Exploring the metabolic profiling of *A. baumannii* for antimicrobial development using genome-scale modeling

Nantia Leonidou[1,2,3,4*], **Yufan Xia**[2], **and Andreas Dräger**[1,2,3,4]

[1]Computational Systems Biology of Infections and Antimicrobial-Resistant Pathogens, Institute for Bioinformatics and Medical Informatics (IBMI), Eberhard Karl University of Tübingen, Sand 14, 72076 Tübingen, Germany
[2]Department of Computer Science, Eberhard Karl University of Tübingen, Sand 14, 72076 Tübingen, Germany
[3]Cluster of Excellence 'Controlling Microbes to Fight Infections', Eberhard Karl University of Tübingen, Germany
[4]German Center for Infection Research (DZIF), partner site Tübingen, Germany
[*]Correspondence: nantia.leonidou@uni-tuebingen.de

ABSTRACT

With the emergence of multidrug-resistant bacteria, the World Health Organization published a catalog of microorganisms in 2017 for which new antibiotics are urgently needed. Within this list, the carbapenem-resistant pathogen Acinetobacter baumannii, belonging to the ESKAPE group, has been granted the "critical" status. Over the years, such isolates have been detected within healthcare units, posing a global threat to upcoming pandemics. One way to facilitate a systemic view of bacterial metabolism and allow the development of new therapeutics based on environmental and genetic alterations is to apply constraint-based modeling on metabolic networks. We developed a versatile workflow to build high-quality and simulation-ready genome-scale metabolic models. We applied our workflow to create a novel metabolic model for A. baumannii and validated its predictive capabilities using experimental nutrient utilization and gene essentiality data. Our analysis showed that our model *i*ACB23LX could recapitulate cellular metabolic phenotypes observed during *in vitro* experiments with an accuracy of over 80%, while positive biomass production rates were observed in growth media relevant to *A. baumannii*. Additionally, we identified putative essential genes with no human counterparts, which could serve as novel antibiotic candidates for the development of future antimicrobial strategies. Finally, we have assembled the first curated collection of available reconstructions for distinct *A. baumannii* strains and analyzed their growth characteristics. The presented models herein are in a standardized and well-curated format, facilitating their usability, while they can be used to guide the reconstruction of multi-strain networks. Ultimately, they serve as a knowledge base for reliable predictions under various perturbations and the development of effective drugs.

Keywords: *i*ACB23LX, *Acinetobacter baumannii*, ATCC 17978, constraint-based modeling, flux balance analysis, flux variability analysis, mathematical modeling, curated collection, Gram-negative, Proteobacteria, multi-drug resistant, ESKAPE, nasal microbiome, antimicrobial resistance, silent pandemic

Introduction

In the 21st century, treating common bacterial infections has become a global health concern. The rapid emergence of pathogens with newly developed resistance mechanisms led to the ineffectiveness of hitherto used antimicrobial drugs. According to their resistance patterns, bacteria are classified into three main categories: multidrug-resistant (MDR, resistant to at least one agent in more than three antibiotic categories), extensively drug-resistant (XDR, non-susceptible to one or two categories), and pandrug-resistant (PDR, non-susceptible to all drugs in all categories) [1]. Pathogens from the last two classes are called "superbugs". In February 2022, Murray et al. develop predictive statistical models within a large-scale global study and estimated 1.27 million deaths directly associated with antimicrobial resistance (AMR) [2]. The same study underlines the highly virulent ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.) as the primary cause of AMR-related deaths, while the World Health Organization (WHO) announced in 2017 the urgent need for novel and effective therapeutic strategies against these microorganisms, assigning them the "critical status".

As already known, Gram-negative bacteria are more resistant than Gram-positive. This is attributed to their outer lipopolysaccharide-rich and protective membrane [3]. Over the years, numerous studies highlighted the Gram-negative human pathogen *Acinetobacter baumannii* of substantial concern in hospital environments attributable to its increasing resistance and ability to survive from the biocides of antimicrobial agents [4, 5, 6, 7]. *A. baumannii* (from the Greek word *akínētos*, meaning "unmoved") is a rod-shaped, non-motile, and strictly aerobic bacterium. It is an opportunistic pathogen whose adaptable genetic apparatus has caused it to become endemic in intensive care units (ICUs), affecting immunocompromised patients, causing pneumonia, bacteremia, endocarditis, and more. Especially the carbapenem-resistant *A. baumannii* poses a serious global threat with high mortality rates [8, 9, 10].

It targets exposed surfaces and mucous tissues, colonizes the human nose [11, 12, 13] and is closely related to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infections [14, 15, 16, 17]. The skin has shown to be a community reservoir for *A. baumannii* in a very small percentage of samples [18, 19], while its prevalence in the soil is a frequent misconception as species from the genus Acinetobacter are ubiquitous in nature [6, 20]. Finally, it shows susceptibility to commonly used drugs, like β-lactams, aminoglycosides, and polymyxins.

Systems biology, and especially the field of genome-scale metabolic network analysis, is the key to exploring genotype-phenotype relationships, better understanding mechanisms of action of such threatening pathogens, and ultimately developing novel therapeutic strategies. Genome-scale metabolic models (GEMs) combined with constraint-based modeling (CBM) provide a well-established, fast, and inexpensive *in silico* framework to systematically assess an organism's cellular metabolic capabilities under varying conditions having only its annotated genomic sequence [21]. As of today, they have numerous applications in metabolic engineering, leading to the formulation of novel hypotheses driving the detection of new potential pharmacological targets [22].

It has been more than a decade since the release of the first mathematical simulation of *A. baumannii* metabolism. Kim et al. integrate biological and literature data to manually build AbyMBEL891, representing the strain AYE [23]. This model was further employed as an essential foundation for future reconstructions; however, its non-standardized and missing identifiers limited its use. Following a tremendous increase in the amount of literature and experimental data regarding *A. baumannii* (over 5,670 articles published between 2010 and 2017 according to PubMed), two novel strain-specific metabolic networks arose, *i*LP844 [24] and the AGORA (Assembly of Gut Organisms through Reconstruction and Analysis) model [25]. Both models were reconstructed in a semi-automated process and simulated the metabolism of two distinct strains: ATCC 19606 and AB0057, respectively. With the help of transcriptomic data of sampled colistin responses and *i*LP844, it was observed that the type strain ATCC 19606 underwent metabolic reprogramming, demonstrating a stress condition as a resistance mechanism against colistin exposure. Alterations in gene essentiality phenotypes between treated and untreated conditions enabled the discovery of putative antimicrobial targets and biomarkers. Moreover, the model for AB0057 was part of an extensive resource of GEMs built to elucidate the impact of microbial communities on host metabolism. The amount of mass- and charge-balanced reactions in these models is very high; however, they carry few to no database references. Norsigian et al. improved and expanded AbyMBEL891 to finally create the high-quality model *i*CN718 that exhibited a prediction accuracy of over 80% in experimentally data [26], while Zhu et al. built a GEM for ATCC 19606 (*i*ATCC19606) integrating multi-omics data [27]. Compared to *i*LP844, *i*ATCC19606 incorporates metabolomics data together with transcriptomic data enabling the deciphering of bactericidal activity upon polymyxin treatment and the interplay of various metabolic pathways. Last but not least, in 2020, the first *in vivo* study on *A. baumannii* infection was published utilizing constraint-based modeling [28]. This time, the collection of strain-specific models was enriched with the first GEM for the hyper-virulent strain AB5075 (*i*AB5075). The model was validated using various experimental data, while transcriptomics data was leveraged to identify critical fluxes leading to mouse bloodstream infections. Our literature search revealed one last metabolic model of *A. baumannii* ATCC 17978, named *i*JS784, which, by the time of writing, has not been officially published in a scientific journal or been deposited in a mathematical models database. Instead, it is currently available solely in the form of a dissertation [29]. Nonetheless, the model cannot produce biomass even when all uptake reactions are open and all medium nutrients are available to the cell, making it unusable and hampering reproducibility.

We expanded this collection by building a novel GEM for the nosocomial strain ATCC 17978, named *i*ACB23LX. The presented model follows the FAIR data principles and community standards and recapitulates experimentally-derived phenotypes with high predictive capability and accuracy scores. We enriched the model with numerous database cross-references, and we computationally inferred the minimal nutritional requirements. Moreover, we used this model to investigate the organism's growth ability in experimentally defined media and within the human nose while we assessed its ability to predict essential genes using two different optimization approaches. Among the discovered strains, ATCC 17978 is one of the most well-studied, with a substantial amount of experimental data available that can be used to direct model refinement and validation. Besides that, we systematically refined and evaluated all pre-existing reconstructions' performance to finally create the first compendium of curated and standardized models for *A. baumannii*. With this, we aim to promote further studies to give new insights into this pathogen and promote strain- and species-specific therapeutic approaches.

Results

Reconstruction process of the novel metabolic network *i*ACB23LX

To build a high-quality model for *A. baumannii* ATCC 17978, we developed a workflow shown in Figure 1 following the community standards [30] (see *Materials and Methods* for detailed description). We named the newly reconstructed network *i*ACB23LX, where *i* stands for *in silico*, ACB is the organism- and strain-specific three-letter code from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, 22 the year of reconstruction, and LX the authors' initials. Our protocol involves eight major stages starting from the attainment of the annotated genomic sequence until the model validation, applies to any organism from the tree of life (Archaea, Bacteria, and Eukarya), and ensures the good quality

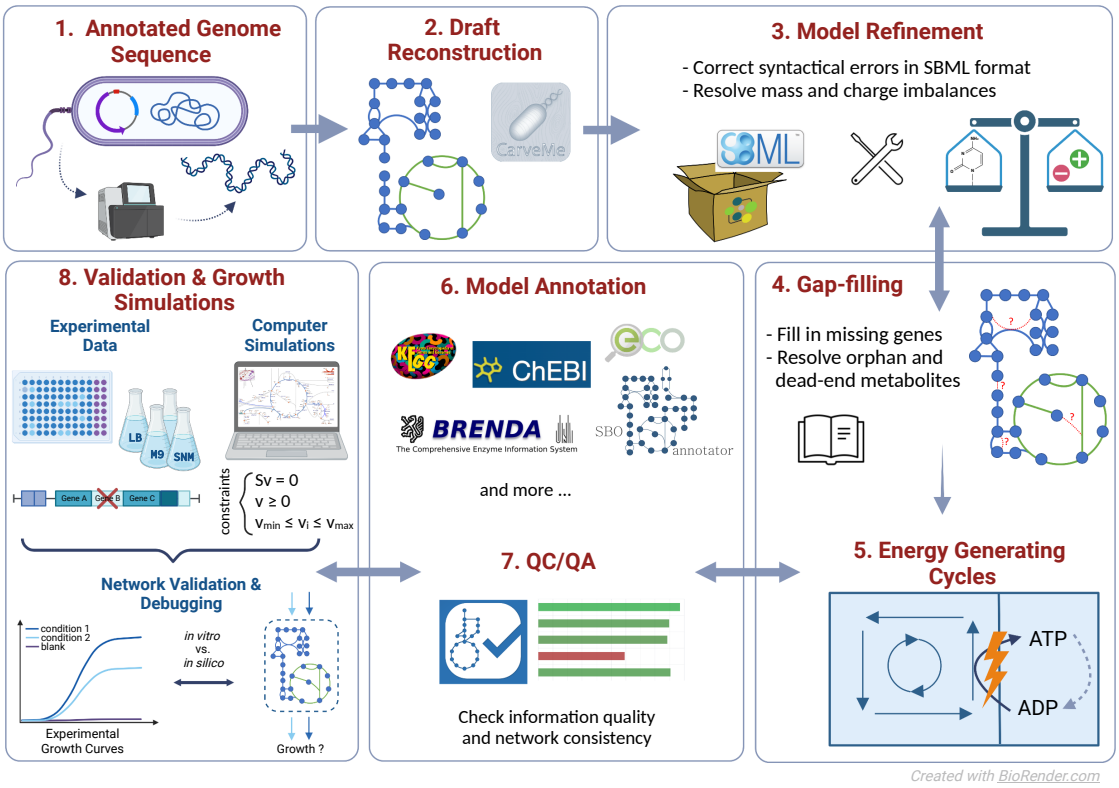

Figure 1 | Workflow developed for the metabolic network reconstruction of *i*ACB23LX. The created workflow consists of eight main steps: extraction of the annotated genome, draft model reconstruction, model refinement, gap-filling, investigation of energy generating cycles (EGCs), model annotation, quality control and quality assurance (QC/QA), and model validation using experimental data. The growth simulations include the examination of growth requirements and the definition of a minimal growth medium. The last six processes are continuously iterated until the model is of good quality and recapitulates known phenotypes.

and correctness of the final model. CarveMe [31] was used to build a preliminary model, which was subsequently extended and curated manually. We resolved syntactical issues and mass and charge imbalances during manual refinement while we defined missing metabolite charges and chemical formulas. Our final model contains no mass-imbalanced reactions, while two reactions remain charge-imbalanced. After extensive effort, resolving all charge imbalances was impossible since all participated metabolites are interconnected to multiple reactions within the network, and any modification in their charge resulted in newly introduced imbalances. The model extension process involved incorporating missing metabolic genes considering the network's connectivity. Dead-end and orphan metabolites do not exist biologically in the species, implying knowledge gaps in metabolic networks. Moreover, reactions including such metabolites are not evaluated in flux balance analysis (FBA). Hence, reactions with zero connectivity and no organism-specific gene evidence were omitted from the gap-filling. We extended the draft model by 138 reactions, 77 genes, and 110 metabolites in three compartments (cytosol, periplasm, and extracellular space). All in all, *i*ACB23LX comprises 2,321 reactions, 1,660 metabolites, and 1,164 genes (Figure 2). It is the most comprehensive model, while its stoichiometric consistency lies at 100% and contains no unconserved metabolites. Over 1,800 reactions have a gene-protein-reaction associations (GPR) assigned, while 149 are catalyzed by enzyme complexes (GPR contains at least two genes connected via a logical AND).

