



A Formal Model for the Simulation and Analysis of Early Biofilm Formation

Antonio Cerone¹ and Enrico Marsili²

¹ Department of Computer Science, School of Engineering and Digital Sciences, Nazarbayev University, Nur-Sultan, Kazakhstan

`antonio.cerone@nu.edu.kz`

² Department of Chemical and Materials Engineering, School of Engineering and Digital Sciences, Nazarbayev University, Nur-Sultan, Kazakhstan

`enrico.marsili@nu.edu.kz`

Abstract. Biofilms are structured communities of bacterial cells adherent to a surface. This bacterial state is called *sessile*.

This paper focuses on the modelling of the transition between planktonic and sessile state using Real-time Maude as the modelling language. With more and more bacteria joining the sessile community, the likelihood of producing a biofilm increases. Once the percentage of bacterial cells that adheres to the surface reaches a threshold, which is specific for the considered bacterium species, a permanent biofilm is formed. An important challenge is to predict the time needed for the formation of a biofilm on a specific surface, in order to plan when the material infrastructure that comprises such a surface needs to be cleaned or replaced. We exploit the model-checking features of Real-time Maude to formally prove that a regular cleaning or replacement of the infrastructure prevents the biofilm formation.

Keywords: Biofilms · Formal methods · Rewriting logic · Real-Time Maude

1 Introduction

Biofilms are microstructured bacterial communities that live at interfaces. They are the most common mode of life for microorganisms in both terrestrial and marine environments. Usually, biofilms thrive at liquid/solid interfaces like the inner surface of water pipes [29] and catheters [28], or in soil and sediments [10]; biofilms play a beneficial role in wastewater treatment, where they increase organic carbon degradation and contaminant removal. However, biofilms harbor pathogens and protect them from antimicrobial agents, thus posing serious threats in health settings. Biofilms consist of bacterial cells encased in self-produced extracellular polymeric substances, collectively termed biofilm matrix.

This work was partially funded by the Faculty Development Competitive Research Grant Program (Grant number 110119FD4537), Nazarbayev University, Kazakhstan.

© The Author(s) 2021

J. Bowles et al. (Eds.): DataMod 2020, LNCS 12611, pp. 134–151, 2021.

https://doi.org/10.1007/978-3-030-70650-0_9

The biofilm matrix is composed of carbohydrates, proteins and extracellular DNA. The relative proportion of these components is species-dependent and varies with biofilm age and nutrient concentration [17].

The biofilm life cycle is illustrated in Fig. 1. Biofilm formation initiates when single planktonic cells enter in contact with a surface, usually solid. Following this initial interaction, several steps can be described, including reversible attachment, irreversible attachment, microcolonies formation, and biofilm maturation, in which complex microstructures are formed and the biofilm reach its maximum thickness. In the planktonic-sessile transition, cells lose their flagella and start producing extracellular polymeric substance. The biofilm life cycle ends with biofilm detachment or dispersal, where part of the cells transitions from biofilm to planktonic state and move downstream to seed other sections of the surface. The biofilm life cycle has been validated with microscopy, transcriptomics and metabolomics analysis [30].

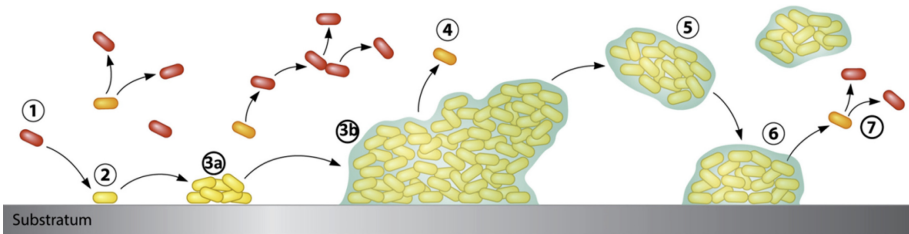


Fig. 1. Biofilm life cycle: (1) Planktonic phenotype; (2) Newly attached planktonic cells (“settler” biofilm); (3a) Fully mature biofilm stage (biofilm phenotype); (3b) Newly single cells dispersed from the biofilm (newly dispersed phenotype); (4) Detached biofilm aggregates (biofilm phenotype); (5) Reattached biofilm aggregates (biofilm phenotype); (6) Newly dispersed phenotypes cells from the biofilm giving rise to planktonic phenotype cells.

The biofilm early formation is largely dependent on the concentration of bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP), a secondary messenger that is ubiquitous in the bacterial world. Enzymes that synthesize c-di-GMP or degrade it are found in great numbers, indicating its importance as a regulator of many bacterial functions. C-di-GMP contributes to determine motility, biofilm formation, and production of multiple proteins in microorganisms. It binds to a large variety of effector components and controls targets involved in transcription and formation of large cellular and extracellular structures.

In the biofilm life cycle, the most important step is the initial attachment, in which the cells transition from free-swimming state (so-called *planktonic*) to *sessile* state, in which they attach to solid surface, lose their motility and start producing the extracellular matrix, in addition to the normal growth process [9]. Modelling of initial biofilm formation is important to predict the extent of biofilm growth on removable biomedical devices (e.g., catheters) and drinking water

pipes, thus allowing for planned cleaning or replacement of biofilm-contaminated parts and minimising the risk of infections [11]. The effectiveness of cleaning treatment depends also on biofilm concentration, thus the formal modelling of initial biofilm formation will contribute to optimising the frequency and the duration of antimicrobial application in biomedical devices and water systems.

A number of biofilm models have been developed during the last 30 years. They can be roughly divided into continuum and discrete models.

