Interdisciplinary Description of Complex Systems 21(6), 623-630, 2023

OPTIMAL STRATEGIES FOR VIRUS PROPAGATION

Soumya Banerjee*

University of New Mexico, Department of Computer Science
Albuquerque, The United States of America

Ronin Institute
Montclair, The United States of America

Complex Biological Systems Alliance
Medford, The United States of America

Broad Institute of MIT and Harvard
Cambridge, The United States of America

DOI: 10.7906/indecs.21.6.7
Regular article

Received: 25 April 2023.
Accepted: 25 June 2023.

ABSTRACT

This article explores a number of questions regarding optimal strategies evolved by viruses upon entry into a vertebrate host. The infected cell life cycle consists of a non-productively infected stage in which it is producing virions but not releasing them and of a productively infected stage in which it is just releasing virions. The study explores why the infected cell cycle should be so delineated, something which is akin to a classic "bang-bang control" or all-or-none principle. The times spent in each of these stages represent a viral strategy to optimize peak viral load. Increasing the time spent in the non-productively infected phase (τ_1) would lead to a concomitant increase in peak viremia. However increasing this time would also invite a more vigorous response from Cytotoxic T-Lymphocytes. Simultaneously, if there is a vigorous antibody response, then we might expect τ_1 to be high, in order that the virus builds up its population and conversely if there is a weak antibody response, τ_1 might be small. These trade-offs are explored using a mathematical model of virus propagation using Ordinary Differential Equations. The study raises questions about whether common viruses have actually settled into an optimum, the role for reliability and whether experimental infections of hosts with non-endemic strains could help elicit answers about viral progression.

KEY WORDS

viral dynamics, optimization, bang-bang control, viral strategies, optimal control theory

CLASSIFICATION

JEL: I19

*Corresponding author, η: soumya.banerjee@ronininstitute.org; -; -

INTRODUCTION

A normal cell upon infection goes through a life cycle characterized by 2 phases: a stage in which it is producing virions but not releasing it (non-productively infected stage) and a stage in which it is releasing virions into the outside environment (productively infected stage). Hence there is a delay between infection and release of virions. In whatever follows, we denote the time spent in the non-productively infected stage as τ_1 and the time spent in the productively infected stage as τ_2. The time in τ_1 is spent in viral penetration, uncoating of viral core, transcription and assembly.

The number of virions produced over the entire infected cell life cycle is directly proportional to $\tau_1 + \tau_2$. It is asked whether the virus might be trying to maximize this quantity in order to optimize "virulence" (a quantity which shall be concretized shortly). The question of why there need be 2 distinct phases and not just one where virion production and release occur simultaneously, also cries out for explanation. Such forms of delineation are called "bang-bang control" or the all-or-none principle and are characterized by a phase of proliferation and then terminal differentiation, and are frequently encountered in optimal biological systems [1].

If the total length of the infected cell lifetime is a measure of "virulence", we can then set a theoretical upper bound on it and then compare it with its actual value from field measurements. This would give us a qualitative understanding of "how far" the virus can still go in optimizing itself e.g. it can be used to determine if the avianinfluenza virus is already as virulent as it can be or is it still sub-optimal.

The rest of the article is organized as follows: Section 2 discusses arguments for optimization in biological systems and Section 3 introduces the principle of "bang-bang control". The hypotheses and questions are posed in Section 4 and Section 5 outlines the mathematical model. Section 6 contains the results and discussions and concluding remarks are presented in Section 7.

