
Mathematical modeling of the genetic regulatory network controlling competence for natural transformation in *Streptococcus thermophilus*

Thèse présentée en vue de l'obtention du grade de docteur en Sciences par

Gabrielle Laurie Haustenne

Promoteurs

Pr. Pascal Hols
Dr. Laetitia Fontaine

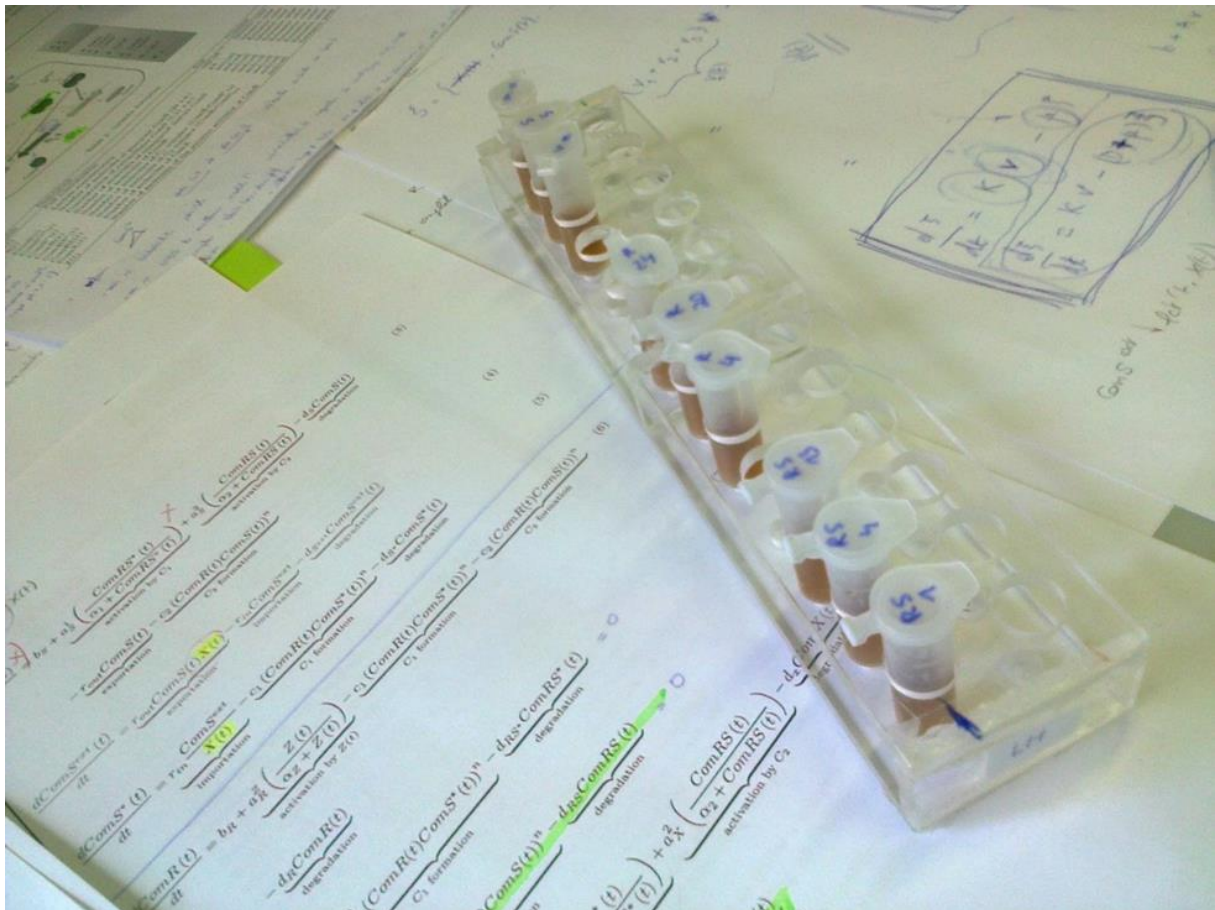
Membres du jury

Pr. Georges Bastin
Pr. Jacques Mahillon
Pr. Xavier De Bolle
Pr. Olivier Bernard

Président du jury

Pr. Bernard Hallet

Louvain-la-Neuve
Janvier 2017



“La magie des mathématiques, c’est de rendre visibles des choses cachées.”

COMPOSITION OF THE JURY

Pr. Georges Bastin (UCL/ICTEAM)

Pr. Jacques Mahillon (UCL/ELI)

Pr. Xavier De Bolle (UNamur)

EXTERNAL MEMBERS:

Pr. Olivier Bernard (INRIA/France)

SUPERVISORS:

Pr. Pascal Hols (UCL/ISV)

Dr. Laetitia Fontaine (UCL/ISV)

PRESIDENT:

Pr. Bernard Hallet (UCL/ISV)

REMERCIEMENTS

Au terme de cette aventure de quatre années, je tiens à remercier plusieurs personnes.

Tout d'abord, mes deux (presque trois) promoteurs : Pascal, Laetitia et Georges, pour m'avoir suivie avec patience tout au long de ce travail de longue haleine.

Merci à Pascal et Laetitia de m'avoir accueillie au sein du labo BBGM, alors que je ne connaissais encore pas grand-chose en génétique, et de m'avoir donné la chance de remédier à cela par un apprentissage « sur le tas » pendant lequel j'ai appris à construire mes souches mutantes, moi qui n'avais jamais vraiment tenu une pipette en main de ma vie. Je vous remercie pour votre encadrement régulier, votre attention et vos remarques constructives, votre patience devant mes erreurs de manipulation... ainsi que pour toutes les belles discussions – parfois mouvementées – que cette thèse nous a permis d'avoir.

