Bacteriophage Infection Dynamics: Multiple Host Binding Sites

H.L. Smith* and R.T. Trevino

School of Mathematical and Statistical Sciences
Arizona State University, 85287 Tempe, AZ, USA

Abstract. We construct a stochastic model of bacteriophage parasitism of a host bacteria that accounts for demographic stochasticity of host and parasite and allows for multiple bacteriophage adsorption to host. We analyze the associated deterministic model, identifying the basic reproductive number for phage proliferation, showing that host and phage persist when it exceeds unity, and establishing that the distribution of adsorbed phage on a host is binomial with slowly evolving mean. Not surprisingly, extinction of the parasite or both host and parasite can occur for the stochastic model.

Key words: bacteriophage, stochastic simulation algorithm, phage adsorption, persistence, basic reproductive number

AMS subject classification: 92D25, 92D40, 34F05

1. Introduction

Bacteriophage, phage for short, are virus that prey on bacteria. The two are abundant in marine and soil environments. Their short generation time and small size make them an ideal model predator-prey system for ecological and evolutionary studies [5, 26, 16, 14]. Laboratory experiments have shown that a mutation in the bacterial host can make it resistant to phage infection, and occasionally the phage can mutate to overcome this resistance. See recent work of Bohannan and Lenski [2, 3] for more on this arms race.

*Corresponding author. E-mail: halsmith@asu.edu
When first discovered early in the previous century, the potential to use phage against pathogenic bacteria to cure infections, so-called “phage therapy”, was immediately grasped but the emergence of antibiotics put this goal on the back burner. However, as multi-drug resistant bacteria have become an enormous public health problem, there is a renewed interest in phage therapy [17, 20, 22, 23, 4, 30].

Mathematical modeling of the phage-bacteria interaction has a long history. The paper by Campbell [5], still very useful reading today, was among the first to write down a mathematical model—a system of delay differential equations for phage and bacteria system. His focus was on “how phages perpetuate themselves under natural conditions”. Later modeling by Levin, Stewart and Chao [14] in a chemostat setting has been very influential. The more recent work of Schrag and Mittler [26] points out that while numerous studies have shown that phage and bacteria may coexist for long periods in laboratory environments, mathematical models tend to lead to phage over-exploiting its bacterial host, dooming itself to extinction as well. They show that spatial refuges may play a role in allowing persistence of predator and prey. Bohannan and Lenski note that “communities of phage and bacteria are often more stable than predicted by theory” [3].

Phage are classified into two types: virulent and temperate, although the type can depend on the host, a particular virus may be virulent against one host and temperate with respect to another. Here, we are interested exclusively with virulent phage so we describe their life style first. Typically, phage specialize to attack only one or a few strains of bacterial host whose cell surface contains an appropriate binding site. Phage attach to a host binding site, a process termed adsorption. Then inject their DNA into the cell which thereafter is called an infected cell. The infected host cell machinery is then co-opted to make new phage particles which are subsequently released in a burst when phage enzymes cause the host cell to lyse and the cycle repeats. The latent period is the time between phage-host binding and subsequent release of phage at cell lysis, usually on the order of 20 minutes to an hour depending on the host-phage system and on the nutritional state of the host. The burst size, ranging between several to thousands, is a measure of the average number of phage progeny resulting from a single infected host cell. Temperate phage have an alternative option to immediate exploitation of the cell; they can integrate their genome into the host genome and be propagated via host cell division.

Adsorption of phage to bacteria has been shown to follow mass-action kinetics [19, 10, 27]. In addition, early experiments have established and recent electron microscope images have vividly shown (see Wikipedia image) that many phage can adsorb to a single bacterium, up to several hundred [27]. What is the effect of multiple infection of a single bacterium? No doubt this varies with the host-phage system but Ellis and Delbrück [10] note that “no significant difference was found either in the latent period or in the size of the burst” for the E.coli-phage system they studied. No mathematical models of phage with which we are familiar account for multiple adsorption of phage to host, excepting [25]. However, in the host-parasite literature, models which explicitly account for the parasite load of a host are known [18, 29]. In a recent paper [25], one of us introduced a model accounting for multiple adsorption of phage to host. The aim of this work was to use this much more complicated model,
which tracks “host + n-phage” complexes, together with time-scale arguments to obtain a simplified low-dimensional model comparable to traditional models, such as Campbell’s and others [1, 21, 8, 22, 23, 4, 24], but with the effects of multiple adsorption taken into account by a modification of mass-action adsorption kinetics. However, the time-scale arguments in [25] are less than convincing. This motivates us to take a different and simpler approach to modeling host-phage complexes in this paper.

As in [25], we assume each bacterium has $N$ binding sites and we account for the various host-phage complexes, both infected and not yet infected. Then we assign natural reaction rates to various reactions that can occur in order to create a stochastic simulation algorithm (SSA) [11, 12]. Among these reactions are the adsorption reactions:

$$\text{host} + n \text{ phage} \rightarrow \text{host} + (n + 1) \text{ phage}, \quad 0 \leq n \leq N$$

for both infected and uninfected host, where the reaction rate decreases to zero as $n$ increases to $N$, and the fast “injection” reactions:

$$\text{uninfected host} + n \text{ phage} \rightarrow \text{infected host} + n \text{ phage}, \quad 1 \leq n \leq N$$

where one of the adsorbed phage injects its DNA into the host (although now empty of DNA, the phage is assumed to hold its binding site). The rate of this reaction increases linearly
with $n$. The lysis reaction

$$\text{infected host} \rightarrow l \text{ virus}$$

is characterized by a random burst size $l$ of viral progeny where $l$ is poisson with mean $L$. We explore this algorithm using computer simulations, primarily to understand the effects on population persistence of demographic-stochasticity. See [6] for modeling the effects of environmental stochasticity on phage-host dynamics.

