Title: Biophysical modeling of the SARS-CoV-2 viral cycle reveals ideal antiviral targets

Authors: Brian T. Castle[1], Carissa Dock[1], Mahya Hemmat[1], Susan Kline[2], Christopher Tignanelli[3], Radha Rajasingham[2], David Masopust[4], Paolo Provenzano[1], Ryan Langlois[4], Timothy Schacker[2], Ashley Haase[4], David J. Odde[1,*]

Affiliations:

[1]Department of Biomedical Engineering, University of Minnesota, Minneapolis, MN 55455

[2]Department of Medicine, Division of Infectious Disease and International Medicine, University of Minnesota, Minneapolis, MN 55455

[3]Department of Surgery, Institute for Health Informatics, University of Minnesota, Minneapolis, MN 55455

[4]Department of Microbiology and Immunology, Center for Immunology, University of Minnesota, Minneapolis, Minnesota, 55414

Correspondence:

[*]oddex002@umn.edu, (612) 626-9980

Abstract:

Effective therapies for COVID-19 are urgently needed. Presently there are more than 800 COVID-19 clinical trials globally, many with drug combinations, resulting in an empirical process with an enormous number of possible combinations. To identify the most promising potential therapies, we developed a biophysical model for the SARS-CoV-2 viral cycle and performed a sensitivity analysis for individual model parameters and all possible pairwise parameter changes ($16^2 = 256$ possibilities). We found that model-predicted virion production is fairly insensitive to changes in viral entry, assembly, and release parameters, but highly sensitive to some viral transcription and translation parameters. Furthermore, we found a cooperative benefit to pairwise targeting of transcription and translation, predicting that combined targeting of these processes will be especially effective in inhibiting viral production.

Main Text:

The ongoing global pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in more than 4 million confirmed cases and 300,000 deaths worldwide (*1*). The most common symptoms of the illness caused by SARS-CoV-2, COVID-19, include fever, cough and fatigue (*2*). The clinical presentation can range from asymptomatic to fatal, with severe cases rapidly progressing to pneumonia, acute respiratory distress syndrome (ARDS), and organ failure (*3*). There is currently no vaccine for the disease, and development could take 12 to 18 months (*4*). Therefore, there exists a critical need for effective therapeutic interventions to minimize the transmission and severity of SARS-CoV-2.

Despite increased expenditures on research and development, less than one out of every ten therapeutic drugs that enter phase I clinical trials eventually gains FDA approval (*5*). To accelerate the development process for COVID-19, the initial focus has been on repurposing approved drugs and biologics. However, owing to the large number of possible therapeutics, and the need for rapid testing, there has not been extensive preclinical testing on SARS-CoV-2 specifically. While these therapeutics have been deployed in combinations of two, three, or four in more than 800 clinical trials globally (*6*), it is not clear which drugs, possibly in combination with others, would, in principle, be most effective. For other RNA viruses, such as HIV, single therapies have not been successful as a result of the virus' ability to rapidly evolve and develop resistance to antivirals (*7*), thus driving the need for combination therapies (*8*). With more than 100 distinct agents currently in trials to treat COVID-19, even a two drug combination has over 10,000 possible combinations that could be tried, raising the question of how to rationally focus clinical trials on the single agents and combinations that are most likely to be effective.

Biophysical modeling has the potential to help rationally guide the development of therapeutic interventions for SARS-CoV-2 by identifying key model parameters for effective targeting. In addition, modeling can potentially be used to identify combination therapies, predict clinical trial outcomes, stratify patients, and identify potential source(s) of variable patient-to-patient outcomes. Here we present a biophysical model for the SARS-CoV-2 viral cycle and identify the single and combination parameters with the highest sensitivity, representing the ideal targets for therapeutic intervention to inhibit virus production.

The biophysical model of the SARS-CoV-2 viral life cycle was constructed based on SARS-CoV literature that included processes underlying viral entry, genome transcription, genome translation, virion assembly, and virion release (Fig. 1; reviewed in (*9*, *10*)). Mass action and chemical rate equations were used to mathematically describe the system using an approach similar to that taken previously to model the life cycle of other viral systems (reviewed in (*11*); see Materials and Methods), and the resulting series of ordinary differential equations solved numerically. As an initial condition we assumed that a single virion was bound to the surface of an individual cell, which was then internalized to initiate the replication cycle. As data specific to the viral cycle to SARS-CoV-2 are limited, we primarily used experimental observations from SARS-CoV to inform model assumptions and parameterization (Table 1). We assume that it is reasonable to use SARS-CoV data based on the high degree of genomic similarity between the two viruses (*12*–*14*). Using the base parameter values, the model reproduces viral production on a timescale and of a quantity consistent with experimental observations (Fig. 2). Importantly, the model reproduces a number of experimental observations without parameter adjustment (Fig. 2B); for example, the model naturally predicts that ~10% of the RNA will be negative sense (Fig. 2). Based on this, we conclude that the model provides a suitable tool to identify points of interest for therapeutic intervention, i.e. those parameters and the associated subprocesses that are particularly sensitive to perturbation.

To assess these points of interest in the viral cycle, we performed a sensitivity analysis for each parameter by systematically increasing and decreasing their values from baseline by up to three orders of magnitude (1000-fold) while holding all other parameters constant (Fig. 3). Parameters specific to virion stoichiometry or RNA segment length were ignored in this analysis as they are likely difficult to target clinically. Using the viral cycle time t_{cycle}, defined as the time to produce 1000 virions, as a readout of viral replication rate, we examined viral production as a function of individual parameter values. We found that parameter perturbation resulted in a variety of responses, ranging from insensitive to highly sensitive (Fig. 3A-C). As seen in Figure 3D, viral production was most sensitive to perturbation of many, but not all, parameters related to genome transcription and translation. Parameters related to viral entry, packaging, and release were comparably insensitive to changes from the base values (Fig. 3D). While it is theoretically possible to inhibit viral production via targeting of these insensitive parameters, it is necessary to achieve a high level of inhibition compared to those that exhibit high sensitivity (Fig. S1). For example, parameters specific to viral entry and packaging required 100-1000x

level of inhibition in order to influence viral production (Figs. 3D and S1A, C), while parameters specific to replication completely inhibited viral production after only a 10x effect (Figs. 3D and S1B). Based on these observations, we conclude that transcription and translation represent high sensitivity targets for therapeutic inhibition.

