**SUPPLEMENTARY INFORMATION**

**Integrated Human-Virus Metabolic Modelling Predicts Host-Based Antiviral Targets Against Chikungunya, Dengue and Zika Viruses**

Sean Aller1, Andrew Scott2, Mitali Sarkar-Tyson2,3, and Orkun S. Soyer1,\*

1 School of Life Sciences, Gibbet Hill Campus, University of Warwick, Coventry, CV4 7ES, UK.

2 Defence Science and Technology Laboratory (Dstl), Porton Down, Salisbury, SP4 0JQ, UK. 3Marshall Center for Infectious Disease Research and Training, School of Biomedical Sciences, University of Western Australia, Perth, Australia.

\* Corresponding author

Aller1, Andrew Scott2, Mitali Sarkar-Tyson2,3, and Orkun S. Soyer1,\*

**CONTENTS**

1. ADDITIONAL ANALYSIS AND METHODS
   1. Comparison of flux balance analysis data with gene expression data from Chikungunya, Dengue and Zika infection studies.
   2. The flux enforcement based predictions of antivirals are robust to changes in the key assumptions on the model.

1.2a. Generation of alternative virus VBOFs and statistical analyses.

1.2b. Measuring impact of alternative host model assumptions.

1.3 Literature review of antiviral targets

1. SUPPLEMENTARY TABLES
2. SUPPLEMENTARY FIGURES
3. LIST OF SUPPLEMENTARY FILES
4. REFERENCES

**1. ADDITIONAL ANALYSIS AND METHODS**

**1.1. Comparison of flux balance analysis data with gene expression data from Chikungunya, Dengue and Zika infection studies.** As discussed in the main text, we used gene expression data from controlled virus infection experiments using Chikungunya, Dengue and Zika viruses. These experiments were conducted using human cell lines as host organisms for the virus infection: Chikungunya virus, HEK293T; Dengue virus, HuH-7; Zika virus, hNPC and SVG. The gene expression data used in Figure S8 is collated from the NCBI’s Gene Expression Omnibus [1,2] and is accessible through the GEO Series accession numbers: Chikungunya virus, GSE49985; Dengue virus, GSE69602; Zika virus, GSE78711 and GSE89915. The macrophage metabolic model utilised in this study [3] contains mappings between reaction identities and enzyme commission (EC) numbers, although this is only for a subset of reactions. We mapped these EC numbers onto human metabolic genes via the HumanCyc metabolic database [4,5]. Due to the nature of gene-reaction mapping, a single reaction can be represented by many genes and vice versa. The gene expression datasets contain unique identifiers which relate to a particular human gene; thus some of the genes in these datasets can be linked back to the model. To do so, the two gene lists (model-derived using EC numbers and GEO-derived) are intersected, leaving a set of genes that are present in the gene expression datasets and also in the model. For these genes, a comparison is made between the change in gene expression upon infection (using the listed GEO databases), and the change in host vs virus optimised reaction fluxes. Both of these changes are presented as log2 of the ratio of the respective values, and the log 2fold-changes are then compared in the Supplementary Figure S8.

**1.2. The flux enforcement based predictions of antivirals are robust to changes on the model.** The results presented in the main text provide support for using integrated host-virus metabolic models for identifying host-based antiviral targets. However, any analysis based on stoichiometric metabolic flux optimisation, as done here, can be dependent on details of model implementation and assumptions [6]. For example, while the host metabolic model and the biomass function that it incorporates are verified against experimental data [7], the model still assumes a specific media composition and uptake fluxes. To test if our predictions are robust against key assumptions in the model, we analysed the effects of variations in the media composition for the host model and the genomic sequence of the virus on the antiviral target prediction. In particular, we repeated the above analysis for 1000 alternative media uptake fluxes, and 1000 point mutations for each of the three virus genomes, affecting the final virus biomass (see *Methods*). We found that the list of reactions with the highest impact on virus biomass, while maintaining the host biomass, are qualitatively not altered with these changes in the model structure, and we still recover known broad antiviral targets (*Supplementary Files S5-S7* and Figures S2-4).

To further probe the significance of these reactions as effectors of virus production, we generated 6000 randomised virus biomass compositions using the three original virus biomass functions as a starting point (see section below). Repeating the enforcement analysis for this randomised set thus allowed us to generate a distribution of effects of reactions on virus-like particle production in the host system, thus acting as a null hypothesis. Results from the enforcement of the original unaltered virus sequences, as well as the point mutated virus sequences (mentioned above), were compared against the population of ‘randomised’ viruses to assess the significance of the antiviral effect (see *Supplementary Figures S5-S7*). This allows derivation of a significance value for the effects resulting from each reaction when the flux enforcement analysis is applied to it. When we rank reactions according to the significance of their effects, we find that the list of reactions shown in *Supplementary File S4* are ranked among the top (*Supplementary Files S5-S7*). These reactions cause biomass reduction for each virus that is statistically significant when compared to their effects on randomised virus-like biomass functions. There are a couple of exceptions only in the case of ZIKV, where the reactions mediated by cystathionine g-lyase (*CYSTGL*) and cystathionine beta-synthase (*CYSTS*) did not show any significance in their effect under flux enforcement. Additional statistical analysis showed that most of the reactions listed in *Supplementary File S4* (27 out of 29) also showed significant differences in the magnitude of their effects among the three different viruses. In other words, while the reactions we highlight are not necessarily unique when comparing amongst CHIKV, DENV and ZIKV, their quantitative effects on virus production is significantly different for each species. Combined with the fact that our randomisation process maintained the key features of stoichiometric differences among the host and virus-like biomass functions, we highlight that the flux-perturbing effects of the identified reactions emerge from the core metabolic stoichiometric differences between host and the viruses. Mainly, the fact that viruses use much higher levels of nucleic acids per biomass unit (see Figure 2 in the main text).

**1.2a. Generation of alternative virus VBOFs and statistical analyses.** The presented approach evaluates the optimisation of host metabolic fluxes for virus production and how specific constraints in the host model can alter such an optimal state. To evaluate the impact of having different VBOFs on the outcome of such an analysis, we have generated a range of VBOFs that were either derived from the original VBOFs or were completely randomly generated. To evaluate impact of small deviations from the original VBOFs, we generated variants of the original virus genomes through nucleotide substitution; for each virus, we generated 1000 genome variants, where each variant contained 1, 2, 3, 4, 5 or 10 nucleotide substitutions. For the subsequent VBOF generation from these variant genomes, the genome and protein copy numbers were kept as in the original. To evaluate more variant VBOFs, we generated another 1000 genomes for each virus that were created from the original genome with a random number (between 0 and the total length of the genome) of nucleotide substitutions and using randomly drawn structural and non-structural polyprotein copy numbers per virus particle. Finally, and in an attempt to generate a set of VBOFs that are far removed from the original ones, both in terms of genome sequence and the structural and non-structural protein numbers, we directly generated random VBOFs. We have implemented this by drawing 1000 sets of individual stoichiometries of biomass components from a uniform distribution on [*a*, *b*], where *a* and *b* are (1) ±99% of the original stoichiometric coefficients of a given virus, or (2) are ±99% of the average of all original stoichiometric coefficients of a given virus. These approaches to generating variant virus genomes give us set of sequences (and associated VBOFs) that are increasingly removed from the original virus VBOFs. For each randomised VBOF created, the host-derived flux enforcement analysis was repeated (with a recalculation of the bounds used for the individual enforcement), and the reactions that perturb virus optima the most when constrained were identified. This whole analysis resulted in 8000 randomised VBOFs and FBA simulations, the results of which are summarised as percentage impact of individual host reactions on virus optima for different sets of VBOFs (see *Supplementary Files S5* and *S7*). To compare results of flux enforcement analysis to that obtained from using randomised biomass functions, we used a one-way ANOVA and Tukey’s honest significance tests for each individual virus against the randomised virus group, and between each individual virus species (*Supplementary File S8*).