OPTIMISATION IN A BIOLOGICAL SYSTEM

Before commencing with the mathematical analysis we state what our modeling philosophy will be and give some justification for employing such an approach. First of all, there is certainly no a priori reason why virus propagation – or any other biological system – should operate in an "optimal" fashion. Indeed there is a substantive issue as to whether the notion of "optimality" can be given an operational meaning for many biological systems. Typically, an organism or a virus is forced to cope with a number of competing influences so that an improvement in one direction involves a sacrifice in another. Thus optimality must be interpreted in a broader sense as a "best compromise" solution. Beyond this consideration, however, there are at least two major reasons why a particular biological system might not be performing its function in the most expeditious fashion. First, despite the fact that one tends to think of natural selection as an inherently optimizing process, improvements on existing mechanisms generally proceed by small modifications of existing structures. Thus, there is ample opportunity for the system to become trapped in "local" maxima; there may be "nearby" structures with higher fitness but to reach them may require a temporary, but fatal, decrease in overall fitness. Second, while the system may be constantly improving, evolution is a slow and erratic process so that any system we examine may not have had time to optimize under existing selective pressures. Both of the objections may be partially circumvented by restricting attention to systems which appear to have been evolutionarily static for a long time. The mammalian immune system and viruses surely fulfil this criterion.

A virus typically only wants to proliferate in a host only so much as to ensure transmission to another host (an exception is the Ebola virus which kills its host so fast as to prevent propagation to another host). Hence it is trying to optimize the basic reproductive ratio R_0 in epidemic models. In vector-borne pathogens, the peak viremia in blood serum is a very good determinant of R_0 [2]. Hence, we assume that the virus is trying to optimize the peak viral load in blood serum (P_v).

BANG-BANG CONTROL

In a seminal paper [1], Perelson et al. examined the mammalian immune system and looked at optimal strategies for B-cell proliferation and differentiation. They used control theoretic principles to analyze the minimum time taken by the immune system to eliminate a fixed amount of antigen in the shortest span of time. The problem briefly stated is as follows: given an initial population of B-cells (which secrete antibodies at a modest rate, proliferate into B-cells or differentiate into plasma cells) and plasma cells (which secrete antibodies at a very high rate but do not proliferate), how do you apportion the total population between B-cells and plasma cells? Does the optimal strategy involve proliferation of B-cells followed by differentiation into plasma cells? Or does it involve simultaneous B-cell proliferation and differentiation? The authors showed using optimal control theory that the optimal strategy for B-cells is to go through a stage of proliferation (to build up their population) and then differentiate into plasma cells. Such a control is called “bang-bang” or all-or-none. It is not immediately evident or intuitive that a strategy of simultaneous B-cell proliferation and differentiation is not optimal.

A parallel is drawn between that work and the problem at hand here, where the infected cell also goes through a phase of production of virions followed by a phase of virion release. The reasons for “bang-bang control” in the infected cell system and its implications are explored in the following sections.

HYPOTHESES AND QUESTIONS

This section explores some of the hypotheses proposed and frames some questions. There are two hypotheses about the non-productively infected stage of the infected cell:

HYPOTHESIS .

The virus is not trying to optimize the duration of the non-productive infected stage (τ_1). Hence this time is exactly equal to the time required for viral penetration, uncoating of viral core, transcription and assembly. The interpretation is that as soon as the first complete virions is assembled, the infected cell immediately proceeds to release the virion i.e. it switches to the next phase of productive infection. The obvious disadvantage of this strategy is that the amount of virions produced would be reduced, compared to an approach in which τ_1 is increased. Clearly this strategy is sub-optimal and we do not explore it further.

HYPOTHESIS 2

The virus is trying to optimize peak viral load and hence viral production. However it cannot increase the duration of the productive infected stage (τ_2). This is so because there are physiological limits imposed by the area and strength of the cell wall, which will constrain the duration of virion release. After a threshold, the cell wall will simply fall apart. It can only increase the duration of the non-productive infected stage (τ_1).

Hypothesis 2a

Having a high τ_1 would imply an increased virion release count. However, this would come at the cost of increased susceptibility to lysis by Cytotoxic T-Lymphocytes (CTLs). A lower τ_1 would reduce the susceptibility to CTL mediated lysis at the expense of a reduced virion count.

Hypothesis 2b

The virus might optimize itself such that it bursts early in the face of a weak antibody response. Conversely, it could burst later (after building up a pool of virions) when confronted with a vigorous antibody response.

There are two attributed research questions:

RQ_1: Why cannot τ1 increase indefinitely?
RQ_1: Why is the optimal control “bang-bang”?

