FBA reveals guanylate kinase as potential target for antiviral therapies against coronavirus SARS-CoV-2

Alina Renz[1,2,*], **Lina Widerspick**[1], **and Andreas Dräger**[1,2,3,*]

[1]Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, 72076 Tübingen, Germany
[2]Department of Computer Science, University of Tübingen, 72076 Tübingen, Germany
[3]German Center for Infection Research (DZIF), partner site Tübingen, Germany
[*]renz@informatik.uni-tuebingen.de; draeger@informatik.uni-tuebingen.de

ABSTRACT

The novel Coronavirus (SARS-CoV-2) currently spreads all over the world as the cause of the coronavirus disease 2019 (COVID-19). Every day, the number of infections increases, and to this date, no approved antiviral therapy exists. The release of the nucleotide sequence of SARS-CoV-2 paved the way for new approaches for the identification of targets for antiviral therapies. One possible approach is the analysis of integrated human-virus metabolic models. Investigations of changed metabolic processes after virus infections and the effect of knock-outs on both the host and the virus can reveal new potential targets for antiviral therapies. We generated an integrated host-virus genome-scale metabolic model of human alveolar macrophages and SARS-CoV-2 using the recently published RNA sequence of SARS-CoV-2. Analyses of stoichiometric and metabolic changes between uninfected and infected host cells using flux balance analysis (FBA) highlighted the different requirements of host and virus. Due to the varying requirements, alterations in the metabolism, such as reaction knock-outs, can have different effects on host and virus, leading to potential antiviral targets. One of these potential targets is guanylate kinase (GK1). In FBA analyses, the knock-out of the guanylate kinase decreased the growth of the virus to zero, while not affecting the host. As GK1 inhibitors have already been described in the literature, its potential therapeutic effect for SARS-CoV-2 infections needs to be verified in *in-vitro* experiments.

Keywords: 2019-nCoV, SARS-CoV-2, Coronavirus, COVID-19, Flux Balance Analysis, Genome-Scale Metabolic Model, Host-Virus Integration, Human Alveolar Macrophage

1 Introduction

In December 2019 an outbreak of pneumonia in Wuhan, Hubei province, in China, has aroused the interest of the international community by showing alarming similarities to the outbreaks caused by other β-*Coronaviruses* (β-CoV) like the Severe Acute Respiratory Syndrome (SARS) virus[1,2].

The febrile respiratory illness caused by the novel Coronavirus (SARS-CoV-2) is thought to have spread as a zoonosis from the Huanan Seafood Wholesale Market, which was as a consequence shut down on January 1st, 2020 to prevent further transmission events[1].

On January 7th, first isolation and subsequent deep-sequencing of SARS-CoV-2 from the human lower respiratory tract samples have made the genetic sequence of the virus available to the public by January 12th, 2020, thus allowing for the identification of the virus as a Group 2B β-CoV[1,2]. SARS-CoV-2 has 82 % sequence similarity with the SARS virus, which has caused an outbreak originating in China in 2002[1,3]. The outbreak in 2002 has peaked at a total of 8098 documented cases with a case fatality rate of 9.6 %[1,2]. In contrast, the ongoing SARS-CoV-2 has reached, at time of writing, a total of 45,171 cases with an estimated fatality rate of 2 % to 5 %[4,5]. The resemblance and the severe global health threat have initiated a swift and determined implementation of public health measures by the Chinese Authorities[1,6].

While human-to-human transmissions in a nosocomial setting were the primary route of transmission of the SARS virus outbreak, Chinese horseshoe bats have been identified as putative primary reservoir and source of zoonosis for SARS[1,2,7]. Moreover, Himalayan palm civets, raccoon dogs, and Chinese ferret badgers from the Guangdong wet markets (live animal retail markets) were identified as intermediate reservoirs of zoonosis[1,8]. In the case of SARS, the switch to the human host was allowed by an adaption of the receptor binding domain (RBD) of the spike (S) protein, conferring to binding capabilities to the human angiotensin-converting enzyme 2 (ACE2)[2,8]. For the SARS-CoV-2, there was initially no proof of efficient human-to-human transmission, however, the rapid increase in cases and distinct clustering of the disease have made it clear that an efficient transmission route in-between humans exists[1,9]. Moreover, Paraskevis *et al.* have suggested, based on full-genome evolutionary analysis, a zoonosis of SARS-CoV-2 from bats to humans[7]. Accordingly, Letko *et al.* have recently identified ACE2 as the SARS-CoV-2 entry receptor[10].

As mentioned above, the SARS-CoV-2 belongs to the subfamily of β-*Coronaviruses*[1,2]. The members of the family are enveloped, single-stranded ribonucleic acid (RNA) viruses with a positive polarity genome of up to 34 kb[2,3,6]. Replication of the RNA genome is performed via an RNA-

dependent RNA polymerase (RdRP) in double-membrane vesicles (DMVs), modified to form a reticulovesicular network (RVN)[6,8,11]. In the DMVs, the 16 non-structural proteins (NSPs) are directly expressed as polyproteins pp1a and pp1ab from the (+)RNA genome[6]. Processing of pp1a/1ab by the viral main protease (M^{pro}) is essential to form the replication-transcription complex (RTC) for the subsequent expression of the viral structural proteins[6,8].

In total, the non-structural proteins constitute two thirds of the genome's coding capacity[6]. The rest encodes for structural proteins, such as the spike (S), membrane (M), and envelope (E) proteins, which cover a helical nucleocapsid made up of the nucleocapsid (N) proteins[6,8]. Unlike non-structural proteins, the structural and additional accessory proteins of *Coronaviruses* require synthesis as subgenomic messenger RNA s (mRNAs) via discontinuous transcription from (-)RNA templates[6].

Clinically, the SARS-CoV-2 seems to be milder than the former SARS virus, although they share the symptoms of febrile illness and pneumonia[1,2,6]. It was recently shown that the SARS-CoV-2 does not only infect lower respiratory tract cells, but also upper cells in the pharyngeal region[2,6]. Moreover, many *Coronaviruses* also infect macrophages and subsequently inhibit an interferon-stimulated genes (ISG)-mediated antiviral response, increasing immune evasion and pathogenicity[2,6,12,13].

By now, no antiviral treatment for *Coronaviruses* has been proven efficacious in a clinical trial[2,6]. Recently, Zhang *et al*. have found α-ketoamides to be broad-spectrum inhibitors of *Coronaviruses* by tissue-dependently inhibiting the main protease of SARS in Vero cells[3]. This inhibition has occurred in a micromolar EC_{50} range, indicating α-ketoamides to be a potential antiviral for the SARS-CoV-2[3]. However, since no therapy or vaccination is available for clinical use, the current standard of care for a SARS-CoV-2 infection is limited to the supportive treatment of symptoms[2,6].