Merci à Georges de m'avoir épaulée et soutenue vaillamment jusque dans les tréfonds abyssaux des équations différentielles pas du tout ordinaires, jusqu'à Viennes et jusqu'à Dakar ☺ Merci d'avoir apporté votre touche de génie à notre originale approche de modélisation. Merci pour tout, M'sieur !

Merci aussi à Bernard Hallet, pour sa bonne humeur et son soutien moral !

Je remercie également Patrice Soumilion, pour son aide et sa contribution à la rédaction de notre article Frontiers, pour son œil critique sur le modèle, et les nombreuses discussions sur la modélisation.

Merci aussi à Véronique Lebrun pour son aide précieuse et pour sa patience devant la catastrophe logistique que je suis !

Merci également à tous mes collègues de BBGM, et en particulier à mes voisins de bureau Cédric et Morgane ☺

Merci aussi à mes voisines de rédaction, Caroline et Oksana, ainsi qu'à Amandine, avec lesquelles j'ai partagé les douleurs de la dernière ligne droite.

Merci à ma famille et à ma belle-famille, et à tous mes amis, qui m'ont soutenue durant toute la thèse, quand parfois j'avais des moments de doute...

Et enfin, merci à mon chéri pour son soutien inconditionnel, pour la thèse, pour la rédaction de l'article (je peux vous dire qu'il en a vraiment bavé avec l'article) et dans tout ce que j'entreprends dans la vie. Je crois que s'il n'avait pas été là pour moi comme il l'a été, le manuscrit que vous tenez en main n'existerait pas ! Merci <3

Table of Contents

Abstract.....	13
List of abbreviations.....	14
Part I - Introduction	17
GENERAL CONTEXT	18
1. Competence for natural DNA transformation.....	21
A. Discovery	21
B. Definition.....	22
C. Dedicated machineries.....	25
I. Binding of the environmental dsDNA.....	26
II. Processing of the recruited dsDNA	27
III. Internalization of the ssDNA pieces	27
a. Crossing the outer membrane.....	27
b. Crossing the cytoplasmic membrane	28
D. Origin of the exogenous DNA.....	29
E. Functions	32
I. Genetic plasticity.....	32
II. Nutrition	33
III. Repair of DNA	33
F. Prevalence.....	33
I. Gram-negative bacteria.....	34
a. Proteobacteria.....	34
b. Deinococcus-Thermus	34
c. Cyanobacteria.....	34
II. Gram-positive bacteria.....	35
a. Actinobacteria	35
b. Firmicutes	35
G. Our model organism: <i>S. thermophilus</i>	35

2. Regulation of competence development	39
A. Observations	39
B. Regulation of competence by communication peptides	40
C. Competence regulation in Bsu	43
I. The ComK transcription factor.....	43
II. Transcriptional control	44
III. Post-transcriptional control	45
IV. Post-translational control	45
D. Competence regulation in streptococci	45
I. Proximal regulation of ComX abundance	46
a. Transcriptional control of <i>comX</i>	46
i. The ComCDE system: signal transduction by phosphorelay	46
ii. The ComRS system: signal transduction by pheromone reimportation	49
b. Post-translational control of ComX abundance.....	52
II. Distal regulation of ComX abundance.....	52
a. Distal regulatory pathway specific to Smu	54
b. Distal regulatory pathway common to Smu and Spn	56
c. Distal regulatory pathway in our model Sth.....	56
E. Subpopulation heterogeneity for competence development	57
F. Open questions about the ComRS system	61
3. Mathematical modelling of regulatory networks.....	62
A. What is a system ? What is a model ?	62
B. Why building a mathematical model ?.....	63
C. Descriptive <i>versus</i> mechanistic models.....	63
D. Particularities of biological models	65
E. How do we build a model ?	66
I. Conceptualization.....	66
II. Choice of the formalism.....	67
III. Transcription into equations	67
IV. Learning step	68

V. Validation step	69
VI. Optimization and use of the model.....	70
F. State of the art: what kind of models do exist ?	70
I. Model of the ComCDE system in Spn	71
II. Model of the BlpCHR/ComRS system in Smu	73
III. What is missing ? What do we want ?.....	73
Part II - Objectives of the thesis.....	75
Part III - Results.....	79
Chapter I – Frontiers in Microbiology	81
INTRODUCTION	85
MATERIALS AND METHODS	89
RESULTS	102
DISCUSSION.....	111
SUPPLEMENTARY MATERIAL	124
Chapter II - MATHMOD.....	141
INTRODUCTION AND MATHEMATICAL MODEL.....	143
FREQUENCY DOMAIN STABILITY	144
LYAPUNOV STABILITY	146
Chapter III – Application of the methodology	153
1. Framework.....	154
2. The analyzed ODE system.....	155
3. Finding the equilibria of the system	156
4. Stability analysis of the steady-points	157
5. Biological relevance of the analysis.....	159
Part IV - Discussion & perspectives	161
1. Modeling choices	163
2. The critical role of ComR on competence level.....	165
3. Oscillations in competence development	166
4. Competence shut-off	167
5. Bistability in competence development	168
6. Model limitations and future improvements	169
7. Conclusion	172
REFERENCES.....	173

Abstract

Bacteria are everywhere, and can be either highly beneficial or dangerous for humans. To evolve, they have developed sophisticated systems enhancing their genetic plasticity. Competence for natural transformation is a horizontal gene transfer mechanism enabling competent cells to stably acquire exogenous DNA. It is a transient physiological state, characterized in streptococci by a high level of ComX alternative sigma-factor, the master regulator of competence, which re-directs the transcription program towards the activation of genes involved in DNA acquisition. Because of the energetic cost of competence, ComX production is tightly regulated. The ComRS signaling system controls *comX* expression in the industrial dairy species *Streptococcus thermophilus*. Interestingly, the transformability of this species varies dramatically between different strains.