Due to the ability to rapidly evolve, viral diseases are often treated with combination therapies. While an exhaustive combinatorial analysis is difficult or impossible to conduct experimentally due to time and resource limitations, it can quickly be implemented *in silico*. To identify potential combinations that could cooperate to inhibit viral production in the model, we performed a pairwise sensitivity analysis of the model parameters. Parameter pairs were either coordinately increased or decreased from their base value and then scored based on the results (see Materials and Methods). To score parameter combinations we defined a sensitivity (S) and range (R) value (Fig. 4A), similar to our previous approach for another biophysical model (*15*). The sensitivity S is a measure of how viral production scales with changes to the base parameter values, while R is a measure of the magnitude of perturbation necessary to produce a maximal effect. Ideally, interventional therapies would have a strong effect (large S) with minimal perturbation (small R). Therefore, we scored parameter pairs by dividing the resulting S value by the value of R (S/R). As seen in Figures 4B and S2, several parameter combinations completely eliminated viral production with only a <10-fold effect. Similar to the single parameter sensitivity analysis, parameters related to transcription and translation once again emerged as the highest sensitivity targets; the highest scores were parameter combinations specific to the host ribosomes and the viral RNA-dependent RNA polymerase (RdRp). Other noteworthy pairs, which completely eliminated viral production, were the rate of proteolytic cleavage of pp1a/1ab by the main protease (k_{Cleav}) and the half-life of pp1a/1ab (h_{pp1}) paired with the binding rate of the RdRp (k_{on}^{RdRp}). Interestingly, the majority of high scoring combinations were not simply the result of targeting two high sensitivity parameters, but rather resulted from cooperating effects between the two parameters (Fig. 4C). Based on these results, we conclude combination targeting of transcription and translation can lead to enhanced effects on viral production, beyond targeting each individually.

Overall, our modeling of the SARS-CoV-2 life cycle, parameterized using published SARS-CoV literature, shows that theoretically there are opportunities for therapeutic interventions that significantly inhibit the viral cycle. In particular, the sensitivity analysis identified several parameters in the middle of the viral cycle, specific to genome transcription and translation, that

present the best opportunities for inhibiting viral production. By comparison, parameters specific to viral entry, virion assembly, and virion release were less sensitive and therefore are less promising as targets for inhibiting viral production. The model further identifies potential combination targets that would cooperatively inhibit viral production. For example, the combined effects of targeting both the stepping rate of the host ribosome, k_{Step}^{Rib}, and the binding rate of RdRp, k_{on}^{RdRp}, may halt the viral cycle even with modest 10X effects. Such a pairwise analysis would be difficult to exhaustively test experimentally, especially in the clinic, since there are $16^2 = 256$ possible combinations.

In addition to identifying novel target opportunities within the viral cycle, biophysical modeling may also provide insight into current and emerging therapeutic approaches. Several existing antiviral drugs are being evaluated for their efficacy in treating COVID-19 (*16*). It is interesting to note that remdesivir, which was recently approved for emergency use in the U.S. and Japan, acts to disrupt genome transcription by interfering with the viral RdRp (*17*, *18*), and thus is one example of a therapeutic intervention that interferes with an area of high sensitivity identified by the model. By contrast, hydroxychloroquine, which inhibits viral entry by increasing endosomal pH and affecting terminal glycosylation of the ACE2 receptor (*19*), is predicted by the model to be less likely to be effective since drugs that act on viral entry would require exceptionally high suppression to achieve appreciable effects on viral production. Viral entry inhibitors have been used to treat HIV (*20*) and influenza (*21*), therefore it may yet be feasible to target viral entry. Additionally, we note that therapeutics predicted to be less effective in the viral cycle model, may have potent effects elsewhere in disease progression, e.g. in the immune response. Thus, ineffectiveness in inhibiting viral production does not preclude therapeutic effectiveness overall.

Finally, the recent host-virus protein interaction map reported by Gordon et al. (*22*) identified two classes of SARS-CoV-2 targeting opportunities: 1) inhibitors of protein translation and 2) regulators of Sigma1 and Sigma2 receptors. Based on our modeling, we predict that therapeutic targeting of the Sigma1 and 2 receptors is not likely to be effective as it would presumably interfere with the virion assembly and release steps, which are relatively insensitive (Fig. 3D). By contrast, we predict that targeting protein translation is more likely to be effective due to the high sensitivity of viral cycle time to translation-associated parameters (Fig. 3D). Everything else being equal, the model predicts that transcription inhibition combined with translation inhibition would be an especially effective combination (Fig. 4B-C). Altogether, our model

provides a framework for understanding viral cycle dynamics and identifying the therapeutic opportunities that are most likely to be effective in inhibiting viral production.

References:

1. "Coronavirus disease (COVID-19) Highlights," (available at https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200518-covid-19-sitrep-119.pdf?sfvrsn=4bd9de25_4).
2. N. Chen *et al.*, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. **395**, 507–513 (2020).
3. X. Yang *et al.*, Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir. Med.* **8** (2020), doi:10.1016/S2213-2600(20)30079-5.
4. F. Amanat, F. Krammer, SARS-CoV-2 Vaccines: Status Report. *Immunity*. **52** (2020), pp. 583–589.
5. T. Takebe, R. Imai, S. Ono, The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development. *Clin. Transl. Sci.* **11**, 597–606 (2018).
6. ClinicalTrials.gov, (available at https://clinicaltrials.gov/ct2/results?cond=COVID-19&Search=Apply&recrs=b&recrs=a&recrs=f&recrs=d&recrs=e&age_v=&gndr=&type=Intr&rslt=).
7. B. R. Edlin *et al.*, In-vitro resistance to zidovudine and alpha-interferon in HIV-1 isolates from patients: Correlations with treatment duration and response. *Ann. Intern. Med.* **117**, 457–460 (1992).
8. R. M. Gulick *et al.*, Treatment with Indinavir, Zidovudine, and Lamivudine in Adults with Human Immunodeficiency Virus Infection and Prior Antiretroviral Therapy. *N. Engl. J. Med.* **337**, 734–739 (1997).
9. J. Ziebuhr, Molecular biology of severe acute respiratory syndrome coronavirus. *Curr. Opin. Microbiol.* **7**, 412–419 (2004).
10. T. S. Fung, D. X. Liu, Human Coronavirus: Host-Pathogen Interaction. *Annu. Rev. Microbiol.* **73**, 529–557 (2019).
11. J. Yin, J. Redovich, Kinetic Modeling of Virus Growth in Cells. *Microbiol. Mol. Biol. Rev.* **82** (2018), doi:10.1128/mmbr.00066-17.
12. Y. M. Bar-On, A. Flamholz, R. Phillips, R. Milo, Sars-cov-2 (Covid-19) by the numbers.

Elife. **9** (2020), doi:10.7554/eLife.57309.

13. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. **579**, 270–273 (2020).
14. C. Wu *et al.*, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* (2020), doi:10.1016/J.APSB.2020.02.008.
15. B. L. Bangasser, S. S. Rosenfeld, D. J. Odde, Determinants of Maximal Force Transmission in a Motor-Clutch Model of Cell Traction in a Compliant Microenvironment. *Biophys. J.* **105**, 581–592 (2013).
16. M. Wang *et al.*, Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* **30** (2020), pp. 269–271.
17. T. P. Sheahan *et al.*, Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci. Transl. Med.* **9** (2017), doi:10.1126/scitranslmed.aal3653.
18. M. L. Agostini *et al.*, Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. *MBio*. **9** (2018), doi:10.1128/mBio.00221-18.
19. M. J. Vincent *et al.*, Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol. J.* **2** (2005), doi:10.1186/1743-422X-2-69.
20. J. A. Esté, A. Telenti, HIV entry inhibitors. *Lancet*. **370**, 81–88 (2007).
21. J. A. Oates, A. J. J. Wood, R. G. Douglas, Prophylaxis and Treatment of Influenza. *N. Engl. J. Med.* **322**, 443–450 (1990).
22. D. E. Gordon *et al.*, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, 1–13 (2020).
23. L. Dong, S. Hu, J. Gao, Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discov. Ther.* **14**, 58–60 (2020).
24. W. Dai *et al.*, Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science (80-.).* (2020), doi:10.1126/SCIENCE.ABB4489.
25. A. Bar-Even *et al.*, The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*. **50**, 4402–4410 (2011).
26. M. Dundr *et al.*, A kinetic framework for a mammalian RNA polymerase in vivo. *Science (80-.).* **298**, 1623–1626 (2002).
27. M. L. Ng, S. H. Tan, E. E. See, E. E. Ooi, A. E. Ling, Early events of SARS coronavirus infection in vero cells. *J. Med. Virol.* **71**, 323–331 (2003).
28. S. H. Northrup, H. P. Erickson, Kinetics of protein-protein association explained by

Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 3338–42 (1992).

