PLoS one

Boolean Models of Biosurfactants Production in *Pseudomonas fluorescens*

Adrien Richard[1]*, Gaelle Rossignol[2], Jean-Paul Comet[1], Gilles Bernot[1], Jannine Guespin-Michel[2], Annabelle Merieau[2]

1 Laboratoire d'Informatique, Signaux et Systèmes Complexes de Sophia-Antipolis (I3S), Université de Nice-Sophia Antipolis and Centre National de la Recherche Scientifique (CNRS), Sophia-Antipolis, France, 2 Laboratoire de Microbiologie du Froid - Signaux et Micro-Environnement (LMDF-SME), Université de Rouen, Evreux, France

Abstract

Cyclolipopeptides (CLPs) are biosurfactants produced by numerous *Pseudomonas fluorescens strains*. CLP production is known to be regulated at least by the GacA/GacS two-component pathway, but the full regulatory network is yet largely unknown. In the clinical strain MFN1032, CLP production is abolished by a mutation in the phospholipase C gene (*plcC*) and not restored by *plcC* complementation. Their production is also subject to phenotypic variation. We used a modelling approach with Boolean networks, which takes into account all these observations concerning CLP production without any assumption on the topology of the considered network. Intensive computation yielded numerous models that satisfy these properties. All models minimizing the number of components point to a bistability in CLP production, which requires the presence of a yet unknown key self-inducible regulator. Furthermore, all suggest that a set of yet unexplained phenotypic variants might also be due to this epigenetic switch. The simplest of these Boolean networks was used to propose a biological regulatory network for CLP production. This modelling approach has allowed a possible regulation to be unravelled and an unusual behaviour of CLP production in *P. fluorescens* to be explained.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: richard@i3s.unice.fr

Introduction

Pseudomonas fluorescens are ubiquitous bacteria endowed with a capacity to adapt to many environments, namely by producing a wide range of extracellular molecules such as enzymes, toxins, antibiotics and biosurfactants. In these species, the main characterized biosurfactants are cyclolipopeptides (CLPs). They are involved in swarming motility and, depending on the strain and/or on the nature the considered CLPs, are sometimes required for biofilm formation [1].

Up to now studies have mainly focused on the effects of environmental conditions on CLP production (temperature, substrate, pH…), but data concerning transduction of the signals are quite limited [2]. Some regulators are involved in CLP production control in all *Pseudomonas* species, whereas others seem to be species dependent or even strain dependent. The global regulatory system GacS/GacA controls CLP production at a post-transcriptional level by the Rsm pathway in all *Pseudomonas* spp. [3–5]. On the other hand, in *P. fluorescens*, the requirement of the N-acylhomoserine lactone mediated quorum sensing for CLP production has only been demonstrated in a few strains [6]. In *P. fluorescens* SS101 and SBW25 (strains that do not produce N-acylhomoserine lactone), the LuxR-type transcriptional regulator positively regulates the production of CLPs. These LuxR-type regulators do not harbour the autoinducer-binding domain found in most quorum sensing associated LuxR regulators. Lux R-type genes are under the control of both GacS/GacA, and of the serine protease ClpP in a GacS/GacA-independent way [5].

In a previous study [7], we showed that the *P. fluorescens* clinical strain MFN1032 displays hemolytic activity due both to phospholipase C (PlcC) and CLPs. A knock out mutation in gene *plcC* leads to the expected loss in hemolytic activity, but also to the unexpected loss of CLPs. Furthermore, and even more unexpectedly, complementation of the *plcC* mutant [7] restored phospholipase C activity, but not the production of CLPs. We suggested that a not yet identified regulation takes place between PlcC and CLPs. In addition, CLP production in MFN1032 is subjected to phenotypic variation. After several generations, in specific conditions at 37°C, 4.10^{-3} variants per generation are no longer able to produce CLPs. Fifty percent of these variants are mutants in the gacA/gacS genes but 50% are not [8].

Heterogeneity within a bacterial population is a well-known phenomenon, which supposedly allow bacteria to invade a wide range of ecosystems and increases the overall fitness of the species. The heterogeneity could result from genetic rearrangement (high rate mutations or inversions of a DNA segment), or DNA methylation, but also to epigenetic switches [9]. Thomas argued that a positive feedback circuit (that may be the self-stimulation of a pivotal transcription factor) is always necessary for such switches [10]. This property has been proved mathematically in continuous [11], Boolean [12] and discrete [13] dynamic frameworks. This property has been documented in numerous bacterial regulatory systems, and demonstrated in some of them using mathematical modelling of the network's dynamics. In a previous study [14], we demonstrated, using Thomas generalised logic modelling followed

by experiments suggested by the model, that the Type 3 secretion system (T3SS) in *Pseudomonas aeruginosa* is subjected to an epigenetic switch between a non-inducible and an inducible state. The positive feedback is due to the self-regulating gene *exsA*, a general positive regulator of the regulon. More recently, MacLeane and Studholme [15] used a Boolean model to predict the dynamic behaviour of the regulation of the T3SS in *Pseudomoans syringae*, displaying a bistability under the control of the GacA/GacS node.

Here we used a Boolean approach to construct all the minimal regulatory networks of CLP production, consistent with the observations previously reported. We show that the simplest explanation for these behaviours is an epigenetic switch of the production of CLPs, the pivotal self regulating gene being unknown.

Methods

2.1 Boolean network

A *Boolean network* with n components is here defined as function f associating with each vector x of n Boolean components a vector $f(x)$ of n Boolean components:

$$f : \{0,1\}^n \rightarrow \{0,1\}^n, \quad x = (x_1, \ldots, x_n) \mapsto f(x) = (f_1(x), \ldots, f_n(x)).$$

In the context of genetic regulatory systems, each component i is seen as a gene that encodes a protein P_i ($i = 1, \ldots, n$). The vector x describes the state of the system in terms of the presence or absence of the proteins: protein P_i is present if $x_i = 1$, and absent if $x_i = 0$. Function f describes the activity of the genes according to the state x of the system: gene i is "on" if $f_i(x) = 1$, and "off" if $f_i(x) = 0$. The evolution of the system is driven by the expression of the genes: if $x_i < f_i(x)$ then x_i is called to increase (protein absent, gene "on"); if $x_i > f_i(x)$ then x_i is called to decrease (protein present, gene "off"); and if $x_i = f_i(x)$, then x_i is not called to change.

