatic compounds and can produce significant amounts of triacylglycerol (TAG). Here, we present iGR1773, the first genome-scale model (GSM) of *R. opacus* PD630 metabolism based on its genomic sequence and associated data. The model includes 1773 genes, 3025 reactions, and 1956 metabolites, was developed in a reproducible manner using CarveMe, and was evaluated through Metabolic Model tests (MEMOTE). We combine the model with two Constraint-Based Reconstruction and Analysis (COBRA) methods that use transcriptomics data to predict growth rates and fluxes: E-Flux2 and SPOT (Simplified Pearson Correlation with Transcriptomic data). Growth rates are best predicted by E-Flux2. Flux profiles are more accurately predicted by E-Flux2 than flux balance analysis (FBA) and parsimonious FBA (pFBA), when compared to 44 central carbon fluxes measured by ¹³C-Metabolic Flux Analysis (¹³C-MFA). Under glucose-fed conditions, E-Flux2 presents an R^2 value of 0.54, while predictions based on pFBA had an inferior R^2 of 0.28. We attribute this improved performance to the extra activity information provided by the transcriptomics data. For phenol-fed metabolism, in which the substrate first enters the TCA cycle, E-Flux2's flux predictions display a high R^2 of 0.96 while pFBA showed an R^2 of 0.93. We also show that glucose metabolism and phenol metabolism function with similar relative ATP maintenance costs. These findings demonstrate that iGR1773 can help the metabolic engineering community predict aromatic substrate utilization patterns and perform computational strain design.

KEYWORDS: ATP maintenance, genome-scale models, omics data, ¹³C-metabolic flux analysis, predictive biology

1. INTRODUCTION

Rhodococcus opacus PD630 (hereafter, *R. opacus*) is a Gram-positive aerobic bacterium known for its pronounced ability to produce triacylglycerol, a biofuel precursor, from aromatic monomers.^{1,2} *R. opacus* can be used as a 'biological funnel' to convert heterogeneous mixtures of aromatic compounds from the thermal or catalytic deconstruction of lignin into lipid-based biofuels.³ Its natural tolerance toward the aromatic compounds from lignin deconstruction is partially attributed to a high-flux β -ketoadipate pathway that facilitates aromatic catabolism. The β -ketoadipate pathway converts aromatic compounds into acetyl-CoA and succinyl-CoA,⁴ both of which enter central metabolism via the TCA cycle. High TCA cycle flux produces large amounts of ATP and NADH, and as a result, *R. opacus* can synthesize highly reduced products.^{2,5}

Previous work on *R. opacus* has identified aromatic tolerance and utilization mechanisms based on transcript profile changes that do not cause large amounts of flux rewiring and that are not dependent on many genetic mutations. The transcriptome and fluxome of the wild type were examined when grown with sugars and model lignin monomers (i.e., aromatics) for a base strain as well as for adaptively evolved mutants.^{5–7} A key finding is that

the adaptive mutants could achieve optical densities (OD_{600}) up to 1900% higher than the wild-type strain when grown on high concentrations of aromatics, despite a limited number of mutations (~12 single nucleotide polymorphisms on average) and limited flux rewiring.^{5,7} The mutants, however, show big differences in their transcriptomic profiles when compared to the wild-type strain, which may account for their abilities to tolerate and utilize higher concentrations of aromatics.^{6,7} In addition, the molecule-level mechanisms for aromatic substrate utilization and regulation have been elucidated.⁸ Despite these advances in understanding the metabolism and gene regulation in *R. opacus*, a predictive genome-scale model derived from its genome has yet to be developed.

Received: November 16, 2022

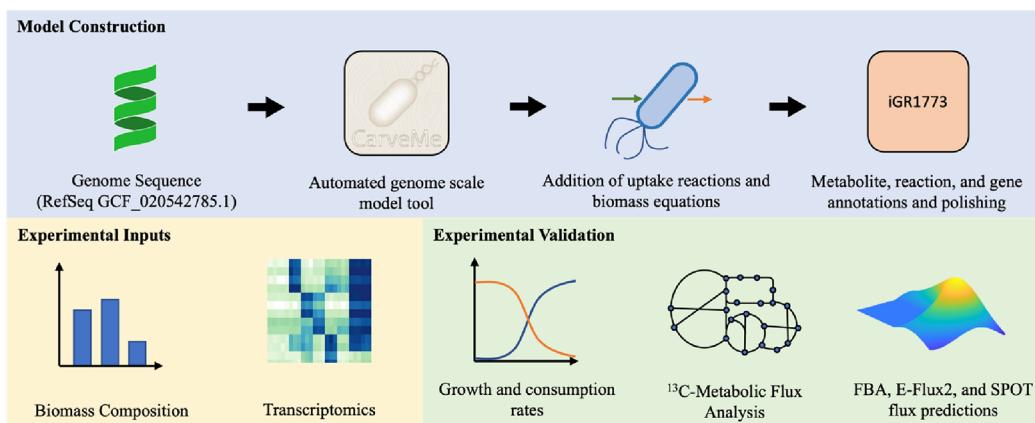


Figure 1. Reconstruction details and model validation. A draft version of the model was created through CarveMe, which was then augmented with relevant uptake and biomass reactions and then manually curated to yield the iGR1773 *R. opacus* GSM.

59 Genome-scale models (GSMs) are comprehensive mathematical summaries of the reactions encoded in an organism's
60 genome. For example, flux balance analysis (FBA) uses GSM to
61 optimize metabolic fluxes through mass balance constraints
62 under the assumption that these fluxes maximize biomass
63 production (i.e., produce the maximum growth rate).⁹ The FBA
64 method has been successful when modeling fast-growing lab-
65 adapted species, but it is less accurate for organisms with slower
66 growth rates.¹⁰ Using data reflecting the internal state of the cell
67 (e.g., omics data) is expected to improve the accuracy of flux
68 predictions. In contrast to input and output flux measurements,
69 omics data are not as straightforward to integrate.¹¹ A variety of
70 Constraint-Based Reconstruction and Analysis (COBRA)
71 methods that integrate omics data have been developed
72 including iMAT,^{12,13} GIMME,^{14,15} E-Flux,¹⁶ E-Flux2 and
73 SPOT,¹⁷ tFBA,¹⁸ GX-FBA,¹⁹ FCG,²⁰ and CORDA.²¹ Such
74 methods may be used to leverage high throughput tran-
75 scriptomics data to improve model predictions.^{22–25} There is,
76 however, no 'best' method to guarantee the most accurate
77 predictions under all circumstances, so care must be taken to
78 identify differences, benefits, and drawbacks of each prediction
79 method in order to apply the method that is most suited to a
80 particular system.²⁶

82 Here, we present and validate iGR1773, the first GSM for *R.*
83 *opacus* derived from its genome, providing a comprehensive
84 description of its internal metabolism and a valuable tool to
85 integrate omics data into metabolic flux predictions. iGR1773
86 consists of 3025 reactions and 1956 metabolites obtained from
87 annotating its completed genome,²⁷ adding the corresponding
88 metabolic reactions, and testing the predictions derived by it.
89 Although previous publications have reported a GSM for *R.*
90 *opacus* PD630,^{28,29} this model did not use an annotation of the
91 *R. opacus* PD630 genome. This model repurposed the
92 *Rhodococcus jostii* GSM³⁰ by doing some minor modifications
93 including setting fluxes to polyhydroxyalkanoates (PHA),
94 polyhydroxyvalerates (PHV), glycogen, and acetate to zero
95 and adjusting the TAG reaction to reflect the fatty acid
96 composition of *R. opacus* PD630.²⁸ Notably, iGR1773 was
97 validated in three different ways: via the Metabolic Model Test
98 (MEMOTE) suite,³¹ by checking growth rate predictions, and
99 by comparing flux predictions to ¹³C-metabolic flux analysis
100 (¹³C-MFA) results. Flux and growth rate predictions from the
101 model were obtained through several COBRA methods,
102 including parsimonious FBA (pFBA) and two methods that
103 integrate transcriptomic data: E-Flux2 and SPOT. Briefly,

Flux2 uses transcript measurements as upper and lower bounds for flux values, and SPOT finds the maximum correlation between transcript levels and reaction rates. These methods were chosen because the solutions they produce are non-degenerate, and they have been validated by previous studies.¹⁷ We found that of the three COBRA methods, E-Flux2 provided the best predictions for growth rates and central carbon fluxes, providing, with iGR1773, an accurate predictive method for future *R. opacus* studies.

2. RESULTS AND DISCUSSION

iGR1773 was created through CarveMe³² and manually curated by refining the reversibility of two reactions based on thermodynamics and adding transport reactions needed for ATP synthesis. MEMOTE³¹ was used to ascertain the quality of the reconstruction, testing on par with state-of-the-art models. We tested iGR1773's predictive capabilities in two different ways: by comparing quantitative predictions of growth rates with experimentally measured growth rates and by comparing flux predictions with ¹³C-MFA measurements. Growth rate and flux predictions were obtained through three different methods: pFBA, EFlux-2, and SPOT. FBA works by providing the fluxes that maximize biomass production whereas pFBA adds an extra step, in which the sum of squared fluxes is minimized while the biomass production flux is held at its maximum. EFlux-2 and SPOT work differently: they do not assume maximum biomass production but constrain fluxes based on transcriptomic measurements. E-Flux2 determines fluxes by solving a transcript-adjusted FBA problem, and SPOT constrains fluxes by maximizing the correlation between fluxes and transcripts. Additionally, ¹³C-MFA and pFBA were used to determine that phenol and glucose metabolisms operate at roughly the same maintenance cost (i.e., similar amounts of ATP are lost to non-growth purposes per mmol of substrate consumed).