29. N. Hirano, K. Fujiwara, M. Matumoto, Mouse Hepatitis Virus (MHV-2). *Jpn. J. Microbiol.* **20**, 219–225 (1976).
30. R. Milo, R. Phillips, *Cell Biology by the Numbers* (2015; http://book.bionumbers.org/).
31. A. J. W. te Velthuis, J. J. Arnold, C. E. Cameron, S. H. E. van den Worm, E. J. Snijder, The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. *Nucleic Acids Res.* **38**, 203–214 (2009).
32. S. M *et al.*, Severe Acute Respiratory Syndrome Coronavirus Replication Is Severely Impaired by MG132 Due to Proteasome-Independent Inhibition of M-calpain. *J. Virol.* **86** (2012), doi:10.1128/JVI.01001-12.
33. S. Stertz *et al.*, The intracellular sites of early replication and budding of SARS-coronavirus. *Virology*. **361**, 304–315 (2007).
34. B. Nal *et al.*, Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J. Gen. Virol.* **86**, 1423–1434 (2005).
35. S. G. Sawicki, D. L. Sawicki, S. G. Siddell, A Contemporary View of Coronavirus Transcription. *J. Virol.* **81**, 20–29 (2007).
36. N. Oh, J. H. Park, Endocytosis and exocytosis of nanoparticles in mammalian cells. *Int. J. Nanomedicine*. **9** (2014), pp. 51–63.
37. B. Alberts *et al.*, *Molecular biology of the cell* (Garland Science, ed. 5th, 2007).

Acknowledgements:

The authors thank David Largaespada and Jonathan Sachs for helpful discussions. **Funding:** The study was supported by the University of Minnesota Institute for Engineering in Medicine Medtronic Professorship held by D.J.O., and a COVID-19 Rapid Response Grant to D.J.O. from the University of Minnesota Medical School. **Author contributions:** All authors contributed to the development and the design of the model. B.T.C. wrote the computer code and performed all analyses. B.T.C, C.D., and D.J.O wrote the manuscript. All authors commented on and contributed to the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All computer code developed in this manuscript will be made available upon reasonable request to the corresponding author.

Materials and Methods:

The system of ODEs describing the viral cycle was solved using the *ode15s* solver in Matlab (R2019a; The Mathworks, Natick, MA).

Viral entry

Upon initialization the model is assumed to have a single viral particle (SARS-CoV-2) bound to angiotensin-converting enzyme 2 (ACE2) on the cell membrane (SCoV2-ACE2). SARS-CoV-2 is internalized into endosomes as the spike protein is cleaved by transmembrane protease, serine 2 (TMPRSS2) at a first order-rate (k_{TMP}).

$$\frac{d[SCoV2-ACE2]}{dt} = -k_{TMP}[SCoV2 - ACE2] \quad \text{(Eq. 1)}$$

This reaction and other proteolytic reactions in the model are simplified to the product of a first-order rate constant k and the concentration of the enzyme ($[E]$) where $k = (k_{cat}/K_M)[E]_{SS}$. This assumption is true when the substrate concentration is much less than the Michaelis-Menten constant value, $[S] << K_M$. For each reaction we assumed a catalytic efficiency (k_{cat}/K_M) equal to 150,000 $M^{-1}s^{-1}$, the median value for a range of enzymes (*25*). The steady-state concentration of the enzyme, $[E]_{SS}$, was estimated using the baseline parameter values in Table 1. Endosomal SARS-CoV-2 is then released and the genomic RNA uncoated with the first-order rates k_{Rel} and k_{UC}, respectively.

$$\frac{d[endoSCoV2]}{dt} = k_{TMP}[SCoV2 - ACE2] - k_{Rel}[endoSCoV2] \quad \text{(Eq. 2)}$$

$$\frac{d[cgRNA]}{dt} = k_{Rel}[endoSCoV2] - k_{UC}[cgRNA] \quad \text{(Eq. 3)}$$

Transcription and translation

After the full-length genomic RNA (gRNA) is released and uncoated, host ribosomes bind to open reading frames to begin translation of pp1a and pp1ab. In each instance, transcription and translation is modeled as a multistep process (Figure 1; inset), similar to that outlined in (*26*).

- **i.** Initial binding of the enzyme (Ribosome or RNA-dependent RNA polymerase) to its respective substrate (RNA) is assumed to be irreversible and is the product of the second-order association rate (k_{on}) and the concentrations of the enzyme and its substrate.
- **ii.** The enzyme completes the primer sequence at rate k_{Prime}, which is determined by the stepping rate of the respective enzyme (k_{Step}) and the length of the primer sequence (l); $k_{Prime} = k_{Step}/l$. This first step is fast (~0.5-1 s), and once completed the substrate is

released such that multiple enzymes can bind to a single strand of RNA. The actively polymerizing translation or transcription complex (T-SE) then continues through the elongation process.

iii. For each active T-SE, the rate of termination (k_{Term}) is proportional to the length (l) of the sequence and the stepping rate; $k_{Term} = k_{Step}/l$. Upon termination, the enzyme unbinds from the substrate and the product is released.

For the specific example of translating pp1a/1ab by the host ribosomes, this multistep process results in the following equations

$$\frac{d[Rib-gRNA]}{dt} = k_{on}^{Rib}[Rib][gRNA] - k_{Prime}^{Rib}[Rib - gRNA] \quad \text{(Eq. 4)}$$

$$\frac{d[T-RibgRNA]}{dt} = k_{Prime}^{Rib}[Rib - gRNA] - k_{Term}^{Rib}[T - RibgRNA] \quad \text{(Eq. 5)}$$

$$k_{Prime}^{Rib} = k_{Step}^{Rib}/l_{Prime}^{Rib} \quad \text{(Eq. 6)}$$

$$k_{Term}^{Rib} = k_{Step}^{Rib}/l_{pp1} \quad \text{(Eq. 7)}$$

Additionally, the polypeptides pp1a/1ab are then cleaved at first-order rate k_{Cleav} to release the non-structural proteins making up the replicase-transcriptase complex, specifically the RNA-dependent RNA polymerase (RdRp), and is degraded with a half-life of h_{pp1}.

$$\frac{d[pp1]}{dt} = k_{Term}^{Rib}[T - RibgRNA] - \left(k_{Cleav} + \frac{\ln(2)}{h_{pp1}}\right)[pp1] \quad \text{(Eq. 8)}$$