2.2 Asynchronous state graph

As argued by Thomas, the appearance or disappearance of a protein occurs after a vacillating delay [16]. As a consequence, the simultaneously updating of several variables implies unlikely equalities between delays. It is thus reasonable to assume that at most one variable is updated at each step (this is the so called asynchronous hypothesis). Without additional information on the delays, this leads to indecision in the variable to update, and all the possibilities are considered (if k variables are called to change, then k next states are possible). Formally, as in Thomas' logical method [16], the dynamic of network f is described by a directed graph, called the *asynchronous state graph* of f: the vertex set is the set of states $\{0,1\}^n$; for each state x and each component i such that $x_i \neq f_i(x)$, there is an arc (or transition), from state x to state $x' = (x_1, \ldots, f_i(x), \ldots, x_n)$.

2.3 Attractors

If $f(x) = x$, then x has no outgoing arc in the asynchronous state graph, and it corresponds to a *stable state* of the network. More generally, an *attractor* is a smallest non-empty subset of states that we cannot leave by following the paths of the state graph (stable states are attractors of size one). If we assume that, at each time t, each transition starting from the current state x^t can be taken with a probability strictly greater than 0, then, the probability that x^t belongs to an attractor tends toward 1 when $t \rightarrow \infty$. Then, one can consider that, at equilibrium, the system is always inside an attractor.

2.4 Interaction graph

The *interaction graph* of f, which describes the structure of the network, is the signed directed graph defined by: the vertices are the components $1, \ldots, n$, and there is a positive (resp. negative) arc $j \rightarrow i$ if there is a state x such that the difference

$$f_i(x_1, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_n) - f_i(x_1, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_n)$$

is positive (resp. negative). Hence, there is an interaction from j to i whenever function f_i depends on variable x_j, that is, whenever the activity of gene i depends on the level of protein P_j. Actually, the definition says that a positive (resp. negative) interaction from j to i exists if, in at least one state x, the appearance of P_j turns on (resp. turns off) gene i. It is possible that one may have both a positive and a negative interaction from j to i (in such a case, the sign of the interaction depends on the state of the system).

2.5 Muted network

We denote by f^i the Boolean networks with n components defined by

$$f^i(x) = (f_1(x), \ldots, f_{i-1}(x), 0, f_{i+1}(x), \ldots, f_n(x)).$$

Therefore, f^i and f are identical except that, in f^i, gene i is always "off". One can see f^i as the Boolean network resulting from a knock-out mutation of i in network f. Naturally, the interaction graph of the muted network f^i is obtained from the one of f by deleting all the incoming interactions of i.

2.6 Computation of consistent networks

In the next section, we are interested in the set of all the minimal Boolean networks whose asynchronous state graph satisfies some properties extracted from biological observations. To compute these networks, we used SMBioNet software [17,18], which is dedicated to the generation and verification of discrete dynamical systems. For instance, this tool can take as input an interaction graph G and a temporal property P written in the Computational Tree Logic (CTL), and return the set of all the Boolean networks f such that the interaction graph of f is a subgraph of G, and such that the asynchronous state graph of f satisfies property P. In particular, if G is the complete interaction graph on n vertices, then the output consists in all the Boolean networks with n components whose asynchronous state graph satisfies property P. However, since the number of Boolean networks with n components is $(2^n)^{(2^n)}$, i.e. over exponential with n, this works only for very small n values, typically $n < 7$.

Results and Discussion

3.1 Biological connections between CLPs and PlcC

Table 1 summarizes the results obtained from our previous studies on the role of a phospholipase C in the hemolytic activity of *P. fluorescens* clinical strain MFN1032. This strain displayed hemolytic activity (HA) and PlcC and CLPs (which we previously showed were necessary for hemolytic activity). A mutation in *plcC* abolished not only the hemolytic activity, but also CLP production, whereas complementation of gene *plcC* (by introducing a plasmid carrying the *plcC* gene under the control of a weak promoter in strain MFN1037) restored PlcC production but not CLPs. The same phenotype was observed after reversion of the PlcC deficient mutant.

How can we explain this surprising, but reproducible, process? The aim of this section is to propose a partial answer by exhibiting

Table 1. Summary of experimental results concerning the production of hemolytic activity, phospholipase C and cyclolipopeptides in different genetic backgrounds.

Strains of Pseudomonas fluorescens	Hemolytic activity (HA)	Phospholipase (PlcC)	Cyclolipopeptides (CLPs)
MFN1032 (WT)	+	+	+
MFN1037 (*plcC* mutant)	−	−	−
MFN1038 (*plcC* mutant complemented with *plcC*)	−	+	−
MFN1032 Group 1 variants	−	−	−
MFN1032 (Group 1 variants complemented by *gacA* or *gacS* gene)	+	+	+
MFN1032 Group 2 variants	−	+	−

+: Presence; −: Absence (early stationary growth phase in LB medium at 28°C).
doi:10.1371/journal.pone.0024651.t001

Boolean models that "explain" the mechanism underlying this process by the presence of genetic regulatory systems displaying two attractors, which describe an epigenetic switch. For that, we proceeded in two steps. First, we interpreted the previous observations in terms of Boolean networks, and then, we exhibited Boolean networks that were consistent with this interpretation.

3.1.1 Interpretation in terms of Boolean networks of connections between CLPs and PlcC. For the interpretation, we used the following hypotheses. (1) The production of CLPs in strain MFN1032 is controlled by an asynchronous Boolean network f, which contains at least two components: a component c for *plcC*, and a component b for the production of CLPs. (2) The production of CLPs in a *plcC* mutant (strain MFN1037) is controlled by Boolean network f^c. (3) The production of CLPs in the complemented *plcC* mutant (strain MFN1038) is controlled by the initial Boolean network f. In other words, we assume that the complementation of *plcC* mutation restores the WT network. (4) The experimental results have been obtained at equilibrium, where the current state of network f is inside an attractor.