This work aims at shedding light on the poorly understood inner workings of the ComRS system, by combining mathematical modeling and experimental approaches in order to uncover the topological hierarchy and dynamics of the timing-device function of the system, and to unravel the shut-off mechanism of competence in *S. thermophilus*. Moreover, the model aimed at better understanding the variability in competence levels observed among different strains of this species.

For that purpose, we built a deterministic, population-scaled model of the time-course evolution of proximal players regulating *comX* expression. The model equations were designed to capture all the *a priori* knowledge available about the ComRS proximal system. First, an original forth-and-back learning step was used to calibrate the model parameters, by using a set of experimental data consisting in transcriptional profiles obtained from luciferase reporter strains in various mutant backgrounds. Second, an independent set of experimental data was used in order to assess the model quality, by testing its robustness against varying simulated scenarios, and its ability to predict the system behavior in untested conditions. The validation was performed at three levels: topological organization, kinetic properties, and reactivity of the system upon addition of the ComS signaling peptide. Once fully validated, the model was used to infer new knowledge about the ComRS system.

In silico simulations and experimental data revealed that ComR is the main limiting factor for *comX* expression. Moreover, our analysis strongly suggests that the level of ComR production could explain the observed heterogeneity in competence levels between different strains of *S. thermophilus*. The model was next challenged by a subsequent *in silico* analysis, totally computational, focusing on the stability analysis of the steady points of the system. Two complementary approaches were proposed to determine the behavior of the system around that points. We showed that the system has two steady points: the "ON" state, corresponding to competence, and the "OFF" state, corresponding to the vegetative state. Mathematical work showed that in the case of *S. thermophilus*, the ComRS system ON state was unstable, while its OFF state was stable.

List of abbreviations

ABC	ATP-binding cassette
AI	Auto-inducers
Bp	Base pair(s)
Bsu	<i>Bacillus subtilis</i>
CDM	Chemically defined medium
<i>com</i>	competence genes
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
CSP	Competence stimulating peptide
DDE	Delay Differential Equations
dof	Degrees of freedom
dsDNA	Double-stranded DNA
DUS	DNA uptake sequence
EPS	Exopolysaccharides
Gram ⁺	Gram-positive
Gram ⁻	Gram-negative
GRAS	Generally recognized as safe
HGT	Horizontal gene transfer(s)
IF	Induction factor
LAB	Lactic acid bacteria
me ⁰	Unmethylated
me ⁺	Methylated
me ^{+/-0}	Heteroduplex of methylated/unmethylated DNA

ODE	Ordinary differential equations
QS	Quorum-sensing
RLU	Relative light units
RM	Restriction-Modification system(s)
SIR	Susceptible-Infected-Recovered
ssDNA	Single-stranded DNA
Smu	<i>Streptococcus mutans</i>
Spn	<i>Streptococcus pneumoniae</i>
Sth	<i>Streptococcus thermophilus</i>
TCS	Two-component system
USS	Uptake signal sequence
XIP	ComX-inducing peptide

Part I

Introduction

GENERAL CONTEXT

Competence for natural DNA transformation is a physiological state in which bacteria naturally enter, when the conditions are favorable, in order to integrate extracellular DNA into their genome. This lateral gene exchange mechanism (Skippington and Ragan, 2011) is widely and extensively used among bacteria (Johnsborg *et al.*, 2007). It allows the recipient cell to rapidly acquire genetic information from its neighbors, giving it a huge evolutionary advantage by promoting its genetic plasticity and hence, its capacity to adapt to varying and hostile environmental conditions. For instance, some pathogenic species, like *Streptococcus pneumoniae*, have developed extremely efficient skills for transformation (Alloing *et al.*, 1998; Havarstein *et al.*, 1995; Havarstein *et al.*, 1996), which allows them to quickly switch their phenotype (Havarstein *et al.*, 1997) and to acquire new traits, depending on the stage of host-infection, in order to escape the immune system. So, even if natural transformation is an interesting mechanism in itself to study at the fundamental level, the scope of lateral gene transfer mechanisms is far beyond that, since it also has major health implications.

Some bacterial species have been studied more extensively for their competence for natural transformation. Among Gram-positive bacteria, *Bacillus subtilis*, *Streptococcus mutans* and *S. pneumoniae* are called “model species”, since many researches were performed on these species. But as we will see, many recent studies were conducted on the competence ability of *Streptococcus thermophilus*, our model species for this thesis. Indeed, the ComRS signaling system, which controls competence in most streptococci, was recently discovered in this dairy species (Fontaine *et al.*, 2010a). It seemed thus legitimate to choose *S. thermophilus* as an alternative working model to *S. pneumoniae* to investigate the regulatory network controlling competence in streptococci.

Since the discovery of the ComRS system by Fontaine *et al.*, a series of experimental works have been pursued by many teams in order to characterize the functioning of this system (for a recent review, see (Fontaine *et al.*, 2015)). These investigations have allowed to shed light on the successive steps of competence development, taking place at the molecular level.

The aim of this thesis was to study the ComRS regulatory system of competence development in *S. thermophilus*, with an original strategy consisting in the combination of molecular biology tools with an *in silico* modeling approach. Besides alleviating the time-consuming burden of experimental investigations by efficiently selecting them (experimental design), modeling a system is sometimes the only way to extract useful information from various and complex experimental data.

In this introduction, the first chapter is an overview of natural transformation occurring in different bacteria. It is followed by a more detailed chapter about the regulation of competence in the different model species. Finally, the introduction will end with approaches of modelling biological systems, emphasizing on the utility and challenges of the mathematical formalization.