Furthermore, we tested our model for EGCs to prevent having thermodynamically infeasible internal loops that bias the final predictions [32]. We defined energy dissipation reactions (EDRs) for 15 energy metabolites and evaluated their production with FBA after disabling all nutrients from entering the system (see *Materials and Methods*). In our final model *i*ACB23LX, none of the tested metabolites could be produced; thus, no EGCs are contained. As shown in Figure 1, a plethora of database cross-references was embedded in the model, while Systems Biology Ontology (SBO) Terms were defined for every reaction, metabolite, and gene (Figure 3) [34]. Additionally, each reaction was mapped to an Evidence and Conclusion Ontology (ECO) Term representing the confidence level and the assertion method.

To assess the model quality, Metabolic Model Testing (`MEMOTE`) [36] and the SBML Validator from the libSBML[37] were used. Our metabolic network, *i*ACB23LX, has a `MEMOTE` score of 89%, and all syntactical errors are resolved. Our

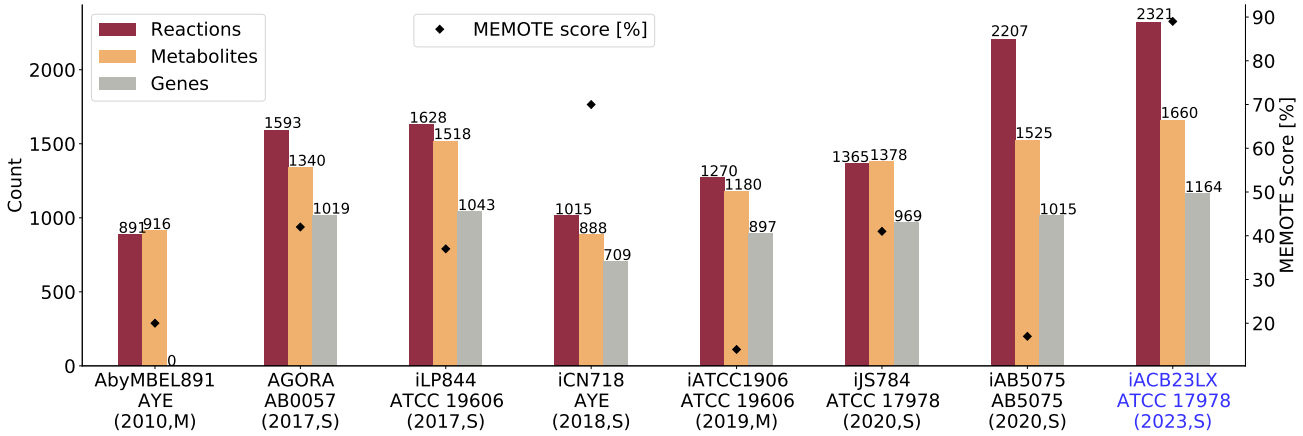

Figure 2 | Properties of all metabolic networks of *A. baumannii*. Blue highlights the novel reconstruction for the strain ATCC 17978 presented in this publication. The ordinate represents the `MEMOTE` scores. The reconstruction process is divided into manual (M, no computational tool was used to reconstruct and refine the model) and semi-automated (S, draft obtained via an automated reconstruction tool, while further extension was done manually) and is written together with the publication year. Also, the respective strains annotate the abscissa labels. Our new model exhibits the highest quality score and is more comprehensive and complete than the preceding reconstructions.

model undoubtedly exhibits the highest quality score among its predecessors (Figure 2). It must be noted here that the `MEMOTE` testing algorithm considers only the parent nodes of the SBO directed acyclic graph and not their respective children. Assigning more representative SBO Terms (Figure 3) does not increase the final score but reduces it by 2%.

The model is available as a supplementary file in the latest SBML Level 3 Version 2 [38] and JSON formats with the flux balance constraints (fbc) and groups plugins available.

*i*ACB23LX is of high quality and exhibits an increased prediction accuracy

Prediction of bacterial growth on various nutritional environments

Constraint-based modeling approaches such as FBA estimate flux rates with which metabolites flow through the metabolic network and predict cellular phenotypes for multiple growth scenarios. *A. baumannii* is known to be strictly aerobic and, compared to the majority of *Acinetobacter* species, it is not considered ubiquitous in nature. As a nosocomial pathogen, it has been mostly detected in hospital environments, particularly in the ICUs, and within the human nasal microbiota [11, 12, 13]. We examined various growth conditions to ensure that our model *i*ACB23LX recapitulates these already known and fundamental phenotypes.

First, we tested our model's capability to simulate a strictly aerobic growth. For this purpose, we examined the directionality of all active oxygen-producing and -consuming reactions when the oxygen uptake was disabled (Figure 4). We observed an accumulation of periplasmic oxygen by reactions that carried remarkably high fluxes and resulted in growth when the oxygen import was turned off. We examined each reaction individually and removed those without evidence to correct this. More specifically, we removed the periplasmic

Table 1 | Composition of the computationally defined minimal growth medium *i*MinMed. It consists of nine transition metals, a carbon source, a nitrogen source, a sulfur source, and a phosphate source. Oxygen was used to simulate aerobic conditions.

	Molecular Formula	Name
Carbon source	$C_6H_{12}O_6$	D-glucose
Nitrogen source	NH_4^{2+}	Ammonium
Sulfur source	SO_4^{2-}	Sulfate
Phosphate source	HPO_4^{2-}	Phosphate
	Ca^{2+}	Calcium
	Cl^-	Chloride
	Cu^{2+}	Copper
	Fe^{3+}	Ferric iron
Transition metals	Co^{2+}	Cobalt
	K^+	Potassium
	Mg^{2+}	Magnesium
	Mn^{2+}	Manganese
	Zn^{2+}	Zinc
Oxygen source	O_2	Oxygen

catalase (`CATpp`), one of bacteria's main hydrogen peroxide scavengers. This enzyme is typically active in the cytosol [40] and was not part of any precursor *A. baumannii* GEM or was found only in cytosol (*i*LP844[24]). To fill the gap and enable the usage of the periplasmic hydrogen peroxide, we included the reaction phenethylamine oxidase (`PEAMNOpp`) in the model. Eventually, our model *i*ACB23LX demonstrated growth only in the presence of oxygen using a rich medium (all exchange reactions are open).

Furthermore, we determined the minimal number of metabolites necessary for growth using our model *i*ACB23LX

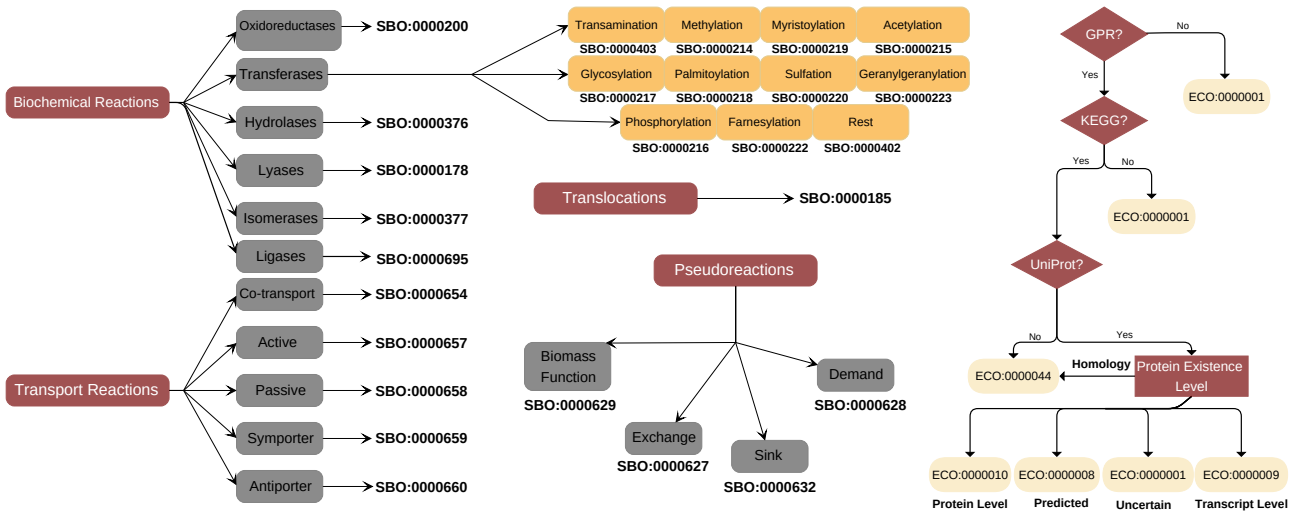

Figure 3 | Schematic representation of the SBO and ECO Terms mapping. It follows the graphs defined in the repository for biomedical ontologies Ontology Lookup Service (OLS) [33]. The SBO terms were added using the SBOannotator tool [34]. The ECO Terms annotated metabolic reactions and were declared based on the presence of GPR along with KEGG and UniProt annotations. Providing UniProt identifiers, the Protein Existence Level guides the mapping to appropriate ECO Terms. Figure created with yEd [35].

Table 2 | Simulated doubling times in various growth media. The tested media are the computationally self-defined minimal medium (*i*MinMed), the Luria-Bertani (LB) medium, the M9 minimal medium, and the synthetic nasal medium (SNM). All flux rates are given in $\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$. The simulated media are available in the supplementary file S1.

*i*MinMed	LB	M9	SNM
lyahoo 0.1677	0.5926	0.3735	0.5099

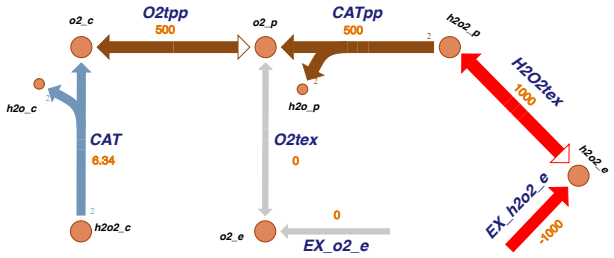

Figure 4 | Oxygen-producing and -consuming reactions found in *i*ACB23LX together with their anaerobic fluxes. All flux rates are written in orange and are given in $\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$. The reaction abbreviations are as follows: `O2tpp`, O_2 transport via diffusion between periplasm and cytosol; `CATpp`, periplasmatic catalase; `H2O2tex`, hydrogen peroxide transport via diffusion; `CAT`, Catalase; `O2tex`, O_2 transport via diffusion between periplasm and extracellular space; `EX_h2o2_e`, hydrogen peroxide exchange and `EX_o2_e`, O_2 exchange. Figure generated with Escher [39].

and the M9 medium (supplementary file S1) as a reference. Minimal growth media typically consist of carbon, nitrogen, phosphate, and sulfate sources, as well multiple inorganic salts and transition metals. These metals are crucial for the growth and survival of all three domains of life; however, they can be transformed into toxic compounds in hyper-availability [41]. The exact composition of our minimal medium (*i*MinMed) is shown in Table 1 (and in the supplementary file S1) and is already defined in the final model. It comprises nine transition metals, D-glucose as the carbon source, ammonium as a nitrogen source, sulfate as a sulfur source, and phosphate as a phosphate source. Oxygen is also a vital component, as *A. baumannii* is known to be a strictly aerobic bacterium. Previous studies have highlighted the importance of the nutrient metals for *A. baumannii* to survive within the host. More specifically, the bacterium utilizes these metals as co-factors for vital cellular processes [42]. Manganese and zinc have also been studied as essential determinants of host defense against *A. baumannii*-acquired pneumonia through their sequestering by calprotectin via a type of bonding called chelation [43]. With *i*MinMed, the optimization of the growth function re-

sulted in a flux rate of 0.1677 mmol/(gDW · h). Like most bacteria, laboratory cultivation of *A. baumannii* is typically done using two pre-defined culture media, the M9 minimal medium and the LB medium.To validate *i*ACB23LX with some of these data, we needed to ensure realistic growth in these two culture media. A realistic growth rate should also be guaranteed in the SNM that mimics the human nasal niche. Growth rates below 2.81 mmol/(gDW · h) are considered to be realistic since the doubling time of the fastest growing organism *Vibrio natriegens* is 14.8 minutes that correspond to 2.81 mmol/(gDW · h) [36]. Table 2 displays the predicted growth rates of *i*ACB23LX in the respective culture media. In the LB medium, our model grew with the highest rate; 0.5926 mmol/(gDW · h). With our self-defined minimal medium, *i*MinMed, our model exhibited the lowest rate; 0.1677 mmol/(gDW · h). Additionally, we examined the growth rate of our model in a rich medium, in which all nutrients are available to the model. With this, the flux through the biomass production was the highest, 2.1858 mmol/(gDW · h), as expected. This is still less than the growth rate of the fastest organism, increasing the confidence in our model's consistency and simulation capabilities. Initially, *i*ACB23LX could not predict any realistic growth rate for the simulated media. Using the gap-filling function of CarveMe [31], we detected three enzymes whose addition into the metabolic network resulted in successful growth in all tested media. These reactions are PHPYROX, OXADC, and LCYSTAT.

Functional validation of i*ACB23LX using experimental data*

Multiple *in silico* approaches have hitherto been employed to predict lethal genes and to assess growth metrics on different carbon/nitrogen sources for severe pathogenic organisms including *Mycobacterium tuberculosis* [44, 45] and *Staphylococcus aureus* [46, 47]. In 2013 and 2014, two studies were published that examined the phenome and gene essentialities of the strain ATCC 17978 [48, 49]. We used these datasets to evaluate the overall performance (functionality and accuracy) of *i*ACB23LX.