Continuum models simulate biofilms in a quantitative and deterministic way [16]. However, such models may result in high computational complexity, particularly for multispecies biofilms and when considering multiple substrates [3]. Early continuum models focused on cell growth and microstructure formation. Modern continuum models concern the effect of fluid dynamics on colony formation and are validated with time-resolved single cell imaging to reveal important details on cell-fluid interactions at different biofilm ages [23].

Discrete models, on the other hand, are very successful in representing the multidimensional heterogeneity of biofilm, but introduce elements of randomness and stochastic effects into the solutions [16]. *Agents-based models*, based on several platforms, eventually integrated (e.g., NetLogo, MatNet), are especially popular because of their simplicity and low computational requirements. However, they may become highly sophisticated when integrating multiscale and constraint-based metabolic modelling [7].

Hybrid biofilms models are the most recent. They support the simulation of discrete bacterial cells within a multiphase continuum consisting of extracellular polymeric substance (EPS) and water as separate interacting phases. These models support the prediction of bacterial colony formation. The distribution of bacterial growth and EPS production is sensitive to the pore spacing between bacteria and the consumption of nutrients within the bacterial colony [12].

Although formal methods have been widely used in systems biology [6], to the best of our knowledge, no formal models of biofilms have been reported in the literature. After Păun's work on P-Systems [24], rule-based systems have been widely used in the modelling of biological systems, due to the natural way in which they can express chemical reactions and biological interactions. The use of a rule-based systems in combination with a formal analysis tool allows us to generate a model that focuses on the individual bacterial cells, whose behaviour can be defined using simple rules, assuming that the cell properties are known, rather than using a computationally expensive multidimensional continuum. We use Real-Time Maude [19,21], a toolset that comprises a rule-based formal modelling language, which uses rewriting logic [18] to model system state transitions, and high-performance simulation and model checking engines, which support formal analysis.

The rest of this paper is organised as follows. Section 1.1 provides a brief highlight of Real-Time Maude and refers to the sections of the paper where the different aspects of the language are illustrated. The Real-Time Maude code for the model, the experiments and the formal analysis can be downloaded from a

GitHub repository¹. Section 2 presents our approach for modelling a bacterial population and its evolution, the process of biofilm formation as well as preventive interventions. Section 3 illustrates our approach using an example inspired by the information about a specific microorganism, *Pseudomonas aeruginosa*, which is available in the literature [1, 25, 31]. It exploits the example to illustrate both in silico experiments and the use of the model-checking features of Real-time Maude to predict the outcome of the experiments and analyse intervention plans. Section 4 concludes the paper with further considerations on our approach and ideas for future work.

1.1 Real-Time Maude

Real-Time Maude [19, 21] is a formal modeling language and high-performance simulation and model checking tool for distributed real-time systems. It is based on Full Maude, the object-oriented extension of Core Maude, which is the basic version of Maude [20]. Real-Time Maude makes use of

- algebraic equational specifications in a functional programming style to define data types;
- labeled rewrite rules to define local and global transitions;
- tick rewrite rules to advance time in the entire system state.

In this paper we do not go into the details of Real-Time Maude syntax but we focus on rewrite rules and commands for simulation and formal analysis. In this section we provide high-level information on the Core Maude data type definition and Full Maude classes, subclasses and objects. Labelled rewrite rules for defining global transitions and Real-Time Maude tick rewrite rules are illustrated in Sect. 2. Commands for simulation and formal analysis are illustrated in Sect. 3.

A Core Maude *data type* is defined using an algebraic *signature*, that is, a set of declarations of *sorts*, *subsorts*, and *operators*. Operators are defined in an equational way as well as by using special tags to declare common properties, such as commutativity, associativity, having a specific identity and constructors. *Constructors* are the carriers of the sort, in the sense that they define in a unique way each member of the sort.

In Full Maude, a *class* declaration

$$\text{class } C \mid att_1 : s_1, \dots, att_n : s_n.$$

declares a class C with attributes att_1, \dots, att_n of sorts s_1, \dots, s_n , respectively. An *object* of class C in a given state is represented as a term

$$\langle O : C \mid att_1 : val_1, \dots, att_n : val_n \rangle$$

of sort `Object`, where O , of sort `Objid`, is the object's *identifier*, and where val_1, \dots, val_n are the current values of the attributes att_1, \dots, att_n , respectively.

¹ github.com/AntonioCerone/Publications/tree/master/2020/DataMod/Biofilms.

A *subclass* inherits all the attributes and rules of its superclasses. The class declaration syntax is also used for subclasses and supports the definition of new attributes specific to the subclass.

Real-Time Maude specifications are organised in *modules*, which can be imported by other modules and, as we will see in Sect. 3, be referred by the commands for simulation and formal analysis.

2 Biofilm Formation Model

We use Real-Time Maude to model a bacterial population as a multiset of objects of a class `Bacterium` with three attributes:

`state` which can be either `planktonic` or `sessile`;
`toDuplication` which is a natural number representing the remaining lifetime in minutes of a cell before duplication;
`c-di-GMP-internal` which is a natural number representing the concentration of c-di-GMP inside a single cell and is expressed in nanomoles (nM)

We also use a single object of a class `Population` to collect the global data of the entire population of bacteria. The class `Population` consists of the following attributes:

`state` which can be one of the following values, whose meaning is illustrated in Fig. 2: `creation`, `next-time`, `ready-to-reproduce`, `ready-to-die` and `attributes-updated`.
`size` which is a natural number representing the number of cells of the population;
`planktonic-cells` which is a natural number representing the number of cells of the population that are in planktonic state;
`sessile-cells` which is a natural number representing the number of cells of the population that are in sessile state;
`c-di-GMP-global` which is a natural number representing the global concentration of c-di-GMP for the entire population and is expressed in nanomoles (nM) with respect to the global volume of the cells;
`biofilm` which is a boolean stating whether the biofilm is formed or not.