1. Competence for natural DNA transformation

A. Discovery

As soon as 1928, the British bacteriologist Frederick Griffith observed and demonstrated for the first time the occurrence and effect of the “transformation principle” (Griffith, 1928). His focus on epidemiology of bacterial pneumonia, which was an important cause of death after World War I, led him to perform experimentations on the pathogenic model species *S. pneumoniae* (Spn). The German bacteriologist Fred Neufeld had studied and characterized three pneumococcal types (types I, II and III). At that time, it was thought that these pneumococcal types were fixed and could not be changed. But F. Griffith proved the contrary. In his striking experiment (see **Figure 1**), F. Griffith used two strains of Spn: one of type II, non-virulent, with a rough phenotype (R strain, for “rough”), and one of type III, virulent, with a smooth phenotype (S strain, for “smooth”) due to its modified cell wall. The strain III-S is able to kill his host because its cell wall is covered by a polysaccharide capsule, allowing the pathogen to escape the host immune system. In contrast, the strain II-R is bereft of such “camouflage clothes” and as a consequence, it cannot kill his host.

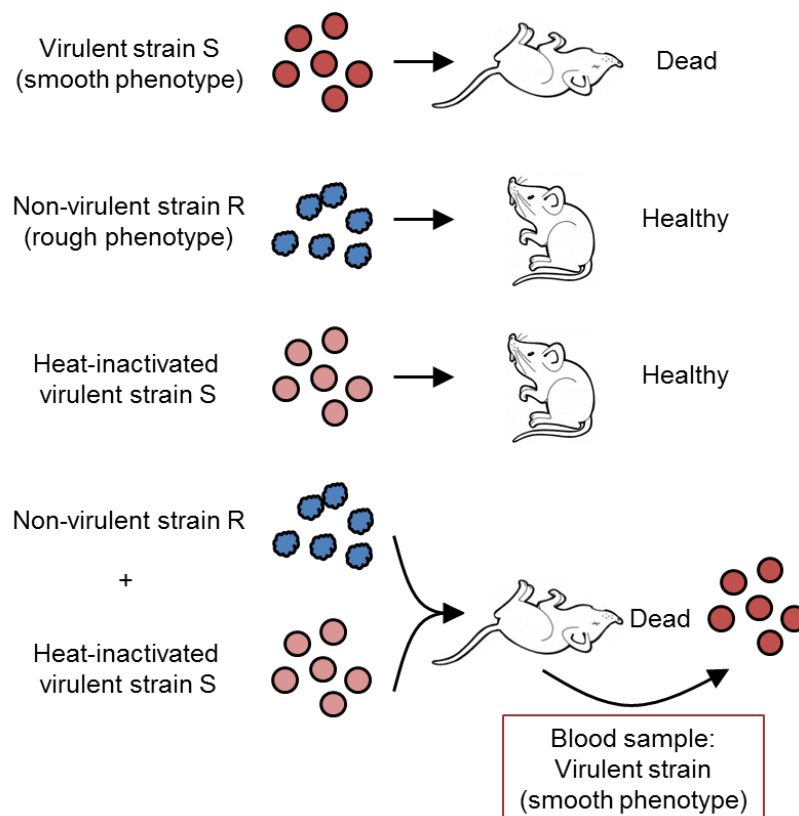


Figure 1. Griffith's experiment shows that the transfer of a substance from one virulent Spn strain to a non-virulent strain transforms the recipient cells into virulent ones. The explanation is that the DNA of the S-cells, containing the information for the virulence trait, has been transferred and acquired by the non-virulent strain, which has become able to express the virulence factors causing the death of the injected mice (figure adapted from (Boutry, 2013)).

The experiment consisted in injecting 4 different inocula to mice:

- 1) A living virulent strain S: the mice die.
- 2) A living non-virulent strain R: the mice live.
- 3) A dead (heat-inactivated) virulent strain S: the mice live.
- 4) A living non-virulent strain R mixed with a dead virulent strain S: the mice die.

The fourth inoculation causes the mice to die, meaning that the non-virulent strain has become virulent because of the contact with the dead virulent strain. Hence, it suggests that the non-virulent strain has acquired the virulence trait from the dead virulent cells. This is what was called “natural transformation”: the spontaneous phenotypic transformation of one strain due to the contact with a substance coming from a dead (lysed) strain. At this time, DNA was not discovered yet, so people did not know what kind of chemical substance had been transferred between the two strains. It was in 1944 that Avery, McLeod and McCarty identified DNA (and not proteins as it was commonly believed) as the carrier of the information coding for virulence factors (Avery *et al.*, 1944).

Since then, a tremendous amount of knowledge has been learned about natural transformation. We are now able to describe what is happening (at least partially) at the molecular level during the natural transformation process, which is also called “**competence for natural transformation**”. Indeed, all bacterial cells are not able to enter into the competence state, so the “lucky” ones that have this capacity are called competent cells.

It is important to note that **all competent cells do not necessarily transform DNA**: only a **fraction of the competent population** is actually transformed after one competence event. Indeed, a bacteria could, for instance, undergo competence development when no sufficient amount of DNA is present in the medium. This could be called a « competence event », in opposite to a « transformation event », in which DNA integration does actually occur.

The ratio of transformed cells varies a lot among species, ranging from 10^{-8} in some strains of *Sth*, up to ~100% in *Spn*.

