This is an author produced version of *Non-lethal control of the cariogenic potential of an agent-based model for dental plaque*.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/80643/

**Article:**

http://dx.doi.org/10.1371/journal.pone.0105012
Non-lethal control of the cariogenic potential of an agent-based model for dental plaque
David A. Head$^{1,*}$, Phil D. Marsh$^{2,3}$, Deirdre A. Devine$^3$

$^1$ School of Computing, University of Leeds, Leeds LS2 9JT, United Kingdom.
$^2$ Microbiology Services, PHE Porton, Salisbury SP4 0JG, United Kingdom.
$^3$ Department of Oral Biology, School of Dentistry, University of Leeds, LS2 9JT, United Kingdom.
* E-mail: d.head@leeds.ac.uk

Abstract

Dental caries or tooth decay is a prevalent global disease whose causative agent is the oral biofilm known as plaque. According to the ecological plaque hypothesis, this biofilm becomes pathogenic when external challenges drive it towards a state with a high proportion of acid-producing bacteria. Determining which factors control biofilm composition is therefore desirable when developing novel clinical treatments to combat caries, but is also challenging due to the system complexity and the existence of multiple bacterial species performing similar functions. Here we employ agent-based mathematical modelling to simulate a biofilm consisting of two competing, distinct types of bacterial populations, each parameterised by their nutrient uptake and aciduricity, periodically subjected to an acid challenge resulting from the metabolism of dietary carbohydrates. It was found that one population was progressively eliminated from the system to give either a benign or a pathogenic biofilm, with a tipping point between these two fates depending on a multiplicity of factors relating to microbial physiology and biofilm geometry. Parameter sensitivity was quantified by individually varying the model parameters against putative experimental measures, suggesting non-lethal interventions that can favourably modulate biofilm composition. We discuss how the same parameter sensitivity data can be used to guide the design of validation experiments, and argue for the benefits of in silico modelling in providing an additional predictive capability upstream from in vitro experiments.

Introduction

Dental caries is a common disease that reduces quality of life globally and represents a substantial economic burden for health organisations [1]. It is caused by acids (primarily lactic) in the oral cavity which initiate demineralisation of tooth enamel, resulting in lesions that can develop into cavities. These acids are produced as a by-product of the glycolysis of dietary sugars by the bacterial community that comprises the oral biofilm known as dental plaque. However, there is no single pathogenic species responsible for caries. Although Streptococcus mutans has been widely implicated and studied in this context, carious lesions can exist in the absence of S. mutans and be absent in its presence [1, 2]. It is instead thought that any acid-producing species that can metabolise sugars at low pH, i.e. those that are both acidogenic and aciduric, can contribute to the caries process. This description includes S. mutans, but also other mutants streptococci, lactobacilli, bifidobacteria, and others. Clinical trials have confirmed the correlation between biofilm composition and caries progression, with healthy and diseased sites associated with distinct subpopulations of bacteria [3–5].

The bacterial composition of dental plaque, and biofilms in general, is the result of a dynamic interplay between microbial physiology and external perturbations from the environment and host, and in this sense can be meaningfully regarded as ecosystems [6–8]. For oral biofilms, this has been formalised into the ecological plaque hypothesis [9]. Varying external conditions therefore alter biofilm composition, potentially increasing or decreasing the fraction of cariogenic species, and in vitro studies have shown that the low pH resulting from pulsing glucose into mixed species biofilms causes population shifts favouring...
S. mutans and lactobacilli [10,11]. This ecological perspective suggests alternative therapeutic strategies: rather than eradication of bacteria, which is both difficult and undesirable as many oral bacteria are also beneficial to the host, it may be possible to instead modulate the biofilm composition to favour bacterial communities with a lower fraction of acidogenic, aciduric species. Indeed, in vitro studies have demonstrated that fluoride below lethal concentrations can reduce the fraction of putative cariogenic organisms [12,13]. The mechanism is thought to be a reduction of their aciduricity and acidogenicity [14–17], and reducing their competitiveness with respect to non-pathogenic species such as S. gordonii [18–20].

Identifying targets to modulate biofilm composition to the benefit of the host is challenging due to the significant complexity of the system, with many coupled mechanisms driving population changes. Systematically varying the many candidate factors in in vitro experiments is costly both in time and expense. It is desirable to introduce an additional predictive layer before in vitro modelling to highlight promising targets, and this can be realised by in silico modelling, i.e. computational simulation of mathematical models. Such models generate quantitative predictions over broad ranges of parameter space in relatively short time. In addition, they are not restricted to specific species of bacteria but can systematically incorporate strain and sub-strain variation by continuously varying parameters related to cellular physiology. They are thus well suited to probing populations of species defined by their function, rather than their genetic identity. Early mathematical models of dental plaque by Dibdin et al. adopted a continuum approach in which concentrations of various dispersed phases varied smoothly with distance from the enamel surface [21–26]. This approach has recently been advanced by Ilie et al. who included numerous coupled fields to more realistically represent acid buffering, polyglucose storage etc. [27]. These studies did not consider the changes in biofilm composition necessary to study population response to perturbations. Such questions can be naturally addressed using agent based modelling, which is established in biofilm research [28–36] but has not yet been applied to plaque.