2.1. Model Attributes and Refinement of Draft Reconstruction. iGR1773 was generated from a recent genome annotation²⁷ and the genome-to-GSM tool CarveMe³² (Figure 1). The draft model produced by CarveMe was accurate but required manual changes: two reactions needed to have their flux bounds adjusted to match known thermodynamic patterns. In the draft model, the succinate dehydrogenase reaction (EC 1.3.5.1; succinate + FAD \leftrightarrow fumarate + FADH₂) allowed flux only in the reverse direction. Based on ¹³C data demonstrating a complete TCA cycle in the forward direction,⁵ this reaction was allowed to have forward and reverse flux. Additionally, the draft

model contained a thermodynamically infeasible cycle that allowed the model to produce unrealistic amounts of ATP. This flaw was traced to two versions of 3-hydroxyadipyl-CoA dehydrogenase (EC 1.1.1.35): one version of the reaction was 3-oxoacyl-CoA + NADH + H⁺ ↔ 3-hydroxyacyl-CoA + NAD⁺ and the other version was 3-hydroxyacyl-CoA → 3-oxoacyl-CoA + H₂. When combined, this reaction pair has the net effect of converting NADH and H⁺ into H₂ and NAD⁺. The resultant H₂ could then be used to pump H⁺ into the periplasm by a hydrogenase reaction (EC 1.12.5.1; H₂ + 2H⁺_{cytosolic} + menaquinone → 2H⁺_{periplasm} + menaquinol), with subsequent periplasmic H⁺ used to drive ATP synthase to produce an unrealistic quantity of ATP. The reaction of 3-hydroxyacyl-CoA → 3-oxoacyl-CoA + H₂ was blocked to prevent this loop from generating ATP. Four reactions were added to the draft model to allow hydrogen ions travel to the periplasm to drive ATP synthase flux. These reactions included cytochrome b6/f complex periplasm, active co2 transporter facilitator (periplasm), cytochrome c oxidase, and cytochrome oxidase bd. These reactions allow reduced energy-carrying molecules, like plastoquinol and ferrocyanochrome, to participate in moving hydrogen ions to the periplasm. After these manual changes, the finalized model contained 3025 reactions and 1956 metabolites (Table 1).

Table 1. iGR1773 Model Statistics

Genes	
total genes	1773
Reactions	
total reactions	3025
transport reactions	824
purely metabolic reactions	1862
Metabolites	
total metabolites	1956
Model Properties	
metabolic coverage	1.71
degrees of freedom	847
compartments	3

2.2. Model Evaluation through MEMOTE. The *R. opacus* GSM was evaluated with MEMOTE,³¹ producing a score commensurate with the best in the field. MEMOTE addresses the problem of assessing the quality of GSMSs, given their complexity (GSMS often include thousands of metabolites and reactions that are assigned to subcellular locations). Adequate model quality tests are critical because mass balance or stoichiometric errors can render erroneous model predictions. The annotated and curated model was determined to have 100% stoichiometric consistency, 100% mass balance, and 100% metabolite connectivity. The annotation scores consist of 79% for metabolites, 77% for reactions, 33% for genes, and 100% for SBO (systems biology ontology). MEMOTE scores are designed to reflect the average completeness of annotations across databases since there are multiple databases for genome-scale model data (e.g., BiGG and KEGG). For each category (e.g., metabolites, reactions, and genes), a score is calculated for each database as a percentage of the category members that contain an annotation corresponding to that database. The overall MEMOTE score for the category is calculated by averaging the database-specific annotation scores. The overall

score for the model was 91%. As a reference, a recent *E. coli* GSM, iML1515, has an overall MEMOTE score of 91%.³³

2.3. Experimental Calculation of Growth Parameters.

R. opacus grown in glucose showed a significantly higher substrate uptake rate ($P < 0.001$, two-tailed Student's *t* test) and yield than when it was grown in phenol ($P < 0.001$, two-tailed Student's *t* test). Paired sets of time course growth and consumption curves were used to determine the growth rate, yield coefficient, and substrate uptake rate of wild type *R. opacus* when grown on phenol or glucose, and for an adapted mutant strain, PVHG6, when grown on phenol. The fitted parameters (Table 2) were confirmed by plotting fitted growth and

Table 2. Fitted Growth Parameters for Wild-Type (WT) and Aromatic-Adapted (PVHG6) Strains^a

	growth rate	yield coefficient	substrate uptake rate
WT phenol	0.065 ± 0.001	0.048 ± 0.005	1.4 ± 0.2
PVHG6 phenol	0.080 ± 0.003	0.040 ± 0.002	2.0 ± 0.1
WT glucose	0.260 ± 0.005	0.073 ± 0.004	3.6 ± 0.2

^aGrowth rate has units of h⁻¹, yield coefficient has units of g biomass/mmol substrate, and substrate uptake rate has units of mmol substrate/g biomass/h. All values are averages and standard deviations of three biological replicates.

consumption curves against measured data (Figures S1–S3). The higher uptake rate and yield contributed to the faster growth rate of *R. opacus* in glucose than in phenol. The aromatic adapted strain, PVHG6, had a faster growth rate in phenol than the wild-type strain ($P = 0.002$, two-tailed Student's *t* test). The mutant was developed through ~30 passages of *R. opacus* grown on a mixture of aromatic substrates including phenol. This process selected for mutations that increased growth rate, so the observed difference between WT and PVHG6 was expected. While the mutant's growth rate in phenol was higher than that of WT, the biomass yield showed no difference between the two strains ($P = 0.09$, two-tailed Student's *t* test).

2.4. Growth Rate Predictions. iGR1773 predicted growth rates in an acceptable, but by no means perfect, manner (Figure 2). The method that provided the most accurate predictions was E-Flux2, with SPOT generating the least accurate predictions. pFBA produced predictions that were somewhat less accurate than those provided by E-Flux2. The fact that enzyme constraints increase the accuracy of growth rate predictions over unbounded pFBA is consistent with recent reports from *Saccharomyces cerevisiae* genome-scale modeling.^{34,35} Growth rates under phenol were lower and better-predicted than those under glucose.

E-Flux2 made the most accurate growth rate predictions, while the other methods either displayed larger errors (pFBA) or completely failed (SPOT) (Figure 2). It is not surprising to see SPOT predict null growth rates since it is based on maximizing the correlation between fluxes and transcripts and not maximizing growth. pFBA and E-Flux2 both typically predict faster growth rates than those that have been measured experimentally. pFBA is expected to overestimate growth rates by aiming to predict the maximum theoretical growth rate. We would expect that the actual growth rate would be less than the theoretical maximum due to other factors. For example, soil bacteria such as *R. opacus* need to consume many carbon sources, and maintaining this ability imposes a cost on the growth rate for any one carbon source. Additionally, pFBA seeks out the most efficient use of carbon resources and does not

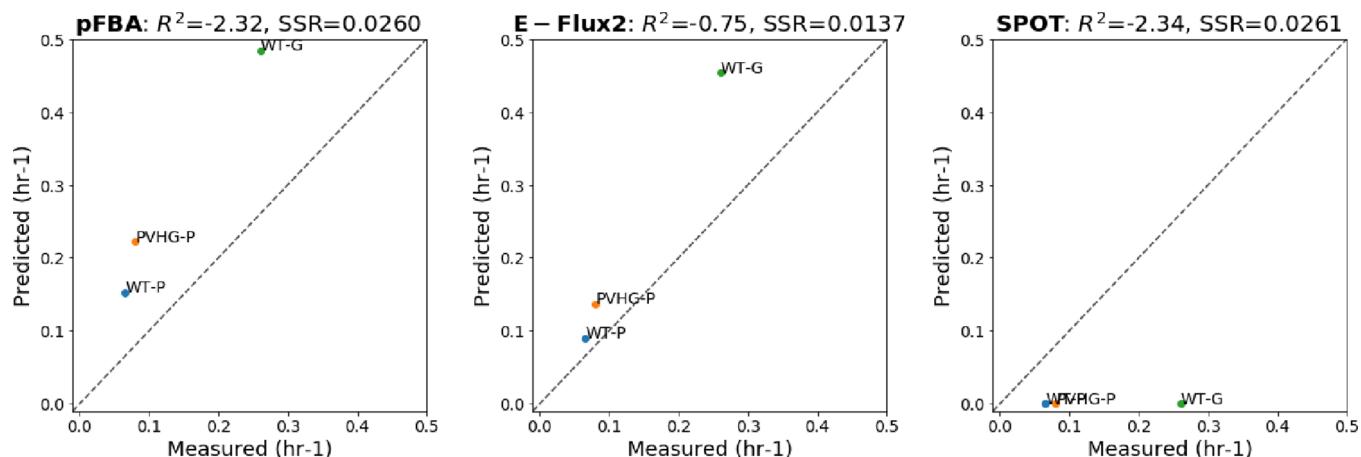


Figure 2. Growth rate predictions. Growth rate predictions are acceptable, but not perfect. Comparison of observed growth rates and model predicted growth rates for wild type consuming glucose (WT-G), wild type consuming phenol (WT-P), and aromatic-adapted strain consuming phenol (PVHG-P). SPOT completely fails. The points represent growth rates with units (h^{-1}). SSR = sum of squared residuals.

242 factor in competing interests, including the cost to make the
243 enzymes. Since enzyme cost is not included in pFBA
244 calculations, pathways with high carbon efficiency are preferred
245 even though these pathways may have low *in vivo* flux due to the
246 overall resource cost in producing the corresponding enzymes.³⁶

247 Growth rates under phenol were lower, and better predicted,
248 than growth rates under glucose. Typically, carbon sources that
249 are consumed through the TCA cycle (e.g., acetate, succinate,
250 and fumarate) result in lower growth rates than for growth on
251 sugars since TCA cycle metabolites are generally more oxidized
252 than sugars. Additionally, when TCA cycle metabolites are used
253 as sole carbon sources, gluconeogenesis is required to produce
254 amino acid precursors. Unlike glycolysis, which produces energy
255 molecules, gluconeogenesis consumes ATP and NADH.
256 Furthermore, phenol is a toxic substance, which imposes an
257 additional metabolic burden via stress response.

258 A possible explanation for why the growth rate predictions are
259 better for phenol than for glucose is that there is only one
260 catabolic pathway for phenol while there are multiple options for
261 glucose. Specifically, phenol degradation into TCA cycle
262 metabolites has low degrees of freedom. Conversely, there are
263 multiple pathways for glucose catabolism, including glycolysis
264 (EMP), Entner–Doudoroff (ED), and pentose phosphate
265 pathways. These pathways can be flexibly regulated and are
266 underdetermined by pFBA.

267 **2.5. Comparison of Model Predictions and ^{13}C -MFA**
268 **Fluxes.** When compared to fluxes measured by ^{13}C -MFA, the
269 flux predictions from the COBRA methods were more accurate
270 for phenol metabolism than for glucose metabolism. Among the
271 COBRA methods we tried, EFlux-2 provided the best
272 predictions, whereas SPOT provided the worst predictions for
273 the phenol uptake case but the second best for glucose. pFBA
274 provided the same results as FBA, which were very good for
275 phenol but not very accurate for glucose. The comparison of
276 predicted fluxes with ^{13}C -MFA flux measurements is the most
277 rigorous test of GSM and COBRA methods since ^{13}C -MFA
278 measurements are the gold standard for quantifying intracellular
279 reaction rates,³⁷ and they provide detailed information about
280 central metabolism instead of aggregated measurements (e.g.,
281 just growth rate). ^{13}C -MFA, however, is an expensive procedure
282 to carry out.³⁸ Thus, it typically provides fewer conditions for
283 comparison than grow/no grow tests or growth rates. However,

the reduction in conditions is more than compensated for by the
284 increased metabolic resolution.