We also formulate the “reaction rate equation”, the ordinary differential equation (ode) model that approximates the stochastic model in the limit where species numbers are large and derive the basic reproductive number for phage proliferation at the virus-free equilibrium. It is shown that when the basic reproductive number exceeds one then virus and host persist. A bifurcation diagram for the ODE model, with the burst size as parameter, shows that the host-virus coexistence equilibrium arises through a transcritical bifurcation from the virus-free equilibrium and subsequently loses stability to a super-critical Hopf bifurcation to a stable periodic orbit. As the bifurcation parameter increases further, the minimum values of host and virus along the periodic orbit become extremely small. Numerical simulations of the SSA model do not support the ODE models predictions of persistence in a variety of cases. For example, if viral removal rates are high, the virus may be eliminated before it can find its host. Also, when the ODE model predicts oscillations with low minimum values, the SSA model may track these oscillations for a few cycles before either the virus or both host and virus become extinct.

Finally, we use time scale arguments on the ODE model, this time in a convincing manner, to obtain a reduced ODE where only absorption, injection, and host lysis are included, and show rigorously that the distribution of phage adsorbed to a single (infected or uninfected) bacterium is binomially distributed with a mean that evolves in time, approaching a constant value. Numerical simulations confirm this analysis. From a theoretical perspective, we confirm the suggestion made in [25] that the effect of multiple viral adsorption is most pronounced when the mean number of potential phage adsorption that can occur during the waiting time for phage injection or during the phage latent period substantially exceeds unity. Our numerical simulations aimed at eliciting an effect on free-virus numbers or on bound virus numbers to an increase in the number of host binding sites were mixed. Effects were substantial in the first few hours of “passive phage therapy” where an initial 100-to-one phage-host ratio results in a significant increase in bound virus as the number of host binding sites increases.

We do not consider here potential mutation of host to phage resistance.

2. Stochastic Simulation Model

We construct a stochastic simulation algorithm, following [11, 12], of bacteriophage and host bacterial interaction. Basically, this algorithm allows simulation of the time-series of species numbers (integer-valued) that is faithful to the underlying chemical master equation that
gives the time-evolution of the state-transition probabilities. Necessary terms and definitions are given below, followed by model assumptions.

Phage bind to their bacterial host, then inject their genetic material into the host. Our model allows multiple phage to adsorb to a bacterium. A host bacteria is said to be uninfected so long as none of its adsorbed phage have injected their DNA; it is infected once injection by one of its adsorbed phage occurs. Although the waiting time for an adsorbed phage to inject its DNA is very short, perhaps a minute or so, we include it in our model.

Model assumptions are:

(i) Uninfected bacteria undergo cell division at a rate independent of the number of adsorbed phage; infected bacteria do not divide. The division rate is taken to be either constant or logistic.

(ii) If an uninfected bacterium divides, any adsorbed phage are evenly distributed among daughter cells.

(iii) A bacterium has $N$ binding sites at which a virus may adsorb.

(iv) Phage adsorption to host follows mass-action kinetics but the adsorption rate decreases linearly as the number of occupied sites increases, vanishing when all sites are occupied.

(v) An adsorbed phage injects DNA into the host following an exponentially distributed waiting time with rate independent of the presence of other adsorbed phage.

(vi) After injecting DNA into the host, the empty virus remains attached to the binding site, excluding that site from subsequent re-occupation.

(vii) The latent period is exponentially distributed.

(viii) At host lysis, the burst size is a Poisson-distributed random variable, independent of the number of phage injections into the host.

(ix) Free (un-adsorbed) phage deactivate at a constant rate.

(x) Bacteria are removed at a constant rate independent of infection status and number of adsorbed phage.

Dependent variables and parameters are described in Table 1. See Table 1 of [9] for some values of key phage parameters. Another source for adsorption rates is [30]. Our value for $\rho$, as used in [25], is supported by [13].

A stochastic simulation algorithm is described in an equation-free manner. All that is required is a list of possible reactions among the model “reactants” and the reaction rates,
Table 1: Dependent variables and parameters

<table>
<thead>
<tr>
<th>quantity</th>
<th>description</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_x^i$</td>
<td>uninfected cells with $i$ adsorbed phage</td>
<td>$0 \leq i \leq N$</td>
</tr>
<tr>
<td>$C_y^i$</td>
<td>infected cells with $i$ adsorbed phage</td>
<td>$1 \leq i \leq N$</td>
</tr>
<tr>
<td>$v$</td>
<td>free phage</td>
<td>na</td>
</tr>
<tr>
<td>$X$</td>
<td>total uninfected cells</td>
<td>$\sum_i C_x^i$</td>
</tr>
<tr>
<td>$Y$</td>
<td>total infected cells</td>
<td>$\sum_i C_y^i$</td>
</tr>
<tr>
<td>$b$</td>
<td>max. adsorption rate</td>
<td>$10^{-6} \text{ hr}^{-1}$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>injection rate</td>
<td>$30 \text{ hr}^{-1}$</td>
</tr>
<tr>
<td>$N$</td>
<td>viral binding sites per host</td>
<td>$1 - 100$</td>
</tr>
<tr>
<td>$L$</td>
<td>mean burst size</td>
<td>$50 - 150$</td>
</tr>
<tr>
<td>$\nu$</td>
<td>lysis rate (latent period)$^{-1}$</td>
<td>$2.5 \text{ hr}^{-1}$</td>
</tr>
<tr>
<td>$a_m$</td>
<td>max. host growth rate</td>
<td>$0.3 \text{ hr}^{-1}$</td>
</tr>
<tr>
<td>$K$</td>
<td>host carrying capacity</td>
<td>variable</td>
</tr>
<tr>
<td>$r$</td>
<td>bacterial removal rate</td>
<td>$0 \text{ hr}^{-1}$</td>
</tr>
<tr>
<td>$m$</td>
<td>phage deactivation rate</td>
<td>variable $\text{ hr}^{-1}$</td>
</tr>
</tbody>
</table>

traditionally written above the reaction arrows. In our case, the reactants are the uninfected host-virus complexes, $C_x^0, \ldots, C_x^N$, the infected host-virus complexes, $C_y^1, \ldots, C_y^N$, and the phage, $v$. The reaction rates are constants except possibly for the cell division of uninfected bacteria, $a$, which may decrease with host numbers in a logistic manner as described below. Here are the reactions:

- adsorption: $C_x^i + v \xrightarrow{b_i} C_x^{i+1}$
- adsorption: $C_y^i + v \xrightarrow{b_i} C_y^{i+1}$
- injection: $C_x^i \xrightarrow{\rho} C_y^i$
- cell division: $C_x^i \xrightarrow{a} 2C_x^{i/2}$ $i$ even
- cell division: $C_x^i \xrightarrow{a} C_x^{(i+1)/2} + C_x^{(i-1)/2}$ $i$ odd
- cell lysis: $C_y^i \xrightarrow{\nu} lv$
- cell removal: $C_x^i \xrightarrow{r} Z_x$
- cell removal: $C_y^i \xrightarrow{r} Z_y$
- virus removal: $v \xrightarrow{m} Z_v$

Some clarification is necessary. The species $Z_x, Z_y$ and $Z_v$ denote, respectively, removed uninfected host, removed infected host and removed phage. The burst size $l$, in the “cell lysis” reaction is poisson distributed with mean $L$. Note that this reaction says that one
infected complex lyses, producing $l$ virus. The adsorption rate of phage to a host having $i$ of its $N$ binding sites occupied is given by:

$$ b_i = b \left(1 - \frac{i}{N}\right), \quad 0 \leq i \leq N $$

(2.2)

It decreases to zero as the number of unoccupied binding sites decreases to zero, in contrast to the assumption $b_i = b$, $1 \leq i \leq N$ in [25] where adsorption is assumed to be independent of the number of occupied sites. The rate of cell division of uninfected host is given by:

$$ a = a(X,Y) = a_m \left(1 - \frac{X + Y}{K}\right) $$

(2.3)

for some $K$ with $0 < K \leq \infty$ and $a_m \geq 0$. $K = \infty$ corresponds to exponential growth rate.

Recall that $X = \sum_i C^i_x x$ and $Y = \sum_i C^i_y$.

3. ODE Model

The ODE model, which in SSA terminology is referred to as the “reaction rate equations”, is obtained directly from the set of reactions in a standard way:

$$ (C^i_x)' = b_{i-1} v C^{i-1}_x - (b_i v + i \rho + r) C^i_x + a(C^{2i+1}_x + 2C^{2i}_x + C^{2i-1}_x - C^i_x), \quad i = 0..N, $$

(3.1)

$$ (C^i_y)' = i \rho C^i_x + b_{i-1} v C^{i-1}_y - (b_i v + \nu + r) C^i_y, \quad i = 1..N, $$

$$ v' = -v \left(m + \sum_{i=0}^{N-1} b_i C^i_x + \sum_{i=1}^{N-1} b_i C^i_y\right) + \nu L \sum_{i=1}^{N} C^i_y $$

with (2.2) and (2.3) and where it is assumed that $C^{-1}_x, C^0_y = 0$.

If $C_x$ denotes the $N+1$-vector of $C^i_x$ and $C_y$ denotes the $N$-vector of $C^i_y$, then the system (3.1) is defined on

$$ D = \{(C_x, C_y, v) \in \mathbb{R}^{2n+2}_+ : X + Y \leq K\} $$

which, we show below, is positively invariant. We note that all solutions are bounded if $K < \infty$ and $m > 0$.

Define the moments of the occupancy distribution as

$$ p(t) = \sum_{i=0}^{N} i \frac{C^i_x}{X} = \text{average # occupied sites for uninfected cells} $$

$$ q(t) = \sum_{i=1}^{N} i \frac{C^i_y}{Y} = \text{average # occupied sites for infected cells} $$

$$ Var_x = \sum_{i=0}^{N} (i - p(t))^2 \frac{C^i_x}{X} = \sum_{i=0}^{N} i^2 \frac{C^i_x}{X} - p^2(t) = \text{Occupancy variance} $$
Straightforward but tedious computations give

\[
X' = (a - r - \rho p(t))X \tag{3.2}
\]
\[
Y' = \rho p(t)X - (\nu + r)Y \tag{3.3}
\]
\[
p' = bv(1 - \frac{p}{N}) - ap - \rho Var_x \tag{3.4}
\]
\[
q' = \rho \frac{X}{Y} (Var_x + p(p - q)) + bv(1 - \frac{q}{N}) \tag{3.5}
\]
\[
v' = -bv(1 - \frac{p}{N}) - bvY(1 - \frac{q}{N}) + \nu LY - mv \tag{3.6}
\]

This system is not closed since \(Var_x\) is not determined. We confront the classical moment

\[
\text{Moment:} \quad M(t) = \int_a^b x^n f(x) dx
\]

By the moment generating function:

\[
M(t) = \int_a^b e^{tx} f(x) dx
\]

We seek to find the expected value of the random variable \(X\) using the moment generating function.
The total virus, free plus adsorbed, is given by

\[ V(t) = v + \sum_{i=1}^{N} i(C_x^i + C_y^i) = v + pX + qY \]

We compute \( V' \)

\[ V'(t) = \nu Y(L - q) - r(pX + qY) - mv \]

in order to confirm that bacterial reproduction has no effect on the total viral population.

Observe that the total host population satisfies:

\[ (X + Y)' = aX - \nu Y - r(X + Y) \quad (3.7) \]

If \( K < \infty \) and \( X + Y = K \), then \( a = 0 \) so \( (X + Y)' \leq 0 \), implying that \( D \) is positively invariant.

Aside from the trivial equilibrium in which virus and host are absent, if \( K < \infty \), the virus-free equilibrium of system (3.1) exists and is given by:

\[ C^0_x = \bar{X}, \; C^i_x = C^i_y = 0, \; 1 \leq i \leq N, \; v \equiv 0, \; a_m > r. \quad (3.8) \]

where \( \bar{X} = K^{a_m-r}_{a_m} \).