The aim of this paper is to describe findings from an agent based model of supragingival plaque, i.e. that component of the oral biofilm above the gumline that is responsible for dental caries, developed to probe the relationship between evolving biofilm composition and cariogenic potential. The model consists of two competing populations of bacteria, one that is pathogenic in that it is both acidogenic and aciduric, and a second non-pathogenic population which, while acidogenic, cannot metabolise sugars at low pH and thus is non-aciduric. These populations are labelled A and NA respectively, for ‘aciduric’ and ‘non-aciduric’. It is found that one population or the other dominates at late times, with a ‘tipping point’ between the two outcomes that depends on a range of parameters relating to both intrinsic physiological processes and external factors such as the frequency of sugar intake. Treatments intended to modify these parameters could therefore drive plaque composition towards a healthy, non-cariogenic state. In addition, sensitivity analysis reveals the relative importance of each model parameter on putative experimental measurements which, as well as highlighting the most important mechanisms relevant to plaque function, can also guide validation experiments by identifying parameter-measurement pairings with high sensitivity, suitable for fixing input parameters from in vitro data.

Results

Interspecific competition

Snapshots of a section of a biofilm for the primary parameters in Table 1 are given in Fig. 1 for the initial conditions at \( t = 0d \), at an early time \( t = 10d \), and at the final time point \( t = 200d \). This time scale is relevant to biofilm accumulation at stagnant sites, such as the approximal surfaces between teeth. Even for the early time point there is visible aggregation of each distinct population compared to the initial condition, and these aggregates evolve into the depth-spanning domains visible at the later time. This is primarily a consequence of daughter particles remaining localised to their mother during division, driving the aggregation through shared proximity of descendants from the same progenitor. Note also that there
is a clear gradient in the concentration of acid produced by the biofilm, with higher concentrations near the enameled surface and lower concentrations as one moves through the biofilm into the saliva layer, as seen in real plaque [37]. Lateral gradients in acid are not visible in these snapshots but are also present, in particular at late times when the large domains of A produce more acid than NA. Larger, colour images and movies are available in Figure S1 and Movies S1 and S2 of the Supplementary Information.

The two populations A and NA are not competing for the sole carbon source (glucose), since there is no mass transfer limitation of glucose as the model is defined. They are however competing for space, in that the fixed system size imposed by the plaque thickness \( h_{\text{plaque}} \) restricts the total number of particles of either population, \( N_{\text{tot}} = N_A + N_{\text{NA}} \), to remain roughly constant in time. Therefore, if one population exhibits a net growth rate faster than the other, it will increase its fraction within the biofilm at the expense of the other, whose fraction will contract. The only other form of direct interaction between the two populations is the acid produced which, particularly during the glucose pulse, severely reduces the growth of NA relative to A. Since A produces more acid than NA, this drives the formation of lateral pH gradients mentioned above.

**Alternating differential growth**

The dominant population at late times is that with the greater net growth rate when averaged over both the inter and intra-pulse periods. If one population exhibited the greater rate in both periods, it would clearly outgrow the other. However, for this study we have chosen parameters for which population A exhibits the greater growth rate during the pulse when the pH is low, but population NA grows the fastest between glucose pulses when the pH is high. This situation is representative of oral bacteria as determined in chemostat experiments, for instance *S. mutans* is more aciduric that *S. gordonii*, but exhibits lower growth around neutral pH [38, 39]. Although the parameters were varied to cover a broad range of values, encompassing functionally similar species and strains, they were limited to ensure this basic alternation between the most competitive population remained true.

An example of this alternation is given in Fig. 2, which shows both the pH measured at the enamel surface, and the fraction of the biofilm that belongs to population A, i.e. \( N_A/N_{\text{tot}} \), during three consecutive glucose pulses. For clarity of presentation the long inter-pulse periods have been compressed in this diagram, but the behaviour of both quantities can be inferred as explained in the caption. It is evident that the fraction \( N_A/N_{\text{tot}} \) increases during the glucose pulse, but decreases again before the start of the subsequent pulse, when it again increases. Since A is more aciduric than NA, more acid should be produced as the fraction \( N_A/N_{\text{tot}} \) increases, and this is also evident in the figure where there is a clear correlation between pH and biofilm composition.

The example shown in Fig. 2 is for a somewhat frequent glucose pulsing with \( t_{\text{durn}} = 4h \). In this case, the increase in \( N_A/N_{\text{tot}} \) during the pulse is greater than the decrease between pulses. The long-term trend is therefore for A to dominate, and correspondingly the pH during pulses to drop. This can be regarded as a ratchet effect whereby each full pulse cycle increases the fraction of A by a small amount. The small shifts in composition and pH in the diagram lead to biologically significant changes when integrated over many such cycles. It might be expected that a longer \( t_{\text{durn}} \), and hence a longer period of time spent in the high-pH environment, will result in a larger drop in \( N_A/N_{\text{tot}} \) between pulses and a net decrease of \( N_A/N_{\text{tot}} \), resulting in a slow drift towards a NA-dominated state. This is indeed observed, but instead we now consider data averaged over whole pulses for which ratcheting is not visible, to better focus on long-term trends.