B. Definition

Competence for natural transformation is usually defined as “*a transient physiological state enabling some bacteria to take up, internalize and stably integrate exogenous DNA in their genome*” (Johnsborg *et al.*, 2007) (see **Figure 2**). The incoming genetic information is thus inheritably transmitted to the next generations. To perform transformation, bacteria have developed specific machineries in order to capture, import and insert the piece of DNA into their genome by a double recombination event. Most of bacteria are only temporarily competent, entering this particular physiological state when certain conditions are met, or when a specific signal is

received (as shown in **Figure 3**). Along with transduction and conjugation, competence is one of the main Horizontal Gene Transfer (HGT) mechanisms (Skipington and Ragan, 2011), which define genetic exchange communities. Here is a short description of the two other main HGT mechanisms (Thomas and Nielsen, 2005):

- i) During **conjugation**, in Gram-negative bacteria, the donor cell expresses a pilus, which is a hair-like appendage composed of oligomeric pilin proteins, and encoded by a conjugative plasmid. This pilus recruits the recipient cell in order to proceed to conjugation (see **Figure 4**). Gram-positive bacteria use a type IV secretion system (T4SS), quite similar to their Gram-negative counterparts. Indeed, the first steps of conjugation are almost identical. However, due to the different organization of their cell envelopes, the subsequent steps (including the actual DNA translocation) are different between Gram-positive (Grohmann *et al.*, 2003) and Gram-negative (Ilangoan *et al.*, 2015) bacteria.
- ii) **Transfection** is the mechanism by which DNA is transferred in cells by phages (see **Figure 5**).

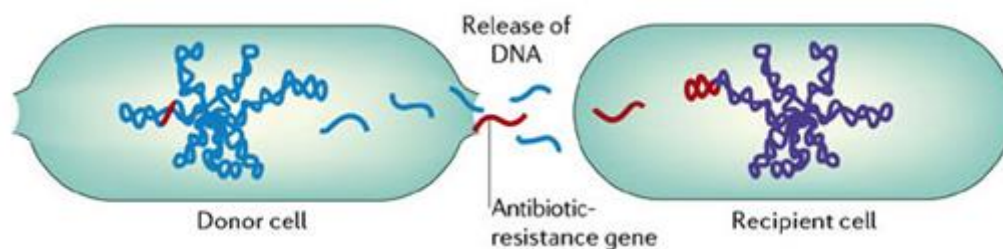


Figure 2. DNA transfer by natural transformation. The competent, recipient cell takes up, internalizes, and integrates a piece of exogenous double-stranded DNA into its genome. The incoming DNA can contain for instance the information coding for an antibiotic-resistance gene (figure adapted from (Furuya and Lowy, 2006)).

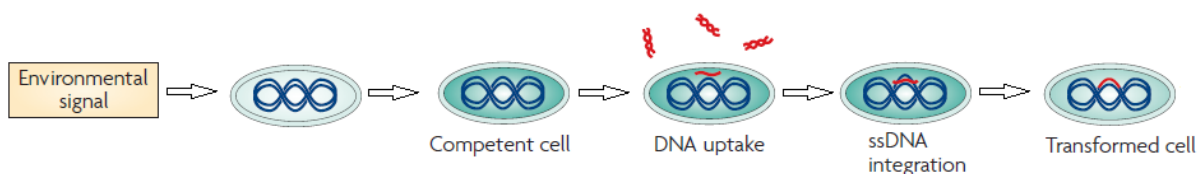


Figure 3. Competence activation upon perception of a specific signal. This signal can be for instance a short peptide, acting like a pheromone. The cells can then enter into the competence state, enabling them to stably integrate exogenous DNA into their genome: they are said to be “transformed” (figure from (Claverys and Havarstein, 2007)).

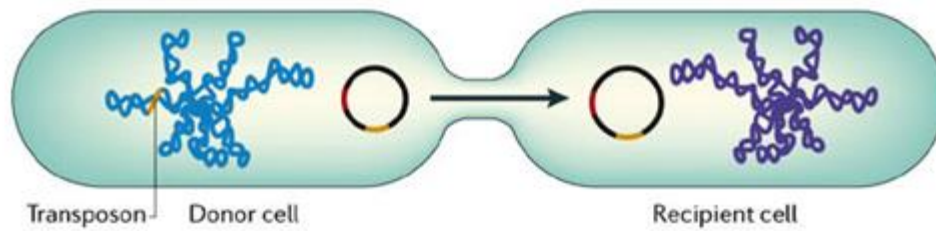


Figure 4. DNA transfer by conjugation. The donor cell expresses a conjugation pilus, which allows it to connect with the recipient cell (figure adapted from (Furuya and Lowy, 2006)).

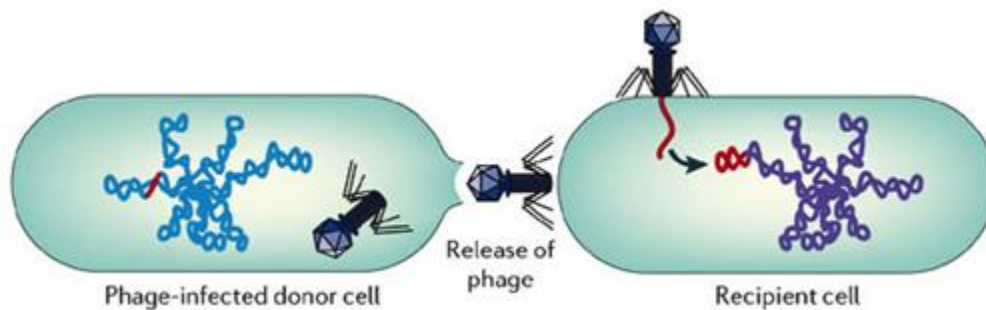


Figure 5. DNA transfer by transfection. An infected cell becomes a donor by releasing phages all around. These infectious agents might carry DNA from the donor cell by accidentally packaging it. They will then enter in the recipient cell, possibly delivering the donor cell's DNA (figure adapted from (Furuya and Lowy, 2006)).