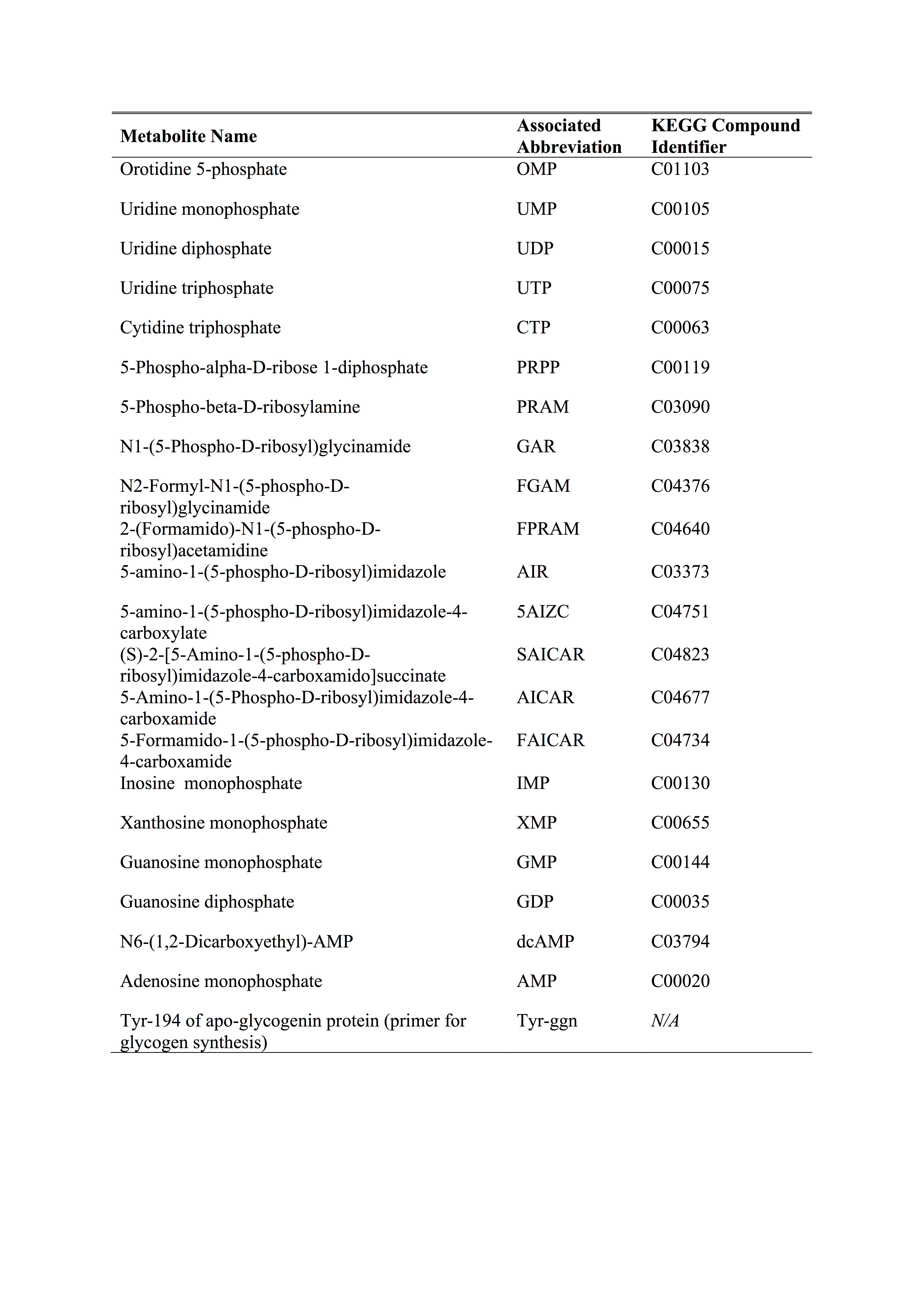
**1.2b. Measuring impact of alternative host model assumptions.** As explained above, the presented analysis involves integrating a VBOF into a host model. While we used here a validated, published human macrophage metabolic model as the host model, it is desirable to evaluate the impact of specific assumptions of such a model on the results of this analysis. In particular, the macrophage model uses specific metabolite uptake fluxes, which are mostly based on experimental observations [7-9], but which can directly influence FBA-based results. To evaluate the potential impact of model uptake bounds on our *in silico* predictions, we systematcially re-analysed virus optimization and its constraining by the host model, using alternative metabolite uptake fluxes (i.e. different media compositions) in the host model. We first identified metabolites that are supplied (via exchange reactions) to the metabolic model with non-arbitrary lower bounds (*lb*), where *0* > *lb > -∞*. Each of the identified reactions are then systematically constrained, such that the *lb* is reduced from the original model values [6] (in steps of 10%) until the *lb* = 0, effectively knocking-out the respective exchange reaction. For each altered (additionally constrained) model, the host-derived enforcement analysis is repeated, with re-calculation of the viable host reaction bounds and optimisation of VBOF. This is done for each of the three viruses (CHIKV, DENV, ZIKV) and the results are summarised in *Supplementary File S6*.

**1.3.      Literature review of antiviral targets.**In order to assess the pre-existing collection of antivirals that targeted reactions, and specifically in the case of RNA viruses, a literature search was conducted. NCBI and Google Scholar were searched for combinations of keywords: “antiviral”, “reaction target”, “molecular”, “metabolism”, “RNA viruses”, “RNA”. Subsequently, papers that described the use of antiviral drug compounds to specifically target metabolic components (reactions), and presented affects against RNA viruses, were selected [1,5,13- 18]. For antivirals with metabolic reaction targets identified, the presented experimental evidence was checked to see if the three viruses in this study (*Chikungunya*, *Dengue* and *Zika*) had been experimentally tested against. A summary of the literature review is included in Table S1, while a full breakdown (including pathway effects) is provided in *Supplementary File* S10.