**Transition between homeostatic and pathogenic biofilms**

Fig. 3 shows the fraction of population A versus time for pulse duration \( t_{\text{durn}} \) increasing from 2h to 10h, where each data point corresponds to the averaged value over a 2d period so that the sawtooth variation in Fig. 2 is not visible. Starting from a 50:50 population of A:NA, the biofilm composition
becomes increasingly A-dominated with time for low $t^{\text{durn}}$ corresponding to frequent glucose pulses, i.e. frequent acid challenges. By contrast, for infrequent pulses with high $t^{\text{durn}}$, $N_A/N_{\text{tot}}$ becomes small, signifying the biofilm becoming predominantly of type NA. Varying the frequency of glucose pulses in the biofilm environment therefore determines the late-time fate of the biofilm, i.e. whether it is A-dominated (pathogenic) or NA-dominated (homeostatic).

It is evident from the figure that the biofilm composition $N_A/N_{\text{tot}}$ at any given time point continuously increases as the frequency of the glucose pulses increase. There is an intermediate value $t^{\text{durn}} \approx 6h$ where the variation in biofilm composition is not discernible over the available data window. It can therefore be hypothesised that there is a transition value of $t^{\text{durn}}$ about which $N_A/N_{\text{tot}}$ remains at 50:50 for all times, and neither population comes to dominate the other. This is supported by quantitative analysis of the late-time variation of $N_A/N_{\text{tot}}$, which can be shown to exponentially increase towards unity for low $t^{\text{durn}}$, and exponentially decrease towards zero for high $t^{\text{durn}}$. This is demonstrated by semi-logarithmic plot in Fig. 4, which shows examples of the exponential decay of A ($t^{\text{durn}} > 6h$) or NA ($t^{\text{durn}} < 6h$). The fits in these figures are to a simple exponential,

$$\frac{N_A}{N_{\text{tot}}}(t) = A \exp^{-kt},$$

where $k$ is the rate at which the minority fraction decays. The inset to the figure shows that this rate decreases continuously to zero as $t^{\text{durn}}$ approaches a critical point close to 6h. The exponential fit (1) is only successful if an initial transient of roughly 10 days is removed prior to fitting; no simple fit, including logistic growth, was found to fit the entire data range.

**Multi-factorial modulation of pathogenicity**

The duration of glucose pulses $t^{\text{durn}}$ is not the only factor that determines the relative competitiveness of the two populations. Any of the model parameters listed in Tables 1 and 2 can, in principle, affect the growth rates of one or both populations over the course of a complete pulse, and therefore influence the selection of the dominant population. This is expected for the metabolic parameters such as the half-concentrations, but equally holds true for environmental factors. For example, increasing the biofilm thickness $h_{\text{plaque}}$ with all other parameters held fixed, results in an increased production of acid and a lower pH, reducing the metabolic rate of population A to a lesser degree than NA. This shift in competitiveness might be enough to make A the dominant population at late times, even when $t^{\text{durn}}$ is greater than 6h.

Confirmation that parameters other than $t^{\text{durn}}$ can promote a transition between pathogenic and homeostatic biofilms is presented in Fig. 5. This shows the variation of both the intra-pulse pH and the biofilm composition as a function of time, for a series of independent runs that differ only in one of the aciduricity parameters, $K^A_{\text{acid}}$ and $K^{NA}_{\text{acid}}$. There is a clear trend from a late-time biofilm composition that is A-dominated with a low pH, arising for high $K^A_{\text{acid}}$ or low $K^{NA}_{\text{acid}}$, to a NA-dominated state with a higher pH for low $K^A_{\text{acid}}$ or high $K^{NA}_{\text{acid}}$. Thus, increasing the aciduricity of one population enhances its competitiveness with respect to the other. Note that in Fig. 5(b) the pH at the final time point first increases with the parameter $K^{NA}_{\text{acid}}$ as it is increased, but then decreases with $K^{NA}_{\text{acid}}$ for the higher values considered. The initial increase is because of the increasing dominance of the NA population as just discussed. The subsequent decrease arises because, although the NA remain dominant, they become increasingly aciduric as $K^{NA}_{\text{acid}}$ is increased, resulting in increased acid production. Indeed, increasing $K^{NA}_{\text{acid}}$ until it equals $K^A_{\text{acid}}$ would result it two populations of equal aciduricity, rendering our labels A and NA meaningless.