The basic reproductive number of phage replication in the environment corresponding to the virus-free equilibrium is easy to deduce. A single virus added to this environment has probability \( \frac{v}{\nu + \rho} \) of adsorbing to a (uninfected) bacterium whereupon it must avoid removal before it can inject. The probability of injecting before removal is \( \frac{\rho}{\nu + \rho} \). The resulting infected cell has probability \( \frac{\nu}{\nu + r} \) of lysing before removal thereby producing \( L \) viral progeny; all other possibilities lead to no viral progeny. Therefore, the expected number of progeny of a virus in the environment is:

\[ R_0 = L \frac{a_m - r}{(a_m - r + ma_m/bK)(1 + r/\nu)(1 + r/\rho)} \quad (3.9) \]

Virus can invade the virus-free equilibrium if \( R_0 > 1 \) and the equilibrium is stable to virus invasion if \( R_0 < 1 \). A linearized stability analysis of the equilibrium leads to the same result. The expression for \( R_0 \) simplifies considerably if \( r = 0 \) to \( R_0 = \frac{L}{1 + a_m/bK} \).

In the appendix, we prove the following result for the full system (3.1).

**Theorem 1.** Let \( 0 < K < \infty \) and \( m > 0, a_m > r \geq 0 \). Then the following hold for system (3.1):

(a) the bacteria persist: there exists \( \epsilon_1 > 0 \) such that if \( X(0) > 0 \), then

\[ \liminf_{t \to \infty} X(t) \geq \epsilon_1. \quad (3.10) \]
(b) The virus-free equilibrium exists if \( a_m > r \) and is asymptotically stable if the reproductive number for virus (3.9) satisfies \( R_0 < 1 \) and is unstable if \( R_0 > 1 \).

(c) If \( R_0 > 1 \), then the virus persist: there exists \( \epsilon_2 > 0 \) such that if \( X(0) > 0 \) and \( Y(0) + v(0) > 0 \), then

\[
\liminf_{t \to \infty} v(t) \geq \epsilon_2
\]

Furthermore, at least one positive equilibrium representing coexistence of host and phage exists.

Theorem 1 says that, for the ODE model, the bacterial host cannot be driven to extinction by phage if it is viable \( (a_m > r) \) and if \( R_0 > 1 \) then both phage and host persist. It does not say what dynamical form this coexistence takes.

Figure 3: Bifurcation diagram with parameter \( L \) for \( N = 20, m = 1.0, K = 100000 \).

The dynamical properties of the ODE system are best summed up by the typical bifurcation diagram (Figure 3), with burst size \( L \) as the bifurcation parameter here. Note that the basic reproductive number in this case is \( R_0 = L/11 \). Starting at \( L = 0 \), the phage free steady state is stable (SSS) until undergoing a transcritical bifurcation at the first critical value \( L_1 \approx 11 \) where \( R_0 = 1 \), at which point it becomes unstable (USS), while the coexistence branch becomes positive and stable. For a still larger critical value \( L_2 \approx 62 \) the coexistence state gives way to a Hopf bifurcation with characteristic stable periodic orbits (SPO) of increasing amplitude around an unstable steady state.
4. Numerical Simulations

In all our simulations, only free virus, $v_i$, and uninfected cells with no attached virus, $C^0_x$, are initially nonzero: $C^i_x(0) = C^i_y(0) = 0$, $1 \leq i \leq N$.

To contrast the ODE properties with that of the stochastic simulation algorithm (SSA), we consider two cases where their respective results may vary widely. Even though the ODE system was derived directly from the reactions, it is a deterministic model with the assumption of sufficiently large quantities that the continuous approximation is valid. Observing the bifurcation diagram, it is clear that this assumption is not valid in two cases: first, near the first critical value $L_1$, where quantities are small, and again, well beyond $L_2$ where oscillations repeatedly take quantities near zero. For such smaller quantities there is sufficient probability that bacteria and/or phage may die-off before quantities become large enough to be sustained.

We illustrate these two cases as follow. In Figure 4, the total bacteria population grows nearly identically for both ODE and SSA (thus only ODE is shown). However, for phage,
compare the assured persistence for ODE to the two typical SSA results: elimination of phage (SSA1), and survival of phage (SSA2). In Figure 5 we see the ODE (upper plots) suggest wide oscillations which are unlikely to continue for long in the stochastic case (lower plots). In this example the phage are eliminated while the host achieves its carrying capacity. It also is possible that both host and parasite will perish, particularly for a low virus removal rate, as suggested by Figure 6. In this example, a particular outcome of SSA (lower plots) is shown in which three oscillations occur before luck runs out for the bacteria, extinction of phage follows. Of course, many other outcomes are possible. In the ODE model (upper plots), both persist.

In Figure 7 we illustrate the effect of binding site quantity in the case of $m = 0$. Thus virus loss is due only to adsorption (to binding sites) and allows us to observe the effect of $N$ independently. This graph suggests that the influence of binding site quantity on outcome is minimal for typical adsorption rates found in the literature [9] but becomes more significant as this attack rate is increased.
In Figure 8 we illustrate the effects on the transient dynamics of the ODE model of an increase in the available host binding sites $N$. Ignoring bacterial growth ($a_m = 0$) and viral replication ($L = 0$), the simulation shows viral attachment to bacteria (total bound virus) during the first few hours after introduction of phage at 100-to-one phage-host ratio. Such phage to host ratios are appropriate for so-called “passive phage therapy” which relies on first generation phage to overwhelm bacteria [22, 23]. Note that increasing levels of bound virus mean decreasing numbers of free virus which mean fewer new infections.