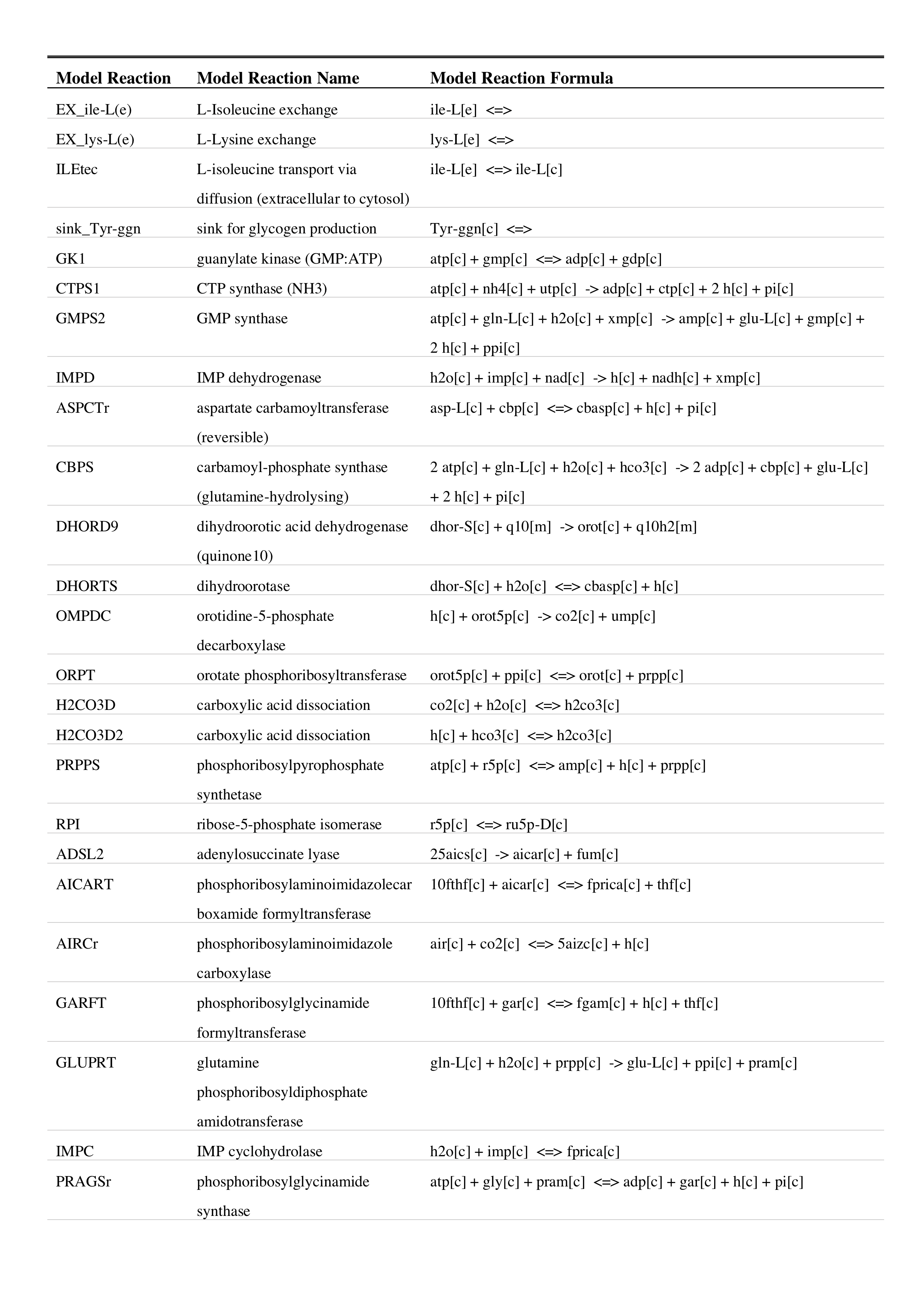
**2. SUPPLEMENTARY TABLES**

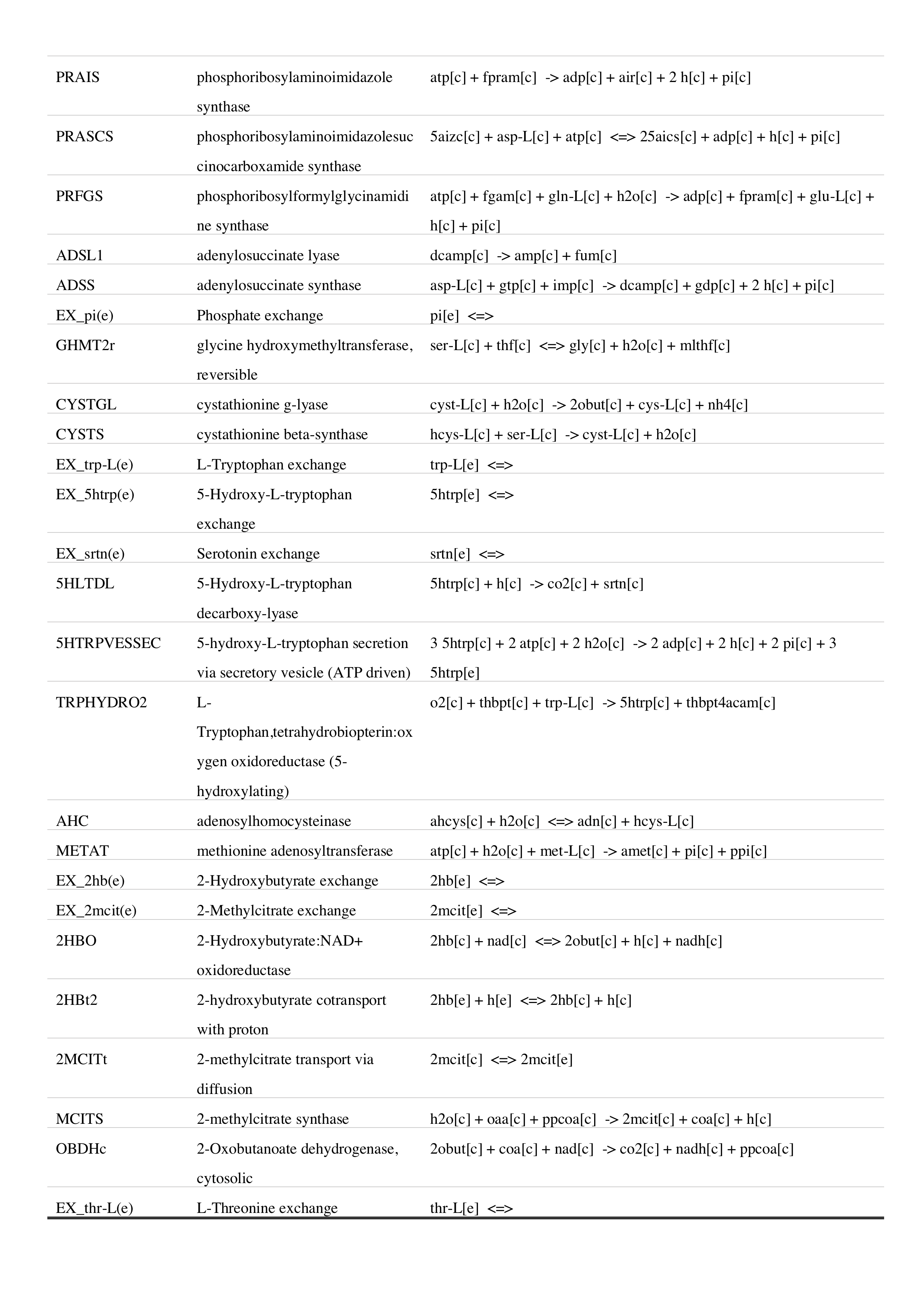
****

**Supplementary Table S1 |** List of antiviral compounds identified, which interact with host metabolic reactions that affect RNA virus production [10,11]. \*CTP Synthase is encoded for by two genes, CTPS1 and CTPS2, and both are affected by the associated antiviral compounds. Only one gene is shown. Viruses from this study are identified if they were experimentally validated against (Chikungunya virus, CHIKV; Dengue virus, DENV; Zika virus, ZIKV). All antiviral compounds, and their associated target reactions, were identified from literature search (see Citations).