The multistep process outlined above repeats for production of negative-sense RNA (nRNA) templates,

$$\frac{d[RdRp-gRNA]}{dt} = k_{on}^{RdRp}[RdRp][gRNA] - k_{Prime}^{RdRp}[RdRp - gRNA] \quad \text{(Eq. 9)}$$

$$\frac{d[T-RdRpgRNA]}{dt} = k_{Prime}^{RdRp}[RdRp - gRNA] - k_{Term}^{RdRp}[T - RdRpgRNA] \quad \text{(Eq. 10)}$$

$$k_{Prime}^{RdRp} = k_{Step}^{RdRp}/l_{Prime}^{RdRp} \quad \text{(Eq. 11)}$$

$$k_{Term}^{RdRp} = k_{Step}^{RdRp}/l_{gRNA} \quad \text{(Eq. 12)}$$

The resulting nRNA templates are then used to transcribe the full-length gRNA and sub-genomic RNA (sgRNA) encoding the structural proteins. Additionally, nRNA is degraded with a half-life of h_{nRNA}

$$\frac{d[nRNA}{dt} = k_{Term}^{RdRp}[T - RdRpgRNA] - \left(k_{on}^{RdRp}[RdRp] + \frac{\ln(2)}{h_{nRNA}}\right)[nRNA] + k_{Prime}^{RdRp}[RdRp - nRNA] \quad \text{(Eq. 13)}$$

$$\frac{d[RdRp-nRNA]}{dt} = k_{on}^{RdRp}[RdRp][nRNA] - k_{Prime}^{RdRp}[RdRp - nRNA] \quad \text{(Eq. 14)}$$

$$\frac{d[T-RdRpnRNA]}{dt} = k_{Prime}^{RdRp}[RdRp-nRNA] - \rho\left(k_{Term}^{RdRp}[T-RdRpnRNA]\right) - \cdots$$
$$(1-\rho)\left(k_{Term,sg}^{RdRp}[T-RdRpnRNA]\right) \quad \text{(Eq. 15)}$$

$$\frac{d[RdRp]}{dt} = k_{Cleav}[pp1] - k_{on}^{RdRp}[RdRp][gRNA] + k_{Term}^{RdRp}[T-RdRpgRNA] - k_{on}^{RdRp}[RdRp][nRNA] + \ldots$$
$$\rho\left(k_{Term}^{RdRp}[T-RdRpnRNA]\right) + (1-\rho)\left(k_{Term,sg}^{RdRp}[T-RdRpnRNA]\right) \quad \text{(Eq. 16)}$$

Here, ρ is the fraction of transcription complexes producing full-length gRNA as opposed to sgRNA. sgRNA is produced at a faster rate due to the shorter sequence length such that

$$k_{Term,sg}^{RdRp} = k_{Step}^{RdRp}/l_{sgRNA} \quad \text{(Eq. 17)}$$

Where l_{sgRNA} is the average length of sgRNA coding the structural proteins, which is weighted by the relative length and stoichiometry of each protein. This then completes the cycle for replicating the full-length gRNA. An additional term is added for the coating of the genome with nucleocapsid (N) protein in preparation for encapsulation in the mature virion, where k_{NCap} is the second-order rate of producing nucleocapsids.

$$\frac{d[gRNA]}{dt} = k_{UC}[cgRNA] - \left(k_{on}^{Rib}[Rib] + k_{NCap}[N] + \frac{\ln(2)}{h_{gRNA}}\right)[gRNA] + k_{Prime}^{Rib}[Rib-gRNA] + \ldots$$
$$k_{Prime}^{RdRp}[RdRp-gRNA] + \rho\left(k_{Term}^{RdRp}[T-RdRpnRNA]\right) \quad \text{(Eq. 18)}$$

The final multistep process is the translation of sgRNA by host ribosomes to produce the structural proteins making up the mature virion, the spike (S), envelope (E), membrane (M), and nucleocapsid protein (N). Additionally, sgRNA is degraded with the same half-life as gRNA (h_{gRNA}).

$$\frac{d[sgRNA]}{dt} = (1-\rho)\left(k_{Term,sg}^{RdRp}[T-RdRpnRNA\right) - \left(k_{on}^{Rib}[Rib] + \frac{\ln(2)}{h_{gRNA}}\right)[sgRNA] + \cdots$$
$$k_{Prime}^{Rib}[Rib-sgRNA] \quad \text{(Eq. 19)}$$

$$\frac{d[Rib-sgRNA]}{dt} = k_{on}^{Rib}[Rib][sgRNA] - k_{Prime}^{Rib}[Rib-sgRNA] \quad \text{(Eq. 20)}$$

$$\frac{d[T-RibsgRNA]}{dt} = k_{Prime}^{Rib}[Rib-sgRNA] - \lambda\left(k_{Term,SEM}^{Rib}[T-RibsgRNA]\right) - \cdots$$
$$(1-\lambda)\left(k_{Term,N}^{Rib}[T-RibsgRNA]\right) \quad \text{(Eq. 21)}$$

$$\frac{d[N]}{dt} = (1-\lambda)\left(k_{Term,N}^{Rib}[T-RibsgRNA]\right) - a\left(k_{NCap}[N][gRNA]\right) \quad \text{(Eq. 22)}$$

$$\frac{d[Rib]}{dt} = -k_{on}^{Rib}[Rib]([gRNA]+[sgRNA]) + k_{Term}^{Rib}[T-RibgRNA] + \cdots$$
$$\lambda\left(k_{Term,SEM}^{Rib}[T-RibsgRNA]\right) + (1-\lambda)\left(k_{Term,N}^{Rib}[T-RibsgRNA]\right) \quad \text{(Eq. 23)}$$

where

$$k_{Term,SEM}^{Rib} = k_{Step}^{Rib}/l_{SEM} \quad \text{(Eq. 24)}$$

$$k_{Term,N}^{Rib} = k_{Step}^{Rib}/l_{N} \quad \text{(Eq. 25)}$$

In the above equations, λ is the ratio of producing S, E, or M proteins relative to N protein. The value of λ is determined by the stoichiometries of each. Furthermore, the translation rate of S, E, or M protein is assumed to be proportional to the final stoichiometry in the mature virion, and therefore they can be lumped together, simplifying the handling equations.