As no antiviral treatment is currently available for *Coronaviruses*, the identification of potential antiviral targets is of great interest. One possibility of identifying new targets is the analysis of metabolic changes in infected cells. Aller *et al*. introduced a procedure for integrated human-virus metabolic models to predict host-based antiviral targets against Chikungunya, Dengue and Zika viruses. The analysis of the integrated human-virus models predicted inhibiting effects of constrained host-reactions on virus production. These predictions included already known targets of existing antiviral drugs, demonstrating the applicability of such analysis methods[14].

In our study, we integrated and analyzed a human genome-scale metabolic model infected with SARS-CoV-2. As it is shown that *Coronaviruses* infect alveolar macrophages[12,13], and Bordbar *et al*.[15] already published an extensive metabolic model of human alveolar macrophages, this model was used as a host model. For SARS-CoV-2, a virus biomass objective function (VBOF) was generated according to Aller *et al*.[14]. Subsequent analysis of the integrated host-virus model revealed potential targets for antiviral therapies.

2 Results

We developed a genome-scale metabolic model (GEM) of a human macrophage infected with the novel coronavirus SARS-CoV-2. To integrate the virus into the human macrophage, a virus biomass reaction representing the virus particle production was added to the model. This virus biomass reaction was generated based on the nucleotide and amino acid sequence (section 4 on page 7).

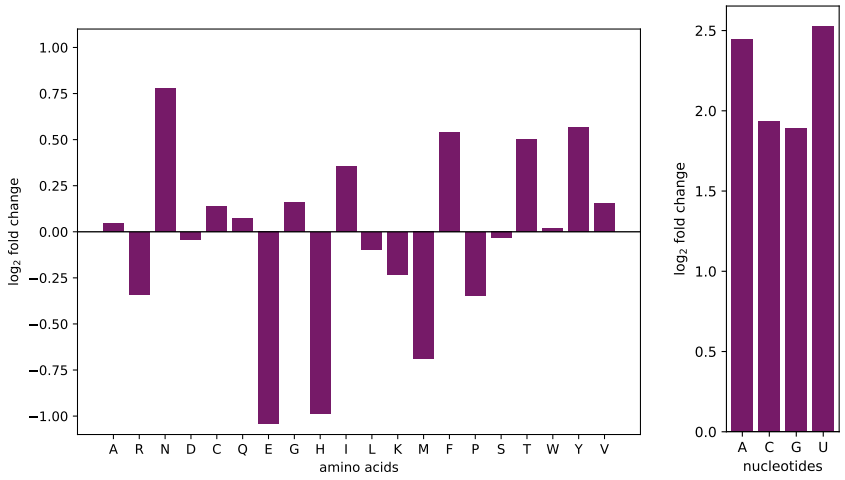
2.1 Stoichiometric differences

The human alveolar macrophage biomass maintenance function is comprised of several macromolecules, including amino acids, deoxyribonucleic acid (DNA) and RNA nucleotides, components for the energy requirements and others, such as fatty acids or phospholipids. In contrast, the virus biomass objective function only consists of amino acids, RNA nucleotides, and the components for the energy requirements. The stoichiometric coefficients of amino acids and nucleotides are compared by calculating the fold change. The result is visualized in figure 1 on the next page. The stoichiometric coefficients of L-asparagine (N), L-phenylalanine (F), L-threonine (T), and L-tyrosine (Y) are increased. In contrast, the coefficients of L-glutamate (E), L-histidine (H), and L-methionine (M) are reduced. These findings might indicate an up- or down-regulation of the respective metabolic pathways.

2.2 Influence of the copy number of the structural proteins

The calculations for obtaining a virus VBOF include the parameter for the copy number of the structural proteins (C_{sp}). The copy number of structural proteins for some viruses, such as *Alpha-* and *Flaviviruses*, is known and ranges from 180 for *Flaviviruses*[16] to 240 for *Alphaviruses*[17]. *Coronaviruses* have four major structural proteins: the envelope (E) protein, membrane (M) protein, nucleocapsid (N) protein, and the spike (S) protein[18–20]. However, the copy number of those structural proteins is not reported so far. To evaluate the influence of the copy number on the modeling results, we varied the parameter C_{sp} between 1 and 1500. As shown in figure 2a on page 4, the optimization result of the VBOF depends on the copy number. The growth rate increases for copy numbers ranging from 1 to 58 and decreases for higher copy numbers. For very high copy numbers of structural proteins, the growth rate seems to reach a steady state.

Coronaviruses have a diameter of approximately 80 nm to 120 nm[21], while *Alpha-* and *Flaviviruses* have only a diameter of approximately 40 nm to 80 nm and 30 nm to 50 nm, respectively[22]. However, not only the virus size is larger, but also the size of its structural proteins. Since no value for the structure protein copy number in *Coronaviruses* was available, we continued our analyses with a C_{sp} value of 500 and verified the results with C_{sp} values of 200, 800 and 1200.



(a) Amino acids **(b) Nucleotides**

Figure 1. Fold-change differences in amino acid and nucleotides usage. The stoichiometric coefficients of the alveolar macrophage biomass maintenance function and the virus biomass objective function (VBOF) are compared using eq. (2). Panel 1a displays the fold-change of all 20 proteinogenic amino acids. The one letter code of the amino acids is used for labeling the x-axis. Panel 1b displays the fold-change of the four RNA nucleotides. The one letter code of the RNA nucleotides is used for labeling the x-axis.

2.3 Optimization of host and virus metabolism

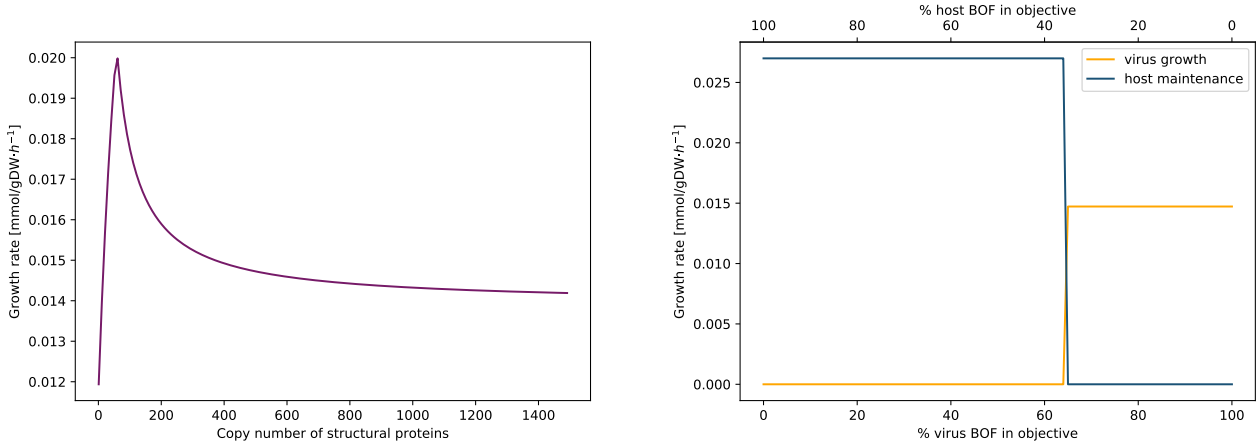
Potential changes in the human alveolar macrophage metabolism during virus infection were first examined by analyzing the following two scenarios: (i) the host cell is not infected and the metabolic system is optimized for the already included maintenance biomass reaction (host optimization). This scenario reflects the normal physiological state of human alveolar macrophages; (ii) the host cell is infected by the virus and is optimized solely for the production of the virus particles (virus optimization). The host optimization results in a biomass maintenance flux of $0.0267\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$, while the virus optimization returns a flux of $0.0147\,\mathrm{mmol}/(\mathrm{gDW}\cdot\mathrm{h})$. When defining both the host maintenance and the virus biomass function as objective functions by assigning both an objective coefficient of 1 simultaneously, the model only optimizes for the host maintenance reaction while the flux through the VBOF is zero. Last, the optimization result of host and virus metabolism was compared using a new constraint. With varying percentages of the host maintenance reaction and the VBOF on the objective expression, the effect on the respective biomass maintenance or growth function was investigated. The model does not predict an equilibrium state, where both the host maintenance reaction and the VBOF are active. As displayed in figure 2b on the next page, the switch between the maintenance of the host metabolism and the growth of the virus is at 65 % of virus biomass objective function (BOF) contribution to the objective expression. This switch in biomass production is rather insensitive to the structural proteins' copy number. For $C_{sp} = 200$, the switch occurs at 63 % virus BOF contribution, and for $C_{sp} = 800$ and $C_{sp} = 1200$ at 66 %.