Some notes on the specifics of the computational tools and methods follow. The SSA algorithm used is that of Gillespie [11, 12] in its most basic form and implemented using MatLab for arbitrary $N$. No tau-leaping or other accuracy compromising modifications were necessary or required, hence the contrast with the ODE model is maximized. Instead, the main loop was coded in the C language and integrated using MatLab Exchange, resulting in sufficiently fast but exact simulation for even large initial conditions and large $N$ ($N = 100$) to be run in a few minutes on a typical desktop. The ODE system was also simulated using
Figure 7: Effects of N on final total virus population for $C_x^0(0) = 10^4, v(0) = 2 \times 10^6, L = 50, m = 0, K = 50000$.

MatLab. Where required, ODEs were generated automatically using Perl; hand coding up to 202 equations at a time would be tedious and error-prone. The bifurcation diagram was obtained using XPPAUT, requiring about 4 hours for $N = 20$ (42 equations).

5. Reduced ODE Models using Time Scale Arguments

If the number of cell divisions, cell removals, and the number of host-virus bindings during a mean injection time is much smaller than one:

$$\frac{a_m}{\rho}, \frac{r}{\rho}, \frac{b v}{\rho} \ll 1$$

then

$$(C_1^1)' \approx b v C_x^0 - \rho C_1^1 = \rho \left( \frac{b v}{\rho} C_x^0 - C_1^1 \right)$$

so, if $\rho \gg 1$, we have

$$C_1^1 \approx \frac{b v}{\rho} C_x^0.$$  \hspace{1cm} (5.2)

Similarly,

$$(C_2^1)' \approx b_1 v C_1^1 - 2 \rho C_1^1 \approx \rho \left( \frac{b_1 v b v}{\rho} C_x^0 - 2 C_2^2 \right)$$
so

\[ C_x^2 \approx \frac{1}{2} \frac{b_1 v \nu}{\rho} C_x^0 \approx 0 \]

Since injection \( C_x^i \rightarrow C_y^i \) dominates host-virus bindings \( C_x^i \rightarrow C_x^{i+1} \) by (5.1), flow in the vertical direction in the first row of Figure 1 dominates flow in the horizontal direction. Very soon, uninfected cells will have at most one attached phage and we can ignore the \( C_y^i \) for \( i \geq 2 \). This means we can drop the equations for \( C_y^i \) for \( 1 \leq i \leq N \) and use the approximation (5.2) for \( C_y^1 \). The input term due to injection, \( i \rho C_x^i \), into the \( C_y^i \) equation can be neglected for \( i \geq 2 \) and for \( i = 1 \) it is approximately \( b v C_x^0 \). On renaming \( X = C_x^0 \), (3.1) is well-approximated by:

\[
X' = (a - r) X - b v X \\
(C_y^1)' = b v X - (b_1 v + \nu + r) C_y^1 \\
(C_y^i)' = b_{i-1} v C_y^{i-1} - (b_i v + \nu + r) C_y^i, \quad 2 \leq i \leq N, \\
v' = -\nu \left( m + b X + \sum_{i=1}^{N-1} b_i C_y^i \right) + \nu L \sum_{i=1}^{N} C_y^i
\]

Note that, with parameters as in Table 1, (5.1) is satisfied in the sense that all quantities are \( O(10^{-2}) \) provided \( v \leq O(10^5) \).
5.1. Fast-time Scaling

It is always useful to scale variables. Here we scale dependent variables as follows:

\[
\tilde{C}_i^x = C_i^x / X(0), \quad \tilde{C}_i^y = C_i^y / X(0), \quad \tilde{v} = v / v(0)
\] (5.4)

There are several natural time scales. Here, our focus is on the fastest time scale, that of injection:

\[
\tilde{t} = \rho t
\] (5.5)

Thus, a unit of time corresponds to the waiting time for injection, roughly a minute. If we let

\[
\tilde{b}_i = b_i v(0), \quad M = v(0) / X(0),
\]

then the scaled system becomes:

\[
\rho \frac{d\tilde{C}_i^x}{dt} = \tilde{b}_{i-1} \tilde{v} \tilde{C}_i^{a-1} - (\tilde{b}_i \tilde{v} + i \rho + r) \tilde{C}_i^x + a(\tilde{C}_i^{2i+1} + 2 \tilde{C}_i^{2i} + \tilde{C}_i^{2i-1} - \tilde{C}_i^x), \quad i = 0..N,
\]

\[
\frac{d\tilde{C}_i^y}{dt} = i \rho \tilde{C}_i^x + \tilde{b}_{i-1} \tilde{v} \tilde{C}_i^{a-1} - (\tilde{b}_i \tilde{v} + r + \nu) \tilde{C}_i^y, \quad i = 1..N,
\]

\[
\frac{d\tilde{v}}{dt} = -\frac{1}{M} \tilde{v} \left( \sum_{i=0}^{N-1} \tilde{b}_i \tilde{C}_i^x + \sum_{i=1}^{N-1} \tilde{b}_i \tilde{C}_i^y \right) + \nu LX(0) \sum_{i=1}^N \tilde{C}_i^y - m \tilde{v}
\] (5.6)

Observe that if \( v(0) = O(10^6) \) then, using parameters as in Table 1, \( \tilde{b} \approx 4 \) is comparable to \( \rho \). Clearly, the reproduction and removal terms in (5.6) are negligible since \( m / \rho, \frac{r}{\rho}, \frac{a_m}{\rho} \ll 1 \). If the multiplicity of infection \( M \) is large, the dominant term in the virus equation is accumulation of fresh virus by host lysis.

Although we will not use the scaled equations (5.6) directly, they motivate us to ignore host cell division and host and virus removal. Our aim is to study the dynamics of phage attachment to host. We begin by ignoring host division and removal. If \( a_m = r = 0 \), then system (3.1) has two families of equilibria. The virus-free family: \( v = 0, \quad C_i^x = C_i^y = 0, \quad 1 \leq i \leq N \) and \( C_i^0 \geq 0 \) arbitrary; and the host-free family: \( C_i^x = 0, C_i^y = 0 \) for all \( i, j \) and \( v \geq 0 \) arbitrary (\( m = 0 \)) or \( v = 0 \) (\( m > 0 \)). As expected in this case, all host are lysed and solutions approach a host-free equilibrium.