Table 1 shows that MFN1032 displays hemolytic activity, and both PlcC and CLPs. The interpretation is that, in MFN1032, the current state of network f belongs to an attractor A such that $x_c = 1$ and $x_b = 1$ for every state x in A. The fact that a mutation of *plcC* abolishes CLP production is then interpreted as follows: originally, the state of the network is in attractor A, and a *plcC* mutation leads the system to an attractor in which CLPs are absent. Hence, the asynchronous state graph of muted network f^c contains a path from A to an attractor B such that $x_c = 0$ for every state x in B. Finally, the fact that complementation of *plcC* mutation restores PlcC production, but not for CLPs', is interpreted as follows: once the *plcC* mutation has led the network to attractor B, a complementation of the mutation leads the network to an attractor in which PlcC is present and CLPs are absent. Therefore, the asynchronous state graph of f contains a path from B to an attractor C such that $x_c = 1$ and $x_b = 0$ for every state x in C.

To be consistent with this interpretation of the results, f has to satisfy the following property P:

The asynchronous state graph of f contains an attractor A such that: (1) $x_c = 1$ and $x_b = 1$ for every state x in A (MFN1032); (2) the asynchronous state graph of f^c contains a path from A to an attractor B such that $x_b = 0$ for every state x in B (MFN1037); and (3) the asynchronous state graph of f contains a path from B to an attractor C such that $x_c = 1$ and $x_b = 0$ for every state x in C (MFN1038).
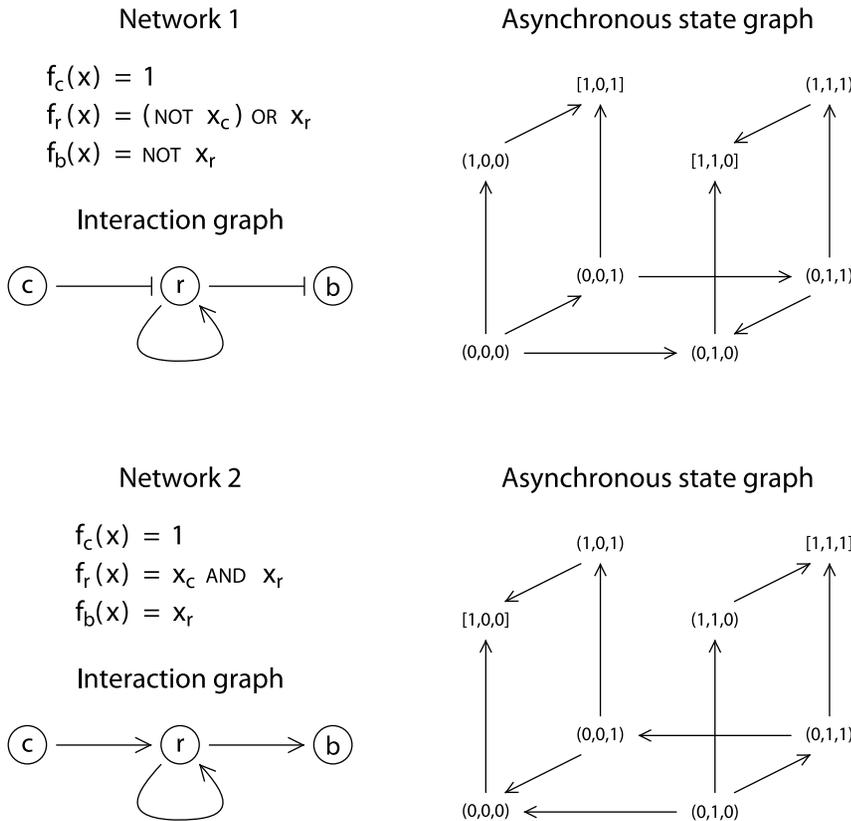
3.1.2 Boolean networks for connections between CLPs and PlcC. Now, we must determine whether Boolean networks exist which satisfy property P. First we observe that property P

implies that the asynchronous state graph of f has at least two distinct attractors A and C such that $x_c = 1$ for all states x in A or C. It has been proved that such a property requires the presence, in the interaction graph of f, of at least one positive circuit, which does not contain vertex c (a positive circuit is a directed cycle with an even number of negative arcs and without repeated vertexes). This theorem is known as the Boolean version of the first Thomas' conjecture [12]. Since it is accepted that CLPs do not have regulatory activity, the interaction graph of a consistent network f cannot have interactions starting from b. So the required positive circuit of the interaction graph of f contains neither c nor b. This means that at least a third component, say r, is needed.

Using intensive computation, we have generated the $(2^{2^2})^3 = 4096$ Boolean networks with three variables (c, r, b) such that b has no outgoing interaction, and demonstrated (by model checking) that exactly 12 of them satisfy property P (this generation/selection procedure takes less than one minute on a common laptop, with 2.5 GHz CPU and 4 GB RAM). These 12 consistent Boolean networks are given in Figure S1. In all cases, the interaction graph contains a positive arc from r to itself, and this positive loop on r corresponds to the required positive circuit. Furthermore, there is always a positive path from c to b, and this is consistent with the fact that the experimental results show that a mutation in *plcC* abolishes CLP production.

There are two minimal consistent Boolean networks with respect to the number of interactions. These two networks contain three interactions, and are shown in Figure 1. According to network 1 of Figure 1, *plcC* is always "on", gene r is "on" if its own product is present or if PlcC is absent, and CLPs are produced when the product of r is absent. Additionally, attractor A corresponds to stable state $(1,0,1)$, and attractor C corresponds to stable state $(1,1,0)$. Initially, the network is in stable state $(1,0,1)$, which is consistent with the phenotype of MFN1032 (PlcC and CLPs are present). Then, as illustrated in Figure 2, a mutation in *plcC* induces, firstly a disappearance of PlcC, and then an increase of r (due to the negative interaction from c to r) that causes a disappearance of CLPs (triggered by the negative interaction from r to b). These three changes lead the network in attractor B, which here corresponds to stable state $(0,1,0)$, and which is consistent with the phenotype of MFN1037 (PlcC and CLPs are absent). Finally, complementation of *plcC* restores the production of PlcC, but the positive loop on r allows r to remain present, and as a consequence, CLPs remain absent. The network then remains in stable state $(1,1,0)$, which is consistent with the phenotype of MFN1038.