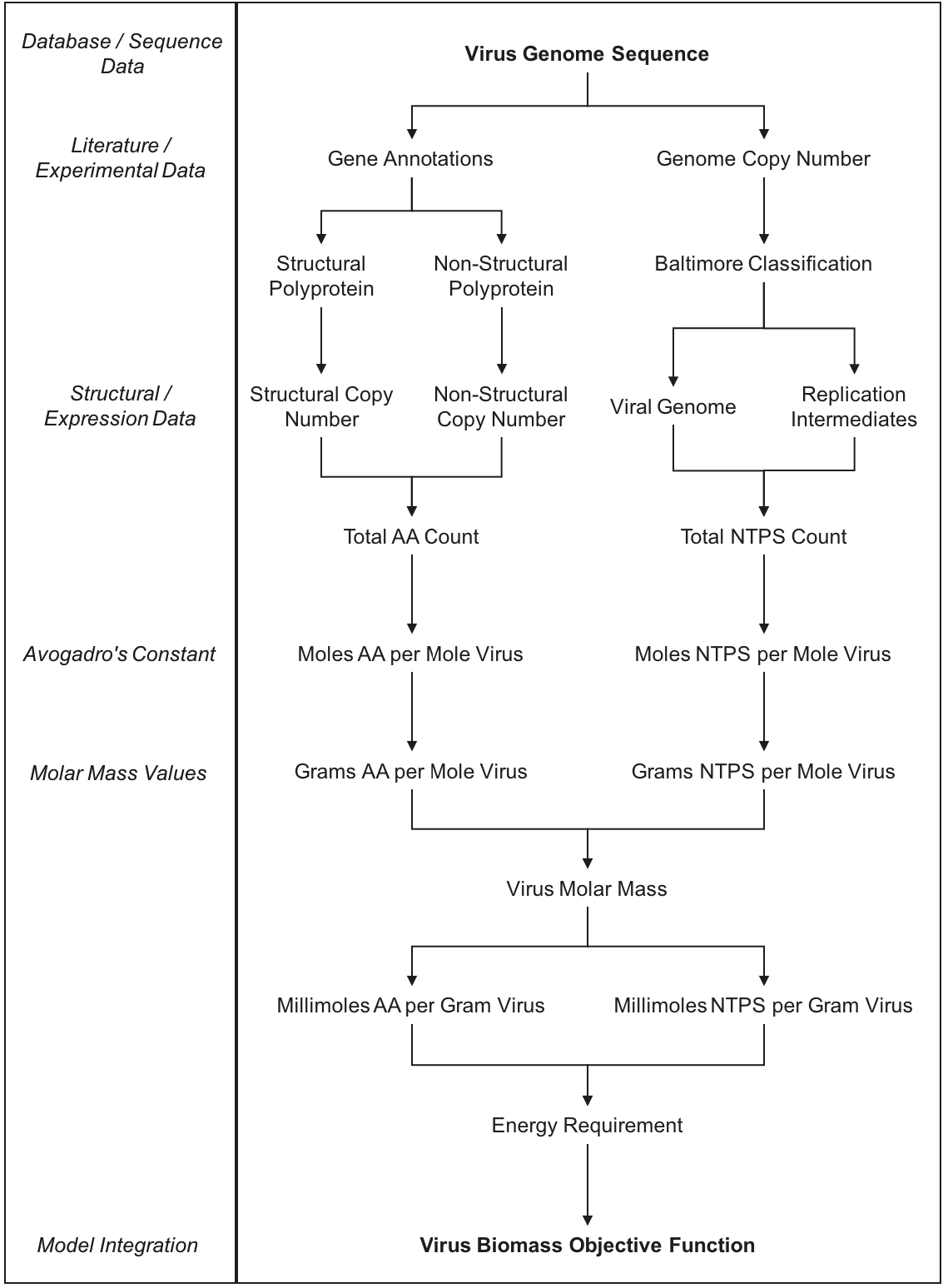
****

**Supplementary Table S2 |** Metabolite abbreviations and full names used in Figure 4, as derived from the iAB-AMØ-1410 Human Alveolar Macrophage metabolic reconstruction [12].

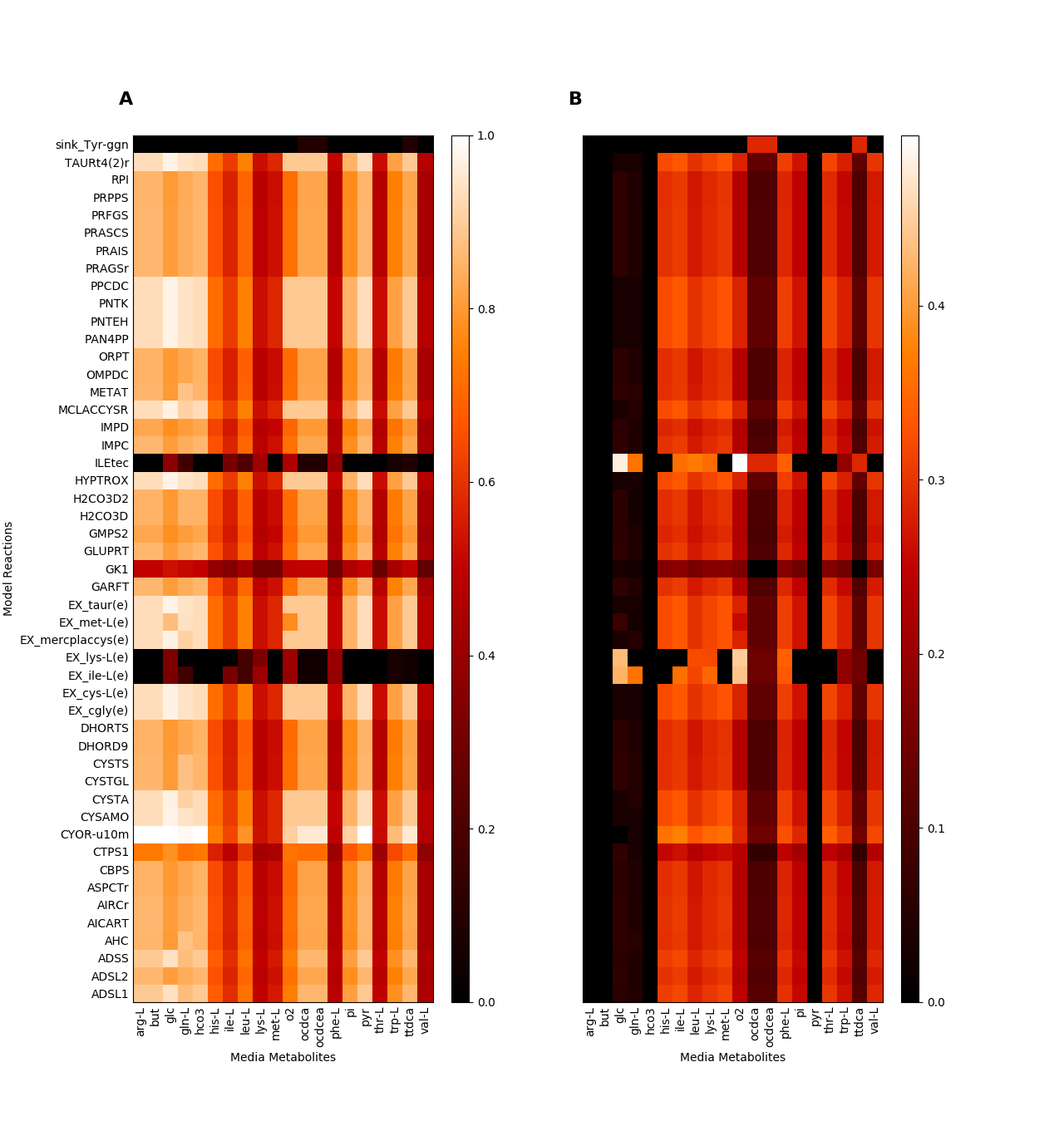
****

**Supplementary Table S3 |** Reaction abbreviations and full names used in Figure 4, as derived from the iAB-AMØ-1410 Human Alveolar Macrophage metabolic reconstruction [12].

