New Workflow Predicts Drug Targets against SARS-CoV-2 via Metabolic Changes in Infected Cells

Nantia Leonidou [1,2,3,*] (iD), **Alina Renz** [1,2,3] (iD), **Reihaneh Mostolizadeh** [1,2,4] (iD) **and Andreas Dräger** [1,2,3,4] (iD)

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

Abstract: COVID-19 is one of the deadliest respiratory diseases, and its emergence caught the pharmaceutical industry off guard. While vaccines have been rapidly developed, treatment options for infected people remain scarce, and COVID-19 poses a substantial global threat. This study presents a novel workflow to predict robust druggable targets against emerging RNA viruses using metabolic networks and information of the viral structure and its genome sequence. For this purpose, we implemented pymCADRE and PREDICATE to create tissue-specific metabolic models, construct viral biomass functions and predict host-based antiviral targets from more than one genome. We observed that pymCADRE reduces the computational time of flux variability analysis for internal optimizations. We applied these tools to create a new metabolic network of primary bronchial epithelial cells infected with SARS-CoV-2 and identified enzymatic reactions with inhibitory effects. The most promising reported targets were from the purine metabolism, while targeting the pyrimidine and carbohydrate metabolisms seemed to be promising approaches to enhance viral inhibition. Finally, we computationally tested the robustness of our targets in all known variants of concern, verifying our targets' inhibitory effects. Since laboratory tests are time-consuming and involve complex readouts to track processes, our workflow focuses on metabolic fluxes within infected cells and is applicable for rapid hypothesis-driven identification of potentially exploitable antivirals concerning various viruses and host cell types.

Keywords: host-virus interactions; tissue-specific model; COVID-19; SARS-CoV-2; antiviral targets; flux balance analysis; flux variability analysis; reaction knockout; host-derived enforcement; metabolic modeling; virus mutations; software engineering; Python

---

Author summary

The recently emerged human coronavirus SARS-CoV-2 spread worldwide, causing severe challenges in health care, the economy, and society. Developing new vaccines and therapies is essential to prevent the next pandemic efficiently. However, vaccines have the disadvantage of decreased immunity over time, while they lose their efficacy against subsequent mutations and variants. Hence, effective pandemic preparedness requires discovering broadly acting antivirals with high resistance barriers. Here, we accelerate the detection of antiviral drug candidates against RNA viruses by analyzing metabolic changes in infected cells. Viruses rely on the host to acquire all macromolecules needed for their replication, re-programming the cellular metabolism according to their needs. For this reason, host-directed approaches are of great importance. We develop software to reconstruct constraint-based models, simulate infections of a cell, and identify host-based metabolic pathways that can be inhibited to suppress viral replication. We identify promising targets with inhibitory effects

across multiple variants facilitating further *in vitro* and *in vivo* experiments. Our workflow can be applied to any RNA virus and aims to rapidly identify antiviral targets to better prepare for the next pandemic.

Introduction

In a study published in October, 2007, scientists studying coronaviruses characterized the situation in China as a ticking "time bomb" for a potential virus outbreak[**?**]. They had three strong indications to worry: the animal-related eating habits in southern China, the previous appearance of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV)-like viruses in horseshoe bats, and the ability of coronaviruses to undergo recombination. Since the first major pandemic of the new millennium in 2002, over 4000 publications on coronaviruses became available, giving insights and leading to the discovery of 36 SARS-related coronaviruses in humans and animals. Eighteen years later, the whole world experiences the realization of this prophecy with the emergence of the Coronavirus Disease 2019 (COVID-19) to be one of the deadliest respiratory disease pandemics since the "Spanish" influenza in 1918[1]. Scientists globally try to understand the host's immunopathological response, how the novel virus Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) adapts, and how it spreads.

Viruses, being infectious agents, replicate only within the cells of a living organism and re-program them to form other virus particles and accelerate their own reproduction. Their life cycle is divided into four main steps: host cell attachment, penetration, reproduction within the host cell (uncoating, gene expression, replication, and assembly), and release[2]. To increase their mass production, they consume energy from the host cell. This dependency is proved by experimental findings showing considerable metabolic flux alterations in host cells upon infection[3]. To this end, engineering the host metabolism to govern viral infections is of great interest. In fact, one of the largest classes of small-molecule antiviral drugs, the nucleoside and nucleotide analogs, target metabolic enzymes in the nucleotide synthesis resulting in a nucleotide pool imbalance[4]. Examples of such analogs that are already used against RNA viruses are ribavirin[5], acyclovir[6], and remdesivir[7]. Systems-level analysis of gene knock-outs upon bacterial infection with bacteriophage lambda also revealed metabolic genes that, when knocked-out, prevented the phage from replication[8], confirming the engineering of host metabolism as a virus growth regulator.

These laboratory findings highlight the impact of viral biosynthesis on host metabolism and the importance of metabolic alterations in the virus growth minimization. Hence, finding a suitable Virus Biomass Objective Function (VBOF) that reflects the functions of the infected cell is of immense interest. The VBOF is a pseudo-reaction simulating the production of the different virus particles and is analogous to the biomass function used for the metabolic models of prokaryotes and eukaryotes. It consists of energy metabolites, nucleotides, and amino acids, essential for the replication and production of genetic material and proteins. In 2018, Aller *et al.* present a computational approach to create viral objective functions and predicted critical host reactions of the human macrophages against epidemic viruses, like the Zika virus[9]. The applicability of their method was verified by recovering antecedent antiviral targets and predicting new ones.

Notwithstanding the recent therapeutic advances and the approval of multiple vaccines, COVID-19 remains a substantial global health threat. Currently, great efforts are initiated to detect effective antiviral treatments for this pathogenic agent. Like all viruses, SARS-CoV-2 continuously evolves over time as modifications in its genome occur during replication. Such alterations are typical for viruses that encode their genome in RNA, as enzymes that copy the ribonucleic acid are prone to making errors leading to the presence of copying mistakes during viral replication[10]. It has been reported that SARS-CoV-2, along with all coronaviruses, has relatively low mutation rates ($\sim 10^{-6}$ per site per cycle) compared to other RNA viruses, like the Human Immunodeficiency Virus (HIV)-1 or influenza viruses[11,12]. This is ascribed to the presence of proofreading and error-correcting enzymes that recognize and repair copying mistakes hindering the development of anti-CoV drugs and

vaccines[13]. SARS-CoV-2 encodes an Exonuclease (ExoN) in the Non-structural Protein 14 (NSP14), which participates in the genome proofreading mechanism and results in low mutation rates (or high viral fidelity)[14]. The 5' region of the SARS-CoV-2 genome encodes for two open reading frames (ORF1a/ORF1ab and ORF1b) which include 16 Non-structural Proteins (NSPs)[15]. These are followed by four structural proteins: nucleocapsid (N), envelope (E), the spike (S) and the membrane (M), and nine accessory proteins (NS)[15].

At the time of writing, five variants of SARS-CoV-2 have been designated as Variants of Concern (VOC) by the World Health Organization (WHO). These are the Alpha (14th December, 2020, United Kingdom (UK), lineage B.1.1.7), Beta (18th December, 2020, South Africa, lineage B.1.351), Gamma (2nd January, 2021, Brazil, lineage P.1), Delta (24th March, 2021, India, lineage B.1.617), and Omicron (24th November, 2021, South Africa/Botswana, lineage B.1.1.529) variants[16]. These differ from the conventional virus in terms of their pathogen properties (e.g., transferability, virulence, or susceptibility to the immune response of recovered or vaccinated people). Mutations on the structural proteins occur most frequently and issue complications en route to pathogenesis. The most common mutation of the S protein is the non-synonymous replacement of aspartate by glycine (D614G), which is found to decrease the virus effectivity[17]. Mutations in the E protein have not been reported in any variants, except the Beta and Omicron. These are the substitution of proline by leucine (P71L)[18], and the exchange of the hydrophilic threonine by the hydrophobic isoleucine (T9I)[19].

Identifying potential targets and druggable compounds is of vast concern, and one way to detect them is by analyzing metabolic changes in infected cells. This can be achieved with the help of systems biology and the reconstruction of cell-specific Genome-scale Metabolic Models (GEMs) that recapitulate the metabolism of particular cell types[20]. Targeting the host metabolism has already been suggested as a prospective novel antiviral approach, given the relevance of metabolism in virus infection[21]. Since the emergence of SARS-CoV-2 and within a year several studies have been published trying to identify antiviral targets using constraint-based metabolic modeling and utilizing various approaches and resources[22–27]. For instance, a recent study by Bannerman *et al.* employs a draft model of the airway epithelial cells built from Recon1[30], refines it using Recon3D[?], and predicts drug targets against SARS-CoV-2 [26]. However, they used pre-existing reconstruction tools and models to obtain a representation of the tissue metabolism.

In 2012 Wang *et al.* publish the metabolic Context-specificity Assessed by Deterministic Reaction Evaluation (mCADRE) algorithm to construct metabolic models based on human gene expression data and network topology information[28]. This tool is implemented in MATLAB[29], and its functionality is based on the first version of the human model, namely Recon1[30]. This resulted in its limited usability in the last few years since MATLAB is a commercial and closed-source software.

Here, we present pymCADRE, a re-implementation of mCADRE in Python striving for a more accessible and updated version of the reconstruction tool. Additionally, we implemented scripts for data pre-processing facilitating relevant curation tasks, such as assigning confidence scores to reactions, binarizing raw transcriptomic data, and calculating gene ubiquity scores. Pathological studies already pointed out that SARS-CoV-2 targets the airways and the lungs. The entry and infectivity of enveloped viruses are strongly regulated by proteolytic cleavage of the viral envelope glycoproteins[31]. In the case of SARS-CoV-2, the S protein, when bound in the cell surface, is susceptible to airway protease cleavage, which results in conformational change favoring the entry of the virus into human bronchial epithelial cells[31]. Further single-cell analyses provided insights into the virus replication and the cell tropism, confirming that infection with SARS-CoV-2 is also localized in the bronchial epithelial cells [32,33]. Hence, we applied pymCADRE to create a novel tissue-specific model of primary Human Bronchial Epithelial Cells (HBECs) based on the already available human metabolic network, Recon1. We updated the model by including a biomass maintenance function that Bordbar *et al.* published in 2010[34].

We subsequently infected this model *in silico* with the novel SARS-CoV-2 virus by constructing a viral biomass reaction derived from its structural information. Therefore, we created a fully automated

computational tool in Python, called Prediction of Antiviral Targets (PREDICATE), which applies the stoichiometric approaches introduced by Aller *et al.* on a metabolic network, constructs a single VBOF, and creates an integrated host-virus model[9]. Subsequently, our tool predicts exploitable cellular metabolic pathways that can be inhibited to suppress virus replication with minimal or no effect on the cell. This is attained using two approaches: the host-derived enforcement (HDE)[9] and single-reaction knock-outs. We applied our automated script to our tissue-specific model Recon1-HBEC and detected potential host-based targets for future COVID-19 therapeutic strategies. We further used PREDICATE and validated the robustness of our predicted targets against all five variants of concern. We underline the identified metabolic reactions as experimentally exploitable drug targets for suppressing SARS-CoV-2 replication in human bronchial epithelial cells. We syntactically validated our model and compared it against the corresponding model reconstructed using mCADRE.

Altogether, our novel workflow can be summarized in a four-step process, as shown in Figure 1, which is fully transferable to any existing RNA virus and any host cell. With this, we aim to support further the development of effective therapies against emerging viruses and their mutations and create a library of drugs to design broad-spectrum antiviral therapies as an essential resource for pandemic preparedness.

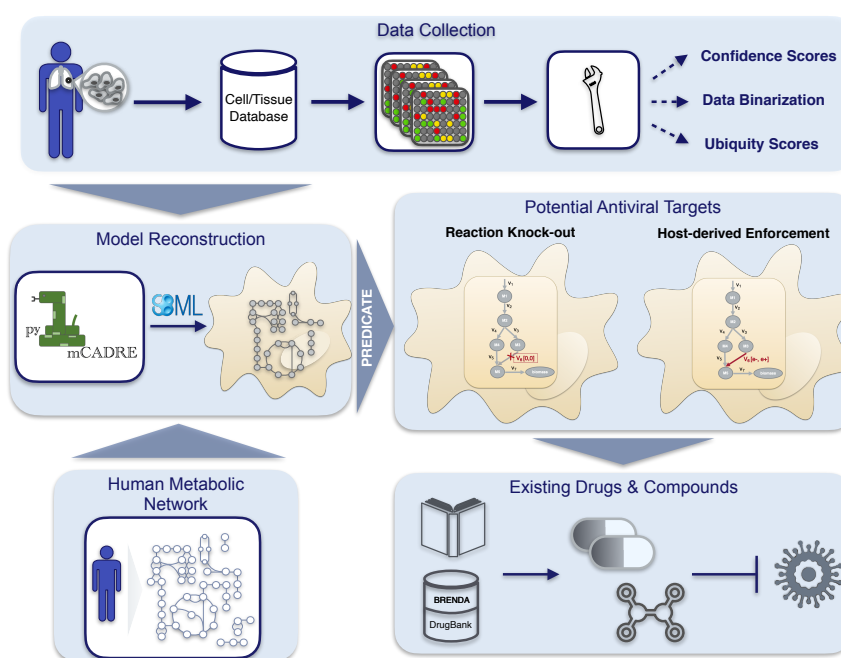



Figure 1. Workflow overview to reconstruct integrated host-virus genome-scale models and detect promising compounds with an antiviral activity. After collecting and curating the required data (the gene expression data and the human metabolic network), pymCADRE reconstructs a tissue-specific model using information from the network topology. The reconstructed metabolic network is then infected *in silico* with the virus of interest and is used to detect promising antiviral targets in an automated process. Detailed description of the process and the respective algorithm, called PREDICATE, is provided in Figure 7. Reaction knock-outs and the host-derived enforcement are used to detect exploitable enzymatic targets that keep the host maintenance at 100 %, while suppressing the virus replication. The resulted top hits are further inspected manually in terms of already existing drugs and compounds in different databases, such as BRENDA[35] and DrugBank[36].

Results

Tissue-specific reconstruction using pymCADRE

The pymCADRE tool was developed to reconstruct tissue-specific metabolic models based on human gene expression data and topological information from the metabolic network. Like mCADRE, pymCADRE leverages gene expression microarray data, literature-derived evidence, and information from the network topology to build context-specific metabolic models. More accurately, it uses a fully automated way to determine core reactions by setting a threshold to expression-based evidence. Therefore, reactions with scores above this threshold are characterized as core reactions, while the rest constitute the non-core set (more details in the Material and Methods section). To test the functionality of pymCADRE and increase its ability to create multiple models of human cells, specifically related to the current outbreak of SARS-CoV-2, we applied pymCADRE to a microarray expression profile dataset of the primary HBEC. Prior to reconstruction, we incorporated a Biomass Objective Function (BOF) to the first version of the human metabolic network, Recon1[30], and used it as a generic host human model.