**Proposition 2.** Let \( a_m = 0 \). If \( m = 0 \), then every solution converges to a virus-only equilibrium:

\[
v(t) \to v_\infty, \quad C_i^x(t) \to 0, \quad C_i^y(t) \to 0
\]

If \( m > 0 \), then every solution converges to zero:

\[
v(t) \to 0, \quad C_i^x(t) \to 0, \quad C_i^y(t) \to 0.
\]
5.2. The dynamics of multiple phage attachments to host

The viral inactivation rate $m$ is small relative to other rates and it is useful to ignore it, as well as host growth and removal ($m = a_m = r = 0$). In this case, we see from Proposition 2 that all host cells are eventually infected, lyse and release phage particles, leading to a limiting value $v_\infty$ of phage. We seek to estimate the total free phage at the end of the “epidemic”, $v_\infty$. Integrating the equation for $X + Y$:

$$-(X(0) + Y(0)) = -\nu \int_0^\infty Y dt$$

and by integrating $V$ equation:

$$V(\infty) - V(0) = v_\infty - V(0) = -\nu \int_0^\infty Y(L - q(t)) dt$$

so

$$\frac{v_\infty - V(0)}{X(0) + Y(0)} = \frac{\int_0^\infty Y(L - q(t)) dt}{\int_0^\infty Y dt} = L - \hat{q}$$

for some $\hat{q}$ belonging to the closure of the range of $q$, by the mean value theorem for integrals. In the special case that $X(0) = C_x^0(0), Y(0) = 0$, then $V(0) = v(0)$ so we have

$$v_\infty = v(0) + X(0)(L - \hat{q})$$

(5.7)

This estimate is crude. $Y$ converges to zero while $q$ rises from zero to $N$ so $\int_0^\infty Y q dt$ is hard to estimate. We would like to have an explicit formula for $v_\infty$ given the initial data.

The uninfected phage occupancy fraction $p_i = C_x^i / X$ satisfies

$$p_i' = b_i - vp_i + \rho p_i(p - i)$$

(5.8)

The numerics suggests that bacterial loading of cells appears to follow a binomial distribution. That is, $p_i = C_x^i / X$ satisfies

$$p_i = \binom{N}{i} \theta^i(1 - \theta)^{N-i}$$

(5.9)

where

$$\theta = \theta(t) = \text{Probability a binding site occupied } \in (0, 1)$$

If true, then the mean and variance are known:

$$p = N\theta, \quad Var_x = N\theta(1 - \theta)$$

(5.10)

Putting the ansatz (5.9) and $p = N\theta$ into (5.8) and using

$$p_i' = \binom{N}{i} \theta^{i-1}(1 - \theta)^{N-i-1}[i - N\theta] \theta'$$

127
dividing out the common factor \( \binom{N}{i} \theta^{i-1}(1-\theta)^{N-i-1} \) from both sides results, after some algebra, in the differential equation

\[
(i - N\theta)\theta' = \rho(1 - \theta)(i - N\theta) \left( \frac{bv}{N\rho} - \theta \right)
\]

so, on dividing out the common factor (the equation holds for each \( i, \ 0 \leq i \leq N \)), we have

\[
\theta' = \rho(1 - \theta) \left( \frac{bv}{N\rho} - \theta \right) \tag{5.11}
\]

Since \( v(t) \to v_\infty \), \( \theta(t) \) will converge to the solution of the autonomous equation obtained by substituting \( v_\infty \) into the above equation. Thus, if \( \frac{bv_\infty}{\rho N} \in (0, 1) \), then

\[
\begin{align*}
\theta(t) & \to \frac{bv_\infty}{\rho N} \\
p(t) & \to \frac{bv_\infty}{\rho} \\
Var_x(t) & \to \frac{bv_\infty}{\rho} \left( 1 - \frac{bv_\infty}{\rho N} \right) \tag{5.14}
\end{align*}
\]

Now consider the fraction of uninfected host \( u = \frac{X}{X+Y} \). We find that

\[
u u + \left[ \nu - \rho p \nu - u \right]
\]

As \( p \to \frac{bv_\infty}{\rho} \), we find that \( u' < 0 \) for large \( t \) if the number of bindings per latent period exceed unity:

\[
\frac{bv_\infty}{\nu} > 1
\]

According to Table 1, this should hold in all cases considered here. Thus, \( u \to 0 \) implying that

\[
X/Y \to 0 \tag{5.15}
\]

The equation for the infected phage occupancy fraction \( q_i = C_i/Y \) is given by:

\[
q_i' = \rho \frac{X}{Y} (ip_i - q_ip) + b_{i-1}v_{q_{i-1}} - b_i v q_i \tag{5.16}
\]

Making the analogous binomial ansatz for \( q_i \) as for \( p_i \):

\[
q_i = \binom{N}{i} \alpha^i (1-\alpha)^{N-i} \tag{5.17}
\]

where

\[(0, 1) \cdot \alpha(t) = \text{Probability binding site is occupied}\]
and using that \( X/Y \to 0 \) to drop the first term in (5.16), we find the analogous equation to (5.11) is:
\[
\alpha' = 1 - \frac{\alpha}{N}
\]  
Consequently:
\[
\alpha(t) \to 1 \\
q(t) = N\alpha(t) \to N \\
Var_y(t) = N\alpha(1 - \alpha) \to 0
\]  
The following figures depict simulations supporting our analysis. The second row of plots in Figure 9 give the distribution of \( C^x_i \) and \( C^y_i \) with respect to \( i \) at the final time. The estimate \( v_\infty = 2 \times 10^6 \) from Figure 9 allows us to estimate the asymptotic moments from (5.12) as:
\[
p \approx 0.27, \ Var_x \approx 0.27, \ q = 20
\]
These values are corroborated in Figure 10. Figure 11 depicts the time series of the phage occupancy distribution on infected host.