We preferred to keep the multiset of `Bacterium` objects ‘outside’ the population data, rather than incorporating it as a field of the `Population` object. This choice aims at considering the ‘global population’ and the ‘multiset of bacteria’ as a sort of pair of ‘interacting agents’ whereby, in a prospective extension of our approach, global data on the population behaviour may actually effect the development of the multiset of agents. In fact, it is easier in biology to have availability of data on the behaviour of the population as a whole rather than on the behaviour of the single cells that make up the population.

Therefore, the system that describes the bacterial ecosystem consists of one object of class `Population` and a multiset of objects of class `Bacterium`. Figure 2 illustrates the initial creation of the bacterial population and the evolution of the

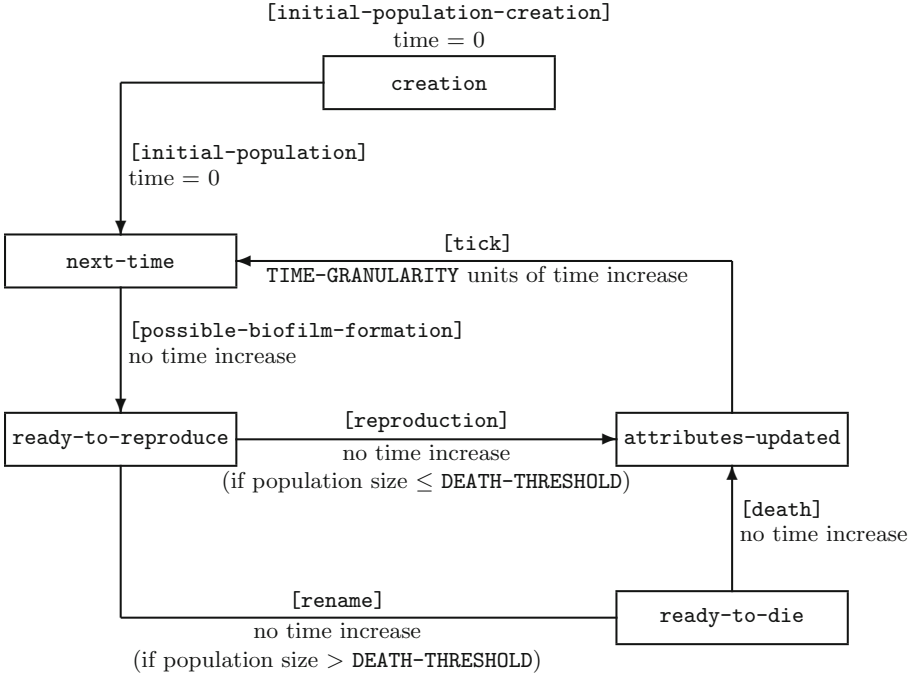


Fig. 2. State transitions in the model of the bacterial population.

population. Evolution consists of `TIME-GRANULARITY` unit cycles. During each cycle the one object of class `Population` goes through the different states that update the attributes of all objects (bacteria and population) and terminate the cycle with a `TIME-GRANULARITY` unit time increment.

With reference to Fig. 2, in Sect. 2.2, 2.3, 2.4, 2.5 and 2.6 we illustrate the rewrite rules that drive the population initial creation and its evolution. Section 2.7 illustrate how our model may be extended to include intervention preventing biofilm formation. Section 2.1 set the context for the next sections by briefly explaining the notion of system state and overviewing the syntax and semantics of rewrite rules.

2.1 Real-Time Maude Configuration and Rewrite Rules

The *global system state* is a term of pre-defined sort `Configuration` and is a *multiset* of objects. Multiset union is denoted by an associative and commutative juxtaposition operator, so that rewriting is *multiset rewriting*. Transitions between global system states are defined using *rewrite rules*.

Maude labeled rewrite rules

$$r1 [l] : \{t\} \Rightarrow \{t'\} \quad \text{or} \quad cr1 [l] : \{t\} \Rightarrow \{t'\} \text{ if } cond$$

define transitions from the global state pattern t to the global state pattern t' . The transition is enabled, and the rule is applied, if the state pattern t matches the current global system state and, for *conditional rules* (`cr1`), the condition *cond* is met. Complex conditions may be built using the conjunction operator \wedge . Conditions may also be of the form $v := t''$ in order to assign a term t'' , containing only variables occurring in t , to variable v , which can then be used to define the new term t' . These special conditions are evaluated to `true`.

Real-Time Maude *tick rewrite rules*

```
r1 [l] : {t} => {t'} in time Δ or cr1 [l] : {t} => {t'} in time Δ if cond
```

have the additional function of advancing time in the global state t by Δ time units.

2.2 Creation of the Initial Population

In order to start with an initial population of planktonic cells uniformly distributed in term of cell age, we consider a constant value `INITIAL-POPULATION` for the initial number of bacteria and use the following two rewrite rules:

`initial-population-creation` which cyclically assigns an age between 1 and the age at duplication (i.e., the duration of the cell life cycle) for the specific species to the newly created cell;

`initial-population` which stops the creation process by changing the population state from `create` to `next-time`, when the bacterial population reaches the value `INITIAL-POPULATION`, thus disabling the previous rule.

The first rule also initialises the concentration of c-di-GMP internal to the cell. This concentration depends on the age of the cell and on food availability.

2.3 Biofilm Formation

The transition of the object of class `Population` from state `next-time` to state `ready-to-reproduce` updates the information about the biofilm formation and the changes of state of the single cells (from `planktonic` to `sessile` and vice versa). We consider two threshold values for the concentration of c-di-GMP: `C-DI-GMP-THRESHOLD-BIOFIM` and `C-DI-GMP-THRESHOLD-SESSILE`.