MATHEMATICAL FORMULATION

A standard mathematical model of virus propagation adapted from Baccam et al. [3] is constructed to test the hypothesis. Ordinary Differential Equations (ODEs) are used to represent populations of virus, infected cells and normal cells. The equations are as follows:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\beta T V,$$
$$\frac{\mathrm{d}I_1}{\mathrm{d}t} = \beta T V - k I_1 - \omega_{CTL} I_1,$$
$$\frac{\mathrm{d}I_2}{\mathrm{d}t} = k I_1 - \delta I_2,$$
$$\frac{\mathrm{d}V}{\mathrm{d}t} = p I_2 - c V,$$

where T is target cell population, I_1 – non-productively infected cell population, I_2 – productively infected cell population, V – virus population, β – rate constant of infection, k – rate of death of non-productively infected cells, δ – rate of death of productively infected cells, ω_{CTL} = rate of CTL-mediated lysis of non-productively infected cells, p – number of virions produced per productively infected cell per time step and c – rate of clearance of free virus particles.

We also get $\tau_1 = 1/k$ and $\tau_2 = 1/\delta$.

In this simple ODE model, the population of target cells (normal and uninfected cells) are represented by the variable T. They are also lost due to infection, which is represented by the term $-\beta \cdot T \cdot V$. The non-productively infected cells (I_1) are supplied by the loss from the target-cell pool and die at a rate proportional to their number density with constant of proportionality k. They are also lysed by (Cytotoxic T-Lymphocytes) CTLs at a rate proportional to their density and with a constant of proportionality of ω_{CTL}. Productively infected cells (I_2) are replenished from the non-productive pool and die at a rate proportional to their density and with a constant of proportionality of δ. New virions (V) are produced by infected cells at the rate $p \cdot I_2$ and virions are lost at a rate proportional to the virus concentration with constant of proportionality c (representing antibody-mediated virion clearance).

The variation in ω_{CTL} has been modelled in a time-dependent fashion. Namely it is made to mimic the clonal expansion of a pool of effector CTLs after day 4.

$$\omega_{CTL} = \begin{cases} 0, & t < 4, \\ \Omega \cdot \exp[\theta(t-4)], & t \geq 4. \end{cases}$$

The model was parameterized from a study of experimental infection of Influenza A virus in humans [3]. The model was implemented in the Berkeley Madonna package [14] and the code is freely available for download [15]. The model parameters are shown in Table 1.

Table 1. Estimated parameter values from Baccam et al. [3].

Parameter	β, ml/($TCID_{50}$·day)	δ, day^{-1}	p, ml/($TCID_{50}$·day)	k, day^{-1}	c, day^{-1}	T_0	V_0, $TCID_{50}$/ml
Value	$4{,}9 \cdot 10^{-5}$	4,2	$2{,}8 \cdot 10^{-2}$	3,9	4,3	$4 \cdot 10^{8}$	$4{,}3 \cdot 10^{-2}$

RESULTS AND DISCUSSION

The model as outlined in the previous section, thus parameterized, was used to test the hypothesis.

TEST OF HYPOTHESIS 2A

Restating, having a high τ_1 would imply an increased virion release count with cost of an increased susceptibility to lysis by Cytotoxic T-Lymphocytes (CTLs). A lower τ_1 would reduce the susceptibility to CTL mediated lysis at the expense of a reduced virion count.

As a result, we observed that the optimal strategy was to increase τ_1 till a threshold (in this particular case it was found to be just less than 4 days). Incidentally, day 4 is also the time at which CTL action is initiated. Hence, the optimal strategy for the virus is to continue the non-productively infected phase till just before CTL initiation. Till CTL action is initiated, the virus will continue to build its population. Increasing τ_1 beyond 4 days would lead to loss of produced virions due to CTL-mediated infected cell lysis. Any decrease below 4 days would reduce the total virus production and hence peak viremia. Hence the optimal control is "bang-bang", Figure 1. Bang-bang control strategies have also been known to be optimal in other biological systems like differentiation of B-cells and production of plant seeds [1]. Note that due to the use of a continuous ODE system (which mimics biology more closely) as opposed to a delay-differential equation, some infected cells do burst earlier than day 4.

