Systematic sensitivity analyses of growth modelling to evaluate the robustness of Genome Scale Metabolic Network models — case study with the filamentous fungus *Penicillium rubens*

Nègre Delphine[1,2], Larhlimi Abdelhalim[2], Bertrand Samuel*[1]

[1] Nantes Université, Institut des Substances et Organismes de la Mer, ISOMer, UR 2160, F-44000 Nantes, France.
[2] Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France.

*Corresponding author
Correspondence: samuel.bertrand@univ-nantes.fr

ABSTRACT

The functionality—and thereby the quality—of a Genome-Scale Metabolic Network (GSMN) model is commonly evaluated by using a biomass reaction as the objective function to determine flux distributions under optimal network states through popular analyses such as Flux Balance Analysis (FBA) or Flux Variance Analysis (FVA). However, assessing model functionality solely by its ability to produce biomass is reductive; the predicted growth must also be biologically consistent. The definition and integration of this artificial reaction should theoretically reflect the organism-specific macromolecular composition and stoichiometry under defined environmental conditions. In practice, however, reconstruction decisions — whether documented, reasoned, or arbitrary — often lead to model refinements aimed at ensuring functionality, which can potentially introduce biases. Such accommodations, along with limited transparency in model development, may compromise the reliability of simulation outcomes. To mitigate this issue, we advocate for the systematic implementation of rapid and simple sensitivity analyses during the final stages of GSMN reconstruction. This approach enables the evaluation of model robustness and the predictive reliability of flux distributions with respect to the biomass reaction formulation. To illustrate this framework, we apply two sensitivity tests to the *i*Prub22 GSMN of the filamentous fungus *Penicillium rubens*, an organism known for its capacity to produce a broad range of high-value natural products. The first test evaluates sensitivity to the quantitative definition of the biomass reaction, and the second examines model behaviour under alternative nutrient conditions. Adopting this approach enables a quantitative assessment of the degree of independence exhibited by a model's flux productions with respect to the composition of biomass reactions.

Keywords: Network Validation, Systems Biology, Biomass Objective Function, Predictive Modelling, Flux Balance Analysis (FBA), Flux Variance Analysis (FVA), Stoichiometric coefficients, Media definition

Introduction

Since 2010, the reconstruction of Genome-Scale Metabolic Networks (GSMNs) has been guided by a reference protocol (Thiele & Palsson, 2010). However, the diversity of metabolic databases, the wide range of reconstruction and analysis tools, and the inconsistent definition and application of standards (Ebrahim *et al.*, 2015; Carey *et al.*, 2020) have often led to models functioning as “black boxes”, with key reconstruction decisions insufficiently documented. This lack of transparency limits their reusability and complicates the assessment of their predictive accuracy, as the reproducibility of results remains a notable obstacle (Tiwari *et al.*, 2021). Among the various criteria used to evaluate GSMN quality, the capacity to simulate organism growth — typically through a functional biomass reaction — remains central, as the predictive reliability of a model is closely tied to its ability to reproduce this fundamental biological process.

Traditional analyses of GSMNs typically rely on flux-based approaches such as flux balance analysis (FBA) (Orth *et al.*, 2010) and flux variability analysis (FVA) (Gudmundsson & Thiele, 2010). These methods are applied under defined conditions and assume the system operates in a steady state. The problem is formulated mathematically as a linear optimisation, where the objective function generally represents organism growth. FBA aims to predict a single optimal flux distribution that maximises the objective, while FVA explores the range of feasible fluxes for each reaction that is consistent with this optimum. These approaches are widely used to study metabolic capabilities, identify essential reactions, and explore network robustness.

The growth simulation relies on the formulation of the biomass reaction, an artificial construct representing the sum of macromolecules required for cellular structure and the energy necessary for their synthesis (Feist & Palsson, 2010). The biomass reaction, thus, reflects the dynamic interplay between catabolic pathways, releasing energy and precursors, and anabolic pathways that utilise these for biosynthesis and maintenance. Biomolecules, mainly carbon-based, form the cellular framework, with functional properties defined by specific groups. Carbohydrates, lipids, proteins, and nucleic acids constitute the central macromolecular classes, most of which are polymers of monomers. Accurate modelling of metabolism and growth requires a detailed understanding of their composition and functions.

Once the nature of these biomolecules is established, their quantities — as molar fractions and thus expressed in the stoichiometric coefficients of the biomass reaction — must be determined. Ideally, they are based on experimentally measured data from defined culture conditions (*e.g.,* continuous growth in a chemostat), with coefficients normalised per gram of dry biomass. This detailed characterisation enables precise calculation of biomass yields from stoichiometry. Beyond quantitative aspects, the qualitative composition of the biomass reaction — *i.e.,* the inclusion or exclusion of specific precursors — is critical for *in silico* gene knock-out simulations and identifying essential genes and reactions. In recent years, despite significant progress in improving the accuracy of biomass reaction definitions (Lachance *et al.*, 2019), the need to standardise the acquisition and integration of this reaction persists (Chan *et al.*, 2017).

According to this definition, and as expected, the biomass reaction is highly specific to the organism under study. Nevertheless, its formulation is often inherited from similar species due to the experimental challenges associated with its accurate determination (Beck *et al.* 2018). This aspect represents one of the main sources of uncertainty in defining biomass composition, an issue that is often underestimated (Bernstein *et al.* 2021). Indeed, biomass reactions from more or less related organisms are regularly used to characterise the organism of interest without necessarily considering the variations between different species (Xavier *et al.* 2017). When the biomass reaction employed in a model is derived from another reconstruction — whether from the same organism or a related species — a quality check of the biomass reaction can be performed using sensitivity analyses.

Beyond differences in biomass composition between organisms, variability may also exist within the same species. Changes in growth rate, nutrient availability, or physical parameters can affect organism growth (Bernstein *et al.* 2021), and cells respond through metabolic rearrangements that influence the biosynthesis of metabolic precursors required for building cellular components. Therefore, it is essential to consider the dependency of biomass composition on environmental conditions (Schulz *et al.* 2021).

The fundamental objective of sensitivity analyses is to assess model robustness and identify critical parameters for biomass production. These analyses examine how an optimal output state of a mathematical model responds, qualitatively or quantitatively, to various changes in model inputs. For example, sensitivity analysis of the biomass objective function to variations in biomass precursor coefficients (Nanda *et al.* 2020) helps verify whether *in silico* growth is sensitive to all biomass precursors. Similarly, sensitivity analysis of exchange reaction boundary variations (Lachance *et al.* 2021) allows testing model responses to environmental changes and, where relevant, determining optimal ranges for designing different culture media.

Therefore, we propose to systematically apply these two approaches to easily and rapidly assess how the definition of the biomass reaction influences the predictive capabilities of metabolic models. We illustrate this using *i*Prub22, the GSMN of the filamentous fungus *Penicillium rubens* (Nègre *et al.*, 2023), which, like many filamentous fungi, exhibits substantial biosynthetic capacity for specialised metabolism production (Eagan & Keller, 2025). As a complement, we extend the investigation to the production reactions of such compounds, which are typically not directly involved in biomass synthesis and are often considered functionally decoupled from growth. Exploring how the definition of the biomass reaction affects the predicted production of such compounds provides insight into model robustness and versatility. This evaluation bridges primary metabolism, growth, and specialised metabolism, enabling a more comprehensive assessment of model performance in capturing physiologically relevant phenotypes. Ultimately, the predictive quality of *i*Prub22 depends on its ability to simulate the various phenotypes observed and documented during experimentation. These simulations serve to evaluate the impact of biomass formulation on the independence of predictions and on model sensitivity, criteria that, when met, attest to the quality of the model used. Such insights will help advance understanding of the relationships between environment, metabolism, and biomass composition in the organism under study.

Materiels & Methods

Model and Computational Environment

The GSMN model used in this study corresponds to the most recent reconstruction of *P. rubens*, *i*Prub22 (Nègre *et al.*, 2023). The model is publicly available *via* BioModels under the identifier MODEL2306150001.