The constant `C-DI-GMP-THRESHOLD-BIOFIM` represents the global concentration of c-di-GMP that triggers the transition of the bacterial colony to a biofilm, when the threshold is reached or exceeded, and the transition back to a dispersed colony, when the concentration drops below the threshold. The transition is recorded in the boolean attribute `biofilm` of the object of class `Population`.

The constant `C-DI-GMP-THRESHOLD-SESSILE` represents the internal concentration of c-di-GMP of a single cell that triggers the transition of the state of that cell to `sessile`, when the threshold is reached or exceeded, and the transition back to `planktonic`, when the concentration drops below the threshold.

The rewrite rule `possible-biofilm-formation` models the transition to biofilm and back:

```

crl [possible-biofilm-formation] :
  {BACTERIA
    < P : Population | state : next-time,
                        size : N,
                        planktonic-cells : NP,
                        sessile-cells : NS,
                        c-di-GMP-global : C,
                        biofilm : BOOL >}
=>
  {NEW-BACTERIA
    < P : Population |
      state : ready-to-reproduce,
      size : N,
      planktonic-cells : countCells(NEW-BACTERIA, planktonic),
      sessile-cells : countCells(NEW-BACTERIA, sessile),
      c-di-GMP-global : C,
      biofilm : (C >= C-DI-GMP-THRESHOLD-BIOFIM) >}
  if NEW-BACTERIA := changeState(BACTERIA) .

```

Note that this rewrite rule is always enabled in the state `next-time` in order to update the state of each single cell after the concentration of c-di-GMP internal to the cell has been updated during the last time increment. The operator `changeState` recursively changes the state of the entire multiset of bacteria by comparing the concentration of c-di-GMP internal to the cell with the threshold `C-DI-GMP-THRESHOLD-SESSILE`. The operator `countCells` recursively counts the number planktonic and sessile cells of the `NEW-BACTERIA` multiset, which is the bacterial colony after updating the states of the single cells.

2.4 Cell Reproduction

Cell reproduction may only occur when the bacterial population does not exceed a given `REPRODUCTION-THRESHOLD` threshold. The reproduction process is modeled by the operator `mitoses` which, for each cell that has reached the reproduction age, i.e., whose `toDuplication` attribute equals 0,

1. resets the attribute `toDuplication` of that cell to a given `LIFE-DURATION` value, which represents the duration of the life cycle of the cell;
2. creates a new cell identical to the previous one apart from a 0 value for the concentration of c-di-GMP internal to the cell;
3. uses the size of the population, passed as the second argument, to define the object identifiers of the new cell by recursively incrementing such argument.

We have chosen to leave the entire concentration of c-di-GMP within one of the two cells. This is consistent with recent literature, which has shown that asymmetric division results in a better colonization of the surface [8], with the formation of multiple microcolonies [14].

The rewrite rule **reproduction** models the cell reproduction process and its impact on the attributes of the **Population** object:

```

crl [reproduction] :
  {BACTERIA
    < P : Population | state : ready-to-reproduce,
                        size : N,
                        planktonic-cells : NP,
                        sessile-cells : NS,
                        c-di-GMP-global : C,
                        biofilm : BOOL >}

=>
  {NEW-BACTERIA
    < P : Population |
      state : attributes-updated,
      size : NEW-NP + NEW-NS,
      planktonic-cells : NEW-NP,
      sessile-cells : NEW-NS,
      c-di-GMP-global : NEW-C-DI-GMP,
      biofilm : (NEW-C-DI-GMP >= C-DI-GMP-THRESHOLD-BIOFIM) >}

if N <= REPRODUCTION-THRESHOLD /\
  NEW-BACTERIA := mitoses(BACTERIA, N) /\
  NEW-NP := countCells(NEW-BACTERIA, planktonic) /\
  NEW-NS := countCells(NEW-BACTERIA, sessile) /\
  NEW-C-DI-GMP := count-c-di-GMP(NEW-BACTERIA) .

```

All attributes of the object of the class **Population** are updated to take into account the newly created cells. In particular, the global concentration of **c-di-GMP** is updated by using the operator **count-c-di-GMP**, which sums up the concentrations of **c-di-GMP** internal to all the single cells in multiset **NEW-BACTERIA** (after the reproduction has occurred).

2.5 Cell Death

When the bacterial population exceeds the given **REPRODUCTION-THRESHOLD** threshold, cell reproduction can non longer occur. Instead, cells start dying at a rate that has been estimated around 5% for each life cycle [2].

The rewrite rule **death** models such a form of death:

```

crl [death] :
  {BACTERIA
    < P : Population | state : ready-to-die,
                        size : N,
                        planktonic-cells : NP,
                        sessile-cells : NS,
                        c-di-GMP-global : C,
                        biofilm : BOOL >}

=>
  {NEW-BACTERIA
    < P : Population | state : attributes-updated,

```

```

    size : NEW-NP + NEW-NS,
    planktonic-cells : NEW-NP,
    sessile-cells : NEW-NS,
    c-di-GMP-global : NEW-C-DI-GMP,
    biofilm : (C >= C-DI-GMP-THRESHOLD-BIOFIM) >}
if M := ((N * DEATH-RATE * TIME-GRANULARITY) quo 100) quo LIFE-DURATION /\
NEW-BACTERIA := starvation(BACTERIA, M, 0) /\
NEW-NP := countCells(NEW-BACTERIA, planktonic) /\
NEW-NS := countCells(NEW-BACTERIA, sessile) /\
NEW-C-DI-GMP := count-c-di-GMP(NEW-BACTERIA) .