The objective function originates from the human alveolar macrophage model published by Bordbar *et al.* in 2010 (supplementary file S1 Table)[34]. We updated the resulting model by adding subsystems to all the missing metabolic reactions from Recon1. A subsystem-wise classification in Figure 2 indicates that most reactions in the final Recon1-HBEC model belong to the class of transport reactions, while the biosynthesis of other secondary metabolites is the least represented subsystem. Moreover, in Recon1, there is no growth medium defined, and all extracellular transport reactions are open, i.e., lower fluxes equal $-1000.0\,\mathrm{mmol\,gDW^{-1}\,h^{-1}}$. Hence, we used the already-defined blood medium[37], and we computationally specified a minimal growth medium using the Constraints-Based Reconstruction and Analysis for Python (COBRApy)[38] package. The exact medium compositions are provided in the supplementary file S5 Table.

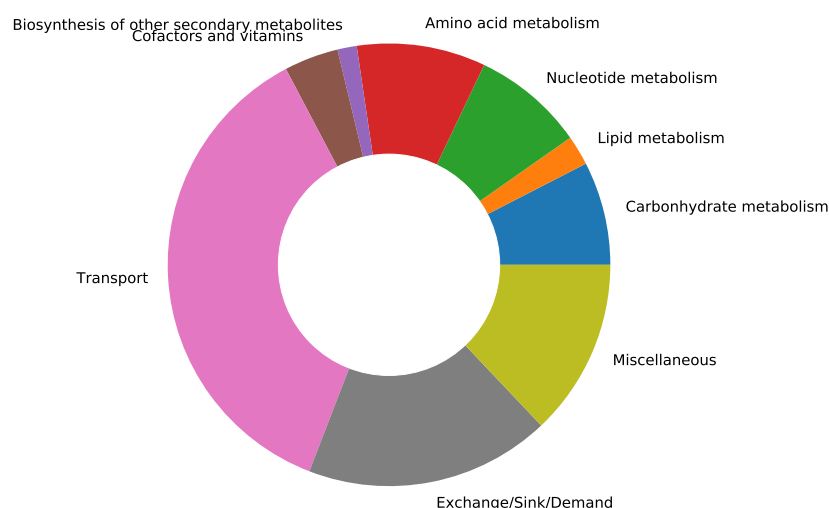



Figure 2. Subsystem-wise classification of all reactions included in Recon1-HBEC. The reaction pathways are merged based on metabolic pathways and according to Kyoto Encyclopedia of Genes and Genomes (KEGG) [39]. The biomass reaction was assigned to "Miscellaneous." The majority of reactions in the final Recon1-HBEC model are transport reactions, while the least amount of reactions is assigned to the biosynthesis of other secondary metabolites subsystem.

The new integrated tissue-specific model Recon1-HBEC contains 1973 reactions, 1442 metabolites, and 1381 genes (Table 1). Almost 70 % of all reactions is associated to a gene-protein-reaction rule (1391; 1086 metabolic and 305 transport reactions), while 391 metabolic and 264 transport

reactions are not related to any gene. To perform internal consistency checks, the user can choose between two COBRApy's[38] tailored optimization functions, Flux Variability Analysis (FVA) [40] and FASTCC[41]. Both methods detect blocked reactions and deliver consistent networks by resolving linear programming problems. Hence, the final pruned model does not contain any blocked reactions. We observed that pymCADRE reduces the pruning time while maintaining the highest possible accuracy compared to the model created with mCADRE (Table 1). With a 3.3 GHz processor and 16 GB Random-access Memory (RAM) on a local computer, mCADRE with FVA demanded ∼6 CPU-hours, while pymCADRE ∼5 CPU-hours. Totally 1272 blocked reactions were eliminated from Recon1 during the consistency check. Furthermore, 498 reactions (9 core and 489 non-core) were inactive in the cell type of interest and removed from the generic model during pruning. Inconsistencies were encountered in the performance of FASTCC as implemented in COBRApy. After multiple runs, the function detected a variable number of blocked reactions. This affected the final pruned model, which differed from the ground truth. However, internal optimizations with FASTCC were executed faster compared to FVA. Duplicating the available RAM can reduce the computational time of the pymCADRE twofold.

Table 1. Analysis results of the HBEC-specific reconstructions using FVA for internal optimizations. The reaction overlap between both models is over 99.5 %.

	Pruned Model			Removed Reactions	
	Reactions	Metabolites	Genes	Cores	Non-cores
mCADRE	1,977	1,442	1905	9	487
pymCADRE	1,973	1,442	1,381	9	489

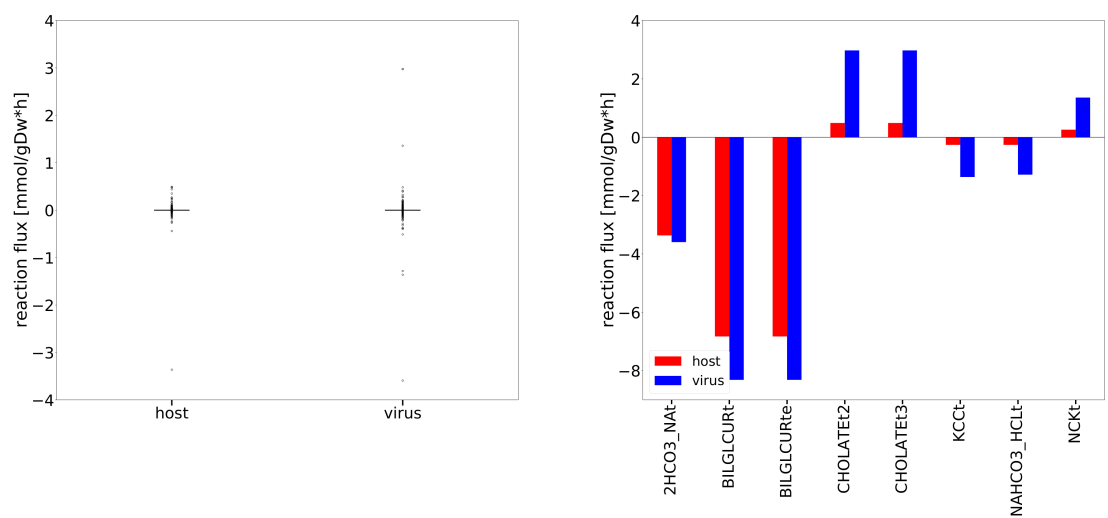
We further extended the models by adding missing exchange reactions to all extracellular metabolites (71 in the mCADRE and 73 in the pymCADRE model). The final reconstructions shared over 2040 reactions, meaning an overlap of over 99.5 % of all reactions in each model. Hence, we have a considerable convergence between the tools, indicating the high quality of models generated with pymCADRE. Table 2 lists the symmetric difference between both models.

Table 2. Symmetric difference of reactions in the models created by mCADRE and pymCADRE.

mCADRE	
ARTPLM1	R group to palmitate conversion
ARTPLM2	R group to palmitate conversion
PE_HStm	Phosphatidylethanolamine scramblase
RETFA	Retinol acyltransferase

pymCADRE	
Htx	Peroxisomal transport of hydrogen
LRAT	Lecithin retinol acyltransferase

Additional analysis using Flux Balance Analysis (FBA) allowed us to study the flux dispersion between the host and virus and conclude which reactions are vital for both host maintenance and virus growth. Explanatory Data Analysis (EDA) showed that non-zero fluxes are mostly fluctuating above zero (Figure 3a). Totally 12 numerically distant values (outliers) were observed (Figure 3b). Inspection of the flux distribution vector showed higher viral flux through transporters of essential metabolites, like K^+ and Na^+, and Adenosine 5'-triphosphate (ATP)-binding cassette (ABC)

transporter. Furthermore, the bicarbonate transporter (`2HCO3_NAt`) and the bilirubin beta-diglucuronide transporters (`BILDGLCURt` and `BILDGLCURte`) are used remarkably more by the host and virus to maintain its optimal state, compared to the virus.

(a) The flux distributions were computed based on a five-number summary (S7 Table). Remarkable outliers with a flux value greater than $1.0\,\text{mmol}/(\text{gDW}\cdot\text{h})$ or less than $-1.0\,\text{mmol}/(\text{gDW}\cdot\text{h})$ were investigated separately (b).

(b) The fluxes through `2HCO3_NAt`, `BILDGLCURt`, and `BILDGLCURte` are remarkably higher when the model is optimized for both the host and the virus. Overall, all displayed reactions are essential for host maintenance and virus growth.

Figure 3. Flux dispersion among host and virus in the Recon1-HBEC model. Distribution of host and virus fluxes as derived from FBA.

Similar to mCADRE, pymCADRE encompasses functionality tests to ensure the fulfillment of the resulting models' basic cellular metabolic capabilities. These tests include the production of various metabolites, such as amino acids and compounds from the Tricarboxylic Acid Cycle (TCA) when the uptake of glucose is enabled[28]. Additional to this, we tested our model for internal cycles that result in erroneous energy production by testing the production of different energy metabolites when no nutrients are available.[42] Our final model did not include any futile cycles since none of the metabolites could be generated.

The new tissue-specific model created with pymCADRE was converted into SBML Level 3 Version 2 [43] format using the Systems Biology Format Converter (SBFC) [44] and passed the syntactical validation using libSBML[45]. Additionally, the Metabolic Model Testing (MEMOTE) suite Version 0.11.1 was used to assess the GEM quality[46]. MEMOTE reports for a given GEM an independent and comparable score along with a comprehensive overview. This test reported a score of 70 % for our integrated model, which indicates a well-annotated model of high quality. Metabolic networks of the same or different tissue possess lower quality scores. For instance, the integrated model of macrophages has a MEMOTE score of 44 %[22], while the model of airway epithelial cells from Bannerman *et al.* has a score of 51 %[26]. While these models contain a wide range of reactions and metabolites annotations, the mass and charge imbalances are still high resulting in lower scores. Nevertheless, they contain over 1,800 blocked reactions, while the integrated macrophage-virus model contains over 140 dead-end and orphan metabolites. The automatically reconstructed models of bronchial and airway epithelial cells from Wang *et al.* have a lower MEMOTE score of 19 %[28]. It is mainly attributed to lacking database cross-references and missing Systems Biology Ontology (SBO)[47] terms.

To examine whether pymCADRE functions as expected, we implemented test scripts, which are available at https://github.com/draeger-lab/pymCADRE/.

Since we purposed to use the model to detect possible anti-SARS-CoV-2 targets, we also included a VBOF that imitates the production of virus particles from its different constituents. Following

the pipeline developed by Aller *et al.* and extended by Renz *et al.*, we created this pseudo-reaction and used it to infect the new model (Recon1-HBEC) *in silico*. The human bronchial epithelial cell's biomass maintenance function (BOF) encompasses amino acids, DNA and RNA nucleotides, and compounds vital for energy supply, and other macromolecules like fatty acids and phospholipids. Similarly, the VBOF contains amino acids, RNA, lipids, and energy-related compounds (S1 Table, S4 Fig.), as well necessary lipids. Analysis of both functions highlights leucine as the most-used amino acid (highest stoichiometric coefficient) in the SARS-CoV-2 growth and the maintenance of the host bronchial cells, while both host and virus utilize only a few tryptophan (Figure 4). Moreover, the same amount of asparagine and phenylalanine is required for the maintenance of the host cell, while the virus needs less phenylalanine. Similar pattern was observed for tyrosine and histidine. Using FBA, optimization of the Recon1-HBEC for the host resulted in a flux for the biomass maintenance function of 0.2344 mmol gDW^{-1} h^{-1}, while optimizing the SARS-CoV-2 growth function resulted in a flux of 0.1575 mmol gDW^{-1} h^{-1}.

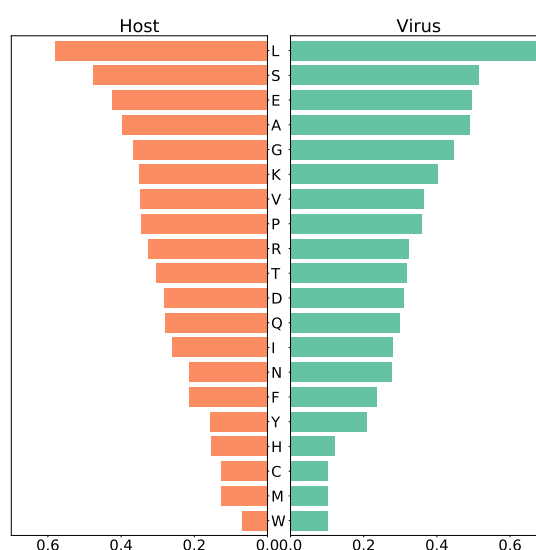
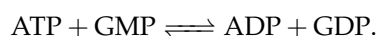



Figure 4. Amino acid usage between host and virus based on the stoichiometric coefficients. The two panels show the amino acid composition of the host maintenance function (left) and the virus biomass (right). The amino acids are annotated using the one-letter code (S6 Table). Both host and virus use mostly leucine (L) for their maintenance/growth, while tryptophan (W) is needed at least. The same amount of asparagine (N) and phenylalanine (F) is required for the maintenance of the host cell, while the virus needs less phenylalanine. Similar pattern can be observed for tyrosine (Y) and histidine (H).