Network 2 could be discussed similarly. *plcC* is always "on", gene r is "on" if its own product is present and if PlcC is present, and CLPs are produced when the product of r is present. In

Network 1

$$f_c(x) = 1$$
$$f_r(x) = (\text{NOT } x_c) \text{ OR } x_r$$
$$f_b(x) = \text{NOT } x_r$$

Interaction graph

Asynchronous state graph

Network 2

$$f_c(x) = 1$$
$$f_r(x) = x_c \text{ AND } x_r$$
$$f_b(x) = x_r$$

Interaction graph

Asynchronous state graph

Figure 1. The two minimal consistent Boolean networks with 3 components. The graphical conventions used to described the interaction graph of each network are as follows: normal arrows correspond to positive interactions, and T-end arrows correspond to negative interactions. Asynchronous state graphs are represented in a 3D-grid, with c in the vertical axis (first component), r in the horizontal axis (second component), and b in the third axis (third component).

addition, attractor A corresponds to stable state $(1,1,1)$, and attractor C corresponds to stable state $(1,0,0)$. Initially, the network is in stable state $(1,1,1)$, which is consistent with the phenotype of MFN1032 (PlcC and CLPs are present). A mutation in $plcC$ induces, firstly a disappearance of PlcC, and then a disappearance of r that induces a disappearance of CLPs. These
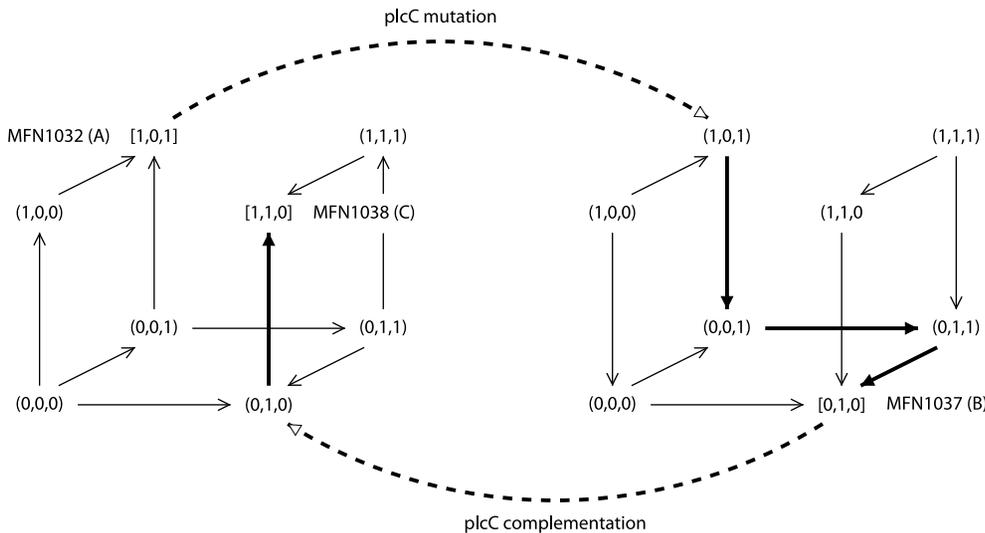
Figure 2. Illustration of property P on network 1. Asynchronous state graphs are represented in a grid, as in Figure 1. The left graph described the dynamics of network 1. The right graph described the dynamics of network 1 under $plcC$ mutation. Stable states A, B and C correspond to strains MFN1032, MFN1037 and MFN 1038, respectively. Dotted arrows represent the effect of mutation or complementation in $plcC$. Bold lines represent trajectories followed by the systems.

three changes lead the network in attractor *B*, which here corresponds to stable state (0,0,0), and which is consistent with the phenotype of MFN1037 (PlcC and CLPs are absent). Finally, complementation of *plcC* restores the production of PlcC, but *r* remains absent (PlcC alone is not sufficient to restore *r* expression), and as a consequence, CLPs remain absent. The network then remains in stable state (1,0,0), which is consistent with the phenotype of MFN1038.

Since PlcC is absent in the early growth phase and is not sufficient to restore *r* expression, this model seems less likely from a biological point of view, so we shall only discuss Network 1 in the following sections.

3.2 Connections between CLPs, PlcC and the GacS/GacA system

In a previous study, we have shown that strain MFN1032 displays phenotypic variations when incubated in specific conditions. The results are also summarized in Table 1.

Frequently, *gacS/gacA* mutants occur (as in other *Pseudomonas* spp.) and generate group 1 variants, where the WT phenotype is restored by complementation by one of the two genes encoding the Gac system. A second class of variants, group 2, is composed of variants, which are not restored to WT by complementation by the *gacA* or *gacS* genes. Group 2 variants have a phenotype similar to the *plcC* complemented mutant (strain MFN1038 or to the reverted *plcC* mutant) with respect to PlcC and CLP production. They neither produced CLPs nor hemolytic activity. The mechanism involved in this variation has not been elucidated [8].

3.2.1 Interpretation in terms of Boolean networks of connections between CLPs, PlcC and the GacS/GacA system. For the interpretation of group 1 variants, we consider a network *f* with at least three components: *c*, *b*, and an additional component *a* for the *gac* gene. The modelling hypotheses we used are similar (in particular, the production of CLPs in a *gac* mutant is controlled by f^a, and under a complementation of *gac* mutation, it is controlled by the initial Boolean network *f*).