```

As shown in Fig. 2, before applying the rewrite rule **death**, we need to use the rewrite rule **rename** to rename all objects of the class **Bacteria**. This purely technical manipulation, which has no biological meaning, allows us to reuse the object names and prevents the number of system states to grow arbitrarily. Since, the balance between reproduction and death keeps the population below some size threshold, the behaviour of the model has a finite (though very large) number of states and can, in principle, be analysed using model checking with no time limitations. However, in practice, we normally introduce a time upper bound when using model checking to avoid the state explosion problem.

The rewrite rule **death** is enabled by the population state **ready-to-die** which, as shown in Fig. 2, is changed from **ready-to-reproduce** by the rewrite rule **rename**, when the attribute **size** of the **Population** object exceeds the threshold **REPRODUCTION-THRESHOLD**. The rewrite rule **death** uses the operator **starvation** to remove a **DEATH-RATE** per cent of the population over a duration corresponding to the cell life cycle (i.e., **LIFE-DURATION**).

Let **TIME-GRANULARITY** be the number of units time advances at each increment. The number **M** of cells to remove during one time unit is calculated by considering the number of cells $N * DEATH-RATE \text{ quo } 100$ to be removed during an entire life cycle and dividing it by **LIFE-DURATION**. This value has to be multiplied by **TIME-GRANULARITY** to get the number of cells to be removed at each time increment. Note that the terms of the expression have been rearranged in the rewrite rule condition to avoid the inclusion of term $N * DEATH-RATE \text{ quo } 100 \text{ quo } LIFE-DURATION$, which may become 0 for small values of **N** even for a coarse time granularity (values of **TIME-GRANULARITY** greater than 1).

In addition to removing **M** cells from the multiset **BACTERIA** independently of their ages, the reduction of expression **starvation(BACTERIA, M, 0)** recursively increments its third argument by 1 and uses it to rename cells that do not have to be removed from the multiset. In this way, the resultant multiset of bacteria will be named with all natural numbers between 1 and the size of the bacterial population. Furthermore, the operator **starvation** resets the **toDuplication** attribute of each bacteria that is not removed to **LIFE-DURATION**. This ensures that the uniform distribution of the cell ages is preserved.

2.6 Time Increment

Both rewrite rules **reproduction** and **death** change the state of the object of class **Population** to **attributes-updated**. This is the final state for the current

time interval and features the updated values for all attributes of the objects of the system configuration.

In the state `attributes-updated` time is incremented discretely as intervals of `TIME-GRANULARITY` units by the following tick rewrite rule:

```

rl [tick] :
  {BACTERIA
    < P : Population | state : attributes-updated >}
=>
  {idle(BACTERIA, TIME-GRANULARITY)
    < P : Population | state : next-time >}
  in time TIME-GRANULARITY .

```

The operator `idle` not only advances time for all cells but also calculates the internal concentration of c-di-GMP, depending on the age of the cell after the increment, and changes the `Population` state to `next-time` to be ready for updating all the object attributes during the new time interval.

2.7 Intervention to Prevent Biofilm Formation

In this section we show how to model a simple form of intervention to prevent the formation of a biofilm on a specific surface: the replacement of the material infrastructure that comprises such a surface. To this purpose we consider a constant time `TIME-BETWEEN-INTERVENTIONS`, we extend the class `Population` by defining a subclass `Intervention` which has the additional following attribute:

`next-intervention` which is the possibly infinite time before the next intervention;

and we replace the `tick` rewrite rule introduced in Sect. 2.6 with the following `tick-without-intervention` tick rewrite rule:

```

crl [tick-without-intervention] :
  {BACTERIA
    < I : Intervention | next-intervention : T,
                        state : attributes-updated >}
=>
  {idle(BACTERIA, TIME-GRANULARITY)
    < I : Intervention | next-intervention : T minus 1,
                        state : next-time >}
  in time TIME-GRANULARITY
  if T > 0 .

```

When the `next-intervention` attribute of the `Intervention` object has become as low as 0, the following `tick-with-intervention` rewrite rule reset the system to the initial population:

```

r1 [tick-with-intervention] :
  {BACTERIA
    < I : Intervention | next-intervention : 0,
                      state : attributes-updated >}
=>
  {< I : Intervention | next-intervention : TIME-BETWEEN-INTERVENTIONS,
                      state : creation,
                      size : 0,
                      planktonic-cells : 0,
                      sessile-cells : 0,
                      c-di-GMP-global : 0,
                      biofilm : false >}
  in time TIME-GRANULARITY .

```

Note that when `TIME-BETWEEN-INTERVENTIONS` is set to infinite value `INF`, rule `tick-with-intervention` can never be applied.

3 In Silico Experiments and Formal Analysis

Maude modules are executable and the Real-Time Maude toolset provides a variety of formal analysis methods.

The *timed rewriting* command

```
(tfrew [r] in m : s0 in time <= t .)
```

simulates *one* of the system behaviours of module *m* by rewriting the initial state *s*₀ using up to *r* term rewrites and up to duration *t*.

The timed rewriting command provides us with an important tool to perform in silico experiments, by simulating experiments that would last for hours within a few seconds. Although only one of the possible system behaviour of a nondeterministic system is shown during simulation, this may provide important information on the biological system evolution.

The *timed search* command supports the analysis of all possible behaviours from a given initial state, relative to the chosen time sampling strategy. The command performs *reachability analysis* by searching for the states that match a search pattern and are reachable in a given time interval (if indicated). There are several variants for the syntax of the search command, such as:

```

(tsearch [n] in m : s with no time limit .)
(tsearch [n] in m : s in time <= t .)
(tsearch [n] in m : s in time >= t .)