6. Appendix

The Jacobian matrix \( J \) at the virus-free equilibrium \( (\bar{X},0, \cdots, 0) \) \( (\bar{X} = K \frac{a_m - r}{a_m}) \) has \( 2N + 2 \) rows and columns. It’s a quasipositive matrix (off-diagonal entries nonnegative) so it has a real eigenvalue having maximal real part among all eigenvalues. We focus on this dominant real eigenvalue. The structure of \( J \) is easily described by first removing the first and last rows and columns corresponding to variables \( C^0_x \) and \( v \) to obtain the matrix with \( 2N \) rows and columns corresponding to variables \( C^i_x, C^i_y, 1 \leq i \leq N \). The latter, \( JC \), decomposes into four size \( N \) sub-matrices:
\[
JC = \begin{pmatrix}
\text{upperdiag} & 0 \\
\text{diag}_1 & \text{diag}_2
\end{pmatrix}
\]
where upperdiag is upper diagonal with \(-r - \rho, -r - 2\rho, -2r - 3\rho, \cdots, -2r - N\rho\) on the diagonal, \( \text{diag}_1 \) has \( \rho, 2\rho, \cdots, N\rho \) on the diagonal, and \( \text{diag}_2 = -(\nu + r)I \). The first and last rows and columns of \( J \) are:

row 1 of \( J \) = \((-a_m - r), -a_m + 2r, -a_m - r, \cdots, -(a_m - r), 0, \cdots, 0, -b\bar{X})

row \( 2N+2 \) of \( J \) = \((0, \cdots, 0, -\nu L, \cdots, -\nu L, -m - b\bar{X})

column 1 of \( J \) = \((-a_m - r), 0, \cdots, 0)^T

column \( 2N+2 \) of \( J \) = \((-b\bar{X}, b\bar{X}, 0, \cdots, 0, -m - b\bar{X})^T

It is easily seen from the first column of \( J \) that \(-a_m - r\) is an eigenvalue; we may now focus on the size \( 2N + 1 \) sub-matrix \( J' \) obtained by removing the first row and column of
Figure 9: Time series and occupancy distribution for $N = 20, C_x(0) = 10^4, v(0) = 10^6, a_m = 0, b = 4 \times 10^{-6}, L = 100, m = 0, K = \infty$.

\[ J. \] By expanding $J' - \lambda I$ along its last column, starting with the lower right entry, and expanding the co-factor of the upper right entry along the second row, then along the $N+1$ column, we obtain the polynomial whose roots give the remaining eigenvalues:

\[
P(\lambda) = (-\lambda - m - b\bar{X})(-\lambda - \nu - r)^N(-\lambda - \rho - r)(-\lambda - 2\rho - r) \cdots (-\lambda - N\rho - r)
+ b\bar{X}\nu L(-\lambda - 2\rho - r) \cdots (-\lambda - N\rho - 2r)(-\lambda - \nu - r)^{N-1} = (-\lambda - 2\rho - r) \cdots (-\lambda - N\rho - 2r)(-\lambda - \nu - r)^{N-1}Q(\lambda)
\]

where

\[
Q(\lambda) = (-\lambda - m - b\bar{X})(-\lambda - \nu - r)(-\lambda - \rho - r) + b\bar{X}\rho\nu L \quad (6.1)
\]

Clearly, we may focus on the cubic $Q$. $-Q$ is strictly increasing for $\lambda \geq 0$ and therefore there is a positive real root if and only if $-Q(0) < 0$ while all roots have negative real part.
if and only if \(-Q(0) > 0\). As \(Q(0) = (m + b\bar{X})(\nu + r)(\rho + r) - b\bar{X}\nu L\), all eigenvalues of \(J\) have negative real part if and only if \(R_0 < 1\), and a positive eigenvalue exists if and only if \(R_0 > 1\).

**Proof of Proposition 2.** We first show that solutions are bounded. The total number of host cells \(X + Y\) is decreasing by (3.7); integrating the differential equation for \(X + Y\) shows that \(\int_{0}^{\infty} Y(t)dt < \infty\). As \(v' \leq \nu LY\), we may conclude that

\[ v(t) \leq v(0) + \nu L \int_{0}^{t} Y(s)ds \leq v(0) + \nu L \int_{0}^{\infty} Y(s)ds. \]

Thus, solutions are bounded.

It follows from the Lyapunov-LaSalle Theorem using the Lyapunov function \(L = X + Y\) that all solutions converge to the set of virus-only equilibria. The fact that solutions converge to a single equilibrium, in the case that \(m = 0\), follows from center manifold theory because all eigenvalues of the linearization at any point of this manifold of virus-only equilibria have negative real part except for a simple zero eigenvalue. See Theorem 2 (b) of [7]. Indeed, if we order the variables \((C_0^x, C_1^x, \cdots, C_N^x, C_1^y, \cdots, C_N^y, v)\), then the \(2N + 2 \times 2N + 2\)-Jacobian at a virus-only equilibrium is lower triangular and given in block form as:

\[
J = \begin{pmatrix}
B_x & 0 & 0 \\
D & B_y & 0 \\
\alpha & c & 0
\end{pmatrix}
\]
where $B_x$ is size $N + 1$ square, $B_y$ is size $N$ square, $D$ is $N \times N + 1$, $a$ is a $1 \times N + 1$, $c$ is a $1 \times N$, and the 0 in lower right corner is a $1 \times 1$ scalar. Matrix $B_x$ is lower triangular with its eigenvalues:

$$-(b_i v + i \rho), \ 0 \leq i \leq N$$
on the diagonal. Here, $v$ is the sole nonzero coordinate of the equilibrium. Similarly, $B_y$ is lower triangular with its eigenvalues:

$$-(b_i v + \nu), \ 1 \leq i \leq N$$

Therefore, all eigenvalues are negative except for a simple zero eigenvalue.