Virion assembly and release

S, E, and M proteins are polymerized in the cytoplasm but are then translocated through the endoplasmic reticulum (ER) and ER-golgi intermediate complex (ERGIC) at first-order rates $k_{Cyt\text{-}ER}$ and $k_{ER\text{-}ERGIC}$, respectively.

$$\frac{d[SEM_{cyto}]}{dt} = \lambda\left(k_{Term,SEM}^{Rib}[T - RibsgRNA]\right) - k_{Cyt-ER}[SEM_{cyto}] \quad \text{(Eq. 26)}$$

$$\frac{d[SEM_{ER}]}{dt} = k_{Cyt-ER}[SEM_{cyto}] - k_{ER-ERGIC}[SEM_{ER}] \quad \text{(Eq. 27)}$$

Budding from the ERGIC at rate k_{Bud} forms the mature virion from the encapsulation of nucleocapsids. Mature virions are then released via exocytosis at first-order rate k_{Exo}.

$$\frac{d[SEM_{ERGIC}]}{dt} = k_{ER-ERGIC}[SEM_{ER}] - c(k_{Bud}[SEM_{ERGIC}][NCap]) \quad \text{(Eq. 28)}$$

$$\frac{d[NCap]}{dt} = k_{NCap}[N][gRNA] - d(k_{Bud}[SEM_{ERGIC}][NCap]) \quad \text{(Eq. 29)}$$

$$\frac{d[Vir]}{dt} = k_{Bud}[SEM_{ERGIC}][NCap] - k_{Exo}[Vir] \quad \text{(Eq. 30)}$$

Sensitivity analysis

To initially assess model sensitivity, we used the viral cycle time t_{cycle}, defined as the time to produce 1000 virions, as a measure of viral production; the longer the cycle time, the lower the viral production. For single parameter changes (Figure 3), individual parameters were adjusted via multiplication by a scaling constant (c), which was assigned values between 0.001 and 1000 on a logarithmic scale.

$$p = cp_0 \quad \text{(Eq. 31)}$$

where p_0 is the baseline value of the specific parameter.

Pairwise sensitivity analysis was performed using a similar approach to that described in (*15*), where a sensitivity (S) and range (R) value were assigned based on how model output scales

with changes in parameter values. To perform the pairwise analysis, we defined viral production level by

$$v_p = \frac{t_{cycle,0}}{t_{cycle,ij}} \quad \text{(Eq. 32)}$$

where $t_{cycle,0}$ is the baseline cycle time and $t_{cycle,ij}$ is the cycle time resulting after perturbation of the paired parameters. Parameter combinations were scaled using a constant (c), which was assigned values between 0.1 and 1 on a logarithmic scale. Both parameters were scaled down by multiplication with the constant

$$p_i = cp_{i,0} \quad \text{(Eq. 33)}$$

$$p_j = cp_{j,0} \quad \text{(Eq. 34)}$$

or scaled up by dividing by the same constant value.

$$p_i = \frac{p_{i,0}}{c} \quad \text{(Eq. 35)}$$

$$p_j = \frac{p_{j,0}}{c} \quad \text{(Eq. 36)}$$

A sensitivity value was assigned by measuring the change in viral production as a function of the change in the parameter values.

$$S = \frac{d\, v_p}{d \log(c)} \quad \text{(Eq. 37)}$$

Thus, S is the slope of the plot of v_p versus $\log(c)$ (Figure 5A). Assigning values using Eqs. S30 and S35 resulted in positive values for parameters that inhibited viral production, increasing cycle time, and negative values for parameter changes that promoted viral production, decreasing cycle time. The R values were defined as the scale value between the maximum and minimum viral production level.

$$R = \log(c(v_{p,max})) - \log\left(c\left(v_{p,min}\right)\right) = \log\left(\frac{c(v_{p,max})}{c(v_{p,min})}\right) \quad \text{(Eq. 38)}$$

A minimum of a 10% change in viral production was required to assign a value, such that if $v_{p,max} - v_{p,min} \leq 0.1$, then $R = 1$. We scored pairwise combinations by multiplying the sensitivity value by the reciprocal of the range value.

$$\frac{S}{R} = \frac{d\, v_p}{d \log(c)} \cdot \frac{1}{\log\left(\frac{c(v_{p,max})}{c(v_{p,min})}\right)} \approx \frac{v_{p,max} - v_{p,min}}{\log(c(v_{p,max})) - \log(c(v_{p,min}))} \cdot \frac{1}{\log(c(v_{p,max})) - \log(c(v_{p,min}))} = \frac{d\, v_p}{d^2 \log(c)} \quad \text{(Eq. 39)}$$

Doing so assured higher scores were assigned to those combinations exhibiting high sensitivity over a small range of values, the ideal characteristic for targets of therapeutic intervention. Individually, some parameters are insensitive while others are highly sensitive (Figure 3A-C). Therefore, we sought to assess whether high-scoring pairwise combinations were simply the

combined effects of two sensitive parameters or if there was some amount of cooperativity occurring when both were targeted. We approximated parameter cooperativity as the difference between the pairwise score and the sum of the individual scores according to

$$C = [S/R]_{ij} - \left([S/R]_i + [S/R]_j\right) \quad \text{(Eq. 40)}$$

Thus, cooperative pairs will have a positive value according to Eq. 40, while those that are anti-cooperative will have a negative value.

Modeling small molecule inhibitors

To simulate the effects of a small molecule inhibitor on individual parameters we scaled individual parameters (p) according to a Hill-function

$$p = p_0 - \frac{p_0[D]}{K_D + [D]} \quad \text{(Eq. 41)}$$

where p_0 is the baseline value, $[D]$ is the concentration of the theoretical small molecule inhibitor, and K_D is the dissociation constant, or the affinity, of the inhibitor for its target parameter. Simulated values for K_D are indicated in Figure 4. Inhibitor concentrations were varied between 0.01 and 100 μM on a logarithmic scale. For each case, viral production was estimated according to Eq. 32 above. Levels of inhibition at each concentration come from the Hill-function in Eq. 41 and are equal to p/p_0 for a given inhibitor concentration and K_D value.

Figures:

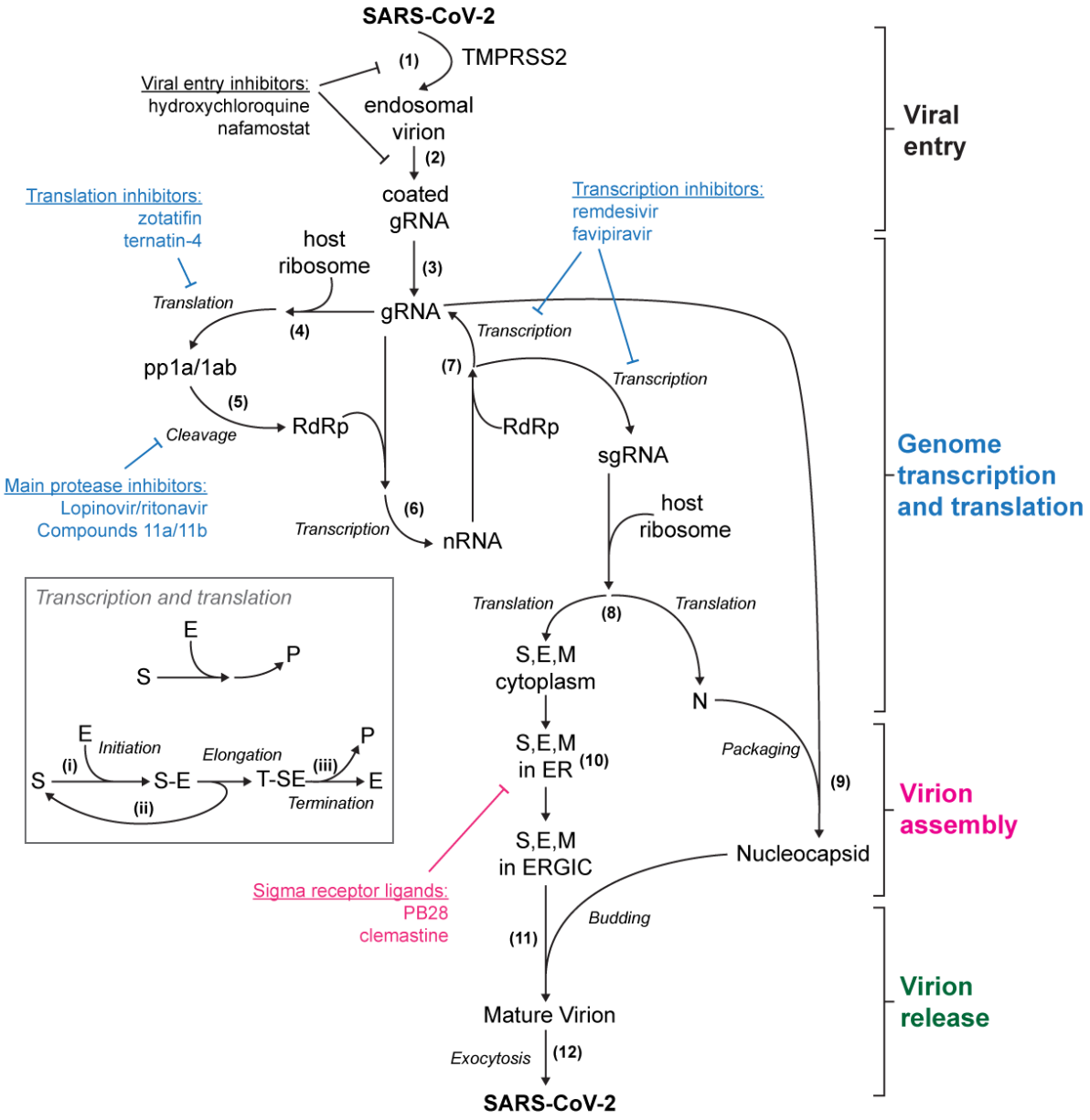
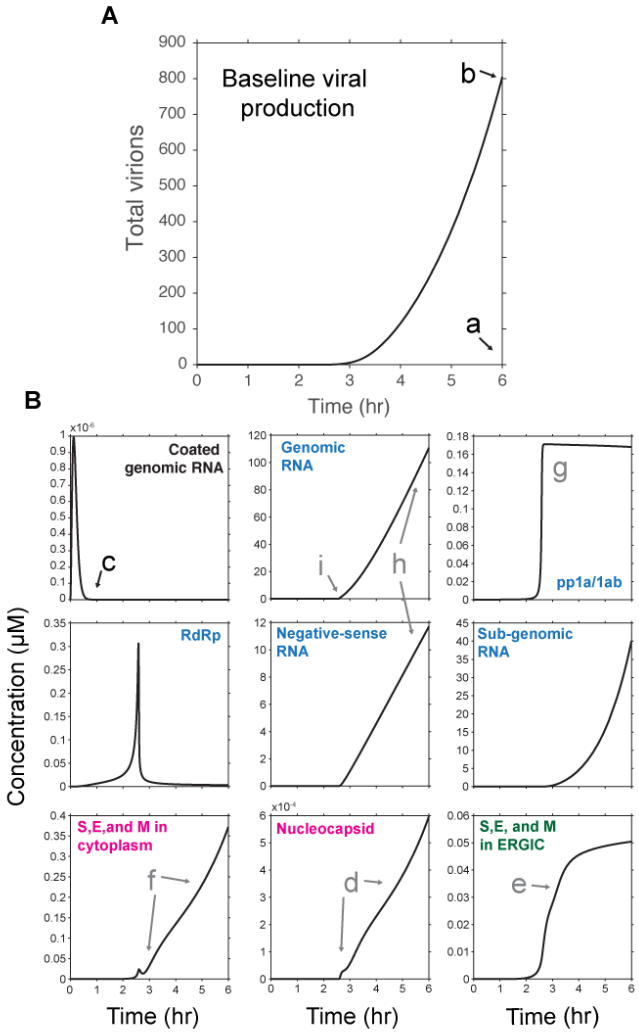

SARS-CoV-2
(1)
TMPRSS2
Viral entry inhibitors:
hydroxychloroquine
nafamostat
endosomal virion
(2)
coated gRNA
Translation inhibitors:
zotatifin
ternatin-4
host ribosome
(3)
Transcription inhibitors:
remdesivir
favipiravir
Translation
gRNA
(4)
Transcription
pp1a/1ab
(7)
Transcription
(5)
RdRp
RdRp
Cleavage
sgRNA
Main protease inhibitors:
Lopinovir/ritonavir
Compounds 11a/11b
(6)
Transcription
nRNA
host ribosome
Transcription and translation
E
S
P
E
Initiation
Elongation
P
S
(i)
S-E
T-SE
(iii)
E
(ii)
Termination
Translation
(8)
Translation
S,E,M cytoplasm
N
S,E,M in ER
(10)
Packaging
(9)
S,E,M in ERGIC
Nucleocapsid
Sigma receptor ligands:
PB28
clemastine
Budding
(11)
Mature Virion
Exocytosis
(12)
SARS-CoV-2
Viral entry
Genome transcription and translation
Virion assembly
Virion release

Figure 1 - Model for the viral cycle of SARS-CoV-2. A single SARS-CoV-2 virion bound to the membrane of a target cell is internalized via TMPRSS2-mediated cleavage of the spike protein **(1)** followed by endosomal release **(2)** and uncoating **(3)** of the viral genome. Once the full-length genome is internalized, host ribosomes bind to open reading frames to translate the polypeptides pp1a and pp1ab **(4)**. Part of pp1a/1ab encodes the main protease (Mpro) involved in cleaving the polypeptides into the non-structural proteins that make up the replicase-transcriptase complex involved in replicating the virus, including the RNA-dependent RNA polymerase (RdRp) **(5)**. **(6)** The RdRp first creates negative-sense RNA (nRNA) templates from the positive-sense genomic RNA (gRNA), which are then used to replicate the full-length gRNA as well as sub-genomic messenger RNAs (sgRNA) **(7)** encoding the structural proteins necessary to build the mature virion. **(8)** sgRNA is translated by the host ribosomes to make the structural spike (S), envelope (E), membrane (M), and nucleocapsid (N) proteins. **(9)** Multiple N proteins bind to the gRNA to package the nucleocapsids. **(10)** While initially translated in the cytoplasm, S, E, and M proteins are translocated through the endoplasmic-reticulum (ER) and ER-Golgi intermediate complex (ERGIC), where buds from the ERGIC eventually encapsulate nucleocapsids to form the mature virion **(11)**. **(12)** Mature virions are then released into the surrounding tissue through exocytosis. Inset: detailed diagram of steps in transcription and translation. **(i)** Initially the RdRp or ribosome (enzyme) binds to the primer region of the respective RNA segment (substrate). **(ii)** Once the primer segment is completed, the substrate is released such that it is available for binding to the primer region by another enzyme. Meanwhile, the initial enzyme continues with elongation steps, polymerizing the reaction product. **(iii)** Upon complete polymerization of the product (RNA segment or protein), transcription/translation is terminated, and the enzyme and resulting product are released. Figure is annotated with example therapeutics to show their approximate point of influence on the viral cycle (*16*, *22–24*). Examples are not exhaustive. Color of the text indicates the part of the viral cycle each example is associated with.