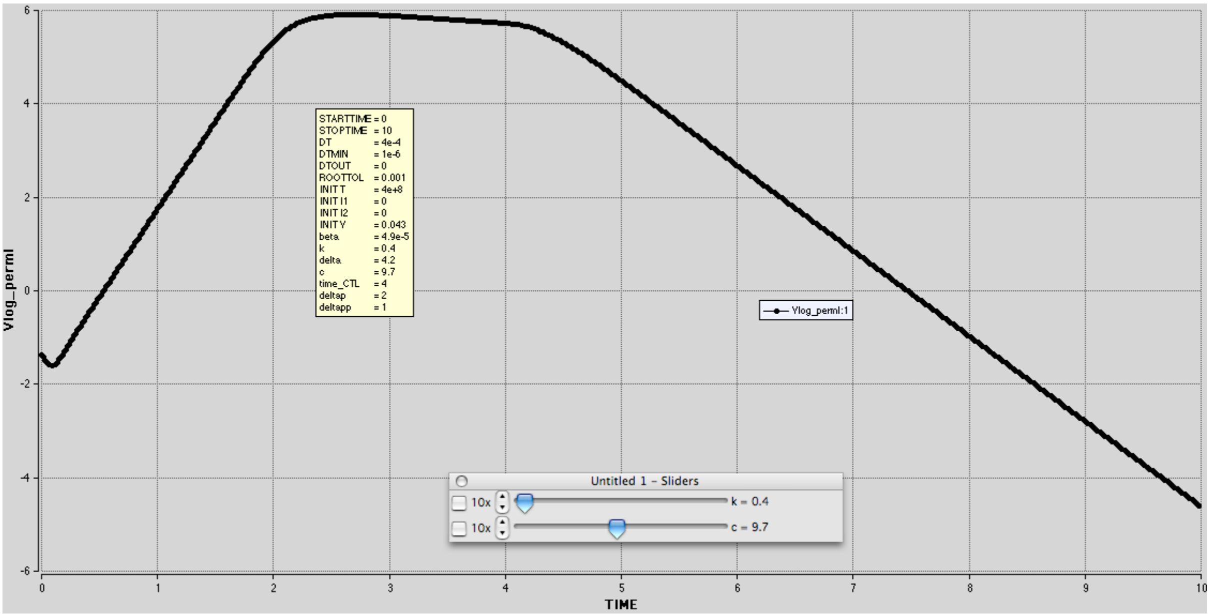


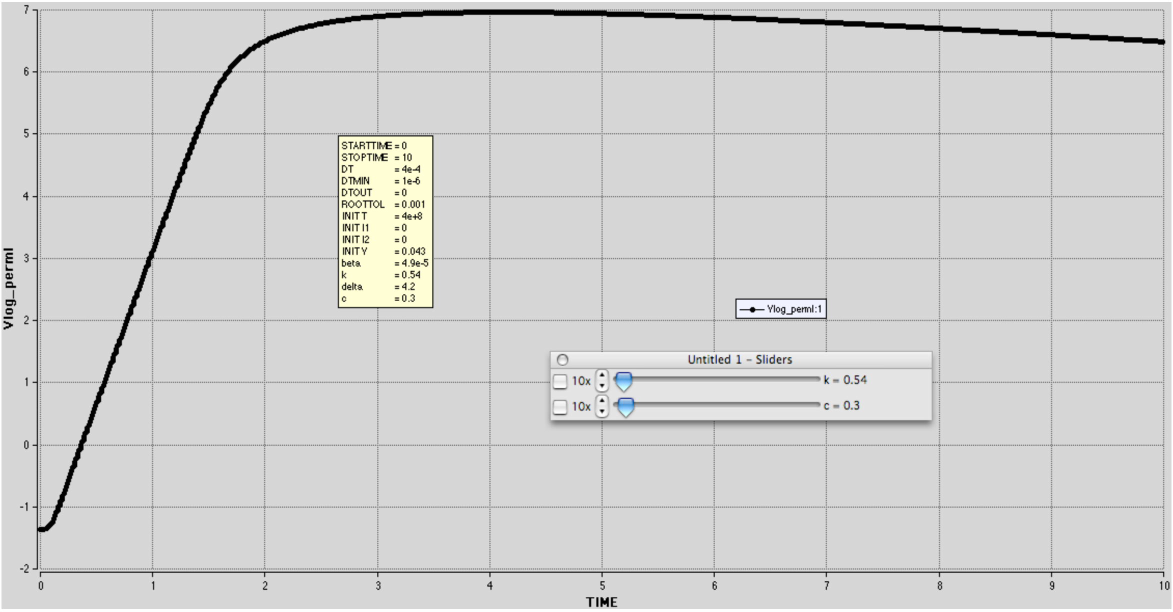
Figure 1. Predicted plot of logged viremia ($\log_{10}TCID_{50}$/ml) versus time in days for Hypothesis 2a.