All simulations were primarily conducted in MATLAB (MathWorks Inc., Natick, MA, USA — R2021a) using the COBRA Toolbox v3.4 (Heirendt *et al.*, 2019) and the Gurobi optimiser (v12.0.2). Simulation results were exported to R (v4.3.3) and visualised in RStudio using the ggplot2 package (v3.4.2) (Wickham, 2016). Alternatively, simulations and visualisations can be reproduced using Python (v3.12.7) and COBRApy (v0.29.0) (Ebrahim *et al.*, 2013), thereby avoiding reliance on proprietary software. All scripts required for both workflows (*i.e.,* MATLAB–R and Python-only) are archived on Zenodo under the URL https://doi.org/10.5281/zenodo.17091974. To ensure full reproducibility, Conda environment files are supplied for both the R-based visualisation and the Python-based simulation and plotting pipelines. Interactive files are provided in their native formats—MATLAB LiveScripts, R Markdown, and Jupyter Notebooks—as well as in HTML-rendered versions to support transparency, reproducibility, and exportability. The MATLAB/R process requires approximately 22 min from model loading to data generation and plotting, whereas the Python-based pipeline takes approximately 26 min. All simulations were executed on a Dell Precision 3591 mobile workstation running Ubuntu 22.04.5 LTS (Linux kernel 6.8.0-60-generic, x86_64 architecture). The machine was equipped with an Intel® Core™ Ultra 7 165H processor (16 physical and 22 logical cores) and 66.85 GB of RAM.

Constraint Specification

The default model constraints were maintained. The minimal nutrient set required for model functionality included a carbon source, a nitrogen source, molecular oxygen (Uptake_136), ferrous iron (Uptake_062), sulphur (Uptake_169), inorganic phosphate (Uptake_146), thiamine (Uptake_171), and riboflavin (Uptake_157). The latter two were necessary to activate the cofactor (COF) subsystem by enabling the biosynthesis of thiamine diphosphate and flavin adenine dinucleotide (FAD). It should be noted that flavin mononucleotide (Uptake_065) and hydrogen peroxide (Uptake_077) can be used as substitutes for riboflavin and the oxygen source, respectively. Oxygen uptake was left unconstrained. The upper bounds for carbon and nitrogen sources were 15 and 5 mmol gDW^{-1} h^{-1}, respectively. All other mentioned uptakes were left to 10 mmol gDW^{-1} h^{-1}.

Uniquely, for specialised metabolite production analyses, three additional uptake reactions were enabled to account for cofactor and precursor requirements: Uptake_155 (*i.e.,* Red-NADPH-Hemoprotein-Reductases, a class of enzymes involved in redox reactions using NADPH to reduce haemoproteins), Uptake_044 (decanoate), and Uptake_144 (phenoxyacetate). Each was constrained to 10 mmol gDW^{-1} h^{-1}.

Stoichiometric Sensitivity of the Biomass Reaction

To assess the model's sensitivity to quantitative variations in the biomass reaction (Biomass_rxn), the stoichiometric coefficient of each reactant in the biomass equation, except for ATP, was independently modified by ±50% of its original value. To refine the resolution of the analysis, the same perturbation strategy was subsequently applied to the seven metabolic subsystems responsible for the biosynthesis of each biomass precursor: cofactors (r1465), DNA (r1458), RNA (r1457), phospholipids (r1460), cell wall components (r1455), proteins (r1456), and the free amino acid pool (r1459). In all FBA simulations, the biomass objective function was maximised at 100%.

For specialised metabolite production, FVA was performed under the constraint of 80% maximal biomass flux to estimate theoretical upper bounds for key precursors (*i.e.,* metabolic branch points) or final products of specialised metabolites known to be produced by *P. rubens*. Only stoichiometric coefficients from the biomass reaction itself were tested. The relevant production reactions were those leading to the production of: isopenicillin N (Production_010) and 6-aminopenicillanate (Production_001) for penicillins; 3-methylphenol (Production_016) for isoepoxydon and yanuthones; histidyltryptophyldiketopiperazine (HTD) (Production_022) for roquefortines; YWA1 (Production_033) for melanin; aristolochene (Production_036) for PR-toxin; 3,5-dimethylorsellinate (Production_037) for andrastin; ferrichrome (Production_034) and ferrichrome A (Production_035); averufin (Production_038); and versiconol (Production_040).

Sensitivity of Biomass Yield to Nutrient Availability

Robustness analysis was performed using the robustnessAnalysis function in the COBRA Toolbox. This method evaluates the sensitivity of a target flux — here, biomass production — to variations in individual reaction — here, nutrient uptake fluxes (Becker *et al.*, 2007). For each selected uptake reaction, the flux was fixed to 500 evenly spaced values between its minimum and maximum bounds. At each fixed value, FBA was conducted to maximise biomass production, and the resulting objective values were recorded. These data were exported to R CRAN and used to generate robustness profiles, representing biomass production as a function of nutrient availability. Carbon and nitrogen sources yielding strictly identical robustness profiles (*i.e.*, identical biomass values across all tested uptake rates) were grouped. A total of 31 carbon sources and 28 nitrogen sources (Table S1 in Appendices) were tested as potential substitutes for glucose and ammonia. Ammonia served as the fixed nitrogen uptake during carbon source evaluations; glucose was used as the fixed carbon source during nitrogen source analyses. The values displayed by a vertical bar on the robustness plots (Figure 3.A) correspond to the nutrient uptake rates required to reach the organism's theoretical growth rate, set at 0.20 h^{-1} (Nielsen, 1997).

Results

Model Sensitivity to Stoichiometric Perturbations in Biomass Reaction Reactants

- *Effect on Predicted Growth Rate*

To assess the robustness of the GSMN *i*Prub22, we examined the model's sensitivity to variations in the stoichiometric coefficients of the biomass reaction and its subsystems (Figure 1). Specifically, we aimed to quantify the dependency of predicted growth rate on biomass composition.

Each reactant within the biomass equation, except for ATP, was individually perturbed by ±50% of its original stoichiometric coefficient. FBA was then used to estimate the new optimal predicted specific growth rate change. Under default conditions, the model predicts a growth rate of 0.29 h^{-1}, which is consistent with experimental observations. Reactants whose perturbation led to marked deviations in the optimal biomass flux value are considered critical and thus require careful determination. Conversely, minor effects on growth rate indicate a degree of resilience to such stoichiometric variations.

Across the biomass reaction components, the maximum observed variation in predicted growth rate was 0.06 h^{-1} (Figure 1.A), indicating overall robustness. The most pronounced effects were associated with the PROTEIN and CELLWALL subsystems. In contrast, negligible changes—ranging from 10^{-3} to 10^{-6} h^{-1} — were recorded for synthetic precursor subsystems such as COF (Figure 1.B), DNA (Figure 1.C), RNA (Figure 1.D), PLIPIDS (Figure 1.E), AAPOOL (Figure 1.G), and PROTEIN (Figure 1.H). However, two CELLWALL precursors — UDP-*α*-glucose (CPD-12575) and UDP-*N*-acetyl-*α*-**D**-glucosamine (UDP-N-ACETYL-D-GLUCOSAMINE) — yielded a predicted growth rate interval that does not include the model's optimal reference value (Figure 1.F), suggesting potential vulnerabilities requiring further scrutiny.

- *Effect on Maximum Theoretical Production of Specialised Metabolites*

We further investigated whether perturbations in biomass precursor coefficients affect the model's capacity to produce selected specialised metabolites related to pathways included in the model. Using FVA, we quantified the maximum production fluxes of ten targeted metabolites or their immediate precursors under both default conditions and after ±50% variation in the stoichiometric coefficients of biomass components (Figure 2). All the minimum values of the flux were equal to 0. This analysis served as an indicator of the phenotypic stability of the model in response to changes in biomass composition.