285 **2.5.1. Comparison of Phenol Flux Predictions and ^{13}C -MFA**
286 **Fluxes.** ^{13}C -MFA of phenol metabolism was obtained from a
287 previous publication.⁵ The glucose ^{13}C -MFA data was obtained
288 following the same procedure as discussed in that publication.
289 The transcriptomics data and growth curves for phenol came
290 from Henson et al.⁷ The glucose growth curves and
291 consumption data are new in this work, and they were generated
292 from cultures grown under the same conditions as the Henson et
293 al. data (except for the carbon source). For comparisons with
294 ^{13}C -MFA data, the carbon source uptake rates for pFBA, E-
295 Flux2, and SPOT were normalized to 100 units (instead of the
296 experimentally determined mmol substrate/g biomass/h), in
297 accordance with ^{13}C -MFA convention.
298

299 For the phenol case, intracellular fluxes were accurately
300 predicted by the COBRA methods (Figures 3 and 4). Fluxes
301 predicted by E-Flux2 were very close to the fluxes measured
302 through ^{13}C -MFA ($R^2 = 0.96$ without considering ATP
303 maintenance). pFBA predicted fluxes that were slightly less
304 accurate than those predicted by E-Flux2 ($R^2 = 0.93$). Though
305 minor compared to the other methods, the largest divergences
306 between E-flux2 predictions and ^{13}C -MFA measurements were
307 found in anaplerotic reactions and transport reactions. For
308 pFBA, the trend continued with the largest divergences coming
309 from anaplerotic reactions and transport reactions. The
310 prediction errors for anaplerotic reactions may be a result of
311 their underdetermined nature in ^{13}C -MFA due to reactions with
312 matching labeling patterns. pFBA and E-Flux2 both under-
313 predicted the flux of CO_2 out of the cell, which is a direct
314 consequence of growth rate overprediction. Since these
315 methods assume complete carbon efficiency to maximize
316 biomass, it is expected that they would underestimate the
317 amount of carbon lost as CO_2 .

318 For phenol, SPOT's predictions were the least accurate ($R^2 =$
319 0.66). Despite the decent R^2 value, a closer analysis of SPOT's
320 predictions, guided by biochemical knowledge, shows that it
321 generates an unrealistic metabolic flux profile. Most notably,
322 SPOT critically underestimates TCA cycle fluxes, especially with
323 respect to isocitrate dehydrogenase, α -ketoglutarate dehydro-
324 genase, and succinyl-CoA synthetase. When phenol uptake was
325 normalized to 100 units, each of these reactions had errors over
326 100 units. Particularly noteworthy are isocitrate dehydrogenase
327

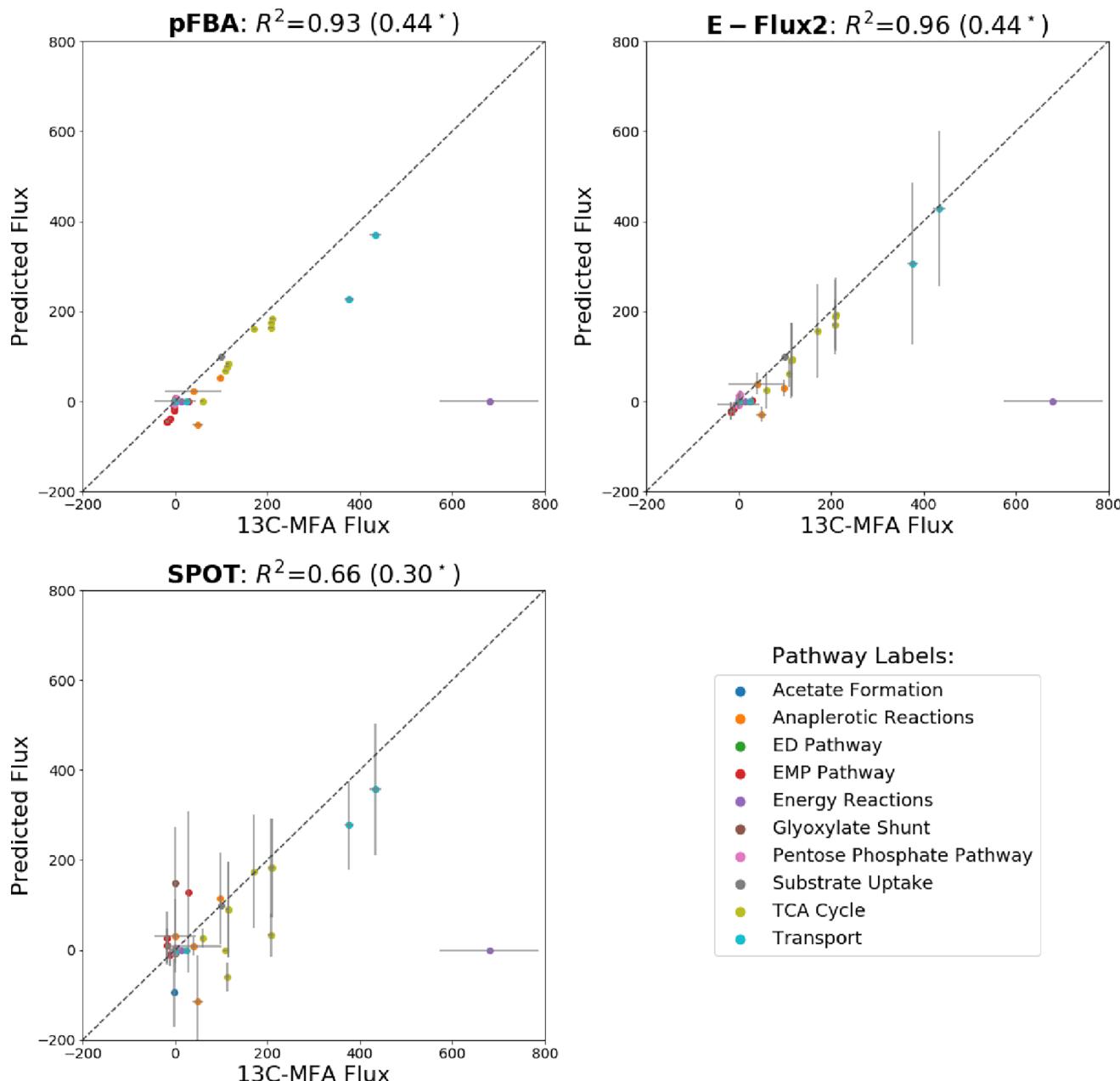


Figure 3. Flux predictions for phenol metabolism. Predictions are accurate for the three COBRA methods. The *y* axis represents the predicted flux by each of the COBRA methods (pFBA, E-Flux2, and SPOT) and the *x* axis represents the flux measured via ^{13}C -MFA. The fluxes are normalized to the carbon source uptake (units are mmol reaction/100 mmol phenol uptake). The first R^2 value does not include ATP maintenance reaction and the R^2 value in parentheses includes the ATP maintenance reaction. The *x* axis error bars are 90% confidence intervals as determined via ^{13}C -MFA, and if applicable, the *y* axis error bars are standard deviations of flux predictions made from three biological replicates of transcriptomics data.

and α -ketoglutarate dehydrogenase, which were predicted to have negative and zero flux, respectively (Table S1). To compensate for the underpredictions of the TCA cycle reactions, the flux through the glyoxylate shunt was overpredicted. SPOT predicted the flux of isocitrate lyase to be ~ 150 units, while the ^{13}C -MFA determined its flux to be only 0.4 (Figure 4). This discrepancy casts doubt on the viability of SPOT as a widely applicable standalone method for predicting fluxes from transcript data.

E-Flux2 and SPOT were also applied to phenol metabolism in the PVHG6 strain. Since pFBA does not take transcript measurements into account, its predictions are the same for the wild type and mutant strains. Overall, the transcript profiles

of the two strains on phenol were very similar,⁷ so it was expected that the mutant strain flux predictions from EFlux-2 and SPOT would be similar to the wild type predictions. Indeed, EFlux-2 makes accurate flux predictions for phenol metabolism in the mutant strain (wild type EFlux-2, $R^2 = 0.96$; mutant EFlux-2, $R^2 = 0.95$; Figure S4). Interestingly, despite similar transcriptomics measurements, SPOT's predictions of fluxes in the mutant strain are different from the wild type (Table S2) (wild type SPOT, $R^2 = 0.66$; mutant SPOT, $R^2 = 0.39$ Figure S4). The greater difference of SPOT's predictions between the strains compared to E-Flux2 demonstrates that E-Flux2 is more robust to small changes in transcript values than SPOT. As in the