Our first validation experiment evaluated the accuracy of our model's carbon and nitrogen catabolism potentials. More specifically, we detected compounds that could serve as sole carbon and nitrogen sources using the large-scale phenotypic data by Farrugia et al.[48]. Although the authors tested more compounds in total, we could examine only 80 compounds as carbon sources and 48 metabolites as sole nitrogen sources. For the remaining molecules, either no Biochemical, Genetical, and Genomical (BiGG) identifier could be found, or they were not part of the metabolic network. According to the experimental protocol followed by Farrugia et al., we applied the M9 minimal medium and enabled D-xylose as the sole carbon source for the nitrogen testings. As D-xylose was initially not part of the reconstructed network, we conducted extensive literature and database search to include missing reactions. This improved the prediction accuracy, especially for the carbon sources, where an amelioration of 18% was achieved. In more detail, despite the comprehensive manual curation, the first draft model was reconstructed using the automated tool CarveMe [31]. This may result in the incorrect inclusion of transport reactions, which were consequently removed to reduce the number of false positive predictions. In both cases, our main objective was to improve the accuracy while keeping the number of orphan and dead-end metabolites low and removing only reactions with no gene evidence (lack of assigned GPR). Similarly, missing reactions were identified and included in the network to eliminate the false negative predictions. For instance, in accordance with the phenotypic data, the strain ATCC 17978 should not be able to grow when utilizing D-trehalose as the sole carbon source. Our model initially predicted a growth phenotype for this carbon source. To overcome this conflict, we deleted the reaction TREP with no organism-specific gene evidence, meaning no GPR was assigned. However, it was not feasible to resolve all inconsistencies since adding transport reactions to resolve false positives or false negatives in the nitrogen testings led to more false predictions in the carbon sources. More specifically, when adenosine, inosine, L-homoserine, and uridine are utilized as sole carbon sources, the model should not predict growth, while sole nitrogen sources should result in a non-zero objective value. In this case, adding transporters would resolve false predictions in the nitrogen tests, while it would have induced more false predictions in the carbon sources tests. All in all, *i*ACB23LX exhibited an overall accuracy 86.3% for the carbon and 79.2% for the nitrogen sources test (Figure 5c); however, after further curation, the accuracy was remarkably improved and reached 87.5%. By adding their corresponding transport reactions, we resolved discrepancies regarding uridine, inosine, adenosine, and L-homoserine.

Our model was able to catabolize 49 sole carbon and 40 sole nitrogen sources (see Figure 5a) and Figure 5b), recapitulating totally 69 and 38 experimentally-derived phenotypes, respectively.

We further assessed the ability of our model to predict known gene essentialities. First, 1,164 *in silico* single gene deletions were conducted on both LB and rich growth media, respectively, to identify all lethal gene deletions. Subsequently, the ratio between the growth rate after and before the respective knockouts (FC_{gr}) was calculated, and the genes were accordingly classified (see *Materials and Methods*). For the optimization, two mathematics-based approaches from the Constraints-Based Reconstruction and Analysis for Python (COBRApy) [50] package were deployed: the FBA [51] and the Minimization of Metabolic Adjustment (MOMA) [52]. Between the two methods, a similar distribution of the FC_{gr} values was observed (Figure 6a and Figure 6b). Using FBA, 97, 75, and 991 genes were predicted to be essential, partially-essential, and inessential on the LB medium, respectively, whereas optimization with MOMA resulted in 110, 85, and 968 genes (Figure 6c and supplementary files S2 and S3). These genes were primarily associated with the biosynthesis of cofactors and vitamins, the amino acid/nucleotide metabolism, the energy metabolism, and the metabolism of

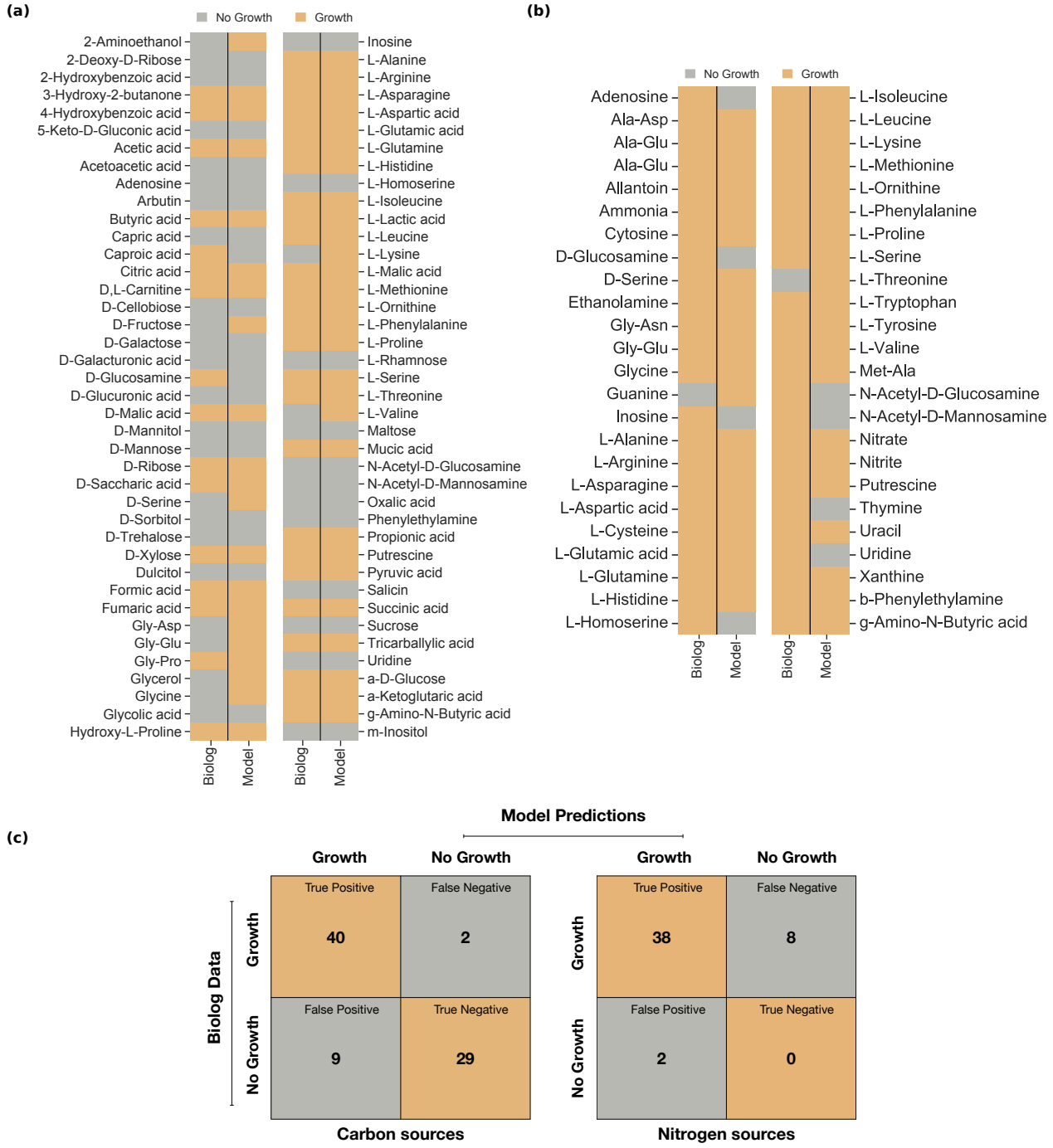

Figure 5 | Model predictions compared to the Biolog experimental measurements for various carbon and nitrogen sources. From the Biolog data, only substances mappable to model metabolites were included, while the M9 minimal medium was applied. a) and b) The model's ability to catabolize various carbon and nitrogen sources was assessed using the strain-specific phenotypic data by Farrugia et al.[48]. Blue indicates no growth, and orange indicates growth. Totally, 80 and 48 compounds were tested as sole carbon and nitrogen sources, respectively. Out of these, 69 and 38 phenotypes were recapitulated successfully by *i*ACB23LX. c) Confusion matrices of model predictions and Biolog experimental measurements. The overall accuracy of *i*ACB23LX is 86.3% for the carbon (left matrix) and 79.2% for the nitrogen (right matrix) testings. Orange represents correct predictions, and grey represents wrong predictions.

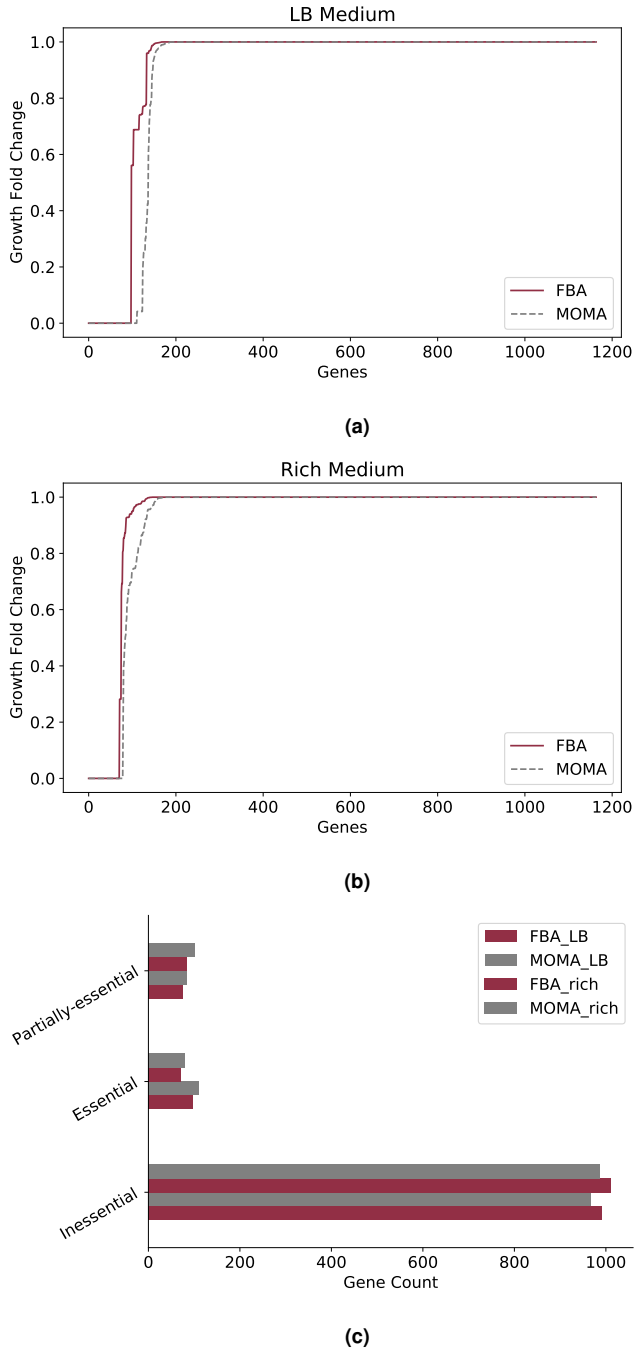

Figure 6 | Distribution of the FC_{gr} values calculated for all genes included in *i*ACB23LX. Red lines represent FBA predictions, while grey is the MOMA-derived growth ratio. Totally 1,164 knockouts were conducted using each method in the **(a)** LB medium and **(b)** rich medium. **(c) Classification of gene essentiality results between LB and rich growth medium.** Both FBA (hatched bars) and MOMA were examined for optimization. The genes were classified as essential, inessential, and partially essential based on their FC_{gr} values.

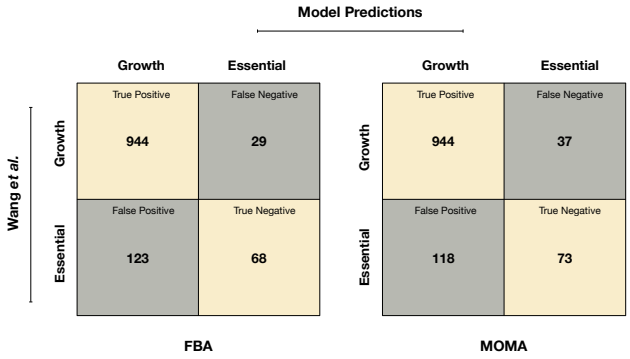

Figure 7 | Accuracy of gene essentiality predictions based on experimental data. *i*ACB23LX was employed to predict essential genes. The *in silico* results were compared the Wang et al. transposon library. The LB medium was applied to mirror the experimental settings. The model exhibited 87% accuracy with FBA (left) and 87% with MOMA (right). Beige indicates correct predictions; grey indicates incorrect predictions.

terpenoids and polyketides. Additionally, we examined in more detail how nutrition availability impacts the gene essentiality by conducting single-gene knockouts in the rich medium. Both optimization methods resulted in more essential genes when the model was required to alter its metabolic behavior due to the absence of nutrients, i.e., with the LB growth medium, compared to the rich medium (Figure 6c and supplementary files S2 and S3). In general, FBA detected more genes to be dispensable for growth in both nutritional environments. On the other hand, MOMA classified more genes as essential or partially-essential (Figure 6c and supplementary files S2 and S3), while genes from FBA build a subset of the essential genes derived by MOMA. Furthermore, we validated the prediction accuracy of *i*ACB23LX using already existing gene essentiality data. At the time of writing, the transposon mutant library by Wang et al. is the only ATCC 17978-specific experimental dataset [49]. With this dataset together and the LB medium, our model demonstrated an accuracy of 87% with both optimization methods (Figure 7). We further analyzed the predicted false negative genes and probed their proteomes to investigate the existence of human orthologs (see *Materials and Methods*). With this, we aimed to eliminate cross-linkings to human-similar proteins since metabolic pathways or enzymes that are missing from the human host have been an important resource of druggable targets against infectious diseases [53]. From the 37 genes that our model predicted to be essential (with FBA and MOMA) contradicting the experimental results, 17 were found to be non-homologous (see Supplementary File S7). Some examples are the genes encoding the enolpyruvylshikimate phosphate (EPSP) synthase (`A1S_2276`), chorismate synthase (`A1S_1694`), riboflavin synthase (`A1S_0223`), phosphogluconate dehydratase (`A1S_0483`), and 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase (`A1S_0484`). The EPSP synthase converts the shikimate-3-phosphate together with phosphoenolpyruvate to 5-O-(1-

carboxyvinyl)-3-phosphoshikimic acid. Subsequently, the chorismate synthase catalyzes the conversion of the 5-O-(1-carboxyvinyl)-3-phosphoshikimic acid to chorismate, the seventh and last step within the shikimate pathway [54]. Chorismate is the common precursor in the production of the aromatic compounds tryptophan, phenylalanine, and tyrosine, as well as folate and menaquinones during the bacterial life cycle. The shikimate pathway is of particular interest due to its absence from the human host metabolome and its vital role in bacterial metabolism and virulence. Moreover, the enzyme riboflavin synthase catalyzes the final step of riboflavin (vitamin B2) biosynthesis with no participating cofactors. Riboflavin can be produced by most microorganisms compared to humans, who have to externally uptake them via food supplements. Also, it plays an important role in the growth of different microbes, especially due to its photosynthesizing property that marks it as a non-invasive and safe therapeutic strategy against bacterial infections [55]. Lastly, the phosphogluconate dehydratase catalyzes the dehydration of 6-Phospho-D-gluconate to KDPG, the precursor of pyruvate and 3-Phospho-D-glycerate [56]. This enzyme is part of the Entner–Doudoroff pathway that catabolizes glucose to pyruvate, similarly to glycolysis, but using a different set of enzymes [57].