Two distinct response patterns emerged. Adjustments to the stoichiometric coefficients of PROTEIN, RNA, and AAPOOL induced substantial changes in the production potential of penicillin (Figure 2.A), roquefortins (Figure 2.C), and ferrichrome-related compounds (Figure 2.E). In particular, modifying the PROTEIN subsystem yielded maximal flux deviations exceeding 0.5 $mmol \cdot gDW^{-1}\ h^{-1}$ for 6-aminopenicillanic acid, 0.3 $mmol \cdot gDW^{-1}\ h^{-1}$ for isopenicillin N, 0.2 $mmol\ gDW^{-1}\ h^{-1}$ for HTD, and 0.1 $mmol\ gDW^{-1}\ h^{-1}$ for ferrichromes.

In contrast, the production of isoepoxydon and yanuthones (Figure 2.B), melanin (Figure 2.D), PR-toxin (Figure 2.F), andrastin (Figure 2.G), averufin, and versicolorin (Figure 2.H) remained highly stable, with variations below 10^{-5} $mmol\ gDW^{-1}\ h^{-1}$.

Interestingly, perturbations in COF, AAPOOL, CELLWALL, and PLIPIDS occasionally led to marginally increased production fluxes. These differences are likely attributable to solver precision, as Gurobi's feasibility tolerance (*i.e.,* 10^{-6}) could explain discrepancies in this order. It is worth noting that variations of the order of 10^{-5} to 10^{-7}, typically observed in results generated under MATLAB using the COBRA Toolbox, were not reproduced in the Python implementation with COBRApy, where differences between minimum and maximum fluxes after perturbation were consistently in the range of 10^{-14}.

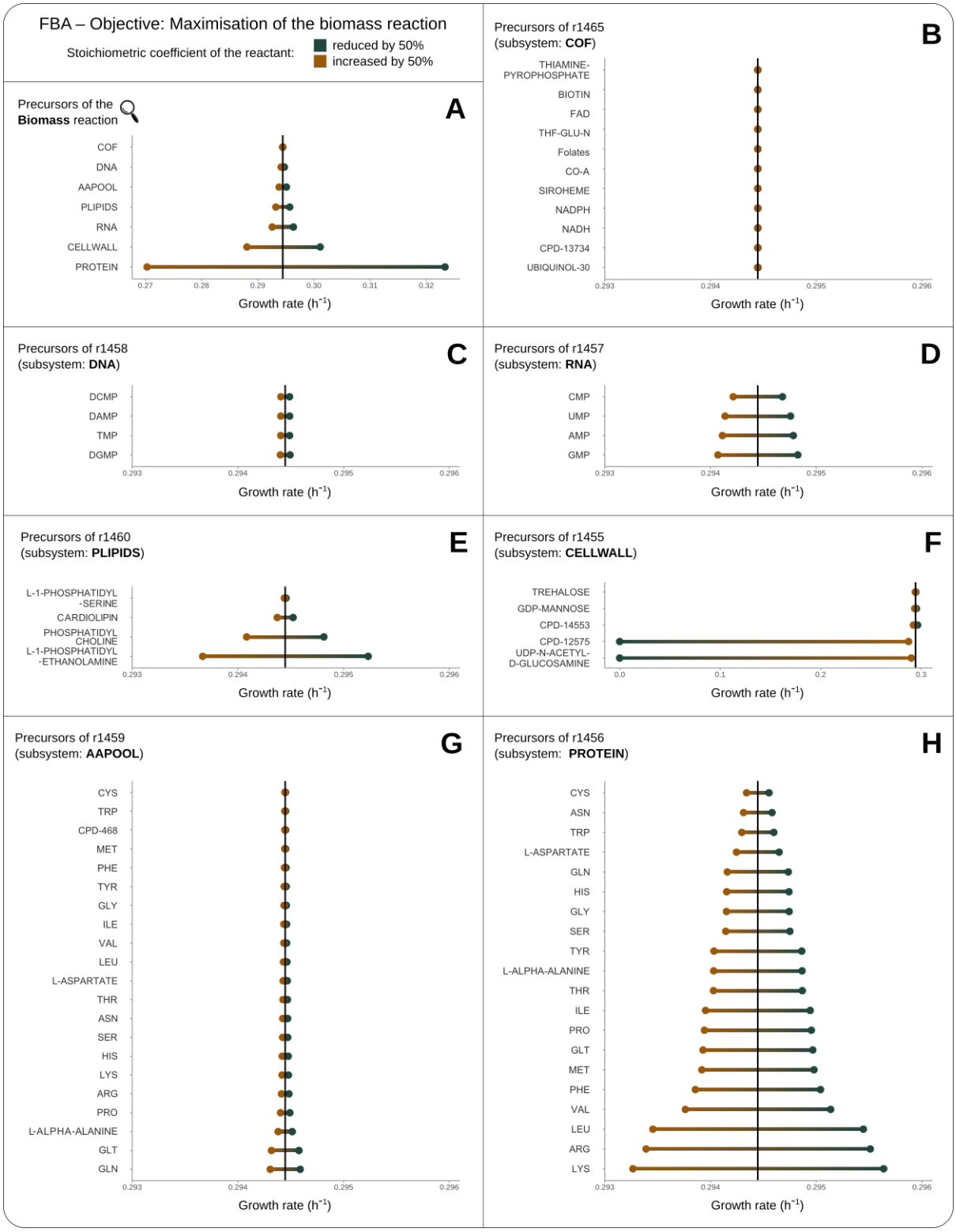

Figure 1 - Sensitivity analysis of the biomass objective function to variations in the coefficients of biomass precursors (A) and its subsystems (B-H). FBA was maximised to 100% for each variation in the coefficient of the precursor considered. The model constraints correspond to the default settings. The scales of the panels were harmonised, except for panels A and F, which display variations of larger magnitude. Respectively, brown (●) and green (●) dots represent the FBA results for a 50% increase and decrease in the coefficient value. The black vertical line indicates the biomass flux under default conditions, *i.e.,* 0.2944472 h^{-1}. Subsystem precursors are identified using their MetaCyc identifiers.