2.4 Metabolic changes in alveolar macrophages after virus infection

The flux distribution for the host- and the virus-optimized states were compared using FBA and fold change calculations. For 256 reactions (7.8 % of all model reactions), a fold change was calculated. As expected from the stoichiometric analysis, reactions related to amino acid and nucleotide metabolism were altered. The main portion of changed reactions consists of transport reactions. However, also several reactions from other subsystems, such as the steroid, fatty acid synthesis (FAS), and central metabolism, are altered. An overview of altered reactions concerning their fold change is given in figure 3 on page 5. One needs to keep in mind that the fold change can only be calculated if neither of the reaction fluxes in the host- and virus-optimized state was zero. Hence, we calculated the absolute change for reactions that have either in the host- or the virus-optimized state a flux of zero. A total of 97 reactions (2.9 % of all model reactions) either turned on a previously turned off reaction or vice versa. For example, in the virus-optimized state, the virus turned off 14 reactions related to lipid metabolism.

2.5 Identification of metabolic targets for antiviral strategies

FBA and flux variability analysis (FVA) can be used to identify metabolic targets for antiviral strategies by comparing the host- and virus-optimized states after alterations in metabolic pathways.



(a) Influence of C_{sp} on the growth rate.

(b) Linear combination of weighted biomass functions.

Figure 2. 2a Influence of C_{sp} on the growth rate. The copy number of structural proteins C_{sp} in *Coronaviruses* is not yet reported in literature. Hence, its influence on the growth rate in the virus-optimized state is evaluated by varying C_{sp} between 11500. For each copy number, the integrated host-virus model was optimized for the virus biomass objective function (VBOF). For C_{sp} values between 58 and 1500, the growth rate decreases. **2b Linear combination of weighted biomass functions.** The objective function was defined as linear combination of the host maintenance and virus biomass objective function by weighting the proportion of the respective biomass functions between 0 and 1. The sum of the two weights always sums up to 1. At 65 % of virus BOF contribution to the objective expression, a switch between host maintenance and virus production occurs.

2.5.1 Knock-out of reactions reveals first promising metabolic target for antiviral strategies

The results from the stoichiometric differences and metabolic changes in alveolar macrophages after virus infections with coronavirus SARS-CoV-2 indicate alternative host- and viral-optimal states. The diverging flux distribution in the two states provides an opportunity for the identification of potential antiviral targets. To identify potential antiviral targets that limit virus production, we implemented two different analysis methods: (i) reaction knock-outs, and (ii) host-derived enforcement. The reaction knock-out revealed exactly one reaction (over all tested copy numbers of structural proteins), whose knock-out reduces the flux of the virus biomass to zero, while maintaining the host biomass maintenance at 100 %: The guanylate kinase (GK1) reaction that converts adenosine triphosphate (ATP) and guanosine monophosphate (GMP) to adenosine diphosphate (ADP) and guanosine diphosphate (GDP):

$$\mathrm{ATP} + \mathrm{GMP} \overset{\mathrm{GK1}}{\rightleftharpoons} \mathrm{ADP} + \mathrm{GDP}\,. \tag{1}$$

2.5.2 Host-derived enforcement substantiates the metabolic target and reveals further points of action

This reaction is also observed as a potential target in the second analysis method, the host-derived enforcement. In this approach, the reaction fluxes are constrained to ranges derived from FVA. With flux ranges outside of the optimal state of the virus, the virus production is perturbed while the host maintenance is not affected. For structural protein copy numbers between 500 and 1200, we identified four possible targets, including the GK1 that reduced the virus growth flux to below a threshold of 80 % of its initial value. Further potential targets concern the availability of the amino acids L-isoleucine (I), and L-lysine (K), either via the alteration of exchange reactions, or, in case of L-isoleucine, also via the L-isoleucine transporter. In contrast to GK1, where the flux through the reaction is down-regulated (or in case of the reaction knock-outs completely knocked out), the alterations of the mentioned exchange, and transport reactions go into the other direction: the uptake of the amino acids is enabled or even enforced, leading to the host's maintenance while decreasing the growth of the virus to 50 % of its initial growth. The enforcement of exchange reactions can be reached by a sufficient supply of the respective amino acid.

If the host has enough L-lysine in its environment, the maintenance of the host cell is ensured, while the growth of the virus is diminished. We repeated the analysis that evaluates the effect on the growth rates based on the percental involvement of the two biomass functions within the setting of L-lysine excess. Figure 4 on page 6 illustrates the virus inhibiting effect of L-lysine availability in the environment. Even when the objective function is solely defined by the VBOF (% of virus BOF in objective = 100 %), the growth rate of the host still exceeds the growth rate of the virus.