We further assessed the druggability of our essential non-homologous proteins and investigated the existence of inhibitors or compounds known to interact with the enzymes. For this, we used the online DrugBank database that contains detailed information on various drugs and drug targets [58]. In all cases, the listed drugs are of unknown pharmacological action, and there is still no evidence indicating the enzymes' association with the molecule's mechanism of action. For instance, the flavin mononucleotide and the cobalt hexamine ion were listed as known inhibitors of yet unknown function against the chorismate synthase, while glyphosate, shikimate-3-Phosphate, and formic acid have been experimentally found to act with EPSP synthase. Six non-homologous genes were marked as hypothetical or putative in the KEGG database and/or lacked enzyme-associated information. We searched for drug leads by aligning the query sequences against the DrugBank's database to find homologous proteins. Two out of six were found to have a protein hit. More specifically, the protein encoded by `A1S_0589` was found to have high sequence identity with the phosphocarrier protein HPr of *Enterococcus faecalis* (Bit-score: 48.5), while the translation product of `A1S_0706` resembles the sugar phosphatase YbiV of *Escherichia coli* (Bit-score: 225.3). According to DrugBank, dexfosfoserine and aspartate beryllium trifluoride have been experimentally determined to bind to these enzymes; however, their pharmacological action is still unknown. The Supplementary Table S7 lists all non-homologous essential genes reported for *i*ACB23LX.

Overall, *i*ACB23LX exhibits high agreement to all validation tests and can, therefore, be used to systematically derive associations between genotypes and phenotypes.

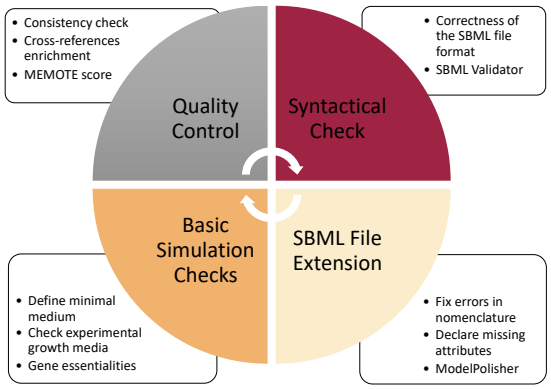

Figure 8 | Debugging workflow to curate and evaluate already published models. Following the community standards, the already published *A. baumannii* models were curated and transformed into re-usable, simulatable, and translatable models. Quality controls and metabolic standardized tests were conducted using `Metabolic Model Testing (MEMOTE)`, while the validity of the file format and syntax were examined with the SBML Validator. ModelPolisher enhanced the models with missing metadata.

A curated collection of already published *A. baumannii* metabolic models

In 2010, Kim et al. publish the first GEM for the multidrug-resistant strain *A. baumannii* AYE [23]. After that, multiple studies provided new data and genomic analyses were published, paving new ways towards its update and refinement [48, 49, 59, 60]. Since then, a variety of GEMs was developed aiming at the empowering of drug development strategies and the enforcement of metabolic engineering by formulating novel and reliable hypotheses (Table 3). However, the amount and format of information contained are inconsistent, with some being syntactically invalid or of older formats. Here, we systematically analyzed the quality of all seven currently existing GEMs, reporting their strengths and weaknesses and debugging them to finally build a curated, standardized, and updated collection. To do so, we developed a workflow with curation steps applicable to all models aiming at the standardization and usability of published GEMs by the community (Figure 8). This closely follows the community-driven workflow published by Carey et al. for the reconstruction of reusable and translatable models [30]. The curation procedure includes a series of stages aiming at modifying data format, data amount, and information quality. It is important to note that no contextual modifications were conducted that could affect the model's prediction capabilities (see *Materials and Methods*).

Five *A. baumannii* strains have been created throughout the years, with AYE and ATCC 19606 having two reconstructions each. All models are publicly stored and can be downloaded either from a database/repository (BioModels, Virtual Metabolic Human (VMH) [61], BiGG [62], and GitHub) or directly from the publication's supplementary material. The use of distinct identifiers prevents the metabolic networks from being compared to each other. More specifically, *i*LP844 and

Table 3 | List of genome-scale metabolic models curated for *A. baumannii*, along with information relevant to the manual refinement. Default growth rates (i.e., model simulated as downloaded), the cellular compartments (C: cytosol, E: extracellular space, P: periplasm, and ER: endoplasmic reticulum), and the reactions and metabolites identifiers are listed in the table. MEMOTE scores before and after manual curation are given in the last column. Blue highlights the novel reconstruction for the strain ATCC 17978 presented in this publication. After manual curation, our model developed following our workflow in Figure 1 has the highest quality score and comes along with a minimal medium defined.

	Availability	Used Identifiers	Growth by default $\mathrm{mmol}/(\mathbf{gDW}\cdot\mathrm{h})$	Compartments	MEMOTE
AbyMBEL891 [23]	BioModels	Customized	119	Cell	20% + 17%
AGORA [25]	VMH	VMH	134	C, E	42% + 37%
***i*LP844** [24]	Suppl. Mat	ModelSeed	15.88	C, E, P	37% + 21%
***i*CN718** [26]	BiGG	BiGG	1.31	C, E, P, ER	70% + 3%
***i*ATCC1906** [27]	Suppl. Mat	KEGG	46.34	C, E	14% + 44%
***i*JS784** [29]	GitHub	ModelSeed	0.0	C, E, P	41% + 18%
***i*AB5075** [28]	Suppl. Mat	BiGG	1.729	C, E, P	17% + 50%
***i*ACB23LX**	BioModels	BiGG	0.1677	C, E, P	89%

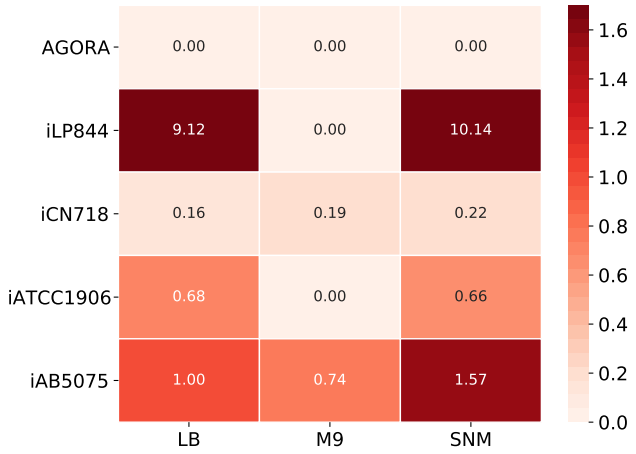

Figure 9 | Model-predicted growth ability in experimentally-defined growth environments. The respective growth rates of *i*ACB23LX are in Table 2.

*i*JS784 carry ModelSEED [63] identifiers for reactions and metabolites, while *i*CN718 and *i*AB5075 BiGG [62] identifiers. AbyMBEL891 uses distinct identifiers not supported by any database, and *i*ATCC1906 includes identifiers derived from KEGG [64].

Most of the models resulted in an unrealistic and inflated growth rate (reference: doubling time of the fastest growing organism *V. natriegens*) in their defined medium, while *i*JS784 showed a zero growth even when all imports were enabled. Hence the model was excluded from further analysis (Table 3). For each of the remaining GEM, we defined the minimum growth requirements that result in a non-zero and realistic objective value. For instance, the AGORA model required at least 21 compounds (mostly metal ions), while oxygen was sufficient for AbyMBEL891 to simulate a non-zero growth (see Supplementary File S4). Since these models should successfully reflect the *A. baumannii* metabolic and growth capabilities, we examined the flux through their biomass reaction in various growth media known to induce *A. baumannii* growth (Figure 9). The majority resulted in a biomass flux of $0.0\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$ with the M9 medium, while the AGORA model could not simulate growth in the LB and SNM as well. Thus, we investigated and identified minimal medium supplementations needed to enable cellular biomass production. As already mentioned, *i*JS784 was excluded from further analysis (Table 3), together with AbyMBEL891 that debilitated the analysis due to its non-standardized identifiers and its missing genes and GPRs. When the medium of *i*ATCC1906 and *i*LP844 was supplemented with D-alanine and glycyl-L-glutamate, respectively, their biomass reactions carried a positive flux rate of $0.4989\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$ and $0.7357\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$. Supplementation of meso-2,6-diaminoheptanedioate, menaquinone-8, niacinamide, heme, siroheme, and spermidine into the LB medium of the AGORA model resulted in a positive growth rate ($1.9430\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$). Similarly, when supplementing the SNM with glycyl-L-asparagine, the derived growth rate was $1.5020\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$, while the M9 medium needed to be extended with glycyl-L-asparagine and thiamine (resulted growth rate: $0.5844\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$). Lastly, like with *i*ACB23LX, the LB medium, together with FBA and MOMA, were applied to detect lethal genes in all models (see Supplementary Files S5 and S6). Despite significant efforts, we could not derive a mapping scheme between the strain-specific gene identifiers of *i*LP844 and *i*ATCC1906 to resolve PROKKA or HMPREF identifiers. Thus, a strain-wise comparison of essential genes could be feasible only for the strain ATCC 17978. As already mentioned, *i*JS784 simulated continuously zero growth and was excluded from the analysis. Consequently, we examined which genes were necessary for growth among the remaining models across three different strains: AYE (*i*CN718), ATCC 17978 (*i*ACB23LX), and AB0057 (AGORA). Totally, 392 genes were identified as essential, while 34 occurred in all three strains. For instance, when the genes encoding for dephospho-coenzyme A (CoA) kinase, phosphopantetheinyl transferase, shikimate

kinase (A1S_3190), or chorismate synthase (A1S_1694) were deleted from the three strains, no growth could be simulated in the LB medium. As already mentioned, the gene encoding the chorismate synthase has no human-like counterpart. This, together with the fact that it was detected to be vital for growth across three distinct strains, increases its potential to be a drug candidate for future therapies. Generally, most essential genes are members of the purine metabolism and encode various transferases. Besides this, the pantothenate and CoA biosynthesis and the amino acid metabolism were found to be a prominent target pathways for further drug development.

Discussion

The historical timeline of past pandemics shows the imposed threat of bacteria in causing repetitive outbreaks with the highest death tolls [65], such as cholera and plague. By 2050, antimicrobial-resistant pathogens are expected to kill 10 million people annually [66], while the antibiotics misuse accompanied by the ongoing Coronavirus Disease 2019 (COVID-19) crisis exacerbated this global threat. It is noteworthy that elevated morbidity rates were ascribed to bacterial co/secondary infections during previous viral disease outbreaks [67, 68, 69]. Hence, developing effective antibiotic regimens is of urgent importance.