Stoichiometric modeling of the integrated host-virus model predicts targets against SARS-CoV-2

To analyze the host-virus interactions from a metabolic point of view, we created an integrated stoichiometric model of human bronchial epithelial cells infected with SARS-CoV-2. We then used our model to detect host-based reactions, which, when constrained, reduce the virus production the most. According to Aller *et al.*, this analysis can be computationally implemented through systematically setting individual lower and upper bounds to zero (i.e., reaction knock-outs). Applying this approach, we identified a single target enzyme, which if knocked-out, completely inhibits the virus while keeping the host maintenance at 100 % of its initial growth rate. This enzyme is called Guanylate Kinase 1 (GK1, EC-Number: 2.7.4.8, KEGG Reaction ID: R00332) and catalyzes the conversion of ATP and Guanosine 5'-monophosphate (GMP) to Adenosine 5'-diphosphate (ADP) and Guanosine 5'-diphosphate (GDP) (KEGG Reaction ID: R00332):

$$ATP + GMP \rightleftharpoons ADP + GDP.$$

To ensure the maintenance of the metabolic network in a host-optimized state while suppressing the viral growth, we applied the HDE (see Materials and Methods)[9,48]. We constrained all reaction fluxes to ranges obtained from FVA, allowing the attainment of host-optimal state and suppressing the virus production at most. This approach verified the enzymatic target GK1 and revealed further possible compounds that could inhibit the viral production without harming the host cell. The most promising novel hit was the CTP synthase 1 (CTP synthase 1 (CTPS1)) from the *de novo* pyrimidine synthesis pathway that, when constrained, inhibited the virus growth by 62 % with no effect on the host's maintenance (100 % of the initial rate). CTPS1 catalyzes the formation of Cytidine 5'-triphosphate (CTP) from Uridine 5'-triphosphate (UTP). It is important to note here that when the activity of CTPS1 is constrained, and therefore the formation of CTP, host cells can use alternative routes through the salvage pathway using Cytidine 5'-diphosphate (CDP) and/or Cytidine 5'-monophosphate (CMP) to restore the CTP levels directly. Similar results were observed for GK1 with adapted bounds. Further 33 enzymatic targets with inhibitory effects on the virus production were reported using the HDE approach. These concerned, for instance, restricting the extracellular exchange of L-proline and phosphate (EX_pro__L_e and EX_pi_e) and constraining the enzymatic activity in the metabolism of purines (PUNP4, IMPD, GMPS2, NDPK8m, and DGNSKm) and pyrimidines (UMPK5, NDPK2, and DTMPK). Moreover, inhibiting the functionality of enzymes in carbohydrate metabolism, more specifically in the amino/nucleotide sugar metabolism and sucrose metabolism (e.g., ACGAMK, UAGDP, and PGMT) as well as in glycolysis/gluconeogenesis (GAPD, PGK, HEX1, FBA2, PGM, and ENO) led to a remarkable decrease in the viral production by 50 % to 58 % of the initial growth. Figure 5 illustrates all antiviral targets predicted using HDE against the percentage of the remaining virus growth after constraining the reaction bounds. Detailed information about all reactions is included in the supplementary material S2 Table.

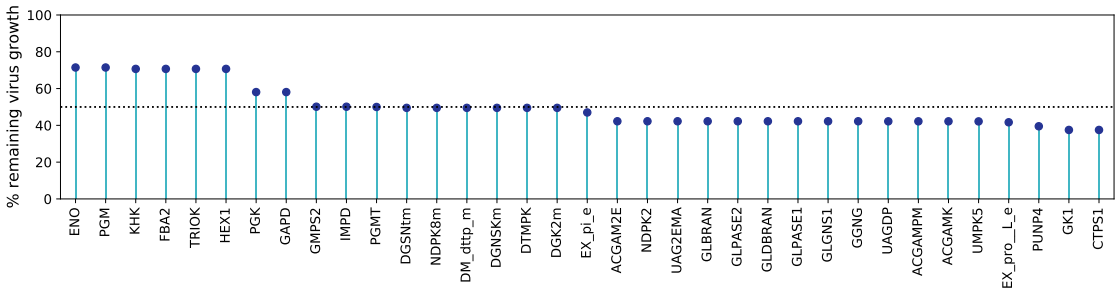



Figure 5. Enzymatic targets of SARS-CoV-2 from the HDE experiments applied to the Recon1-HBEC model. Potential antiviral targets were reported when the virus rate of growth with shifted bounds was beneath the threshold of 90 % of its initial growth rate. Enzymes with adapted bounds from the purine and pyrimidine metabolism led to a remarkable virus decrease, while further promising targets were reported from the carbohydrate metabolism. The dashed line represents the 50 % of the virus remaining.

The GK1 has been recently identified as an essential component for viral propagation and a potential target for antiviral therapies against SARS-CoV-2 in the human alveolar macrophage model[22] and further cell lines[23,27]. Renz *et al.* showed that GK1 could decrease the virus production up to 50 % without damaging the macrophages' maintenance (100 %)[22]. Our host-derived enforcement on the bronchial epithelial cells also reported GK1 as a potential anti-SARS-CoV-2 target, however, with a similar impact on the virus production compared to CTPS1. Similarly, ENO, PGK, and PGM have been predicted as targets to inhibit the production of SARS-CoV-2 by Delattre *et al.*

Interestingly, the already identified robust target GK1 is closely interconnected with two reported promising targets in the purine metabolism (Figure 6). IMPD catalyses the NAD$^+$-dependent oxidation of Inosine 5'-monophosphate (IMP) to Xanthosine 5'-phosphate (XMP) that is subsequently used by GMPS2 to generate GMP. From this, we suggest that focusing on the purine metabolism, and more

specifically on the action of one of these enzymes to inhibit SARS-CoV-2 is a well-established approach that needs to be validated *in vitro* and in cell culture experiments.

Metabolic fluxes are highly affected by the nutrients' availability. Since our approaches mainly focus on studying the metabolic changes in infected cells, fluxes play a major role in the simulation outcomes. So far, we have focused on a chemically defined medium simulating the human blood [37]. Additionally, we examined the virus inhibition that our targets could reach using a minimal growth medium computed with linear optimization[38]. A novel enzymatic target was reported from the pyrimidine metabolism with the minimal medium defined, named NDPK3. Like NDPK2, UMPK5, and GK1, it belongs to the class of phosphotransferases with a phosphate group as an acceptor and was highlighted as a hit target from the single-reaction deletions and the HDE. NDPK3 constrained resulted in a decrease of 44.6 % of the virus growth, which is comparable to the effect of GK1 using the minimal medium. However, NDPK2 and UMPK5 with adapted upper and lower fluxes led to higher viral reduction and appeared as the best hits. The HDE-derived metabolic targets using the minimal medium are shown in S1 Figure and the medium definition is provided in the supplementary file S5 Table.

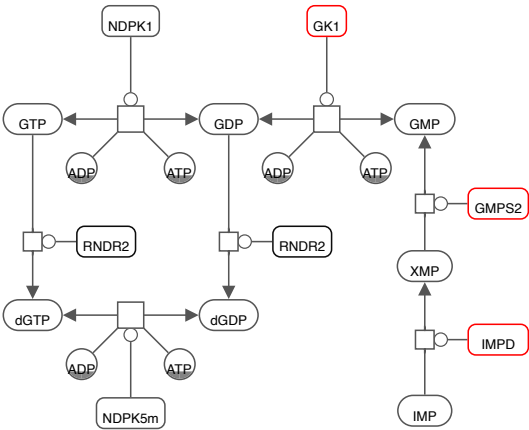



Figure 6. Graphical illustration of the interconnection between promising targets reported from the purine metabolism (red colored) using the Systems Biology Graphical Notation (SBGN)[49]. To annotate reactions and metabolites, Biochemcial, Genetical, and Genomical (BiGG) Identifiers were utilized. Biological map created with Newt[50].

Altogether, we created a fully automated computer tool, which simulates the virus growth in target cells with the help of metabolic networks. Subsequently, our tool applies the above-mentioned host-dependent approaches, HDE, and reaction knock-outs, and predicts enzymatic targets with high inhibitory potency against the virus.

Predicted targets are robust against all known variants of concern

Novel mutations of RNA viruses emerge daily, and as of February, 2022, five SARS-CoV-2 variants have prevailed and spread since its emergence in 2019. These are the Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.717.2), and Omicron (B.1.1.529) variants[16] and have been marked as VOC. Since the beginning of the COVID-19 pandemic, there has been an exponential growth in the number of stored genome sequences within large databases. The WHO asked all scientists around the world to upload their data on the Global Initiative on Sharing All Influenza Data (GISAID) database and help accelerate the response against health threats to humankind[51]. In January, 2020, the GISAID's EpiCoV database launched, becoming the most popular repository for SARS-CoV-2 as it gathers over eight million viral sequences by February, 2022. To examine the variants' effect on the predicted metabolic targets, sequences for all VOC were downloaded from GISAID and investigated further.

We reconstructed a SARS-CoV-2 VBOF using the same approaches as with the reference (wildtype) sequence for each retrieved mutated sequence. We reconstructed 100 individualized biomass functions

and tested each to detect enzymes that inhibit the virus's growth while keeping the host maintenance at maximum. To speed up the reconstruction and analysis processes, we developed an automated script to analyze more than one sequence simultaneously (Figure 7). Additionally, we implemented an algorithm to modify reference protein sequences and introduce amino acid mutations (replacements, insertions, deletions, and duplications) and named this tool PREDICATE. Since RNA viruses are composed of similar building blocks, nucleotides, and proteins, our pipeline can be applied to any single- or double-stranded RNA virus that could infect any cell or tissue type.

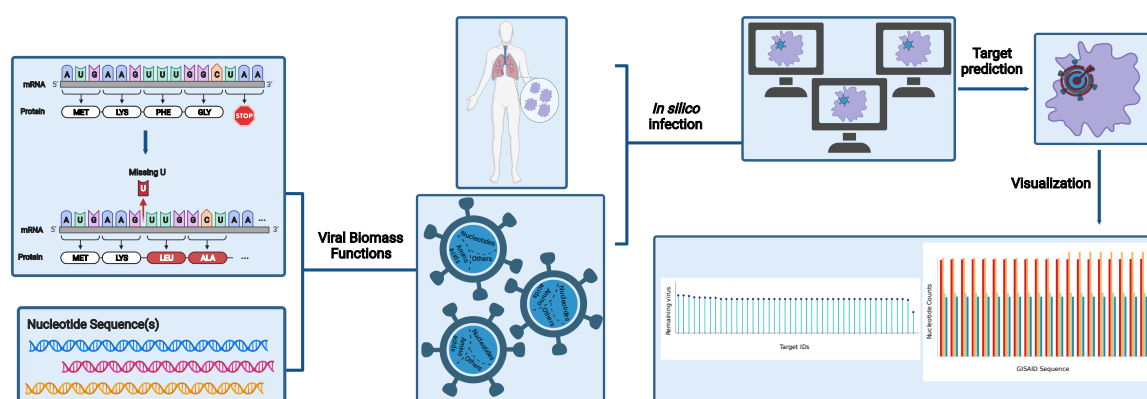



Figure 7. Overview of PREDICATE developed to create viral biomass reactions and predict host-based antiviral targets using host-virus models. First of all, our algorithm, PREDICATE, modifies the amino acids in the protein sequence according to the defined mutations. The mutated protein sequence and the nucleotide sequences are further employed to calculate the stoichiometric coefficients for the virus biomass functions. Reaction knock-outs and the host-derived enforcement are applied to reveal enzymatic reactions that suppress the viral replication. The final step generates various plots, providing insights into the dataset and a better understanding of the results. This pipeline can be applied to either one or more nucleotide sequences and all existing RNA viruses. This makes it particularly advantageous and time-saving when studying multiple variants of a single virus. The number of genomic input sequences equals the number of the calculated VBOF. The Materials and Methods section describes the implemented approaches to predict antiviral targets. Figure created with BioRender (BioRender.com).

To evaluate the mutations' effect on the viral biomass, we calculated the mean of all estimated coefficients across all mutated sequences and compared them against the wildtype (WT) stoichiometries. We did this by looking at the variant-wise differences to the WT. Figure 8 shows how much the variant-wise calculated mean of coefficients deviates from the stoichiometries calculated for the reference sequence. We observed a remarkable increase in the stoichiometric coefficients of ATP and ADP between the Omicron variant and the WT. This pattern is mainly distinct to the ATP and ADP but is observed for the majority of the stoichiometric coefficients. We analyzed the mathematical calculations that led to the stoichiometric coefficients to explain this further. All coefficients depend on the total viral molar mass M_v, which is derived from the sum of the mass of the genome (G_i) and proteome (G_j)[9]. The randomly downloaded genomic sequences of the Omicron variant contained a higher amount of NNN stretches (i.e., nucleotides that could not be determined via sequencing) compared to the other variants. Consequently, the Omicron variant has a decreased count of nucleotides (G_i) and amino acids (G_j), thus a lower total molar mass M_v. Moreover, the overall moles of energy (A^{TOT}) needed to assemble the structural and non-structural proteins strongly influences the stoichiometric coefficient of ATP[9]. The A^{TOT} is related to the total amino acids counts (X_j). Overall, the Alpha variant included more deletions in the spike (H69del, Y144del, and V70del) and NSP6 (F108del, G107del, and S106del) proteins compared to the Omicron variant. The same holds for the Beta, Gamma, and Delta variants. This affected the X_j, which was higher for the Omicron

variant. Altogether, a decreased total viral molar mass and a higher total amino acids count resulted in the apparent rise of the ATP and ADP stoichiometric coefficients for the Omicron variant. Accordingly, the absolute differences between the WT and the Omicron variant are higher than the rest.

When looking at the differences between the amino acids and the WT stoichiometric coefficients, a noticeable increase in the Omicron Variant can be observed for lysine. For this reason, we inspected the respective amino acid mutations. In more than half of the Omicron-related genomic sequences, other amino acids were more often replaced by lysine (Spike N440K, Spike N764K, Spike N856K, Spike N969K, Spike T478K, Spike N679K, Spike T547K, and N R203K). In contrast, the substitution of lysine by other amino acids is rarely occurred (NSP3 K38R and Spike K417N). This also affected the stoichiometric coefficient of asparagine. As most of these mutations emerge in the spike protein, which has the highest copy number, their impact on the amino acid count, and consequently, the stoichiometric coefficient, is considerable. Among the variants for which a higher coefficient was computed for asparagine than the WT, the greatest increase was observed for the Delta variant. This could be justified by the presence of mutations, in which mostly an amino acid is being replaced by asparagine (Spike D950N, M S197N, NSP16 H186N, NSP3 K1693N, NSP8 K37N, and NSP3 K902N). A substitution of asparagine by another amino acid occurs only in three mutation types. Thus, overall there is an increase in the total amount of asparagine, and therefore, in the stoichiometric coefficient for the Delta variant. Lastly, the Omicron variant needs the least glutamine (-0.013) and the most lysine (0.015) compared to the WT.

To verify the validity of our calculations, we searched in the literature to find evidence about the amino acid composition of the different variants. For instance, we observed higher stoichiometric coefficients of charged and hydrophobic residues in the Omicron variant compared to the Delta. Recently, computational analyses indicated in the Omicron variant an increased amount of arginine, lysine, aspartate, and glutamate that contribute to the formation of salt bridges[52]. The same study pointed out the accumulation of the hydrophobic residues, phenylalanine and isoleucine, in the spike protein of same variant[52].