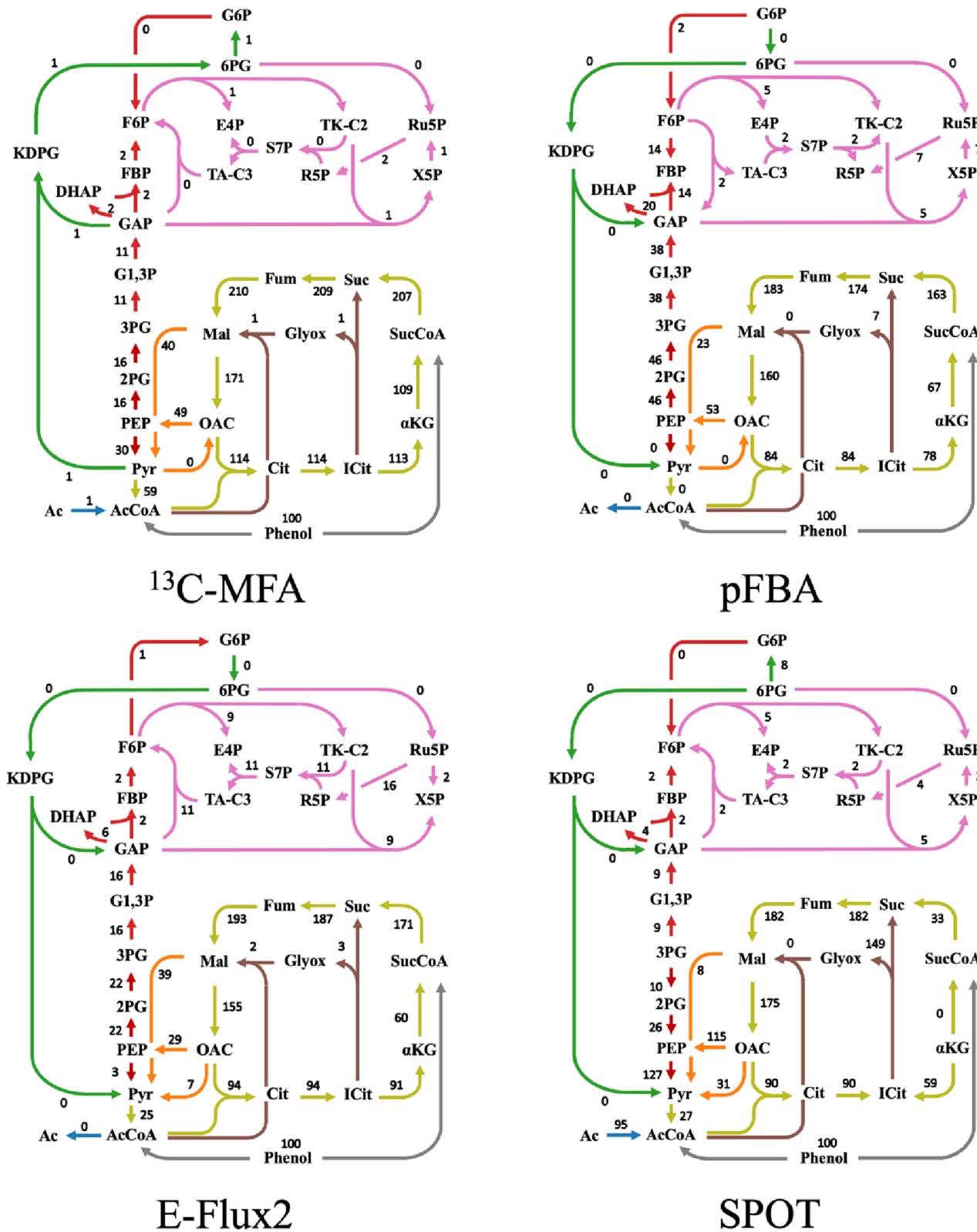


Figure 4. Phenol flux maps. Flux map predictions when phenol is the sole carbon source. The flux values are relative flux distributions based on 100 mmol of phenol consumed by the cell to generate 100 mmol of influx toward both acetyl-CoA and succinyl-CoA. A mapping of abbreviations to metabolite names is given in Table S8.

wild type's phenol condition, the largest errors in SPOT's mutant predictions occurred in the TCA cycle (Figure S5).

2.5.2. Comparison of Glucose Flux Predictions and ^{13}C -MFA Fluxes. In the case of glucose, each of the three predictive

methods show limitations (Figure 5). As observed with the phenol condition, E-Flux2 had the best predictions, though in this case, its predictions only fit moderately well ($R^2 = 0.63$). SPOT's predictions had the second best fit for glucose ($R^2 = 0.56$).

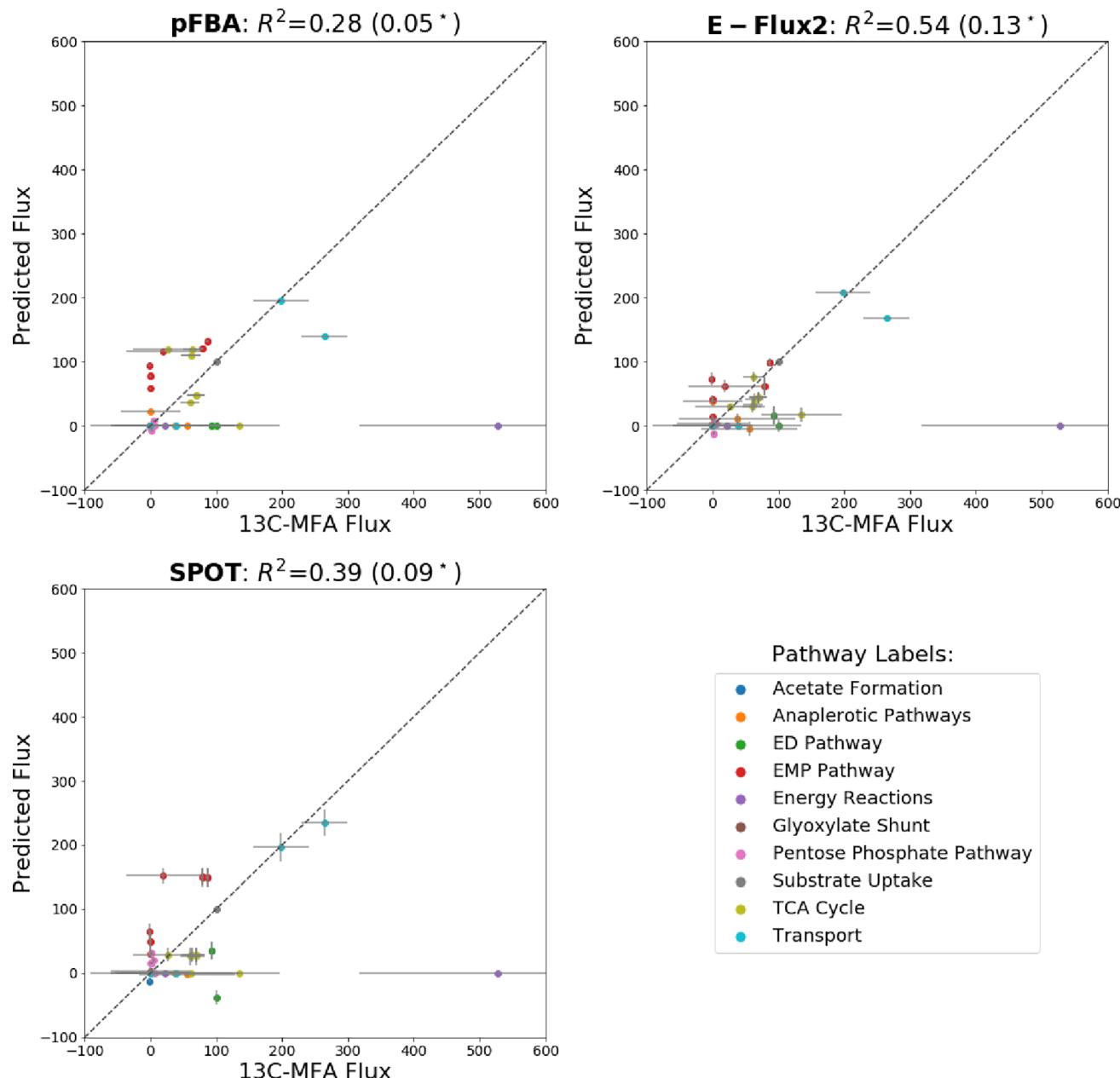
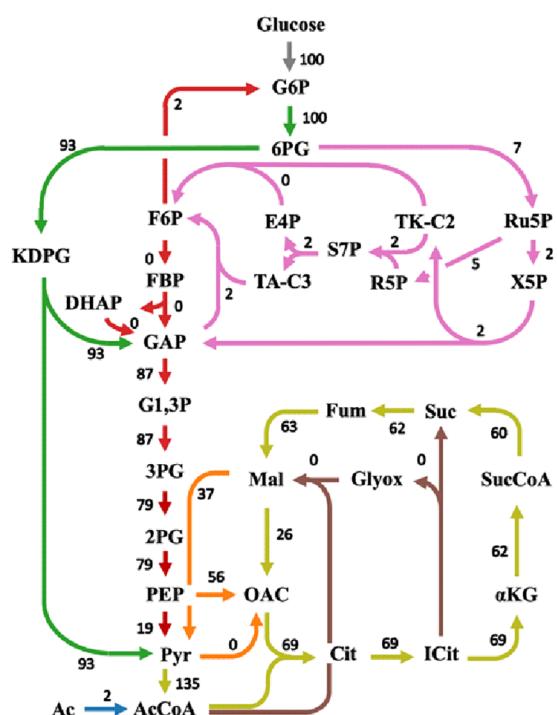
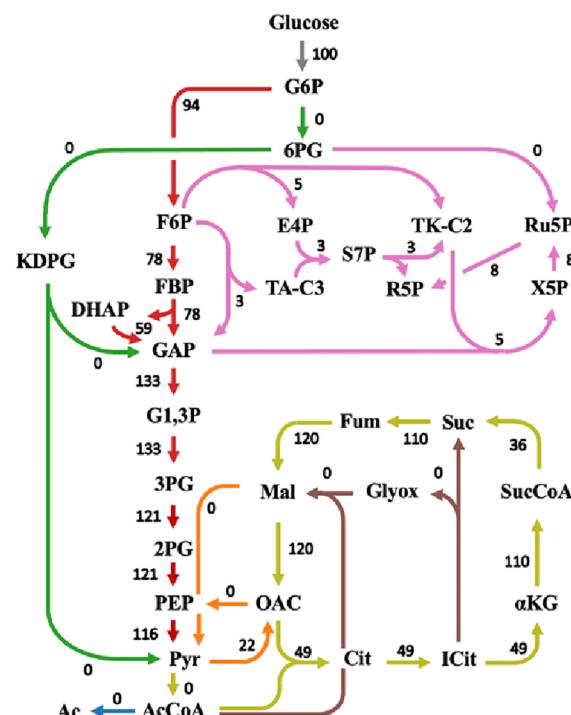


Figure 5. Glucose metabolism flux predictions. Glucose metabolism flux predictions are much less accurate for the COBRA methods considered here. Comparison of ^{13}C -MFA fluxes with model-predicted fluxes for glucose metabolism in the wild type strain. Horizontal and vertical axes and error bars are as described in Figure 3. In the same way, the R^2 value in parentheses is the R^2 value when ATP maintenance is included in the calculation.