Systematically varying each of the parameters in Table 1 reveals that most of them can also promote this transition. This is summarised in Fig. 6, which shows the biofilm composition and pH for 10 of the parameters between the extremes of the ranges given in the table. In all cases the observed variation
is intuitive. For instance, increasing the diffusion $D$ of acid results in more rapid dispersal from the system according to the boundary conditions of Methods, resulting in a higher pH and a lower fraction of $A$. Conversely, reducing the death rate $k_{\text{death}}$ increases acid production, but this effect is weak as demonstrated below. The parameters $d^{\text{max}}$, $\kappa$ and $\sigma^{\text{div}}$ have been omitted from this figure since the signal-to-noise ratio was too small to discern any trend.

**Parameter sensitivity**

It is not readily apparent from Fig. 6 which parameters are the most important for determining the late-time biofilm composition, making it unsuitable for identifying potential targets for controlling biofilm fate. In addition, the time point for the predicted quantities (pH and composition at $t = 200d$) are not suitable for *in vitro* validation, for which a much shorter time scale is convenient. Both of these issues can be resolved by tabulating the changes in a range of putative experimental measurements with respect to each input parameter, allowing pairings that exhibit high or low sensitivity to be immediately identified.

The sensitivity heat map for this model is Fig. 7, where each entry shows the predicted percentage change in a range of measurable outcomes for a 1% change in each input parameter. The measurable outcomes include 4 predictions for single-species biofilms, *i.e.* the concentration of $H^+$ ions during and between pulses for systems comprised purely of $A$ or $NA$, which reach dynamical steady states in the order of days. In addition, 4 outcomes for mixed-species films were considered. This includes the variation in the critical pulse duration $t^{\text{durn}}$ separating pathogenic from homeostatic biofilms, and 3 quantities measured at a time point of $t = 20d$ starting from an initial 50:50 composition of $A:NA$, namely the fraction of $N_A/N_{\text{tot}}$, and the concentration of $H^+$ (again during and between pulses).

Inspection of the table immediately reveals that the quantity $N_A/N_{\text{tot}}|_{t=20d}$ is the most sensitive to a number of parameters, suggesting this would be a useful quantity to measure in experiments. Single-species experiments may be the best to measure each species' metabolic half-constants, although environmental factors are clearly also important. Reading across rows rather than down columns suggests that the final 5 parameters in the diagram give weak or no change to the measurable outcomes, suggesting these mechanisms are not worthy of further investigation. The variations are not necessarily zero, as seen by inspection of the raw data given in Table S1 of the Supplementary Information, but their influence is evidently weak.

**Discussion**

Mathematical modelling represents a powerful tool for biofilm investigation, providing the capability for quantitative prediction upstream from *in vitro* models and *in vivo* trials. All conceivable measures of biomass composition, structure, and associated chemical gradients can be extracted non-invasively from an *in silico* biofilm. In addition, such data are acquired in an accelerated time frame, which for the results presented here translates into 200 days real time in approximately 10 hours of simulation time, and this ratio could be further improved with additional numerical optimisation and parallelisation. This rapidity allows ranges of each input parameter to be systematically assayed and the corresponding effect on the growing biofilm to be determined, both qualitatively in terms of the nature of the changes, and quantitatively in terms of if these changes are significant or slight.

The freedom with which parameters can be varied makes this form of modelling well suited to studying oral biofilms. The microbial composition of these complex multi-species communities varies greatly between human hosts, with phylogenetically distant organisms fulfilling overlapping functional roles. Focussing on microbial function rather than genetic identity is therefore desirable when developing clinical treatments, and mathematical modelling facilitates this by permitting the rapid assaying of parameters that vary between functionally-similar species. This should be contrasted with the equivalent *in vitro*
experiments, which would need many repeats with different species, strains and sub-strains to sample equivalent ranges of metabolic activity.

When applied to competing acidogenic populations varying in their rates of nutrient uptake \( K_{\text{nut}}^{A/N_A} \) and aciduricity \( K_{\text{acid}}^{A/N_A} \) as in this work, this approach demonstrates that these parameters measurably affect the cariogenic potential of (supragingival) oral biofilms; see Fig. 7. This insight can be used to suggest treatments for modulating biofilm composition towards a benign homeostatic state, and indeed reducing aciduricity is one mechanism by which fluoride promotes improved oral health [12, 13]. This highlights the potentially controlling role of sublethal treatments in modulating population dynamics. It also highlights an additional advantage of mathematical modelling, in that it allows us to isolate this one mechanism from many other postulated roles for fluoride, simply because these alternatives were not included in the model.

Confirmation of the predictive capability of this or any model requires model validation. The parameter sensitivity map in Fig. 7 can serve as a framework for model validation, in two respects. Firstly, it helps identify experiments that can be used to estimate specific model parameters. If the sensitivity of any single experimental measure with respect to a given parameter is high, corresponding to bright entries in the table, the resulting estimate will have a low signal-to-noise ratio and thus should be reliable. Secondly, parameters for which the sensitivity is low need not be determined to any degree of precision for reliable model predictions. It is clear from Fig. 7 that the parameters for cell division, stiffness and death rates have little effect on any of the postulated experimental measures, suggesting that rough approximation of their values should suffice.

