



High-Quality Genome-Scale Reconstruction of *Corynebacterium glutamicum* ATCC 13032

Martina Feierabend^{1,2†}, Alina Renz^{1,2†}, Elisabeth Zelle³, Katharina Nöh³, Wolfgang Wiechert^{3,4} and Andreas Dräger^{1,2*}

¹ Computational Systems Biology of Infections and Antimicrobial-Resistant Pathogens, Institute for Bioinformatics and Medical Informatics (BMI), University of Tübingen, Tübingen, Germany, ² Department of Computer Science, University of Tübingen, Tübingen, Germany, ³ Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany, ⁴ Computational Systems Biotechnology (AVT.CSB), RWTH Aachen University, Aachen, Germany

OPEN ACCESS

Edited by:

Yu Wang,
Tianjin Institute of Industrial
Biotechnology, Chinese Academy of
Sciences (CAS), China

Reviewed by:

Hongwu Ma,
Tianjin Institute of Industrial
Biotechnology, Chinese Academy of
Sciences (CAS), China
Ilya R. Akberdin,
Biosoft.ru, Russia

*Correspondence:

Andreas Dräger
andreas.draeger@uni-tuebingen.de

†These authors share first authorship

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 30 July 2021

Accepted: 19 October 2021

Published: 15 November 2021

Citation:

Feierabend M, Renz A, Zelle E, Nöh K,
Wiechert W and Dräger A (2021)
High-Quality Genome-Scale
Reconstruction of *Corynebacterium
glutamicum* ATCC 13032.
Front. Microbiol. 12:750206.
doi: 10.3389/fmicb.2021.750206

Corynebacterium glutamicum belongs to the microbes of enormous biotechnological relevance. In particular, its strain ATCC 13032 is a widely used producer of L-amino acids at an industrial scale. Its apparent robustness also turns it into a favorable platform host for a wide range of further compounds, mainly because of emerging bio-based economies. A deep understanding of the biochemical processes in *C. glutamicum* is essential for a sustainable enhancement of the microbe's productivity. Computational systems biology has the potential to provide a valuable basis for driving metabolic engineering and biotechnological advances, such as increased yields of healthy producer strains based on genome-scale metabolic models (GEMs). Advanced reconstruction pipelines are now available that facilitate the reconstruction of GEMs and support their manual curation. This article presents *i*CGB21FR, an updated and unified GEM of *C. glutamicum* ATCC 13032 with high quality regarding comprehensiveness and data standards, built with the latest modeling techniques and advanced reconstruction pipelines. It comprises 1042 metabolites, 1539 reactions, and 805 genes with detailed annotations and database cross-references. The model validation took place using different media and resulted in realistic growth rate predictions under aerobic and anaerobic conditions. The new GEM produces all canonical amino acids, and its phenotypic predictions are consistent with laboratory data. The *in silico* model proved fruitful in adding knowledge to the metabolism of *C. glutamicum*: *i*CGB21FR still produces L-glutamate with the knock-out of the enzyme pyruvate carboxylase, despite the common belief to be relevant for the amino acid's production. We conclude that integrating high standards into the reconstruction of GEMs facilitates replicating validated knowledge, closing knowledge gaps, and making it a useful basis for metabolic engineering. The model is freely available from BioModels Database under identifier MODEL2102050001.

Keywords: *Corynebacterium glutamicum*, genome-scale metabolic model, constraint-based reconstruction, optimization, metabolic engineering, FAIR, flux balance analysis, MEMOTE

1. INTRODUCTION

The strain *Corynebacterium glutamicum* ATCC 13032 is a Gram-positive, facultatively anaerobic soil bacterium, which produces L-glutamate under particular treatments or growth conditions (Kimura, 2005). The annual production of several tons of L-glutamate (Eggeling and Bott, 2005) as well as other metabolically engineered products, such as other amino acids (Eggeling and Bott, 2015; Wendisch et al., 2016), alcohols (Inui et al., 2004a; Niimi et al., 2011; Yamamoto et al., 2013; Jojima et al., 2015), biopolymers (Liu et al., 2007), organic acids (Hüser et al., 2005; Okino et al., 2008; Takeno et al., 2013), terpenoids (Heider et al., 2014; Kang et al., 2014) or diamines (Kind et al., 2010a,b; Schneider and Wendisch, 2010), have turned *C. glutamicum* into a versatile and enormously relevant biotechnological microorganism. Despite an ongoing biotechnological application of *C. glutamicum* and the resulting knowledge on this bacterium for more than 70 years (Vertes et al., 2013), its metabolic potential not yet exhausted. Due to the prominent role of *C. glutamicum* in biotechnology, obtaining a more profound understanding of its physiology and metabolism is highly desirable.

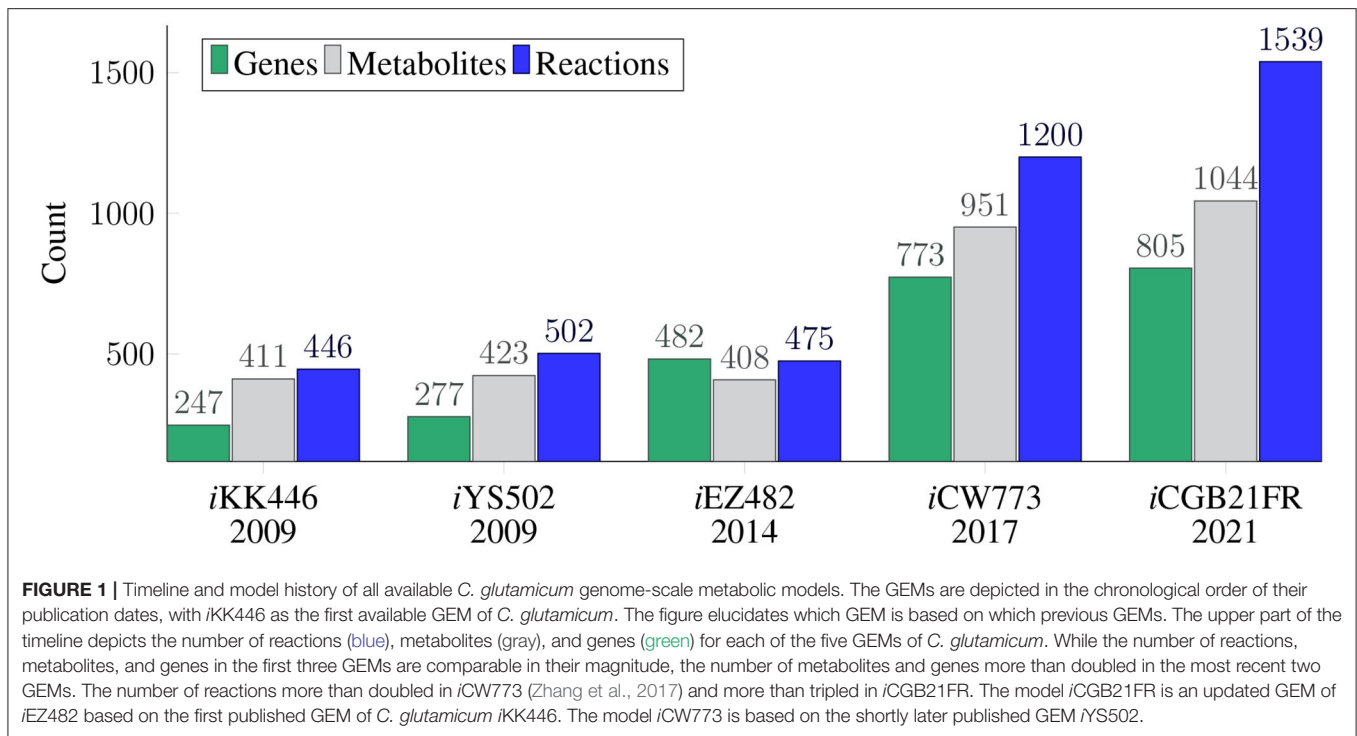
One method of formalizing this knowledge is a genome-scale metabolic network reconstruction. Genome-scale metabolic network reconstructions represent a systematic knowledge base of bibliomic and genomic data of all known metabolic reactions of a specific target organism (Thiele and Palsson, 2010). By creating a mathematical representation of the reconstructed network, the network can be changed into a genome-scale metabolic model (GEM). GEMs enable the qualitative description of the genotype-phenotype relationship and predictions of various phenotypes (Fang et al., 2020).

GEMs can be constructed by mapping the annotated genome sequence with its genes via the encoded proteins to reactions. This step is followed by an intensive curation phase of the computational model and a subsequent analysis phase. Prevalent methods for analyzing GEMs are summarized under the term constraint-based modeling. The main advantage of these modeling techniques over other approaches, such as dynamic modeling (Dräger et al., 2009), lies in their potential to analyze entire metabolic networks at the scale of all enzymatic capabilities of an organism without the necessity of knowing numerical values of all the kinetic parameters therein. Flux sampling can be used as an unbiased way to characterize the space of stoichiometrically feasible fluxes and solutions (Jadebeck et al., 2020). Flux balance analysis (FBA) is a biased method for steady-state analysis of GEMs. By imposing further physiologically realistic, relevant constraints and a target objective function on the computational model, the network's metabolic flux distributions can be simulated (Fang et al., 2020). Nevertheless, increasing network scale results in an increasingly complex process of reconstructing all cellular properties in the form of a coherent computer model.

In recent years, new tools and automated techniques in systems biology have emerged, such as CarveMe (Machado et al., 2018), ModelPolisher (Römer et al., 2016), MEMOTE (Lieven et al., 2020), or BOFdat (Lachance et al., 2019). These tools

support the reconstruction, refinement, and validation of GEMs using Minimal Information Required In the Annotation of Models (MIRIAM) standards (Le Novère et al., 2005). Several GEMs of the *C. glutamicum* have already been published (e.g., Kjeldsen and Nielsen, 2009; Shinfuku et al., 2009, see **Figure 1**). However, these models were curated before the newly developed tools were available. Thus, these new tools were so far not applied to GEMs of *C. glutamicum*. The most recently published GEM of *C. glutamicum* is *iCW773* (Zhang et al., 2017), which is based on Shinfuku et al. (2009). The model *iCW773* can produce all canonical amino acids. The production rates of amino acids are generally lower than experimental results (Eggeling and Bott, 2005). Comparing these production rates to those of other published GEMs of *C. glutamicum* is difficult since neither the composition of the complete medium nor the medium used for the *in silico* experiments is reported. Based on the MEMOTE report of *iCW773*, the model seems to lack stoichiometric consistency and contains no Systems Biology Ontology (SBO) terms (see below for more information on SBO terms Courtot et al., 2011). It contains 98 orphan and 116 dead-end metabolites. In the respiratory chain, the metabolites ubiquinone and its derivatives are used. However, several experimental studies confirmed that the only respiratory quinones in *C. glutamicum* are menaquinone and its derivatives (Kanzaki et al., 1974; Collins et al., 1977, 1979; Bott and Niebisch, 2003; Maeda et al., 2020). After conversion to Systems Biology Markup Language (SBML) Level 3 Version 1 (Hucka et al., 2018), *iCW773* reaches a total MEMOTE score of only 29% (Lieven et al., 2020, see below for more information on MEMOTE and this score). Newly available tools such as MEMOTE have not yet been applied to reconstruct any previous GEM of *C. glutamicum*. The goal of this model is to fill this application gap. Given its importance as a biotechnological microbe, an updated GEM reflecting the current state of knowledge about *C. glutamicum* and incorporating the scope of newly available tools is indispensable.

In this study, we present an updated GEM of high quality for *C. glutamicum* named *iCGB21FR*. It combines the knowledge about *C. glutamicum* from the previous models *iKK446* (Kjeldsen and Nielsen, 2009) and *iEZ482* (Zelle et al., 2015) and extends it by including a broader metabolic coverage than previous models. This GEM was reconstructed using the latest available *in silico* methods and tools and represents a model composed of the most current standards in systems biology. Furthermore, this GEM uses current community standards and follows the best-practice recommendations by Carey et al. (2020). High quality in terms of GEM reconstruction encompasses several aspects, such as a fully annotated GEM in terms of metabolites, reactions, and genes with gene-protein-reaction (GPR) associations. In addition, SBO terms (Courtot et al., 2011) are included in the model. These allow a more fine-grained description of the respective compound. With the aid of the high-quality reconstruction of the GEM, we reproduced experimentally validated findings. This model allows a more accurate *in silico* depiction of the genetic makeup of *C. glutamicum*. The new model *iCGB21FR* contributes to filling knowledge gaps in the metabolism of *C. glutamicum* by providing further information



on relevant pathways used in the production of L-glutamate. Finally, this model uses FAIR data standards (findable, accessible, interoperable, reusable; Wilkinson et al., 2016). Access to all data and metadata used in this model is provided. A highly detailed annotation level within the model is used, and the reconstruction process is described as transparently as possible (Carey et al., 2020).

2. MATERIALS AND METHODS

2.1. The Metabolic Network Reconstruction Process

2.1.1. Strain

The GEM of the strain *Corynebacterium glutamicum* ATCC 13032 was reconstructed using the annotated genome sequence (accession number: NC006958.1), which was downloaded from the National Center for Biotechnology Information (NCBI) at <https://www.ncbi.nlm.nih.gov> (Agarwala et al., 2018).