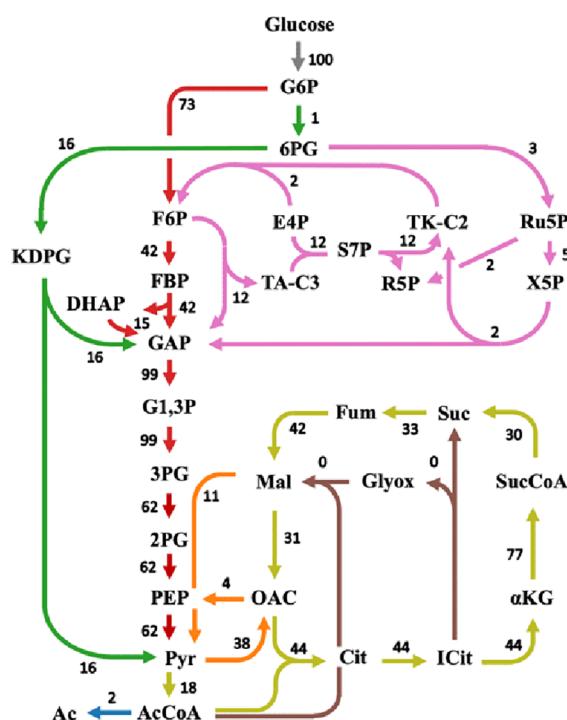
0.45), and pFBA's predictions were largely inaccurate ($R^2 = 0.28$) (Figure 5). One major difference between the three methods occurred in the predictions for the glucose uptake pathways. Two of these pathways, the EMP pathway and the ED pathway, share the enzymes that connect glyceraldehyde-3-phosphate to pyruvate but differ in their initial enzymes. Between the two, *R. opacus* shows a strong preference for the ED pathway, with approximately 95% of glucose consumed via this pathway despite a complete EMP pathway also being present.³⁹ While the two run essentially in parallel, this stark disparity is nonetheless unexpected, as the EMP pathway produces an extra molecule of ATP per molecule of glucose metabolized.³⁶ Potentially, the enzyme efficiency of the ED pathway explains this preference. Predictably, while ^{13}C -MFA determined that 93% of glucose was consumed through the ED pathway, pFBA

predicted that the ED pathway would have zero flux because creating extra ATP helps facilitate reactions including the biomass production reaction. Interestingly, the methods that incorporate transcriptomics into the genome-scale model recapitulate some ED flux. E-Flux2 and SPOT predict 21% and 38% of glucose consumption to occur via the ED pathway, respectively (Figure 6). These non-zero ED flux values contribute to the increased accuracy of the transcriptomics-based methods over FBA-based methods.

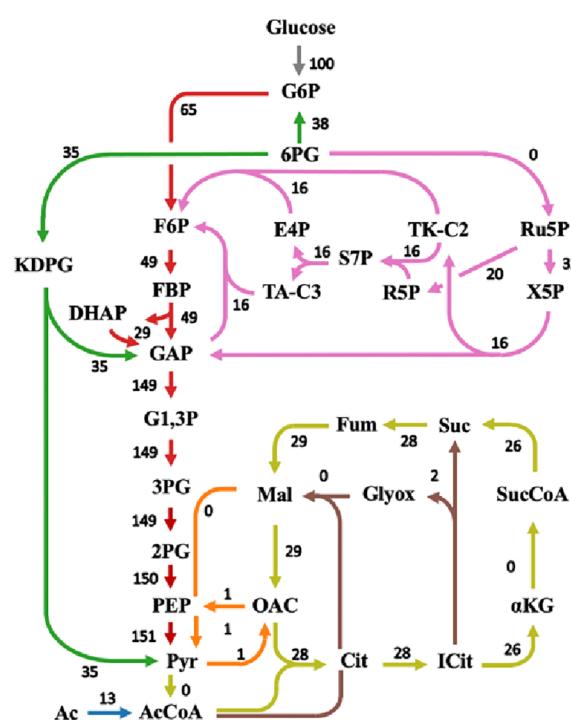
Similar to the predictions made for phenol growth conditions, pFBA predicted TCA cycle fluxes of glucose metabolism with less accuracy than E-Flux2 and SPOT. pFBA overestimated the fluxes of α -ketoglutarate dehydrogenase, succinate dehydrogenase, fumarase, and malate dehydrogenase (Table S3). All of these enzymes, except fumarase, produce reducing equivalents

¹³C-MFA

pFBA



E-Flux2



SPOT

Figure 6. Glucose flux maps. Flux map predictions when glucose is the sole carbon source. The flux values are relative flux distributions based on 100 mmol of phenol consumed by the cell to generate 100 mmol of influx toward both acetyl-CoA and succinyl-CoA.

390 in the form of NADH or FADH₂. FBA and pFBA's
391 overprediction of these TCA cycle reactions results in additional
392 energy molecules and carbon losses.

2.6. ATP Maintenance Flux Upper Bound Estimates. 393
Multiple methods for determining the non-growth associated 394
ATP maintenance flux (NGAM) show that glucose metabolism 395

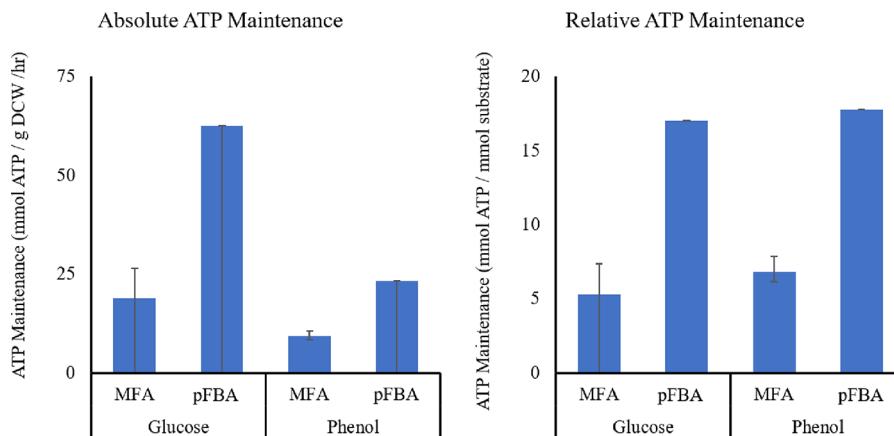


Figure 7. ATP maintenance flux as determined by metabolic flux analysis (MFA) and flux balance analysis (pFBA). Absolute ATP maintenance is the mmol of ATP used by 1 g of dry cell weight per hour, and relative ATP maintenance is the mmol of ATP used per mmol of either glucose or phenol consumed.

and phenol metabolism function with similar efficiency (i.e., relative ATP used for maintenance). NGAM is the amount of ATP generated in a metabolic model that is not consumed by the reactions in the model. It is thought that this excess ATP is used for cellular “housekeeping” tasks such as maintaining ionic gradients and producing enzymes via transcription and translation.⁴⁰ A cell is considered to be operating at higher efficiency when its ATP maintenance flux is low as less ATP is “lost” to non-growth purposes.

The GSM calculated non-growth associated ATP maintenance flux via FBA. When ATP maintenance loss is high, less biomass can be produced because ATP (growth associated) is a reactant in the biomass equation, and ATP is a required cofactor for many reactions that produce biomass precursors. The flux configuration with the maximum growth rate has zero ATP maintenance flux, and the flux configuration with the maximum ATP maintenance flux has zero biomass production (Figure S6). The true ATP maintenance loss can be estimated by mapping the experimental growth rate onto the ATP maintenance flux vs growth rate curve. This method gives the same result as fixing the growth rate and then calculating the maximum ATP maintenance flux (using fixed growth associated ATP maintenance). Using this method, the model predicts that the non-growth associated ATP maintenance flux was 23.4 mmol ATP per gram dry cell weight per hour when consuming phenol and 63.0 mmol ATP per gram dry cell weight per hour when consuming glucose (Figure 7). In ¹³C-MFA, ATP maintenance flux is a fitted variable constrained by amino acid labeling patterns. The ¹³C-MFA ATP maintenance flux was 9.2 mmol ATP per gram dry cell weight per hour when consuming phenol and 18.9 mmol ATP per gram dry cell weight per hour when consuming glucose (Figure 7).

The ATP maintenance flux calculated using FBA is roughly three times greater than the value determined by ¹³C-MFA (Figure 7), a discrepancy that can be traced to FBA’s fundamental assumption that cells are optimized to maximize biomass production. As described above, FBA was used to estimate the ATP maintenance flux by fixing the model’s growth rate to the experimental growth rate and then maximizing the amount of ATP maintenance flux. As a result, the ATP maintenance value predicted by FBA represents the upper bound of possible ATP maintenance values in the same way that FBA’s growth rate predictions represent the theoretical maximum growth rate. Interestingly, while glucose had a higher

absolute ATP maintenance flux per hour than phenol, when the data was normalized per mmol of substrate uptake, this difference was largely eliminated. This indicates that per mole of substrate, both conditions use roughly the same amount of ATP for non-growth activities despite the difference in uptake rates.

3. CONCLUSIONS

In this article, we present a GSM for *R. opacus* PD630: iGR1773. This model provides a tool for predicting this organism’s metabolism and can help fulfill its potential as a platform for converting lignin derivatives into liquid fuels and chemicals. iGR1773 was validated with the Metabolic Model Test (MEMOTE) suite,³¹ by checking growth rate predictions, and through comparisons of flux predictions via COBRA methods to ¹³C-MFA measurements. The COBRA method that provided the most accurate predictions was E-Flux2 followed by pFBA and SPOT. In general, the COBRA methods were more accurate for phenol than for glucose. Additionally, the model was used to demonstrate that *R. opacus*’ metabolic network operates with similar efficiency when consuming phenol or glucose. We expect this GSM to be a stepping-stone toward building progressively more predictive models of *R. opacus* metabolism that will guide future metabolic engineering efforts.

4. MATERIALS AND METHODS

4.1. Strains and Data. The data used in this manuscript originated either in previous publications^{5,7} or are newly reported in this work (Table 3). The experiments in this work used *Rhodococcus opacus* PD630 (DSMZ 44193) as the wild-

Table 3. Sources of the Experimental Data Used in This Paper

	phenol (wild type and PVHG6)	glucose (wild type)
transcript data	Henson et al. (2018) ⁷	Henson et al. (2018) ⁷
growth curves	Henson et al. (2018) ⁷	first published in this paper
substrate consumption curves	Henson et al. (2018) ⁷	first published in this paper
¹³ C-metabolic flux analysis	Roell et al. (2019) ⁵	first published in this paper
biomass composition	first published in this paper	first published in this paper

466 type strain and a *Rhodococcus opacus* PD630 mutant strain
467 PVHG6, which had previously been adaptively evolved on a
468 mixture of phenol, vanillate, guaiacol, 4-hydroxybenzoate, and
469 guaiacol.⁷ All data was generated from fermentation experiments
470 wherein *R. opacus* was cultured in minimal media B with either
471 phenol or glucose as the sole carbon source and 1 g/L
472 ammonium sulfate as the nitrogen source.⁴¹ The transcript data
473 used in this analysis comes from a previous publication⁷ stored
474 in the National Center for Bioinformatics Sequence Read
475 Archive in bioproject PRJNA431604, and the data was
476 reprocessed to count per million (CPM) normalization. The
477 growth curve data for phenol conditions, OD₆₀₀, and substrate
478 consumption data, were from a previous report,⁷ while the
479 glucose data was generated in this work. The ¹³C-MFA data for
480 phenol was previously reported,⁵ and the glucose data was
481 obtained using the same procedure as described therein. The
482 biomass composition data for both phenol and glucose was
483 obtained using a custom spectrophotometry method described
484 in Section 4.4.