Parameter sensitivity can also be used to suggest directions for further model development. It is apparent from Fig. 7 that the parameters for inter-pulse glucose concentration \([\text{Gl}]_{\text{inter}}\) and acid buffering \( K_a \) have a measurable influence on the resulting predictions. Both of these mechanisms were incorporated in an approximate manner in this work, and this sensitivity suggests that further modelling would benefit from expanding each mechanism to include more detail, albeit with the overhead of an increased number of parameters. Indeed, series of reactions for both glucose storage and acid buffering have already been specified for the non-growing biofilms of Ilie et al. [27], and these could be incorporated in an agent-based model such as ours.

The real plaque ecosystem maintains a dynamic equilibrium between multiple species whose relative fractions vary with environmental conditions, but never vanish entirely [1]. By contrast, here one species type is progressively removed from the system, as long as the frequency of glucose intake and low-pH challenges remains constant. This most likely represents the simplicity of the model, and many model extensions are likely to permit subpopulations of functionally dissimilar species to be perpetually maintained. Additional community complexity in the form of more than two distinct species types interacting via a range of interactions, mutualistic, antagonistic and otherwise, should permit a dynamically stable biofilm composition. Heterogenous biofilm composition leads to chemical gradients and an extended habitat range for species that would not be able to persist in a homogenous biofilm, features that can be included in agent based modelling. Other possibilities for maintaining a dynamic equilibrium have been discussed elsewhere [6–8]. Finally, we notice that clonal variation was not included in this version of the model, so all members of a population are phenotypically identical. It is, however, straightforward to introduce such a feature into agent-based modelling, at the expense of additional parameter fitting.

**Analysis**

The mathematical model employed here is based on the Individual based Model (IbM), an established agent-based model for biofilms that has been applied to a range of bacterial communities and environments [28–36]. Such models consist of two coupled phases, a particulate phase where each particle represents a bacterium or bacterial aggregate, and a series of overlapping continuous phases representing the concentration fields of one or more dissolved species, e.g. nutrients or metabolic products. The vari-
ant here admits biomechanics in that the particles are interconnected via springs representing adhesion by the extracellular polymeric substances (EPS). Here we summarise only those features of the model relevant to the subsequent discussion, and direct the reader to [40] for further details.

Model overview

In this study we consider a mixed film consisting of two microbial populations that differ in their ability to metabolise sugars in the presence of low pH. These are referred to as A for aciduric and NA for non-aciduric, and parameters relating to each species are labelled A or NA accordingly. Additionally there is a single scalar field representing the concentration of lactic acid produced by each particle’s glycolysis of dietary sugar. The system domain is schematically shown in Fig. 8(a). A two-dimensional geometry has been employed as this permits biofilms of lateral extent far exceeding their thickness to be simulated within a reasonable timeframe. Periodic boundaries have been assumed in the direction parallel to the enamel surface.

The nutrient (glucose) is not represented as a spatially-varying field, but is instead assumed to have a uniform concentration with no gradients. The concentration does however vary in time as shown in Fig. 8(b). This follows a feast-famine protocol representing the dietary intake of fermentable carbohydrates, whereby the concentration \( [Gl] \) alternates between short periods at a high value \( [Gl]_{\text{intra}} \), interspersed with extended periods at a lower value \( [Gl]_{\text{inter}} \). \( [Gl]_{\text{intra}} \) is taken to be far above the half-concentrations for nutrient uptake (see below), so the metabolism for both populations is saturated during the pulse. The duration of the pulse has been fixed at 15 minutes as this represents a typical removal time of acid from the oral cavity [1, 41]. The duration of a complete pulse cycle (i.e. inter plus intra periods), \( t_{\text{durn}} \), is a key parameter of the model.

The model parameters are listed in two separate tables, where the free parameters that were systematically varied in this study are listed in Table 1, and the fixed parameters whose values were estimated from the literature and not varied are given in Table 2. Additional numerical parameters, such as the convergence parameters for the chemical and mechanical relaxation, were tested to be sufficiently small to not affect the results and are not quoted here.

Cell metabolism

Acid buffering within plaque results in a lower concentration of \( H^+ \) ions, and therefore a higher pH, than for the same concentration of lactic acid in aqueous solution [41]. However, empirical titration curves cannot be easily incorporated into models as discussed elsewhere [23]. Rather than include a series of coupled reactions as in [27], which would slow down our simulations and introduce additional parameters, we instead treat the dissociation of lactic acid to \( H^+ \) as a single-step process with an effective dissociation constant \( K_a \) that is far lower in value than aqueous dissociation. \( K_a \) thus becomes a free parameter which we systematically vary. The concentration of glucose between pulses, which depends on the balance of storage and conversion reactions, is also simplified to a single free parameter \( [Gl]_{\text{inter}} \).