In general, HGT are initiated by the donor cell. For instance, transfection is initiated by the infected donor cells sending phages around (which transfer the DNA), while the recipient cells are more passive, just getting infected. But exceptions occur: in *Spn* for example, a conjugation event might be induced by a recipient cell releasing pheromones in order to prompt the exchange.

Natural transformation has the particularity to be systematically initiated by the recipient cells, which actively internalize exogenous DNA coming from its environment. Moreover, natural transformation does not require any direct interaction between a donor and a receiver cell: naked DNA is directly captured from the extracellular medium. Besides, in opposite to conjugation and transduction, natural transformation is the sole HGT mechanism that is generally programmed by native chromosomal genes (Johnston *et al.*, 2014).

C. Dedicated machineries

In order to acquire exogenous DNA, competent bacteria temporarily express a series of specific machineries, as illustrated in **Figure 6**. This DNA uptake apparatus is sometimes called the “transformasome”.

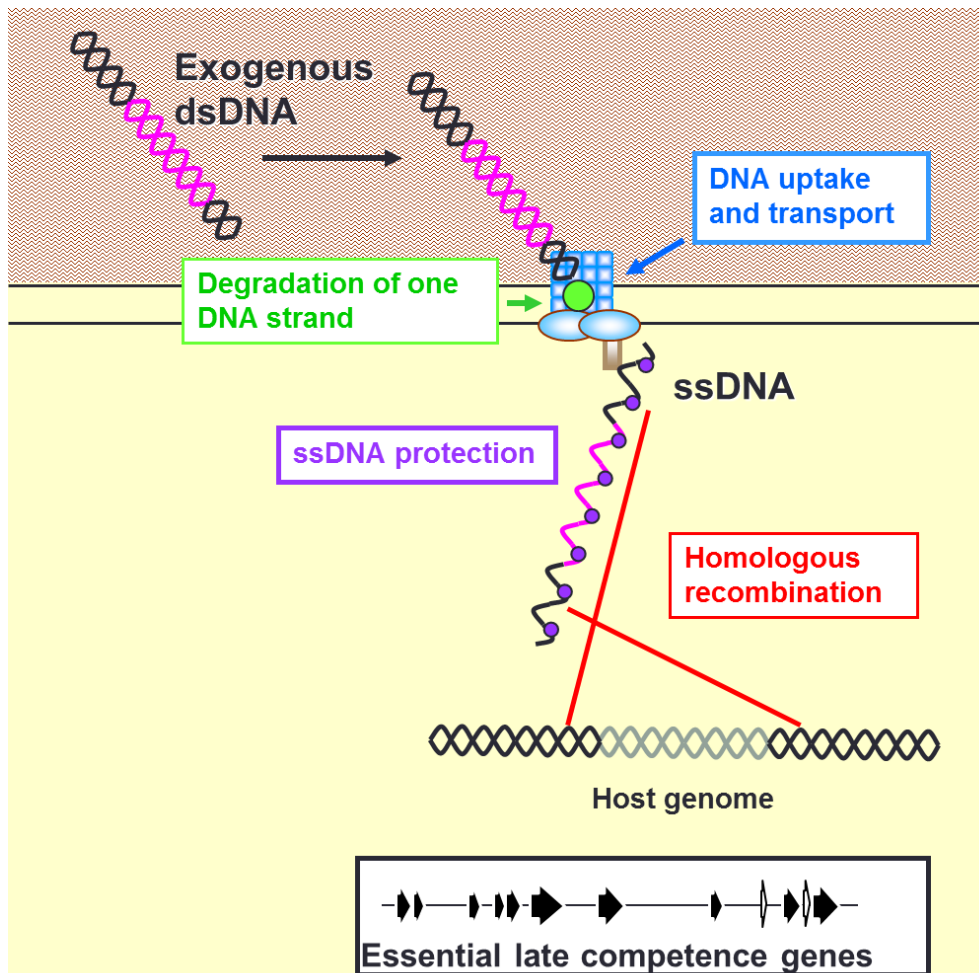


Figure 6. Major steps in natural transformation. The successive steps of natural transformation are dsDNA capture, DNA internalization with degradation of one strand, ssDNA protection and integration by homologous recombination into the recipient chromosome. The proteins required for all these machineries are encoded by the “late competence genes”.

The transformation proteins are highly conserved among Gram-positive and Gram-negative bacteria: the essential part of the transformasome is encoded by 14 necessary genes, but each species possesses additional accessory proteins (Peterson *et al.*, 2004; Claverys *et al.*, 2009). The main difference between Gram-positive and Gram-negative bacteria is that the incoming DNA has to cross one additional membrane in the latter group, so there are additional components for this

step. As explained hereunder, these machineries aim to capture the extracellular double-stranded DNA (dsDNA), to internalize it, to degrade one strand, to protect the resulting single-stranded DNA (ssDNA) inside the cytoplasm by single-strand binding proteins (Ssb), and to recruit the recombination factors in order to stably integrate it by double homologous recombination (Johnsborg *et al.*, 2007).

I. Binding of the environmental dsDNA

The incoming, often dsDNA binds to the recipient cell surface in a non-covalently manner, through interaction with an element of the dedicated uptake apparatus (Claverys *et al.*, 2009; Craig and Li, 2008), called “pilus” in Gram-negative bacteria, and “pseudo-pilus” in Gram-positive bacteria. The (pseudo-)pilus then brings the foreign DNA to the translocation machinery, which internalizes it.