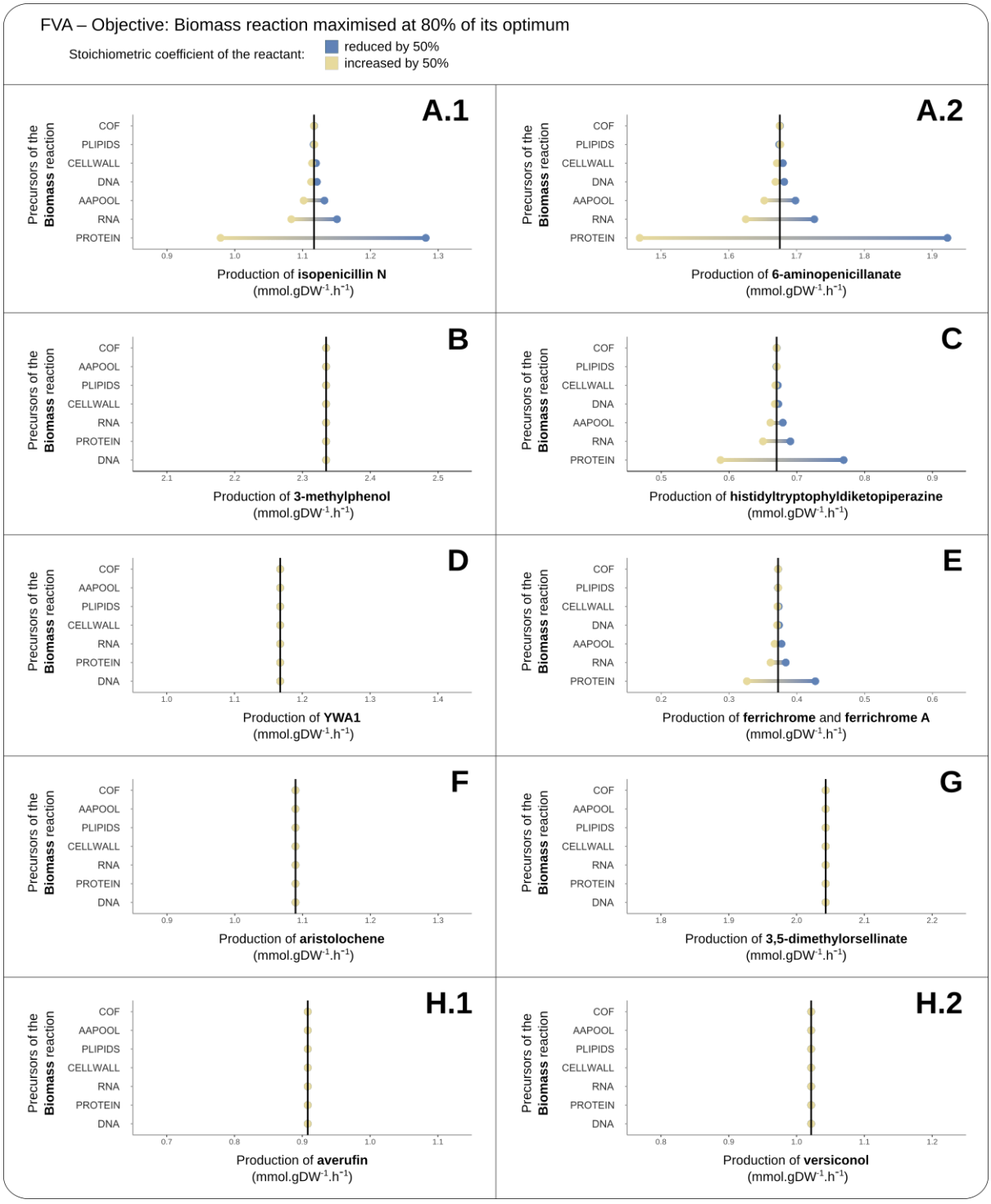

Figure 2 - Sensitivity analysis of specialised metabolite production fluxes in *Penicillium rubens* in response to variations in biomass precursor coefficients. An FVA maximising the biomass reaction to 80% was performed for each variation in the coefficient of a biomass precursor. Yellow (●) and blue (●) dots represent thbe maximum FVA values obtained after a 50% increase and decrease, respectively, in the coefficient of each precursor. Reactants are ranked by decreasing amplitude, with scales standardised across panels. The selected precursors (*i.e.,* branching points) are representative of the main biosynthetic pathways for (A) penicillins, (B) isoepoxydon and yanuthones, (C) roquefortines, (D) melanin, (E) ferrichromes, (F) PR-toxin, (G) andrastins, and (H) averufin and versicolorin B. The black vertical line represents the maximum production flux under the default biomass precursor coefficient value.

Model Sensitivity to Nutritional Environment

To evaluate the metabolic adaptability of *i*Prub22 to environmental changes, we examined how variations in the uptake fluxes of key nutrients affect biomass production. Using single-variable robustness analysis, we characterised the system's response to individual nutrient constraints, aiming to identify thresholds, sensitivities, and potential toxic effects (Figure 3.A). This approach provides insight into the physiological relevance of nutrient availability and the functional limitations embedded in the network structure.

Robustness profiles for phosphate, thiamine, and riboflavin exhibit linear behaviour, indicating proportional biomass response to increased import. However, the required quantities differ significantly: thiamine and riboflavin are essential in trace amounts, as evidenced by early inflexion points followed by saturation plateaus.

The oxygen uptake profile reveals a triphasic pattern: a steep initial rise in biomass flux reflecting high sensitivity, followed by a plateau suggesting oxygen saturation, and finally, a decline indicating inhibitory effects at elevated concentrations. Ferric iron and sulphur display similar biphasic behaviour, excluding the saturation phase. Their initial increase supports growth, confirming their limiting role. However, at higher import rates, a decline in biomass flux suggests toxicity or metabolic imbalance.

Substitutability tests revealed identical profiles for riboflavin and flavin mononucleotide (FMN), indicating functional redundancy in biomass formation. Conversely, the differences observed between oxygen and hydrogen peroxide suggest that, although both can function as electron acceptors, biomass production is more sensitive to oxygen, especially under conditions of low availability.

When analysing all available carbon and nitrogen sources, clustering based on biomass response identified ten and nineteen behavioural groups, respectively. All carbon sources showed proportional biomass increases with no saturation except for group C1, indicating that most tested conditions were suboptimal (Figure 3.B). Among nitrogen sources, groups N1-N15 exhibited biphasic response curves: an initial steep slope indicating strong sensitivity to nitrogen availability, followed by a shallower slope, reflecting diminishing returns (Figure 3.C). Group N16 reached saturation, while groups N17-N19 showed a subsequent decline in biomass production, consistent with inhibitory effects from nitrogen excess.