Analogous to L-lysine, the availability of L-isoleucine and L-tyrosine has a positive effect on the growth of the host while decreasing the growth rate of the virus by 50 %. The last potential metabolic target for antiviral strategies using the host-derived enforcement approach concerns one transport

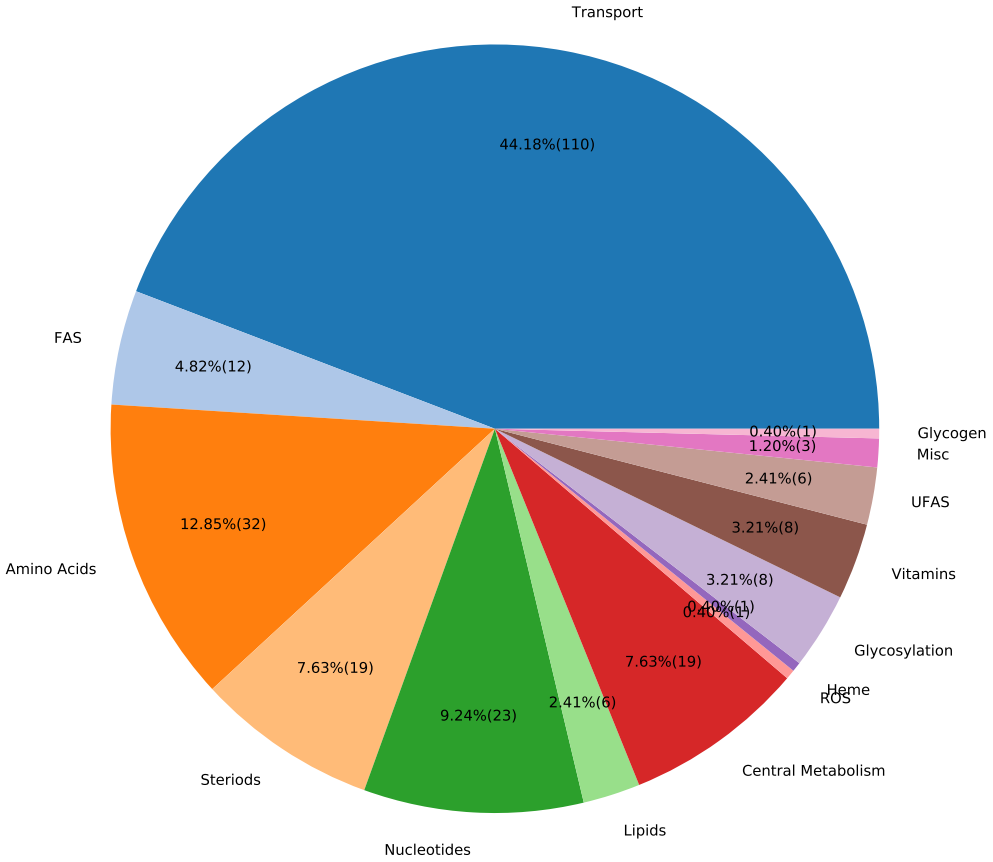


Figure 3. Altered reactions in virus-optimized state sorted by subsystems. The integrated virus-host model was optimized for the host maintenance and virus biomass objective function (VBOF) using FBA. The fold change was calculated using eq. (3) and the reactions with fold changes were grouped into the aggregated subsystems. Many altered reactions belong to the transport subsystem, or amino acid and nucleotide metabolism.

reaction of L-isoleucine (`ILEtec`). Enforcing a minimal import of L-isoleucine via this transport has the same effect as the other three targets and results in a diminution of 50 % of the virus growth. All targets obtained from the host-derived enforcement for the structural protein copy numbers C_{sp} = 500, 800 and 1200 are summarized in table 1 on the following page.

The analysis of the host-derived enforcement for $C_{sp} = 200$ showed among the in table 1 on the next page listed 29 reactions as potential antiviral targets that decrease the virus growth below the threshold of 80 %. While the four already known targets again reduced the growth rate by 50 %, the growth reduction of the other targets varies between 70 % and 80 % of the initial value. Of the 25 new targets, almost half (twelve reactions) are associated with nucleotide metabolism. Only four further reactions are affecting the amino acid metabolism and transport mechanisms, respectively. The other reactions affected are part of the central metabolism and miscellaneous reactions. The full list is provided in the supplementary material.

2.6 Existing drugs can target the predicted reactions

For the identified potential targets for antiviral therapies, we have searched for existing drugs or compounds.

2.6.1 Direct inhibition of GK1

As guanylate kinase (GK1), also known as guanosine monophosphate kinase (GMPK), was found to be an essential factor for viral growth in this study, the inhibition of the enzyme may be a feasible target in SARS-CoV-2 therapy. The enzyme catalyzes the reversible turnover of GMP or deoxyguanosine monophosphate (dGMP) to GDP or deoxyguanosine diphosphate (dGDP) by binding and transferring a phosphoryl from ATP to GMP[23,24]. GK1 has a highly conserved structure, with a core domain, a GMP-binding domain, a lid domain with catalytic residues, as well as an ATP binding (P)-loop[23,24]. Through activation of ganciclovir in

Table 1. Host-derived enforcement of reactions reduced growth of virus. The host-derived enforcement analysis (see section 4.5.2 on page 9) disclosed reactions reducing the virus growth (initial growth rate 0.0147 mmol/(gDW · h)), while maintaining the host's growth rate at 100 %.

Reaction ID	Growth host [mmol/(gDW · h)]	Growth virus [mmol/(gDW · h)]	Regulation
GK1	0.027	0.00736	down
EX_ile__L(e)	0.027	0.00736	up
EX_lys__L(e)	0.027	0.00736	up
ILEtec	0.027	0.00736	up

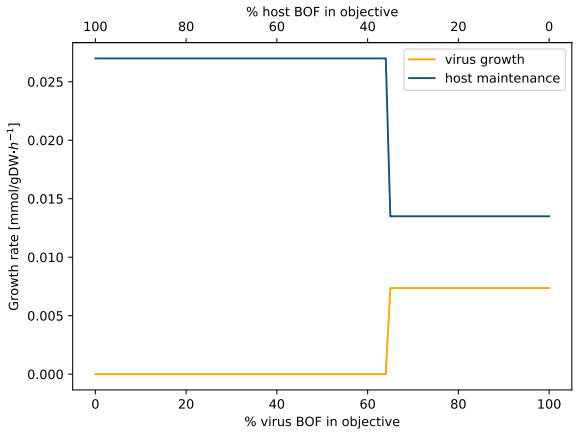


Figure 4. Linear combination of weighted biomass function with adapted boundaries for L-lysine exchange. The objective function was defined as linear combination of the host maintenance and virus biomass objective function by weighting the proportion of the respective biomass functions between 0 and 1. The sum of the two weights always sums up to 1. The upper and lower bounds of the L-lysine exchange were adapted based on the host-derived enforcement results (see section 4.5.2 on page 9). Even when the objective function is solely defined by the VBOF (% of virus BOF in objective = 100 %), the growth rate of the host still exceeds the growth rate of the virus.

herpes virus treatment and 6-thioguanine or 6-mercaptopurine activation in tumor treatment, GK1 plays an essential role in diverse treatment strategies[23,24].

Based on the results of this study, we propose a direct inhibitor of GK1 to harbor the potential of SARS-CoV-2 inhibition. Interestingly, such inhibitors have already been described in the context of antiviral therapies[23–26]. For instance, Hible *et al.* have demonstrated P1-(5'-adenosyl) P5-(5'-guanosyl) pentaphosphate (Ap5G) to be a potent, bi-substrate inhibitor of GK1[23,26]. Although Hible *et al.* have only shown the high potency of the inhibition in *Escherichia coli*, Khan *et al.* have demonstrated the binding of Ap5G to introduce a complete closure of the human GK1, indicating inaccessibility of the substrates upon inhibitor binding[23,24,26].