Here we present the most recent and comprehensive ready-to-use blueprint GEM for the Gram-negative pathogen *A. baumannii*. For this, we developed a workflow that applies to any living organism and ensures the reconstruction of high-quality models following the community standards. Our model, *i*ACB23LX, was able to simulate growth in the human nose (SNM) and in the experimentally defined media LB and M9. With this we assured that our model recapitulates growth phenotypes in media that reflect *Acinetobacter*-associated environments. We also determined the minimal number of compounds (*i*MinMed) needed to achieve non-zero growth. We called this the *i*MinMed medium and contains totally 14 compounds, including transition metals and a carbon source. Transitions metals have been shown to participate in important biological processes and are vital for the survival of living organisms [41]. Furthermore, we validated *i*ACB23LX quantitatively and qualitatively using existing experimental data and observed remarkable improvements compared to precursory models. More specifically, our model predicted experimental Biolog growth phenotypes on various carbon sources [48] with 86.3% overall agreement, which is higher than the predictions capability of *i*ATCC19606 (84.3%) and *i*LP844 (84%), and comparable to that of *i*AB5075 (86.3%). Similarly, *i*ACB23LX exhibited 79.2% predictive accuracy on nitrogen sources tests, while this increases to 87.5% after further refinement. Improving and re-defining the biomass objective function (BOF) based on accurate strain-specific experimental data would be the next step to diminish the number of inconsistent predictions and to further improve the network and its predictive potential. During gene lethality analysis in LB medium, our model predicted 110 genes with MOMA to be essential, while 97 of them were also reported by FBA to impair the growth. Generally, after enriching the nutritional input with all available compounds (rich medium), less lethal genes resulted, meaning that *A. baumannii* undergoes metabolic alterations when nutrients are lacking. Our *in silico* results compared to the strain-specific gene essentiality data [49] resulted in 87% overall accuracy, which is remarkably higher than all GEMs built for *A. baumannii* (e.g., 80.22% for *i*CN718 and 72% for *i*LP844), except *i*AB5075 which performed comparably. Subsequently, we examined more carefully our false negative predictions and searched for putative drug targets that could be employed for future therapeutics. More specifically, we focused on genes found to be essential for growth and encode proteins with no human counterparts (see Supplementary Table S7). Our study highlighted the EPSP and chorismate synthases from the shikimate pathway as prominent target candidates with no correlation to the human proteome. Several knock-out studies have highlighted the significance of enzymes from the shikimate metabolism as potential targets against infections caused by threatening microorganisms, e.g., *Mycobacterium tuberculosis* [70], *Plasmodium falciparum* [71], and *Yersinia enterocolitica* [72]. Umland et al. identify these two genes products as essential in an *in vivo* study using a clinical isolate of *A. baumannii* and a rat abscess infection model [73]. This increases the confidence of our results and indicates that novel genes found to be essential *in silico* should be considered as potential antimicrobial targets. Similarly, numerous studies have suggested one of our further candidates, riboflavin, as a potential antimicrobial agent [55], while the Entner–Doudoroff pathway (in which our candidate targets phosphogluconate dehydratase and KDPG aldolase act to produce pyruvate) is similar to the glycolysis but with different member enzymes, has been firstly discovered in *Pseudomonas saccharophila* [57] and later in *E. coli* [74]. Meanwhile, it is vital for the survival of further pathogenic microorganisms, like *Neisseria gonorrhoeae*, *Klebsiella pneumoniae*, and *Pseudomonas aeruginosa* [75, 76, 77]. However, they have not yet been examined in the context of *Acinetobacter* species and could be a source of antimicrobial therapeutic strategies. Hence, these biosynthetic routes could be a valuable resource for targets to fight bacterial infectious diseases. Finally, we investigated the druggability of our essential non-homologous genes. We searched the DrugBank database to find compounds known to inhibit these genes and that are already approved by the Food and Drug Administration (FDA). Our analysis resulted in only drugs that have been found to interact with the gene product of interest; however their pharmacological action is yet unknown. We further probed the hypothetical and putative non-homologous genes against the DrugBank's sequence database to find homologous proteins and determine their activity. Also in this case, the resulted drugs were listed with still undetermined pharmacological action. These putative and yet unexplored targets with inhibitory potential are of great interest in the context of developing novel classes of antibiotics.

Overall, our model reached a MEMOTE score of 89%, which is the highest score reported for this organism.

Moreover, we improved and assessed all previously published models and created the first curated collection of metabolic networks for *A. baumannii*. We created a debugging workflow consisting of four major steps to systematically analyze and curate constraint-based models focusing on their standardization and the FAIR data principles. We applied this workflow and curated a total of seven metabolic models for *A. baumannii*. In addition, most of the models simulated growth rates by default that were unrealistic when compared to the fastest growing organism (*V. natriegens*) [36]. Hence, we determined the minimal number of components needed for these models to result in non-inflated biomass production rates. The defined minimal media were mostly composed of metal ions (e.g., cobalt, iron, magnesium) that are essential for bacterial growth. For the model *i*JS784, the minimization process was infeasible; thus, the model was not considered for further analysis. We also examined the growth ability of these models in three media (SNM, LB, and M9) and compared them to our model, *i*ACB23LX. When the models simulated a zero flux through the biomass reaction, we continued by detecting the minimal amount of metabolites supplemented in the medium that resulted in a non-zero growth rate. These would enable the detection of gaps and assist in future improvement of the models. It is important to note here that with this curation, we opted for a systematical assessment of the previously reconstructed models and the detection of their assets and liabilities. Consequently, we did not undertake any contextual modification that could alter the models' predictive capabilities. Finally, we *in silico* detected lethal genes among comparable and simulatable models of *A. baumannii*. Our analysis incorporated three strains of *A. baumannii* (AYE, AB0057, and ATCC 17978), and we examined the effect of genetic variation across strains in the gene essentiality. Our analysis highlighted once again the shikimate pathway, as well as the purine metabolism, the pantothenate, and CoA biosynthesis, and the amino acid metabolism as candidate routes to consider for future new classes of antibacterial drugs with potential effect across multiple *A. baumannii* strains. The curated models, together with our novel model, would benefit the future prediction of candidate lethal genes by reducing the considerable resources needed for classical whole-genome essentiality screenings. All in all, this collection of simulation-ready models will forward the selection of a suitable metabolic network based on individual research questions and help define the entire species and new hypothesis.

Our new metabolic reconstruction and the curated collection of further strain-specific models will guide the formulation of ground-breaking and reliable model-driven hypotheses about this pathogen and help examine the diversity in the metabolic behavior of different *A. baumannii* species in response to genetic and environmental alterations. Additionally, they can be utilized to detect critical pathways related to responses against multiple antibiotic treatments. This will ultimately strengthen the development of advanced precision antimicrobial control strategies against multidrug-resistant (MDR) *A. baumannii* strains.

Taken together, our workflows and models can be employed to expand this collection further with additional standardized strain-specific metabolic reconstructions to finally define the core and pan metabolic capabilities of *A. baumannii*.

Materials and Methods

The metabolic model reconstruction workflow

Figure 1 illustrates the workflow we developed to create the novel high-quality genome-scale metabolic network *i*ACB23LX, following the state-of-the-art protocol of Thiele and Palsson [21]. Our workflow consists of eight major steps starting from the extraction of an annotated genome until the model validation using experimental data. Modifications in the model structure, as well as the inclusion of cross-references to multiple functional databases, were done using the libSBML [37] library, while all simulations were conducted via the COBRApy-0.22.1 [50] suite that includes functions commonly used for simulations.

The individual steps are described below in more detail with respect to the reconstruction of *i*ACB23LX.

Draft reconstruction

A first draft model was built with CarveMe 1.5.1 using the annotated genome sequence of the strain ATCC 17978. This was downloaded from the National Centre for Biotechnology Information (NCBI) at https://www.ncbi.nlm.nih.gov and has the assembly accession number ASM1542v1 [78]. Seven strain-specific assemblies are registered in NCBI; however, the chosen entry is also present in the KEGG database facilitating the model extension. The genome is 3.9 Mio bp long and has two plasmids (pAB1 and pAB2). We set the SBML flavor to activate the SBML Level 3 extension for fbc version 2 that allows semantic descriptions for domain-specific elements such as metabolite chemical formulae and charges together with reaction boundaries and GPRs. Moreover, we the CarveMe parameter gramneg to employ the specialized template for the Gram-negative bacteria. Compared to the Gram-positive template, the Gram-negative template comes with phosphatidylethanolamines, murein, and a lipopolysaccharide unit. Its biomass reactions involve membrane and cell wall components resulting in more accurate gene essentiality predictions in the lipid biosynthesis pathways.

Manual refinement and extension

We started the manual refinement of the draft model by resolving syntactical errors within the model file using SBML Validator from the libSBML library[37]. Missing metabolite charges and chemical formulas were retrieved from the BiGG [62] and ChEBI [79] databases, while mass- and charge-imbalanced reactions were corrected. The most intense part of the workflow is the manual network extension and gap-filling. this was done using the organism-specific databases KEGG [64]

and BioCyc [80], together with ModelSEED [63]. We mapped the new gene locus tags to the old ones using the GenBank General Feature Format (GFF) [81] and added missing metabolic genes along with the respective reactions and metabolites into our model. The network's connectivity was ensured by resolving as many dead-ends (can only be produced but not consumed) and orphan (can only be consumed but not produced) metabolites as possible. Also, reactions with no connectivity were not included in the model, while reactions with no organism-specific gene evidence were removed from the model.

Erroneous energy generating cycles

Energy generating cycles (EGC) are thermodynamically infeasible loops found in metabolic networks and have not been experimentally observed, unlike futile cycles. EGCs charge energy metabolites like adenosine triphosphate (ATP) and uridine triphosphate (UTP) without any external source of nutrients and may result in incorrect and unrealistic energy increases. Their elimination is crucial while correcting the energy metabolism since they can inflate the maximal biomass yields and make the predictions unreliable. We checked their existence in *i*ACB23LX applying an algorithm developed by Fritzemeier et al.[32].

We created a Python script that (1) defines and adds energy dissipation reactions (EDRs) in the network:

$$\mathrm{X[c]} + \mathrm{H_2O[c]} \longrightarrow \mathrm{X^+[c]} + \mathrm{H[c]} + \mathrm{Pi[c]},$$

where X is the metabolite of interest and (2) maximizes each EDR while blocking all influxes. This can be formulated as follows:

$$\max(v_{edr})$$

subject to

$$\begin{aligned} & S \cdot v = 0 \\ & \forall i \notin E : v_i^{\min} \leq v_i \leq v_i^{\max} \\ & \forall i \in E : v_i = 0 \end{aligned} \tag{1}$$

where edr is the index of the current dissipation reaction, S is the stoichiometric matrix, v the flux vector, E the set of all exchange reactions, and $v_i^{\min}$ and $v_i^{\max}$ the upper and lower bounds. The existence of EGCs is indicated by a positive optimal value of v_{edr}.

Totally we examined energy 14 energy metabolites: ATP, cytidine triphosphate (CTP), guanosine triphosphate (GTP), UTP, inosine triphosphate (ITP), nicotinamide adenine dinucleotide (NADH), nicotinamide adenine dinucleotide phosphate (NADPH), flavin adenine dinucleotide (FADH2), flavin mononucleotide (FMNH2), ubiquinol-8, menaquinol-8, demethylmenaquinol-8, acetyl-CoA, and L-glutamate. Moreover, we tested the proton exchange between cytosol and periplasm.

In the case of existing EGCs, we examined the directionality and the gene evidence of all participated reactions using the BioCyc organism-specific information as reference [80].

Database annotations

In this stage, the model was enriched with cross-linkings to various functional databases. Reactions and metabolites were annotated with databases (e.g., KEGG [64], BRENDA [82], and UniProt [83]). These were included in the model as controlled vocabulary (CV) Terms following the Minimal Information Required In the Annotation of Models (MIRIAM) guidelines [84] and the resolution service at `https://identifiers.org/`. We used ModelPolisher [85] to complete the missing available metadata for all metabolites and genes. Similarly, metabolic genes were annotated with their KEGG, NCBI Protein, and RefSeq identifiers using the GFF [81]. To reactions, metabolites, and genes SBO Terms were assigned using the SBOannotator [34]. SBO Terms provide unambiguous semantic information and specify the type or role of the individual model component [86]. In addition, ECO Terms were added to every reaction to capture the type of evidence of biological assertions with `BQB_IS_DESCRIBED_BY` as a biological qualifier. They are useful during quality control and mirror the curator's confidence about the inclusion of a reaction. When multiple genes encode a single reaction, an ECO Term was added for every participant gene. Both terms were incorporated into the model according to our mapping in Figure 3.

Finally, reactions were annotated with the associated subsystems in which they participate using the KEGG database and the biological qualifier `BQB_OCCURS_IN`. Moreover, the "groups" plugin was activated [38]. Every reaction that appeared in a given pathway was added as a `groups:member`, while each pathway was created as a group instance with `sboTerm="SBO:0000633"` and `groups:kind="partonomy"`.

Quality control and quality assurance

`MEMOTE`[36] version 0.13.0 was used to assess and track the quality of our model after each modification, providing us with information regarding the model improvement. The final model was converted into the latest SBML Level 3 Version 2 [38] format using the libSBML package, while the SBML Validator tracked syntactical errors and ensured a valid format of the final model[37].

Constraint-based analysis

The most frequently used constraint-based modeling approach is the FBA that determines a flux distribution via optimization of the objective function and linear programming [51]. Prior to this, the metabolic network is mathematically encoded using the stoichiometric matrix S formalism. This structure delineates the connectivity of the network, and it is formed by the stoichiometric coefficients of all participating biochemical reactions. The rows and columns are represented by the metabolites and the mass- and charge-balanced reactions respectively. At steady state, the system of linear equations derived from the network is defined as follows:

$$\mathbf{S} \cdot \vec{v} = 0 \tag{2}$$

with $\mathbf{S}$ being the stoichiometric matrix and $\vec{v}$ the flux vector. With no defined constraints, the flux distribution may be deter-

mined at any point within the solution space. This space must be further restricted since the system is under-determined and algebraically insoluble. An allowable solution space is defined by a series of imposed constraints that are followed by cellular functions. Altogether the FBA maximization problem, with mass balance, thermodynamic, and capacity constraints, is defined as:

$$\begin{aligned} \text{max or min} \quad & Z = c^T \vec{v} \\ \text{subject to:} \quad & \mathbf{S} \cdot \vec{v} = 0 \\ & 0 \leq v_i \qquad\qquad \forall \text{ irreversible reaction } i \\ & v_{\min} \leq v_i \leq v_{\max} \quad \text{for } i = 1, \ldots, n. \end{aligned} \tag{3}$$

Here, n is the amount of reactions, Z represents the linear objective function, and $\vec{c}$ is a vector of coefficients on the fluxes $\vec{v}$ used to define the objective function.