485 **4.2. Draft Model Reconstruction and Gap Filling.** The
486 initial version of the GSM for *R. opacus* was made using
487 CarveMe, an automated tool developed to produce GSMS.³² For
488 this reconstruction, the following versions were used: CarveMe
489 1.5.1, Diamond 0.9.14, and CPLEX 12.10.0.0. CarveMe follows
490 a top-down approach where a universal model and genome
491 sequence are the only required inputs to construct a model in a
492 fast and reproducible manner. The GSM was based on a recent
493 genomic sequence of the *Rhodococcus opacus* PD630 (Refseq ID:
494 GCF_020542785.1).²⁷ The initial model was made using the
495 command line command “carve r_opacus_bologna.faa -u
496 grampos -o r_opacus_bologna_raw.xml”. After the model
497 generation, this initial draft model was also gap-filled to ensure
498 growth on M9 and LB media using the command “gapfill
499 r_opacus_bologna.xml -m M9,LB -o r_opacus_bologna_gap-
500 filled.xml”.

501 **4.3. Addition of Uptake Reactions (Notebook A).** As
502 generated by CarveMe, the GSM did not contain uptake
503 reactions for all the carbon sources *R. opacus* can metabolize, so
504 these reactions were added in notebook A. This initial model
505 contained all the reactions needed for the model to consume
506 several carbon sources including glucose, 4-hydroxybenzoate,
507 vanillate, and benzoate. To account for growth with phenol, the
508 metabolites for extracellular and intracellular phenol were added
509 as well as the reactions for phenol exchange (adding phenol to
510 the medium), phenol transport (phenol entering the cell), and
511 phenol monooxygenase (phenol + NADH + O₂ + H⁺ →
512 catechol + NAD⁺ + H₂O; EC 1.14.13.244). For growth with
513 guaiacol, intracellular and extracellular guaiacol were added, and
514 so were reactions for exchange, transport, and guaiacol-o-
515 demethylase (guaiacol + NADPH + O₂ → catechol +
516 formaldehyde + NADP⁺ + H₂O; EC 1.14.14.-). Additionally,
517 an intracellular metabolite for triacylglycerol (TAG) and
518 reactions for its production from 1,2-diacyl-sn-glycerol and
519 palmitoyl-CoA and transport out of the cell were added to the
520 model. The bounds of two reactions, catalyzed by 3-
521 hydroxyadipyl-CoA dehydrogenase and succinate dehydrogen-
522 ase, were adjusted to avoid thermodynamically infeasible cycles.
523 This notebook also contains tests to ensure that the model can
524 explain the growth in glucose, phenol, vanillate, 4-hydrox-
525 ybenzoate, guaiacol, and benzoate. In addition to these aromatic
526 carbon sources, *R. opacus* PD630 has also been shown to be able
527 to consume mannitol, ribose, xylose, lactose, and maltose as sole
528 carbon sources according to the BacDive page for DSMZ 44193.

The model from CarveMe was able to consume all these carbon
529 sources without the need for manual edits.
530

4.4. Addition of Custom Biomass Reactions (Notebook)

B). The biomass composition of *R. opacus* when grown with
532 various substrates was quantified in terms of carbohydrate, lipid,
533 and protein fractions. Carbohydrates were measured using a
534 hydrolysis procedure. Lipid extraction, purification, and
535 measurement were conducted using the Bligh and Dyer
536 method.⁴² Proteins were measured with an L-8800 AAA Hitachi
537 High-Speed Amino Acid Analyzer. These measurements are
538 summarized in Table S4.
539

The biomass composition data and previously reported amino
540 acid data⁵ were used to make customized biomass equations for
541 the *R. opacus* GSM when grown in glucose or phenol. These
542 biomass equations were based on the *Bacillus subtilis* biomass
543 equation that comes by default with CarveMe for Gram-positive
544 bacteria.⁴³ In the customized *R. opacus* biomass equations, the
545 coefficients for precursors that are not amino acids, lipids, or
546 carbohydrates (e.g., energy molecules and salts) are the same as
547 they are in the *B. subtilis* biomass equation. The coefficients of
548 lipid and carbohydrate precursors were scaled proportionally to
549 the measured amount of lipids or carbohydrates in *R. opacus*.
550 The amino acid coefficients were calculated using the measured
551 milligrams of amino acids per gram of biomass and the measured
552 mole percentage of each amino acid. Table S5 contains a
553 comparison of the biomass equations for *R. opacus* with phenol,
554 *R. opacus* with glucose, and *B. subtilis*.
555

556 **4.5. Addition of Metabolite, Reaction, and Gene
Annotations (Notebook C).** The reconstruction from
557 CarveMe included detailed metabolite and reaction annotations.
558 The only metabolites in the *R. opacus* model that were not
559 included in the BiGG Universal model were guaiacol and
560 triacylglycerol.⁹ All but 25 of the reactions in the *R. opacus* model
561 were found in the Universal model, so these reactions were left
562 unannotated. The model's gene IDs are the NCBI non-
563 redundant protein accession numbers (with the prefix ‘WP_’)
564 from the NCBI database (Refseq ID: GCF_020542785.1).⁴⁴
565 The proper system biology ontology (SBO)⁴⁵ numbers were
566 also added to all metabolites, reactions, and genes. Further, since
567 the annotations in the Universal model are the Python type, List,
568 they were converted into dictionaries with keys to match
569 MEMOTE's requirements.
570

571 **4.6. Experimental Determination of Growth Rate and
Substrate Uptake Rate (Notebook D).** Experimental growth
572 rates were calculated by first collecting time-course OD₆₀₀ data
573 from fermentations with 5 mM phenol or glucose as the carbon
574 source and 1 g/L ammonium sulfate as the nitrogen source. The
575 growth rate was calculated using the slope of the log-
576 transformed OD vs time regression since the growth in the
577 exponential phase follows the equation $X(t) = X_0 e^{\mu t}$, where $X(t)$
578 represents the OD at time t , X_0 is the initial OD, μ is the growth
579 rate in h⁻¹, and t is the time in hours. The yield coefficient (g
580 biomass/mmol substrate) was determined using the slope of the
581 line made when plotting the amount of substrate consumed vs
582 the amount of biomass produced. The substrate consumption
583 rate (mmol substrate/g biomass/h) was calculated by dividing
584 the growth rate (h⁻¹) by the yield coefficient (g biomass/mmol
585 substrate). For each of the three conditions (wild-type phenol,
586 wild-type glucose, and PVHG6 phenol), there were three
587 biological replicates of growth and consumption data. The
588 growth parameters were calculated individually for each trial and
589 then averaged for each condition (Table 2).
590

591 **4.7. Growth Rate Simulations.** The *R. opacus* GSM was
592 used to make growth rate predictions. While GSMS are
593 stoichiometric models without a time component, when the
594 input and output reactions are properly scaled, these models can
595 be used to predict growth rates.⁴⁶ The model was calibrated to
596 simulate the behavior of 1 g of dry cell weight for 1 h. The
597 substrate uptake rate was set to the amount of substrate, in
598 mmol, that 1 g of biomass would consume in 1 h, and the
599 biomass formation reaction was set up so that its flux would
600 equal the amount of biomass in grams produced in 1 h. Growth
601 rate (μ) is defined according to the equation $dX/dt = \mu X$, where
602 dX/dt is the rate of change of biomass and X is the biomass
603 concentration. Translating to the GSM, dX/dt is equal to the
604 biomass flux, and since the model was scaled for 1 g of biomass
605 ($X = 1$), the biomass flux is equal to the growth rate.

606 **4.8. Comparison with ^{13}C -MFA.** Another approach for
607 validating the GSM is to compare its flux predictions with fluxes
608 determined using ^{13}C -MFA. Since the ^{13}C -MFA metabolic
609 network contains ~ 70 reactions and the iGR1773 GSM contains
610 ~ 2300 reactions, reactions from the two cannot be directly
611 compared. A mapping of reactions from the GSM to the ^{13}C -
612 MFA reactions was made to compare genome-scale flux
613 predictions and ^{13}C -MFA measurements (Table S6). Some
614 reactions in the ^{13}C -MFA model involve multiple reactions in
615 the GSM. This can happen when two reactions occur in series or
616 when they occur in parallel. An example of reactions in series is
617 the conversion of 3-phosphoglycerate to phosphoenolpyruvate.
618 In the GSM, 3-phosphoglycerate is converted to 2-phosphoglyc-
619 erate and then to phosphoenolpyruvate, while in the ^{13}C -MFA,
620 3-phosphoglycerate is directly converted to phosphoenolpyr-
621 uvate. The minimum flux value of reactions in series was
622 compared to ^{13}C -MFA flux. Additionally, some reactions in the
623 ^{13}C -MFA have multiple reactions that act in parallel in the GSM.
624 An example is malate dehydrogenase. In the ^{13}C -MFA, there is
625 only a single isozyme (that produces NADH), while in the GSM,
626 there are isozymes that produce NADH, menaquinone, and
627 ubiquinone. The sum of fluxes of parallel reactions was
628 compared to ^{13}C -MFA flux. The quality of GSM fit was
629 determined by calculating the R^2 (coefficient of determination)
630 between the GSM model fluxes and the ^{13}C -MFA fluxes with
631 and without the ATP maintenance flux.⁴⁷

632 **4.9. Methods to Predict Fluxes from Transcripts.** E-
633 Flux2 predicts fluxes from transcripts by solving an FBA problem
634 where the upper and lower bounds for each reaction have been
635 modified according to the absolute expression for the
636 corresponding gene.¹⁷ The underlying idea is that, given a
637 limited translational efficiency and enzyme accumulated over
638 the time, the mRNA level can be considered as an approximate
639 upper bound on the maximum amount of metabolic enzyme
640 available and hence as a bound on reaction rates. Briefly, after a
641 suitable flux bound normalization, the upper bound for each flux
642 with transcript information is substituted by the absolute
643 expression for the corresponding gene (for a positive upper
644 bound, zero otherwise). If the reaction is reversible, the lower
645 bound is substituted by the negative value of the absolute
646 expression for the corresponding gene (if lower bound is
647 negative, zero otherwise). An FBA problem is solved using these
648 bounds and, as a last step, which differentiates E-Flux2 from its
649 previous version of E-Flux,¹⁶ the norm of the resulting flux is
650 minimized. This ensures a single solution, unlike E-Flux. SPOT,
651 instead of optimizing growth, maximizes the correlation
652 between fluxes and the measured transcript profile, as
653 determined through the Pearson correlation coefficient.¹⁷ The

654 assumption is that enzymatic transcript concentrations and 655
656 fluxes tend to be as proportional to each other as allowed by 655
657 stoichiometric constraints and enzyme presence. SPOT trans- 656
658 forms the problem into an equivalent semi-definite program- 657
659 which is the version we use here. 659