2.1.2. Draft Reconstruction

The reconstruction process closely followed the protocol by Thiele and Palsson (2010). In short, an automated draft reconstruction was created using CarveMe (Machado et al., 2018), version 1.2.2, and stored in the SBML Level 3 Version 1 format (Hucka et al., 2018). The SBML Level 3 extension for flux balance constraints (fbc) version 2 by Olivier and Bergmann (2018) was enabled and used under default settings for the draft reconstruction. SBML represents a machine-readable exchange format that allows manipulating computational models of biological processes (Keating et al., 2020; Renz et al., 2020). The

fbc plugin enables adding structured, semantic descriptions for domain-specific model components such as charges, annotations, flux bounds, GPR rules, or chemical formulas of metabolites (Lieven et al., 2020). This initial draft contained 1496 reactions, 1030 metabolites, and 782 genes in the three compartments: extracellular, cytosol, and the periplasm.

Further automated and manual refinement of the reconstruction of *C. glutamicum* was performed using libSBML (Bornstein et al., 2008), version 5.18.0, and COBRApy (Ebrahim et al., 2013), version 0.17.1. All simulations were run using the CPLEX optimizer, version 12.10 by IBM (<https://www.ibm.com/analytics/cplex-optimizer>). Metabolic pathways were visualized using the Escher software (King et al., 2015). To support the display as standardized Process Description (PD) map (Rougnny et al., 2019) in Systems Biology Graphical Notation (SBGN) enabled software (Touré et al., 2020), the Escher maps were converted to the SBGN Markup Language (SBGNML) format (Bergmann et al., 2020) using EscherConverter (<https://github.com/draeger-lab/EscherConverter>).

2.1.3. Annotations

Cross-references of the model's instances to other databases were shifted from the notes to the annotations field. Additional metadata, such as annotations and cross-references, was added using the ModelPolisher (Römer et al., 2016). The model's genes were annotated using the old and new locus tags from NCBI and the NCBI protein identifier. SBO terms (Courtot et al., 2011) further annotate the model's instances. SBO terms represent controlled vocabularies, which provide semantic information about model components. For metabolites and genes, the general SBO-terms for simple chemical (SBO:0000247) and genes

(SBO:0000243) were used, respectively. The SBO terms for the reactions were chosen as precisely as possible using a new curation pipeline (Fritze, 2020).

2.1.4. Refinement of Metabolite Attributes

The draft was curated to include the correct positioning of the metabolites' chemical formulas and charges. All charges were obtained, if more than one charge per compound was available, in the Biochemically, Genetically, and Genomically structured (BiGG) Models database (Norsigian et al., 2019). In the following verification step, the most appropriate charge for a given reaction in a specific compartment was manually chosen and added to the model. Dead-end metabolites and orphan metabolites were identified and, when appropriate, removed.

2.1.5. Manual Extension

Intensive manual curation was done using the databases BiGG (Norsigian et al., 2019), MetaCyc (Caspi et al., 2020), BioCyc (Karp et al., 2019), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2019), and new bibliomic data. This draft was then revised using the *iEZ482* model (Zelle et al., 2015) as a reference. The model *iEZ482* is an updated version of the *iKK446* model (Kjeldsen and Nielsen, 2009) and contains 475 reactions, 408 metabolites, and 482 genes. Reactions, metabolites, and genes present in *iEZ482* but not in *iCGB21FR* were manually checked in MetaCyc (Caspi et al., 2020) or BioCyc (Karp et al., 2019) for their biochemical relevance in the model and, if appropriate, added. Altogether, 50 new reactions, 14 new metabolites, and 23 new genes were added to *iCGB21FR*. BiGG identifiers (IDs) and annotations were included in the model for all newly added compounds, thus enabling easier comparison with other models. If BiGG IDs were not yet existent, BioCyc IDs (Karp et al., 2019) and additional annotations such as SBO terms were added to the new instance.

2.1.6. Mass and Charge Imbalances

The chemical formulas of all participating metabolites were verified. All mass and charge imbalanced reactions were manually checked. Pseudo-reactions, including exchange, sink, or biomass reactions, were excluded from this curation step. For reactions with imbalanced charge, the charge of every participating metabolite was verified and, if necessary, adapted. Mass imbalanced reactions were checked for missing metabolites, such as protons.

2.1.7. Energy-Generating Cycles

Energy-generating cycles represent thermodynamically infeasible states. Charging of energy metabolites without any energy source causes such cycles (Fritze et al., 2017). If left undetected in the model, these can result in erroneous increases in maximal yields in the biomass (Fritze et al., 2017). The following 13 carrier metabolites for energy or redox equivalent were tested for their ability to form thermodynamically infeasible cycles: adenosine triphosphate (ATP), cytidine triphosphate (CTP), guanosine triphosphate (GTP), uridine triphosphate (UTP), inosine triphosphate (ITP), reduced nicotinamide adenine dinucleotide (NADH), reduced nicotinamide adenine dinucleotide phosphate (NADPH), flavin

adenine mononucleotide (FMN), flavin adenine dinucleotide (FAD), menaquinol-8, 2-demethylmenaquinol 8, acetyl-CoA, and L-glutamate. All exchange reactions of the model were set to $0 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to investigate the presence of energy-generating cycles. Energy dissipating reactions were created for each of the 13 individual metabolites. These allow the corresponding metabolite to be removed from the system. Each reaction was added one-at-a-time to the model and then used as the objective function. If the optimization returned a result unequal to zero, an energy-generating cycle was detected and subsequently removed. Additionally, the proton exchange between cytosol and periplasm was included.

2.1.8. Biomass Objective Function

The initial biomass objective function (BOF) of *iCGB21FR* was created using CarveMe (Machado et al., 2018). It represents a universal bacterial biomass objective function (BOF). The species-specific biomass objective function (BOF) was further refined using BOFdat (Lachance et al., 2019). BOFdat allows calculating and refining a pseudo-reaction for the biomass function without using any pseudo-metabolites or macromolecules, such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA), or protein. The nucleotide sequence of *C. glutamicum* ATCC 13032 was used to refine the DNA nucleotides in the BOF. Coenzymes and inorganic ions were identified and specifically adapted for *C. glutamicum* in the BOF within the second step of BOFdat. As the model initially did not simulate growth on the minimal medium CGXII (see section 2.2.2), trace elements in the BOF were compared to the elemental composition of *C. glutamicum* cells (Liebl, 2005). Based on this comparison, cobalt was removed from the BOF.

2.1.9. Subsystems and Groups Plugin

Biological pathways were obtained from the KEGG database (Kanehisa et al., 2019) using the old locus tags in the genes' annotations. Pathways associated with a reaction were added to the reaction's annotations based on genes in the GPR association. The pathways were added as a biological qualifier with the attribute OCCURS_IN. Additionally, the groups plugin was enabled, available for SBML Level 3. The groups plugin in libSBML (Bornstein et al., 2008) allows a more flexible grouping of specific connected components in the metabolic model (Hucka and Smith, 2016). The groups plugin was used to add every metabolic pathway or subsystem as a group. Participating reactions were then added to the groups as members.

2.1.10. Quality Control

The quality of the GEM was tested performing a FROG analysis (König, 2020) and using MEMOTE, version 0.11.1. MEMOTE is a platform to test standardized measures of metabolic models and outputs quality scores ranging from 0% for poor model quality to 100% for excellent model quality (Lieven et al., 2020). The measures that generate the MEMOTE scores evaluate the model's consistency and annotations within different categories. These categories include basic information about the model, the metabolites and reactions, the degree of annotations for metabolites, reactions, genes, and SBO terms. MEMOTE also

checks the presence of GPRs, a realistic biomass function, energy metabolism, and appropriate network topology. Apart from these individual MEMOTE scores for the different subcategories, MEMOTE also reports an overall score. This overall score represents an overall measurement of how well the model scored within all individual categories. To evaluate the consistency of the model, the stoichiometric consistency, mass and charge balances, metabolite connectivity, and unbounded fluxes in the default medium were used. Within the evaluation of the annotations, MEMOTE checks for the presence and conformity of various databases and the presence of specific SBO terms. All categories are scored individually. The overall MEMOTE quality score is calculated based on the individual category scores (Lieven et al., 2020).

2.1.11. Curation of *iCW773*

The model *iCW773* (Zhang et al., 2017) was downloaded in Microsoft Excel format as the supplementary published and converted to Character-Separated Value (CSV) format. The application Table2Model (Dräger, 2021) was developed based on JSBML (Rodriguez et al., 2015) to parse the CSV files and convert the information to SBML Level 3 Version 1 (Hucka et al., 2018). Since the original publication did not explicitly define any units, these had to be added to the model. For consistency reasons, the units were defined in the same way as for *iCGB21FR*. The generated SBML Level 3 Version 1 file was syntactically validated using a combination of JSBML (Rodriguez et al., 2015) and libSBML (Bornstein et al., 2008), including unit consistency validation. MEMOTE version 0.11.1 (Lieven et al., 2020) was used for semantic model checking. Annotation of the model *iCW773* was performed using the same curation pipeline described above with the help of ModelPolisher (Römer et al., 2016) and SBO term addition (Fritze, 2020). The model was wrapped in an (Bergmann et al., 2014) OMEX archive file (Neal et al., 2018) together with a metadata file and uploaded to BioModels Database (Malik-Sheriff et al., 2020), where it is available under accession MODEL2110010001 (see Availability).

2.2. Model Validation

All model validations were performed with a physiological pH of 7.0. The growth behavior was tested in several media with access to varying carbon sources under aerobic and anaerobic conditions to validate the predictive power of the curated model *iCGB21FR*.

2.2.1. Definition of the Growth Unit

The growth rate is defined as the flux through the biomass objective function, which corresponds to the system's biomass-producing reaction. In their fundamental work from Varma and Palsson (1994) explain that “ V_{gro} is the growth flux (grams of biomass produced), which with the basis of 1 g (dry weight) per h reduces to the growth rate (grams of biomass produced per gram [dry weight] per hour).” It should be noted that 1 gDW corresponds to 1 g with a semantic annotation regarding the dry weight fraction of a probe. Gottstein et al. (2016) explain that the metabolic fluxes are typically given in $\text{mmol gDW}^{-1} \text{h}^{-1}$ and confirm (Varma and Palsson, 1994) that the growth rate

μ has the unit $\text{g gDW}^{-1} \text{h}^{-1}$. Gottstein et al. (2016) also state that the biomass objective function describes the accumulation of biomass components per hour and relative to the amount of biomass in gDW. Consequently, all molecular species need to be expressed in the unit mmol gDW^{-1} , which corresponds to the amount of the biomass component per gram of biomass (cf. section 2.1; Gottstein et al., 2016). Since all stoichiometric coefficients have dimensionless units, the biomass forming reaction can be considered a summation of components in mmol gDW^{-1} , each times a dimensionless factor. Consequently, the rate of this reaction, which defines a change per time, results in $\text{mmol gDW}^{-1} \text{h}^{-1}$.

Accordingly, the SBML specification defines that the units of all reactions in a model have to be identical and are defined in units of *extent per time* (see Hucka et al., 2018, section 4.2.5;). According to the specification of SBML Level 3 Version 1 Release 2 (see Hucka et al., 2018, Table 9), the extent units should be substance units or a combination of units derived from those. Here, the extent of the reactions and the substance units of all compounds are defined in units of mmol gDW^{-1} (note that in contrast to Varma and Palsson (1994), we here define the biomass in units of mmol instead of in g). The time units are defined in h (or 3600 s). Hence, all reactions have the unit $\text{mmol gDW}^{-1} \text{h}^{-1}$. It should be noted that the upper and lower bounds of all reactions have the same unit and are therefore consistently defined with the flux through the biomass reaction. In this way, these parameters already implicitly define the flux units because the flux's upper and lower bounds must have the same unit as the flux itself.

For more information, readers may also consider the specification of the SBML extension package fbc (Olivier and Bergmann, 2018), which provides similar examples in its appendix, and the detailed analysis on this matter outlined by Gottstein et al. (2016). To improve the units' definition, *iCGB21FR* and *iCW773* explicitly declare the attributes `extentUnits` and `timeUnits` within the model element in their SBML files. It also declares `substanceUnits` in mmol gDW^{-1} and the `volumeUnits` in `fl` so that all compounds and compartments inherit defined units from the model container.

Experimentally observed growth rates μ may be given in the unit 1/h. In this case, directly comparing the calculated growth rate to the experimentally obtained value is possible if the biomass consistency of a GEM approaches $1 \text{ mmol gDW}^{-1} \text{h}^{-1}$ because then its produced biomass has a molecular weight of 1 g mmol^{-1} . With this, the conversion $1 \text{ g gDW}^{-1} \text{h}^{-1} = 1 \text{ g g}^{-1} \text{h}^{-1} = 1 \text{ h}^{-1}$ can be performed because the biomass of the GEM is, in this case, *standardized*. A direct comparison of growth rates is then valid, because with a biomass consistency close to $1 \text{ mmol gDW}^{-1} \text{h}^{-1}$, the different units of the growth rate μ converge.