TEST OF HYPOTHESIS 2B

Restating, the virus might optimize itself such that it bursts early in the face of a weak antibody response. Conversely, it could burst later (after building up a pool of virions) when confronted with a vigorous antibody response.

Result: The antibody response was varied by manipulation of the virion clearance term c in the ODE system. It was found that the optimal strategy remained conserved under variations in the antibody response i.e. the optimal strategy for the virus was always to burst at $\tau_1 = 4$ days. We can reason about this in the following manner: increasing τ_1 beyond 4 days would lead to loss of produced virions due to CTL-mediated infected cell lysis and any decrease below 4 days would reduce the total virus production and hence peak viremia. Hence antibody response has no effect on τ_1 – a fact that is perhaps not intuitively obvious.

a)



b)

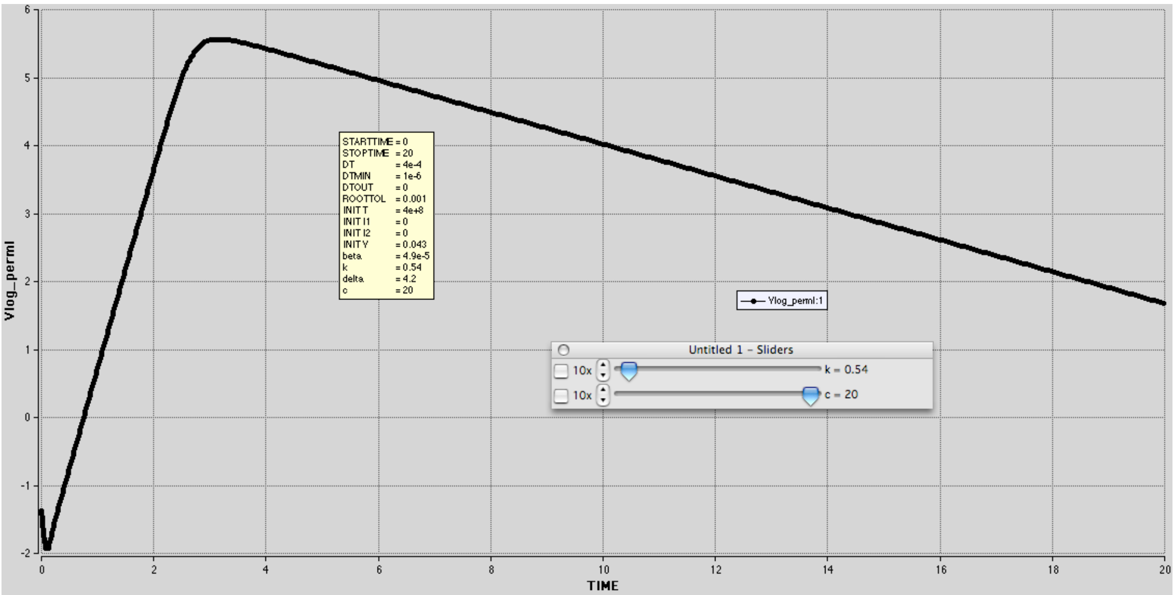


Figure 2. Predicted plot of logged viremia ($\log_{10}TCID_{50}$/ml) versus time in days for Hypothesis 2b, a) low antibody response with τ_1 = 4 days, b) high antibody response with τ_1 = 4 days. The optimal strategy remains the same (burst just before time to CTL initiation).