As previously demonstrated, the WT phenotype is interpreted by the presence, in the asynchronous state graph of *f*, of an attractor *A* such that $x_c = 1$ and $x_b = 1$ for all states *x* in *A*. Resulting in the interpretation of the group 1 phenotype: originally, the state of the system is in *A*, and a *gacA* mutation leads the system to an attractor in which PlcC and CLPs are absent. The asynchronous state graph of f^a contains a path from *A* to an attractor *B'* such that $x_c = 0$ and $x_b = 0$ for all states *x* in *B'*. Finally, the fact that a complementation of a mutation in *gac* restores the wild type phenotype in group 1 variants is interpreted as follows: once the *gac* mutation has led the system to attractor *B*, a complementation of the mutation leads the system to an attractor in which PlcC and CLPs are present. The asynchronous state graph of *f* then contains a path from *B'* to an attractor *A'* such that $x_c = 1$ and $x_b = 1$ for every *x* in *A'*.

To be consistent with both this interpretation of the *gac* gene influence and the previous property *P*, the hypothetical network *f* has to satisfy the following property *P'*:

The asynchronous state graph of f contains an attractor A verifying points (1–3) of property P, and two additional points: (4) the asynchronous state graph of f contains a path from A to an attractors B' such that $x_b = 0$ and $x_c = 0$ for all states x in B' (Group 1); (5) the asynchronous state graph of f contains a path from B' to an attractor A' such that $x_c = 1$ and $x_b = 1$ for all states x in A (complementation of gac gene in Group 1).

3.2.2 Boolean networks for connections between CLPs, PlcC and the GacS/GacA system. We now exhibit Boolean

networks satisfying property *P'*. We assume that *b* has no outgoing interaction (as previously stated), and that *a* is an input of the system, i.e., has no incoming interaction. Since property *P'* implies property *P*, in order to satisfy *P'*, the interaction graph of *f* has to contain a positive circuit that does not contain vertex *c*. Since neither *b* nor *a* belong to a circuit (*b* has no outgoing interaction, and *a* no incoming interaction), the additional component *r* is again needed.

Using intensive computation, as previously done, we have generated, among the among the $2^{64} \sim 10^{19}$ Boolean networks with four components (*a*, *b*, *c*, *r*), the $(2^{2^4})^3 = 16\,777\,216$ Boolean networks such that *b* has no outgoing interaction, and such that $f_a(x) = 1$ (this is clearly a necessary condition for satisfying *P'*). Using model checking, we show that exactly 2400 of them satisfy *P'* (this generation/selection procedure takes about 82 hours on an usual laptop, with 2.5 GHz CPU and 4 GB RAM). The asynchronous state graph of each consistent network contains exactly two attractors, which are stable states. Furthermore, the interaction graph of each consistent network contains at least 6 interactions, and 12 consistent networks have an interaction graph with exactly 6 interactions. One of them is given in Figure 3. We only discuss this minimal consistent network since it can be seen as an extension of Network 1 given in Figure 1 (and, as argued in the previous section, this network is the most likely minimal 3-component network). Indeed, we can see that, in the hyper-plane "$x_a = 1$" (i.e. in the "top 3-dimensional cube") the dynamics of the network is the one of Network 1 given in Figure 1. As a consequence, the 4-component network is consistent with property *P*, as illustrated in Figure 4. Additionally, a *gac* mutation leads the network to stable state (0,0,0,0) where no product is present, which is consistent with group 1 phenotype, see Figure 5. Then, complementation (or reversion) of *gac* leads the network to the state (1,0,0,0) from which the original stable state (1,1,0,1), which is consistent with both the WT and the complemented group 1 phenotypes, can be reached. But from state (1,0,0,0), a path leading to stable state (1,1,1,0) also exists, which is consistent with both the complemented plcC mutant and group 2 phenotypes. According to this model, group 2 variants could be due to a *gac* mutation followed by a reversion. This is not unlikely since *gac* mutations and reversions are frequent in Pseudomonas spp [19].

Surprisingly, each of the 2400 consistent Boolean networks is in agreement with this hypothesis. Indeed, for each consistent Boolean network, once a *gac* mutation has led the network in attractor *B'*, a complementation induces two possible futures: either (1) the system reaches attractor *A*, which is consistent with both the WT and the complemented group 1 phenotypes, or (2) the system reaches attractor *C*, which is consistent with both the complemented *plcC* mutant and surprisingly, group 2 phenotypes.

3.2.3 From Boolean networks to biological regulatory networks. The simplest biological regulatory network (with a negative interaction) deduced from the model that accounts for all observations is shown in Figure 6. This regulatory network is likely to be an accurate model, not only for Boolean logic but also for generalized logic [16] where different thresholds or different delays are to be considered. For instance, literature underlines the fact that GacA/GacS system acts on a post-transcriptional level via the Gac/Rsm pathway [20] and via transcriptional mechanisms [5,21]. If the GacA-PlcC interactions were faster than the GacA-R interactions, the complementation of a *gacA* mutant would result in a wild type phenotype. If there are different thresholds of GacA for these two interactions, GacA = 0 when the *gacA* mutation is stable and the system reaches the group 1 attractor. But if there is only a transient decrease of GacA (stochastic event or fast mutation/reversion) with a threshold of
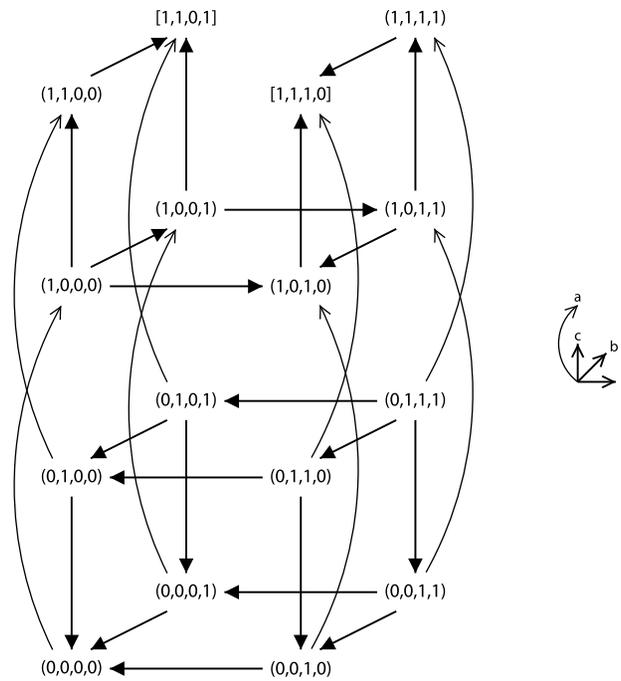
4–components consistent network

$$f_a(x) = 1$$
$$f_c(x) = x_a$$
$$f_r(x) = x_a \text{ AND } ((\text{NOT } x_c) \text{ OR } x_r)$$
$$f_b(x) = x_a \text{ AND } (\text{NOT } x_r)$$