The glycolysis of glucose to lactic acid is assumed to depend on two factors, the concentration of glucose \( [Gl] \) and the local acidity \( [H^+] \). Both quantities modulate the overall reaction rate as independent Monod factors with half-concentrations \( K_{\text{nut}} \) and \( K_{\text{acid}} \) for nutrient uptake and acid inhibition respectively, as shown in Fig. 8(c). The full expression for the reaction rate \( r_i \) (mass per unit time) for the particle with label \( i \) and mass \( m_i \) is

\[
r_i = m_i q_{\text{max}} \left( \frac{[Gl]}{[Gl] + K_{\text{nut}}} \right) \left( \frac{K_{\text{acid}}}{[H^+] + K_{\text{acid}}} \right),
\]

where \( [Gl] \) is in units of mass per unit volume, and \( [H^+] \) in molarity. Each half-concentration has the same units as its corresponding concentration, and have separate values for A and NA, e.g. \( K_{\text{nut}}^A \) and \( K_{\text{nut}}^N \).
The spatial distribution of lactic acid, which is both produced by (2) and modulates it by determining $[H^+]$, obeys the standard reaction-diffusion equation in which local production of acid is balanced by diffusion away from the source. The primary value for the diffusion coefficient $D_{\text{acid}}$ is taken to be that for lactic acid in water, but was also systematically varied. The reaction-diffusion equation was numerically solved using geometric multi-grid on a rectangular mesh [40], with no-flux boundary conditions at the enamel surface and the requirement that the acid at the upper surface of the saliva layer is zero.

Once the $r_i$ for each particle $i$ is determined, the resulting rate of increase in mass is computed as $Y^c r_i$, where $Y^c$ is the dimensionless yield factor; here we fix $Y^c = 0.1$ for both bacterial populations, comparable to representative oral bacteria [38,39]. EPS is assumed to be produced at a rate proportional to the cell mass, i.e. the mass of the EPS increases at a rate $Y_{\text{rel}}^\text{EPS} Y^c r_i$, where $Y_{\text{rel}}^\text{EPS}$ is a free parameter that we systematically vary. Particle and EPS masses are converted to physical size by assuming fixed densities of both components as listed in Table 2.

When a particle diameter exceeds the critical value $d_{\text{max}}$, it divides into two daughter particles. The mass of the mother particle, $m_m$, is conserved but is distributed asymmetrically to the daughters, i.e. with masses $m_1^d$ and $m_2^d$ obeying
\begin{equation}
    m_1^d, m_2^d = f m_m, (1-f)m_m,
\end{equation}
where $f$ is a random number drawn from a Normal distribution with mean 0.5 and width $\sigma_{\text{div}}$. The mass of the EPS is distributed similarly, with the same $f$. The two parameters $d_{\text{max}}$ and $\sigma_{\text{div}}$ are treated as free and systematically varied.

**Redistribution and removal of biomass**

After particle growth and division, the particles are rearranged so as to ensure the biofilm as a whole is mechanically stable. This procedure, described in detail in [40], involves constructing a network of springs connecting nearby particles. Each of these springs has a stiffness that is proportional to the local EPS mass and the model parameter $\kappa_{\text{char}}$. In addition, springs between particles and the enamel surface have a stiffness that is also proportional to $\kappa_{\text{char}}$. The free parameter $\kappa_{\text{char}}$ is here varied around values suggested by AFM experiments [42] as shown in Table 1. The numerical procedure involves converting the requirement of mechanical equilibrium to a sparse matrix equation and solving using the conjugate gradient method. This matrix approach has also been adopted for plant biofilms, motivated by numerical performance [43].

The plaque biofilm is limited in thickness to $h_{\text{plaque}}$, in that any particles whose centre exceeds this height is removed from the system. This can be thought of as a simplistic representation of biomass removal due to fluid shear by flowing saliva [32], and can be achieved in vitro by a constant depth film fermenter [13,44–46]. Above the plaque biofilm is the saliva layer of thickness $h_{\text{saliva}}$ in which there is no biomass, but the concentration of lactic acid still continuously varies until vanishing at the upper saliva surface.

Additionally, since bacteria deep within a biofilm exhibit a lower viability than those near the surface, we include a second mechanism for particle removal, namely that cells are ‘killed’ at a rate that depends on their distance from the exposed biofilm surface, with the highest death rate near the enamel. The death rate per unit depth per unit time is denoted $k_{\text{death}}$ and is a free parameter that is varied here. This parameter is the same for both populations.

**Parameter sampling**

Given the large number of parameters, a systematic investigation of all combinations was not feasible. Instead, each free parameter was assigned a primary value based on chemostat experiments of oral bacteria [38,39], known environmental conditions for oral biofilms [1,7,8] and related models [27]. The primary value for the effective dissociation constant $K_a$ was selected to give a realistic pH during the
glucose pulse, and that for $[Gl]_{\text{inter}}$ to give a critical value of $t^{\text{dur}}$ (see below) around 6h. Each parameter was then varied over the range specified in Table 1 with all other parameters held fixed.

The sensitivity heat map discussed in Results refers to linear response, i.e. small changes in the input parameters for which the corresponding change in each output parameter was proportional. For each model parameter, a range of variations was tested to ensure the corresponding variation in the outcome was linear (unlike the data in Fig. 6, which is concerned with non-linear trends over a broad variation of parameters), which was typically the case for a 5-10% change in the model parameter. The sensitivity was then scaled to a 1% value. The raw data (with errors) are given in Table S1 of the Supplementary Information.