WHY CANNOT τ_1 INCREASE INDEFINITELY?

From the preceding discussion, it becomes evident that if τ_1 were to increase indefinitely beyond the time to CTL initiation, then there would be a concomitant decrease in virion output due to CTL-mediated infected cell lysis. Hence the time to CTL initiation sets an upper bound on τ_1.

WHY IS THE OPTIMAL CONTROL "BANG-BANG"?

Due to physiological limits on cell wall integrity, the time spent in the productively infected phase (τ_1) must be limited. Any attempt to increase it beyond a threshold would merely cause the whole cell wall to break down. Hence, in order to increase virus production, the only "recourse" the virus has is to increase the time spent in the non-productively infected phase and build up the virus population until onset of CTLs. This naturally gives rise to 2 delineated phases ("bang-bang control"). Any intermediate graded response i.e. virion production occurring simultaneously with release is essentially equivalent to the productively infected phase and since the time that can be spent in it is severely limited, we see that it is a sub-optimal strategy. Such strategies are also optimal in diverse biological systems ranging from differentiation of B-cells in the immune system to allocation of energy to seeds in plants [1].

CONCLUSIONS

This work visits virus proliferation from an optimization viewpoint. A few basic assumptions are made: a) the time spent in the productively infected phase is constant and cannot be subjected to optimization beyond a threshold, and b) the virus is trying to optimize virion production and hence peak viremia. Starting from these assumptions, it is posited that the optimal strategy for virus proliferation is to delay burst till onset of Cytotoxic T-Lymphocytes. This so called "bang-bang control" or all-or-none principle is exhibited in many other biological systems like ant colonies and annual plants [1]. However, optimization may not be the only principle at work. In fact, considerations of reliability may be invoked to explain the presence of long-lived latently infected cells (e.g. HIV). These long-lived cells evade detection by CTLs and ensure a prolonged viremia in hosts.

Another conclusion, which is not intuitively obvious, is the fact that the optimal strategy of allocating the maximum time in the non-productively infected phase remains invariant even in the face of a varying antibody response. This strategy is insensitive to the humoral response and depends only on the time to CTL initiation.

The total length of the infected cell lifetime is a measure of "virulence", and a theoretical upper bound has been set on it. Comparing this value to the actual value from field measurements would give us a qualitative understanding of "how far" the virus can still go in optimizing itself e.g. it can be used to determine if the avian-influenza virus is already as virulent as it can be or is it still sub-optimal. In the case of the Influenza A virus from the Baccam et al. study [3], the theoretical upper bound on τ_1 is around 4 days, whereas the observed is around 12 hours, suggesting that the virus is still operating sub-optimally and still has scope to improve by mutating itself. Insights like these could be crucial for bio-surveillance efforts and help inform strategies to cope with future pandemics caused by virus mutations.

Lastly, it is instructive to note that experimental infections of hosts with non-endemic strains (viral strains that have not co-evolved with the host and hence are not operating in an optimal manner) could affect experiment outcome. This would elicit a lower than normal viral response, since the viral strategy would now be characterized by Hypothesis 1 i.e. the time spent in the productively infected phase (τ_1) would just constitute the time required for viral penetration, uncoating of viral core, transcription and assembly and no more.

Clearly more work needs to be done to verify these arguments and an extensive analytical treatment of these arguments coupled with more experimental work will be the subject of future work. The current work highlights the significance of simple mathematical and dynamical models that reveal insights into biological processes as has been done previously in immunology and cell biology [4-13].

ACKNOWLEDGMENT

The author wishes to acknowledge fruitful discussions in the 2008 Santa Fe Institute Complex Systems Summer School.