Interaction graph
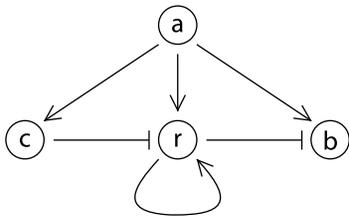
Asynchronous state graph

Figure 3. A minimal consistent Boolean network with 4 components. The asynchronous state graph is represented in a 4D-grid, with a in the curved axis (first component), c in the vertical axis (second component), r in the horizontal axis (third component), and b in the last axis (fourth component).
doi:10.1371/journal.pone.0024651.g003
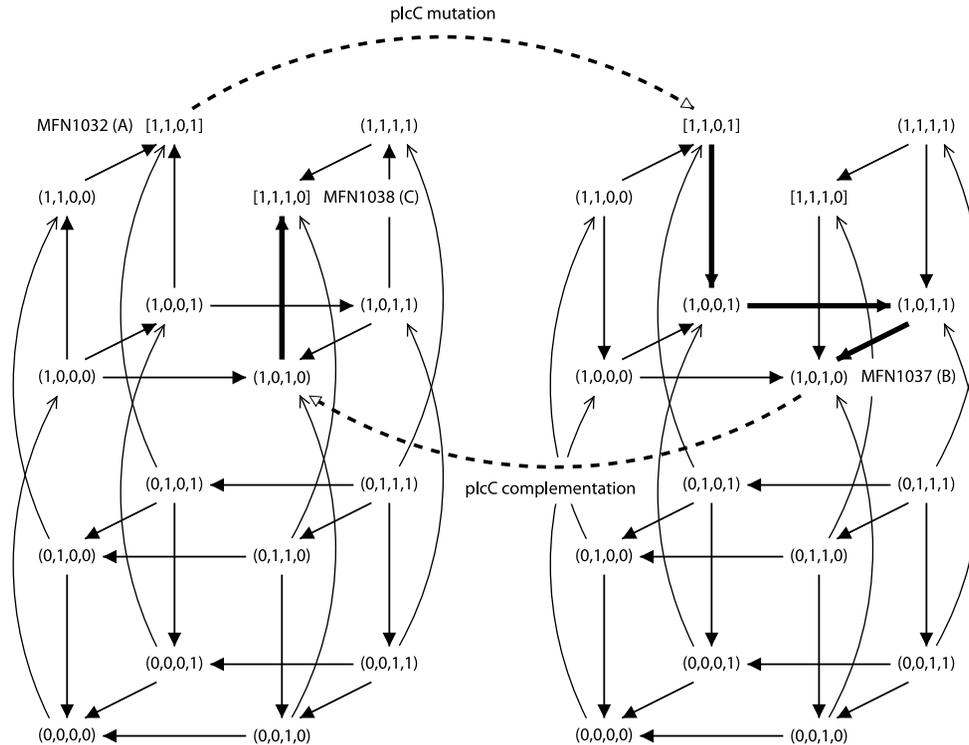
Figure 4. Illustration of property P on the consistent network of Figure 3. Asynchronous state graphs are represented in a 4D-grid, as in Figure 3. The left graph described the dynamics of network of Figure 3. The right graph described the dynamics of this network under $plcC$ mutation. Stable states A, B and C corresponds to strains MFN1032, MFN1037 and MFN1038, respectively. Dotted arrows represent the effect of mutation or complementation in $plcC$. Bold lines represent trajectories followed by the systems.
doi:10.1371/journal.pone.0024651.g004
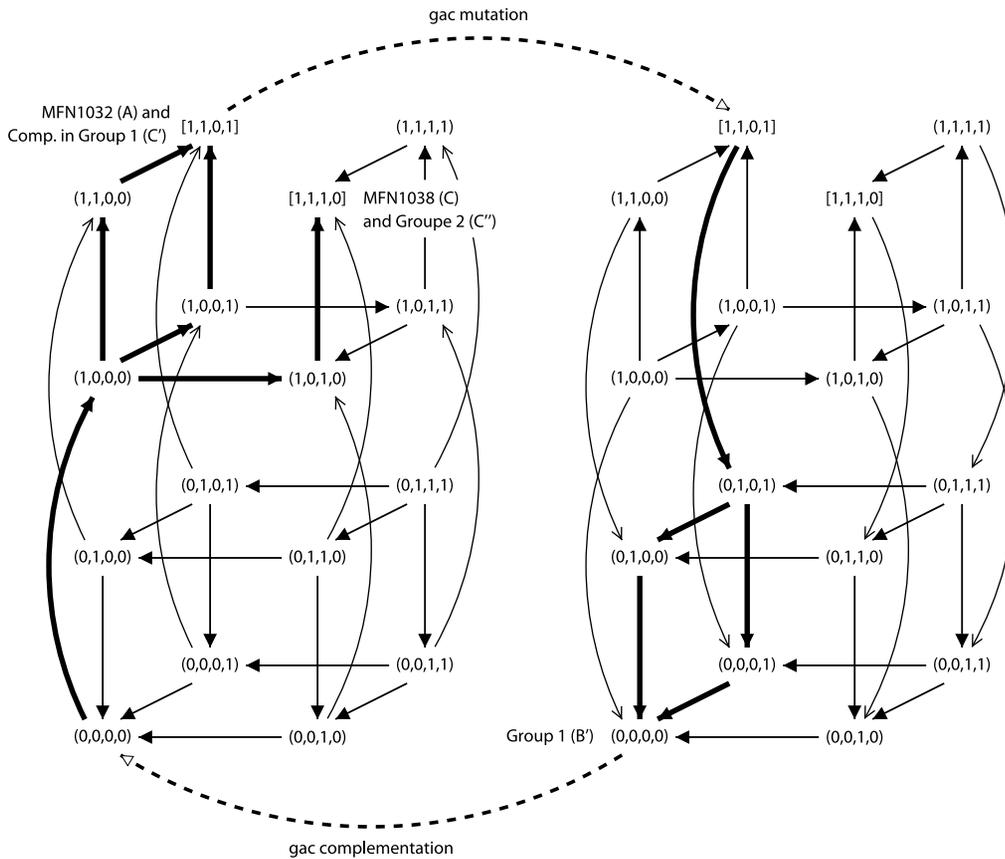
Figure 5. Illustration of property P' on the consistent network of Figure 3. Asynchronous state graphs are represented in a 4D-grid, as in Figure 3. The left graph described the dynamics of network of Figure 3. The right graph described the dynamics of this network under *gac* mutation. Stable states A and C' correspond to strains MFN1032 and group 1 variants complemented by *gacA* or *gacS* gene, stable state B' corresponds to group 1 variants, and stable states C corresponds to strains MFN1038 and group 2 variants. Dotted arrows represent the effect of mutation or complementation in *gac*. Bold lines represent trajectories followed by the systems.
doi:10.1371/journal.pone.0024651.g005