660 **4.10. Summary of Jupyter Notebooks in This Pub- 660**
661 **lication.** Table S7 contains the list of the Jupyter notebooks 661
662 used for creating the figures in this paper. 662

■ ASSOCIATED CONTENT 663

Data Availability Statement 664

The data used in this project is publicly available at <https://github.com/LBLQMM/RhodococcusGSM>. 665
666

SI Supporting Information 667

The Supporting Information is available free of charge at 668
<https://pubs.acs.org/doi/10.1021/acssynbio.2c00618>. 669

Table S1: wild-type phenol ^{13}C -MFA vs genome-scale 670
model predicted fluxes; Table S2: PVHG6 phenol ^{13}C - 671
MFA vs genome-scale model predicted fluxes; Table S3: 672
wild-type glucose ^{13}C -MFA vs genome-scale model 673
predicted fluxes; Table S4: biomass composition of *B.* 674
subtilis and *R. opacus* with glucose and *R. opacus* with 675
phenol; Table S5: *R. opacus* PD630 biomass reactants for 676
phenol and glucose conditions; Table S6: mapping of 677
reactions from ^{13}C -MFA to genome-scale model reac- 678
tions; Table S7: description of the notebooks used to 679
generate and test iGR1773; Table S8: mapping of 680
abbreviations to metabolite names; Figure S1: WT 681
phenol growth and consumption data; Figure S2: 682
PVHG6 phenol growth and consumption data; Figure 683
S3: WT glucose growth and consumption data; Figure S4: 684
phenol metabolism flux predictions based on PVHG6 685
strain transcripts; Figure S5: phenol flux maps based on 686
PVHG6 transcripts; Figure S6: FBA prediction of ATP 687
maintenance vs growth rate (PDF) 688

■ AUTHOR INFORMATION 689

Corresponding Authors 690

Tae Seok Moon – Department of Energy, Environmental and 691
Chemical Engineering, Washington University in St. Louis, St. 692
Louis, Missouri 63130, United States; Email: hgmartin@lbl.gov 693
694

Yinjie J. Tang – Department of Energy, Environmental and 695
Chemical Engineering, Washington University in St. Louis, St. 696
Louis, Missouri 63130, United States; Email: yinjie.tang@wustl.edu 697
698

Hector García Martín – BCAM - Basque Center for Applied 699
Mathematics, Bilbao 48009, Spain; DOE Agile BioFoundry, 700
Emeryville, California 94608, United States; Biological 701
Systems and Engineering Division, Lawrence Berkeley National 702
Lab, Berkeley, California 94720, United States; DOE Joint 703
BioEnergy Institute, Emeryville, California 94608, United 704
States; [© orcid.org/0000-0002-4556-9685](http://orcid.org/0000-0002-4556-9685); 705
Email: tsmoon7@gmail.com 706

Authors 707

Garrett W. Roell – Department of Energy, Environmental and 708
Chemical Engineering, Washington University in St. Louis, St. 709
Louis, Missouri 63130, United States 710

Christina Schenk – BCAM - Basque Center for Applied 711
Mathematics, Bilbao 48009, Spain; Biological Systems and 712

713 Engineering Division, Lawrence Berkeley National Lab,
714 Berkeley, California 94720, United States;  orcid.org/0000-
715 0002-7817-6757

716 Winston E. Anthony — The Edison Family Center for Genome
717 Sciences and Systems Biology, Washington University in St.
718 Louis School of Medicine, St. Louis, Missouri 63110, United
719 States; Department of Pathology and Immunology,
720 Washington University in St. Louis School of Medicine, St.
721 Louis, Missouri 63108, United States

722 Rhiannon R. Carr — Department of Energy, Environmental and
723 Chemical Engineering, Washington University in St. Louis, St.
724 Louis, Missouri 63130, United States

725 Aditya Ponukumati — Department of Energy, Environmental and
726 Chemical Engineering, Washington University in St. Louis,
727 St. Louis, Missouri 63130, United States

728 Joonhoon Kim — DOE Agile BioFoundry, Emeryville,
729 California 94608, United States; DOE Joint BioEnergy
730 Institute, Emeryville, California 94608, United States

731 Elena Akhmatkaya — BCAM - Basque Center for Applied
732 Mathematics, Bilbao 48009, Spain; Biological Systems and
733 Engineering Division, Lawrence Berkeley National Lab,
734 Berkeley, California 94720, United States; IKERBASQUE,
735 Basque Foundation for Science, Bilbao 48009, Spain

736 Marcus Foston — Department of Energy, Environmental and
737 Chemical Engineering, Washington University in St. Louis, St.
738 Louis, Missouri 63130, United States;  orcid.org/0000-
739 0002-4227-0362

740 Gautam Dantas — The Edison Family Center for Genome
741 Sciences and Systems Biology, Washington University in St.
742 Louis School of Medicine, St. Louis, Missouri 63110, United
743 States; Department of Pathology and Immunology and
744 Department of Molecular Microbiology, Washington
745 University in St. Louis School of Medicine, St. Louis, Missouri
746 63108, United States; Department of Biomedical Engineering,
747 Washington University in St. Louis, St Louis, Missouri 63130,
748 United States; Department of Pediatrics, Washington
749 University School of Medicine in St Louis, St Louis, Missouri,
750 United States;  orcid.org/0000-0003-0455-8370

751 Complete contact information is available at:

752 <https://pubs.acs.org/10.1021/acssynbio.2c00618>

753 Author Contributions

754 T.S.M., Y.J.T., and H.G.M. conceived the project. G.W.R., C.S.,
755 and H.G.M. conceived the methods of GSM construction and
756 application of omics data. W.E.A., R.R.C., and A.P. prepared data
757 for modeling. All authors wrote and proofread the paper.

758 Notes

759 The authors declare no competing financial interest.
760 The code for this project and the model itself are open source.
761 They are available at <https://github.com/LBLQMM/RhodococcusGSM>.

763 ■ ACKNOWLEDGMENTS

764 This project was supported by the US DOE (DE-SC0018324)
765 and the Office of Science Graduate Student Research (SCGSR)
766 Program. C.S. and E.A. were also supported by the Basque
767 Government through the BERC 2018–2021 program and by
768 the Spanish Ministry of Economy and Competitiveness
769 MINECO: BCAM Severo Ochoa excellence accreditation
770 SEV-2017-0718.

■ REFERENCES

- (1) Gani, A.; Naruse, I. Effect of cellulose and lignin content on pyrolysis and combustion characteristics for several types of biomass. *Renewable Energy* **2007**, *32*, 649–661.
- (2) Alvarez, H. M.; Mayer, F.; Fabritius, D.; Steinbüchel, A. Formation of intracytoplasmic lipid inclusions by *Rhodococcus opacus* strain PD630. *Arch. Microbiol.* **1996**, *165*, 377–386.
- (3) Chatterjee, A.; DeLorenzo, D. M.; Carr, R.; Moon, T. S. Bioconversion of renewable feedstocks by *Rhodococcus opacus*. *Curr. Opin. Biotechnol.* **2020**, *64*, 10–16.
- (4) Fuchs, G.; Boll, M.; Heider, J. Microbial degradation of aromatic compounds — from one strategy to four. *Nat. Rev. Microbiol.* **2011**, *9*, 780 803–816.
- (5) Roell, G. W.; Carr, R. R.; Campbell, T.; Shang, Z.; Henson, W. R.; Czajka, J. J.; Martín, H. G.; Zhang, F.; Foston, M.; Dantas, G.; Moon, T. S.; Tang, Y. J. A concerted systems biology analysis of phenol metabolism in *Rhodococcus opacus* PD630. *Metab. Eng.* **2019**, *55*, 787 120–130.
- (6) Yoneda, A.; Henson, W. R.; Goldner, N. K.; Park, K. J.; Forsberg, K. J.; Kim, S. J.; Pesesky, M. W.; Foston, M.; Dantas, G.; Moon, T. S. Comparative transcriptomics elucidates adaptive phenol tolerance and utilization in lipid-accumulating *Rhodococcus opacus* PD630. *Nucleic Acids Res.* **2016**, *44*, 2240–2254.
- (7) Henson, W. R.; Campbell, T.; DeLorenzo, D. M.; Gao, Y.; Berla, B.; Kim, S. J.; Foston, M.; Moon, T. S.; Dantas, G. Multi-omic elucidation of aromatic catabolism in adaptively evolved *Rhodococcus opacus*. *Metab. Eng.* **2018**, *49*, 69–83.
- (8) Diao, J.; Carr, R.; Moon, T. S. Deciphering the transcriptional regulation of the catabolism of lignin-derived aromatics in *Rhodococcus opacus* PD630. *Commun. Biol.* **2022**, *5*, 1109.
- (9) King, Z. A.; Lu, J.; Dräger, A.; Miller, P.; Federowicz, S.; Lerman, J. A.; Ebrahim, A.; Palsson, B. O.; Lewis, N. E. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **2016**, *44*, D515–D522.
- (10) Feist, A. M.; Palsson, B. O. The biomass objective function. *Curr. Opin. Microbiol.* **2010**, *13*, 344–349.
- (11) Dahal, S.; Yurkovich, J. T.; Xu, H.; Palsson, B. O.; Yang, L. Synthesizing Systems Biology Knowledge from Omics Using Genome-Scale Models. *Proteomics* **2020**, *20*, 1900282.
- (12) Zur, H.; Ruppert, E.; Shlomi, T. iMAT: an integrative metabolic analysis tool. *Bioinformatics* **2010**, *26*, 3140–3142.
- (13) Shlomi, T.; Cabili, M. N.; Herrgård, M. J.; Palsson, B. Ø.; Ruppert, E. Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **2008**, *26*, 1003–1010.
- (14) Becker, S. A.; Palsson, B. O. Context-Specific Metabolic Networks Are Consistent with Experiments. *PLoS Comput. Biol.* **2008**, *4*, No. e1000082.
- (15) Schmidt, B. J.; Ebrahim, A.; Metz, T. O.; Adkins, J. N.; Palsson, B. Ø.; Hyduke, D. R. GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data. *Bioinformatics* **2013**, *29*, 2900–2908.
- (16) Colijn, C.; Brandes, A.; Zucker, J.; Lun, D. S.; Weiner, B.; Farhat, M. R.; Cheng, T.-Y.; Moody, D. B.; Murray, M.; Galagan, J. E. Interpreting Expression Data with Metabolic Flux Models: Predicting *Mycobacterium tuberculosis* Mycolic Acid Production. *PLoS Comput. Biol.* **2009**, *5*, No. e1000489.
- (17) Kim, M. K.; Lane, A.; Kelley, J. J.; Lun, D. S. E-Flux2 and SPOT: Validated Methods for Inferring Intracellular Metabolic Flux Distributions from Transcriptomic Data. *PLoS One* **2016**, *11*, No. e0157101.
- (18) van Berlo, R. J. P.; de Ridder, D.; Daran, J.-M.; Daran-Lapujade, P. A. S.; Teusink, B.; Reinders, M. J. T. Predicting Metabolic Fluxes Using Gene Expression Differences As Constraints. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2011**, *8*, 206–216.
- (19) Mahadevan, R.; Schilling, C. H. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* **2003**, *5*, 264–276.
- (20) Kim, H. U.; Kim, W. J.; Lee, S. Y. Flux-coupled genes and their use in metabolic flux analysis. *Biotechnol. J.* **2013**, *8*, 1035–1042.