REFERENCES

[1] Perelson, A.S.; Mirmirani, M. and Oster, G.F.: *Optimal Strategies in Immunology: B-Cell Differentiation and Proliferation*.
Journal of Mathematical Biology **3**, 325-367, 1976,
http://dx.doi.org/10.1007/BF00275065,

[2] Komar, N., et al.: *Experimental Infection of North American Birds with the New York 1999 Strain of West Nile Virus*.
Emerging Infectious Diseases **9**(3), 311-322, 2003,
http://dx.doi.org/10.3201/eid0903.020628,

[3] Baccam, P., et al.: *Kinetics of Influenza A Virus Infection in Humans*.
Journal of Virology **80**(15), 7590-7599, 2006,
http://dx.doi.org/10.1128/JVI.01623-05,

[4] Banerjee, S. and Moses, M.: *Scale invariance of immune system response rates and times: perspectives on immune system architecture and implications for artificial immune systems*.
Swarm Intelligence **4**, 301-318, 2010,
http://dx.doi.org/10.1007/s11721-010-0048-2,

[5] Banerjee, S.: *Scaling in the Immune System*. Ph.D. Thesis.
University of New Mexico, 2013,
https://digitalrepository.unm.edu/cs_etds/26,

[6] Banerjee, S.; van Hentenryck, P. and Cebrian, M.: *Competitive dynamics between criminals and law enforcement explains the super-linear scaling of crime in cities*.
Palgrave Communications **1**, No. 15022, 2015,
http://dx.doi.org/10.1057/palcomms.2015.22,

[7] Peng, L., et al.: *A bioorthogonal small-molecule switch system for controlling protein function in cells*.
Angewandte Chemie **53**(38), 10049-10055, 2014,
http://dx.doi.org/10.1002/anie.201403463,

[8] Graessl, M., et al.: *A mechano-sensitive excitable system causes local Rho activity oscillations*.
The Journal of Cell Biology **216**(12), 4271-4285, 2017,
http://dx.doi.org/10.1083/jcb.201706052,

[9] Banerjee, S., et al.: The value of inflammatory signals in adaptive immune responses.
In: Lio, P., et al., eds.: *Artificial Immune Systems*. Lecture Notes in Computer Science **6825**.
Springer Verlag, Berlin, pp.1-14, 2011,
http://dx.doi.org/10.1007/978-3-642-22371-6_1,

[10] Banerjee, S. and Moses, M.E.: *A hybrid agent based and differential equation model of body size effects on pathogen replication and immune system response*.
In: Andrews, P.S., et al., eds.: *Artificial Immune Systems*. Lecture Notes in Computer Science **5666**.
Springer Verlag, Berlin, pp.14-18, 2009,
http://dx.doi.org/10.1007/978-3-642-03246-2_5,

[11] Banerjee, S.: *Analysis of a Planetary Scale Scientific Collaboration Dataset Reveals Novel Patterns*.
Preprint. arXiv:1509.07313v2[cs.SI], 2015,
http://dx.doi.org/10.48550/arXiv.1509.07313,

[12] Banerjee, S.: *An Immune System Inspired Approach to Automated Program Verification*.
Preprint. arXiv:0905.2649v1[cs.NE], 2009,
http://dx.doi.org/10.48550/arXiv.0905.2649,

[13] Banerjee, S. and Moses, M.: *Modular RADAR: An immune system inspired search and response strategy for distributed systems*.
In: Hart, E., et al., eds.: *Artificial Immune Systems*. Lecture Notes in Computer Science **6209**,
Springer Verlag, Berlin, pp.116-129, 2010,
http://dx.doi.org/10.1007/978-3-642-14547-6_10,

[14] Macey, R.; Oster, G. and Zahnley, T.: *Berkeley Madonna User's Guide*. Version 8.0.
https://www.berkeleymadonna.com/downloads/BM_Users_Guide_8.0.pdf,

[15] Banerjee, S.: *Model file for Berkeley Madonna for simulations (text file): Optimal Strategies for Virus Propagation*.
http://dx.doi.org/10.13140/RG.2.1.3183.4963.