In Gram-negative bacteria, the uptake apparatus involves a type IV pilus structure (Chen and Dubnau, 2004; Chen *et al.*, 2005; Johnsborg *et al.*, 2007). Pili are long filamentous complexes (Craig and Li, 2008), visible under the electron microscope. Besides natural transformation, pili are also used for adhesion and motility purposes. This structure is a polymerization of pilin proteins whose depolymerization triggers the retraction of the pilus. The pilus recruits the DNA by direct interaction (Craig and Li, 2008), non-covalently, without any sequence specificity, except in some species like *Haemophilus influenzae* and *Neisseria gonorrhoeae*, which have specific USS (uptake signal sequences) and DUS (DNA uptake signal sequences) recognition sequences (Seitz and Blokesch, 2013). It then retracts itself in order to convey the DNA towards the pore at the cell surface, where it will be internalized through the translocation machinery (Burton and Dubnau, 2010).

In Gram-positive bacteria, the uptake apparatus is called a pseudo-pilus, because until recently it had never really been observed to protrude from the cell surface. It is composed of orthologous proteins to those of Gram-negative bacteria (Chen and Dubnau, 2004). In *Bacillus subtilis* (Bsu), the pseudo-pilus structure is made of pilin-like proteins (Chen *et al.*, 2006), and was shown to be located at the cell poles (Hahn *et al.*, 2005). In 2013, the team of Laurenceau discovered a new appendage at the cell surface of Spn. They showed that this appendage is the primary DNA receptor of Spn, similar to a type IV-like pilus, composed of ComGC pilins, directly interacting with DNA, and essential for competence (Laurenceau *et al.*, 2013). A subsequent study proposed a totally different structure (short, plaited polymers, see (Balaban *et al.*, 2014)) as transformation appendages. However, Laurenceau *et al.* brought experimental evidence that these plaited structures, which they called spiroosomes, are not related to natural transformation: they are macromolecular complexes of the acetaldehyde-alcohol dehydrogenase enzyme AdhE (Laurenceau *et al.*, 2015).

II. Processing of the recruited dsDNA

Concomitantly to its importation, the environmental dsDNA is processed in two ways: it is first fragmented into smaller pieces by specific endonucleases (EndA in *Spn*, see (Berge *et al.*, 2013;Dubnau, 1999;Lacks *et al.*, 1974)), and then one of its two strands is degraded. These two steps yield to relatively small pieces of ssDNA. In *Bsu*, the fragmentation step yields linear fragments up to 20 kb in size (Provvedi *et al.*, 2001). It is performed by the surface endonuclease NucA (Dubnau, 1999), a membrane-localized protein which introduces double-strand breaks. After its fragmentation, the degradation of one strand releases phosphorylated fragments into the medium. The transforming strand is then internalized in the cytoplasm and protected against DNAses.

III. Internalization of the ssDNA pieces

The polarity of the entry of ssDNA is 3' -> 5'. In Gram-negative bacteria, which contain an extra outer membrane, dsDNA has first to cross this additional outer membrane before its conversion into ssDNA (Collins *et al.*, 2005;Frye *et al.*, 2006).

a. Crossing the outer membrane

The dsDNA is actively transported through the outer membrane by the PilQ channel protein (see **Figure 7**), which is a homododecameric structure, with an affinity for DNA, forming a pore (Assalkhou *et al.*, 2007;Collins *et al.*, 2001;Collins *et al.*, 2003;Collins *et al.*, 2004;Collins *et al.*, 2005;Frye *et al.*, 2006).

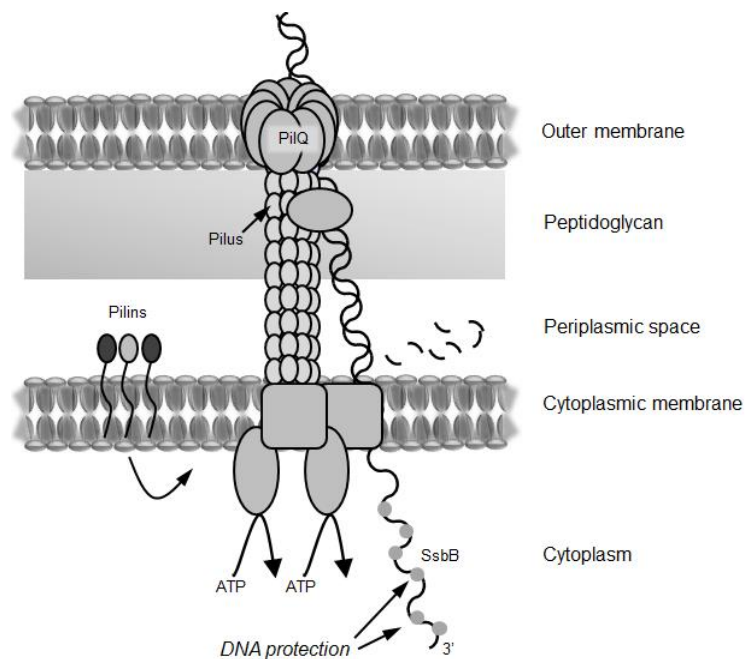


Figure 7. Crossing the outer membrane in Gram-negative bacteria. The dsDNA is actively transported through the outer membrane by the PilQ channel protein, which is a homododecameric structure, with an affinity for DNA, forming a pore through the outer membrane (figure adapted from (Boutry, 2013)).

Helicobacter pylori presents the particularity to use a type IV secretion system, which conveys dsDNA through the outer membrane via the ComB protein (Fischer *et al.*, 2010).

b. Crossing the cytoplasmic membrane

The DNA is brought by ComEA, a membrane-anchored DNA binding protein (Seitz *et al.*, 2014), to ComEC, a polytopic membrane protein that forms an aqueous channel crossing the cytoplasmic membrane (Draskovic and Dubnau, 2005; Dubnau, 1999) (see **Figure 8**). The energy needed for the translocation step of ssDNA is given by the ATP-hydrolysis of ComFA (Chen and Dubnau, 2004).