2.2.2. Growth in Different Media and Conditions

Following common laboratory practice in cultivating *C. glutamicum*, the complete lysogeny broth (LB) medium (Bertani, 1951) and the two minimal media M9 (Sambrook et al., 1989) and CGXII (Keilhauer et al., 1993; Eggeling and

Bott, 2005) were chosen to simulate *in silico* aerobic growth of *C. glutamicum*. Transporters for the inorganic ions nickel and calcium had to be added to allow growth on the M9 minimal medium. As protocatechuic acid is a component of the CGXII medium (Keilhauer et al., 1993), all necessary exchange and transport reactions were added to model the uptake this compound. The model *i*CGB21FR did initially not grow on the minimal medium CGXII. Literature research pointed toward cobalt in the BOF as a potential issue. Removing cobalt from the BOF allowed growth on CGXII.

D-glucose served as the predominant carbon source in the two minimal media. The composition of each medium was used to constrain the model's exchange reactions with the environment. For simulating growth in the three different media, the lower bounds of the metabolites' exchange reactions available in the respective medium were set to the default value $-10 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to enable the uptake. All other exchange reactions' lower bounds were set to $0 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. While applying these medium-specific constraints, the BOF was set as the objective function. If the model did not simulate growth on one of the experimentally confirmed media, literature was queried to identify missing metabolites or reactions hampering growth. These were then added to *i*CGB21FR.

C. glutamicum is a facultative anaerobe microbe (Eggeling and Bott, 2005, 440). The growth under anaerobic conditions was evaluated to demonstrate the validity of *i*CGB21FR. The model initially created with CarveMe (Machado et al., 2018) did not simulate growth when applying anaerobic conditions by blocking the oxygen uptake. The model was evaluated using flux balance analysis (FBA) to identify relevant oxygen-carrying reactions to identify potential reasons for this. Additionally, literature was searched to find alternative or missing reactions. Furthermore, the gap-filling option of CarveMe was used for the M9 minimal medium under anaerobic conditions. To this end, a novel draft model with CarveMe was created, where the gap-filling option was enabled during the curation step. The reaction set of the gap-filled model was compared to our extended *i*CGB21FR model's reaction set, and the missing reactions were added. These six missing reactions include the catalase reaction (CAT), the succinate dehydrogenase (SUCDi), the phosphoribosylformylglycinamide synthase (PRFGS_1), a different calcium transporter (CAT4), the fumarate reductase (FRD7), and the glycolate transport via proton symport (GLYCLTt2rpp). With the inclusion of these reactions, the model simulated anaerobic growth on all three tested media.

The *C. glutamicum*-specific CGXII minimal medium was used to test the model's growth behavior on different carbon sources. The metabolites glucose, fructose, sucrose, ribose, gluconate, pyruvate, acetate, lactate, and propionate were tested under aerobic and anaerobic conditions as sole carbon sources since experimental data confirmed their role as carbon sources (Michel et al., 2015). All tested compounds could serve as the sole carbon source under aerobic conditions. Under anaerobic conditions, however, only glucose, fructose, sucrose, and ribose could serve as carbon and energy sources (Michel et al., 2015). Therefore, all nine carbon sources were tested *in silico* under aerobic and anaerobic conditions using the CGXII minimal

medium. If *i*CGB21FR did not simulate growth on one of the experimentally verified carbon sources, missing exchange and transport reactions were added based on results from a literature search. These included adding a pyruvate exchange and transport reaction and a lactate transporter for the aerobic condition. Further gap-filling steps were performed when necessary.

2.2.3. Verifying Capabilities for Amino Acid Production

The model was further validated by simulating the production of all 20 canonical amino acids in the CGXII medium and D-glucose as the predominant carbon source under aerobic conditions. The availability of D-glucose was restricted to the default uptake rate of $10 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. The growth rate was fixed to $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to ensure the microbe's maintenance during the amino acid production. Subsequently, a sink reaction was created for each amino acid, set as the objective function, and optimized. The relative amino acid production was calculated by dividing the total amino acid production rate by the glucose uptake rate. The same approach was taken for the CO_2 production rate, which was set in relationship to the amino acid production rate. The efflux of the CO_2 exchange reaction (EX_co2_e) was taken as the CO_2 production rate. The ATP production rate was calculated by summing up the fluxes of all ATP-producing reactions. These were then correlated with the amino acid production rate, analogously to the CO_2 production rate.

2.3. Model Application: New Insights for Metabolic Engineering

As *C. glutamicum* is widely used in biotechnology, the model's capabilities can be used to yield hints for metabolic engineering. All subsequent analyses were performed using the CGXII minimal medium with D-glucose as the sole carbon source under aerobic conditions.

2.3.1. Relation Between Growth and L-glutamate Production

A sink reaction (sink_glu_L) was added to optimize the L-glutamate production (see Figure 5). This reaction was then set as the objective function. As L-glutamate is also part of the biomass objective function, a potential association between these two reactions (BOF and L-glutamate sink reaction) was evaluated by varying the BOF between 0 and the maximum growth rate $0.57 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ while maximizing the sink reaction. $0.57 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ is the maximum *in silico* growth rate of *i*CGB21FR under aerobic growth conditions with D-glucose as the sole carbon source on CGXII (see section 3.2).

2.3.2. Relevance of the pyruvate carboxylase (PC)

PC is a pivotal enzyme in the L-glutamate production in *C. glutamicum* (Peters-Wendisch et al., 2001). A metabolic map was drawn, which depicts the primary reactions relevant for the L-glutamate production starting at D-glucose as the predominant carbon source using the tool Escher (King et al., 2015, see Figure 5). The model was optimized for the sink reaction for L-glutamate (sink_glu_L) using FBA while fixing the growth

rate to $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. This growth rate is an intentionally chosen growth rate within the interval where the L-glutamate production is only marginally affected by growth. The resulting flux distribution is depicted in **Figure 6**. This analysis was repeated after knocking out the PC's reaction to elucidate further the PC's effect on the metabolic flux distribution.

2.3.3. Identification of Relevant Reactions for L-glutamate Production

A loopless flux variability analysis (FVA) was used to identify reactions relevant to L-glutamate production. FVA represents a standard method to evaluate the range of feasible steady-state fluxes for each reaction by sequentially minimizing and maximizing each reaction (Schellenberger et al., 2011). Loop reactions are a subset of reactions with unbounded fluxes. Loopless FVA eliminates thermodynamically infeasible loops by not allowing the model to use these loops (Schellenberger et al., 2011). After running the loopless FVA, reactions with almost identical minimal and maximal allowed flux values were extracted (relative tolerance of 10^{-5} , absolute tolerance of 10^{-8}).

3. RESULTS

3.1. The Model *iCGB21FR* Is of High Quality

The new GEM of *Corynebacterium glutamicum* constructed in this work is named *iCGB21FR*. This name follows the latest recommended naming conventions, which are part of the community standardization of metabolic models (Carey et al., 2020). The lower-case “*i*” in italics means *in silico*, followed by the species indicator “CG” for *C. glutamicum*. “B” represents the city where the particular strain ATCC 13032 was sequenced (Bielefeld, Germany, see also Kalinowski et al., 2003). The three-letter code “CGB” also serves the corresponding strain identifier in the KEGG pathway database (Kanehisa et al., 2019). It follows an iteration identifier, in this case, the year 21 of this century. The last two characters, “FR,” refer to the last names of the primary model curators.

The model *iCGB21FR* is available in the SBML Level 3 Version 1 format (Hucka et al., 2019) with the fbc plugin (Olivier and Bergmann, 2018) and the groups plugin (Hucka and Smith, 2016) enabled. It contains 1042 metabolites, 1539 reactions, and 805 genes. Thus, further 442 reactions, 13 metabolites, and 25 genes were added to the model following the initial draft reconstructed with CarveMe (Machado et al., 2018). All metabolites and reactions have a human-readable, descriptive name and a chemical formula. The model comprises the cytosolic, periplasmic, and extracellular compartments.

Its overall MEMOTE score amounts to 87%. The MEMOTE score of the initial draft model created with CarveMe was 33%. With intensive manual curation, the number of mass and charge imbalanced reactions could be diminished from an initial 170 to 19 imbalanced reactions. These represent 1.2% of the total number of reactions. The model has a stoichiometric consistency of 99.7% and does not contain any energy-generating cycles, dead-end metabolites, nor orphan metabolites.

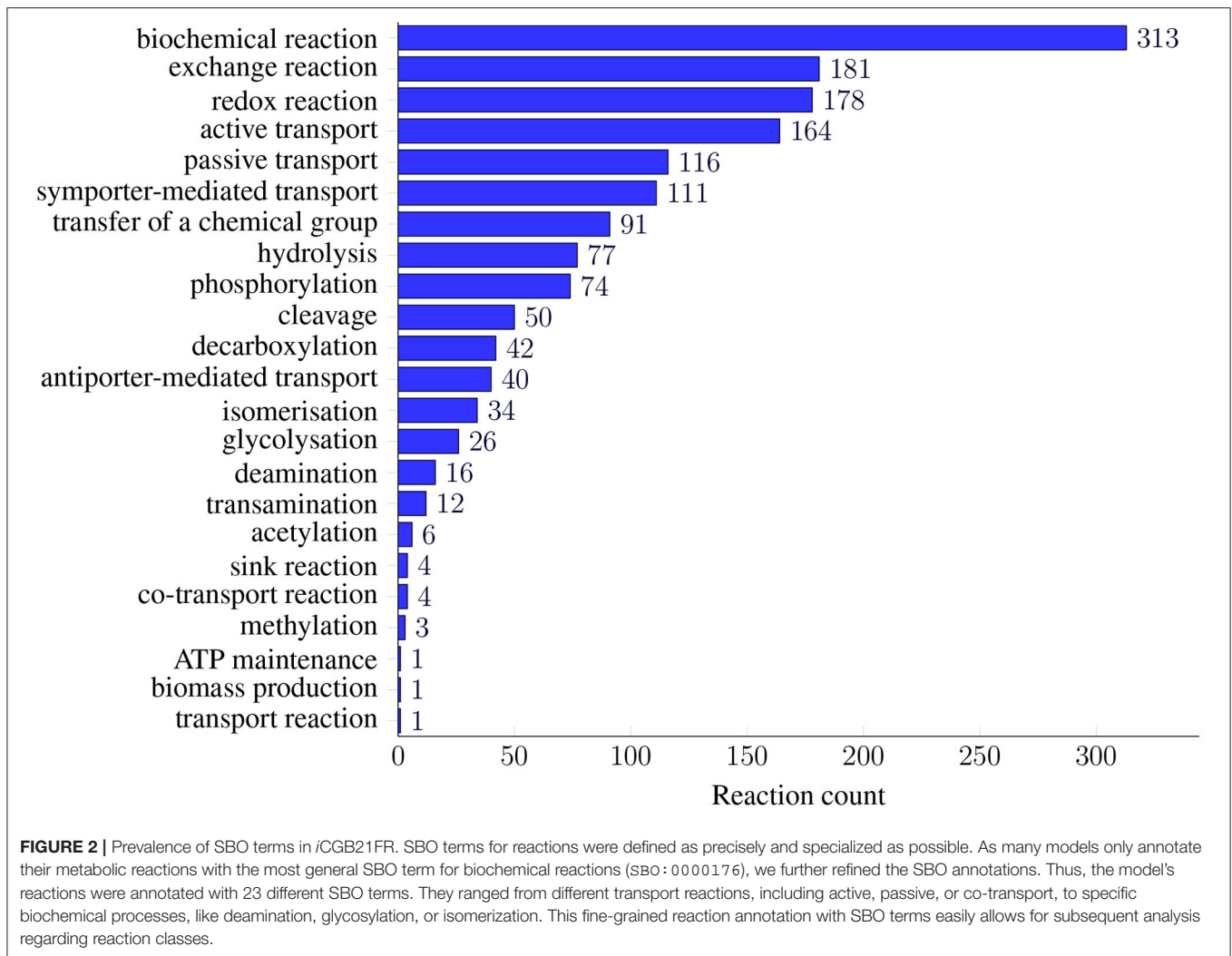
Seventeen different databases are cross-referenced in the model's instances, yielding a MEMOTE annotation score of 84%

for reactions, 84% for metabolites, and 49% for genes. Genes include cross-references to the three databases KEGG (Kanehisa et al., 2019), NCBI genes (Maglott et al., 2005), and NCBI proteins (Pruitt et al., 2005). Metabolites and reaction annotations contain cross-references to 13 and seven different databases, respectively. The databases BiGG (Norsigian et al., 2019), BioCyc (Karp et al., 2019), KEGG (Kanehisa et al., 2019), MetaNetX (Moretti et al., 2021), Reactome (Croft et al., 2010), and ModelSEED (Henry et al., 2010) are referenced for metabolites and reactions. Reactions also have cross-references to the RHEA database (Lombardot et al., 2019) and EC numbers. Metabolites have additional cross-references to the ChEBI database (Hastings et al., 2016), the Human Metabolome Database (HMDB) (Wishart et al., 2007), BioPath (Brandenburg et al., 2004), InChIKey (Heller et al., 2015), UniPathway (Morgat et al., 2011), lipid maps structure database (Sud et al., 2007), and the University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) (Ellis et al., 2003).