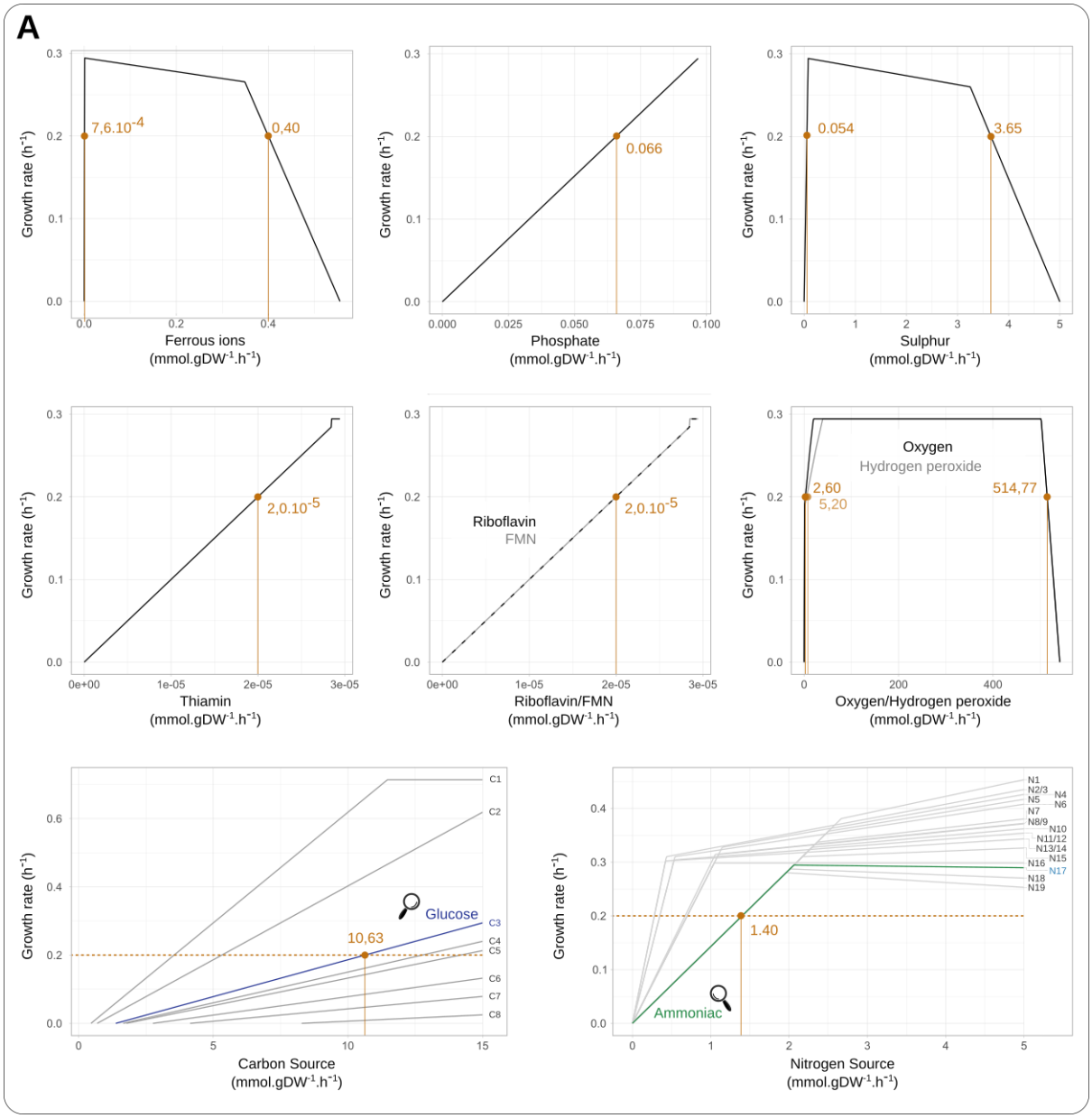

Figure 3 - Robustness analyses of the required uptake rates required for the functionality of *i*Prub22. (A) Robustness profiles for the minimum import fluxes are listed in the Materials & Methods section. Orange dots indicate the import flux values required to reach a biomass production rate of 0.20 h^{-1}, corresponding to a classical experimentally determined growth rate (Nielsen, 1997). (B) Sensitivity analysis of various carbon sources, grouped into eight categories according to carbon content (C1-C8). Except for groups C5, C7, and C8, the results show that a lower carbon-to-oxygen ratio (*i.e.,* lower carbon availability) and a lower carbon number generally require higher uptake rates to reach the target biomass value. The imposed upper limit of 15 mmol gDW^{-1} h^{-1} does not allow saturation plateaus to be reached. (C) Sensitivity analysis of various nitrogen sources, grouped into nineteen categories according to nitrogen content (N1-N19). For nitrogen sources, the carbon-to-nitrogen ratio appears to have a less pronounced effect on biomass production than the carbon-to-oxygen ratio observed for carbon sources

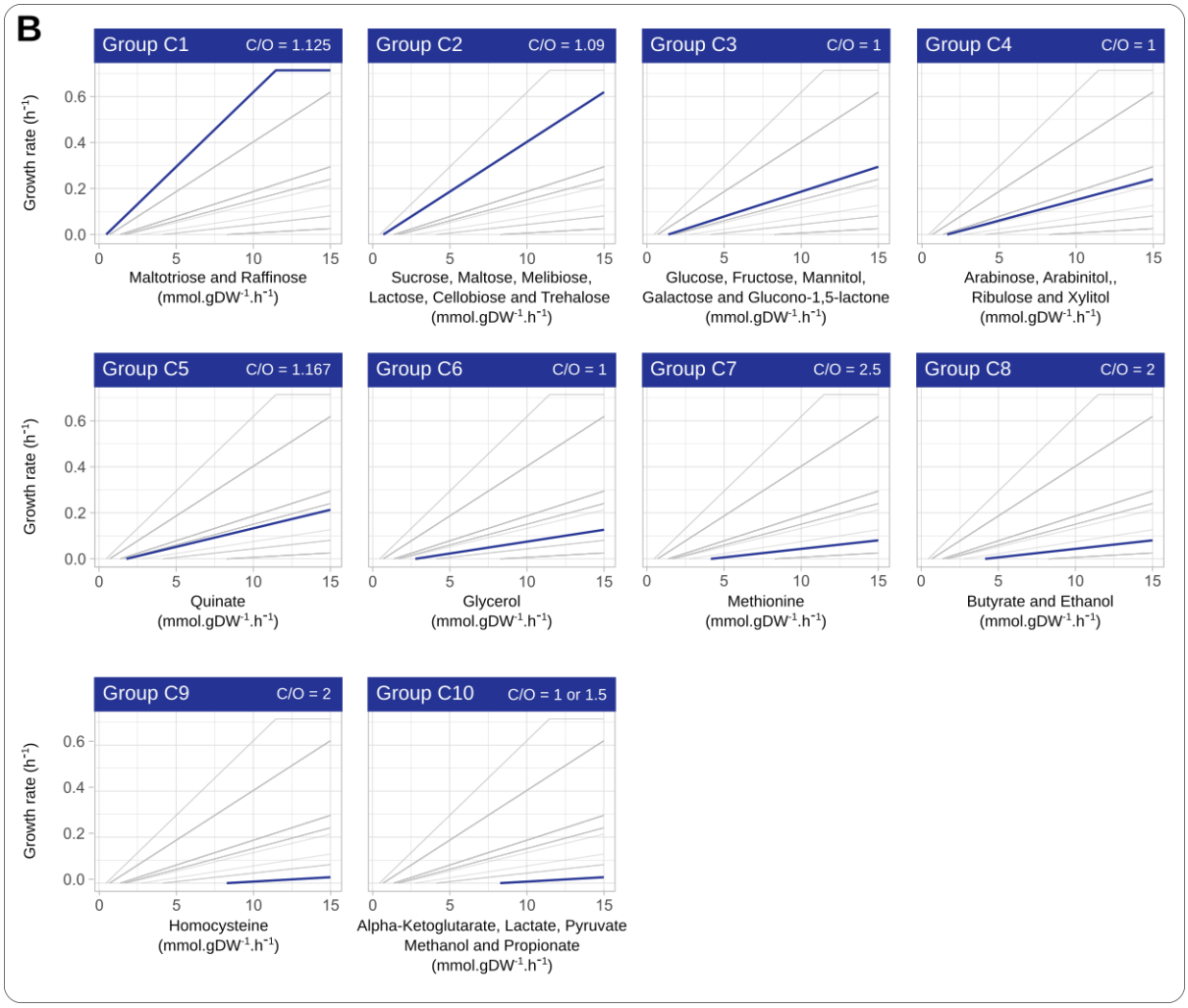

Figure 3 - (continued)

Figure 3 - (continued)

Discussion

Implications of Biomass Stoichiometry on Growth Rate Robustness

The sensitivity analysis of the *i*Prub22 biomass reaction highlights the contribution of individual precursors to the model's growth phenotype. Specifically, variations in stoichiometric coefficients reveal both the essentiality and potential limiting nature of specific metabolites. An increase in growth rate following the upscaling of a precursor's coefficient suggests that this compound constrains biomass production under default conditions. Conversely, a decline in growth upon a decrease in the coefficient confirms its essential role in maintaining viability.

However, while a linear framework may predict reduced growth under increased biosynthetic demand, metabolic networks exhibit non-linear behaviour. Intracellular flux redistribution and regulatory mechanisms can modulate or exacerbate these effects, complicating direct interpretations. Nonetheless, results obtained from such simplified analyses provide valuable indications for identifying potential bottlenecks in simulations and highlighting critical control points within the system.

UDP-related compounds exemplify this complexity, as shown for UDP-*α*-glucose (CPD-12575) and UDP-*N*-acetyl-*α*-D-glucosamine (UDP-N-ACETYL-D-GLUCOSAMINE), two CELLWALL precursors involved in glycogenesis and glycoprotein/glycolipid biosynthesis, respectively. Alterations in their coefficients lead to suboptimal biomass production. In both cases, increased coefficients reduce growth, likely due to elevated metabolic costs. Decreasing the coefficients results in a complete loss of biomass production, indicating that these precursors are mandatory for cell viability. These findings suggest that the UDP-related compounds coefficient used in *i*Prub22 — based on empirical data from *i*AL1006 (Agren *et al.*, 2013) — lies within a narrow range that supports optimal performance. The observed thresholds reflect the critical importance of maintaining precise stoichiometric balances.

By applying ±50% variations to the baseline stoichiometric values, the model was deliberately subjected to extreme perturbations. The resulting patterns suggest that, within the metabolic network of *P. rubens*, the precise quantitative specification of biomass — while empirically informed — may be of lesser importance than its qualitative composition. Indeed, the biomass reaction in *i*Prub22 displays broad robustness to most stoichiometric perturbations yet exhibits pronounced sensitivity to a subset of precursors involved in energy metabolism and cell wall biosynthesis. These metabolites define narrow tolerance thresholds, beyond which both depletion and excess compromise cellular growth. Accurate experimental calibration of their stoichiometric coefficients is, therefore, critical to preserve the model's predictive reliability.

Implications of Biomass Stoichiometry on Specialised Metabolite Production

Specialised metabolites — such as non-ribosomal peptides, polyketides, terpenes, and hybrid compounds — confer ecological and competitive advantages to their producer organism (Rokas *et al.*, 2020). The simulations performed with *i*Prub22 demonstrate a consistent and biologically plausible response of these biosynthetic pathways to perturbations in biomass composition.