**Proof of Theorem 1.** the system (3.1) is defined on $D = \{(C_x, C_y, v) \in \mathbb{R}_+^{2n+2} : X+Y \leq K\}$ which is positively invariant since the vector-field points inward or tangential to the boundary.
hyperplanes. As a consequence, solutions are defined for all \( t \geq 0 \) and, since \( X + Y \leq K \), we have \( v' \leq -mv + \nu LK \). It follows that there is a compact global attractor.

We use the persistence function \( P(C_x, C_y, v) = X \), a measure of the total uninfected host, to establish host persistence. The set \( X_0 = \{(C_x(0), C_y(0), v(0)) : X(t) = 0, t \geq 0\} \) consists of points where \( C_x(0) = 0 \), or equivalently, \( X(0) = 0 \). All solutions starting in this positively invariant set, where only infected host and virus exist, converge to the trivial equilibrium. Indeed, \( Y' = -(\nu + r)Y \) and \( v' \leq -mv + \nu LY \) together imply this. Now it suffices to show that the trivial equilibrium is uniformly weakly \( P \)-repelling in the sense that there exists \( \epsilon > 0 \) such that \( \limsup_{t \to \infty} (X(t) + Y(t) + v(t)) \geq \epsilon \) whenever \( X(0) > 0 \). If this were false, then for any \( \epsilon > 0 \) there is a solution with \( X(0) > 0 \) such that \( \limsup_{t \to \infty} (X(t) + Y(t) + v(t)) < \epsilon \) and, since \( X(0) > 0 \) implies \( C_x^0(t) > 0 \) for all \( i \) and \( t > 0 \), by translating the solution in time, we can assume that \( X(t) + Y(t) + v(t) < \epsilon \), \( t \geq 0 \) and that \( C_x^0(t) > 0 \), \( t \geq 0 \). Now, \( (C_x^0)' \geq (a_m(1 - \frac{X + Y}{K}) - bv - r)C_x^0 \) so, by choosing \( \epsilon \) small enough, \( (a_m(1 - \frac{X + Y}{K}) - bv - r) > \delta \) for some \( \delta > 0 \). This leads to the contradiction that \( C_x^0(t) \nearrow \infty \). Therefore, we conclude that the zero equilibrium is uniformly weakly \( P \)-repelling and it follows from the acyclicity theorem, see [28], that uniform-\( P \) persistence holds.

We use the persistence function \( P(C_x, C_y, v) = \min\{X, v\} \) to establish phage persistence. The set \( X_0 = \{(C_x(0), C_y(0), v(0)) \in D : \min\{X(t), v(t)\} = 0, t \geq 0\} \) consists of points of the \( C_x^0 \)-axis that belong to \( D \) (where \( v(t) \equiv 0 \)) together with points where \( X(0) = 0 \) (where \( X(t) \equiv 0 \)). On the positively invariant set \( X_0 \), solutions converge to either the trivial equilibrium (TE) or the virus-free equilibrium (VFE), the latter only if they start on the positive \( C_x^0 \)-axis. We claim that the set \{VFE, TE\} is an acyclic covering of the asymptotic dynamics on \( X_0 \). While there is a heteroclinic orbit in \( X_0 \) connecting TE to VFE there is no heteroclinic orbit connecting VFE to TE in \( X_0 \). In the previous paragraph, we showed that TE is uniformly weakly \( P \)-repelling for \( P = X \) so it must also be uniformly weakly \( P \)-repelling for \( P = \min\{X, v\} \). We must show that the VFE is uniformly weakly \( P \)-repelling in the sense that there exists \( \epsilon > 0 \) such that \( \limsup_{t \to \infty} d((C_x(t), C_y(t), v(t)), VFE) \geq \epsilon \) for every solution starting in \( D \) with \( X(0), v(0) > 0 \). Arguing as above by contradiction, for arbitrarily small \( \epsilon > 0 \), there is a solution with \( X(0), v(0) > 0 \) such that \( |C_x^0(t) - \bar{X}| < \epsilon \) and \( C_x^1(t), v(t) < \epsilon \), \( z = x, y, i = 1, \cdots, N \) for all \( t \geq 0 \). For such a solution, we have the estimate:

\[
\begin{align*}
(C_x^1)' & \geq b(\bar{X} - \epsilon)v - (\rho + r + b_1\epsilon)C_x^1 \\
(C_y^1)' & \geq \rho C_x^1 - (\nu + r + b_1\epsilon)C_y^1 \\
v & \geq -v \left( m + b(\bar{X} + \epsilon) + 2(N-1)\epsilon \right) + \nu LC_y^1
\end{align*}
\]

If we let \( z = (C_x^1, C_y^1, v)^T \) then this estimate can be written as

\[
z' \geq (A + \epsilon B)z
\]

where

\[
A = \begin{pmatrix}
-\rho - r & 0 & b\bar{X} \\
\rho & -\nu - r & 0 \\
0 & \nu L & -m - b\bar{X}
\end{pmatrix}
\]

133
is a quasipositive and irreducible matrix whose characteristic polynomial is precisely the cubic (6.1). As $R_0 > 1$, we know that the eigenvalue of $A$ with largest real part, $\lambda_0$, is positive and that there exists a positive left eigenvector $u = (u_1, u_2, u_3)$, $u_i > 0$. We have $uA = \lambda_0 u$. If $\epsilon > 0$ is small enough, these same properties hold for the quasipositive and irreducible matrix $A + \epsilon B$: there is a positive left eigenvector $u^*$ corresponding to a positive eigenvalue $\lambda_0^*$. Multiplying both sides of our differential inequality for $z$ by $u^*$, we obtain

$$(u^* z)' \geq \lambda_0^*(u^* z)$$

since $(u^* z)(0) > 0$ and $\lambda_0^* > 0$ we obtain the contradiction that $u^* z \rightarrow \infty$. This contradiction proves that VFE is uniformly weakly $P$ repelling. $P$ persistence follows again by the acyclicity theorem [28].

The existence of at least one positive equilibrium at which bacteria and virus coexist follows from persistence and [31, Theorem 1.3.6].

References