```

where *s* is the search pattern, *n* is the number of solutions searched for and *t* is the time at which to stop the search.

Section 3.1 introduces the case study that we use to illustrate our approach. Section 3.2 illustrates the use of timed rewriting to perform in silico experiments and Sect. 3.3 illustrates the use of timed search to perform formal analysis.

3.1 *Pseudomonas Aeruginosa*

Pseudomonas aeruginosa is a Gram-negative microorganism, an opportunistic pathogen, actually one of the most important pulmonary pathogens and the predominant cause of morbidity and mortality in cystic fibrosis. *P. aeruginosa* forms rapidly biofilms on plastic surfaces, agar medium, graphite, metals and other materials. The composition of its extracellular polymeric substance has been studied in details and high quality data are available on the c-di-GMP distribution in cells, division upon replication and concentration with time measured through fluorescence in recombinant laboratory strains.

The life cycle of *P. aeruginosa* has a duration of 120 min [31] and the concentration of c-di-GMP internal to a cell in relation to the cell age is given in Table 1 [1].

3.2 Simulation

In order to illustrate our approach we consider the simulation reported in Table 2. We have defined Real-Time Maude module P-AERUGINOSA, which models a simplified bacterial ecosystem with values of the parameters taken from the literature on *P. aeruginosa* [1,25,31]. The model is defined without intervention. The results in Table 2 have been produced using the command

```
(tfrew [1000000] in P-AERUGINOSA : {init} in time <= t .)
```

with t representing the considered time in minutes (second column of Table 2).

The duration of the simulation is 12 h. We use an initial population of 120 cells in order to have a uniform distribution of the cells within the 120 min of cell life cycle. We adopt a 10 min time granularity to be able to observe cell death at each time increment, when the population size is above the REPRODUCTION-THRESHOLD threshold. In fact, due to the small population size we considered, using a finer granularity would result in a 0% death rate for each time increment. The small population size has been chosen due to the illustrative purpose of the simulation and to the need to have a computational response time of the order of seconds.

The chosen 10 min time granularity is a sufficient time step to model the biofilm formation phenomenon and is within the range adopted in previous studies. In general the time step is chosen depending on the bacterial species considered and on the phenomenon under investigation (e.g., biofilm initial attachment, biofilm detachment, viscoelastic modification of the matrix, etc.). It is commonly accepted that the time step for microbial phenomena is much higher than for hydrodynamic phenomena. The time step used in previous studies range from five minutes in *P. aeruginosa* biofilm growth [7] to one hour for oral biofilm formation [15] and hours in biofilm formation and bacteria decay [26].

The simulation in Table 2 shows a growth of the population size and of the global concentration of c-di-GMP due to cell reproduction, with a biofilm formation at time 2:50, when the c-di-GMP global concentration threshold is reached. The growth of both values then continues until time 3:00, when the population

Table 1. Concentration of c-di-GMP internal to a cell of *A. aeruginosa* [1].

Cell age	0–20	20–40	40–60	60–120
c-di-GMP (nM)	40	270	110	40

Table 2. In silico experiments performed with time granularity 10 for an initial population of 120 cells with 120 min time to reproduction, reproduction threshold of 350 cells, death rate 5% and c-di-GMP concentration: 200 for cell transition to sessile and back to planktonic and 40,000 for biofilm transition and back to disperse community.

Time		Population size			Global c-di-GMP	Biofilm presence	From death or reproduction
hr:min	min	total	planktonic	sessile			
00:00	0	120	100	20	15 500	No	Initial population
00:10	10	130	110	20	15 500	No	Reproduction
00:30	30	150	130	20	16 300	No	Reproduction
01:00	60	180	140	40	23 400	No	Reproduction
02:00	120	240	200	40	30 600	No	Reproduction
02:30	150	300	260	40	32 600	No	Reproduction
02:40	160	320	260	60	38 000	No	Reproduction
02:50	170	340	260	80	43 400	Yes	Reproduction
03:00	180	360	280	80	46 800	Yes	Reproduction
03:10	190	359	279	80	51 490	Yes	Death
03:30	210	357	277	80	54 070	Yes	Death
04:00	240	354	316	38	45 840	Yes	Death
04:30	270	351	312	39	41 530	Yes	Death
04:40	280	350	290	60	44 920	Yes	Death
04:50	290	389	310	79	45 290	Yes	Reproduction
05:00	300	388	310	78	50 501	Yes	Death
06:00	360	382	346	36	48 680	Yes	Death
10:00	600	358	326	32	45 440	Yes	Death
11:00	660	352	280	72	45 650	Yes	Death
11:10	670	351	279	72	46 680	Yes	Death
11:20	680	350	239	111	56 940	Yes	Death
11:30	690	366	255	111	57 220	Yes	Reproduction
12:00	720	363	317	46	48 310	Yes	Death

threshold that stops reproduction (**REPRODUCTION-THRESHOLD**) is exceeded (population size $360 > 350$ threshold). Cells start then to die until the population size drops again down to the threshold level at time 4:40. We have then an alternation of long time intervals of cell death and short time intervals of reproduction, with the biofilm persisting throughout. This behaviour mimics on a smaller scale real cell behaviour and is based on biological data for *P. aeruginosa* [1, 25, 31]. However, we can also consider a single cell as a representative of a cluster of cells that, globally, exhibits a homogeneous behaviour. This approach is commonly used and considered realistic [22, 27].