As anticipated, increasing the stoichiometric coefficients of biomass precursors reduces the metabolic space available for specialised metabolite synthesis, while decreasing them tends to enhance it. This trend is particularly pronounced for non-ribosomal peptides, such as penicillins, roquefortines, and ferrichromes, whose production strongly depends on the availability of amino acids. Accordingly, the most significant shifts in predicted fluxes are observed following perturbations in the PROTEIN and AAPOOL subsystems, which represent the primary sources of amino acids in the model.

By contrast, only negligible perturbations were observed for specialised metabolites such as polyketides and terpenes, whose biosynthesis mainly relies on acetyl-CoA and malonyl-CoA. This outcome suggests an almost complete independence of these pathways from biomass composition, although minor variations may arise from solver sensitivity or marginal shifts in model behaviour.

Moreover, when the organism seeks to maximise growth, it does not prioritise the production of specialised metabolites, as evidenced by FVA minimum fluxes consistently equal to zero. This supports the notion of a certain orthogonality between growth and specialised metabolite production. It indicates that the biosynthetic pathways for specialised metabolites are non-essential for growth and represent facultative or alternative routes. In contrast, the theoretical maximum fluxes reveal the system's flexibility, illustrating the range of metabolite production that can be accommodated without compromising growth.

Despite the magnitude of the stoichiometric variations applied (±50%), the resulting changes in metabolite production largely fall within the expected range of biological variability. These results confirm both the sensitivity and the flexibility of *i*Prub22, while also mirroring potential fluctuations observed *in vitro*. However, caution is warranted when interpreting these variations in the context of simulations involving environmental constraints, particularly for pathways that rely on nitrogen-rich precursors.

Taken together, these observations indicate that the quantitative formulation of the biomass reaction exerts minimal influence on the production capacity of most specialised metabolites — except for non-ribosomal peptides —, which remain more susceptible due to their dependence on amino acid availability. The model's overall robustness across perturbations reinforces its utility for simulating specialised metabolism and highlights amino acid metabolism as a key determinant of biosynthetic responsiveness.

Metabolic Response of Biomass Production to Nutrient Variability

Fungal metabolism requires a set of essential nutrients, notably carbon and nitrogen, in addition to oxygen and hydrogen, which are indispensable for energy production and redox balance. Carbon and nitrogen sources may be chemically independent (*e.g.,* glucose and ammonia) or co-associated within a single molecule (*e.g.,* amino acids, which contain both carbon and nitrogen moieties). Phosphorus and sulphur are also fundamental, typically assimilated as inorganic phosphate and sulphate ions.

Robustness analysis of the *i*Prub22 model revealed distinct metabolic responses to variations in nutrient uptake rates. Linear profiles suggest proportional responses in biomass flux to increasing nutrient availability, while saturable responses indicate the existence of a threshold beyond which further increases no longer enhance growth. Inhibitory profiles, characterised by decreasing biomass production beyond a critical uptake rate, suggest metabolic overload or toxic accumulation effects. These patterns point to key physiological constraints within the network and highlight nutrient-specific tolerances. Plateaus represent regions of insensitivity, whereas abrupt changes in slope — critical points — indicate high sensitivity to nutrient limitation or excess. Such features are essential for identifying metabolic bottlenecks and informing targeted optimisation strategies in strain design or media formulation.

Qualitatively, most simulations align with biological expectations. For instance, the oxygen uptake profile is likely attributable to oxidative stress or broader metabolic disruption, whereas the iron uptake pattern may involve distinct mechanisms. These include oxidative damage (*e.g.,* free radical formation), disturbance of intracellular homeostasis and storage pathways, or cofactor depletion driven by elevated biosynthetic demands. However, discrepancies remain at the quantitative level, particularly regarding amino acid utilisation. Amino acids differ in their carbon-to-nitrogen ratios, degradation pathways, and energetic yields. Despite this, the model predicts unexpectedly low growth rates for substrates such as alanine, glutamate, or aspartate — commonly efficient in fungal metabolism — while complex amino acids like lysine yield comparatively higher growth. This lack of consistency with empirical observations suggests

limitations in reaction curation, cofactor coupling, or flux partitioning that require further refinement to improve the model's predictive resolution.

Conversely, other nutrient simulations reproduce experimental patterns more accurately. For example, in line with the findings of Allam and Elzainy (Allam & Elzainy, 1969), the model identifies hypoxanthine and adenine as favourable nitrogen sources, followed by xanthine, urea, and nitrate. These compounds are grouped into similar robustness clusters, reflecting shared assimilation routes and metabolic efficiency, thereby supporting the model's structural coherence in selected sub-networks.

Altogether, the analysis demonstrates that *i*Prub22 is capable of capturing fundamental metabolic responses to variations in nutrients. The identification of linear, threshold, and inhibitory uptake-growth relationships allows for a detailed characterisation of network flexibility and constraint. Nonetheless, the present study considers only single-variable perturbations. Further analyses incorporating combinatorial flux variations — such as double robustness tests or phenotype phase plane analysis — are needed to reveal interaction effects and refine the model's applicability for systems-level predictions.

Is Metabolic Database Content a Limiting Factor in Accurate Growth Modelling and Simulations?

In this study, we focused on the quantitative definition of the biomass reaction rather than its qualitative composition — an essential yet frequently overlooked aspect of GSMN reconstruction (Bernstein *et al.*, 2021). This limitation is relatively less critical for *P. rubens*, given the extensive experimental knowledge available since the 1950s (Nielsen, 1997). Nonetheless, compared with the earlier model *i*AL1006 (Agren *et al.*, 2013), the updated reconstruction *i*Prub22 (Nègre *et al.*, 2023) lacks several metabolic subsystems, notably those involved in the biosynthesis of fatty acids, phospholipids, neutral lipids, sterol esters, and glycerides. These omissions are primarily attributable to the absence of specific or compatible identifiers in MetaCyc. More broadly, the lack of standardised cross-references, the use of generic compound names, and the limited coverage of certain metabolic classes constrain the integration of several biosynthetic pathways (Pham *et al.*, 2019). This phenomenon is particularly evident in lipid metabolism, where reconstruction remains challenging despite recent efforts to improve annotation and ontology coverage (Poupin *et al.*, 2020). These gaps may limit the model’s ability to capture growth dynamics accurately.

Conclusion

Transparency in data acquisition and biomass reaction definition is a key factor for ensuring the reliability of a metabolic model’s predictive capabilities. In this study, we assessed the sensitivity of *i*Prub22, the last GSMN available for *P. rubens,* to quantitative variations in the biomass reaction and its subsystems. The model proved robust to moderate perturbations in biomass composition, suggesting a level of biological consistency that supports its reliability.

We further performed single-reaction robustness analyses to assess metabolic tolerance to nutrient flux variations, identify critical thresholds affecting viability, and evaluate metabolic flexibility under alternative nutritional scenarios. These tests revealed variation patterns consistent with known physiology and highlighted key exchange reactions shaping growth predictions, thereby attesting to the robustness and overall quality of the *i*Prub22 model.

Through these examples, we aimed to underline the relevance of simple, rapid, and pre-existing analyses for GSMN evaluation and refinement. Beyond these results, we propose that systematically investigating the degree of independence between model flux distributions and the biomass reaction composition represents a valuable step in GSMN validation. Such an approach would help identify models whose predictive properties are not excessively sensitive to arbitrary choices in biomass formulation, thereby strengthening confidence in their biological relevance.

Thus, while simulations were conducted in MATLAB and figures generated in R, we also provide a lightweight Python-based implementation to facilitate wider adoption of these practices. This version bypasses licensing constraints and is designed for easy integration into diverse modelling workflows. Despite their low computational cost and straightforward implementation, such tests remain, in our opinion, predominantly underutilised in model reconstruction workflows. Their systematic inclusion would enhance the transparency, reliability, and biological interpretability of published GSMNs, supporting both their validation and broader applicability in predictive systems biology.