All model instances were further annotated using SBO terms (see **Figure 2**). While genes and metabolites received general SBO terms for genes and simple chemicals, the model's reactions were annotated with 23 different SBO terms. The most prominent ontology group is “biochemical reactions”: 313 reactions in the model hold the general SBO term for reactions. The number of biochemical reactions is followed by the group of exchange reactions with 181 reactions. The transport reactions are described more precisely by the SBO terms for active, passive, co-, symporter-mediated, antiporter-mediated, or general transport. For all other reactions, we identified SBO terms that describe the occurring biochemical reaction more precisely. In terms of ontology, these SBO terms are child nodes of the SBO term for biochemical reactions. These terms include, for example, redox reactions, the transfer of a chemical group, hydrolysis, or phosphorylation. The SBO terms for ATP maintenance and biomass production occur only once in the model. **Figure 2** gives an overview of all 23 added reaction SBO terms and their occurrence in the model.

The plugins fbc (Olivier and Bergmann, 2018) and groups (Hucka and Smith, 2016) are enabled in *iCGB21FR*, thus allowing information such as metabolic charges, chemical formulas, or gene products to be stored. All identified KEGG pathways (Kanehisa et al., 2019) were added as a group to the model, with all reactions participating in the pathway as group members. In total, 102 groups were added to the model. The group with the most members is the “metabolic pathways” group with 563 members, followed by the “biosynthesis of secondary metabolites” group with 297 members. Other groups with more than 100 members are the “biosynthesis of amino acids” with 104 associated reactions, the “biosynthesis of cofactors” with 151 members, and the group of reactions associated with “microbial metabolism in diverse environments” with 171 members. With the help of these groups, reactions of a particular pathway can easily be extracted and analyzed.

The biomass objective function (BOF) created by CarveMe was refined in several steps to obtain realistic growth rates for the tested media. With the help of BOFdat and the nucleotide sequence of *C. glutamicum* ATCC 13032, the stoichiometric

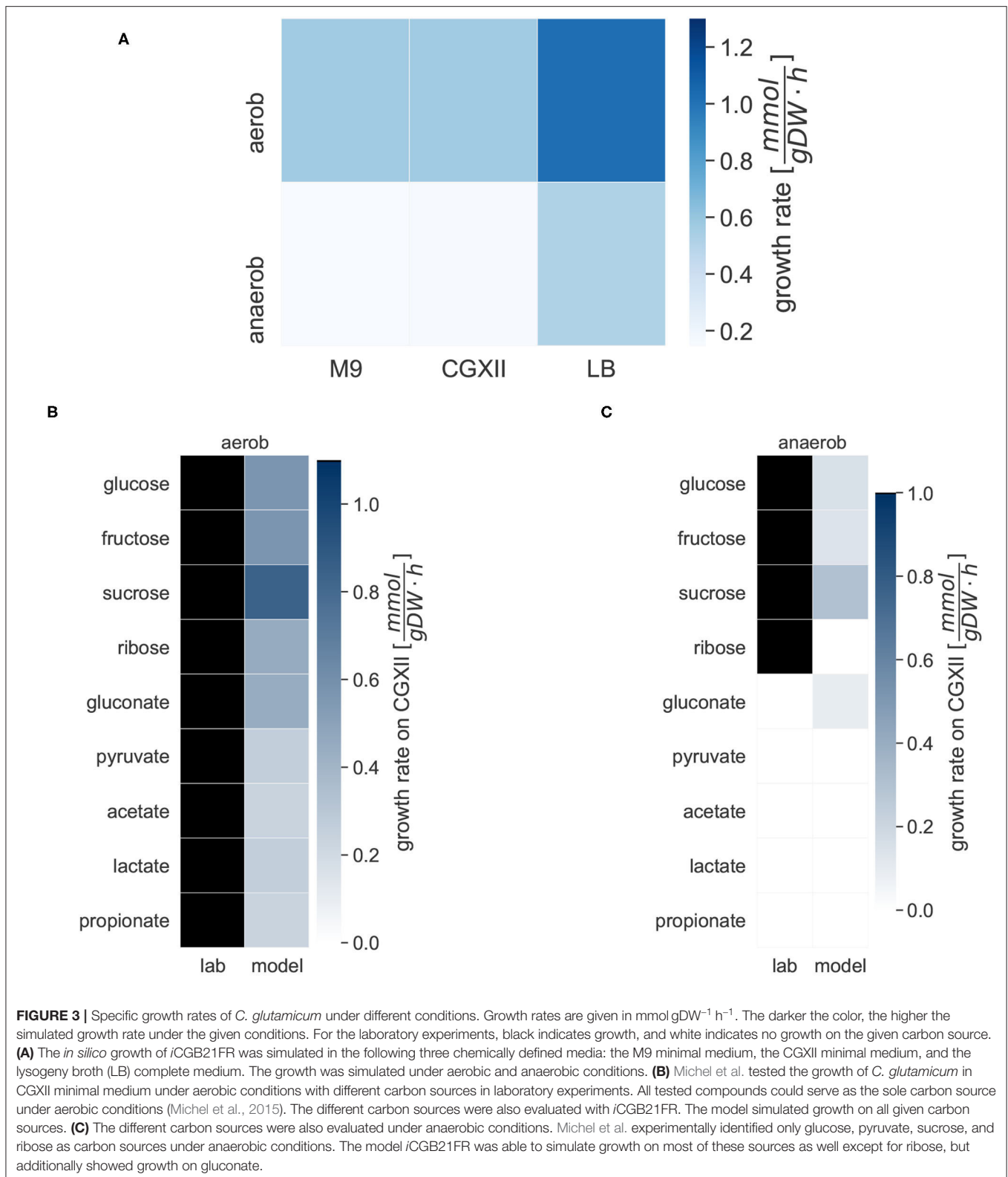


coefficients of the DNA nucleotides were adapted. The following seven metabolites were added as coenzymes and inorganic ions to the biomass objective function: NADH, NADPH, adenosine monophosphate (AMP), pyruvate, ammonium, sodium, and nickel. The stoichiometric coefficients of 16 further coenzymes and inorganic ions were adapted using BOFdat. The inorganic ion cobalt was removed from the BOF based on the elemental composition of *C. glutamicum* ATCC 13032 cells, as described by Eggeling and Bott (2005, p.16, 18). After including these changes, the simulated biomass production is in the range of a reasonable growth rate with no blocked biomass precursors in both the default and the complete medium.

3.2. Simulations of *iCGB21FR* Are Consistent With Experimental Data

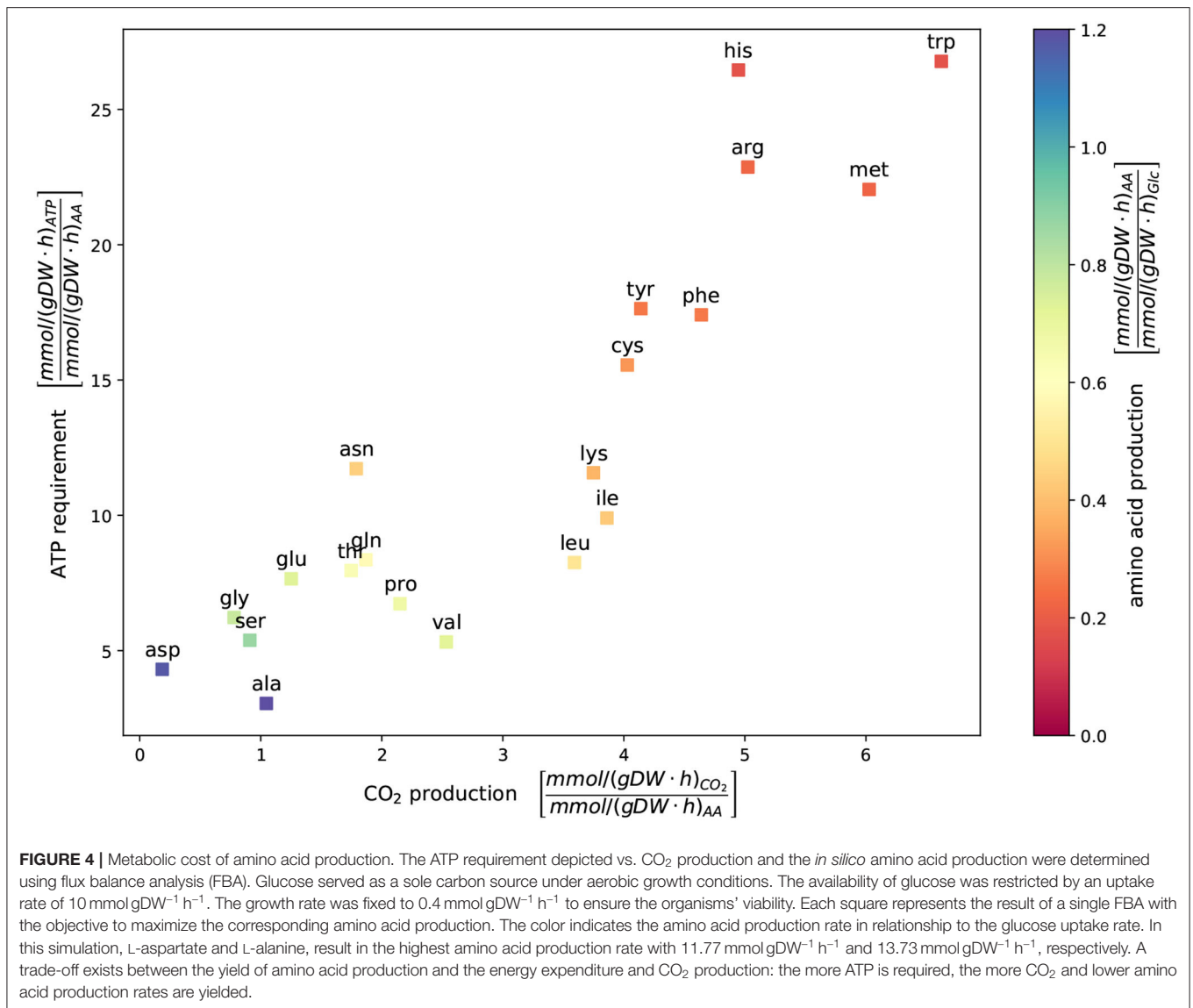
We simulated the growth of *iCGB21FR* in different media under aerobic and anaerobic conditions, and with access to different carbon sources (see Figure 3). Growth was tested on the two minimal media, M9 and CGXII, and the complete

LB medium. The heat map in Figure 3A gives an overview of the growth behavior of *C. glutamicum* in the three different media under aerobic and anaerobic conditions. With $1.0266 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ the biomass consistency of *iCGB21FR* is close to $1 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. Consequently, it is approximately possible to directly compare the *in-silico* growth rate to an experimentally obtained growth rate given in 1 h^{-1} . The simulated aerobic model growth on the minimal medium M9 with glucose as a single carbon source resulted in a maximal growth rate of $0.57 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. A maximal realistic aerobic growth rate of $0.57 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ on CGXII was obtained using the simulation tools COBRApy (Ebrahim et al., 2013) and confirmed with the Systems Biology Simulation Core Library (SBSCCL) (Panchiwala et al., 2021). The simulated value is only slightly lower than the growth rate of 0.61 h^{-1} that Unthan et al. (2014) could experimentally obtain. The model simulates growth on the complex LB medium with a growth rate of $1.0214 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ under aerobic conditions without further adjustments or refinements. As expected, the growth rate



in the complex medium (LB) is approximately twice as high as in the two minimal media (M9 and CGXII). The aerobic growth conditions in all three media show a higher simulated

growth rate compared to the anaerobic conditions, as anticipated. All growth rates are within a realistic range (Unthan et al., 2014).



The growth of *C. glutamicum* in CGXII minimal medium under aerobic and anaerobic conditions with varying carbon sources was tested. Under aerobic conditions, the model simulated biomass production on all carbon sources. Aerobic growth was possible on all carbon sources. The growth rates varied between 0.8437 mmol gDW⁻¹ h⁻¹ on sucrose and 0.2401 mmol gDW⁻¹ h⁻¹ on acetate. In our anaerobic *in silico* experiments, biomass production was also possible on three of the experimentally validated carbon sources, but not on ribose. Additionally, gluconate can be used as carbon sources. The biomass production on gluconate yielded a rate of 0.0945 mmol gDW⁻¹ h⁻¹.

The new model of *C. glutamicum* can simulate the production of all 20 canonical amino acids while growing on the CGXII medium with D-glucose as the carbon source under aerobic conditions. In **Figure 4**, each square represents the result of a

single FBA with the objective to maximize the corresponding amino acid production. The color indicates the amino acid production rate with respect to the glucose uptake rate. The positioning represents the ATP requirements and CO₂ production in relation to the amino acid production rate. The two amino acids, L-aspartate (asp) and L-alanine (ala), have the highest absolute amino acid production rates with 11.77 mmol gDW⁻¹ h⁻¹ and 13.73 mmol gDW⁻¹ h⁻¹, respectively (see also Zelle et al., 2015). In contrast, the amino acids L-histidine, L-arginine, and L-tryptophan have the lowest amino acid production rate and the highest ATP requirements. A relationship exists between the yield of amino acid production, energy expenditure, and CO₂ production: The more ATP is required, the more CO₂ produced and the lower is the amino acid production rate. L-glutamate is of particular interest for metabolic engineering in *C. glutamicum*. Its total production rate under the selected conditions yields 8.7 mmol gDW⁻¹ h⁻¹.