3.3 Formal Analysis Using Model Checking

The search command can be used to *predict* the biofilm formation in an in silico experiments. With reference to the experiment considered in Sect. 3.2, we can use the search command

```
(tsearch [3] in P-AERUGINOSA init =>*
  {C:Configuration
    < P-aeruginosa : Population | state : attributes-updated,
                                     biofilm : true >}
  with no time limit .)
```

to find the solutions for times 170, 180 and 190 in Table 2. Note that we must include the state `attributes-updated` in the search pattern to make sure that the solution has all attributes updated for the specific time unit.

Moreover, the results of the experiment shown in Table 2 seem to suggest that the biofilm, once formed, will persist forever. Such a conjecture cannot be validated through an experiment, since we do not know for how long we need to continue the experiment to possibly observe the biofilm disappearing. However, using the search command

```
(tsearch [1] in P-AERUGINOSA init =>*
  {C:Configuration
    < P-aeruginosa : Population | state : attributes-updated,
                                     biofilm : false >}
  in time >= 170 .)
```

we find out that after 860 min (14 h and 20 min) the biofilm disappears and the dispersed population consists of 385 cells (357 planktonic and 28 sessile) with 38 670 as a global c-di-GMP concentration. This proves that the conjecture suggested by the experiment results was actually false.

We have also defined module P-AERUGINOSA-INTERVENTION, which extends module P-AERUGINOSA by modelling a replacement intervention to be repeated every 160 min, namely TIME-BETWEEN-INTERVENTIONS equals 160. Then the search command

```
(tsearch [1] in P-AERUGINOSA-INTERVENTION init =>*
  {C:Configuration
    < P-aeruginosa : Intervention | state : attributes-updated,
                                     biofilm : true >}
  in time <= 3000 .)
```

will return ‘No Solution’ thus predicting that no biofilm would be formed even if we could extend the duration of the experiment to 3000 min (50 h).

4 Conclusion and Future Work

We have used Real-Time Maude to define an approach for the formal modelling and analysis of early biofilm formation. Our approach supports both simulation to mimic lab experiments and formal analysis by exploiting Real-Time Maude timed search command to predict

1. a specific outcome of an experiment;
2. the absence of a given outcome from all possible results of an experiments.

Prediction 2 is essential for planning effective interventions to prevent the formation of a biofilm. Although the kind intervention considered in this paper is a simple replacement, as part of our future work we plan to consider more sophisticated forms of intervention, such as the use of antibacterial and quenching.

Quenching refers to the capture/deactivation of *c*-di-GMP and other messengers or signalling molecules involved in the process of biofilm formation and maturation. Recent examples of *c*-di-GMP quenching agents include the aromatic compound coumarin [32], the fungal-produced antimicrobial terrein [13], the immunosuppressive drug azathioprine [5] and the antimetabolite drug sulfathiazole [4]. These and other *c*-di-GMP quenching compounds can be used to reduce biofilm formation without eliciting antimicrobial resistance in the target microorganisms.

We also plan to extend our model by including more accurate biological mechanisms and by improving the precision of the generated behaviour. For example, in terms of biological mechanisms, we have neglected the cell grows through time and normalised the *c*-di-GMP concentration with respect to the growing volume of the cell. Separating the two mechanisms, cell growth and *c*-di-GMP concentration, would result in a more accurate model. For the case study of *P. aeruginosa*, detailed microbial growth parameters are reported by Schleheck *et al.* [25]: duration of the log growth phase, stationary phase, and death phase. The precision of the generated behaviour could be improved by using the `Population` class to record further global aspects in order to support the mapping of a global behaviour observed in lab experiments to the behaviour of individual cells. For example, we could record the death toll over a period of time that equals the life cycle and make adjustments during the rewrite process to keep it within two given thresholds in accordance with the given birth rate. This would allow us to avoid the problem discussed in Sect. 2.5 and make the `death` rewrite rule work with higher precision and with no constraints on the size of the population and time granularity. Finally we plan to consider a time interval rather than an exact time for cell life-cycle duration and, at a later stage, introduce stochasticity into our model.

References

1. Abel, S., et al.: Bi-modal distribution of the second messenger *c*-di-GMP controls cell fate and asymmetry during the *Caulobacter* cell cycle. *PLOS Genet.* **9**(9), e1003744 (2013)
2. Ahmadi, M., Jorfi, S., Kujlu, R., Ghafari, S., Soltani, R.D.C., Haghhighifard, N.J.: A novel salt-tolerant bacterial consortium for biodegradation of saline and recalcitrant petrochemical wastewater. *J. Environ. Manag.* **191**, 198–208 (2017). <https://doi.org/10.1016/j.jenvman.2017.01.010>
3. Alpkvist, E., Klapper, I.A.: Multidimensional multispecies continuum model for heterogeneous biofilm development. *Bull. Math. Biol.* **69**, 765–789 (2007). <https://doi.org/10.1007/s11538-006-9168-7>