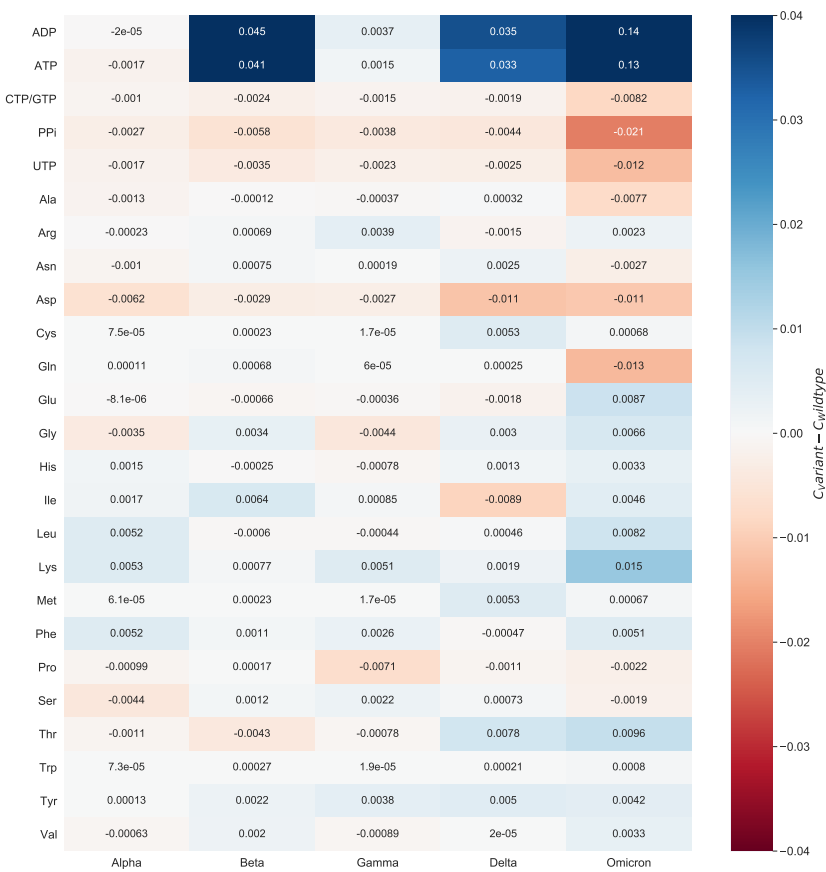



Figure 8. Variant-wise comparison of stoichiometric coefficients derived directly mutated sequences and the wildtype. The difference between the average stoichiometric coefficients of the individual variants and the reference sequence was computed. Red color highlights decreased stoichiometric coefficients in the variants, while increased coefficients are colored by blue. A remarkable increase can be observed in the stoichiometric coefficients of ATP and ADP between the Omicron variant and the wildtype. The stoichiometries of charged and hydrophobic amino acids were higher for the Omicron variant. All in all, the variations between mutants and wildtype are very small.

After investigating the mutations' impact on the viral stoichiometric coefficients, we tested the effectiveness of the previously identified targets against the SARS-CoV-2 variants repeating the single-reaction deletions and HDE experiments. Our single-reaction knock-outs indicated GK1 to be the only potent antiviral inhibitor. All host-based targets detected from the HDE analysis to have an inhibitory effect on SARS-CoV-2 for all variants are shown in Figure 9. Targets were reported as potentially effective when the virus growth rate with altered bounds was lower than the threshold of 90 % of its initial growth rate. The CTPS1 was reported to have the highest virus inhibitory effect across all Variants of Concern. After its inhibition, the virus growth dropped to 24.4 % to 37.5 % of its initial growth in the host cell. Further possible compounds were found to inhibit the viral production while keeping the host at maximum. Eight targets in total were detected to be WT-specific: ACGAM2E, DGK2m, DGNSKm, DGSNtm, HEX1, NDPK8m, PUNP4, and UAG2EMA. Except for CTPS1, GK1 was a common target, which constraint led to a reduced virus growth, however not as effective as CTPS1. Moreover, the five SARS-CoV-2 variants shared twelve additional hits with the WT that reported inhibitory effects (S5 Fig.). Our integrated host-virus model suggested the supplementation of L-proline and phosphate in the host's environment as potential targets ensuring the cell's maintenance. Moreover, four targets from the carbohydrate metabolism (UAGDP, ACGAMPM, ACGAMK, and PGMT) showed a remarkable inhibitory effect in all studied variants, while once more targeting the metabolism of purines and pyrimidines seemed promising for all virus variants.

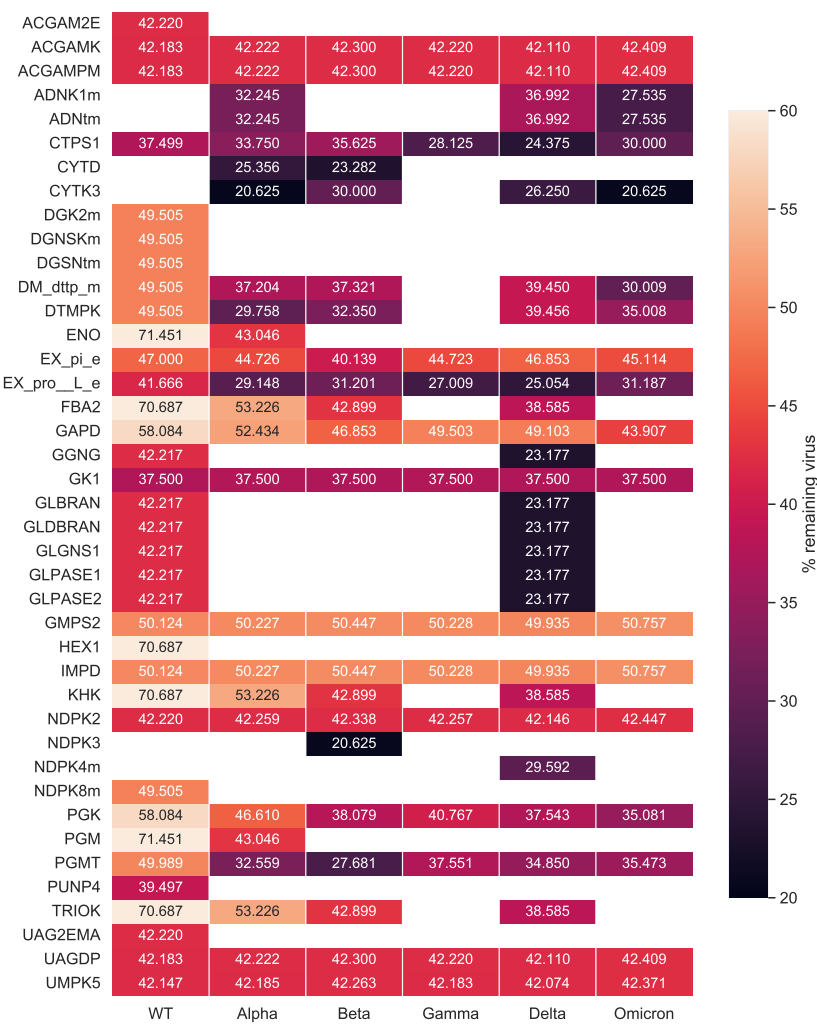



Figure 9. Results of the host-derived enforcement applied to all known variants of concern. The range and effect of reaction inhibitions on the VBOF were calculated while keeping the host's maintenance at 100 %. Only targets predicted across all retrieved sequences for a single variant were considered robust and were examined further. Empty cells in the heatmap represent targets that were not predicted as potential inhibitors for the corresponding variant. `CTPS1` showed the highest inhibitory effect against the virus at all studied variants, followed by `GK1`. Targeting the amino sugar and nucleotide carbohydrate metabolism highlighted to be a robust hit against all studied variants.

Existing drugs and inhibitors could target predicted enzymes to hinder the growth of SARS-CoV-2

The computational approaches used here allowed the prediction of diverse enzymatic targets that could inhibit the SARS-CoV-2 replication in primary human bronchial epithelial cells. For these targets, we evaluated their corresponded enzymes considering already existing approved drugs using the BRENDA[35] and DrugBank[36] databases. We found various hitherto approved drugs and compounds that target some of the predicted reactions and could inhibit them, including those targeting the very promising enzymes `CTPS1` and `GK1`. Table 3 lists examples of already existing drugs that inhibit our predicted anti-SARS-CoV-2 target reactions. These compounds and drugs could be used as an indication to validate the computational predictions made here experimentally.

Like all living cells, virus-infected cells require nucleotides to synthesize deoxyribonucleic and ribonucleic acid to strengthen their proliferation. Hence, nucleotide metabolism is regulated to establish constant pools of pyrimidines and purines. Various drugs targeting the nucleotide metabolism in viral infections represent a therapeutic approach to limit viral replication. There are two main strategies

to rewire the nucleotide metabolism: via purine and pyrimidine analogs (i.e., modified nucleosides used to stop DNA or RNA polymerase) or directly inhibiting the enzymes involved in DNA and RNA synthesis. The majority of our predicted targets are involved in the purine and pyrimidine metabolism.

Table 3. Exemplary selection of already approved drugs and compounds that act against proteins associated with our predicted anti-SARS-CoV-2 target reactions and could possibly used for antiviral therapies. All listed drugs have known pharmacological action and are sorted based on the predicted percentage of virus reduction in the wildtype sequence.

Reaction	EC-Number	Approved drug	Reference (PubMed ID)	Predicted % virus reduction
CTPS1	6.3.4.2	CPEC	10930994[53]	62.5
GK1	2.7.4.8	Acyclovir	1316735[54]	62.5
PUNP4	2.4.2.1	Ganciclovir	24107682[55]	60.5
IMPD	1.1.1.205	Ribavirin	4197928[56]	49.9

We conducted extensive literature research and highlighted the cytidine analog Cyclopentenyl cytosine (CPEC) as an already known competitive inhibitor of the CTP synthetase CTPS1 (Table 3). In CPEC, the ribose is substituted by an unsaturated carbocyclic ring and must undergo three times phosphorylation to form CPEC-TP that finally acts as an inhibitor of CTPS1[57]. A nucleoside analog named acyclovir is an approved drug that acts against GK1 (Table 3). In acyclovir, the sugar in the deoxyguanosine is substituted by an acyclic side chain, a (2-hydroxyethoxy)methyl substituent, at position nine. The viral DNA polymerase is competitively inhibited by acyclovir which acts as an analog to deoxyguanosine 5'-triphosphate (dGTP). This results in chain termination since the adherence of further nucleosides is prevented by the absence of the 3'-hydroxyl group[58].

The second approach is the direct inhibition of enzymes related to nucleotide synthesis. In the past few years, diverse enzyme inhibitors have been known to treat viral infections. One such antiviral, merimepodib, targets the action of inosine-5'-monophosphate dehydrogenase (IMPD) and has already been tested against emerging RNA viruses (e.g., Zika, Ebola, Lassa, Junin, and Chikungunya viruses)[59]. Merimepodib has also been examined in the context of SARS-CoV-2 and has demonstrated *in vitro* suppression of viral inhibition [60]. Our methods reported the IMPD as a promising hit for therapies against all SARS-CoV-2 variants with 49.9 % virus reduction. Together with merimepodib, DrugBank and BRENDA list ribavirin as an inhibitor with known pharmacological action. Several studies have postulated that ribavirin's mechanism of action lies on various not mutually exclusive pathways[61]. Lastly, with our methods, we identified the purine-nucleoside phosphorylase (PUNP4) for which ganciclovir has known inhibitory effects (Table 3).

Discussion

Studying human metabolism guides the understanding of diverse diseases by determining the cells' health. The existence of high-quality genome-scale reconstructions facilitates systems-based insights into metabolism. As complex organisms, humans embody multiple cell and tissue types, each with different functions and metabolisms, leading to the essential use of cell- or tissue-specific metabolic networks to enable the accurate prediction of the cells' metabolic behavior. Here, we presented pymCADRE, a re-implementation of mCADRE [28] in Python that allows the reconstruction of tissue-specific human models based on human gene expression data and network topology information. Similar to the original mCADRE algorithm, pymCADRE consists of three parts:

(1) ranking
(2) consistency check
(3) pruning, enabling the user to choose between two optimization methods, FVA and FASTCC, to check for model consistency (S3 Figure) .

We enriched our implementations with data pre-processing scripts that simplify multiple data curation tasks.

We used our tool to create a tissue-specific model of primary Human Bronchial Epithelial Cell (HBEC) to investigate SARS-CoV-2 infections. We used the human metabolic network Recon1 as a generic model to test our tool to avoid high computational costs. When FVA was used, pymCADRE proceeded faster than mCADRE, maintaining the highest possible similarity to the ground truth, i.e., the mCADRE-derived model. The two models showed a reaction overlap of almost 100 %, suggesting a substantial similarity between both implementations and demonstrating confidence about the quality of the pymCADRE models. Since we did not modify the initial mCADRE algorithm, the varying amount of reactions in the final tissue-specific models suggests variable performance among built-in functions in COBRAToolbox[62] and COBRApy[38]. More specifically, we observed divergent results among the two programming languages when FASTCC was employed. In both cases, the function is implemented as described by Vlassis *et al.*[41]; however the pythonic version detected a varying number of blocked reactions after multiple runs. The bug has already been reported and awaits resolve. Additionally, the detected inactive reactions were dissimilar compared to the reactions in the mCADRE model. This was not the case when the COBRApy methods,

```
flux_variability_analysis()
```

or

```
find_blocked_reactions()
```

from the package cobra.flux_analysis were employed. Moreover, the current version of FVA in MATLAB only supports the industrial proprietary CPLEX versions older than V 12.10[63]. The latest solver release, V 20.1 (released in December, 2020), does not yet include MATLAB-related binaries, and hence, FVA from the COBRAToolbox is of restricted use. This problem is resolved by pymCADRE, as the latest version of COBRApy enables the users to choose among the open-source GLPK package and the CPLEX solver from IBM to perform optimization tasks. Another reason for the divergent performance among both tools could be the implementation of organic exchange/demand reactions detection. We achieved this in a more powerful and fully automated script. Thus, pymCADRE detected four additional organic exchange/demand reactions in Recon1, affecting the result of consistency checks. The utilized human generic model, Recon1, does not include a BOF. We updated the generic human model by including a BOF extracted from the macrophage model published by Bordbar *et al.*[34].

Furthermore, we used our model and simulated an infection with the SARS-CoV-2 to better understand the host's impact on the virus and vice versa. For this purpose, we generated a SARS-CoV-2 VBOF based on the protocol of Aller *et al.* to create an integrated metabolic model aiming the analysis of host-virus interactions and the identification of effective targets for antiviral therapeutic strategies[9]. They recovered already known antiviral targets for the Chikungunya, Dengue, and Zika viruses within the human macrophage cell, verifying their approach's robustness. As Aller *et al.* suggested, FBA and FVA can be employed to predict essential host reactions, especially in cases of novel viruses. Two different computational experiments achieved this: single-reaction knock-outs and host-derived enforcement. Both approaches verified GK1 as a target to restrict the SARS-CoV-2 growth without harming the host. GK1 has already been reported to show inhibitory effects in the macrophage and the lung models[23,48]. Similarly, our results confirmed enzymatic hits from the glycolysis (ENO, PGM, and PGK) that have been previously described in the literature[23].