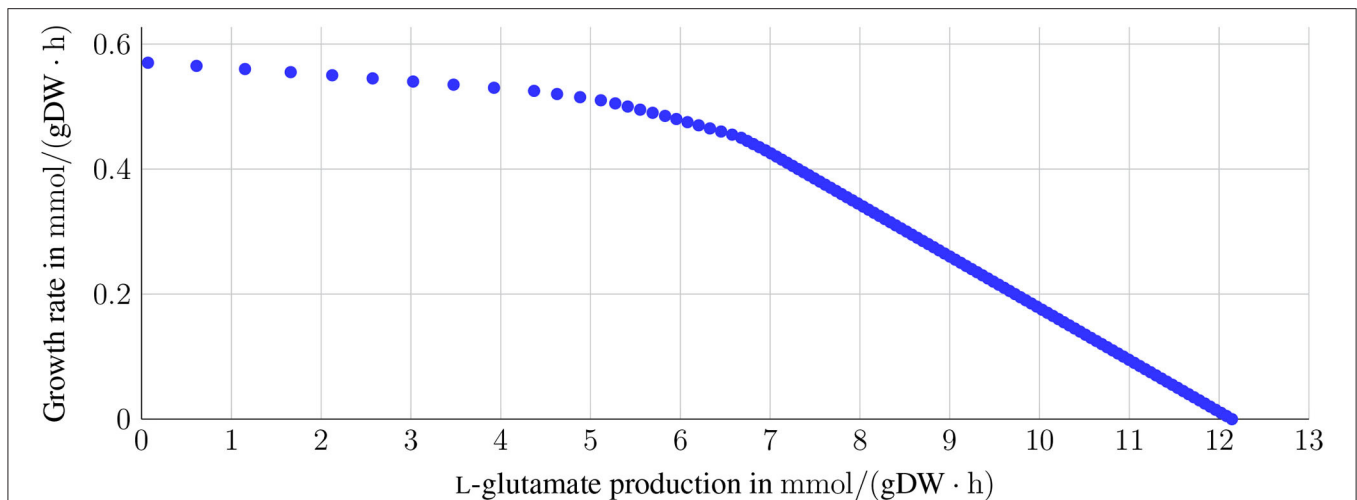


FIGURE 5 | Trade-off between L-glutamate production and growth. The sink reaction for L-glutamate was set as the objective function to investigate the relation between the production of L-glutamate and growth. The growth rate was varied between 0 and the maximum growth rate of $0.57 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ with glucose as the sole carbon source on CGXII. A dependency between growth and L-glutamate production is expected, as L-glutamate is part of the growth function with a stoichiometric coefficient of 0.0149. For growth rates between 0 and up to $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, the L-glutamate production only decreases slightly (from 12 to approximately $7 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). In contrast, the production of L-glutamate decreases drastically for higher growth rates.

3.3. Pointers to Metabolic Engineering for the L-glutamate Production

C. glutamicum is a well-known L-glutamate producer. However, L-glutamate is also required for the growth or maintenance function. L-glutamate accounts for the growth function with a stoichiometric coefficient of 0.0149. Thus, a trade-off between growth requirement and the production of L-glutamate is expected. This trade-off is depicted in **Figure 5**. For growth rates between $0 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ and $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, the L-glutamate production rate remains comparably high. It only decreases by $5 \text{ mmol gDW}^{-1} \text{ h}^{-1}$. With increasing growth rates greater than $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$, the L-glutamate production rate decreases rapidly.

The PC plays a pivotal role in L-glutamate production (Peters-Wendisch et al., 2001). The effect of a knock-out of the PC on the flux distribution is depicted in **Figure 6**. Knocking out the PC decreases the L-glutamate production only to a small extent (from $7.31 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to $7.26 \text{ mmol gDW}^{-1} \text{ h}^{-1}$). The limiting factor in L-glutamate production is the availability of a carbon source (in this example, D-glucose) and, as shown above, the growth rate. The PC *in silico* knock-out experiment indicates that *C. glutamicum* can compensate for the knocked-out reaction.

Performing FVA helps to identify potential the ranges of each flux. Reactions relevant for optimizing the objective function can be identified by filtering for reactions with almost identical minimal and maximal flux values. With loopless FVA, we identified six highly relevant reactions for L-glutamate production in *C. glutamicum*. Among these six reactions were two pseudo-reactions: the exchange reaction of D-glucose and the sink reaction for L-glutamate. Glucose is the sole carbon source in the *in silico* experiment. Therefore, its strong influence on L-glutamate production is apparent.

The same holds for the sink reaction that was used as the objective function in the FVA. The other four relevant reactions include the aconitate hydratase (ACONT), which converts citrate to isocitrate, the citrate synthase (CS), which converts acetyl-CoA and oxaloacetate to citrate and coenzyme A, the glucose transport via phosphoenolpyruvate, and the isocitrate dehydrogenase (ICDHyr), which converts isocitrate to 2-oxoglutarate (see also **Table 1**). The reactions ACONT, CS, and ICDHyr represent the fragile connection between glycolysis and L-glutamate biosynthesis. This connection can additionally be seen in **Figure 6**, where the three mentioned reactions are also illustrated.

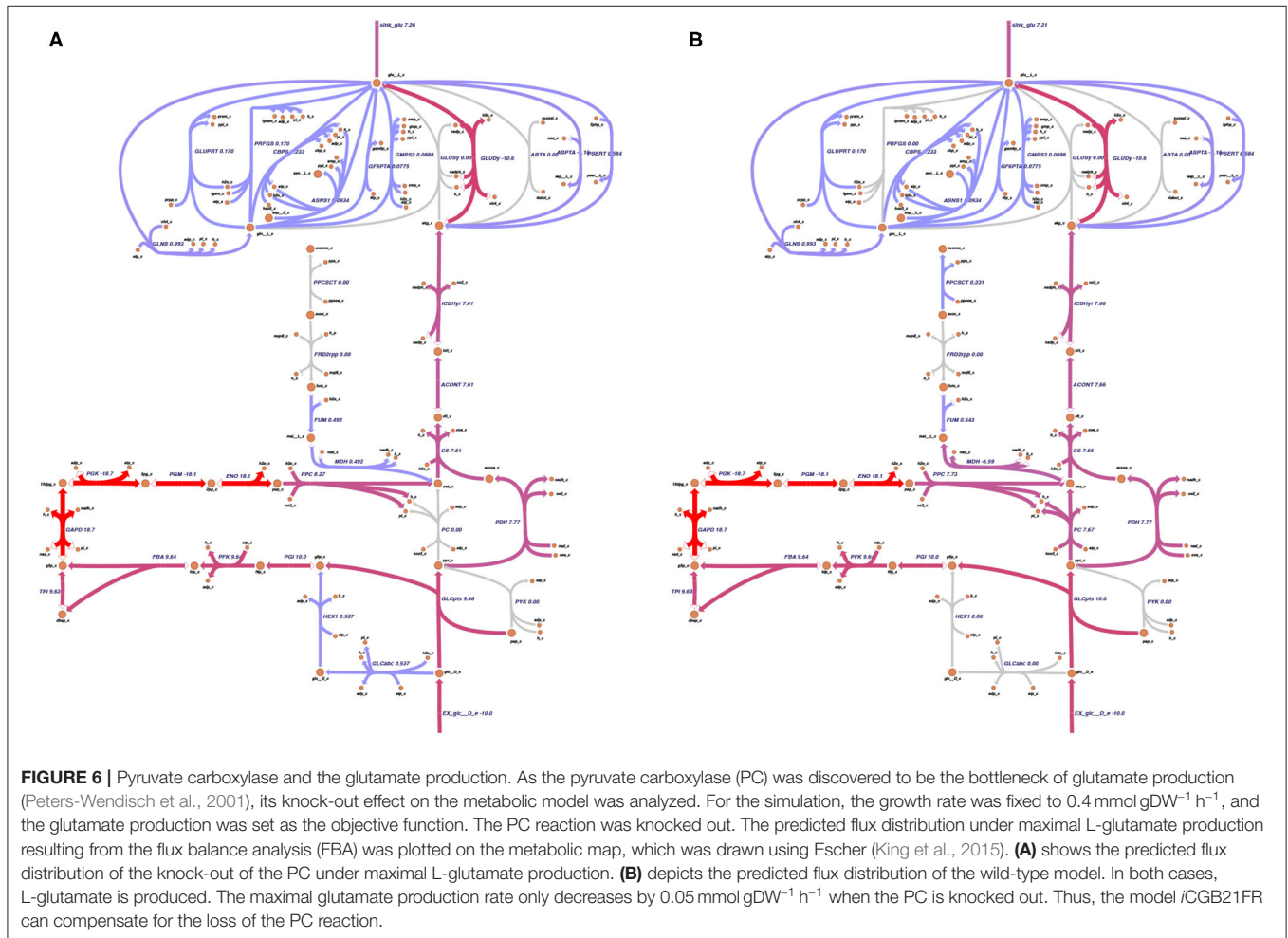
4. DISCUSSION

An updated genome-scale metabolic model *iCGB21FR* of *C. glutamicum* ATCC 13032 was reconstructed and validated using newly available specialized reconstruction tools. Using recent tools, the phenotypic prediction of the model's metabolism allows a more accurate depiction of the metabolic capabilities of *C. glutamicum*. This GEM was created using current community standards for high-quality reconstructions. The new *in silico* model reproduces experimentally validated data. In addition, we also curated *iCW773* to meet systems biology standards (see also Carey et al., 2020). Initially, the model *iCW773* was only available as a spreadsheet in Microsoft Excel format. According to the profound debate within the systems biology community (see Ebrahim et al., 2015), using this format is no longer recommended because it does not support unambiguous interpretation and direct reuse in further model analysis, especially for non-computational scientists. Generally, models in spreadsheet files do not fully support the principles of findable, accessible, interoperable, reusable data in science (Wilkinson

TABLE 1 | Reactions of particular relevance for the production of L-glutamate in *C. glutamicum*.

Reaction ID	Reaction name	Reaction
ACONT	aconitate hydratase	Citrate \rightleftharpoons Isocitrate
CS	citrate synthase	Acetyl-CoA + H ₂ O + Oxaloacetate \rightarrow Citrate + Coenzyme A + H ⁺
EX_glc_D_e	D-glucose exchange	D-Glucose \rightleftharpoons \emptyset
GLCpts	D-glucose transport via PEP:Pyr	D-Glucose + Phosphoenolpyruvate \rightarrow D-Glucose 6-phosphate + Pyruvate
ICDHyr	isocitrate dehydrogenase (NADP)	Isocitrate + NADP \rightleftharpoons 2-Oxoglutarate + CO ₂ + NADPH
sink_glu_L	sink reaction for L-glutamate	L-glutamate \rightarrow \emptyset

Reactions identified by flux variability analysis (FVA) to be highly relevant for L-glutamate production.



et al., 2016) because using them in computational analyses requires converting these files to a standardized format. After converting *i*CW773 to SBML Level 3 Version 1 and performing several curation steps, it now contains SBO terms and has a MEMOTE score that was increased from initially 29% to 70%. However, precaution is advised when using *i*CW773 as it contains reconstruction inconsistencies and incorrect metabolites. The curated *i*CW773 is available in SBML Level 3 Version 1 on the BioModels Database (Malik-Sheriff et al., 2020) under the accession number MODEL2110010001 (see Availability below).

4.1. Reconstruction Is of High Quality

The comprehensive annotations of all model components, including metabolites, reactions, and genes, contribute to the high quality of the reconstruction of *i*CGB21FR. Each instance is uniquely referenced to at least one database, thus providing a permanent link to clearly and uniquely identify this instance with its attributes (Juty et al., 2012). Almost all model instances are annotated with references to a minimum of one database, allowing more precise cross-referencing and interoperability between different databases. This high level of annotations is advantageous, as findable, accessible, interoperable, reusable

(FAIR) data principles allow fellow scientists to conduct research on and with this model continuously (Dräger and Palsson, 2014; Wilkinson et al., 2016). Erroneous or missing information, incompatible data formats, or missing annotations significantly hamper the reuse of GEMs (Ravikrishnan and Raman, 2015). Missing annotations can lead to identification problems of compounds and reactions. GPRs are added for different reactions, and all instances are equipped with SBO terms, facilitating FAIR data principles.

Generally, the high degree of annotations in *iCGB21FR* is confirmed by the high MEMOTE scores within the different categories. In terms of the presence of annotations, almost all MEMOTE scores of the annotation categories rank close to 100%. This high score implies that almost all suggested standards concerning annotations are met for this GEM (see also Carey et al., 2020). The current version of MEMOTE does not include all *C. glutamicum*-specific databases, while other organism-specific databases with less relevance for our model are incorporated. One example can be found in the annotations section of the genes, where cross-references to different *Escherichia coli* databases are checked. In *iCGB21FR*, the MEMOTE score for the presence of SBO terms for biochemical reactions sticks out due to its comparatively low value. The current version of MEMOTE checks every reaction for the annotation with the most general SBO term (SBO:0000176), “biochemical reaction.” This check implies that MEMOTE can not yet capture the fine-grained description of biochemical reactions in this model. Thus, the score of the metabolic reactions of 33.4% diminishes the overall MEMOTE SBO term annotation score.