Growth simulations

Strict aerobic growth check

At the time of writing, the utilized draft reconstruction tool, CarveMe [31], does not include reconstruction templates to differentiate between aerobic and anaerobic species. The directionality of reactions that produce or consume oxygen may affect the model's ability to grow anaerobically. *A. baumannii* is defined to be a strictly aerobic species. Hence, we tested whether our model could grow with no oxygen supplementation. For this purpose, we examined all active oxygen-producing reactions under anaerobic conditions. We corrected their directionality based on the organism-specific information found in BioCyc [80] and kept only those with associated gene evidence.

Defining a minimal growth medium

To determine the minimal number of nutrients needed for the bacterium to grow, we defined a minimal medium using *i*ACB23LX. We determined the minimal amount of metabolites needed for growth using the M9 medium (supplementary file S1) as a reference. We modeled growth on the minimal medium by enabling the uptake of all metabolites that constitute the medium (lower bound of exchange reactions set to $-10\,\text{mmol}/(\text{gDW}\cdot\text{h})$). The lower bound for the rest of the exchanges was set to $0\,\text{mmol}/(\text{gDW}\cdot\text{h})$. The final minimal medium is listed in Table 1 and the supplementary file S1. It consists of nine transition metals, a carbon source, a nitrogen source, a sulfur source, and a phosphate source. The aerobic environment was simulated by setting the lower bound for the oxygen exchange to $-10\,\text{mmol}/(\text{gDW}\cdot\text{h})$.

Growth in chemically defined media

We utilized experimentally verified growth media to examine the growth capabilities of *i*ACB23LX. The M9 minimal medium supplemented with various carbon sources and the LB medium are often used in experimental studies to culture *A. baumannii*.Hence, we tested the ability of our model to grow using both media. Additionally, we inspected the growth of our model in the human nasal niche, as *A. baumannii* have been isolated from nasal samples within ICUs [11, 12, 13]. For this purpose, we utilized the SNM that imitates the human nasal habitat [87]. In all cases, if macromolecules or mixtures were present, we considered the constitutive molecular components for the medium definition. As our model was initially unable to reproduce growth on the applied media, we deployed the gap-filling option from CarveMe to detect missing reactions and gaps in the network [31]. All simulated media are available in S1.

Rich medium definition

To investigate our model's growth rate when all nutrients are available to the bacterial cell, we defined the rich medium. For this purpose, we enabled the uptake of all extracellular metabolites by the model setting the lower bound of their exchange reactions to $-10\,\text{mmol}/(\text{gDW}\cdot\text{h})$.

Model validation

Evaluation of carbon and nitrogen utilization

We employed the previously published Biolog Phenotypic Array data by Farrugia et al. for *A. baumannii* ATCC 17978 to validate the functionality of our model [48]. According to the experimental guidelines provided by Farrugia et al., we utilized the M9 minimal medium for all simulations. The medium was then supplemented with D-xylose as a carbon source for the nitrogen testings, while ammonium served as the only nitrogen source for the carbon tests. As D-xylose was initially not part of the model, we conducted an extensive search in the organism-specific databases KEGG [64] and BioCyc[80] to include missing reactions.

The phenotypes were grouped by their maximal kinetic curve height. A trait was considered positive ("growth") if the height exceeded the 115 and 101 OmniLog units for a nitrogen and carbon source, respectively. The prediction accuracy was evaluated by comparing the *in silico*-derived phenotypes to the Biolog results. More specifically, the overall model's accuracy (ACC) was calculated by the overall agreement:

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where true positive (TP) and true negative (TN) are correct predictions, while false positive (FP) and false negative (FN) are inconsistent predictions. Discrepancies were resolved via iterative manual curation of the model.

Gene perturbation analysis

We performed *in silico* single-gene deletions on *i*ACB23LX to detect essential genes. For this purpose, we utilized the `single_gene_deletion` function from the COBRApy [50] package. A gene is considered to be essential if a flux of $0.0\,\text{mmol}/(\text{gDW}\cdot\text{h})$ was observed through the biomass reaction after setting the lower and upper bounds of the associated reaction(s) to $0.0\,\text{mmol}/(\text{gDW}\cdot\text{h})$.

Additionally, we examined the effect of gene deletions using two different optimization approaches: FBA [51] and

MOMA [52]. Contrary to FBA, MOMA is based on quadratic programming, and the involved optimization problem is the Euclidean distance minimization in flux space. Moreover, it approximates the metabolic phenotype and relaxes the assumption of optimal growth flux for gene deletions [52].

The results were compared to the ATCC 17978-specific gene essentiality dataset from 2014 [49]. Wang et al. generated a random mutagenesis dataset including 15,000 unique transposon mutants using insertion sequencing (INSeq) [49]. By the time of writing, this is the only *A. baumannii* ATCC 17978 library presenting gene essentiality information. Analogously to the experimental settings, the nutrient uptake constraints were set to the LB medium. From the 453 genes identified as essential by Wang et al., 191 could be compared to our predictions. The rest were not part of *i*ACB23LX due to their non-metabolic functions. To measure the effect of a single deletion, we calculated the fold change (FC) between the model's doubling time after (gr_{KO}) and before (gr_{WT}) a single knockout. This is formulated as follows:

$$FC_{gr} = \frac{gr_{KO}}{gr_{WT}} \tag{5}$$

To this end, if $FC_{gr} = 0$, the deleted gene is classified as essential, meaning its removal prevented the network from producing at least one key biomass metabolite predicting no growth. Similarly, if $FC_{gr} = 1$, the deletion of the gene from the network did not affect the growth phenotype (labeled as inessential), while when $0 < FC_{gr} < 1$, the removal of this gene affected partially the biomass production (labeled as partially-essential). The complete lists of the gene essentiality results are available in the supplementary files S2 and S3.

To examine the potential of the *in silico* determined essential genes on becoming novel drug candidates to fight *A. baumannii* infections, we probed the queries of predicted false negative candidates against the human proteome using Basic Local Alignment Search Tool (BLAST) [88]. The protein sequences were aligned to the human protein sequences using the default settings of the NCBI BLASTp tool (word size: 6, matrix: BLOSUM62, gap costs: 11 for existence and 1 for extension). To eliminate adverse effects and ensure no interference with human-like proteins, queries with any non-zero alignment score with the human proteome were not considered. Lastly, we searched in the DrugBank database version 5.1.9 to find inhibitors or ligands known to act with the enzymes encoded by the non-homologous genes [58].

Curation of existing metabolic networks

Previously reconstructed models of *A. baumannii* for multiple strains (Figure 2, Table 3) were collected and curated following community standards and guidelines. For this, we created a workflow shown in Figure 8 that consists of four main steps and utilizes model validation and annotation tools. This can be applied to any metabolic network in SBML format and follows the community "gold standards" strictly as proposed by Carey et al. [30]. The curation steps involved changes in the format, amount, and quality of the included information. The context has not been altered in any way that could impact the models' prediction capabilities. We employed a combination of already existing tools to analyze, simulate, and quality-control the models (COBRApy [50], `MEMOTE` [36], and the SBML Validator [37]). Different database cross-references were incorporated in the models using ModelPolisher [85] and following the MIRIAM guidelines [84], while the libSBML library[37] was used to manipulate the file format and convert to the latest version. To resolve inflated growth rates, we determined computationally-defined minimal growth media. The growth capabilities were examined with respect to various experimentally-derived growth media, while the LB medium was applied to identify lethal genes. A strain-wise comparison was not feasible due to strain-specific identifiers, no successful growth, or missing genes. Hence, we investigated the essential genes across all models with identifiers that could not be mapped with the Pathosystems Resource Integration Center (PATRIC) ID mapping tool [89].

To begin with the debugging, we examined the syntactical correctness and internal consistency of the downloaded files using the SBML Validator from the libSBML library[37]. Two models (*i*CN718 and *i*JS784) could not pass the validator check and reported errors since they were not in a valid SBML format right after their attainment. We made *i*CN718 valid by deleting the reaction `DNADRAIN` for which neither a reactant nor a product was assigned since the associated metabolite was not part of the model. Similarly, the empty `groups` attribute was removed from *i*JS784, converting the file into a valid format. Warnings were detected for *i*ATCC1906, and *i*AB5075 due to missing definition of the fbc extension (became available at the latest Level 3 release [90]) and the non-alphanumeric chemical formulas. We resolved these issues by defining the fbc list `listOfGeneProducts` and the species attribute `chemicalFormula`. In more detail, we extracted the given GPR from the notes field and defined individual `geneProduct` classes with `id`, `name`, and `label`. The attribute `chemicalFormula` was set equal to the species chemical formulas extracted from the notes and is particularly essential in reaction's validation and balancing. Following the SBML specifications regarding its constitution, in case of ambiguous formulas separated by a semicolon (`;`), the first molecular representation was chosen. With this, the genes and metabolites' chemical formulas became part of the file's main structure. Since *i*ATCC1906 carried KEGG identifiers, we could extract the metabolites' chemical formulas from the database and add them to the model. Moving on with the file extension, we declared the remaining missing attributes from reactions, metabolites, and genes that are required according to the SBML language guidelines [38]. More specifically, we defined the `metaid` attribute when missing, while we fixed any errors regarding the identifiers nomenclature. Further extension involved the annotation of reactions, metabolites, and genes with a plethora of database cross-references following the MIRIAM guidelines [84]. For this, we employed Model-

Polisher that complements and annotates SBML models with additional metadata using the BiGG Models knowledgebase as reference [85]. We also defined precise SBO Terms with the `sboTerm` attribute using the SBOannotator [34]. The final step of debugging involved the conversion of all models to the newest available format SBML Level 3 Version 2 [38], as well as the quality control using `MEMOTE` [36].

Data availability

Supplementary tables in Microsoft Excel format are available along with this article. The model *i*ACB23LX along with all curated and refined models are available at the BioModels Database[91] as an SBML Level 3 Version 2 file[38] distributed as Open Modeling EXchange format (OMEX) archive[92].

Author contributions

Conceptualization and idea, N.L.; model reconstruction and analysis, N.L. and Y.X.; curation of models, N.L.; manuscript writing, N.L.; manuscript revision, N.L., Y.X., and A.D.; supervision and funding acquisition, A.D. All authors approved the publishing of the manuscript.

Acknowledgments

This work was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2124 – 390838134 and supported by the Cluster of Excellence 'Controlling Microbes to Fight Infections' (CMFI). A.D. is supported by the German Center for Infection Research (DZIF, doi: 10.13039/100009139) within the *Deutsche Zentren der Gesundheitsforschung* (BMBF-DZG, German Centers for Health Research of the Federal Ministry of Education and Research (BMBF)), grant № 8020708703. The authors acknowledge the support by the Open Access Publishing Fund of the University of Tübingen (`https://uni-tuebingen.de/en/216529`). One figure was created with BioRender.com.

Competing interests

The authors declare no conflict of interest.

List of Abbreviations

AGORA	Assembly of Gut Organisms through Reconstruction and Analysis
AMR	antimicrobial resistance
ATP	adenosine triphosphate
BiGG	Biochemical, Genetical, and Genomical
BLAST	Basic Local Alignment Search Tool
BMBF	Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*)
BMBF-DZG	*Deutsche Zentren der Gesundheitsforschung*
BOF	biomass objective function
CBM	constraint-based modeling
CMFI	Controlling Microbes to Fight Infections
COBRApy	Constraints-Based Reconstruction and Analysis for Python
CoA	coenzyme A
COVID-19	Coronavirus Disease 2019
CTP	cytidine triphosphate
CV	controlled vocabulary
DFG	*Deutsche Forschungsgemeinschaft*
DZIF	German Center for Infection Research
ECO	Evidence and Conclusion Ontology
EDR	energy dissipation reaction
EGC	energy generating cycle
EPSP	enolpyruvylshikimate phosphate
FADH2	flavin adenine dinucleotide
FAIR	Findable, Accessible, Interoperable, and Reusable
FC	fold change
FMNH2	flavin mononucleotide
FBA	flux balance analysis
fbc	flux balance constraints
FDA	Food and Drug Administration
FN	false negative
FP	false positive
GEM	genome-scale metabolic model
GFF	General Feature Format
GPR	gene-protein-reaction associations
GTP	guanosine triphosphate
ICU	intensive care unit
INSeq	insertion sequencing
ITP	inosine triphosphate
KDPG	2-keto-3-deoxy-6-phosphogluconate
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani
MDR	multidrug-resistant
MEMOTE	Metabolic Model Testing
MIRIAM	Minimal Information Required In the Annotation of Models
MOMA	Minimization of Metabolic Adjustment
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NCBI	National Centre for Biotechnology Information
OLS	Ontology Lookup Service
OMEX	Open Modeling EXchange format
PATRIC	Pathosystems Resource Integration Center
PDR	pandrug-resistant
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SBO	Systems Biology Ontology

SNM	synthetic nasal medium
TN	true negative
TP	true positive
UTP	uridine triphosphate
VMH	Virtual Metabolic Human
WHO	World Health Organization
XDR	extensively drug-resistant

References

1 A.-P. Magiorakos, A. Srinivasan, R. B. Carey, Y. Carmeli, M. Falagas, C. Giske, S. Harbarth, J. Hindler, G. Kahlmeter, B. Olsson-Liljequist, et al. "Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance." In: *Clinical microbiology and infection* 18.3 (2012), pp. 268–281. DOI: 10.1111/j.1469-0691.2011.03570.x.

2 C. J. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. R. Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, et al. "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis." In: *The Lancet* 399.10325 (2022), pp. 629–655. DOI: 10.1016/S0140-6736(21)02724-0.

3 R. M. Epand, C. Walker, R. F. Epand, and N. A. Magarvey. "Molecular mechanisms of membrane targeting antibiotics." In: *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1858.5 (2016), pp. 980–987. DOI: 10.1016/j.bbamem.2015.10.018.