a. Viral cycle is completed on the order of 7-10 hrs (Schneider et al. 2012)
b. Cycle produces a burst of ~1000 virions (Hirano et al. 1976)
c. Viral binding, entry, and uncoating are completed within 30 mins (Ng et al. 2003)
d. Co-localization of N protein and genomic RNA observed between 3 and 5 hrs post-infection (Stertz et al. 2007)
e. Budding from ERGIC observed as early as 3 hrs post-infection (Stertz et al. 2007)
f. Little to no S protein observed at 3 hrs but large increase between 3 and 5 hrs post-infection (Stertz et al. 2007)
g. Nsp3 observable at 3 hrs post-infection (Stertz et al. 2007)
h. Majority of the viral RNA in the cell positive-sense, ~1-10% is negative-sense (Masters et al. 2006)
i. RNA synthesis detectable 1.5-3 hrs post-infection (Sawicki et al. 2007)

Figure 2 - Baseline parameter values reproduce viral replication consistent with previous observations. A) Viral production as a function of time for the baseline parameters shown in Table 1. B) Concentration as a function of simulated time is shown for multiple species within the model. Specific species of interest are indicated within each subpanel. Color corresponds to the point of the viral cycle indicated in Figure 1. Letters indicate examples of model output that is consistent with experimental observations. Black text indicates those observations used to fit model parameters, while those in gray are observations that naturally occur as outputs of the model without parameter adjustment.

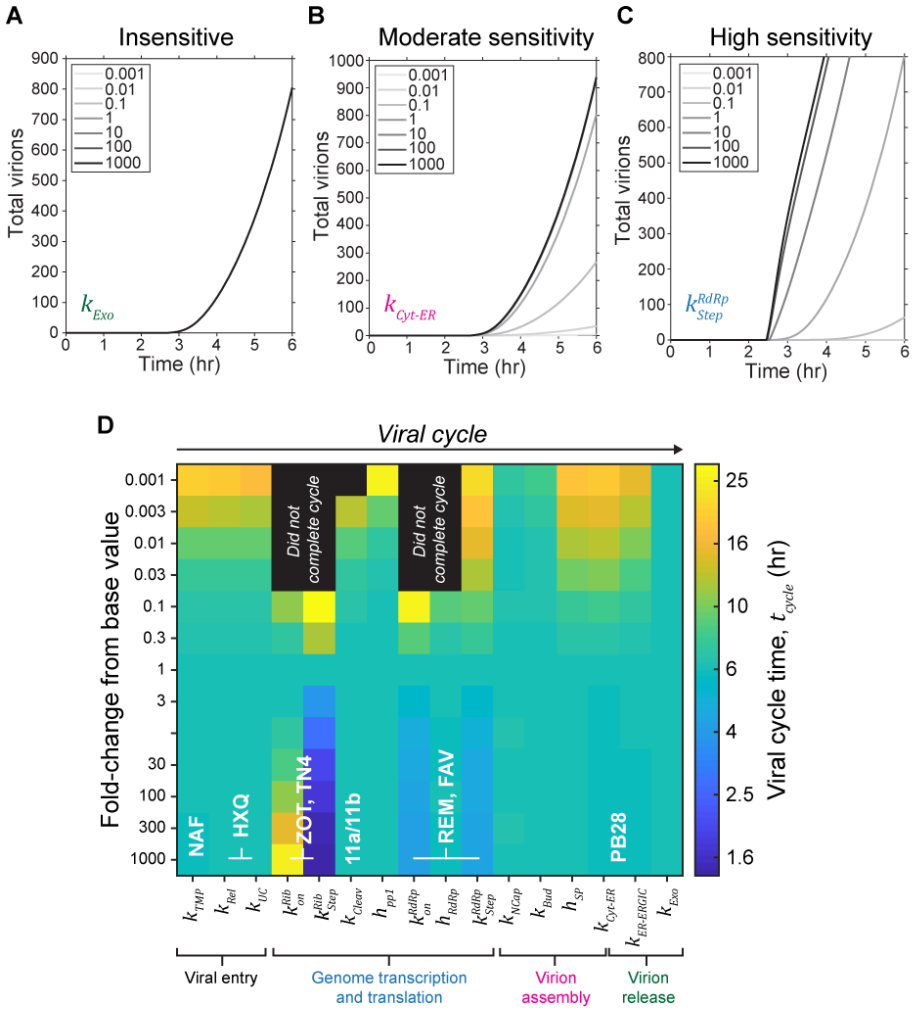


Figure 3 - Sensitivity analysis reveals parameter targets of high interest. (A-C) Examples of insensitive (A), moderate sensitivity (B), and high sensitivity (C) parameters are shown. Inset: fold-change in parameter value relative to the baseline value. Color of the parameter text represents the point of the viral cycle indicated in Figure 1. D) Viral cycle time, estimated as the time to produce 1000 virions is shown as a function of varying parameter values. From left to right, parameters are sorted in the order in which they occur in the viral cycle. Black squares indicate parameter values where the model did not produce 1000 virions within 48 hrs. Figure is annotated with example therapeutics from Figure 1 to show their approximate point of influence on the model parameters. Examples are not exhaustive. NAF - nafamostat; HXQ - hydroxychloroquine; ZOT - zotatifin ; TN4 - ternatin-4; 11a/11b - compounds 11a and 11b; REM - remdesivir; FAV - favipiravir.