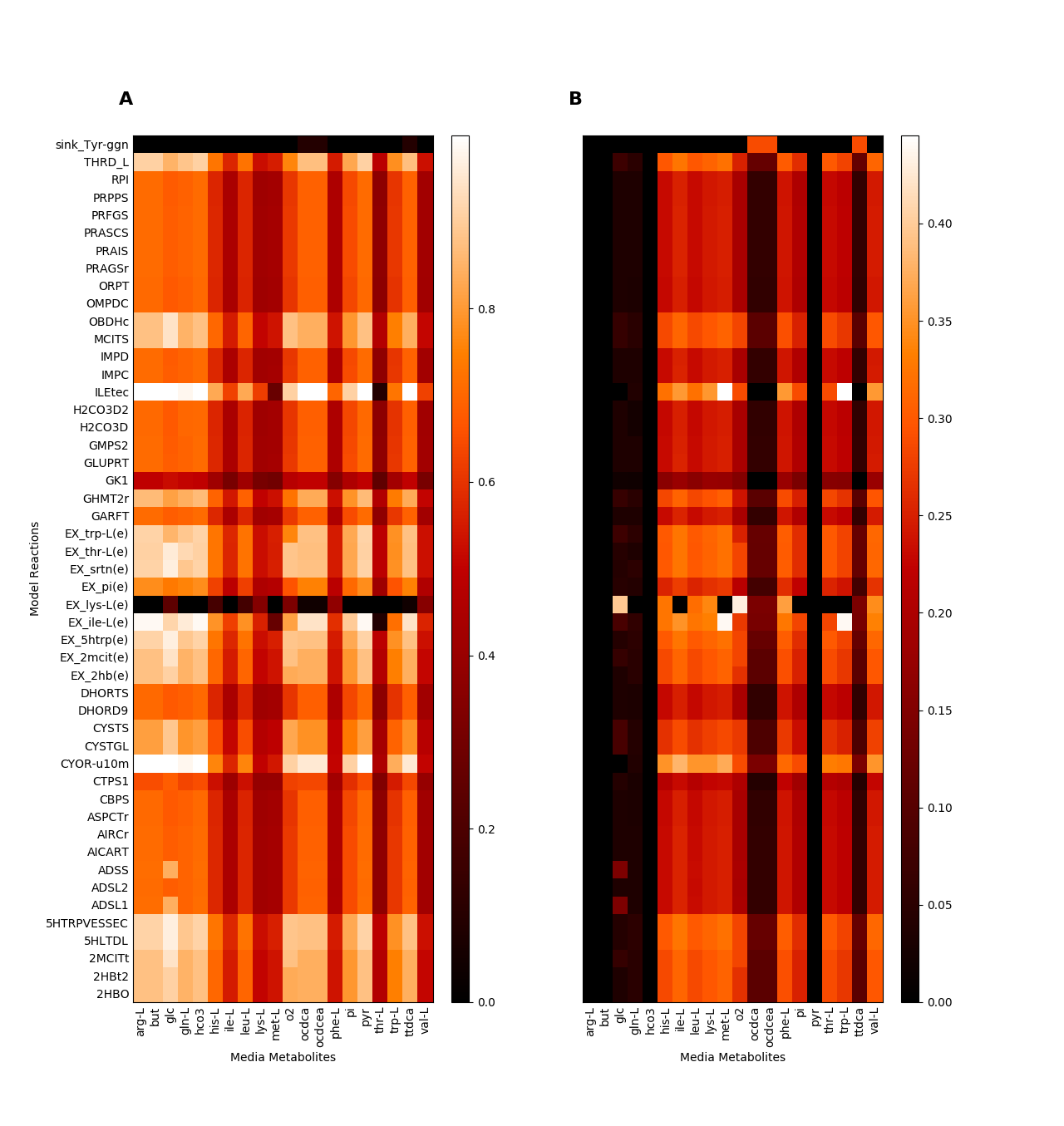
**3. SUPPLEMENTARY FIGURES AND LEGENDS**

****

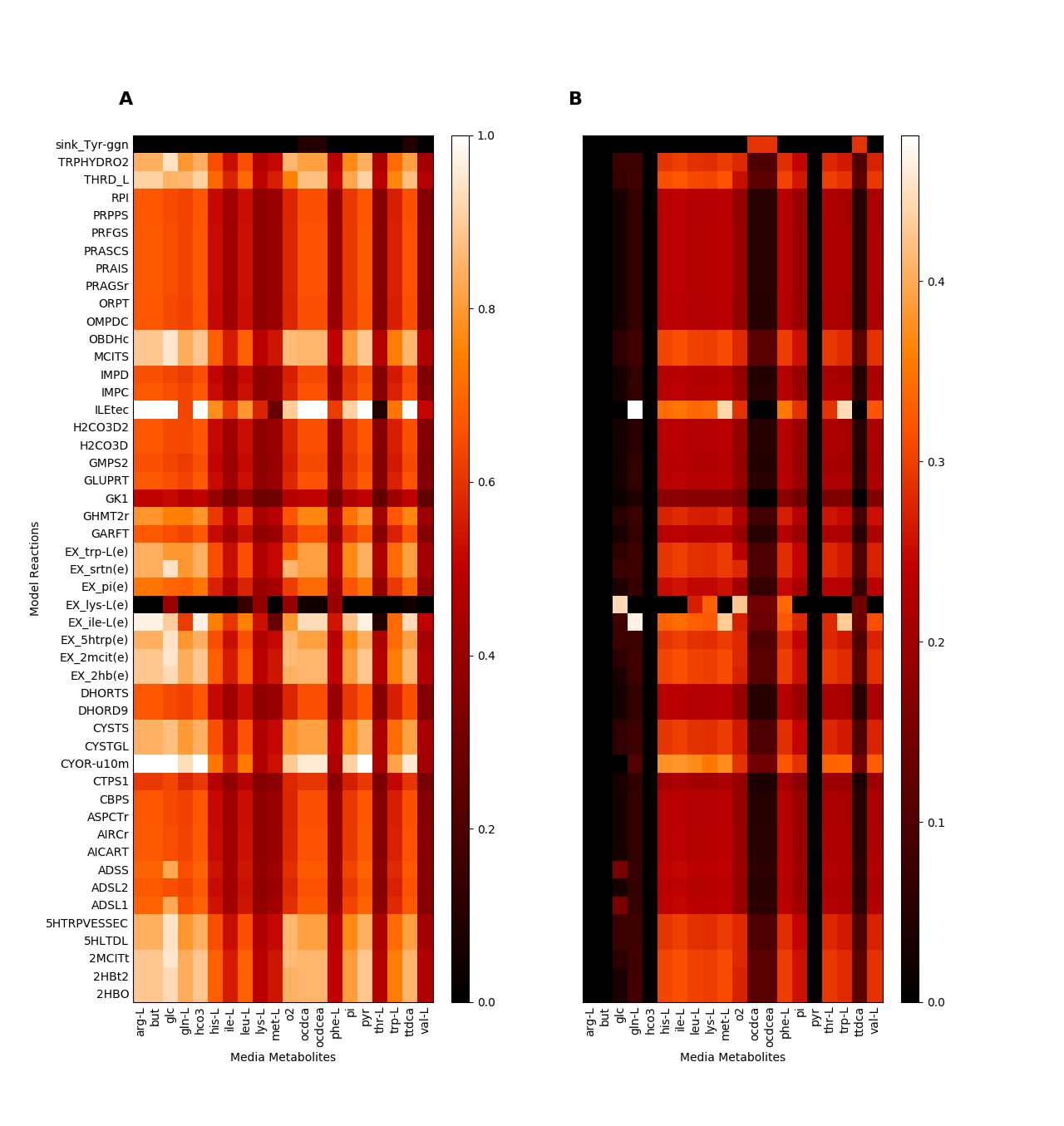
**Figure S1 | Schematic of virus biomass objective function (VBOF) generation.** Diagram outlines the process of forming the necessary components for the pseudo-reaction that represents the production of virus particles (biomass).