4 P. E. Fournier and H. Richet. "The epidemiology and control of Acinetobacter baumannii in health care facilities." In: *Clinical Infectious Diseases* 42 (5 Mar. 2006), pp. 692–699. ISSN: 10584838. DOI: 10.1086/500202/2/42-5-692-TBL003.GIF.

5 L. Dijkshoorn, A. Nemec, and H. Seifert. "An increasing threat in hospitals: multidrug-resistant Acinetobacter baumannii." In: *Nature Reviews Microbiology 2007 5:12* 5 (12 Dec. 2007), pp. 939–951. ISSN: 1740-1534. DOI: 10.1038/nrmicro1789.

6 A. Y. Peleg, H. Seifert, and D. L. Paterson. "Acinetobacter baumannii: Emergence of a Successful Pathogen." In: *Clinical Microbiology Reviews* 21 (3 July 2008), p. 538. ISSN: 08938512. DOI: 10.1128/CMR.00058-07.

7 S. Ibrahim, N. Al-Saryi, I. M. Al-Kadmy, and S. N. Aziz. "Multidrug-resistant Acinetobacter baumannii as an emerging concern in hospitals." In: *Molecular Biology Reports* 48 (10 Oct. 2021), p. 6987. ISSN: 15734978. DOI: 10.1007/S11033-021-06690-6.

8 P. Bogaerts, T. Naas, I. Wybo, C. Bauraing, O. Soetens, D. Piérard, P. Nordmann, and Y. Glupczynski. "Outbreak of infection by carbapenem-resistant Acinetobacter baumannii producing the carbapenemase OXA-58 in Belgium." In: *Journal of clinical microbiology* 44.11 (2006), pp. 4189–4192. DOI: 10.1128/JCM.00796-06.

9 B. Cai, R. Echols, G. Magee, J. C. Arjona Ferreira, G. Morgan, M. Ariyasu, T. Sawada, T. Nagata, et al. "Prevalence of carbapenem-resistant Gram-negative infections in the United States predominated by Acinetobacter baumannii and Pseudomonas aeruginosa." In: *Open forum infectious diseases*. Vol. 4. 3. Oxford University Press. 2017. DOI: 10.1093/ofid/ofx176.

10 A. Iovleva, M. M. Mustapha, M. P. Griffith, L. Komarow, C. Luterbach, D. R. Evans, E. Cober, S. S. Richter, K. Rydell, C. A. Arias, et al. "Carbapenem-Resistant Acinetobacter baumannii in US hospitals: diversification of circulating lineages and antimicrobial resistance." In: *Mbio* 13.2 (2022), e02759–21. DOI: 10.1128/mbio.02759-21.

11 Q. Damaceno, J. R. Nicoli, and A. Oliveira. "Variability of cutaneous and nasal population levels between patients colonized and infected by multidrug-resistant bacteria in two Brazilian intensive care units." In: *SAGE Open Medicine* 3 (2015), p. 2050312114566668. DOI: 10.1177/2050312114566668.

12 M.-L. Liou, K.-H. Chen, H.-L. Yeh, C.-Y. Lai, and C.-H. Chen. "Persistent nasal carriers of Acinetobacter baumannii in long-term-care facilities." In: *American journal of infection control* 45.7 (2017), pp. 723–727. DOI: 10.1016/j.ajic.2017.02.005.

13 C.-H. Chen, M.-L. Liou, C.-Y. Lee, M.-C. Chang, H.-Y. Kuo, and T.-H. Chang. "Diversity of nasal microbiota and its interaction with surface microbiota among residents in healthcare institutes." In: *Scientific reports* 9.1 (2019), pp. 1–10. DOI: 10.1038/s41598-019-42548-5.

14 K. Rangel, T. P. G. Chagas, and S. G. De-Simone. "Acinetobacter baumannii infections in times of COVID-19 pandemic." In: *Pathogens* 10.8 (2021), p. 1006. DOI: 10.3390/pathogens10081006.

15 D. Contou, A. Claudinon, O. Pajot, M. Micaëlo, P. Longuet Flandre, M. Dubert, R. Cally, E. Logre, M. Fraissé, H. Mentec, et al. "Bacterial and viral co-infections in patients with severe SARS-CoV-2 pneumonia admitted to a French ICU." In: *Annals of intensive care* 10.1 (2020), pp. 1–9. DOI: 10.1186/s13613-020-00736-x.

16 W. G. Lima, J. C. M. Brito, and W. S. da Cruz Nizer. "Ventilator-associated pneumonia (VAP) caused by carbapenem-resistant Acinetobacter baumannii in patients with COVID-19: Two problems, one solution?" In: *Medical hypotheses* 144 (2020), p. 110139. DOI: 10.1016/j.mehy.2020.110139.

17 A. Russo, F. Gavaruzzi, G. Ceccarelli, C. Borrazzo, A. Oliva, F. Alessandri, E. Magnanimi, F. Pugliese, and M. Venditti. "Multidrug-resistant Acinetobacter baumannii infections in COVID-19 patients hospitalized in intensive care unit." In: *Infection* 50.1 (2022), pp. 83–92. DOI: 10.1007/s15010-021-01643-4.

18 H. Seifert, L. Dijkshoorn, P. Gerner-Smidt, N. Pelzer, I. Tjernberg, and M. Vaneechoutte. "Distribution of Acinetobacter species on human skin: comparison of phenotypic and genotypic identification methods." In: *Journal of clinical microbiology* 35.11 (1997), pp. 2819–2825. DOI: 10.1128/jcm.35.11.2819-2825.1997.

19 L. Dijkshoorn, E. Van Aken, L. Shunburne, T. Van Der Reijden, A. Bernards, A. Nemec, and K. Towner. "Prevalence of Acinetobacter baumannii and other Acinetobacter spp. in faecal samples from non-hospitalised individuals." In: *Clinical microbiology and infection* 11.4 (2005), pp. 329–332. DOI: 10.1111/j.1469-0691.2005.01093.x.

20 A. Howard, M. O'Donoghue, A. Feeney, and R. D. Sleator. "Acinetobacter baumannii: an emerging opportunistic pathogen." In: *Virulence* 3.3 (2012), pp. 243–250. DOI: 10.4161/viru.19700.

21 I. Thiele and B. Ø. Palsson. "A protocol for generating a high-quality genome-scale metabolic reconstruction." In: *Nature protocols* 5.1 (2010), pp. 93–121. DOI: 10.1038/nprot.2009.203.

22 M. A. Oberhardt, B. Ø. Palsson, and J. A. Papin. "Applications of genome-scale metabolic reconstructions." In: *Molecular systems biology* 5.1 (2009), p. 320. DOI: 10.1038/msb.2009.77.

23 H. U. Kim, T. Y. Kim, and S. Y. Lee. "Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen Acinetobacter baumannii AYE." In: *Molecular BioSystems* 6.2 (2010), pp. 339–348. DOI: 10.1039/b916446d.

24 L. Presta, E. Bosi, L. Mansouri, L. Dijkshoorn, R. Fani, and M. Fondi. "Constraint-based modeling identifies new putative targets to fight colistin-resistant A. baumannii infections." In: *Scientific reports* 7.1 (2017), pp. 1–12. DOI: 10.1038/s41598-017-03416-2.

25 S. Magnúsdóttir, A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jäger, J. Baginska, P. Wilmes, et al. "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota." In: *Nature biotechnology* 35.1 (2017), pp. 81–89. DOI: 10.1038/nbt.3703.

26 C. J. Norsigian, E. Kavvas, Y. Seif, B. O. Palsson, and J. M. Monk. "iCN718, an updated and improved genome-scale metabolic network reconstruction of Acinetobacter baumannii AYE." In: *Frontiers in genetics* 9 (2018), p. 121. DOI: 10.3389/fgene.2018.00121.

27 Y. Zhu, J. Zhao, M. H. M. Maifiah, T. Velkov, F. Schreiber, and J. Li. "Metabolic responses to polymyxin treatment in Acinetobacter baumannii ATCC 19606: integrating transcriptomics and metabolomics with genome-scale metabolic modeling." In: *Msystems* 4.1 (2019), e00157–18. DOI: 10.1128/mSystems.00157-18.

28 J. Zhao, Y. Zhu, J. Han, Y.-W. Lin, M. Aichem, J. Wang, K. Chen, T. Velkov, F. Schreiber, and J. Li. "Genome-scale metabolic modeling reveals metabolic alterations of multidrug-resistant acinetobacter Baumannii in a murine bloodstream infection model." In: *Microorganisms* 8.11 (2020), p. 1793. DOI: 10.3390/microorganisms8111793.

29 J. S. C. Barbosa. "Genome-scale reconstruction of the metabolic network iJS784 for acinetobacter baumannii strain ATCC 17978 to address drug target prioritization." dissertation. National Laboratory for Scientific Computing, Brazil, 2020.

30 M. A. Carey, A. Dräger, M. E. Beber, J. A. Papin, and J. T. Yurkovich. "Community standards to facilitate development and address challenges in metabolic modeling." In: *Molecular Systems Biology* 16.8 (2020), e9235. DOI: 10.15252/msb.20199235.

31 D. Machado, S. Andrejev, M. Tramontano, and K. R. Patil. "Fast automated reconstruction of genome-scale metabolic models for microbial species and communities." In: *Nucleic acids research* 46.15 (2018), pp. 7542–7553. DOI: 10.1093/nar/gky537.

32 C. J. Fritzemeier, D. Hartleb, B. Szappanos, B. Papp, and M. J. Lercher. "Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal." In: *PLoS computational biology* 13.4 (2017), e1005494. DOI: 10.1371/journal.pcbi.1005494.

33 S. Jupp, T. Burdett, C. Leroy, and H. E. Parkinson. "A new Ontology Lookup Service at EMBL-EBI." In: *SWAT4LS* 2 (2015), pp. 118–119. DOI: 10.1093/nar/gkz943.

34 N. Leonidou, E. Fritze, A. Renz, and A. Dräger. "SBOannotator: a Python Tool for the Automated Assignment of Systems Biology Ontology Terms." In: *Bioinformatics* (July 2023), btad437. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btad437.

35 yWorks GmbH. "yEd." Version 3.19. In: (Mar. 19, 2019).

36 C. Lieven, M. E. Beber, B. G. Olivier, F. T. Bergmann, M. Ataman, P. Babaei, J. A. Bartell, L. M. Blank, S. Chauhan, K. Correia, et al. "MEMOTE for standardized genome-scale metabolic model testing." In: *Nature biotechnology* 38.3 (2020), pp. 272–276. DOI: 10.1038/s41587-020-0446-y.

37 B. J. Bornstein, S. M. Keating, A. Jouraku, and M. Hucka. "LibSBML: an API library for SBML." In: *Bioinformatics* 24.6 (2008), pp. 880–881. DOI: 10.1093/bioinformatics/btn051.

38 M. Hucka, F. T. Bergmann, C. Chaouiya, A. Dräger, S. Hoops, S. M. Keating, M. König, N. Le Novère, C. J. Myers, B. G. Olivier, S. Sahle, J. C. Schaff, R. Sheriff, L. P. Smith, D. Waltemath, D. J. Wilkinson, and F. Zhang. "Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 2 Core Release 2." In: *Journal of Integrative Bioinformatics* 16.2 (May 2019), p. 1. ISSN: 1613-4516. DOI: 10.1515/jib-2019-0021.

39 Z. A. King, A. Dräger, A. Ebrahim, N. Sonnenschein, N. E. Lewis, and B. O. Palsson. "Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways." In: *PLoS computational biology* 11.8 (2015), e1004321. DOI: 10.1371/journal.pcbi.1004321.

40 S. Mishra and J. Imlay. "Why do bacteria use so many enzymes to scavenge hydrogen peroxide?" In: *Archives of biochemistry and biophysics* 525.2 (2012), pp. 145–160. DOI: 10.1016/j.abb.2012.04.014.

41 M. I. Hood and E. P. Skaar. "Nutritional immunity: transition metals at the pathogen–host interface." In: *Nature Reviews Microbiology* 10.8 (2012), pp. 525–537. DOI: 10.1038/nrmicro2836.

42 B. L. Mortensen and E. P. Skaar. "The contribution of nutrient metal acquisition and metabolism to Acinetobacter baumannii survival within the host." In: *Frontiers in cellular and infection microbiology* 3 (2013), p. 95. DOI: 10.3389/fcimb.2013.00095.

43 M. I. Hood, B. L. Mortensen, J. L. Moore, Y. Zhang, T. E. Kehl-Fie, N. Sugitani, W. J. Chazin, R. M. Caprioli, and E. P. Skaar. "Identification of an Acinetobacter baumannii zinc acquisition system that facilitates resistance to calprotectin-mediated zinc sequestration." In: *PLoS pathogens* 8.12 (2012), e1003068. DOI: 10.1371/journal.ppat.1003068.

44 K. Raman, P. Rajagopalan, and N. Chandra. "Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs." In: *PLoS computational biology* 1.5 (2005), e46. DOI: 10.1371/journal.pcbi.0010046.

45 S. Ma, K. J. Minch, T. R. Rustad, S. Hobbs, S.-L. Zhou, D. R. Sherman, and N. D. Price. "Integrated modeling of gene regulatory and metabolic networks in Mycobacterium tuberculosis." In: *PLoS computational biology* 11.11 (2015), e1004543. DOI: 10.1371/journal.pcbi.1004543.

46 D.-S. Lee, H. Burd, J. Liu, E. Almaas, O. Wiest, A.-L. Barabási, Z. N. Oltvai, and V. Kapatral. "Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple Staphylococcus aureus genomes identify novel antimicrobial drug targets." In: *Journal of bacteriology* 191.12 (2009), pp. 4015–4024. DOI: 10.1128/JB.01743-08.