Acknowledgments

References


Tables

<table>
<thead>
<tr>
<th>Label</th>
<th>Meaning</th>
<th>Primary value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{\text{durn}}$</td>
<td>Duration of pulse cycle</td>
<td>6h</td>
<td>2 – 10h</td>
</tr>
<tr>
<td>$K_{A_{\text{acid}}}$</td>
<td>Half concentration for acid inhibition for A</td>
<td>$2 \times 10^{-5}$ mol/L</td>
<td>$5 \times 10^{-6} - 10^{-4}$ mol/L</td>
</tr>
<tr>
<td>$K_{A_{\text{acid}}}$</td>
<td>Half concentration for acid inhibition for NA</td>
<td>$2 \times 10^{-7}$ mol/L</td>
<td>$5 \times 10^{-8} - 10^{-6}$ mol/L</td>
</tr>
<tr>
<td>$K_{A_{\text{nut}}}$</td>
<td>Half concentration for nutrient uptake for A</td>
<td>5 g/L</td>
<td>0.1 – 20 g/L</td>
</tr>
<tr>
<td>$K_{A_{\text{nut}}}$</td>
<td>Half concentration for nutrient uptake for NA</td>
<td>20 mg/L</td>
<td>5 – 40 mg/L</td>
</tr>
<tr>
<td>$K_a$</td>
<td>Effective dissociation constant</td>
<td>$10^{-9}$ mol/L</td>
<td>$10^{-10} - 10^{-8}$ mol/L</td>
</tr>
<tr>
<td>$h_{\text{plaque}}$</td>
<td>Thickness of the plaque layer</td>
<td>150 µm</td>
<td>50 - 250 µm</td>
</tr>
<tr>
<td>$h_{\text{saliva}}$</td>
<td>Thickness of the saliva layer</td>
<td>100 µm</td>
<td>25 - 350 µm</td>
</tr>
<tr>
<td>$[\text{Gl}]_{\text{inter}}$</td>
<td>Concentration of sugar between pulses</td>
<td>50 g/L</td>
<td>1 - 15 mg/L</td>
</tr>
<tr>
<td>$D_{\text{acid}}$</td>
<td>Diffusion coefficient for the acid (uniform)</td>
<td>$10^{3}$ µm²/s</td>
<td>140 - 2700 µm²/s</td>
</tr>
<tr>
<td>$Y_{\text{EPS}}$</td>
<td>Relative yield factor for EPS</td>
<td>0.4</td>
<td>0.2 - 1.5</td>
</tr>
<tr>
<td>$k_{\text{death}}$</td>
<td>Linear factor in the kill rate</td>
<td>$10^{-4}$/µm h</td>
<td>$2 \times 10^{-5} - 5 \times 10^{-4}$ / µm h</td>
</tr>
<tr>
<td>$d_{\text{max}}$</td>
<td>Threshold diameter for cell division</td>
<td>5 µm</td>
<td>4µm - 10µm</td>
</tr>
<tr>
<td>$\kappa_{\text{char}}$</td>
<td>Characteristic stiffness</td>
<td>50 pN/µm</td>
<td>30 – 200 pN/µm</td>
</tr>
<tr>
<td>$\sigma_{\text{div}}$</td>
<td>Width of daughter mass ratios after division</td>
<td>0.1</td>
<td>0.05 - 0.2</td>
</tr>
</tbody>
</table>

Table 1. Physical and biological parameters that were systematically varied in this study. When one parameter was varied over the range given in the fourth column, the remaining parameters took the values given in the third column. See Methods for further details.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{\text{intra}}$</td>
<td>Duration of carbohydrate pulse</td>
<td>15 m</td>
<td>[41]</td>
</tr>
<tr>
<td>$L$</td>
<td>Longitudinal film width</td>
<td>2 mm</td>
<td>-</td>
</tr>
<tr>
<td>$[\text{Gl}]_{\text{intra}}$</td>
<td>Concentration of sugar during a pulse</td>
<td>50 g/L</td>
<td>[38, 39, 47–50]</td>
</tr>
<tr>
<td>$\rho^c$</td>
<td>Cell density (excluding water)</td>
<td>0.2 pg/µm³</td>
<td>[31]</td>
</tr>
<tr>
<td>$\rho^e$</td>
<td>EPS density (excluding water)</td>
<td>$4 \times 10^{-2}$ pg/µm³</td>
<td>[30]</td>
</tr>
<tr>
<td>$q_{\text{max}}$</td>
<td>Base reaction rate</td>
<td>5/h</td>
<td>[38, 39]</td>
</tr>
<tr>
<td>$Y_c$</td>
<td>Yield factor for cell mass</td>
<td>0.1</td>
<td>[38, 39]</td>
</tr>
<tr>
<td>$M_{\text{acid}}$</td>
<td>Molecular weight of (lactic) acid</td>
<td>90.08 g/mol</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Physical and biological parameters that were not varied in this study and kept at the values shown.
Supporting Information Legends

**Figure S1. Snapshots of examples biofilms.** (a) $t_{\text{dur}} = 8h$ at a time point of 10 days, (b) the same after 200 days, and (c) $t_{\text{dur}} = 4h$ after 200 days. Green (blue) discs correspond to populations of A (NA), respectively. The red in the background corresponds to the lactic acid, with red (black) corresponding to high (low) concentrations respectively. All other parameters are the same as Fig. 1 in the main article.