****

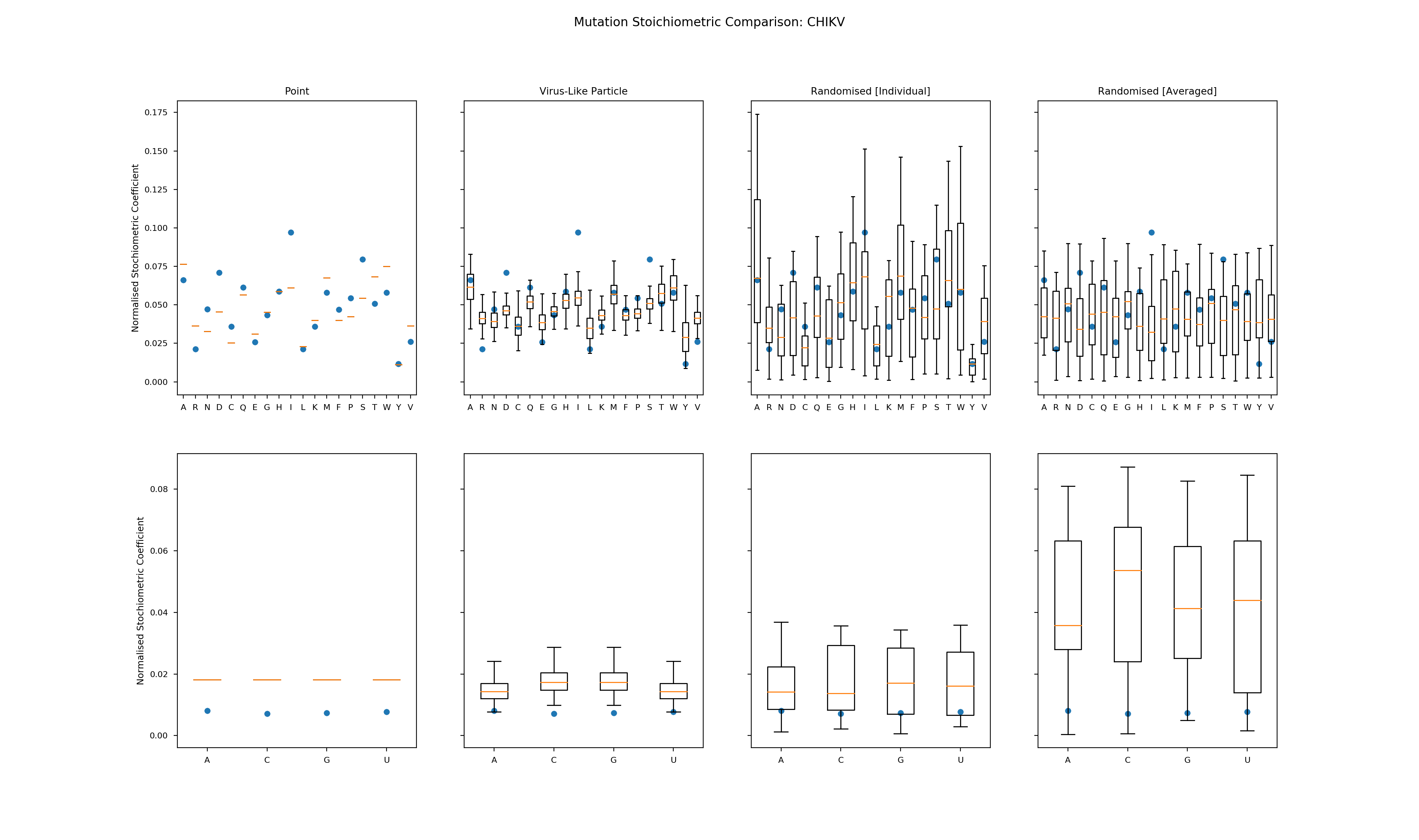
**Figure S2 | Host-derived enforcement reactions analysed over varying model media compositions, for given in silico media metabolites, for CHIKV.** A) average normalised virus optima for each media alteration condition. B) standard deviation of virus optima for each media alteration condition.



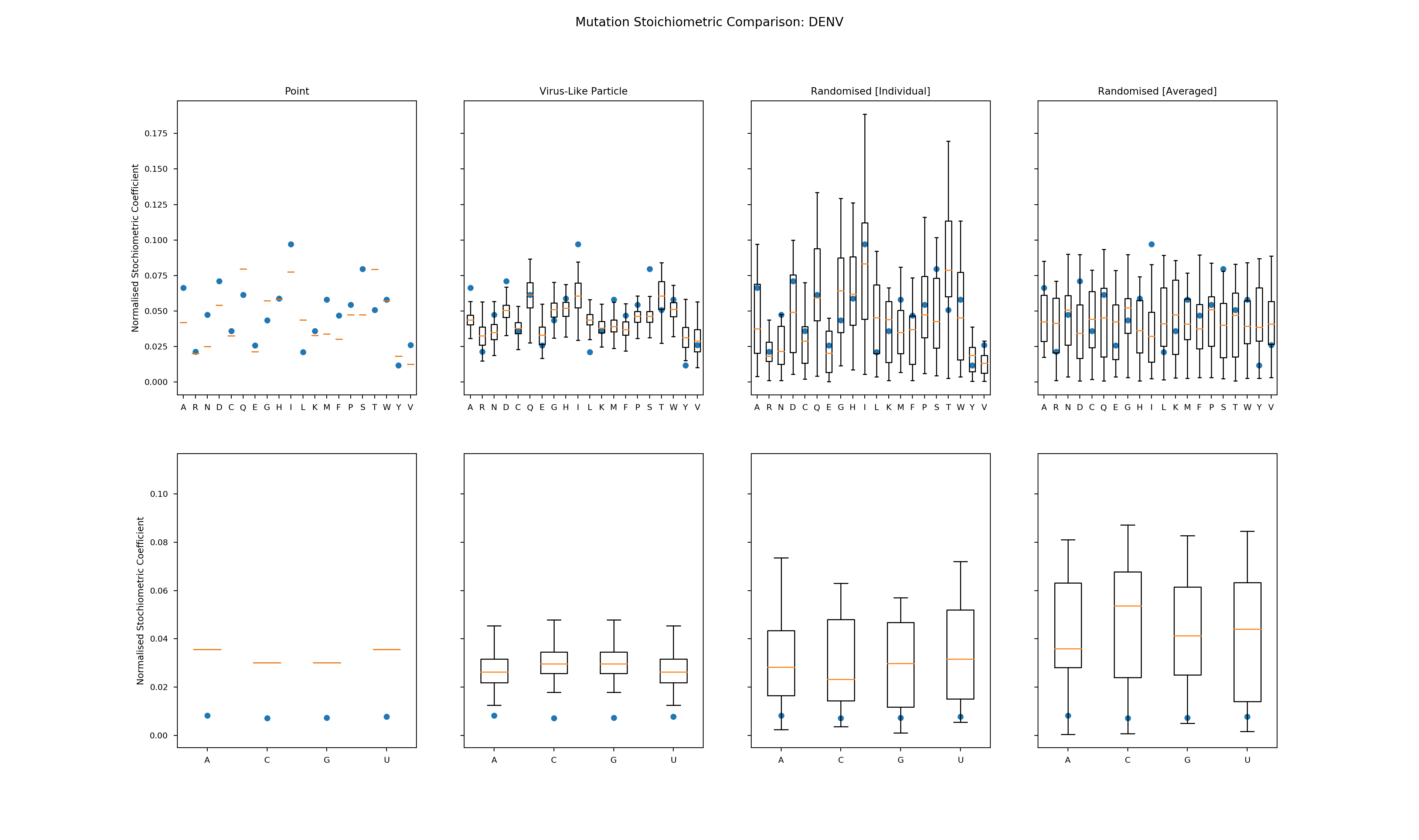
**Figure S3 | Host-derived enforcement reactions analysed over varying model media compositions, for given in silico media metabolites, for DENV.** A) average normalised virus optima for each media alteration condition. B) standard deviation of virus optima for each media alteration condition.