GacA sufficient for GacA-R interaction and not for GacA-PlcC interaction, the network would reach the group 2 attractor. So far, this regulatory model explains all our previous observations but the discovery of *r* remains the bottleneck of the experimental evidence.
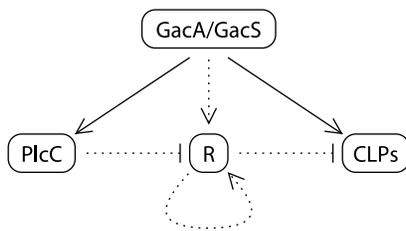
Figure 6. Hypothetic biological regulatory network of CLP production (with negative interactions) deduced from the Boolean model. Circles represent the proteins in the network and edges represent regulatory interaction (bold lines: known interactions, dotted lines: hypothetical interactions). Arrow headed edges represent positive regulatory interaction, and T-headed edges represent negative interaction. PlcC: phospholipase C; CLPs: cyclolipopeptides production; R: unknown regulator (with positive feedback).
doi:10.1371/journal.pone.0024651.g006

From a biological standpoint, this approach suggests at least three experiments to validate the model with a negative *r* interaction.

According to the dynamics of *gacA* interactions, a *gacA* mutant complementation can lead to a wild type phenotype or to a group 2 phenotype. We introduced *gacA* gene in a *gacA* mutant and screened 1 500 clones for CLP production. All the clones were restored in the wild type phenotype. This could indicate that, in our conditions, GacA-PlcC interactions is favoured instead of the GacA-R interaction.

Secondly, a pulse of R in the wild type strain could lead to group 2 variants. In literature, no negative regulator for CLP production has yet been described. Indeed, identification of such regulators is difficult with the routinely used random mutagenesis. The CLP phenotype of an *r* mutant in wild type MFN1032 must be of a wild type according to our model. Only R over-expression in MFN1032 should be detectable in terms of CLP phenotype. We analyzed data from transcriptomic consequences of a *gacA* mutation in *Pseudomonas fluorescens* strain Pf-5, a CLP producer [22]. We looked for a putative negative regulator under positive GacA transcriptional control. Among the 5 transcriptional factors positively regulated by GacA during the stationary phase, the only negative identifiable transcriptional regulator belonged to the GntR family. It seems especially relevant, since a *gntR* gene was identified downstream *plcC* gene in MFN1032. GntR regulators

have been reported to maintain their own expression [23], so a positive feedback loop regulating GntR levels may exists. We have promoted overexpression of GntR in the wild type strain MFN1032 by cloning *gntR* in an expression plasmid with *ptac* promoter. However this does not impair CLP production, which shows that GntR is not R.

And last, without other valid data from literature, random mutagenesis on *plcC* mutant or a group 2 variant may, by a mutation in *r*, restore wild type CLP production. Such an experiment might help to identify R.

Conclusion

From a biological point of view, the Boolean approach has proven particularly well suited for this study for several reasons: (i) the qualitative aspect of the observations, which are expressed in terms of the presence/absence of CLPs and PlcC; (ii) the lack of information about the genetic network controlling CLP production; (iii) reproducible but not easily nor understood straightforward observations.

The particularity of our modelling study consists in translating the observations in a formal way, and in computing all the Boolean networks which satisfy the property resulting from this translation, without any assumption on the topology of the considered network (except taking into account the minimal number of genes required). By considering only the minimal Boolean networks (those in which the number of interactions is also minimal) consistent with the observations, we are able to propose new hypotheses concerning the regulatory network of CLP production. There is a strong probability that a yet unknown regulator, endowed with positive self-regulation is negatively controlled by PlcC and negatively controls the production of CLPs.

By contrast, the hypothesis that type 2 variants are due to a mutation/reversion in one of the *gac* genes (which are known to be highly mutable) is strongly suggested by the models. Finally, the models show the consistency of our hypothesis of a bistability (or epigenetic switch) in CLP production. It is shown more and more often that bistability is a means for bacteria to maintain diversity in their populations. It could account for the diversity of CLP phenotypes observed in *Pseudomonas* strains, in which numerous non-CLP producer strains and CLP producers are found in the same environment [24,25].