- 840 (21) Schultz, A.; Qutub, A. A. Reconstruction of Tissue-Specific
841 Metabolic Networks Using CORDA. *PLoS Comput. Biol.* **2016**, *12*,
842 No. e1004808.
- 843 (22) Fondi, M.; Liò, P. Multi -omics and metabolic modelling
844 pipelines: Challenges and tools for systems microbiology. *Microbiol.*
845 *Res.* **2015**, *171*, S2–64.
- 846 (23) Hyduke, D. R.; Lewis, N. E.; Palsson, B. Ø. Analysis of omics data
847 with genome-scale models of metabolism. *Mol. BioSyst.* **2013**, *9*, 167–
848 174.
- 849 (24) Reed, J. L. Shrinking the Metabolic Solution Space Using
850 Experimental Datasets. *PLoS Comput. Biol.* **2012**, *8*, No. e1002662.
- 851 (25) Blazier, A. S.; Papin, J. A. Integration of expression data in
852 genome-scale metabolic network reconstructions. *Front. Physiol.* **2012**,
853 *3*, 299.
- 854 (26) Machado, D.; Herrgård, M. Systematic Evaluation of Methods
855 for Integration of Transcriptomic Data into Constraint-Based Models
856 of Metabolism. *PLoS Comput. Biol.* **2014**, *10*, No. e1003580.
- 857 (27) Firrincieli, A.; Grigoriev, B.; Dostálová, H.; Cappelletti, M. The
858 Complete Genome Sequence and Structure of the Oleaginous
859 Rhodococcus opacus Strain PD630 Through Nanopore Technology.
860 *Front. Bioeng. Biotechnol.* **2022**, *9*, No. 810571.
- 861 (28) Sundararaghavan, A.; Mukherjee, A.; Sahoo, S.; Suraishkumar, G.
862 K. Mechanism of the oxidative stress-mediated increase in lipid
863 accumulation by the bacterium, R. opacus PD630: Experimental
864 analysis and genome-scale metabolic modeling. *Biotechnol. Bioeng.*
865 **2020**, *117*, 1779–1788.
- 866 (29) Kim, H. M.; Chae, T. U.; Choi, S. Y.; Kim, W. J.; Lee, S. Y.
867 Engineering of an oleaginous bacterium for the production of fatty acids
868 and fuels. *Nat. Chem. Biol.* **2019**, *15*, 721–729.
- 869 (30) Tajparast, M.; Frigon, D. Genome-scale metabolic model of
870 Rhodococcus jostii RHA1 (iMT1174) to study the accumulation of
871 storage compounds during nitrogen-limited condition. *BMC Syst. Biol.*
872 **2015**, *9*, 43.
- 873 (31) Lieven, C.; Beber, M. E.; Olivier, B. G.; Bergmann, F. T.; Ataman,
874 M.; Babaei, P.; Bartell, J. A.; Blank, L. M.; Chauhan, S.; Correia, K.;
875 Diener, C.; Dräger, A.; Ebert, B. E.; Edirisinghe, J. N.; Faria, J. P.; Feist,
876 A. M.; Fengos, G.; Fleming, R. M. T.; García-Jiménez, B.;
877 Hatzimanikatis, V.; van Helvoort, W.; Henry, C. S.; Hermjakob, H.;
878 Herrgård, M. J.; Kaafarani, A.; Kim, H. U.; King, Z.; Klamt, S.; Klipp, E.;
879 Koehler, J. J.; König, M.; Lakshmanan, M.; Lee, D.-Y.; Lee, S. Y.; Lee,
880 S.; Lewis, N. E.; Liu, F.; Ma, H.; Machado, D.; Mahadevan, R.; Maia, P.;
881 Mardinoglu, A.; Medlock, G. L.; Monk, J. M.; Nielsen, J.; Nielsen, L. K.;
882 Nogales, J.; Nookae, I.; Palsson, B. O.; Papin, J. A.; Patil, K. R.;
883 Poolman, M.; Price, N. D.; Resendis-Antonio, O.; Richelle, A.; Rocha,
884 I.; Sánchez, B. J.; Schaap, P. J.; Malik Sheriff, R. S.; Shoae, S.;
885 Sonnenschein, N.; Teusink, B.; Vilaça, P.; Vik, J. O.; Wodke, J. A. H.;
886 Xavier, J. C.; Yuan, Q.; Zakhartsev, M.; Zhang, C. MEMOTE for
887 standardized genome-scale metabolic model testing. *Nat. Biotechnol.*
888 **2020**, *38*, 272–276.
- 889 (32) Machado, D.; Andrejev, S.; Tramontano, M.; Patil, K. R. Fast
890 automated reconstruction of genome-scale metabolic models for
891 microbial species and communities. *Nucleic Acids Res.* **2018**, *46*,
892 7542–7553.
- 893 (33) Monk, J. M.; Lloyd, C. J.; Brunk, E.; Mih, N.; Sastry, A.; King, Z.;
894 Takeuchi, R.; Nomura, W.; Zhang, Z.; Mori, H.; Feist, A. M.; Palsson, B.
895 O. iML1515, a knowledgebase that computes Escherichia coli traits.
896 *Nat. Biotechnol.* **2017**, *35*, 904–908.
- 897 (34) Lu, H.; Li, F.; Sánchez, B. J.; Zhu, Z.; Li, G.; Domenzain, I.;
898 Marcisauskas, S.; Anton, P. M.; Lappa, D.; Lieven, C.; Beber, M. E.;
899 Sonnenschein, N.; Kerkhoven, E. J.; Nielsen, J. A consensus S.
900 cerevisiae metabolic model Yeast8 and its ecosystem for comprehen-
901 sively probing cellular metabolism. *Nat. Commun.* **2019**, *10*, 3586.
- 902 (35) Sánchez, B. J.; Zhang, C.; Nilsson, A.; Lahtvee, P.-J.; Kerkhoven,
903 E. J.; Nielsen, J. Improving the phenotype predictions of a yeast
904 genome-scale metabolic model by incorporating enzymatic constraints.
905 *Mol. Syst. Biol.* **2017**, *13*, 935.
- 906 (36) Flamholz, A.; Noor, E.; Bar-Even, A.; Liebermeister, W.; Milo, R.
907 Glycolytic strategy as a tradeoff between energy yield and protein cost.
908 *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 10039–10044.
- 909 (37) Tang, Y. J.; Martin, H. G.; Myers, S.; Rodriguez, S.; Baidoo, E. E.
910 K; Keasling, J. D. Advances in analysis of microbial metabolic fluxes via
911 13C isotopic labeling. *Mass Spectrom. Rev.* **2009**, *28*, 362–375.
- 912 (38) Nießer, J.; Müller, M. F.; Kappelmann, J.; Wiechert, W.; Noack,
913 S. Hot isopropanol quenching procedure for automated microtiter plate
914 scale 13C-labeling experiments. *Microb. Cell Fact.* **2022**, *21*, 78.
- 915 (39) Hollinshead, W. D.; Henson, W. R.; Abernathy, M.; Moon, T. S.;
916 Tang, Y. J. Rapid metabolic analysis of Rhodococcus opacus PD630 via
917 parallel 13C-metabolite fingerprinting. *Biotechnol. Bioeng.* **2016**, *113*,
918 91–100.
- 919 (40) Varma, A.; Palsson, B. O. Stoichiometric flux balance models
920 quantitatively predict growth and metabolic by-product secretion in
921 wild-type Escherichia coli W3110. *Appl. Environ. Microbiol.* **1994**, *60*,
922 3724–3731.
- 923 (41) DeLorenzo, D. M.; Henson, W. R.; Moon, T. S. Development of
924 Chemical and Metabolite Sensors for Rhodococcus opacus PD630.
925 *ACS Synth. Biol.* **2017**, *6*, 1973–1978.
- 926 (42) Bligh, E. G.; Dyer, W. J. A rapid method of total lipid extraction
927 and purification. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
- 928 (43) Oh, Y.-K.; Palsson, B. O.; Park, S. M.; Schilling, C. H.;
929 Mahadevan, R. Genome-scale Reconstruction of Metabolic Network in
930 Bacillus subtilis Based on High-throughput Phenotyping and Gene
931 Essentiality Data*. *J. Biol. Chem.* **2007**, *282*, 28791–28799.
- 932 (44) Kitts, P. A.; Church, D. M.; Thibaud-Nissen, F.; Choi, J.; Hem,
933 V.; Sapojnikov, V.; Smith, R. G.; Tatusova, T.; Xiang, C.; Zherikov, A.;
934 DiCuccio, M.; Murphy, T. D.; Pruitt, K. D.; Kimchi, A. Assembly: a
935 resource for assembled genomes at NCBI. *Nucleic Acids Res.* **2016**, *44*,
936 D73–D80.
- 937 (45) Courtot, M.; Juty, N.; Knüpfer, C.; Waltemath, D.; Zhukova, A.;
938 Dräger, A.; Dumontier, M.; Finney, A.; Golebiewski, M.; Hastings, J.;
939 Hoops, S.; Keating, S.; Kell, D. B.; Kerrien, S.; Lawson, J.; Lister, A.; Lu,
940 J.; Machne, R.; Mendes, P.; Pocock, M.; Rodriguez, N.; Villegas, A.;
941 Wilkinson, D. J.; Wimalaratne, S.; Laibe, C.; Hucka, M.; Le Novère, N.
942 Controlled vocabularies and semantics in systems biology. *Mol. Syst.*
943 *Biol.* **2011**, *7*, 543.
- 944 (46) Österlund, T.; Nookae, I.; Bordel, S.; Nielsen, J. Mapping
945 condition-dependent regulation of metabolism in yeast through
946 genome-scale modeling. *BMC Syst. Biol.* **2013**, *7*, 36.
- 947 (47) Castro, A. R.; Rocha, I.; Alves, M. M.; Pereira, M. A. Rhodococcus
948 opacus B4: a promising bacterium for production of
949 biofuels and biobased chemicals. *AMB Express* **2016**, *6*, 35.