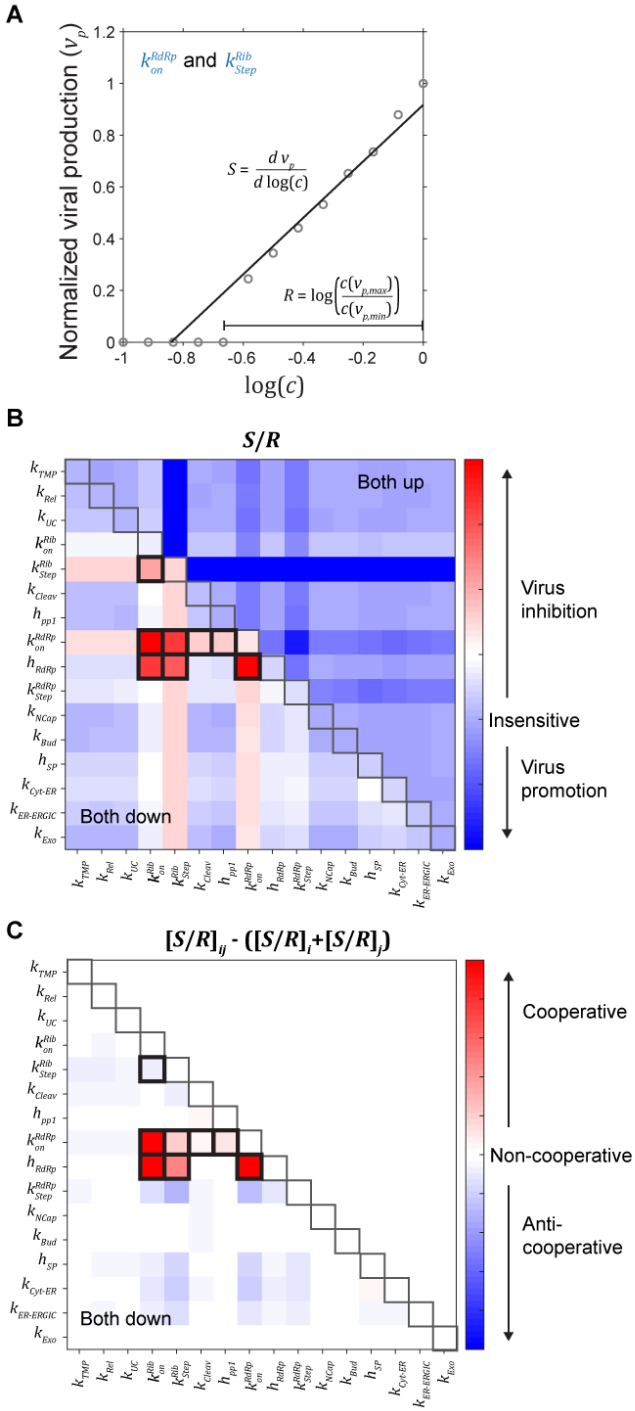

A
Normalized viral production (v_p)
k_{on}^{RdRp} and k_{Step}^{Rib}
$S = \frac{d\,v_p}{d\log(c)}$
$R = \log\left(\frac{c(v_{p,max})}{c(v_{p,min})}\right)$
log(c)
B
S/R
Both up
Both down
Virus inhibition
Insensitive
Virus promotion
C
$[S/R]_{ij}$ - ($[S/R]_i$+$[S/R]_j$)
Both down
Cooperative
Non-cooperative
Anti-cooperative

Figure 4 - Pairwise sensitivity analysis reveals combinations of high interest. A) Example output used to score sensitivity (S) and range (R) is shown. S is estimated as the slope of the normalized viral production (p) versus the logarithm of the scaling parameter (c). As a result of the analysis approach, viral inhibition (as shown here) results in a positive slope, or S value, while promotion of viral production results in a negative S value. R is the difference between the scaling parameter values where maximum and minimum viral production occur. B) Pairwise scoring of sensitivity and range (S/R). Gray boxes indicate the diagonal where only a single parameter is scaled. Relative to the diagonal, the upper-right region is where both parameter values were increased, while the lower-left region is where both parameter values were decreased. Bold boxes indicate parameter combinations that eliminated viral production (p_{min} = 0). C) Cooperative parameter effects are estimated by the difference between the combined effect ($[S/R]_{ij}$) and the sum of the individual effects ($[S/R]_i + [S/R]_j$). Since increasing parameter values (upper-right) did not result in viral inhibition (B), this region was omitted. Positive values (red) resulted when the combined effect was stronger than the sum of the individual effects. Negative values (blue) resulted when the combined effect was weaker than the individual effects. Gray and bold boxes are the same as in B.

Tables:

Table 1- Model parameters

Parameter	Definition	Value	Source
k_{TMP}	TMPRSS2-mediated cleavage of viral spike protein	0.0043 s^{-1}	Estimated using catalytic efficiency of 150,000 $M^{-1}s^{-1}$ (*25*) and assuming 2000 TMPRSS2 enzymes per cell
k_{Rel}	Release of internalized virion from endosome	0.005 s^{-1}	Fit to observations in (*27*)
k_{UC}	Uncoating of genomic RNA	0.005 s^{-1}	Fit to observations in (*27*)
k_{on}^{Rib}	Binding rate of host ribosomes to viral RNA	0.5 $\mu M^{-1}s^{-1}$	Estimated using typical bimolecular association rate of ~10^6 $M^{-1}s^{-1}$ (*28*) and a 10% efficiency of initiation (*26*)
k_{Step}^{Rib}	Stepping rate of host ribosomes	6 aa/s	Fit using total number of virions per cycle (*29*)
k_{Cleav}	Proteolytic cleavage of pp1a and 1ab	0.021 s^{-1}	Estimated using catalytic efficiency of 150,000 $M^{-1}s^{-1}$ (*25*) and steady-state concentration of pp1a/1ab
h_{pp1}	Half-life of pp1a and 1ab polypeptide	1 hr	Approximated from (*30*)
k_{on}^{RdRp}	Binding rate of RNA-dependent RNA polymerase	0.09 $\mu M^{-1}s^{-1}$	Fit to data in (*31*)
h_{RdRp}	Half-life of RNA-dependent RNA polymerase	2 hr	Fit to viral cycle time (*32*)
k_{Step}^{RdRp}	Stepping rate of RNA-dependent RNA polymerase	20 nt/s	Fit to data in (*31*)
k_{NCap}	Nucleocapsid packaging rate	0.00001 $\mu M^{-1}s^{-1}$	Fit to observations in (*32*, *33*)
k_{Bud}	Budding rate from ERGIC	0.01 $\mu M^{-1}s^{-1}$	Fit to observations in (*33*)
h_{SP}	Half-life of structural proteins	30 min	(*34*)
h_{gRNA}	Half-life of positive-sense RNA, genomic and sub-genomic	1 hr	Approximated from (*30*)
h_{nRNA}	Half-life of negative-sense RNA	5 min	Unstable compared to positive-sense RNA (*35*)
k_{Cyt-ER}	Translocation from cytoplasm to ER	0.002 s^{-1}	Fit to observations in (*32*)
$k_{ER-ERGIC}$	Translocation from ER to ERGIC	0.002 s^{-1}	Fit to observations in (*32*)
k_{Exo}	Exocytosis of mature virions	0.0002 s^{-1}	~ 1 hr to release half of the material through exocytosis (*36*)

ρ	Production ratio of genomic RNA:sub-genomic RNA	0.95	Model-fit
λ	Translation ratio of structural proteins (SEM:N)	0.7	(*12*)
a	genomic-RNA per nucleocapsid	1	(*12*)
b	Nucleocapsid protein (N) per nucleocapsid	500	(*12*)
c	Spike (S), envelope (E), and membrane (M) protein per virion	2320	(*12*)
d	Nucleocapsids per virion	2	Approximated based on virion size (*12*)
l_{Primer}^{Rib}	Primer length for protein translation	3 aa	Model-fit based on stepping rate, primer step ≤ 1 s
l_{Primer}^{RdRp}	Primer length for transcription	20 nt	Model-fit based on stepping rate, primer step ≤ 1 s
l_{gRNA}	Length of the full genome	30,000 nt	(*13*, *14*)
l_{sgRNA}	Average length of sub-genomic RNA	1120 nt	(*14*)
l_{pp1}	Length of pp1a and pp1ab segment	6600 aa	(*13*)
l_{SEM}	Average length of the S, E, and M structural proteins	354 aa	(*12*, *14*)
l_N	Length of N structural protein	420 aa	(*12*, *14*)
$[Rib]_0$	Total ribosome concentration in the cell	4 μM	Fit to the total number of virions (*29*), constrained approximating ~10^6 total ribosomes and a host cell volume of ~1000 μm^3 (*30*, *37*)