Potent GK1 inhibitors have also been presented by Navé *et al.*[25]. Accordingly, 9-phosphonoalkyl derivates, such as 9-(6-phosphonohexyl)guanine and 9-(6,6-difluoro-6- phosphonohexyl)guanine impede GK1 activity via competitive inhibition of GMP and non-competitive inhibition of ATP[25].

2.6.2 Acyclonucleotide analogues require GK1 activation

In subsequent work, Navé *et al.* have tested the antiviral activity of 9-phosphonopentenyl derivatives of guanine operating as acyclonucleotide analogues[27]. Acyclonucleotide analogs are pro-drugs, which require the activation by GK1 to form nucleoside triphosphate analogues[27]. These are known to inhibit viral DNA polymerases via chain termination in diverse herpes virus and retrovirus infections[27]. In accordance with this, the authors have identified vinyl phosphonates (E)-9-(5-Phosphonopent-4-enyl)guanine and (E)-9-[3-(hydroxymethyl)-5-phosphonopent-4-enyl]guanine to be inhibitors of the human immunodeficiency virus 1 (HIV-1) and human cytomegalovirus (HCMV)[27].

As this study has found increased flux through GK1 in SARS-CoV-2 infected cells, a process of pro-drug activation relying on GK1 may allow for increased activation of terminating nucleoside triphosphate analogs in infected cells compared to healthy cells. However, *Coronaviruses* are (+)single-stranded RNA (ssRNA) viruses replicated by an RdRP, unlike most viruses targeted by available acyclonucleotide analogs, such as ganciclovir and acyclovir[6,25,27]. Moreover, *Coronaviruses* lack a viral kinase required for activation of these acyclonucleotide analogs[6,23]. Despite these limitations, some market-available analogs, such as cidofovir, brincidofovir or favipiravir, do not require a viral kinase activation, and have shown *in vitro* activity against other RNA viruses (Ebola virus) and retroviruses (HIV-1)[27–31]. Thus, they may be drug candidates worth an exploration in the face of the findings in this study and the current SARS-CoV-2 outbreak.

3 Discussion

As proposed by Aller *et al.*[14], computational approaches combining FBA and FVA to recover new metabolic antiviral targets are useful, especially in cases of new and emerging viruses, such as the novel Coronavirus SARS-CoV-2. In this study, we presented a host-virus integrated genome-scale metabolic model using the human alveolar macrophage model *i*AB-AMØ-1410 as host cells and SARS-CoV-2 as virus. We

identified potential targets for antiviral therapies using reaction knock-outs and host-derived enforcement approaches and analyzing their metabolic effects on host- and virus-optimized states by optimizing either for the host maintenance or the virus biomass objective function (VBOF).

However, the VBOF constructed in this study only considers amino acids, nucleotides, and energy requirements. It does not consider or include virus-host cell recognition, viral entry, or the lipid envelope production or release[32]. Especially the metabolic process of lipid envelope production of viruses can give further insight into potential targets for antiviral therapies. First studies with other *Coronaviruses*, such as the human coronavirus 229E (HCoV-229E), suggest elevated and perturbed glycerophospholipids and fatty acids production rates in infected cells[33]. Yan *et al.* suggest the lipid metabolism regulation as a potential druggable target for coronavirus infections[33]. Further information about the lipid metabolism of *Coronaviruses* can enable the integration of lipids into the VBOF. Analyses with the adapted VBOF could highlight additional potential targets for antiviral therapies.

As their coding capacity is limited, *Coronaviruses* strategically regulate host immune response, cell cycle, signaling and metabolism to create a favorable environment for viral replication[8,11,34]. Accordingly, the viroids depend on cellular enzymes for the formation of progeny, which makes host cell resources a potential target to limit virion production[11,34]. The viral hijacking of the cellular metabolic pathways, such as glycolysis, nucleotide and lipid biosynthesis, may shift the environment of the virus to a proliferation promoting environment[35]. In this study, we have also demonstrated that SARS-CoV-2 interferes with the host cell organisms, more precisely, the purine biosynthesis pathway to provide for the production of its biomass and, thus, replication.

In our study, we used a genome-scale metabolic model. This type of model enables the analysis of metabolic changes under certain constraints. However, only the metabolic changes can be investigated. Further network reconstructions, such as dynamic signaling, regulatory, or kinetic network models, can give further insight into changes in signal transduction, regulatory processes, or kinetic properties[36] of virus infections. Ravindran *et al.*, for example, analyzed the effect of human immunodeficiency virus 1 (HIV-1) and hepatitis C virus (HCV) infections using a large human signaling network[37]. They demonstrated how the infecting virus could bring the dynamically organized host system into its control. Tan *et al.* developed a mathematical model describing the virus-induced interferon (IFN) signaling process. Dynamic analysis and numerical simulations led to the suggestion that a balance between viral replication and IFN-induced regulation is responsible for the dynamic behavior of virus-triggered signaling and also for antiviral responses[38]. Dynamic modeling of infections with *Coronaviruses*, especially with SARS-CoV-2, could broaden the understanding of its effects on the host and give further insight into potential targets for antiviral therapies.

In this study, we used the already published *i*AB-AMØ-1410 GEM of human alveolar macrophages. This model does currently not include any genes or annotations. It is built upon the first human reconstruction Recon 1[39]. By now, the human reconstruction Recon3D is available with more than 10,000 reactions, 2000 genes and almost 6000 metabolites[40], including numerous annotations. Following the protocol of Bordbar *et al.*[15], a new model of the human alveolar macrophage could be generated based on the newer version of the human reconstruction, Recon3D. The newly curated model could then be used to verify the findings from this study and to identify further targets.

As *Coronaviruses* are reported to infect human alveolar macrophages, and a model for these cells was available, we integrated SARS-CoV-2 into this model. Nevertheless, SARS-CoV-2 is reported also to infect upper and lower respiratory tract cells, including pharyngeal regions[2,6]. Genome-scale metabolic models (GEMs) for bronchial epithelial cells (BEC) and airway epithelial cells (AEC) are already available. Wang *et al.* reconstructed 126 human tissue-specific GEMs using the metabolic Context-specificity Assessed by Deterministic Reaction Evaluation (mCADRE) algorithm. Those models are also built upon Recon 1. Furthermore, the models include fewer numbers of reactions (1242 and 1296, respectively) and lack a biomass maintenance function. Following the protocol for generating high-quality genome-scale metabolic reconstructions by Thiele and Palsson, meaningful models of human bronchial and airway epithelial cells could be generated[41]. However, generating a biomass maintenance function requires much data. Tools, such as BOFdat, can be beneficial for the generation of an appropriate biomass objective function (BOF)[42]. These models can then be used to verify the potential antiviral targets that were found in alveolar macrophages.

The integrated host-virus model suggested the supplementation of L-isoleucine and L-lysine as a potential target for antiviral therapies, as well as the inhibition of the guanylate kinase (GK1). Since compounds that directly inhibit GK1 are already known, their evaluation and verification in cell culture experiments are required for fast responses to the current spread of SARS-CoV-2 worldwide.