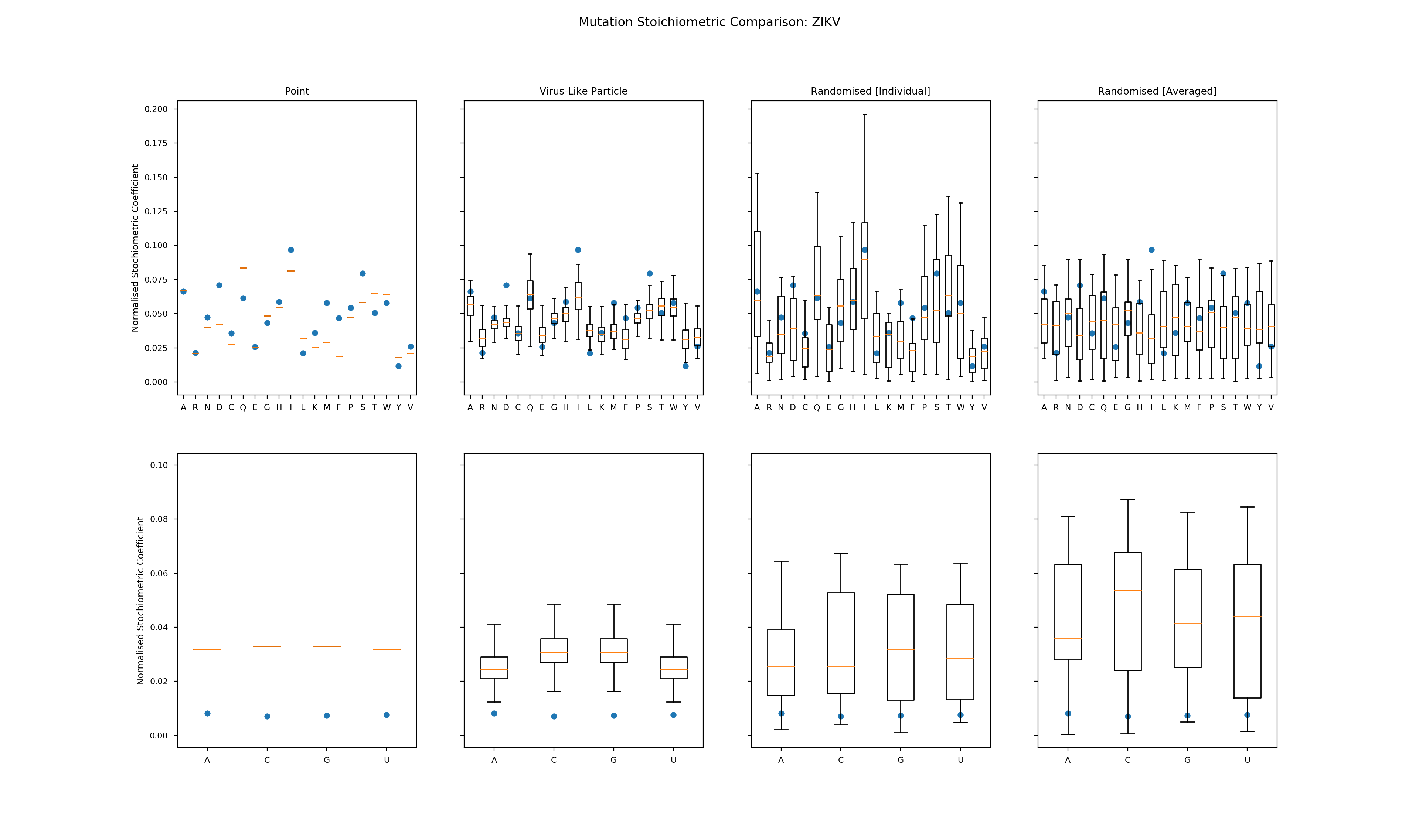
**Figure S4 | Host-derived enforcement reactions analysed over varying model media compositions, for given in silico media metabolites, for ZIKV.** A) average normalised virus optima for each media alteration condition. B) standard deviation of virus optima for each media alteration condition.



**Figure S5 | Comparison of stoichiometric coefficients for multiple mutation conditions for CHIKV.** Normalised stoichiometric coefficients for amino acids and nucleotides. Original stoichiometric coefficients (blue circle) are compared to the mutant stoichiometric coefficients (box and whiskers). Chikungunya virus, CHIKV.

****

**Figure S6 | Comparison of stoichiometric coefficients for multiple mutation conditions for DENV.** Normalised stoichiometric coefficients for amino acids and nucleotides. Original stoichiometric coefficients (blue circle) are compared to the mutant stoichiometric coefficients (box and whiskers). Dengue virus, DENV.

****

**Figure S7 | Comparison of stoichiometric coefficients for multiple mutation conditions for ZIKV.** Normalised stoichiometric coefficients for amino acids and nucleotides. Original stoichiometric coefficients (blue circle) are compared to the mutant stoichiometric coefficients (box and whiskers). Zika virus, ZIKV.

M:\RS\Figshare\Approved\rsif20180125\Supplementary Figure S8.tif

**Figure S8 | Comparison of gene expression data and metabolic model flux predictions.** Comparisons were made between gene expression data sets (denoted in legends) and the fluxes predicted for the respective virus-optimised metabolic network, for each virus: **a,** CHIKV; **b,** DENV; **c,** ZIKV. Comparisons shown are the log-2 fold-change of virus vs host fluxes and infected vs uninfected [virus vs host] gene expression data at different time points in hours (h). CHIKV, chikungunya virus; DENV, dengue virus; ZIKV, zika virus.