Two typical ways exist to calculate the biomass objective function (BOF) of an organism. These are the macromolecular-based and the sequence-based approach. A typical biomass objective function (BOF) comprises the cell’s primary macromolecules, essential coenzymes, inorganic ions, and species-specific metabolites, including the cell wall components. Additionally, the energy requirements for growth and non-growth associated maintenance costs are included (Lachance et al., 2019). Using an experimentally derived biomass composition implies that its cellular composition depends on the experimental conditions under which it was obtained. For example, the availability of nutrients and the resulting growth rate influence the ratio between DNA, RNA, and proteins (Scott et al., 2010). It thus represents a biased approach to compute the biomass. When no species-specific experimental data for the sequence-based approach is available to calculate a species-specific biomass function, a universal bacterial biomass function is included. Adapted biomass composition of a highly developed and curated GEM of *E. coli* is often used (Orth et al., 2011; Xavier et al., 2017). The new model possesses a biomass function adapted specifically to *C. glutamicum*. This conceptual approach differs from previous works by Kjeldsen and Nielsen (2009) and Zelle et al. (2015). For the BOF of Kjeldsen and Nielsen (2009), a biomass equation and the corresponding energy consumption associated with each reaction were formulated for each macromolecule. No *C. glutamicum*-specific data for the energy requirement of the polymerization of macromolecules was available; thus, *E. coli* data was used instead. The BOF of

iEZ482 is based on the BOF of Kjeldsen and Nielsen (2009). Using species-specific data forms the basis for models with high predictive value. BOFdat enables the curation and refinement of a species-specific BOF by incorporating various -omics data into its calculation. In this study, genomics data were available and applied to refine the species-specific BOF.

4.2. *iCGB21FR* Reproduces Experimentally Obtained Data

The model *iCGB21FR* was validated by simulating growth on three different media under aerobic and anaerobic conditions. *C. glutamicum* is also known to grow in the brain heart infusion (BHI) medium. Modeling requires chemically defined media for growth simulations. We could not test the growth of *iCGB21FR* in BHI as no exact composition of the chemical definition of this medium exists. Simulating aerobic growth on LB complete medium was possible without any additional refinement of the model. Aerobic growth on the two minimal media was only possible by adapting the biomass function for growth on CGXII and adding missing reactions to the model for growth on M9. Missing reactions were identified by literature research. Anaerobic growth was enabled after adding six reactions. By this, the model was refined step by step to simulate and confirm already known growth conditions.

In silico growth rates were higher on the complete medium compared to the two minimal media. The aerobic growth rates were higher than the anaerobic growth rates, with all growth rates within a realistic range (Unthan et al., 2014; Michel et al., 2015). Both findings are expected, as complete media provide more nutrients and biomass precursors than minimal media. As their name suggest, minimal media only provide minimal required nutrients for the organism to grow. *C. glutamicum* uses oxygen and the more efficient aerobic respiration. It is even often regarded as aerobe (Takeno et al., 2007). However, as *C. glutamicum* is facultatively anaerobe, it can also switch to fermentation and anaerobic respiration if oxygen is absent. Anaerobic growth by nitrate respiration is limited, as nitrate accumulates and inhibits growth. Additionally, glucose is converted to L-lactate and succinate without the growth of the organism (Inui et al., 2004b; Koch-Koerfges et al., 2013).

Two observations stick out from these growth results. First, the current tools for the reconstruction of GEMs still demand subsequent manual refinement. Even though automated tools, such as CarveMe (Machado et al., 2018), reduce the amount of time spent on the reconstruction dramatically, manual refinement remains a pivotal part of the reconstruction process. The necessity of manual curation becomes particularly apparent when comparing the growth predictions of our model for the different media. The draft model created by CarveMe enabled growth on the complete medium without further ado. However, manual refinement was essential for the simulation of growth on the two minimal media. The second interesting observation is the organism-specific gap-filling, which appears to be more fruitful when applied to media that specifically compensate for certain physiological or metabolic oddities of the organism. In our case, knowledge gap-filling was most fruitful

on the CGXII medium which, for example, compensates for the limited ability of *C. glutamicum* to synthesize and excrete siderophores (Budzikiewicz et al., 1997). This makes sense, as the minimal medium provides the microbe's bare necessities to grow. Potentially lacking compounds could be compensated by the composition of the complete medium.

We validated our model by testing the growth rate on the CGXII minimal medium under aerobic and anaerobic conditions using different experimentally validated carbon sources. Aerobic growth was possible on all experimentally validated carbon sources. In addition to the experimentally confirmed anaerobic growth on glucose, fructose, sucrose as carbon sources, our *in silico* model also grew on gluconate. We verified that all genes required to utilize gluconate as carbon sources exist in *C. glutamicum*. The rate of NADPH reoxidation could be a potential explanation for the *in silico* growth on the additional anaerobic carbon source. The enzyme 6-phosphogluconate dehydrogenase oxidizes 6-phosphogluconate to ribulose 5-phosphate. This enzyme is inhibited by NADPH, which is essential for the cellular control of the NADPH synthesis (Moritz et al., 2000). The rate of NADPH re-oxidation represents a critical element of this process. Gluconate is phosphorylated after uptake and then catabolized in the pentose phosphate pathway. If NADPH re-oxidation was too low under anaerobic conditions, NADPH could accumulate and result in complete inhibition of 6-phosphogluconate dehydrogenase activity. This accumulation would lead to a stop in growth on gluconate, as was shown by experimental data (Michel et al., 2015). If, however, NADPH re-oxidation is sufficiently fast and no NADPH accumulates, the activity of the 6-phosphogluconate dehydrogenase could remain active and allow anaerobic growth on gluconate by simulation studies with this *in silico* model.

As a final validation step of the metabolic model, its ability to produce amino acids was examined. In complex bacterial systems, amino acid production co-occurs with growth (Marx et al., 1996). The biosynthesis of amino acids requires a lot of the carbon source's total budget, usually used for bacterial growth (Neidhardt et al., 1990). The growth rate was fixed to $0.4 \text{ mmol gDW}^{-1} \text{ h}^{-1}$ to ensure the microbe's viability while producing amino acids. With increasing CO_2 production rate and ATP requirements, the amino acid production yield decreases. Especially smaller amino acids with only a few carbon atoms, like L-alanine with only three carbon atoms, or glycine with two carbon atoms, have a low CO_2 production and ATP requirement rate.

In contrast, amino acids with more carbon atoms, such as L-tryptophan, have a much higher ATP requirement and CO_2 production rate, while the amino acid yield is relatively low. This leads to the conclusion that building more extensive and more complex amino acids needs more energy. An increasing proportion of the consumed sugar will be used to produce more energy, is then lost as CO_2 , and cannot be used for amino acid formation (Gourdon et al., 2003).

4.3. L-glutamate Production: New Insights for Metabolic Engineering

One might expect a linear correlation between the L-glutamate production and the growth rate, where the slope is related to the amino acid's stoichiometric coefficient in the biomass function. The trade-off between the production of L-glutamate and the growth rate was investigated by fixing the growth rates and maximizing the L-glutamate production (see Figure 5). The system moves between these two boundaries: the maximal possible growth rate and the maximal possible production of L-glutamate. The closer the values get to either maximum, the greater is the influence on the respective other value. When ceasing growth, the theoretical production of L-glutamate would reach a maximum since the available metabolic capacity is invested in the L-glutamate production. The reverse situation occurs with ceasing L-glutamate production and maximizing growth where all energetic demand is invested. L-glutamate production reaches its maximum when no growth occurs, and the available glucose is completely used to produce L-glutamate. Thus, the growth rate is the limiting factor for our *in silico* model, independent of the L-glutamate production.

The PC was first investigated to study relevant reactions for L-glutamate production (see Figure 6). The PC has been described as the bottleneck in the production of L-glutamate (Peters-Wendisch et al., 2001). Knocking out the PC in laboratory experiments leads to ambivalent results: Both drastic decrease (Peters-Wendisch et al., 2001) and increase (Sato et al., 2008) in L-glutamate production were reported after a disruption of the PC. Pyruvate is part of the complex network responsible for carboxylation and decarboxylation reactions, which connect the glycolysis and TCA cycle (Becker and Wittmann, 2020). In *C. glutamicum*, the PC represents one of the carboxylation enzymes, the other being the phosphoenolpyruvate carboxylase (Eikmanns, 2005). The carboxylation and decarboxylation enzymatic complex in *C. glutamicum* is a highly flexible network that enables several pathways to respond to varying metabolic circumstances (Möllney et al., 2000; Becker et al., 2008; Becker and Wittmann, 2020)—knocking out the PC in our model still allowed L-glutamate production. We also found that the amount of produced L-glutamate does not vary significantly with the PC being knocked. According to our simulation, the limiting factors in L-glutamate production are access to carbon sources and the growth rate.

Four reactions were identified that play a pivotal role in the production of L-glutamate in our model. The first is a glucose transporter, which uses the phosphoenolpyruvate (PEP)-dependent sugar phosphotransferase system. The L-glutamate yield decreased with sugar consumption rates in laboratory experiments (Gourdon et al., 2003). This seems reasonable, as glucose is the sole carbon source and starting point for glutamate production. Increasing glucose availability is only expedient if the glucose transporters' capacity is given to take up the enhanced supply of glucose. The three remaining reactions, aconitase, citrate synthase, and isocitrate dehydrogenase, are all part of the tricarboxylic acid (TCA) cycle. The TCA cycle is a complex

regulated amphibolic pathway with L-glutamate and L-lysine as derived intermediate products (Bott, 2007).

The aconitase gene is regulated by four transcriptional regulators, indicating a tight control of this enzyme. A *C. glutamicum* mutant lacking the aconitase gene was glutamate auxotrophic in the CGXII minimal medium with glucose as the carbon source (Baumgart et al., 2011). The model iCGB21FR could replicate the finding that aconitase is essential for L-glutamate production. It remains to be further experimentally validated how the interplay between the aconitase within the TCA cycle in terms of L-glutamate production can be optimized.

The citrate synthase catalyzes the initial reaction of the TCA cycle. Overexpression of the citrate synthase can redirect more carbon flux into the cycle and result in higher L-arginine production (Man et al., 2016). L-arginine is synthesized from the precursor L-glutamic acid (Utagawa, 2004). Thus, higher production of L-glutamate might also be dependent upon the activity of the citrate synthase. The role of the citrate synthase in L-glutamate production might be an interesting topic to investigate since it might represent a target for metabolic engineering of *C. glutamicum*'s TCA cycle. Since the citrate synthase is the initial reaction of the TCA cycle with L-glutamate and L-lysine as intermediates, its activity might prove particularly fruitful.

Isocitrate dehydrogenase catalyzes the oxidative decarboxylation of isocitrate. Becker et al. (2009) found in their investigation of the effects of the isocitrate dehydrogenase on L-lysine production that decreased activity of the isocitrate dehydrogenase improves the L-lysine production. This decrease induced a flux shift from the TCA cycle to anaerobic carboxylation (van Ooyen et al., 2012). However, the PC functions as an anaerobic enzyme in L-glutamate production (Peters-Wendisch et al., 1997). In other words, the isocitrate dehydrogenase has different functions in L-glutamate production than in L-lysine production. This differing function becomes more apparent when looking at the effects of isocitrate dehydrogenase inactivation in *C. glutamicum*: Inactivation of the NADP-dependent isocitrate dehydrogenase in *C. glutamicum* leads to L-glutamate auxotrophy (Eikmanns et al., 1995). This connection between the PC and the isocitrate dehydrogenase in L-glutamate production might be an interesting target for metabolic engineering.

5. CONCLUSION AND OUTLOOK

The new model iCGB21FR represents an GEM of high quality of the biotechnologically relevant microorganism *Corynebacterium glutamicum* ATCC 13032. We reconstructed this metabolic

model with an adapted, species-specific biomass composition and realistic growth rates in different environments, which were validated using experimentally derived data. Furthermore, alternative metabolic pathways for the production of L-glutamate were shown in our *in silico* model. Particularly, these alternative pathways could be of interest for further investigation in terms of metabolic engineering. Biotin is a key player for the L-glutamate production in *C. glutamicum* since its limitation triggers L-glutamate production. Despite the inclusion of biotin in iCGB21FR and its participation in five biochemical reactions, its role in L-glutamate production is currently not included. The influence of biotin on the PC and the reactions involved in the alternative pathway for L-glutamate production with the pyruvate carboxylase knocked-out should be further investigated in subsequent GEMs of *C. glutamicum*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MF and AR curated and refined the model and conducted the study. MF, AR, and AD wrote the manuscript. EZ, KN, and WW revised the manuscript. AD supervised the study. All authors reviewed and approved the final manuscript.

FUNDING

The authors acknowledge support by the Open Access Publishing Fund of the University of Tübingen (<https://uni-tuebingen.de/en/58988>).