However, our methods revealed further novel targets with promising results. CTPS1 constrained, resulted in a remarkable suppression of the viral replication in the cell, similar to GK1. The pre-print from Rao *et al.* experimentally describes how the SARS-CoV-2 activates CTPS1 deploying several proteins to induce the synthesis of CTP, suppressing thus the interferon production and downstream immune response[64]. The authors suggest targeting the inhibition of the host enzyme CTPS1 as a potential approach to restore the interferon induction and, therefore, to hinder the SARS-CoV-2 replication. Notwithstanding, CPEC has been previously described to exhibit antiviral and anti-tumor

effects. More specifically, it is known to be active against a wide range of RNA and DNA viruses (e.g., influenza, herpes simplex, and yellow fever) *in vitro*[65,66]. Similarly, modulating the pyrimidine ribonucleotide levels has been a widely studied approach in treating cancer. As of today, it has been examined most extensively in leukemic cell lines, but also in the context of colorectal carcinoma, brain tumor, and neuroblastoma[67]. Although dosis-related hypotension events occurred in patients with colon carcinoma treated with CPEC in Phase I trials[68], the cardiotoxicity could not be reproduced in established rat models[69]. Studies have reported the deaminated product of CPEC, CPEU, as well as cytidine as potential modulators of the cytotoxic activity of CPEC[70,71]. It remains to be investigated *in vivo* to which extent it is possible to establish antiviral activities with CPEC without toxic side-effects, but also in combination with other drugs.

Most of our hits fall into the purine and pyrimidine metabolism and are tightly coupled. This implies and verifies that drugs targeting the nucleotide metabolism exemplify a common therapeutic strategy to restrict SARS-CoV-2 replication. We conducted extensive literature and database search and found acyclovir that targets GK1 from the purine synthesis pathway. We conducted extensive literature and database search and found acyclovir that targets GK1 from the purine synthesis pathway. So far, acyclovir is the standard gold treatment of infections with the herpes virus and the Varicella-Zoster Virus (VZV)[6,72]. In the context of SARS-CoV-2, acyclovir has been proposed in studies as a drug with an antiviral potential against coronaviruses[73], more specifically SARS-CoV-2 concurrently with signs of reactivation of VZV[74]. The authors assumed that this reactivation is coupled to the unusually low count of lymphocytes (lymphopenia) in the COVID-19 patients' blood. Its mechanism of action resembles molnupiravir, which has been granted the Food and Drug Administration (FDA)-Emergency Use Authorization against SARS-CoV-2 infections[75]. Both drugs target the viral replication by mimicking ribonucleosides and causing mutagenic effects. Compared to acyclovir, which leads to immediate chain termination, molnupiravir continues incorporating of nucleobases until a mismatch occurs, resulting in an error catastrophe. The only FDA-approved drug called remdesivir acts similarly and is an ATP analog and causes delayed chain termination. Hence, acyclovir's mechanism of action indicates a high potential for successful use against SARS-CoV-2 infections. Intravenous ritonavir-boosted nirmatrelvir (Paxlovid) has also received the Emergency Use Authorizations by FDA medication used in COVID-19 patients. Antiviral effects occur in an earlier stage as it prevents viral replication by inhibiting protein synthesis. Its disadvantage is that it may cause adverse effects upon drug-drug interactions since ritonavir can be dangerous for patients taking antiarrhythmics, blood thinners, and further medications[76]. Nevertheless, several monoclonal Antibodies (mABs), such as bebtelovimab, bamlanivimab-etesevimab, and tocilizumab, have been authorized for intravenous administration and subsequently revised with the emergence of the Omicron variant[77]. On the contrary, acyclovir can be administered orally, making it easier for self-use.

Besides that, we predicted enzymatic candidate targets from the carbohydrate metabolism. In more detail, reactions from the amino/nucleotide sugar and sucrose metabolism demonstrated higher antiviral effects than targets from the glycolysis or the fructose and mannose metabolism. Among others, carbohydrates are essential components of viral particles, with some playing a crucial role in their attachment and penetration into host cells[78]. They have been extensively studied as therapeutic targets against viral infections, while two of the already FDA-approved drugs to treat SARS-CoV-2 [79], remdesivir and molnupiravir, belong into the class of carbohydrate-based antiviral drugs. Our results demonstrated that targeting the path leading to the production of the sialic acid N-Acetyl-D-mannosamine (ManNAc), could result in up to 57.9 % SARS-CoV-2 inhibition.

Moreover, we tested two different growth media to validate the robustness of our predicted targets. GK1 was shown to be more effective against the virus with the blood medium defined, compared to the minimal defined medium. Using both media, NDPK2 demonstrated the same inhibitory effect as UMPK5, while nucleoside diphosphate kinase 3 (NDPK3) constrained with the minimal medium showed a higher effect on virus replication.

We further validated the robustness of our host-based targets against all five variants of concern (Alpha, Beta, Gamma, Delta, and Omicron). To accelerate the VBOF reconstruction, we developed PREDICATE to analyze multiple sequences for a single variant rapidly and in an automated way. Within this tool, we also implemented an algorithm to modify reference protein sequences and introduce amino acid mutations. Our implementations are transferable to all RNA viruses, as they are composed of the same building blocks. Firstly, we evaluated the mutations' effect on the computed stoichiometric coefficients variant-wise for the corresponding mutations. The high stoichiometric coefficients for ATP and ADP are consequences of decreased total viral molar masses and increased total amino acid counts. We observed increased use of lysine in the Omicron variant because most mutations replace amino acids with lysine. The opposite effect was observed in Omicron for asparagine. All single-reaction deletions across all variants highlighted NDPK1 as a potential robust antiviral inhibitor. The NDPK1 also proved by HDE to have the highest inhibitory effect against SARS-CoV-2, without harming the host cell. Besides that, supplementation of L-histidine, L-threonine, L-lysine, L-proline, and L-tryptophan in the host's environment shown to interrupt the virus's growth in all five SARS-CoV-2 variants.

Future improvements need to be done to make pymCADRE computationally feasible with more complex and more comprehensive models, including Recon2.2[?] and Recon3D[?]. Currently, pymCADRE and mCADRE need a large amount of computational time to complete the ranking of reactions when a more complex generic model, like Recon3D, is used. Both tools are automatically killed during pruning as there is no sufficient memory for them to process further reactions. However, we used Recon3D to fill missing knowledge in our model Recon1-HBEC. Our targets' effectiveness needs to be verified in more updated networks that better represent the human metabolism. So far, we tested the results of our pipeline using gene expression data from cell lines originating from primary cells that are easier to handle and analyze. With these, we verified targets already described for SARS-CoV-2 and ensured prediction accuracy. Further future target validation step would be to employ RNA-Seq data of primary cells that retain more traits of living cells and capture the entire transcriptome, consequently, the gene and transcript abundance. This will enable the detection of further unknown enzymatic targets guiding novel antiviral therapies.

Our integrated bronchial-specific metabolic model could be further expanded and investigated regarding the consequences of any upcoming mutation in the predicting antiviral targets. Models created by pymCADRE could be utilized to simulate the interaction of bacterial pathogens or symbionts and detect potential antiviral targets for drugs against emerging viruses on different host cells quickly. This new software provides the basis for systematic studies of a wide range of integrated computer models for host-pathogen interaction. It reduces the time for creating such models maintaining the highest possible similarity compared to the ground truth model. Our methods are based on the metabolic fluxes of infected cells and the interactions between the host cell and the virus. The latter remain unaffected by evolutionary changes. This, together with the fact that virus replication generally depends on conserved cellular pathways, drastically increases the likelihood of identifying druggable targets with broad antiviral activity. In addition, our predicted host-based targets are derived based on human patient data increasing thus their clinical relevance and their potential to achieve higher efficacy in COVID-19 therapies. Our database-derived drug compounds are experimentally supported and have already been suggested for other single-stranded RNA viruses, opening up the potential of experimentally verifying their safety, toxicity, and efficacy in cell culture experiments and in *in vitro* assays. Moreover, their optimum dosage and route of administration at different infection stages must be determined , since metabolic approaches do not consider that.

Altogether, we propose a complete workflow to create constraint-specific models and use them to predict host-based antiviral targets based on metabolic changes in infected cells. Targeting the host cell metabolic pathways has the benefit of robustness and evolutionary stability while it enables the re-purposing of already available drugs and leads to broad-spectrum putative therapeutics. For some viral infections, such as those caused by enveloped viruses, e.g., HIV, hepatitis B, or the human

papillomavirus, it can be effective to target viral proteins with enzymatic activity (e.g., the protease or viral polymerase). However, focusing on viral proteins enhances the evolution of resistance, mainly when used in monotherapy, while new variants carry underlying resistances. Additionally, these direct-acting antivirals are highly virus-specific, preventing from pan-viral efficacy and hindering pandemic preparedness. With that, targeting the host's metabolism using our approaches restrains the emergence of resistance. It reveals host pathways and enzymes essential for viral replication but dispensable for cellular maintenance and survival. Our pipeline has the advantage that applies to all RNA viruses that infect host cells, remarkably reduces the duration of target identification and compound selection, and accelerates the pre-clinical phase. Focusing on the metabolic changes of infected cells, we aim to apply our methods for rapid identification of potential antiviral targets to efficiently prevent future pandemics concerning various viruses and host cell types, promoting pandemic preparedness.

Materials and Methods

Overview of pymCADRE

The tool can be executed via the command line using:

```
python pymcadre.py
```

or using the provided Jupyter notebook named:

```
main_pymCADRE.ipynb
```

The package can also be found on the Python Package Index[80] (https://pypi.org/project/pymCADRE/) and can be installed using:

```
pip install pymcadre
```

Ranking of reactions

The first step in the pymCADRE pipeline is the ranking of all reactions found in the generic model, as Wang *et al.* proposed[28]. The ranking relies on three criteria: expression-based evidence, connectivity of reactions within the model, and confidence-based evidence. The assignment of evidence scores to reactions aims their division to cores and non-cores.

After binarizing the gene-expression data, the frequency of a gene's expression across all experiments of the same tissue is computed; this is the ubiquity score $U(g)$ for each gene g:

$$\forall g \in G : U(g) = \frac{1}{\mid N \mid} \sum_{n \in N} X_{g,n} \tag{1}$$

where N is the total number of samples and $X_{g,n} \in \{0,1\}$ denotes the absence or presence of the gene g in sample $n \in N$. For instance, if a gene is expressed in three of five samples, its ubiquity score will be 0.6. These scores are mapped to the corresponding reactions based on Gene-Protein-Reaction associationss (GPRs). That is the expression-based evidence $E_x(r)$ and can be either the minimum or maximum of two ubiquity scores depending on the respective GPRs rule: AND or OR. The expression-based evidence ranges from zero to one, indicating how likely a reaction is present in the selected tissue. More specifically, a score of zero represents a non-active reaction, while reactions with $E_x(r) > 0.9$ define the core set.

Afterwards, non-core reactions are ranked based on the connectivity-based evidence $E_c(r)$, using the network topology information of the generic model. This score defines in which order the reactions should be removed during pruning. The stoichiometric relationships in matrix **S** are applied to determine whether two reactions are connected. A pair of reactions are considered to be linked if they share at least one metabolite. For this purpose, a so-called weighted influence $WI(r)$ is calculated

as the ratio between $E_x(r)$ and the outgoing influence of each reaction, i.e., the number of reactions connected to it. Then, the actual connectivity-based evidence is determined by the sum of the weighted influences of all reactions adjacent to reaction r. In the supplementary file S2 Fig. we graphically illustrate the computation of each score using a toy metabolic network comprising six reactions and four genes coming from four samples. Lastly, the confidence level-based evidence $E_l(r)$ is the third measure of evidence for non-core reactions and indicates the level of biological evidence for the generic model.

Check Model Function

After classifying the reactions into cores and non-cores, pymCADRE tests the model's ability to produce key metabolites from glucose. These include compounds in the TCA and glycolysis, non-essential amino acids, and more. Totally, 38 metabolites were tested based on previously described criteria and used to evaluate similar models by the authors of mCADRE [28]. This list can be expanded and modified utilizing metabolomics data to include tissue-specific metabolites or known abilities of the tissue of interest.

Model Pruning

The last step of pymCADRE is to sequentially remove each non-core reaction in a reversed order, i.e., beginning from those with the lowest calculated evidence[28]. The respective reaction will be removed if, and only if, its elimination does not prevent the model from producing key metabolites and the set of core reactions remains consistent. This consistency is tested by determining each reaction's minimum and maximum flux while ensuring that at least one is zero.

More specifically, firstly, the production of precursor metabolites is checked. If this test fails, there is no need to check for model consistency with FVA or FASTCC (time-saving step). If the test leads to successful results, the set of inactive cores and non-cores is determined, and the algorithm moves on with the removal of reactions. Reactions with zero expression will be removed with their corresponding inactive core reactions if sufficiently more non-cores are pruned. On the other hand, if the reaction has expression evidence, pymCADRE only attempts to remove inactive non-cores.

Integration of transcriptomic data in a human genome-scale metabolic model

The functionality of pymCADRE was tested using gene expression data of primary Human Bronchial Epithelial Cell (HBEC) downloaded from the Gene Expression Omnibus (GEO) database (accession number: GSE5264)[81]. All data obtained from GEO underwent manual curation and pre-treatment with scripts that we provided together with the pymCADRE source code. Firstly, the expression data were binarized based on the associated RMA signal intensity values and an absolute call value (i.e., P = present, A = absent, and M = on the borderline detection) was defined. This indicates whether messenger Ribonucleic Acid (mRNA) has been detected for that specific gene or not, meaning whether it is expressed or not. The second curation step involved collecting confidence scores from the Virtual Metabolic Human (VMH) database, assigned to all reactions in the model. Then, the raw sample transcriptomic data was enhanced with two new information, gene symbol and Entrez identifiers. During binarization, genes present in the sample took the value of one, while marginal and absent calls were assigned to zero. Lastly, the essential ubiquity scores were calculated to represent a single gene's expression frequency across all samples.

The literature-based Recon1[30] was obtained from the BiGG database[82] and was used as a generic host human model. It consists of 3741 reactions, 2766 metabolites, 1905 transcripts, and 1497 unique genes. We also incorporated a BOF to Recon1 since it does not include one. For this purpose, we used the objective function from the human alveolar macrophage model published by Bordbar *et al.* in 2010[34]. The biomass reaction with the identifier `biomass_hbec` represents the cellular maintenance requirements such as the ATP maintenance.