**Table S1. Raw data for the sensitivity heap map.** Values correspond to the percentage change in the measurable outcome (columns) for a 1% change in the model parameter (rows). Error bars are either given explicitly, or as variation in the final 2 digits (shown in brackets).

**Movie S1. Movie corresponding to Fig. S1(b).** Extends from $t = 0$ to $t = 200$ days. Colour coding and parameters as per Fig. S1(b).

**Movie S2. Movie corresponding to Fig. S1(c).** Extends from $t = 0$ to $t = 200$ days. Colour coding and parameters as per Fig. S1(c).
Figures

Figure 1. Snapshots of a section of a biofilm taken at time points (a) $t = 0d$, (b) $t = 10d$ and (c) $t = 200d$ for the primary parameters in Table 1 (the full system size is roughly 10 times wider). Light grey (black) discs denote aggregates of A (NA), respectively, encased in EPS shells (grey). The shaded field in the background corresponds to the acid, with light (dark) regions for high (low) concentrations. The white regions near the base correspond to high concentrations of acid, visible in the voids created by cell death (see text). Larger, colour images and movies for $t^{\text{turn}} = 4h$ and $t^{\text{turn}} = 8h$ are available in Figure S1 and Movies S1 and S2 of the Supplementary Information.
Figure 2. Example of evolution of the pH averaged over the tooth surface (left axis; solid lines) and fraction of biofilm that belongs to population A (right axis; broken lines) for 3 consecutive glucose pulses. This example is for the primary parameters of Table 1 except the total pulse cycle duration $t_{\text{durn}} = 4h$ here. The pH is observed to decrease for successive pulses, concomitant with an increase in the fraction of A. During the inter-pulse period, which has been compressed for clarity, the fraction of A decreases slowly (shown schematically by the diagonal arrows), and the pH takes much higher values around 6.0 (not shown).

Figure 3. Fraction of system that belong to population A versus time for (from top to bottom) $t_{\text{durn}} = 2h, 3h, 4h, 5h, 5.5h, 6h, 6.5h, 7h, 8h, 10h$, respectively. Error bars show scatter over at least 10 independent runs.
Figure 4. Plot of the fraction of the minority population (NA for open symbols, A for closed symbols) versus time, for the pulse cycle durations $t_{\text{durn}}$ shown in the legend. Solid straight lines are the fit to exponential decay for times $t > 10d$. (Inset) Fitted rates versus $t_{\text{durn}}$. 
Figure 5. Variation of biofilm composition and pH during the glucose pulse as a single parameter is varied, for (a) $K_{\text{acid}}^{A}$ and (b) $K_{\text{acid}}^{NA}$. These parameters are plotted as $-\log_{10}(K_{\text{acid}}^{A/NA})$, corresponding to the pH at which cell metabolism is 50% inhibited by acidity. Each line segment shows the variation in pH, starting from 2d and finishing on 200d, with the arrow showing the direction of increasing time. The numbers show the percentage of biofilm occupied by A (i.e. $N_{A}/N_{\text{tot}}$) at the start and end points.
Figure 6. Variation of pH during the pulse at $t = 200d$ as the first 12 parameters of Table 1 are varied. Each bar corresponds to a single parameter being varied with the remaining held fixed at their primary values. The shading corresponds to the fraction of the biofilm occupied by population A as in the calibration bar.

Figure 7. (Colour online) Sensitivity heat map between input parameters (rows) and measurable outputs (columns). Brightness corresponds to the relative change (increase or decrease) in the output quantity as the parameter is increased by 1%, as indicated by the colour bar. Actual values (with error bars) provided in Table S1 of the Supplementary Information.
Figure 8. Schematic of the model. (a) The system geometry. (b) The temporal variation of the environmental glucose follows a feast-famine protocol with a total cycle time $t_{d\text{urn}}$, including the 15 minute pulse. (c) Relative metabolic rates due to nutrient (left) and acid inhibition (right) as per (2), so 100% means saturation and 1% means a significant reduction. For both plots, the leftmost solid curve corresponds to the primary value for NA, and the right to A. Each curve also has a shaded region giving the total range considered for each parameter. For the left plot, the parameters for $[Gl]_{\text{intra}}$ and $[Gl]_{\text{intra}}$ are shown as vertical dashed lines, with the shaded region for the former showing the range considered. For the right plot, the vertical dashed curves show typical pH values during (right) and in between (left) glucose pulses for the primary values in Table 1.