ACKNOWLEDGMENTS

We thank Elisabeth Fritze for providing access to the program she designed as part of her bachelor's requirements. Her algorithm allowed the assignment of hierarchically differentiated SBO terms to our model.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.750206/full#supplementary-material>

REFERENCES

- Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., et al. (2018). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 46, D8–D13. doi: 10.1093/nar/gkx1095
- Baumgart, M., Mustafi, N., Krug, A., and Bott, M. (2011). Deletion of the aconitase gene in *Corynebacterium glutamicum* causes strong selection pressure for secondary mutations inactivating citrate synthase. *J. Bacteriol.* 193, 6864–6873. doi: 10.1128/JB.05465-11
- Becker, J., Klopprogge, C., Schröder, H., and Wittmann, C. (2009). Metabolic engineering of the tricarboxylic acid cycle for improved lysine production by *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* 75, 7866–7869. doi: 10.1128/AEM.01942-09

- Becker, J., Klopprogge, C., and Wittmann, C. (2008). Metabolic responses to pyruvate kinase deletion in lysine producing *Corynebacterium glutamicum*. *Microb. Cell. Fact.* 7, 1–15. doi: 10.1186/1475-2859-7-8
- Becker, J., and Wittmann, C. (2020). “Pathways at work: metabolic flux analysis of the industrial cell factory *Corynebacterium glutamicum*,” in *Corynebacterium glutamicum* (Berlin; Heidelberg: Springer), 227–265.
- Bergmann, F. T., Adams, R., Moodie, S., Cooper, J., Glont, M., Golebiewski, M., et al. (2014). COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics* 15:369. doi: 10.1186/s12859-014-0369-z
- Bergmann, F. T., Czauderna, T., Dorusoz, U., Rougny, A., Dräger, A., Touré, V., et al. (2020). Systems biology graphical notation markup language (SBGNML) version 0.3. *J. Integr. Bioinform.* 17:20200016. doi: 10.1515/jib-2020-0016
- Bertani, G. (1951). Studies on lysogeny. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 62, 293–300. doi: 10.1128/jb.62.3.293-300.1951
- Bornstein, B. J., Keating, S. M., Jouraku, A., and Hucka, M. (2008). LibSBML: an API library for SBML. *Bioinformatics* 24, 880–881. doi: 10.1093/bioinformatics/btn051
- Bott, M. (2007). Offering surprises: TCA cycle regulation in *Corynebacterium glutamicum*. *Trends Microbiol.* 15, 417–425. doi: 10.1016/j.tim.2007.08.004
- Bott, M., and Niebisch, A. (2003). The respiratory chain of corynebacterium glutamicum. *J. Biotechnol.* 104, 129–153. doi: 10.1016/S0168-1656(03)00144-5
- Brandenburg, F. J., Forster, M., Pick, A., Raitner, M., and Schreiber, F. (2004). “BioPath” Exploration and visualization of biochemical pathways,” in *Graph Drawing Software. Mathematics and Visualization*, ed M. M. P. Jünger (Berlin; Heidelberg: Springer), 215–235.
- Budzikiewicz, H., Bösenkamp, A., Taraz, K., Pandey, A., and Meyer, J.-M. (1997). Corynebactin, a Cyclic Catechol Siderophore from: *Corynebacterium glutamicum* ATCC 14067 (*Brevibacterium* sp. DSM 20411). *Zeitschrift für Naturforschung C* 52, 551–554. doi: 10.1515/znc-1997-7-820
- Carey, M. A., Dräger, A., Beber, M. E., Papin, J. A., and Yurkovich, J. T. (2020). Community standards to facilitate development and address challenges in metabolic modeling. *Mol. Syst. Biol.* 16, e9235. doi: 10.15252/msb.20199235
- Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., et al. (2020). The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res.* 48, D445–D453. doi: 10.1093/nar/gkz862
- Collins, M., Goodfellow, M., and Minnikin, D. (1979). Isoprenoid quinones in the classification of coryneform and related bacteria. *Microbiology* 110, 127–136. doi: 10.1099/00221287-110-1-127
- Collins, M., Pirouz, T., Goodfellow, M., and Minnikin, D. (1977). Distribution of menaquinones in actinomycetes and corynebacteria. *Microbiology* 100, 221–230. doi: 10.1099/00221287-100-2-221
- Courtot, M., Juty, N., Knüpfer, C., Waltemath, D., Zhukova, A., Dräger, A., et al. (2011). Controlled vocabularies and semantics in systems biology. *Mol. Syst. Biol.* 7, 543. doi: 10.1038/msb.2011.77
- Croft, D., O’Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., et al. (2010). Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39(Suppl_1):D691–97. doi: 10.1093/nar/gkq1018
- Dräger, A., Kronfeld, M., Ziller, M. J., Supper, J., Planatscher, H., Magnus, J. B., et al. (2009). Modeling metabolic networks in *C. glutamicum*: a comparison of rate laws in combination with various parameter optimization strategies. *BMC Syst. Biol.* 3:5. doi: 10.1186/1752-0509-3-5
- Dräger, A., and Palsson, B. Ø. (2014). Improving collaboration by standardization efforts in systems biology. *Front. Bioeng.* 2:61. doi: 10.3389/fbioe.2014.00061
- Ebrahim, A., Almaas, E., Bauer, E., Bordbar, A., Burgard, A. P., Chang, R. L., et al. (2015). Do genome-scale models need exact solvers or clearer standards? *Mol. Syst. Biol.* 11, 831. doi: 10.15252/msb.20156548
- Ebrahim, A., Lerman, J. A., Palsson, B. O., and Hyduke, D. R. (2013). COBRApy: constraints-based reconstruction and analysis for python. *BMC Syst. Biol.* 7:74. doi: 10.1186/1752-0509-7-74
- Eggeling, L., and Bott, M. (2005). *Handbook of Corynebacterium glutamicum*. Boca Raton, FL: CRC Press.
- Eggeling, L., and Bott, M. (2015). A giant market and a powerful metabolism: L-lysine provided by *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 99, 3387–3394. doi: 10.1007/s00253-015-6508-2
- Eikmanns, B. (2005). “Central metabolism: tricarboxylic acid cycle and anaplerotic reactions,” in *Handbook of Corynebacterium glutamicum*, eds L. Eggeling and M. Bott (Boca Raton, FL: CRC Press; Taylor & Francis Group), 241–276.
- Eikmanns, B. J., Rittmann, D., and Sahn, H. (1995). Cloning, sequence analysis, expression, and inactivation of the *Corynebacterium glutamicum* icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme. *J. Bacteriol.* 177, 774–782. doi: 10.1128/jb.177.3.774-782.1995
- Ellis, L. B., Hou, B. K., Kang, W., and Wackett, L. P. (2003). The University of Minnesota biocatalysis/biodegradation database: post-genomic data mining. *Nucleic Acids Res.* 31, 262–265. doi: 10.1093/nar/gkg048
- Fang, X., Lloyd, C. J., and Palsson, B. O. (2020). Reconstructing organisms *in silico*: genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* 8, 731–743. doi: 10.1038/s41579-020-00440-4
- Fritze, E. (2020). *Automating the Assignment of SBO-Terms* (Bachelor thesis). University of Tübingen.
- Fritzemeier, C. J., Hartleb, D., Szappanos, B., Papp, B., and Lercher, M. J. (2017). Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal. *PLoS Comput. Biol.* 13:e1005494. doi: 10.1371/journal.pcbi.1005494
- Gottstein, W., Olivier, B. G., Bruggeman, F. J., and Teusink, B. (2016). Constraint-based stoichiometric modelling from single organisms to microbial communities. *J. R. Soc. Interface* 13, 20160627. doi: 10.1098/rsif.2016.0627
- Gourdon, P., Raherimandimby, M., Dominguez, H., Coccain-Bousquet, M., and Lindley, N. D. (2003). Osmotic stress, glucose transport capacity and consequences for glutamate overproduction in *Corynebacterium glutamicum*. *J. Biotechnol.* 104, 77–85. doi: 10.1016/S0168-1656(03)00165-2
- Hastings, J., Owen, G., Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., et al. (2016). ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res.* 44, D1214–D1219. doi: 10.1093/nar/gkv1031
- Heider, S. A., Wolf, N., Hofemeier, A., Peters-Wendisch, P., and Wendisch, V. F. (2014). Optimization of the IPP precursor supply for the production of lycopene, decaprenoxanthin and astaxanthin by *Corynebacterium glutamicum*. *Front. Bioeng. Biotechnol.* 2:28. doi: 10.3389/fbioe.2014.00028
- Heller, S. R., McNaught, A., Pletnev, I., Stein, S., and Tchekhovskoi, D. (2015). InChI, the IUPAC international chemical identifier. *J. Cheminform.* 7, 23. doi: 10.1186/s13321-015-0068-4
- Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., and Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi: 10.1038/nbt.1672
- Hucka, M., Bergmann, F. T., Chaouiya, C., Dräger, A., Hoops, S., Keating, S. M., et al. (2019). Systems biology markup language (SBML) level 3 version 2 core release 2. *J. Integr. Bioinform.* 16, 1. doi: 10.1515/jib-2019-0021
- Hucka, M., Bergmann, F. T., Dräger, A., Hoops, S., Keating, S. M., Le Novère, N., et al. (2018). systems biology markup language (SBML) level 3 version 1 core. *J. Integr. Bioinform.* 15, 1. doi: 10.1515/jib-2017-0080
- Hucka, M., and Smith, L. P. (2016). SBML Level 3 package: groups, version 1 release 1. *J. Integr. Bioinform.* 13, 1. doi: 10.1515/jib-2016-290
- Hüser, A. T., Chassagnole, C., Lindley, N. D., Merkmann, M., Guyonvarch, A., Elišáková, V., et al. (2005). Rational design of a *Corynebacterium glutamicum* pantothenate production strain and its characterization by metabolic flux analysis and genome-wide transcriptional profiling. *Appl. Environ. Microbiol.* 71, 3255–3268. doi: 10.1128/AEM.71.6.3255-3268.2005
- Inui, M., Kawaguchi, H., Murakami, S., Vertès, A. A., and Yukawa, H. (2004a). Metabolic engineering of *Corynebacterium glutamicum* for fuel ethanol production under oxygen-deprivation conditions. *J. Mol. Microbiol. Biotechnol.* 8, 243–254. doi: 10.1159/000086705
- Inui, M., Murakami, S., Okino, S., Kawaguchi, H., Vertès, A. A., and Yukawa, H. (2004b). Metabolic analysis of *Corynebacterium glutamicum* during lactate and succinate productions under oxygen deprivation conditions. *J. Mol. Microbiol. Biotechnol.* 7, 182–196. doi: 10.1159/000079827
- Jadebeck, J. F., Theorell, A., Leweke, S., and Nöh, K. (2020). Hops: high-performance library for (non-) uniform sampling of convex-constrained models. *Bioinformatics* 37, 1776–1777. doi: 10.1093/bioinformatics/btaa872
- Jojima, T., Noburyu, R., Sasaki, M., Tajima, T., Suda, M., Yukawa, H., et al. (2015). Metabolic engineering for improved production of ethanol by *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 99, 1165–1172. doi: 10.1007/s00253-014-6223-4