Funding

This work was supported by the French National Research Agency (ANR) under project number ANR-18-CE43-0013.

Data, scripts, code, and supplementary information availability

The model used in this study is the GSMN *i*Prub22 of *Penicillium rubens*, publicly accessible via BioModels (MODEL2306150001) https://doi.org/10.1371/journal.pone.0289757 (Nègre *et al.*, 2023).

Scripts and computational workflows (*i.e.,* scripts and result files) required to reproduce all analyses are archived on Zenodo at the following URL: https://doi.org/10.5281/zenodo.17091974.

Appendices

Table S1 - Metabolites supporting growth as sole carbon or nitrogen sources in *i*Prub22. This table presents metabolites tested individually as sole carbon or nitrogen sources under defined minimal conditions, while maintaining the *i*Prub22 model functionality. Model constraints were consistent with those determined by default. Only compounds that sustain a non-zero biomass flux are included. Carbon and nitrogen sources are categorised accordingly and identified by their compound name and corresponding uptake reaction. Elements are presented in decreasing order of their potential efficiency in supporting biomass production in *i*Prub22 simulations. Terms in bold followed by an asterisk (*) are derived from published methodologies, such as the patent on the fermentative production of industrially relevant compounds using chemically defined media (Laat *et al.*, 2002).

Type	Metabolite Name	Chemical Formula	Metabolite ID	Uptake ID
Carbon sources				
Dicarboxylates	*α*-ketoglutarate	$C_5H_4O_5$	2-KETOGLUTARATE	Uptake_005
Non-Proteinogenic Amino Acids	Homocysteine	$C_4H_9NO_2S$	HOMO-CYS	Uptake_098
Monocarboxylate	Lactate	$C_3H_5O_3$	L-LACTATE	Uptake_001
Alcohol	Methanol*	CH_4O	METOH	Uptake_123
Fatty acid	Propionate*	$C_3H_5O_2$	PROPIONATE	Uptake_149
Monocarboxylate	Pyruvate	$C_3H_3O_3$	PYRUVATE	Uptake_151
Fatty acid	Butyrate	$C_4H_7O_2$	BUTYRIC_ACID	Uptake_028
Alcohol	Ethanol*	C_2H_6O	ETOH	Uptake_060
Proteinogenic Amino Acids	Methionine	$C_5H_{11}NO_2S$	MET	Uptake_104
Alcohol	Glycerol*	$C_3H_8O_3$	GLYCEROL	Uptake_072
Monocarboxylate	Quinate	$C_7H_{11}O_6$	QUINATE	Uptake_152
Monosaccharide carbohydrate	Arabinose	$C_5H_{10}O_5$	CPD-15699	Uptake_016
Monosaccharide carbohydrate	D-arabinitol	$C_5H_{12}O_5$	CPD-355	Uptake_043
Monosaccharide carbohydrate	Ribulose	$C_5H_{10}O_5$	L-RIBULOSE	Uptake_108
Monosaccharide carbohydrate	Xylitol	$C_5H_{12}O_5$	XYLITOL	Uptake_182
Monosaccharide carbohydrate	Fructose*	$C_6H_{12}O_6$	BETA-D-FRUCTOSE	Uptake_045
		$C_6H_{12}O_6$	CPD-15382	Uptake_046
Monosaccharide carbohydrate	Galactose	$C_6H_{12}O_6$	ALPHA-D-GALACTOSE	Uptake_048
Lactone	Glucono-1,5-lactone	$C_6H_{10}O_6$	GLC-D-LACTONE	Uptake_049
Monosaccharide carbohydrate	Glucose*	$C_6H_{12}O_6$	GLC	Uptake_026
		$C_6H_{12}O_6$	ALPHA-GLUCOSE	Uptake_015
Monosaccharide carbohydrate	Mannitol	$C_6H_{14}O_6$	MANNITOL	Uptake_054
Disaccharide carbohydrate	Cellobiose	$C_{12}H_{22}O_{11}$	CELLOBIOSE	Uptake_033
Disaccharide carbohydrate	Lactose*	$C_{12}H_{22}O_{11}$	CPD-15972	Uptake_086
Disaccharide carbohydrate	Maltose	$C_{12}H_{22}O_{11}$	MALTOSE	Uptake_118
Disaccharide carbohydrate	Melibiose	$C_{12}H_{22}O_{11}$	MELIBIOSE	Uptake_122
Disaccharide carbohydrate	Sucrose*	$C_{12}H_{22}O_{11}$	SUCROSE	Uptake_165
Disaccharide carbohydrate	Trehalose	$C_{12}H_{22}O_{11}$	TREHALOSE	Uptake_014
Trisaccharide carbohydrate	Maltotriose	$C_{18}H_{32}O_{16}$	MALTOTRIOSE	Uptake_119
Trisaccharide carbohydrate	Raffinose	$C_{18}H_{32}O_{16}$	CPD-1099	Uptake_154

Table S2 (continued)

Nitrogen sources

Inorganic compound	Nitrate	NO_3	NITRATE	Uptake_133
Inorganic compound	Cyanate	CNO	CPD-69	Uptake_042
Inorganic compound	Ammonia*	H_3N	AMMONIA	Uptake_130
Organic compound	Urea*	CH_4N_2O	UREA	Uptake_176
Inorganic compound	Ammonium*	H_4N	AMMONIUM	Uptake_019
Proteinogenic Amino Acids	Aspartate	$C_4H_6NO_4$	L-ASPARTATE	Uptake_091
Organic compound	Ethyl-nitronate	C_2H_5NO	CPD-320	Uptake_061
Proteinogenic Amino Acids	Glycine	$C_2H_5NO_2$	GLY	Uptake_073
Proteinogenic Amino Acids	Alanine	$C_3H_7NO_2$	L-ALPHA-ALANINE	Uptake_087
Proteinogenic Amino Acids	Glutamate*	$C_5H_8NO_4$	GLT	Uptake_095
Proteinogenic Amino Acids	Serine	$C_3H_7NO_3$	SER	Uptake_109
Nucleic Acid Components - Purine Bases	Adenine*	$C_5H_5N_5$	ADENINE	Uptake_012
Nucleic Acid Components - Purine Bases	Guanine*	$C_5H_5N_5O$	GUANINE	Uptake_076
Proteinogenic Amino Acids	Asparagine	$C_4H_8N_2O_3$	ASN	Uptake_090
Nucleic Acid Components - Purine Bases	Hypoxanthine*	$C_5H_4N_4O$	HYPOXANTHINE	Uptake_080
Nucleic Acid Components - Purine Bases	Xanthine	$C_5H_4N_4O_2$	XANTHINE	Uptake_180
Non-Proteinogenic Amino Acids	L-2-aminoadipate	$C_6H_{10}NO_4$	CPD-468	Uptake_084
Non-Proteinogenic Amino Acids	*γ*-aminobutyrate	$C_4H_9NO_2$	4-AMINO-BUTYRATE	Uptake_069
Proteinogenic Amino Acids	Threonine	$C_4H_9NO_3$	THR	Uptake_111
Proteinogenic Amino Acids	Glutamine	$C_5H_{10}N_2O_3$	GLN	Uptake_096
Cofactor	Glutathione	$C_{10}H_{16}N_3O_6S$	GLUTATHIONE	Uptake_071
Proteinogenic Amino Acids	Proline	$C_5H_9NO_2$	PRO	Uptake_107
Proteinogenic Amino Acids	Arginine	$C_6H_{15}N_4O_2$	ARG	Uptake_089
Organic compound	Choline*	$C_5H_{14}NO$	CHOLINE	Uptake_039
Proteinogenic Amino Acids	Lysine*	$C_6H_{15}N_2O_2$	LYS	Uptake_103
Amino sugars	Glucosamine	$C_6H_{14}NO_5$	GLUCOSAMINE	Uptake_050
Nucleic Acid Components - Purine ribonucleoside	Adenosine*	$C_{10}H_{13}N_5O_4$	ADENOSINE	Uptake_013
Proteinogenic Amino Acids	Phenylalanine	$C_9H_{11}NO_2$	PHE	Uptake_106
Proteinogenic Amino Acids	Tyrosine	$C_9H_{11}NO_3$	TYR	Uptake_113