**4. LIST OF SUPPLEMENTARY FILES**

**Supplementary File S1.** Biomass objective function stoichiometric coefficients and associated metabolites for the [host] macrophage (iAB-AMØ-1410) biomass maintenance objective function and the Chikungunya (CHIKV), Dengue (DENV) and Zika (ZIKV) virus biomass objective functions.

**Supplementary File S2.** Flux variability analysis results for the host-virus integrated models for CHIKV, DENV and ZIKV. Both the host- and virus-optimal state results are shown. Model reactions, model subsystems and the associated aggregated subsystems are also detailed.

**Supplementary File S3.** Results for the reaction knockout and host-derived enforcement analyses. Model reactions and model subsystems are listed. The resulting virus optima, for CHIKV, DENV and ZIKV, are shown as a percentage of the original model (without the additional constraints imposed by the analyses). The lower and upper flux bounds used for the host-derived enforcement analysis are listed for the relevant reaction and associated virus optimisation result.

**Supplementary File S4.** A list of model reactions identified from the host-enforcement analysis that are able to reduce one of the three viruses, CHIKV, DENV and ZIKV, to below 80% of the original virus optima in the original model (without additional constraints). Inhibitors are listed for the respective reactions and were identified through the Brenda Enzymatic Database via the Enzyme Commission Number associated with the reaction.

**Supplementary File S5.** Summary of results for the host-derived enforcement analysis for different conditions: altered media conditions; virus genome point mutations; virus biomass objective function randomised stoichiometric coefficients. Reactions are selected when they are able to reduce virus optima to below 95% of the original virus optima in the original model (without additional constraints).

**Supplementary File S6.** Results from the *in-silico* media variation sensitivity analyses to evaluate the effectiveness of the host-derived enforcement antiviral targets over a range of varied media compositions for CHIKV, DENV and ZIKV viruses.

**Supplementary File S7.** Results from the mutation-based sensitivity analyses to evaluate the effectiveness of the host-derived enforcement antiviral targets for mutated CHIKV, DENV and ZIKV viruses.

**Supplementary File S8.** Results from the one-way ANOVA and tukey HSD tests on the effect of host-derived enforcements across point and randomised mutation conditions for CHIKV, DENV and ZIKV. Chikungunya virus, CHIKV; Dengue virus, DENV; Zika virus, ZIKV.

**Supplementary File S9.** Results from the comparison of gene expression data and the *in silico* model predictions, for CHIKV, DENV and ZIKV viral infections. Chikungunya virus, CHIKV; Dengue virus, DENV; Zika virus, ZIKV.

**Supplementary File S10.** Results from the literature analysis of antiviral compounds with metabolic targets (reactions). Compound names and descriptions are provided, as identified from the literature (references column). Targets are described in terms of the metabolic pathway and reaction they target (along with the reaction abbreviation from the metabolic model for the aforementioned reaction). Where identifiable, virus taxonomy is defined (species, genera and family), genomic material type (RNA) and genome sense (positive [+] or negative [-]) are shown.

**5. REFERENCES**

1. Littler, E. & Oberg, B. 2005 Achievements and challenges in antiviral drug discovery. *Antivir. Chem. Chemother.* 16, 155–168. (doi:10.1177/095632020501600302)

2. Mlakar, J. et al. 2016 Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* 374, 951–958. (doi:10.1056/NEJMoa1600651)

3. Kotzamanis, K., Angulo, A. & Ghazal, P. 2015 Infection homeostasis: implications for therapeutic and immune programming of metabolism in controlling infection. *Med. Microbiol. Immunol.* 204, 395–407. (doi:10.1007/s00430-015-0402-5)

4. Maynard, N. D., Birch, E. W., Sanghvi, J. C., Chen, L., Gutschow, M. V. & Covert, M. W. 2010 A forward-genetic screen and dynamic analysis of lambda phage host-dependencies reveals an extensive interaction network and a new anti-viral strategy. *PLoS Genet.* 6, e1001017. (doi:10.1371/journal.pgen.1001017)

5. Merino-Ramos, T., Vázquez-Calvo, Á., Casas, J., Sobrino, F., Saiz, J.-C. & Martín-Acebes, M. A. 2015 Modification of the Host Cell Lipid Metabolism Induced by Hypolipidemic Drugs Targeting the Acetyl Coenzyme A Carboxylase Impairs West Nile Virus Replication. *Antimicrob. Agents Chemother.* 60, 307–315. (doi:10.1128/AAC.01578-15)

6. Zhu, Y., Yongky, A. & Yin, J. 2009 Growth of an RNA virus in single cells reveals a broad fitness distribution. *Virology* 385, 39–46. (doi:10.1016/j.virol.2008.10.031)

7. Yu, Y., Clippinger, A. J. & Alwine, J. C. 2011 Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection. *Trends in Microbiology* 19, 360–367. (doi:10.1016/j.tim.2011.04.002)

8. El-Bacha, T., Menezes, M. M. T., Azevedo e Silva, M. C., Sola-Penna, M. & Da Poian, A. T. 2004 Mayaro virus infection alters glucose metabolism in cultured cells through activation of the enzyme 6-phosphofructo 1-kinase. *Mol. Cell. Biochem.* 266, 191–198.

9. Jain, R. & Srivastava, R. 2009 Metabolic investigation of host/pathogen interaction using MS2-infected Escherichia coli. *BMC Syst Biol* 3, 121. (doi:10.1186/1752-0509-3-121)

10. Molenaar, D., van Berlo, R., de Ridder, D. & Teusink, B. 2009 Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology* 5, 323. (doi:10.1038/msb.2009.82)

11. Weiße, A. Y., Oyarzún, D. A., Danos, V. & Swain, P. S. 2015 Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl. Acad. Sci. U.S.A.* 112, E1038–47. (doi:10.1073/pnas.1416533112)

12. Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø. & Jamshidi, N. 2010 Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. *Molecular Systems Biology* 6, 422. (doi:10.1038/msb.2010.68)

13. Leyssen, P., De Clercq, E. & Neyts, J. Molecular strategies to inhibit the replication of RNA viruses. *Antiviral Research* 78, 9–25 (2008).

14. Cyclopentenylcytosine. A carbocyclic nucleoside with antitumor and antiviral properties. 31, 1687–1694 (1988).

15. Lucas-Hourani, M. *et al.* Inhibition of Pyrimidine Biosynthesis Pathway Suppresses Viral Growth through Innate Immunity. *PLoS Pathog.* 9, e1003678 (2013).

16. Wang, Q.-Y. *et al.* Inhibition of dengue virus through suppression of host pyrimidine biosynthesis. *Journal of Virology* 85, 6548–6556 (2011).

17. Hoffmann, H.-H., Kunz, A., Simon, V. A., Palese, P., & Shaw, M. L. (2011). Broad-spectrum antiviral that interferes with de novo pyrimidine biosynthesis. Proceedings of the National Academy of Sciences of the United States of America, 108(14), 5777–5782. http://doi.org/10.1073/pnas.1101143108

18. Neyts, J., Meerbach, A., McKenna, P., De Clercq, E., 1996. Use of the yellow fever virus vaccine strain 17D for the study of strategies for the treatment of yellow fever virus infections. Antivir. Res. 30 (2–3), 125–132.