- Juty, N., Le Novère, N., and Laibe, C. (2012). Identifiers.org and MIRIAM Registry: community resources to provide persistent identification. *Nucleic Acids Res.* 40, D580–D586. doi: 10.1093/nar/gkr1097
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., et al. (2003). The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* 104, 5–25. doi: 10.1016/S0168-1656(03)00154-8
- Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* 47, D590–D595. doi: 10.1093/nar/gky962
- Kang, M.-K., Eom, J.-H., Kim, Y., Um, Y., and Woo, H. M. (2014). Biosynthesis of pinene from glucose using metabolically-engineered *Corynebacterium glutamicum*. *Biotechnol. Lett.* 36, 2069–2077. doi: 10.1007/s10529-014-1578-2
- Kanzaki, T., Sugiyama, Y., Kitano, K., Ashida, Y., and Imada, I. (1974). Quinones of *brevibacterium*. *Biochim. Biophys. Acta* 348, 162–165. doi: 10.1016/0005-2760(74)90102-7
- Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinform.* 20, 1085–1093. doi: 10.1093/bib/bbx085
- Keating, S. M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., et al. (2020). SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol. Syst. Biol.* 16, e9110. doi: 10.15252/msb.20199110
- Keilhauer, C., Eggeling, L., and Sahm, H. (1993). Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.* 175, 5595–5603. doi: 10.1128/jb.175.17.5595-5603.1993
- Kimura, E. (2005). “19 L-glutamate production,” in *Handbook of Corynebacterium glutamicum* (Boca Raton, FL), 439.
- Kind, S., Jeong, W. K., Schröder, H., and Wittmann, C. (2010a). Systems-wide metabolic pathway engineering in *Corynebacterium glutamicum* for bio-based production of diaminopentane. *Metab. Eng.* 12, 341–351. doi: 10.1016/j.ymben.2010.03.005
- Kind, S., Jeong, W. K., Schröder, H., Zelder, O., and Wittmann, C. (2010b). Identification and elimination of the competing N-acetyldiaminopentane pathway for improved production of diaminopentane by *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* 76, 5175–5180. doi: 10.1128/AEM.00834-10
- King, Z. A., Dräger, A., Ebrahim, A., Sonnenschein, N., Lewis, N. E., and Palsson, B. O. (2015). Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways. *PLoS Comput. Biol.* 11:e1004321. doi: 10.1371/journal.pcbi.1004321
- Kjeldsen, K. R., and Nielsen, J. (2009). *In silico* genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network. *Biotechnol. Bioeng.* 102, 583–597. doi: 10.1002/bit.22067
- Koch-Koerfges, A., Pfelzer, N., Platzen, L., Oldiges, M., and Bott, M. (2013). Conversion of *Corynebacterium glutamicum* from an aerobic respiring to an aerobic fermenting bacterium by inactivation of the respiratory chain. *Biochim. Biophys. Acta* 1827, 699–708. doi: 10.1016/j.bbabi.2013.02.004
- König, M. (2020). *matthiaskoenig/fbc_curation*: 0.1.6. doi: 10.5281/zenodo.3708271
- Lachance, J.-C., Lloyd, C. J., Monk, J. M., Yang, L., Sastry, A. V., Seif, Y., et al. (2019). BOFdat: Generating biomass objective functions for genome-scale metabolic models from experimental data. *PLoS Comput. Biol.* 15:e1006971. doi: 10.1371/journal.pcbi.1006971
- Le Novè, N., Finney, A., Hucka, M., Bhalla, U. S., Campagne, F., Collado-Vides, J., et al. (2005). Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat. Biotechnol.* 23, 1509–1515. doi: 10.1038/nbt1156
- Liebl, W. (2005). “*Corynebacterium* taxonomy,” in *Handbook of Corynebacterium glutamicum* (Boca Raton, FL: CRC Press), 9–34.
- Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., et al. (2020). MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* 38, 272–276. doi: 10.1038/s41587-020-0446-y
- Liu, Q., Ouyang, S.-P., Kim, J., and Chen, G.-Q. (2007). The impact of PHB accumulation on L-glutamate production by recombinant *Corynebacterium glutamicum*. *J. Biotechnol.* 132, 273–279. doi: 10.1016/j.jbiotec.2007.03.014
- Lombardot, T., Morgat, A., Axelsen, K. B., Aimo, L., Hyka-Nouspikel, N., Niknejad, A., et al. (2019). Updates in Rhea: SPARQLing biochemical reaction data. *Nucleic Acids Res.* 47, D596–D600. doi: 10.1093/nar/gky876
- Machado, D., Andrejev, S., Tramontano, M., and Patil, K. R. (2018). Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.* 46, 7542–7553. doi: 10.1093/nar/gky537
- Maeda, T., Koch-Koerfges, A., and Bott, M. (2020). Relevance of nadh dehydrogenase and alternative two-enzyme systems for growth of *corynebacterium glutamicum* with glucose, lactate, and acetate. *Front. Bioeng. Biotechnol.* 8:621213. doi: 10.3389/fbioe.2020.621213
- Maglott, D., Ostell, J., Pruitt, K. D., and Tatusova, T. (2005). Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.* 33(Suppl_1):D54–D58. doi: 10.1093/nar/gki031
- Malik-Sheriff, R. S., Glont, M., Nguyen, T. V. N., Tiwari, K., Roberts, M. G., Xavier, A., et al. (2020). BioModels–15 years of sharing computational models in life science. *Nucleic Acids Res.* 48, D407–D415. doi: 10.1093/nar/gkz1055
- Man, Z., Xu, M., Rao, Z., Guo, J., Yang, T., Zhang, X., et al. (2016). Systems pathway engineering of *Corynebacterium crenatum* for improved L-arginine production. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep28629
- Marx, A., de Graaf, A. A., Wiechert, W., Eggeling, L., and Sahm, H. (1996). Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111–129. doi: 10.1002/(SICI)1097-0290(19960120)49:2<andgt;111::AID-BIT1andgt;3.0.CO;2-T
- Michel, A., Koch-Koerfges, A., Krumbach, K., Brocker, M., and Bott, M. (2015). Anaerobic growth of *Corynebacterium glutamicum* via mixed-acid fermentation. *Appl. Environ. Microbiol.* 81, 7496–7508. doi: 10.1128/AEM.02413-15
- Dräger, A. (2021). *Table2model*: 1.0.1.
- Moretti, S., Tran, V. D. T., Mehl, F., Ibberson, M., and Pagni, M. (2021). MetaNetX/MNXref: unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.* 49, D570–D574. doi: 10.1093/nar/gkaa992
- Morgat, A., Coissac, E., Coudert, E., Axelsen, K. B., Keller, G., Bairoch, A., et al. (2011). UniPathway: a resource for the exploration and annotation of metabolic pathways. *Nucleic Acids Res.* 40, D761–D769. doi: 10.1093/nar/gkr1023
- Moritz, B., Striegel, K., de Graaf, A. A., and Sahm, H. (2000). Kinetic properties of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases from *Corynebacterium glutamicum* and their application for predicting pentose phosphate pathway flux *in vivo*. *Eur. J. Biochem.* 267, 3442–3452. doi: 10.1046/j.1432-1327.2000.01354.x
- Neal, M. L., König, M., Nickerson, D., Mısırlı, G., Kalbasi, R., Dräger, A., et al. (2018). Harmonizing semantic annotations for computational models in biology. *Brief. Bioinform.* 20, 540–550. doi: 10.1101/246470
- Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990). *Physiology of the Bacterial Cell: A Molecular Approach, Vol. 20*. Sunderland, MA: Sinauer Associates.
- Niimi, S., Suzuki, N., Inui, M., and Yukawa, H. (2011). Metabolic engineering of 1,2-propanediol pathways in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 90, 1721–1729. doi: 10.1007/s00253-011-3190-x
- Norsigian, C. J., Pusarla, N., McConn, J. L., Yurkovich, J. T., Dräger, A., Palsson, B. O., et al. (2019). BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Res.* 48: gkz1054. doi: 10.1093/nar/gkz1054
- Okino, S., Suda, M., Fujikura, K., Inui, M., and Yukawa, H. (2008). Production of L-lactic acid by *Corynebacterium glutamicum* under oxygen deprivation. *Appl. Microbiol. Biotechnol.* 78, 449–454. doi: 10.1007/s00253-007-1336-7
- Olivier, B. G., and Bergmann, F. T. (2018). SBML level 3 package: flux balance constraints version 2. *J. Integr. Bioinform.* 15:20170082. doi: 10.1515/jib-2017-0082
- Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., et al. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism” 2011. *Mol. Syst. Biol.* 7, 535. doi: 10.1038/msb.2011.65
- Panchiwala, H., Shah, S., Planatscher, H., Zakharchuk, M., König, M., and Dräger, A. (2021). The systems biology simulation core library. *Bioinformatics* doi: 10.1093/bioinformatics/btab669. [Epub ahead of print].
- Petersen S, de Graaf AA, Eggeling L, Möllney M, Wiechert W, Sahm H. *In vivo* quantification of parallel and bidirectional fluxes in the

- anaplerosis of *Corynebacterium glutamicum*. *J Biol Chem.* (2000) 275:35932–41. doi: 10.1074/jbc.M908728199
- Peters-Wendisch, P. G., Schiel, B., Wendisch, V. F., Katsoulidis, E., Möckel, B., Sahm, H., et al. (2001). Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by *Corynebacterium glutamicum*. *J. Mol. Microbiol. Biotechnol.* 3, 295–300.
- Peters-Wendisch, P. G., Wendisch, V. F., Paul, S., Eikmanns, B. J., and Sahm, H. (1997). Pyruvate carboxylase as an anaplerotic enzyme in *Corynebacterium glutamicum*. *Microbiology* 143, 1095–1103. doi: 10.1099/00221287-143-4-1095
- Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 33(Suppl_1):D501–D504. doi: 10.1093/nar/gki025
- Ravikrishnan, A., and Raman, K. (2015). Critical assessment of genome-scale metabolic networks: the need for a unified standard. *Brief. Bioinform.* 16, 1057–1068. doi: 10.1093/bib/bbv003
- Renz, A., Mostolizadeh, R., and Dräger, A. (2020). “Clinical applications of metabolic models in SBML format,” in *Systems Medicine, Vol. 3*, ed O. Wolkenhauer (Oxford: Academic Press), 362–371.
- Rodriguez, N., Thomas, A., Watanabe, L., Vazirabad, I. Y., Kofia, V., Gómez, H. F., et al. (2015). JSBML 1.0: providing a smorgasbord of options to encode systems biology models. *Bioinformatics* 31, 3383–3386. doi: 10.1093/bioinformatics/btv341
- Römer, M., Eichner, J., Dräger, A., Wrzodek, C., Wrzodek, F., and Zell, A. (2016). ZBIT bioinformatics toolbox: a web-platform for systems biology and expression data analysis. *PLoS ONE* 11:e0149263. doi: 10.1371/journal.pone.0149263
- Rougy, A., Touré, V., Moodie, S., Balaur, I., Czauderna, T., Borlinghaus, H., et al. (2019). Systems biology graphical notation: process description language level 1 version 2.0. *J. Integr. Bioinform.* 16:20190022. doi: 10.1515/jib-2019-0022
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, 2nd Edn.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sato, H., Orishimo, K., Shirai, T., Hirasawa, T., Nagahisa, K., Shimizu, H., et al. (2008). Distinct roles of two anaplerotic pathways in glutamate production induced by biotin limitation in *Corynebacterium glutamicum*. *J. Biosci. Bioeng.* 106, 51–58. doi: 10.1263/jbb.106.51
- Schellenberger, J., Lewis, N. E., and Palsson, B. Ø. (2011). Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophys. J.* 100, 544–553. doi: 10.1016/j.bpj.2010.12.3707
- Schneider, J., and Wendisch, V. F. (2010). Putrescine production by engineered *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 88, 859–868. doi: 10.1007/s00253-010-2778-x
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. *Science* 330, 1099–1102. doi: 10.1126/science.1192588
- Shinfuku, Y., Sorpitiporn, N., Sono, M., Furusawa, C., Hirasawa, T., and Shimizu, H. (2009). Development and experimental verification of a genome-scale metabolic model for *Corynebacterium glutamicum*. *Microb. Cell Fact.* 8, 1–15. doi: 10.1186/1475-2859-8-43
- Sud, M., Fahy, E., Cotter, D., Brown, A., Dennis, E. A., Glass, C. K., et al. (2007). Lmsd: Lipid maps structure database. *Nucleic Acids Res.* 35(Suppl_1):D527–D532. doi: 10.1093/nar/gkl838
- Takeo, S., Ohnishi, J., Komatsu, T., Masaki, T., Sen, K., and Ikeda, M. (2007). Anaerobic growth and potential for amino acid production by nitrate respiration in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 75, 1173–1182. doi: 10.1007/s00253-007-0926-8
- Takeo, S., Takasaki, M., Urabayashi, A., Mimura, A., Muramatsu, T., Mitsuhashi, S., et al. (2013). Development of fatty acid-producing *Corynebacterium glutamicum* strains. *Appl. Environ. Microbiol.* 79, 6776–6783. doi: 10.1128/AEM.02003-13
- Thiele, I., and Palsson, B. Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93. doi: 10.1038/nprot.2009.203
- Touré, V., Dräger, A., Luna, A., Dogrusoz, U., and Rougny, A. (2020). The “Systems biology graphical notation: current status and applications in systems medicine,” in *Systems Medicine, Vol. 3*, ed O. Wolkenhauer (Oxford: Academic Press), 372–381.
- Unthan, S., Grünberger, A., van Ooyen, J., Gätgens, J., Heinrich, J., Paczia, N., et al. (2014). Beyond growth rate 0.6: What drives *Corynebacterium glutamicum* to higher growth rates in defined medium. *Biotechnol. Bioeng.* 111, 359–371. doi: 10.1002/bit.25103
- Utawata, T. (2004). Production of arginine by fermentation. *J. Nutr.* 134, 2854S–2857S. doi: 10.1093/jn/134.10.2854S
- van Ooyen, J., Noack, S., Bott, M., Reth, A., and Eggeling, L. (2012). Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity. *Biotechnol. Bioeng.* 109, 2070–2081. doi: 10.1002/bit.24486
- Varma, A., and Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* 60, 3724–3731. doi: 10.1128/aem.60.10.3724-3731.1994
- Vertes, A. A., Inui, M., and Yukawa, H. (2013). “The biotechnological potential of *Corynebacterium glutamicum*, from Umami to Chemurgy,” in *Corynebacterium glutamicum* (Berlin; Heidelberg: Springer), 1–49. doi: 10.1007/978-3-642-29857-8_1
- Wendisch, V. F., Jorge, J. M., Pérez-García, F., and Sgobba, E. (2016). Updates on industrial production of amino acids using *Corynebacterium glutamicum*. *World J. Microbiol. Biotechnol.* 32, 105. doi: 10.1007/s11274-016-2060-1
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3, 1–9. doi: 10.1038/sdata.2016.18
- Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., et al. (2007). HMDB: the human metabolome database. *Nucleic Acids Res.* 35(Suppl_1):D521–D526. doi: 10.1093/nar/gkl923
- Xavier, J. C., Patil, K. R., and Rocha, I. (2017). Integration of biomass formulations of genome-scale metabolic models with experimental data reveals universally essential cofactors in prokaryotes. *Metab. Eng.* 39, 200–208. doi: 10.1016/j.ymben.2016.12.002
- Yamamoto, S., Suda, M., Niimi, S., Inui, M., and Yukawa, H. (2013). Strain optimization for efficient isobutanol production using *Corynebacterium glutamicum* under oxygen deprivation. *Biotechnol. Bioeng.* 110, 2938–2948. doi: 10.1002/bit.24961
- Zelle, E., Nöh, K., and Wiechert, W. (2015). “Growth and production capabilities of *Corynebacterium glutamicum*: interrogating a genome-scale metabolic network model,” in *Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications*, ed A. Burkowski (Poole: Caister Academic Press), 39–54. doi: 10.21775/9781910190050.04
- Zhang, Y., Cai, J., Shang, X., Wang, B., Liu, S., Chai, X., et al. (2017). A new genome-scale metabolic model of *Corynebacterium glutamicum* and its application. *Biotechnol. Biofuels* 10:169. doi: 10.1186/s13068-017-0856-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Feierabend, Renz, Zelle, Nöh, Wiechert and Dräger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.