4 Methods

The methods used in this paper are based on the paper by Aller *et al.*[14]. Methods and analyses were adapted for the coronavirus SARS-CoV-2.

4.1 Generation of virus biomass objective function.

Analogous to the biomass production or maintenance function in prokaryotic or eukaryotic metabolic models, the virus biomass objective function (VBOF) is a pseudo-reaction simulating the production of virus particles. It comprises nucleotides, amino acids, and associated energy metabolites required for the production of the virus particles. Due to a lack of knowledge, the stoichiometric information of the virus envelope, and the dynamic information of lipids are not included in the VBOF. Hence, virus entry or exit cannot be modeled.

The generation of the VBOF was performed following the seven steps described by Aller *et al.*[14].

(1) Virus genome and protein information. The recently published virus genome and protein sequence were obtained from the NCBI database[43] with the accession number `NC_045512.2` in February 2020. Essential for the calculation of the nucleotide count is the classification of the SARS-CoV-2 in the Baltimore Classification System[44] that characterizes viruses based on the replication method. *Coronaviruses* fall into the Group IV classification. These viruses replicate their positive single-stranded RNA (+ssRNA) genomes via a complementary negative single-stranded RNA (-ssRNA). The nucleotide counts of the positive strand can be taken from the genome sequence. The nucleotides of the negative strand can be calculated by counting the complementary nucleotides of the positive strand. Both nucleotides are summed up to receive the total nucleotide count.

SARS-CoV-2 has a total of twelve proteins, four structural proteins, and eight non-structural proteins. Structural and non-structural proteins are not expressed equally, and this ratio needs to be considered during the calculation of the VBOF. To this date, no information about the copy number of structural proteins in *Coronaviruses* is reported.

(2) Nucleotide investment. The nucleotide count of the virus genome and its replication intermediate are summed per nucleotide. The genome copy number (Cg) was assumed to be one. According to Aller *et al*., the moles of nucleotides were converted into grams of nucleotides per mole of the virus. After similar calculations of the amino acids and the calculation of the total molar weight of the virus based on nucleotide and amino acid content, the stoichiometric coefficients of each nucleotide in the VBOF were calculated.

(3) Amino acid investment. Analogous to the nucleotide investment, the stoichiometric coefficient of each amino acid was calculated. Instead of the genome copy number, a copy number for structural and non-structural proteins is required. As already mentioned, the copy number of structural proteins is not reported to date. Therefore, we repeated all analyses described in this paper for copy numbers of structural proteins C_{sp} = 200, 500, 800 and 1200. With the information of the total count of each amino acid, weighted by the copy number of structural proteins, and the inclusion of the respective molar mass, the total viral molar mass was calculated (see also Step 6) to eventually calculate the stoichiometric coefficient for every amino acid.

(4) ATP requirements, and (5) Pyrophosphate (PPi) liberation. The calculations of the ATP requirements for the polymerization of amino acids and the calculation of pyrophosphate liberation from the nucleotide polymerization were performed as described by Aller *et al*. based on the results of the previous steps. The constants $k_{\mathrm{ATP}} = 4$ and $k_{\mathrm{PPi}} = 1$ were chosen, as suggested by Aller *et al*.. k_{ATP} is the required amount of four ATP molecules for the polymerization of amino acid monomers. The constant k_{PPi} stands for the liberation of one diphosphate molecule per polymerization of nucleotide monomers.

(6) Total viral molar mass and (7) Final construction of the VBOF. The total molar mass of SARS-CoV-2 and the final VBOF were calculated and constructed in accordance with Aller *et al*.

4.2 Integration of the SARS-CoV-2 Virus into *i*AB-AMØ-1410

Since Coronaviridae infect alveolar macrophages[12,13], the cell-specific GEM of human alveolar macrophages *i*AB-AMØ-1410 was utilized. This GEM was published in 2010 by Bordbar *et al*.[15] to give insight into human alveolar macrophages infected with *Mycobacterium tuberculosis*. The cell-specific GEM was constructed based on the human metabolic reconstruction Recon 1[39]. It consists of 3394 reactions and 2572 metabolites. The model was downloaded in SBML Level 2 Version 4 format[45] and upgraded to SBML Level 3 Version 1 format[46] using libSBML[47]. The VBOF was incorporated into the existing macrophage model, using flux bounds of 0 and 1000 as lower and upper bounds. Additionally, aggregated subsystems were added to the reactions as additional information for further analyses[14,48].

4.3 Identification of stoichiometric differences

The human alveolar macrophage model *i*AB-AMØ-1410 included now two pseudo-reactions for the production or maintenance of the virus and hosts biomass, respectively. The stoichiometric coefficients of shared metabolites within these pseudo-reactions can be compared to identify differences in nucleotide or amino acid requirements. First, the individual amino acid and nucleotide stoichiometric coefficients were normalized against the sum of all metabolites present in the respective biomass objective function, except for the metabolites for energy requirements. Then, the fold change (FC) of the normalized amino acid or nucleotide was calculated:

$$\mathrm{FC}_i = \log_2\left(\frac{S_i^V/\sum_k S_k^V}{S_i^H/\sum_k S_k^H}\right), \tag{2}$$

with index i over nucleotides or amino acids and k over all biomass precursors (except for energy requirements). Subscripts H and V represent the use of either the host or virus biomass function. Positive values imply higher usage of nucleotides or amino acids in the virus compared to the host, while negative values imply a lower usage.

4.4 FBA and FVA for the comparison of host- and virus-optimized states

The integration of the VBOF into the *i*AB-AMØ-1410 model paved the way for the analysis of metabolic changes between host and virus optimized states. To identify an optimal state, an objective function that is optimized needs to be defined. In general, every reaction in the GEM can serve as objective function. However, biologically meaningful objective functions depend on the research question. For analyses of the growth or survival of the studied organism or cell, biomass production or maintenance reactions can be introduced into the model as pseudo-reactions. In industrial settings, the production of a specific metabolite might be of interest, and hence its production reaction can be set as the objective function. The defined objective function is then optimized with flux balance analysis (FBA). FBA is a mathematical approach using linear optimization to analyze the flow of metabolites through a metabolic network while optimizing for an objective function[49]. This objective function is optimized under a set of constraints. These constraints are, on the one hand, defined by the stoichiometry of the reactions, and, on the other hand, by limitations of reaction fluxes through upper and lower bounds. Hence, not only the biomass production or maintenance reaction can be adapted to a specific organism or cell type, but also the environment, in which it occurs, can be adapted accordingly, e.g., by

constraining exchange fluxes. In this work, we optimized the host-virus integrated *i*AB-AMØ-1410 model for either the host biomass maintenance function or the VBOF using FBA.

Since the solutions calculated by FBA are often not unique, the flux variability analysis (FVA) is a method to identify alternate optimal solutions. In FVA, the maximum and minimum possible flux for each reaction in the network is evaluated while constraining the objective value equal or close to the optimal flux[49]. As for the FBA, the FVA was conducted for both the host-optimized and virus-optimized state. The results of the FVA were used in the subsequent host-derived enforcement to define the upper and lower bounds.