4. Antoniani, D., Maciag, P.B.A., et al.: Monitoring of diguanylate cyclase activity and of cyclic-di-GMP biosynthesis by whole-cell assays suitable for high-throughput screening of biofilm inhibitors. *Appl. Microbiol. Biotechnol.* **85**, 1095–1104 (2010). <https://doi.org/10.1007/s00253-009-2199-x>
5. Antoniani, D., Rossi, E., Rinaldo, S., Bocci, P., Lolicato, M., Paiardini, A., et al.: The immunosuppressive drug azathioprine inhibits biosynthesis of the bacterial signal molecule cyclic-di-GMP by interfering with intracellular nucleotide pool availability. *Appl. Microbiol. Biotechnol.* **97**, 7325–7336 (2013). <https://doi.org/10.1007/s00253-013-4875-0>
6. Bartocci, E., Lió, P.: Computational modeling, formal analysis, and tools for systems biology. *PLoS Comput. Biol.* **12**(1), e1004591 (2016). <https://doi.org/10.1371/journal.pcbi.1004591>
7. Biggs, M.B., Papin, J.A.: Novel multiscale modeling tool applied to *Pseudomonas aeruginosa* biofilm formation. *PLoS ONE* **8**(10), e78011 (2013). <https://doi.org/10.1371/journal.pone.0078011>
8. Christen, M., Kulasekara, H.D., Christen, B., Kulasekara, B.R., Hoffman, L.R., Miller, S.I.: Asymmetrical distribution of the second messenger c-di-GMP upon bacterial cell division. *Science* **328**, 1295–1297 (2010)
9. Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S.A., Kjelleberg, S.: Biofilms: an emergent form of bacterial life. *Nat. Rev. Microbiol.* **14**, 563–575 (2016)
10. Flemming, H.-C., Wuertz, S.: Bacteria and archaea on earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247–260 (2019)
11. Francolini, I., Donelli, G.: Prevention and control of biofilm-based medical-device-related infections. *FEMS Immun. Med. Microbiol.* **59**(3), 227–238 (2010)
12. Jin, X., Marshall, J.S., Wargo, M.J.: Hybrid model of bacterial biofilm growth. *Bull. Math. Biol.* **82**(27), 1–32 (2020). <https://doi.org/10.1007/s11538-020-00701-6>
13. Kim, B., Park, J.S., Choi, H.Y., et al.: Terrein is an inhibitor of quorum sensing and c-di-GMP in *Pseudomonas aeruginosa*: a connection between quorum sensing and c-di-GMP. *Sci. Rep.* **8**(8617), 1–13 (2018). <https://doi.org/10.1038/s41598-018-26974-5>
14. Laventie, B.-J., et al.: A surface-induced asymmetric program promotes tissue colonization by *Pseudomonas aeruginosa*. *Cell Host Microbe* **25**(1), 140–152 (2019). <https://doi.org/10.1016/j.chom.2018.11.008>
15. Martin, B., Tamanai-Shacoori, Z., Bronsard, J., Ginguené, F., Meuric, V., et al.: A new mathematical model of bacterial interactions in two-species oral biofilms. *PLoS ONE* **12**(3), e0173153 (2017). <https://doi.org/10.1371/journal.pone.0173153>
16. Mattei, M.R., Frunzo, L., D’Acunto, B., et al.: Continuum and discrete approach in modeling biofilm development and structure: a review. *J. Math. Biol.* **76**, 945–1003 (2018). <https://doi.org/10.1007/s00285-017-1165-y>
17. McDougald, D., Rice, S.A., Barraud, N., Steinberg, P.D., Kjelleberg, S.: Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nat. Rev. Microbiol.* **10**, 39–50 (2012)
18. Meseguer, J.: Twenty years of rewriting logic. *J. Logic Algebr. Program.* **81**, 721–781 (2012)
19. Ölveczky, P.C.: Real-Time Maude and its applications. In: Escobar, S. (ed.) *WRLA 2014*. LNCS, vol. 8663, pp. 42–79. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12904-4_3
20. Ölveczky, P.C.: *Designing Reliable Distributed Systems*. UTCS. Springer, London (2017). <https://doi.org/10.1007/978-1-4471-6687-0>

21. Ölveczky, P.C., Meseguer, J.: Semantics and pragmatics of Real-Time-Maude. High-Order Symbol. Comput. **20**(1–2), 161–196 (2007)
22. Paula, A.J., Hwang, G., Koo, H.: Dynamics of bacterial population growth in biofilms resemble spatial and structural aspects of urbanization. Nat. Commun. **11**(1354), 1–14 (2020). <https://doi.org/10.1038/s41467-020-15165-4>
23. Pearce, P., et al.: Flow-induced symmetry breaking in growing bacterial biofilm. Phys. Rev. Lett. **123**(25), 258101–258106 (2019). <https://link.aps.org/doi/10.1103/PhysRevLett.123.258101>
24. Păun, G.: P systems with active membranes: attacking NP-complete problems. Autom. Lang. Comb. **6**(1), 65–90 (2001)
25. Schleheck, D., et al.: Pseudomonas aeruginosa PAO1 preferentially grows as aggregates in liquid batch cultures and disperses upon starvation. PLoS ONE **4**(5), e5513 (2009)
26. Skoneczny, S.: Cellular automata-based modelling and simulation of biofilm structure on multi-core computers. Water Sci. Technol. **72**(11), 2071–2081 (2015)
27. Stewart, P., Franklin, M.: Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. **6**, 199–210 (2008). <https://doi.org/10.1038/nrmicro1838>
28. Stickler, D.: Bacterial biofilms in patients with indwelling urinary catheters. Nat. Rev. Urol. **5**, 598–608 (2008)
29. Wang, H., Masters, S., Edwards, M., Falkinham, J., Pruden, A.: Effect of disinfectant, water age, and pipe materials on bacterial and eukaryotic community structure in drinking water biofilm. Environ. Sci. Technol. **48**(3), 1426–1435 (2014)
30. Whiteley, M., Bangerter, M., Bumgarner, R., et al.: Gene expression in Pseudomonas aeruginosa biofilms. Nature **413**, 860–864 (2001). <https://doi.org/10.1038/35101627>
31. Yang, L., et al.: In situ growth rates and biofilm development of Pseudomonas aeruginosa populations in chronic lung infections. J. Bacteriol. **190**(8), 2767–2776 (2008)
32. Zhang, Y., et al.: Coumarin reduces virulence and biofilm formation in Pseudomonas aeruginosa by affecting quorum sensing, type III secretion and c-di-GMP levels. Front. Microbiol. **9**, 1952 (2018). <https://www.frontiersin.org/article/10.3389/fmicb.2018.01952>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