In the Recon1 model, there is no constraint growth medium defined; thus, all extracellular transport reactions have a minimum flux value of $-1000.0\,\mathrm{mmol\,gDW^{-1}\,h^{-1}}$. This means that all exchanges are allowed to carry a flux (rich medium), resulting to unusually high cell growth rates. We have defined here a minimal growth medium using the COBRApy built-in function[38], which contains only essential components for growth. Since the availability of nutrients has a major impact on the metabolic fluxes, we re-ran our simulations using the blood medium[37]. The exact compositions of both media are provided in the supplementary file S5 Table.

We manually expanded our model by adding missing exchange reactions to all extracellular metabolites. We also updated all reaction annotations in our tissue-specific model, Recon1-HBEC, by assigning KEGG IDs[39] and retrieving the corresponding pathways using the KEGG Representational State Transfer (REST) Application Programming transfer Interface (API). These subsystems were incorporated into the model as additional annotations to each reaction with the biological qualifier type BQB_OCCURS_IN. The reaction pathways were merged into main classes based on the KEGG classification system (https://www.kegg.jp/kegg/pathway.html). Additionally to the functionality checks incorporated into the mCADRE and consequently into pymCADRE, we examined the presence of futile cycles in our final tissue-specific model. As Fritzemeier *et al.* propose, we tested the production of energy-generating compounds by including energy dissipation reactions and disabling the external uptake of all metabolites[42]. Our final model could not produce any of the tested metabolites, meaning no futile cycles were included. The tested compounds are listed in the supplementary file S1 Table.

The reconstructions were conducted using a 3.3 GHz processor and 16 GB RAM, while MEMOTE [46] and the SBML Validator from the libSBML [45] were employed to assess the model's quality.

Stoichiometric reconstruction of SARS-CoV-2 biomass objective function

Similar to the biomass production function used for microbial metabolic models, the VBOF is a single pseudo-reaction imitating the production of different virus particles. It consists of nucleotides, amino acids, and components necessary for energy supply. The SARS-CoV-2 virus biomass objective function was created as proposed by Aller *et al.* and as extended by Renz *et al.*. The approach considers the viral structure and its genome sequence, the subsequently encoded proteins, and their copy number, as well as the energy requirements for nucleotide and peptide bonds[9]. The viral genome and protein sequences were downloaded from the National Centre for Biotechnology Information (NCBI) nucleotide database[83] (accession number: NC_045512.2, accessed in May, 2020). The genome copy number (G_g) and the number of copies of each of the non-structural proteins (C_{np}) was assumed to be one[9]. Moreover, the copy number of structural proteins was set to 1000 for membrane proteins (C_m), 456 for nucleocapsid phosphoproteins (C_n), 120 for spike proteins (C_s), and 20 for envelope proteins (C_e)[48].

The SARS-CoV-2 falls into the fourth Baltimore group of viruses (Group IV, positive-sense single-stranded RNA viruses)[84], i.e., it synthesizes mRNA with the help of a template "-" single RNA antisense strand. Thus, the count of nucleotides in the positive strand equals the number of nucleotides in the complementary negative strand. The total moles of each nucleotide in a mole of virus particle were obtained by summing up the nucleotides in the positive and negative strand and multiplying this by the genome copy number. The moles were then converted into grams of nucleotide per mole of the virus by multiplying them with the respective molar mass of the nucleotides[9]. Similar calculations were conducted for the amino acids, as well. Eventually, the stoichiometric coefficients of each nucleotide and amino acid in the VBOF were calculated using the total viral molar mass[9].

For the estimation of the energetic requirements, the ATP requirement per amino acid polymerization and the pyrophosphate liberation during the polymerization of nucleotide monomers were considered. As proposed by Aller *et al.*, four ATP molecules and one pyrophosphate molecule are participating in the formation of nucleotide and amino acid polymers, respectively[9]. Subsequently, the total molar mass of the virus was calculated as the sum of all genome and proteome components.

Finally, to account for the lipid requirements we included phosphatidylcholine (`pchol_hs_c`), phosphatidylethanolamine (`pe_hs_c`), phosphatidylinositol (`pail_hs_c`), phosphatidylserine (`ps_hs_c`), cholesterol (`chsterol_c`), and sphingomyelin (`sphmyln_hs_c`) into the viral biomass function. Renz *et al.* examine the influence of lipids with various stoichiometric coefficients in the viral biomass function and the prediction of antiviral targets. However, they did not incorporated the lipid composition of a single virion into their final viral function[22]. We computed stoichiometric coefficients for these lipids from the surface area of a virion as suggested by Nanda and Ghosh[24].

The generated final VBOF was appended into Recon1-HBEC, with a lower bound of zero and an upper bound of 1000. The individual VBOF components and their stoichiometric coefficients are listed in S1 Table.

Prediction of host-based antiviral targets

Subsequent analysis of Recon1-HBEC allowed us to identify metabolic targets for antiviral therapies. As proposed by Aller *et al.*, FBA and FVA can be used to predict essential host reactions, especially in cases of novel emerging viruses[9]. This can be computationally achieved in two different ways: via single knock-out analysis or via HDE.

The single-reaction knock-out analysis investigates the effect of individual reactions with no flux. Both lower and upper bounds were systematically set to zero once with BOF as the objective function and once with the VBOF. Metabolic targets were reported when the host growth rate was higher than the virus growth rate and when more than 99 % of the initial host growth rate was maintained.

A less harmful approach for the cell is the host-derived enforcement. As Aller *et al.* suggest, herein method, the reaction fluxes are constraint to FVA-derived ranges so that the maintenance of the optimal host state is achieved while reducing the virus propagation[9]. For our analysis, we used an updated version of this method as modified by Renz *et al.*[48]. The re-calculated flux ranges for every reaction were then utilized, and the model was optimized for the VBOF. The resulting optima for the virus production were compared to the original optimal value. Hence, potential antiviral targets were reported when the virus growth rate with altered bounds was beneath the threshold of 90 % of the initial growth rate. Additionally, to ensure a reduction of the virus replication, we keep only targets that had a non-zero flux when the VBOF was optimized. Our Recon1-HBEC model was examined for potential antiviral targets using both methods.

Testing targets' robustness against all known variants of concern

To test our targets' robustness, we examined the consequences of concerning SARS-CoV-2 mutations on our predicted metabolic targets. As of February, 2022, five SARS-CoV-2 VOC are known to differ from the conventional virus in terms of their pathogen properties (e.g., transferability, virulence, or susceptibility to the immune response of recovered or vaccinated people). These are the Alpha, Beta, Gamma, Delta, and Omicron variants[16]. Genomic sequences of patients infected with SARS-CoV-2 were retrieved from the GISAID's EpiCoV database[51]. For each variant, we randomly selected 20 sequences adjusting only the location and variants filters as follows:

(i)   Europe/United Kingdom for VOC Alpha GRY (B.1.1.7+Q.*)
(ii) Africa/South Africa for VOC Beta GH/501Y.V2 (B.1.351+B.1.351.2+B.1.351.3)
(iii) South America/Brazil for VOC Gamma GR/501Y.V3 (P.1+P.1.*)
(iv) Asia/India for VOC Delta GK (B.1.617.2+AY.*)
(v)   Africa/Botswana and Africa/South Africa for VOC Omicron GRA (B.1.1.529) .

We investigated 100 sample sequences in total. To calculate the amino acid investment per virus, we used the annotated protein sequence of the SARS-CoV-2 reference genome (NCBI accession: `NC_045512.2`) and the mutation information extracted from GISAID. All used datasets and tested mutations are provided in the supplementary material S3 Table.

We calculated the stoichiometric coefficients of growth-related constituents for each mutated sequence and reconstructed for each one a VBOF as described in the previous sections. To speed up

the calculations, we implemented PREDICATE, an automated script, which takes as input one or more genome sequences and computes the metabolic stoichiometry using information from the viral genome, the encoded proteins and their copy numbers, and the energetic requirements (Figure 7). The amino acid coefficients are calculated using the reference protein sequence, which our algorithm mutates by introducing all reported mutations (replacements, insertions, deletions, and duplications) extracted from the metadata. Afterwards, each VBOF is integrated into a given cell-specific metabolic network, in our study Recon1-HBEC, to create a host-virus model. Lastly, PREDICATE applies single-reaction knock-outs and HDE to the integrated model resulting in experimentally testable and robust metabolic virus-suppressing targets. Our script also generates different plots, providing insights into the dataset and a better understanding of the results. To evaluate the mutations' effect on the viral biomass, we computed the mean of all estimated coefficients across all mutated sequences and compared them against the WT coefficients.

PREDICATE can be applied to either one or more nucleotide sequences and all existing RNA viruses. This makes it particularly advantageous and time-saving to simultaneously study multiple viruses and variants.

Data and Code availability

The computational host-virus model Recon1-HBEC, as well as the source code of pymCADRE and PREDICATE, test scripts, and test dataset are available in a git repository at https://github.com/draeger-lab/pymCADRE/. Supplementary tables in Microsoft Excel format are available along with this article. The SBML model[85] of the SARS-CoV-2-infected bronchial epithelial cell is available at the BioModels Database [86] as an SBML Level 3 Version 2 file [43] distributed as Open Modeling EXchange format (OMEX) archive[87]. Access the model at https://identifiers.org/biomodels.db/MODEL2202240001.

Supplementary Materials: S1 Fig. Results of the host-derived enforcement after defining the minimal growth medium. After constraining the fluxes of NDPK2 and UMPK5, 49.8 % of the initial virus remained in the host. Compared to the blood medium, these targets proved to have a greater impact on the virus growth leading to a higher decrease than GK1. **S2 Fig. Overview of the evidence-based ranking of reactions in pymCADRE.** The evidence-based ranking of reactions in pymCADRE is conducted similarly to mCADRE and consists of three main parts: (A) After binarizing tissue-specific data, the frequency of a gene's expression across all experiments of the same tissue is computed; this is the ubiquity score $U(g)$ for each gene g. The expression-based evidence $E_x(r)$ is computed for each gene-associated reaction r from ubiquity scores. Reactions with a sufficiently high $E_x(r)$ value are denoted as core reactions. Non-active reactions have zero expression-based evidence. (B) Non-core reactions are ranked based on the connectivity-based evidence $E_c(r)$, using the generic models' network topology and the weighted influence $WI(r)$. Figure re-created from Wang *et al.*[28]. **S3 Fig. Hierarchical organization of the pymCADRE code and its dependencies.** The three main scripts are colored with purple, while intermediate scripts are orange-colored. First of all, the rank_reactions.py module is executed, followed by prune_model.py. The module check_model_function.py is connected to main and intermediate scripts and is used multiple times within a single run. Figure created with yEd [88]. **S4 Fig. Categorization of the compounds needed for the growth of SARS-CoV-2.** The VBOF includes totally four nucleotides, five energy-related metabolites, 20 proteinogenic amino acids, and six fatty acids. **S5 Fig. Hits from the host-derived enforcement with inhibitory effect across all examined variants of concern.** Only hits shared by all virus variants are displayed. The range and effect of reaction inhibitions on the VBOF were calculated while keeping the host's maintenance at 100 %. **S1 Table. Overview of compounds and their stoichiometric coefficients in the host and viral biomass functions together with the energy-generating compounds.** From the listed metabolites, adp_c, h_c, pi_c and ppi_c are the reaction products, while the rest the reactants. **S2 Table. Detailed information of all antiviral targets predicted using the HDE. S3 Table. Summary of datasets used for all variants of concern.** All variants are listed along with their GISAID accession number, the associated mutations and submission details . **S4 Table. The stoichiometric coefficients all the molecules included in the VBOFs created for all five examined variants of concern. S5 Table. Growth media definitions. S6 Table. Amino acids and their three-letter and one-letter codes, and their molecular weight used to construct the SARS-CoV-2 VBOF.** The molecular weights were derived from the Chemicals of Biological Interest (ChEBI) database[89] **S7 Table. Five-number summary of reaction fluxes in host and virus.** The summary consists of five values: minimum, first quartile (25th percentile), median (50th percentile), third quartile (75th percentile), and maximum. **S1 File. Python script of the PREDICATE tool. S2 File. The SBML model of the integrated host-SARS-CoV-2 bronchial epithelial cell.**

Author Contributions: Conceptualization: Nantia Leonidou, Alina Renz, Reihaneh Mostolizadeh, Andreas Dräger. Data curation: Nantia Leonidou. Formal analysis: Nantia Leonidou. Funding acquisition: Andreas

Dräger. Investigation: Nantia Leonidou. Methodology: Nantia Leonidou. Project administration: Andreas Dräger. Resources: Nantia Leonidou. Software: Nantia Leonidou. Supervision: Andreas Dräger. Validation: Nantia Leonidou. Visualization: Nantia Leonidou. Writing– original draft: Nantia Leonidou. Writing– review & editing: Nantia Leonidou, Alina Renz, Reihaneh Mostolizadeh, Andreas Dräger.

Acknowledgments: This work was funded by Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of the Excellence Strategy of the German Federal and State Governments within the project "Identification of robust antiviral drug targets against SARS-CoV-2" as well as 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 BMBF), grant № 8020708703. The authors acknowledge the use of de.NBI cloud and the support by the High Performance and Cloud Computing Group at the *Zentrum für Datenverarbeitung* (ZDV) of the University of Tübingen and the BMBF through grant № 031 A535A. The authors acknowledge the support by the Open Access Publishing Fund of the University of Tübingen (https://uni-tuebingen.de/en/216529).