4.4.1 Copy number analysis

As the copy number of structural proteins in SARS-CoV-2 or *Coronaviruses*, in general, is not reported so far, we evaluated the effect of different C_{sp} values on the growth rate in the virus-optimized state. To do so, we varied the copy number C_{sp} between 1 to 1500. The variation of C_{sp} leads to changing stoichiometric coefficients in the VBOF. Since the VBOF is optimized in the virus-optimized state, changing VBOF reactions can lead to changing optimization results.

4.4.2 Optimization analysis

The host-virus integrated model holds both the host biomass maintenance and the VBOF reaction. First, each biomass reaction was optimized individually using FBA. The objective function is defined by setting the objective coefficient to 1. In general, only one reaction (the objective function) has an objective coefficient of 1. However, one can also set the objective coefficient of several reactions to 1 to optimize for both reactions. Hence, the objective coefficient of both biomass functions from host and virus was set to 1 to optimize for both functions. Last, we used COBRApy's[50] tailored objectives option to weight the influence of the two biomass reactions on the objective function. To do so, we created a linear combination of the two biomass reactions weighted by factors varying between 01, and not exceeding 1 in total. When the host maintenance function weighted 0.2, the VBOF had a weight of 0.8 in the objective function. For each combination, the model was optimized using FBA, and the flux through the two biomass reactions was reported. This analysis was repeated using the results from the host-derived enforcement analysis by adapting the bounds of the reported reactions (see section 4.5.2).

4.4.3 Metabolic analysis

The metabolic analysis was performed for the host- and virus-optimized states. The fluxes of each reaction in the two optimized states were compared by calculating the fold change (FC_r):

$$FC_r = \log_2\left(\frac{F_r^V}{F_r^H}\right), \tag{3}$$

where the indexation r is over all reactions of the model, and the superscript H indicates flux values from the host and V from the virus-optimized models. The fold changes were aggregated into the included subsystems.

4.5 Antiviral target identification

For the identification of potential antiviral targets that preferentially alter the virus growth rate while maintaining the hosts biomass maintenance, each reaction in the model was evaluated using two different approaches.

4.5.1 Reaction knock-out

The first approach is a reaction knock-out that is already implemented in COBRApy. The `single_gene_deletion` function subsequently sets both bounds of every reaction to zero and optimizes for the chosen objective function. We ran this function twice, once with the host maintenance reaction as the objective function and once with the VBOF. Possible targets were only reported when the growth of the virus was diminished (< 99 % of its initial growth rate) and when the growth rate of the virus was below the growth rate of the host-optimized state.

4.5.2 Host-derived enforcement

In the second approach, the effects of changes in flux ranges of the reactions on the VBOF are analyzed, while the metabolic system of the host-optimized state is maintained. To change the flux bounds ensuring maximal biomass maintenance of the host while optimizing for the VBOF, the results of the FVA were utilized. For each reaction, the flux range was defined as described by Aller *et al.*. The corresponding minimum (F^-) and maximum (F^+) flux values from the FVA of the host (H) and virus (V) optimization and their relation define the minimum (ε^-) and maximum (ε^+) flux values. Since the cases suggested by Aller *et al.* did not cover all reactions, we added the two more conditional arguments cases (4) and (8):

$$\begin{cases} \text{continue} & \text{if } F_H^+ = F_V^+ \wedge F_H^- = F_V^- \quad (4) \\ \varepsilon^+ = F_H^+ \\ \varepsilon^- = F_H^+ - \frac{F_H^+ - F_V^+}{2} & \text{if } F_H^+ > F_V^+ \wedge F_H^- \geq F_V^- \quad (5) \\ \varepsilon^+ = F_H^- - \frac{F_H^- - F_V^-}{2} \\ \varepsilon^- = F_H^- & \text{if } F_H^+ \leq F_V^+ \wedge F_H^- < F_V^- \quad (6) \\ \varepsilon^+ = F_H^+ \\ \varepsilon^- = F_H^- & \text{if } F_H^+ < F_V^+ \wedge F_H^- > F_V^- \quad (7) \\ \varepsilon^+ = F_H^+ - \frac{F_H^+ - F_V^+}{2} \\ \varepsilon^- = F_H^- - \frac{F_H^- - F_V^-}{2} & \text{if } F_H^+ \leq F_V^+ \wedge F_H^- \geq F_V^- \quad (8) \end{cases}$$

The flux values ε^- and ε^+ are set as upper and lower bounds of the corresponding reactions and the model was optimized for the VBOF. The resulting optimization result was compared to the original VBOF growth rate. When the growth rate with adapted bounds was below the threshold of 80 % of the initial growth rate, the reaction was reported as potential antiviral target.

5 Data availability

The computational model is accessible at `https://www.ebi.ac.uk/biomodels/MODEL2003020001`.

References

1. Hui, D. S. *et al.* The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int. J. Infect. Dis.* **91**, 264–266 (2020).
2. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* (2020).
3. Zhang, L. *et al.* Alpha-ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication Structure-based design, synthesis, and activity assessment. *J. Med. Chem.* (2020).