References

Agren R, Otero JM, Nielsen J (2013) Genome-scale modeling enables metabolic engineering of *Saccharomyces cerevisiae* for succinic acid production. Journal of Industrial Microbiology & Biotechnology, 40, 735–747. https://doi.org/10.1007/s10295-013-1269-3

Allam AM, Elzainy TA (1969) Degradation of xanthine by *Penicillium chrysogenum*. Journal of General Microbiology, 56, 293–300. https://doi.org/10.1099/00221287-56-3-293

Beck AE, Hunt KA, Carlson RP (2018) Measuring Cellular Biomass Composition for Computational Biology Applications. Processes, 6, 38. https://doi.org/10.3390/pr6050038

Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nature Protocols, 2, 727–738. https://doi.org/10.1038/nprot.2007.99

Bernstein DB, Sulheim S, Almaas E, Segrè D (2021) Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. Genome Biology, 22, 64. https://doi.org/10.1186/s13059-021-02289-z

Carey MA, Dräger A, Beber ME, Papin JA, Yurkovich JT (2020) Community standards to facilitate development and address challenges in metabolic modeling. Molecular Systems Biology, 16, e9235. https://doi.org/10.15252/msb.20199235

Chan SHJ, Cai J, Wang L, Simons-Senftle MN, Maranas CD (2017) Standardizing biomass reactions and ensuring complete mass balance in genome-scale metabolic models. Bioinformatics, 33, 3603–3609. https://doi.org/10.1093/bioinformatics/btx453

Eagan JL, Keller NP (2025) Fungal secondary metabolism. Current Biology, 35, R503–R508. https://doi.org/10.1016/j.cub.2025.02.029

Ebrahim A, Almaas E, Bauer E, Bordbar A, Burgard AP, Chang RL, Dräger A, Famili I, Feist AM, Fleming RM, Fong SS, Hatzimanikatis V, Herrgård MJ, Holder A, Hucka M, Hyduke D, Jamshidi N, Lee SY, Le Novère N, Lerman JA, Lewis NE, Ma D, Mahadevan R, Maranas C, Nagarajan H, Navid A, Nielsen J, Nielsen LK, Nogales J, Noronha A, Pal C, Palsson BO, Papin JA, Patil KR, Price ND, Reed JL, Saunders M, Senger RS, Sonnenschein N, Sun Y, Thiele I (2015) Do genome-scale models need exact solvers or clearer standards? Molecular Systems Biology, 11, 831. https://doi.org/10.15252/msb.20156157

Ebrahim A, Lerman JA, Palsson BO, Hyduke DR (2013) COBRApy: Constraints-Based Reconstruction and Analysis for Python. BMC Systems Biology, 7, 74. https://doi.org/10.1186/1752-0509-7-74

Feist AM, Palsson BO (2010) The Biomass Objective Function. Current opinion in microbiology, 13, 344–349. https://doi.org/10.1016/j.mib.2010.03.003

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

Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V, Magnusdóttir S, Ng CY, Preciat G, Žagare A, Chan SHJ, Aurich MK, Clancy CM, Modamio J, Sauls JT, Noronha A, Bordbar A, Cousins B, Assal DCE, Valcarcel LV, Apaolaza I, Ghaderi S, Ahookhosh M, Guebila MB, Kostromins A, Sompairac N, Le HM, Ma D, Sun Y, Wang L, Yurkovich JT, Oliveira MAP, Vuong PT, Assal LPE, Kuperstein I, Zinovyev A, Hinton HS, Bryant WA, Artacho FJA, Planes FJ, Stalidzans E, Maass A, Vempala S, Hucka M, Saunders MA, Maranas CD, Lewis NE, Sauter T, Palsson BØ, Thiele I, Fleming RMT (2019) Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nature Protocols, 14, 639–702. https://doi.org/10.1038/s41596-018-0098-2

Laat WD, Preusting J, Koekman B (2002) Fermentative production of valuable compounds on an industrial scale using chemically defined media. *US patent application US20020039758A1*. (en)

Lachance J-C, Lloyd CJ, Monk JM, Yang L, Sastry AV, Seif Y, Palsson BO, Rodrigue S, Feist AM, King ZA, Jacques P-É (2019) BOFdat: Generating biomass objective functions for genome-scale metabolic models from experimental data. PLoS Computational Biology, 15, e1006971. https://doi.org/10.1371/journal.pcbi.1006971

Lachance J, Matteau D, Brodeur J, Lloyd CJ, Mih N, King ZA, Knight TF, Feist AM, Monk JM, Palsson BO, Jacques P, Rodrigue S (2021) Genome-scale metabolic modeling reveals key features of a minimal gene set. Molecular Systems Biology, 17, e10099. https://doi.org/10.15252/msb.202010099

Nanda P, Patra P, Das M, Ghosh A (2020) Reconstruction and analysis of genome-scale metabolic model of weak Crabtree positive yeast Lachancea kluyveri. Scientific Reports, 10, 16314. https://doi.org/10.1038/s41598-020-73253-3

Nègre D, Larhlimi A, Bertrand S (2023) Reconciliation and evolution of *Penicillium rubens* genome-scale metabolic networks–What about specialised metabolism? PLoS One, 18, e0289757. https://doi.org/10.1371/journal.pone.0289757

Nielsen J (1997) Physiological Engineering Aspects of *Penicillium chrysogenum*. World Scientific. https://doi.org/10.1142/3195

Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? Nature biotechnology, 28, 245–248. https://doi.org/10.1038/nbt.1614

Pham N, van Heck RGA, van Dam JCJ, Schaap PJ, Saccenti E, Suarez-Diez M (2019) Consistency, Inconsistency, and Ambiguity of Metabolite Names in Biochemical Databases Used for Genome-Scale Metabolic Modelling. Metabolites, 9, 28. https://doi.org/10.3390/metabo9020028

Poupin N, Vinson F, Moreau A, Batut A, Chazalviel M, Colsch B, Fouillen L, Guez S, Khoury S, Dalloux-Chioccioli J, Tournadre A, Le Faouder P, Pouyet C, Van Delft P, Viars F, Bertrand-Michel J, Jourdan F (2020) Improving lipid mapping in Genome Scale Metabolic Networks using ontologies. Metabolomics, 16, 44. https://doi.org/10.1007/s11306-020-01663-5

Rokas A, Mead ME, Steenwyk JL, Raja HA, Oberlies NH (2020) Biosynthetic gene clusters and the evolution of fungal chemodiversity. Natural Product Reports, 37, 10.1039.C9NP00045C. https://doi.org/10.1039/C9NP00045C

Schulz C, Kumelj T, Karlsen E, Almaas E (2021) Genome-scale metabolic modelling when changes in environmental conditions affect biomass composition. PLoS Computational Biology, 17, e1008528. https://doi.org/10.1371/journal.pcbi.1008528

Thiele I, Palsson BØ (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. Nature Protocols, 5, 93–121. https://doi.org/10.1038/nprot.2009.203

Tiwari K, Kananathan S, Roberts MG, Meyer JP, Sharif Shohan MU, Xavier A, Maire M, Zyoud A, Men J, Ng S, Nguyen TVN, Glont M, Hermjakob H, Malik-Sheriff RS (2021) Reproducibility in systems biology modelling. Molecular Systems Biology, 17, e9982. https://doi.org/10.15252/msb.20209982

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. 2nd edition. Springer, Cham. https://doi.org/10.1007/978-3-319-24277-4

Xavier JC, Patil KR, Rocha I (2017) Integration of Biomass Formulations of Genome-Scale Metabolic Models with Experimental Data Reveals Universally Essential Cofactors in Prokaryotes. Metabolic Engineering, 39, 200–208. https://doi.org/10.1016/j.ymben.2016.12.002