47 Y. Seif, J. M. Monk, N. Mih, H. Tsunemoto, S. Poudel, C. Zuniga, J. Broddrick, K. Zengler, and B. O. Palsson. "A computational knowledge-base elucidates the response of Staphylococcus aureus to different media types." In: *PLoS computational biology* 15.1 (2019), e1006644. DOI: 10.1371/journal.pcbi.1006644.

48 D. N. Farrugia, L. D. Elbourne, K. A. Hassan, B. A. Eijkelkamp, S. G. Tetu, M. H. Brown, B. S. Shah, A. Y. Peleg, B. C. Mabbutt, and I. T. Paulsen. "The complete genome and phenome of a community-acquired Acinetobacter baumannii." In: *PloS one* 8.3 (2013), e58628. DOI: 0.1371/journal.pone.0058628.

49 N. Wang, E. A. Ozer, M. J. Mandel, and A. R. Hauser. "Genome-wide identification of Acinetobacter baumannii genes necessary for persistence in the lung." In: *MBio* 5.3 (2014), e01163–14. DOI: 10.1128/mBio.01163-14.

50 A. Ebrahim, J. A. Lerman, B. O. Palsson, and D. R. Hyduke. "COBRApy: constraints-based reconstruction and analysis for python." In: *BMC systems biology* 7.1 (2013), pp. 1–6. DOI: 10.1186/1752-0509-7-74.

51 J. D. Orth, I. Thiele, and B. Ø. Palsson. "What is flux balance analysis?" In: *Nature biotechnology* 28.3 (2010), pp. 245–248. DOI: 10.1038/nbt.1614.

52 D. Segre, D. Vitkup, and G. M. Church. "Analysis of optimality in natural and perturbed metabolic networks." In: *Proceedings of the National Academy of Sciences* 99.23 (2002), pp. 15112–15117. DOI: 10.1073/pnas.232349399.

53 D. Barh, S. Tiwari, N. Jain, A. Ali, A. R. Santos, A. N. Misra, V. Azevedo, and A. Kumar. "In silico subtractive genomics for target identification in human bacterial pathogens." In: *Drug Development Research* 72.2 (2011), pp. 162–177. DOI: 10.1002/ddr.20413.

54 R. Bentley and E. Haslam. "The shikimate pathway—a metabolic tree with many branche." In: *Critical reviews in biochemistry and molecular biology* 25.5 (1990), pp. 307–384. DOI: 10.3109/10409239009090615.

55 N. Farah, V. K. Chin, P. P. Chong, W. F. Lim, C. W. Lim, R. Basir, S. K. Chang, and T. Y. Lee. "Riboflavin as a promising antimicrobial agent? A multi-perspective review." In: *Current Research in Microbial Sciences* (2022), p. 100111. DOI: 10.1016/j.crmicr.2022.100111.

56 H. Meloche and W. Wood. "The mechanism of 6-phosphogluconic dehydrase." In: *Journal of Biological Chemistry* 239.10 (1964), pp. 3505–3510. DOI: 10.1016/S0021-9258(18)97751-3.

57 N. Entner and M. Doudoroff. "Glucose and gluconic acid oxidation of Pseudomonas saccharophila." In: *J Biol Chem* 196.2 (1952), pp. 853–862. DOI: 10.1016/S0021-9258(19)52415-2.

58 D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, et al. "DrugBank 5.0: a major update to the DrugBank database for 2018." In: *Nucleic acids research* 46.D1 (2018), pp. D1074–D1082. DOI: 10.1093/nar/gkx1037.

59 C. W. Dorsey, A. P. Tomaras, and L. A. Actis. "Genetic and phenotypic analysis of Acinetobacter baumannii insertion derivatives generated with a transposome system." In: *Applied and Environmental Microbiology* 68.12 (2002), pp. 6353–6360. DOI: 10.1128/AEM.68.12.6353-6360.2002.

60 L. A. Gallagher, E. Ramage, E. J. Weiss, M. Radey, H. S. Hayden, K. G. Held, H. K. Huse, D. V. Zurawski, M. J. Brittnacher, and C. Manoil. "Resources for genetic and genomic analysis of emerging pathogen Acinetobacter baumannii." In: *Journal of bacteriology* 197.12 (2015), pp. 2027–2035. DOI: 10.1128/JB.00131-15.

61 A. Noronha, J. Modamio, Y. Jarosz, E. Guerard, N. Sompairac, G. Preciat, A. D. Daníelsdóttir, M. Krecke, D. Merten, H. S. Haraldsdóttir, et al. "The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease." In: *Nucleic acids research* 47.D1 (2019), pp. D614–D624. DOI: 10.1093/nar/gky992.

62 C. J. Norsigian, N. Pusarla, J. L. McConn, J. T. Yurkovich, A. Dräger, B. O. Palsson, and Z. King. "BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree." In: *Nucleic Acids Research* 48.D1 (Nov. 2019). DOI: 10.1093/nar/gkz1054.

63 C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay, and R. L. Stevens. "High-throughput generation, optimization and analysis of genome-scale metabolic models." In: *Nature biotechnology* 28.9 (2010), pp. 977–982. DOI: 10.1038/nbt.1672.

64 M. Kanehisa and S. Goto. "KEGG: kyoto encyclopedia of genes and genomes." In: *Nucleic acids research* 28.1 (2000), pp. 27–30. DOI: 10.1093/nar/28.1.27.

65 J. Piret and G. Boivin. "Pandemics throughout history." In: *Frontiers in microbiology* 11 (2021), p. 631736. DOI: 10.3389/fmicb.2020.631736.

66 J. O'Neill et al. "Review on antimicrobial resistance." In: *Antimicrobial resistance: tackling a crisis for the health and wealth of nations* 4 (2014).

67 Y.-W. Chien, K. P. Klugman, and D. M. Morens. "Bacterial pathogens and death during the 1918 influenza pandemic." In: *New England Journal of Medicine* 361.26 (2009), pp. 2582–2583. DOI: 10.1056/NEJMc0908216.

68 Z.-M. Sheng, D. S. Chertow, X. Ambroggio, S. McCall, R. M. Przygodzki, R. E. Cunningham, O. A. Maximova, J. C. Kash, D. M. Morens, and J. K. Taubenberger. "Autopsy series of 68 cases dying before and during the 1918 influenza pandemic peak." In: *Proceedings of the National Academy of Sciences* 108.39 (2011), pp. 16416–16421. DOI: 10.1073/pnas.1111179108.

69 D. E. Morris, D. W. Cleary, and S. C. Clarke. "Secondary bacterial infections associated with influenza pandemics." In: *Frontiers in microbiology* 8 (2017), p. 1041. DOI: 10.3389/fmicb.2017.01041.

70 T. Parish and N. G. Stoker. "The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis." In: *Microbiology* 148.10 (2002), pp. 3069–3077. DOI: 10.1099/00221287-148-10-3069.

71 S. Tapas, A. Kumar, S. Dhindwal, P. Kumar, et al. "Structural analysis of chorismate synthase from Plasmodium falciparum: a novel target for antimalaria drug discovery." In: *International journal of biological macromolecules* 49.4 (2011), pp. 767–777. DOI: 10.1016/j.ijbiomac.2011.07.011.

72 F. Bowe, P. O'Gaora, D. Maskell, M. Cafferkey, and G. Dougan. "Virulence, persistence, and immunogenicity of Yersinia enterocolitica O: 8 aroA mutants." In: *Infection and immunity* 57.10 (1989), pp. 3234–3236. DOI: 10.1128/iai.57.10.3234-3236.1989.

73 T. C. Umland, L. W. Schultz, U. MacDonald, J. M. Beanan, R. Olson, and T. A. Russo. "In vivo-validated essential genes identified in Acinetobacter baumannii by using human ascites overlap poorly with essential genes detected on laboratory media." In: *MBio* 3.4 (2012), e00113–12. DOI: 10.1128/mBio.00113-12.

74 R. C. Eisenberg and W. J. Dobrogosz. "Gluconate metabolism in Escherichia coli." In: *Journal of Bacteriology* 93.3 (1967), pp. 941–949. DOI: 10.1128/jb.93.3.941-949.1967.

75 B. E. Britigan, Y. Chai, and M. Cohen. "Effects of human serum on the growth and metabolism of Neisseria gonorrhoeae: an alternative view of serum." In: *Infection and immunity* 50.3 (1985), pp. 738–744. DOI: 10.1128/iai.50.3.738-744.1985.

76 R. Hommes, P. Postma, D. Tempest, and O. Neijssel. "The influence of the culture pH value on the direct glucose oxidative pathway in Klebsiella pneumoniae NCTC 418." In: *Archives of microbiology* 151.3 (1989), pp. 261–267. DOI: 10.1007/BF00413140.

77 T. Lessie and P. Phibbs Jr. "Alternative pathways of carbohydrate utilization in pseudomonads." In: *Annual review of microbiology* 38.1 (1984), pp. 359–388. DOI: 10.1146/annurev.mi.38.100184.002043.

78 M. G. Smith, T. A. Gianoulis, S. Pukatzki, J. J. Mekalanos, L. N. Ornston, M. Gerstein, and M. Snyder. "New insights into Acinetobacter baumannii pathogenesis revealed by high-density pyrosequencing and transposon mutagenesis." In: *Genes & development* 21.5 (2007), pp. 601–614. DOI: 10.1101/gad.1510307.

79 J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes, and C. Steinbeck. "ChEBI in 2016: Improved services and an expanding collection of metabolites." In: *Nucleic acids research* 44.D1 (2016), pp. D1214–D1219. DOI: 10.1093/nar/gkv1031.

80 P. D. Karp, R. Billington, R. Caspi, C. A. Fulcher, M. Latendresse, A. Kothari, I. M. Keseler, M. Krummenacker, P. E. Midford, Q. Ong, et al. "The BioCyc collection of microbial genomes and metabolic pathways." In: *Briefings in bioinformatics* 20.4 (2019), pp. 1085–1093. DOI: 10.1093/bib/bbx085.

81 K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. "GenBank." In: *Nucleic acids research* 44.D1 (2016), pp. D67–D72.

82 A. Chang, L. Jeske, S. Ulbrich, J. Hofmann, J. Koblitz, I. Schomburg, M. Neumann-Schaal, D. Jahn, and D. Schomburg. "BRENDA, the ELIXIR core data resource in 2021: new developments and updates." In: *Nucleic Acids Research* 49.D1 (2021), pp. D498–D508. DOI: 10.1093/nar/gkaa1025.

83 "UniProt: the universal protein knowledgebase in 2021." In: *Nucleic acids research* 49.D1 (2021), pp. D480–D489. DOI: 10.1093/nar/gkaa1100.

84 N. Juty, N. Le Novere, and C. Laibe. "Identifiers. org and MIRIAM Registry: community resources to provide persistent identification." In: *Nucleic acids research* 40.D1 (2012), pp. D580–D586. DOI: 10.1093/nar/gkr1097.

85 M. Römer, J. Eichner, A. Dräger, C. Wrzodek, F. Wrzodek, and A. Zell. "ZBIT bioinformatics toolbox: a web-platform for systems biology and expression data analysis." In: *PloS one* 11.2 (2016), e0149263. DOI: 10.1371/journal.pone.0149263.

86 M. Courtot, N. Juty, C. Knüpfer, D. Waltemath, A. Zhukova, A. Dräger, M. Dumontier, A. Finney, M. Golebiewski, J. Hastings, et al. "Controlled vocabularies and semantics in systems biology." In: *Molecular systems biology* 7.1 (2011), p. 543. DOI: 10.1038/msb.2011.77.

87 B. Krismer, M. Liebeke, D. Janek, M. Nega, M. Rautenberg, G. Hornig, C. Unger, C. Weidenmaier, M. Lalk, and A. Peschel. "Nutrient limitation governs Staphylococcus aureus metabolism and niche adaptation in the human nose." In: *PLoS pathogens* 10.1 (2014), e1003862. DOI: 10.1371/journal.ppat.1003862.

88 S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. "Basic local alignment search tool." In: *Journal of molecular biology* 215.3 (1990), pp. 403–410. DOI: Basiclocalalignmentsearchtool.

89 J. J. Davis, A. R. Wattam, R. K. Aziz, T. Brettin, R. Butler, R. M. Butler, P. Chlenski, N. Conrad, A. Dickerman, E. M. Dietrich, et al. "The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities." In: *Nucleic acids research* 48.D1 (2020), pp. D606–D612. DOI: 10.1093/nar/gkz943.

90 B. G. Olivier and F. T. Bergmann. "SBML level 3 package: flux balance constraints version 2." In: *Journal of integrative bioinformatics* 15.1 (2018). DOI: 10.1515/jib-2017-0082.

91 R. S. Malik-Sheriff, M. Glont, T. V. N. Nguyen, K. Tiwari, M. G. Roberts, A. Xavier, M. T. Vu, J. Men, M. Maire, S. Kananathan, E. L. Fairbanks, J. P. Meyer, C. Arankalle, T. M. Varusai, V. Knight-Schrijver, L. Li, C. Dueñas-Roca, G. Dass, S. M. Keating, Y. M. Park, N. Buso, N. Rodriguez, M. Hucka, and H. Hermjakob. "BioModels—15 years of sharing computational models in life science." In: *Nucleic Acids Research* 48 (2020), pp. D407–D415. DOI: 10.1093/nar/gkz1055.

92 F. T. Bergmann, R. Adams, S. Moodie, J. Cooper, M. Glont, M. Golebiewski, M. Hucka, C. Laibe, A. K. Miller, D. P. Nickerson, B. G. Olivier, N. Rodriguez, H. M. Sauro, M. Scharm, S. Soiland-Reyes, D. Waltemath, F. Yvon, and N. Le Novère. "COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project." In: *BMC Bioinformatics* 15 (Dec. 2014), p. 369. ISSN: 1471-2105. DOI: 10.1186/s12859-014-0369-z.