Supporting Information

Figure S1 The 12 Boolean networks f with three components (c, r and b) that are consistent with property P and with an interaction graph in which b has no outgoing interaction. The graphical conventions used to described the interaction graph of each network are as follows: arrows of form $->$ correspond to positive interactions; arrow of form $-|$ correspond to negative interactions; and arrows of form $->|$ indicate the presence of both positive and negative interaction. Asynchronous state graphs are represented in a 3D-grid, with c in the vertical axis (first component), r in the horizontal axis (second component), and b in the third axis (third component). (EPS)

Acknowledgments

Author Contributions

Conceived and designed the experiments: AM GR JGM. Performed the experiments: AM GR. Analyzed the data: AM AR GB GR JGM JPC. Wrote the paper: AM AR GR JGM.

References

1. Raaijmakers JM, Bruijn I, Nybroe O, Ongena M (2010) Natural functions of lipopeptides from bacillus and pseudomonas: more than surfactants and antibiotics. FEMS Microbiol Rev 34: 1037–1062.
2. Raaijmakers JM, Bruijn I, Kock MJ (2006) Cyclic lipopeptide production by plant-associated pseudomonas spp.: diversity, activity, biosynthesis, and regulation. Mol Plant Microbe Interact 19: 699–710.
3. Koch B, Nielsen TH, Sorensen D, Andersen JB, Christophersen C, et al. (2002) Lipopeptide production in pseudomonas sp. strain dss73 is regulated by components of sugar beet seed exudates via the gac two-component regulatory system. Appl Environ Microbiol 68: 4509–4516.
4. Bruijn I, de Kock MJ, Yang M, de Waard P, van Beek TA, et al. (2007) Genome-based discovery, structure prediction and functional analysis of cyclic lipopeptide antibiotics in pseudomonas species. Mol Microbiol 63: 417–428.
5. de Bruijn I, de Kock MJ, de Waard P, van Beek TA, Raaijmakers JM (2008) Massetolide a biosynthesis in pseudomonas fluorescens. J Bacteriol 190: 2777–2789.
6. Cui X, Harling R, Mutch P, Darling D (2005) Identification of n-3-hydroxyoctanoyl-homoserine lactone production in pseudomonas fluorescens 5064, pathogenic to broccoli, and controlling biosurfactant production by quorum sensing. Eur J Plant Pathol.
7. Rossignol G, Merieau A, Guerillon J, Veron W, Lesouhaitier O, et al. (2008) Involvement of a phospholipase c in the hemolytic activity of a clinical strain of pseudomonas fluorescens. BMC Microbiol.
8. Rossignol G, Sperandio D, Guerillon J, Duclairoir Poc C, Soum-Soutera E, et al. (2009) Phenotypic variation in the pseudomonas fluorescens clinical strain mfn1032. Res Microbiol 160: 337–344.
9. Hallet B (2001) Playing dr jekyll and mr hyde: combined mechanisms of phase variation in bacteria. Curr Opin Microbiol 4: 570–581.
10. Thomas R (1981) On the relation between the logical structure of systems and their ability to generate multiple steady states and sustained oscillations. In: Series in Synergetics Springer. volume 9. pp 180–193.
11. Soulé C (2003) Graphical requirements for multistationarity. ComPlexUs 1: 123–133.
12. Remy E, Ruet P, Thieffry D (2008) Graphic requirement for multistability and attractive cycles in a boolean dynamical framework. Advances in Applied Mathematics 41: 335–350.
13. Richard A, Comet JP (2007) Necessary conditions for multistationarity in discrete dynamical systems. Discrete Applied Mathematics.
14. Filopon D, Merieau A, Bernot G, Comet JP, Leberre R, et al. (2006) Epigenetic acquisition of inducibility of type iii cytotoxicity in p. aeruginosa. BMC Bioinformatics 7: 272.
15. MacLean D, Studholme DJ (2010) A boolean model of the pseudomonas syringae hrp regulon predicts a tightly regulated system. PLoS One 5: e9101.
16. Thomas R, d'Ari R (1990) Biological feedback CRC Press.
17. Bernot G, Comet JP, Richard A, Guespin-Michel J (2004) Application of formal methods to biological regulatory networks: Extending thomas' asynchronous logical approach with temporal logic. Journal of Theoretical Biology 229: 339–347.
18. Khalis Z, Comet JP, Richard A, Bernot G (2009) The smbionet method for discovering models of gene regulatory networks. Genes, Genomes and Genomics.
19. van den Broek D, Bloemberg GV, Lugtenberg B (2005) The role of phenotypic variation in rhizosphere pseudomonas bacteria. Environ Microbiol 7: 1686–1697.
20. Lapouge K, Schubert M, Allain FH, Haas D (2008) Gac/rsm signal transduction pathway of gamma-proteobacteria: from rna recognition to regulation of social behaviour. Mol Microbiol 67: 241–253.
21. Brencic A, Lory S (2009) Determination of the regulon and identification of novel mrna targets of pseudomonas aeruginosa rsma. Mol Microbiol 72: 612–632.
22. Hassan KA, Johnson A, Shaffer BT, Ren Q, Kidarsa TA, et al. (2010) Inactivation of the GacA response regulator in Pseudomonas fluorescens Pf-5 has far-reaching transcriptomic consequences. Environ Microbiol 4: 899–915.
23. Rigali S, Schlicht M, Hoskisson P, Nothaft H, Merzbacher M, et al. (2004) Extending the classification of bacterial transcription factors beyong the helix-turn-helix motif as an alternative approach to discover new cis/trans relationships. Nucleic Acids Res 32: 3418–3426.
24. Tran H, Kruijt M, Raaijmakers JM (2009) Diversity and activity of biosurfactant-producing Pseudomonas in the rhizosphere of black pepper in vietnam. J Appl Microbiol 104: 839–851.
25. Nielsen TH, Sorensen D, Tobiasen C, Andersen JB, Christophersen C, et al. (2002) Antibiotic and biosurfactant properties of cyclic lipopeptides produced by fluorescent pseudomonas spp. from the sugar beet rhizosphere. Appl Environ Microbiol 68: 3416–3423.