Conflicts of Interest: The authors declare no conflict of interest

dGTP	deoxyguanosine 5'-triphosphate
mCADRE	metabolic Context-specificity Assessed by Deterministic Reaction Evaluation
mRNA	messenger Ribonucleic Acid
ABC	ATP-binding cassette
ADP	Adenosine 5'-diphosphate
API	Application Programming transfer Interface
ATP	Adenosine 5'-triphosphate
HBEC	Human Bronchial Epithelial Cell
BiGG	Biochemcial, Genetical, and Genomical
BMBF	Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*)
BMBF-DZG	*Deutsche Zentren der Gesundheitsforschung*
BOF	Biomass Objective Function
BRENDA	Brunswick Enzyme Database
ChEBI	Chemicals of Biological Interest
CMFI	Controlling Microbes to Fight Infections
COBRApy	Constraints-Based Reconstruction and Analysis for Python
COVID-19	Coronavirus Disease 2019
CDP	Cytidine 5'-diphosphate
CMP	Cytidine 5'-monophosphate
CPEC	Cyclopentenyl cytosine
CTP	Cytidine 5'-triphosphate
CTPS1	CTP synthase 1
DFG	*Deutsche Forschungsgemeinschaft*
DNA	Deoxyribonucleic Acid
DZIF	German Center for Infection Research
EC	Enzyme Commission
EDA	Explanatory Data Analysis
ExoN	Exonuclease
FBA	Flux Balance Analysis
FDA	Food and Drug Administration
FVA	Flux Variability Analysis
GDP	Guanosine 5'-diphosphate
GEM	Genome-scale Metabolic Model
GEO	Gene Expression Omnibus
GISAID	Global Initiative on Sharing All Influenza Data
GK1	Guanylate Kinase 1
GMP	Guanosine 5'-monophosphate
GPR	Gene-Protein-Reaction associations
HDE	host-derived enforcement
HIV	Human Immunodeficiency Virus
ID	Identifier
IMP	Inosine 5'-monophosphate
IMPD	Inosine 5'-monophosphate dehydrogenase
KEGG	Kyoto Encyclopedia of Genes and Genomes
MEMOTE	Metabolic Model Testing
NCBI	National Centre for Biotechnology Information
NDPK3	Nucleoside Diphosphate Kinase 3
NSP	Non-structural Protein
NSP14	Non-structural Protein 14
OMEX	Open Modeling EXchange format
PREDICATE	Prediction of Antiviral Targets
RAM	Random-access Memory
RMA	Robust Multiarray Analysis
RNA	Ribonucleic Acid
REST	Representational State Transfer
SARS	Severe Acute Respiratory Syndrome

SARS-CoV	Severe Acute Respiratory Syndrome Coronavirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SBFC	Systems Biology Format Converter
SBO	Systems Biology Ontology
TCA	Tricarboxylic Acid Cycle
UTP	Uridine 5'-triphosphate
VMH	Virtual Metabolic Human
VBOF	Virus Biomass Objective Function
VOC	Variants of Concern
VZV	Varicella-Zoster Virus
WHO	World Health Organization
WT	wildtype
XMP	Xanthosine 5'-phosphate
ZDV	*Zentrum für Datenverarbeitung* (Center for Data Processing)

References

1. Taubenberger, J.K.; Morens, D.M. 1918 Influenza: the mother of all pandemics. *Revista Biomedica* **2006**, *17*, 69–79. doi:10.3201/eid1201.050979.

2. Ryu, W.S. Part III. RNA Viruses. In *Molecular Virology of Human Pathogenic Viruses*; Ryu, W.S., Ed.; Academic Press: Boston, 2017; pp. 149–150. doi:https://doi.org/10.1016/B978-0-12-800838-6.00044-8.

3. Maynard, N.D.; Gutschow, M.V.; Birch, E.W.; Covert, M.W. The virus as metabolic engineer. *Biotechnology journal* **2010**, *5*, 686–694. doi:10.1002/biot.201000080.

4. Leyssen, P.; De Clercq, E.; Neyts, J. Molecular strategies to inhibit the replication of RNA viruses. *Antiviral research* **2008**, *78*, 9–25. doi:10.1016/j.antiviral.2008.01.004.

5. Feld, J.J.; Hoofnagle, J.H. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature* **2005**, *436*, 967–972. doi:10.1038/nature04082.

6. Engel, J.P.; Englund, J.A.; Fletcher, C.V.; Hill, E.L. Treatment of resistant herpes simplex virus with continuous-infusion acyclovir. *Jama* **1990**, *263*, 1662–1664.

7. Warren, T.K.; Jordan, R.; Lo, M.K.; Ray, A.S.; Mackman, R.L.; Soloveva, V.; Siegel, D.; Perron, M.; Bannister, R.; Hui, H.C.; others. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* **2016**, *531*, 381–385. doi:10.1038/nature17180.

8. Maynard, N.D.; Birch, E.W.; Sanghvi, J.C.; Chen, L.; Gutschow, M.V.; Covert, M.W. A forward-genetic screen and dynamic analysis of lambda phage host-dependencies reveals an extensive interaction network and a new anti-viral strategy. *PLoS genetics* **2010**, *6*, e1001017. doi:10.1038/nature17180.

9. Aller, S.; Scott, A.; Sarkar-Tyson, M.; Soyer, O.S. Integrated human-virus metabolic stoichiometric modelling predicts host-based antiviral targets against Chikungunya, Dengue and Zika viruses. *Journal of The Royal Society Interface* **2018**, *15*, 20180125. doi:10.1098/rsif.2018.0125.

10. Smith, E.C. The not-so-infinite malleability of RNA viruses: Viral and cellular determinants of RNA virus mutation rates. *PLoS pathogens* **2017**, *13*, e1006254. doi:10.1371/journal.ppat.1006254.

11. Drake, J.W. Rates of spontaneous mutation among RNA viruses. *Proceedings of the National Academy of Sciences* **1993**, *90*, 4171–4175. doi:10.1073/pnas.90.9.4171.

12. Bar-On, Y.M.; Flamholz, A.; Phillips, R.; Milo, R. Science Forum: SARS-CoV-2 (COVID-19) by the numbers. *elife* **2020**, *9*, e57309. doi:10.7554/eLife.57309.

13. Domingo, E.; Holland, J. RNA virus mutations and fitness for survival. *Annual review of microbiology* **1997**, *51*, 151–178. doi:10.1146/annurev.micro.51.1.151.

14. Robson, F.; Khan, K.S.; Le, T.K.; Paris, C.; Demirbag, S.; Barfuss, P.; Rocchi, P.; Ng, W.L. Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. *Molecular cell* **2020**, *79*, 710–727. doi:10.1016/j.molcel.2020.07.027.

15. Kim, D.; Lee, J.Y.; Yang, J.S.; Kim, J.W.; Kim, V.N.; Chang, H. The architecture of SARS-CoV-2 transcriptome. *Cell* **2020**, *181*, 914–921. doi:10.1016/j.cell.2020.04.011.

16. World Health Organization. COVID-19 weekly epidemiological update 76– 25 January 2022. Accessed on February 8, 2022.

17. Korber, B.; Fischer, W.M.; Gnanakaran, S.; Yoon, H.; Theiler, J.; Abfalterer, W.; Hengartner, N.; Giorgi, E.E.; Bhattacharya, T.; Foley, B.; others. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **2020**, *182*, 812–827. doi:10.1016/j.cell.2020.06.043.

18. Timmers, L.F.S.M.; Peixoto, J.V.; Ducati, R.G.; Bachega, J.F.R.; de Mattos Pereira, L.; Caceres, R.A.; Majolo, F.; da Silva, G.L.; Anton, D.B.; Dellagostin, O.A.; others. SARS-CoV-2 mutations in Brazil: from genomics to putative clinical conditions. *Scientific reports* **2021**, *11*, 1–14. doi:10.1038/s41598-021-91585-6.

19. Kannan, S.R.; Spratt, A.N.; Sharma, K.; Chand, H.S.; Byrareddy, S.N.; Singh, K. Omicron SARS-CoV-2 variant: Unique features and their impact on pre-existing antibodies. *Journal of autoimmunity* **2022**, *126*, 102779. doi:10.1016/j.jaut.2021.102779.

20. Becker, S.A.; Palsson, B.O. Context-specific metabolic networks are consistent with experiments. *PLoS computational biology* **2008**, *4*, e1000082. doi:10.1371/journal.pcbi.1000082.

21. Mayer, K.A.; Stöckl, J.; Zlabinger, G.J.; Gualdoni, G.A. Hijacking the supplies: metabolism as a novel facet of virus-host interaction. *Frontiers in immunology* **2019**, p. 1533. doi:10.3389/fimmu.2019.01533.

22. Renz, A.; Widerspick, L.; Dräger, A. FBA reveals guanylate kinase as a potential target for antiviral therapies against SARS-CoV-2. *Bioinformatics* **2020**, *36*, i813–i821. doi:10.1093/bioinformatics/btaa813.

23. Delattre, H.; Sasidharan, K.; Soyer, O.S. Inhibiting the reproduction of SARS-CoV-2 through perturbations in human lung cell metabolic network. *Life science alliance* **2021**, *4*. doi:10.26508/lsa.202000869.

24. Nanda, P.; Ghosh, A. Genome Scale-Differential Flux Analysis reveals deregulation of lung cell metabolism on SARS-CoV-2 infection. *PLoS computational biology* **2021**, *17*, e1008860. doi:10.1371/journal.pcbi.1008860.

25. Cheng, K.; Martin-Sancho, L.; Pal, L.R.; Pu, Y.; Riva, L.; Yin, X.; Sinha, S.; Nair, N.U.; Chanda, S.K.; Ruppin, E. Genome-scale metabolic modeling reveals SARS-CoV-2-induced metabolic changes and antiviral targets. *Molecular systems biology* **2021**, *17*, e10260. doi:10.15252/msb.202110260.

26. Bannerman, B.P.; Júlvez, J.; Oarga, A.; Blundell, T.L.; Moreno, P.; Floto, R.A. Integrated human/SARS-CoV-2 metabolic models present novel treatment strategies against COVID-19. *Life science alliance* **2021**, *4*. doi:10.26508/lsa.202000954.

27. Kishk, A.; Pacheco, M.P.; Sauter, T. DCcov: Repositioning of drugs and drug combinations for SARS-CoV-2 infected lung through constraint-based modeling. *Iscience* **2021**, *24*, 103331. doi:10.1016/j.isci.2021.103331.

28. Wang, Y.; Eddy, J.A.; Price, N.D. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC systems biology* **2012**, *6*, 1–16. doi:10.1186/1752-0509-6-153.

29. MATLAB. *version R2020a*; The MathWorks Inc.: Natick, Massachusetts, 2020.

30. Duarte, N.C.; Becker, S.A.; Jamshidi, N.; Thiele, I.; Mo, M.L.; Vo, T.D.; Srivas, R.; Palsson, B.Ø. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences* **2007**, *104*, 1777–1782. doi:10.1073/pnas.0610772104.

31. Kam, Y.W.; Okumura, Y.; Kido, H.; Ng, L.F.P.; Bruzzone, R.; Altmeyer, R. Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells *in vitro*. *PloS one* **2009**, *4*, e7870. doi:10.1371/journal.pone.0007870.

32. Ravindra, N.G.; Alfajaro, M.M.; Gasque, V.; Huston, N.C.; Wan, H.; Szigeti-Buck, K.; Yasumoto, Y.; Greaney, A.M.; Habet, V.; Chow, R.D.; others. Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes. *PLoS biology* **2021**, *19*, e3001143. doi:10.1371/journal.pbio.3001143.

33. Ryu, G.; Shin, H.W. SARS-CoV-2 infection of airway epithelial cells. *Immune network* **2021**, *21*. doi:10.4110/in.2021.21.e3.

34. Bordbar, A.; Lewis, N.E.; Schellenberger, J.; Palsson, B.Ø.; Jamshidi, N. Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. *Molecular systems biology* **2010**, *6*, 422. doi:10.1038/msb.2010.68.

35. Jeske, L.; Placzek, S.; Schomburg, I.; Chang, A.; Schomburg, D. BRENDA in 2019: a European ELIXIR core data resource. *Nucleic acids research* **2019**, *47*, D542–D549. doi:10.1093/nar/gky1048.

36. Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research* **2006**, *34*, D668–D672. doi:10.1093/nar/gkj067.

37. Bernardes, J.P.; Mishra, N.; Tran, F.; Bahmer, T.; Best, L.; Blase, J.I.; Bordoni, D.; Franzenburg, J.; Geisen, U.; Josephs-Spaulding, J.; others. Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells, and plasmablasts as hallmarks of severe COVID-19. *Immunity* **2020**, *53*, 1296–1314. doi:10.1016/j.immuni.2020.11.017.

38. Ebrahim, A.; Lerman, J.A.; Palsson, B.O.; Hyduke, D.R. COBRApy: constraints-based reconstruction and analysis for python. *BMC systems biology* **2013**, *7*, 1–6. doi:10.1186/1752-0509-7-74.

39. Kanehisa, M.; Sato, Y.; Kawashima, M. KEGG mapping tools for uncovering hidden features in biological data. *Protein Science* **2021**. doi:10.1002/pro.4172.

40. Gudmundsson, S.; Thiele, I. Computationally efficient flux variability analysis. *BMC bioinformatics* **2010**, *11*, 1–3. doi:10.1186/1471-2105-11-489.

41. Vlassis, N.; Pacheco, M.P.; Sauter, T. Fast reconstruction of compact context-specific metabolic network models. *PLoS computational biology* **2014**, *10*, e1003424. doi:10.1371/journal.pcbi.1003424.

42. Fritzemeier, C.J.; Hartleb, D.; Szappanos, B.; Papp, B.; Lercher, M.J. Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal. *PLoS computational biology* **2017**, *13*, e1005494. doi:10.1371/journal.pcbi.1005494.

43. Hucka, M.; Bergmann, F.T.; Chaouiya, C.; Dräger, A.; Hoops, S.; Keating, S.M.; König, M.; Le Novère, N.; Myers, C.J.; Olivier, B.G.; Sahle, S.; Schaff, J.C.; Sheriff, R.; Smith, L.P.; Waltemath, D.; Wilkinson, D.J.; Zhang, F. Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 2 Core Release 2. *Journal of Integrative Bioinformatics* **2019**, *16*, 1. doi:10.1515/jib-2019-0021.

44. Rodriguez, N.; Pettit, J.B.; Dalle Pezze, P.; Li, L.; Henry, A.; van Iersel, M.P.; Jalowicki, G.; Kutmon, M.; Natarajan, K.N.; Tolnay, D.; others. The systems biology format converter. *BMC bioinformatics* **2016**, *17*, 1–7. doi:10.1186/s12859-016-1000-2.

45. Bornstein, B.J.; Keating, S.M.; Jouraku, A.; Hucka, M. LibSBML: an API library for SBML. *Bioinformatics* **2008**, *24*, 880–881. doi:10.1093/bioinformatics/btn051.

46. Lieven, C.; Beber, M.E.; Olivier, B.G.; Bergmann, F.T.; Ataman, M.; Babaei, P.; Bartell, J.A.; Blank, L.M.; Chauhan, S.; Correia, K.; others. MEMOTE for standardized genome-scale metabolic model testing. *Nature biotechnology* **2020**, *38*, 272–276. doi:10.1038/s41587-020-0446-y.

47. Courtot, M.; Juty, N.; Knüpfer, C.; Waltemath, D.; Zhukova, A.; Dräger, A.; Dumontier, M.; Finney, A.; Golebiewski, M.; Hastings, J.; Hoops, S.; Keating, S.M.; Kell, D.B.; Kerrien, S.; Lawson, J.; Lister, A.; Lu, J.; Machne, R.; Mendes, P.; Pocock, M.; Rodriguez, N.; Villéger, A.; Wilkinson, D.J.; Wimalaratne, S.; Laibe, C.; Hucka, M.; Le Novère, N. Controlled vocabularies and semantics in systems biology. *Molecular Systems Biology* **2011**, *7*, 543. doi:10.1038/msb.2011.77.