4. WHO. Coronavirus disease (covid-2019) situation report - 23. Accessed:2020-02-12.
5. Wu, Y. C., Chen, C. S. & Chan, Y. J. Overview of The 2019 Novel Coronavirus (2019-nCoV): The Pathogen of Severe Specific Contagious Pneumonia (SSCP). *J Chin Med Assoc* (2020).
6. Chen, Y., Liu, Q. & Guo, D. Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J. Med. Virol.* (2020).
7. Paraskevis, D. *et al.* Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. *Infect. Genet. Evol.* **79**, 104212 (2020).
8. Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* **7**, 439–450 (2009).
9. Chan, J. F. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* (2020).
10. Letko, M. & Munster, V. Functional assessment of cell entry and receptor usage for lineage b β-coronaviruses, including 2019-ncov. *bioRxiv* DOI: 10.1101/2020.01.22.915660 (2020). https://www.biorxiv.org/content/early/2020/01/22/2020.01.22.915660.full.pdf.
11. de Wilde, A. H. *et al.* A Kinome-Wide Small Interfering RNA Screen Identifies Proviral and Antiviral Host Factors in Severe Acute Respiratory Syndrome Coronavirus Replication, Including Double-Stranded RNA-Activated Protein Kinase and Early Secretory Pathway Proteins. *J. Virol.* **89**, 8318–8333 (2015).
12. Deng, X. *et al.* Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4251–E4260 (2017).
13. Cheung, C. Y. *et al.* Cytokine responses in severe acute respiratory syndrome coronavirus-infected macrophages in vitro: possible relevance to pathogenesis. *J. Virol.* **79**, 7819–7826 (2005).
14. 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. *J. Royal Soc. Interface* **15**, DOI: 10.1098/rsif.2018.0125 (2018).
15. Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. & Jamshidi, N. Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. *Mol. Syst. Biol.* **6**, DOI: 10.1038/msb.2010.68 (2010).
16. Mukhopadhyay, S., Kuhn, R. J. & Rossmann, M. G. A structural perspective of the Flavivirus life cycle, DOI: 10.1038/nrmicro1067 (2005).
17. Strauss, J. H. & Strauss, E. G. The alphaviruses: Gene expression, replication, and evolution, DOI: 10.1128/mmbr.58.3.491-562.1994 (1994).
18. Masters, P. S. The Molecular Biology of Coronaviruses, DOI: 10.1016/S0065-3527(06)66005-3 (2006).
19. Mortola, E. & Roy, P. Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Lett.* **576**, 174–178, DOI: 10.1016/j.febslet.2004.09.009 (2004).
20. Wang, C. *et al.* MERS-CoV virus-like particles produced in insect cells induce specific humoural and cellular imminity in rhesus macaques. *Oncotarget* **8**, 12686–12694, DOI: 10.18632/oncotarget.8475 (2017).
21. Guy, J. S., Breslin, J. J., Breuhaus, B., Vivrette, S. & Smith, L. G. Characterization of a coronavirus isolated from a diarrheic foal. *J. Clin. Microbiol.* **38**, 4523–4526, DOI: 10.1128/jcm.38.12.4523-4526.2000 (2000).
22. Fraenkel-Conrat, H. & Wagner, R. R. *Comprehensive Virology: Descriptive Catalogue of Viruses* (Springer US, 1974).
23. Hible, G., Daalova, P., Gilles, A. M. & Cherfils, J. Crystal structures of GMP kinase in complex with ganciclovir monophosphate and Ap5G. *Biochimie* **88**, 1157–1164 (2006).
24. Khan, N. *et al.* Solution structure and functional investigation of human guanylate kinase reveals allosteric networking and a crucial role for the enzyme in cancer. *J. Biol. Chem.* **294**, 11920–11933 (2019).
25. Navé, J. F., Eschbach, A. & Halazy, S. 9-(Phosphonoalkyl)guanine derivatives as substrates or inhibitors of guanylate kinase. *Arch. Biochem. Biophys.* **295**, 253–257 (1992).
26. Jain, R. *et al.* Insights into open/closed conformations of the catalytically active human guanylate kinase as investigated by small-angle X-ray scattering. *Eur. Biophys. J.* **45**, 81–89 (2016).
27. Navé, J. F. *et al.* Synthesis, antiviral activity and enzymatic phosphorylation of 9-phosphonopentenyl derivatives of guanine. *Antivir. Res.* **27**, 301–316 (1995).
28. Dunning, J. *et al.* Experimental Treatment of Ebola Virus Disease with Brincidofovir. *PLoS ONE* **11**, e0162199 (2016).
29. McCarthy, S. D. *et al.* A Rapid Screening Assay Identifies Monotherapy with Interferonγ- and combination Therapies with Nucleoside Analogs as Effective Inhibitors of Ebola Virus. *PLoS Negl Trop Dis* **10**, e0004364 (2016).
30. Arias, A., Thorne, L. & Goodfellow, I. Favipiravir elicits antiviral mutagenesis during virus replication in vivo. *Elife* **3**, e03679 (2014).
31. Garcia, C. R., Torriani, F. J. & Freeman, W. R. Cidofovir in the treatment of cytomegalovirus (CMV) retinitis. *Ocul. Immunol. Inflamm.* **6**, 195–203 (1998).
32. Timm, A. & Yin, J. Kinetics of virus production from single cells. *Virology* **424**, 11–17, DOI: 10.1016/j.virol.2011.12.005 (2012).
33. Yan, B. *et al.* Characterization of the lipidomic profile of human coronavirus-infected cells: Implications for lipid metabolism remodeling upon coronavirus replication. *Viruses* **11**, DOI: 10.3390/v11010073 (2019).
34. Dirmeier, S. & D. .
35. Thaker, S. K., Ch'ng, J. & Christofk, H. R. Viral hijacking of cellular metabolism. *BMC Biol.* **17**, 59 (2019).
36. Palsson, B. Ø. *Systems biology: simulation of dynamic network states* (Cambridge University Press, 2011).
37. Ravindran, V. *et al.* Network controllability analysis of intracellular signalling reveals viruses are actively controlling molecular systems. *Sci. Reports* **9**, 1–11, DOI: 10.1038/s41598-018-38224-9 (2019).
38. Tan, J., Pan, R., Qiao, L., Zou, X. & Pan, Z. Modeling and Dynamical Analysis of Virus-Triggered Innate Immune Signaling Pathways. *PLoS ONE* **7**, e48114, DOI: 10.1371/journal.pone.0048114 (2012).
39. Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc.*

Natl. Acad. Sci. United States Am. **104**, 1777–82, DOI: 10.1073/pnas.0610772104 (2007).

40. Brunk, E. *et al.* Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272–281, DOI: 10.1038/nbt.4072 (2018).

41. Thiele, I. & Palsson, B. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121, DOI: 10.1038/nprot.2009.203 (2010).

42. Lachance, J.-C. *et al.* BOFdat: generating biomass objective function stoichiometric coefficients from experimental data. *bioRxiv* 243881, DOI: 10.1101/243881 (2018).

43. Geer, L. Y. *et al.* The NCBI BioSystems database. *Nucleic Acids Res.* **38**, D492–D496, DOI: 10.1093/nar/gkp858 (2010).

44. Baltimore, D. Expression of animal virus genomes. *Bacteriol. reviews* **35**, 235–241, DOI: 10.1128/mmbr.35.3.235-241.1971 (1971).

45. Hucka, M., Finney, A. M., Keating, S. M., Le Novère, N. & Hoops, S. Systems biology markup language (SBML) Level 2: structures and facilities for model definitions. *Nat. Preced.* **58**, DOI: doi:10.1038/npre.2008.2715.1 (2008).

46. Hucka, M. *et al.* Systems Biology Markup Language (SBML) Level 3 Version 1 Core. *J. Integr. Bioinforma.* **15**, 1, DOI: 10.1515/jib-2017-0080 (2018).

47. Bornstein, B. J., Keating, S. M., Jouraku, A. & Hucka, M. LibSBML: an API Library for SBML. *Bioinformatics* **24**, 880–881, DOI: 10.1093/bioinformatics/btn051 (2008).

48. Zielinski, D. C. *et al.* Pharmacogenomic and clinical data link non-pharmacokinetic metabolic dysregulation to drug side effect pathogenesis. *Nat. Commun.* **6**, DOI: 10.1038/ncomms8101 (2015).

49. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. biotechnology* **28**, 245–8, DOI: 10.1038/nbt.1614 (2010).

50. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74, DOI: 10.1186/1752-0509-7-74 (2013).

6 Acknowledgements

This work was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2124 – 390838134 and is supported by supported by the DZIF (German Center for Infection Research).

7 Author contributions statement

A.R. and A.D. conceived the study. A.R. and L.W. constructed the model. A.R. derived the biomass objective function and conducted the analysis with A.D. A.R., L.W., and A.D. wrote the manuscript. A.D. supervised the study. All authors reviewed and approved the final manuscript.

8 Additional information

Competing interests: none.