48. Renz, A.; Widerspick, L.; Dräger, A. Genome-Scale Metabolic Model of Infection with SARS-CoV-2 Mutants Confirms Guanylate Kinase as Robust Potential Antiviral Target. *Genes* **2021**, *12*. doi:10.3390/genes12060796.

49. Touré, V.; Dräger, A.; Luna, A.; Dogrusoz, U.; Rougny, A. The Systems Biology Graphical Notation: Current Status and Applications in Systems Medicine. In *Systems Medicine*; Wolkenhauer, O., Ed.; Academic Press: Oxford, 2020; Vol. 3, pp. 372–381. doi:10.1016/B978-0-12-801238-3.11515-6.

50. Balcı, H.; Siper, M.C.; Saleh, N.; Safarli, I.; Roy, L.; Kılıçarslan, M.; Özaydın, R.; Mazein, A.; Auffray, C.; Babur, Ö.; others. Newt: a comprehensive web-based tool for viewing, constructing and analyzing biological maps. *Bioinformatics* **2021**, *37*, 1475–1477. doi:10.1093/bioinformatics/btaa850.

51. Khare, S.; Gurry, C.; Freitas, L.; Schultz, M.B.; Bach, G.; Diallo, A.; Akite, N.; Ho, J.; Lee, R.T.; Yeo, W.; others. GISAID's Role in Pandemic Response. *China CDC Weekly* **2021**, *3*, 1049. doi:10.46234/ccdcw2021.255.

52. Kumar, S.; Thambiraja, T.S.; Karuppanan, K.; Subramaniam, G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *Journal of medical virology* **2021**. doi:10.1002/jmv.27526.

53. Verschuur, A.C.; Van Gennip, A.H.; Leen, R.; Meinsma, R.; Voute, P.; Van Kuilenburg, A.B. *In vitro* inhibition of cytidine triphosphate synthetase activity by cyclopentenyl cytosine in paediatric acute lymphocytic leukaemia. *British journal of haematology* **2000**, *110*, 161–169. doi:10.1046/j.1365-2141.2000.02136.x.

54. O'Brien, J.J.; Campoli-Richards, D.M. Acyclovir. *Drugs* **1989**, *37*, 233–309.

55. Furihata, T.; Kishida, S.; Sugiura, H.; Kamiichi, A.; Iikura, M.; Chiba, K. Functional analysis of purine nucleoside phosphorylase as a key enzyme in ribavirin metabolism. *Drug Metabolism and Pharmacokinetics* **2014**, *29*, 211–214. doi:10.2133/dmpk.dmpk-13-nt-065.

56. Streeter, D.G.; Witkowski, J.; Khare, G.P.; Sidwell, R.W.; Bauer, R.J.; Robins, R.K.; Simon, L.N. Mechanism of action of 1-β-D-ribofuranosyl-1, 2, 4-triazole-3-carboxamide (Virazole), a new broad-spectrum antiviral agent. *Proceedings of the National Academy of Sciences* **1973**, *70*, 1174–1178. doi:10.1073/pnas.70.4.1174.

57. Clercq, E.D. Antiviral activity spectrum and target of action of different classes of nucleoside analogues. *Nucleosides, Nucleotides & Nucleic Acids* **1994**, *13*, 1271–1295. doi:10.1080/15257779408012151.

58. Elion, G.B. Mechanism of action and selectivity of acyclovir. *The American journal of medicine* **1982**, *73*, 7–13. doi:10.1016/0002-9343(82)90055-9.

59. Tong, X.; Smith, J.; Bukreyeva, N.; Koma, T.; Manning, J.T.; Kalkeri, R.; Kwong, A.D.; Paessler, S. Merimepodib, an IMPDH inhibitor, suppresses replication of Zika virus and other emerging viral pathogens. *Antiviral research* **2018**, *149*, 34–40. doi:10.1016/j.antiviral.2017.11.004.

60. Bukreyeva, N.; Mantlo, E.K.; Sattler, R.A.; Huang, C.; Paessler, S.; Zeldis, J. The IMPDH inhibitor merimepodib suppresses SARS-CoV-2 replication *in vitro*. *BioRxiv* **2020**. doi:https://doi.org/10.1101/2020.04.07.028589.

61. Te, H.S.; Randall, G.; Jensen, D.M. Mechanism of action of ribavirin in the treatment of chronic hepatitis C. *Gastroenterology & hepatology* **2007**, *3*, 218.

62. Heirendt, L.; Arreckx, S.; Pfau, T.; Mendoza, S.N.; Richelle, A.; Heinken, A.; Haraldsdóttir, H.S.; Wachowiak, J.; Keating, S.M.; Vlasov, V.; others. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. *Nature protocols* **2019**, *14*, 639–702. doi:10.1038/s41596-018-0098-2.

63. Cplex, I.I. V12. 1: User's Manual for CPLEX. *International Business Machines Corporation* **2009**, *46*, 157.

64. Rao, Y.; Wang, T.Y.; Qin, C.; Espinosa, B.; Liu, Q.; Ekanayake, A.; Zhao, J.; Savas, A.C.; Zhang, S.; Zarinfar, M.; others. Targeting CTP synthetase 1 to restore interferon induction and impede nucleotide synthesis in SARS-CoV-2 infection. *bioRxiv* **2021**. doi:10.1101/2021.02.05.429959.

65. De Clercq, E.; Murase, J.; Marquez, V.E. Broad-spectrum antiviral and cytocidal activity of cyclopentenylcytosine, a carbocyclic nucleoside targeted at CTP synthetase. *Biochemical pharmacology* **1991**, *41*, 1821–1829. doi:10.1016/0006-2952(91)90120-t.

66. Marquez, V.E.; Lim, M.I.; Treanor, S.P.; Plowman, J.; Priest, M.A.; Markovac, A.; Khan, M.S.; Kaskar, B.; Driscoll, J.S. Cyclopentenylcytosine. A carbocyclic nucleoside with antitumor and antiviral properties. *Journal of medicinal chemistry* **1988**, *31*, 1687–1694. doi:10.1021/jm00117a004.

67. Schimmel, K.J.M.; Gelderblom, H.; Guchelaar, H.J. Cyclopentenyl cytosine (CPEC): an overview of its *in vitro* and *in vivo* activity. *Current cancer drug targets* **2007**, *7*, 504–509. doi:10.2174/156800907781386579.

68. Politi, P.M.; Xie, F.; Dahut, W.; Ford, H.; Kelley, J.A.; Bastian, A.; Setser, A.; Allegra, C.J.; Chen, A.P.; Hamilton, J.M.; others. Phase I clinical trial of continuous infusion cyclopentenyl cytosine. *Cancer chemotherapy and pharmacology* **1995**, *36*, 513–523. doi:10.1007/BF00685802.

69. Schimmel, K.; Bennink, R.; de Bruin, K.; Leen, R.; Sand, K.; van den Hoff, M.; van Kuilenburg, A.; Vanderheyden, J.L.; Steinmetz, N.; Pfaffendorf, M.; others. Absence of cardiotoxicity of the experimental cytotoxic drug cyclopentenyl cytosine (CPEC) in rats. *Archives of toxicology* **2005**, *79*, 268–276. doi:10.1007/s00204-004-0633-5.

70. Blaney, S.M.; Balis, F.M.; Grem, J.; Cole, D.E.; Adamson, P.C.; Poplack, D.G. Modulation of the cytotoxic effect of cyclopentenylcytosine by its primary metabolite, cyclopentenyluridine. *Cancer research* **1992**, *52*, 3503–3505.

71. Ford Jr, H.; Cooney, D.A.; Ahluwalia, G.S.; Hao, Z.; Rommel, M.E.; Hicks, L.; Dobyns, K.A.; Tomaszewski, J.E.; Johns, D.G. Cellular pharmacology of cyclopentenyl cytosine in Molt-4 lymphoblasts. *Cancer research* **1991**, *51*, 3733–3740.

72. Balfour Jr, H.H.; McMonigal, K.A.; Bean, B. Acyclovir therapy of varicella-zoster virus infections in immunocompromised patients. *Journal of Antimicrobial Chemotherapy* **1983**, *12*, 169–179. doi:10.1093/jac/12.suppl_b.169.

73. Tan, E.L.C.; Ooi, E.E.; Lin, C.Y.; Tan, H.C.; Ling, A.E.; Lim, B.; Stanton, L.W. Inhibition of SARS coronavirus infection *in vitro* with clinically approved antiviral drugs. *Emerging infectious diseases* **2004**, *10*, 581. doi:10.3201/eid1004.030458.

74. Nofal, A.; Fawzy, M.M.; Deen, S.M.S.E.; El-Hawary, E.E. Herpes zoster ophthalmicus in COVID-19 patients. *International Journal of Dermatology* **2020**, *59*, 1545. doi:10.1111/ijd.15240.

75. Kabinger, F.; Stiller, C.; Schmitzová, J.; Dienemann, C.; Kokic, G.; Hillen, H.S.; Höbartner, C.; Cramer, P. Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nature structural & molecular biology* **2021**, *28*, 740–746. doi:10.1038/s41594-021-00651-0.

76. Marzolini, C.; Kuritzkes, D.R.; Marra, F.; Boyle, A.; Gibbons, S.; Flexner, C.; Pozniak, A.; Boffito, M.; Waters, L.; Burger, D.; others. Recommendations for the management of drug-drug interactions between the COVID-19 antiviral nirmatrelvir/ritonavir (Paxlovid) and comedications. *Clinical Pharmacology & Therapeutics* **2022**. doi:https://doi.org/10.1002/cpt.2646.

77. Cavazzoni, P. Coronavirus (COVID-19) update: FDA limits use of certain monoclonal antibodies to treat COVID-19 due to the Omicron variant. *US Food and Drug Administration* **2022**.

78. Mathez, G.; Cagno, V. Viruses like sugars: how to assess glycan involvement in viral attachment. *Microorganisms* **2021**, *9*, 1238. doi:10.3390/microorganisms9061238.

79. Cao, X.; Du, X.; Jiao, H.; An, Q.; Chen, R.; Fang, P.; Wang, J.; Yu, B. Carbohydrate-based drugs launched during 2000- 2021. *Acta Pharmaceutica Sinica B* **2022**. doi:10.1016/j.apsb.2022.05.020.

80. Python Package Index - PyPI. https://pypi.org/.

81. Ross, A.J.; Dailey, L.A.; Brighton, L.E.; Devlin, R.B. Transcriptional profiling of mucociliary differentiation in human airway epithelial cells. *American journal of respiratory cell and molecular biology* **2007**, *37*, 169–185. doi:10.1165/rcmb.2006-0466OC.

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

83. Geer, L.Y.; Marchler-Bauer, A.; Geer, R.C.; Han, L.; He, J.; He, S.; Liu, C.; Shi, W.; Bryant, S.H. The NCBI BioSystems database. *Nucleic acids research* **2010**, *38*, D492–D496. doi:10.1093/nar/gkp858.

84. Baltimore, D. Expression of animal virus genomes. *Bacteriological reviews* **1971**, *35*, 235–241. doi:10.1128/br.35.3.235-241.1971.

85. Keating, S.M.; Waltemath, D.; König, M.; Zhang, F.; Dräger, A.; Chaouiya, C.; Bergmann, F.T.; Finney, A.; Gillespie, C.S.; Helikar, T.; Hoops, S.; Malik-Sheriff, R.S.; Moodie, S.L.; Moraru, I.I.; Myers, C.J.; Naldi, A.; Olivier, B.G.; Sahle, S.; Schaff, J.C.; Smith, L.P.; Swat, M.J.; Thieffry, D.; Watanabe, L.; Wilkinson, D.J.; Blinov, M.L.; Begley, K.; Faeder, J.R.; Gómez, H.F.; Hamm, T.M.; Inagaki, Y.; Liebermeister, W.; Lister, A.L.; Lucio, D.; Mjolsness, E.; Proctor, C.J.; Raman, K.; Rodriguez, N.; Shaffer, C.A.; Shapiro, B.E.; Stelling, J.; Swainston, N.; Tanimura, N.; Wagner, J.; Meier-Schellersheim, M.; Sauro, H.M.; Palsson, B.; Bolouri, H.; Kitano, H.; Funahashi, A.; Hermjakob, H.; Doyle, J.C.; Hucka, M.; Adams, R.R.; Allen, N.A.; Angermann, B.R.; Antoniotti, M.; Bader, G.D.; Červený, J.; Courtot, M.; Cox, C.D.; Dalle Pezze, P.; Demir, E.; Denney, W.S.; Dharuri, H.; Dorier, J.; Drasdo, D.; Ebrahim, A.; Eichner, J.; Elf, J.; Endler, L.; Evelo, C.T.; Flamm, C.; Fleming, R.M.T.; Fröhlich, M.; Glont, M.; Gonçalves, E.; Golebiewski, M.; Grabski, H.; Gutteridge, A.; Hachmeister, D.; Harris, L.A.; Heavner, B.D.; Henkel, R.; Hlavacek, W.S.; Hu, B.; Hyduke, D.R.; Jong, H.; Juty, N.; Karp, P.D.; Karr, J.R.; Kell, D.B.; Keller, R.; Kiselev, I.; Klamt, S.; Klipp, E.; Knüpfer, C.; Kolpakov, F.; Krause, F.; Kutmon, M.; Laibe, C.; Lawless, C.; Li, L.; Loew, L.M.; Machne, R.; Matsuoka, Y.; Mendes, P.; Mi, H.; Mittag, F.; Monteiro, P.T.; Natarajan, K.N.; Nielsen, P.M.F.; Nguyen, T.; Palmisano, A.; Pettit, J.B.; Pfau, T.; Phair, R.D.; Radivoyevitch, T.; Rohwer, J.M.; Ruebenacker, O.A.; Saez-Rodriguez, J.; Scharm, M.; Schmidt, H.; Schreiber, F.; Schubert, M.; Schulte, R.; Sealfon, S.C.; Smallbone, K.; Soliman, S.; Stefan, M.I.; Sullivan, D.P.; Takahashi, K.; Teusink, B.; Tolnay, D.; Vazirabad, I.; Kamp, A.v.; Wittig, U.; Wrzodek, C.; Wrzodek, F.; Xenarios, I.; Zhukova, A.; Zucker, J. SBML Level 3: an extensible format for the exchange and reuse of biological models. *Molecular Systems Biology* **2020**, *16*, e9110. doi:10.15252/msb.20199110.

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

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

88. yWorks GmbH. yEd **2019**.

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