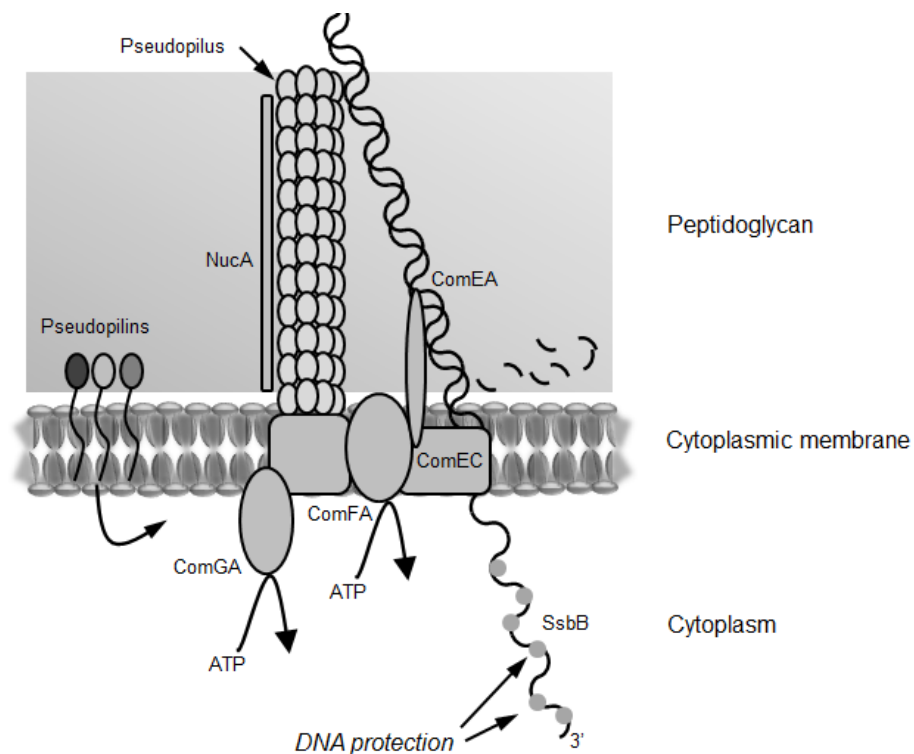


Figure 8. Crossing the cytoplasmic membrane. The DNA is brought by ComEA, a membrane-anchored DNA binding protein to ComEC, a polytopic membrane protein that forms an aqueous channel crossing the cytoplasmic membrane. The energy needed for the translocation step of ssDNA is given by the ATP-hydrolysis of ComFA (figure adapted from (Boutry, 2013)).

IV. Protection of the internalized ssDNA pieces

In *Spn*, once in the cytoplasm, the ssDNA is protected against endogenous DNAses by Single-strand binding (Ssb) proteins: SsbB, RecA, CoiA and DprA (Attaiech *et al.*, 2011; Berge *et al.*, 2003; Morrison *et al.*, 2007). It forms a complex which is thought to yield pools of integrable ssDNA which could be used to execute successive DNA transformation events. Moreover, DprA also triggers the loading of RecA on the ssDNA (Mortier-Barriere *et al.*, 2007).

V. Recombination of the recipient ssDNA

Conveyed by DprA, the 3' extremity of the ssDNA initiates the RecA-mediated recombination with the resident chromosome, yielding heritable integration of the recipient information (Mortier-Barriere *et al.*, 2007). It first results in the formation of an hetero-duplex consisting in one exogenous strand and one endogenous strand, stabilized by non-covalent bonds, before being resolved either by mismatch repair or by replication. The efficiency of transformation is directly proportional to the length of the homology regions contained in the DNA (Kung *et al.*, 2013).

D. Origin of the exogenous DNA

The extracellular DNA, which is used by competent cells for their natural transformation, is generally double-stranded, though it can be single-stranded, but then the transformation efficiency is highly reduced (Miao and Guild, 1970). It can be chromosomal or plasmidic DNA, but plasmidic DNA transformation is less efficient (Nielsen *et al.*, 1997).

The DNA comes from neighboring dead cells that are lysed. There are two possibilities:

- i) These cells can be dead because of **independent reasons**. After having lysed, their DNA is released in the medium. It has been shown that free DNA is present in large amounts in biofilms, to which it contributes by forming the biofilm matrix (Perry *et al.*, 2009). The stability of DNA in biofilms is quite high, so free DNA is a common compound for bacteria to find.
- ii) The donor cells can be **actively killed** by competent cells, through a lysis which is co-programmed with competence development (Claverys *et al.*, 2007). Indeed, the production of bacteriocins and lysins is often triggered along with natural transformation, in order to ensure that the recipient cells will have DNA to take up. This is called “**fratricide**” or “**sobrinicide**”, because the bacteriocins/lysins in general target bacteria that are respectively of the same population (*i.e.* fratricide) or phylogenetically close (*i.e.* sobrinicide) from the killers (Claverys *et al.*, 2007).

The stable integration of the transforming ssDNA will first rely on the efficiency of the recombination and repair machineries of the host. They will limit the incorporation in the host chromosome of mutations present on the transforming DNA.

Moreover, the exogenous DNA has to be “compatible” with the recipient bacterium. In general, it means that the bacteria must be phylogenetically close enough. But some bacteria have developed strategies in order to select their incoming DNA:

- i) the DUS and USS selective sequences (Seitz and Blokesch, 2013), which are short sequences spread in the genome, and represent a necessary condition for the DNA to be integrated by natural transformation in a few bacterial species belonging to Gram-negative bacteria (Findlay and Redfield, 2009; Lorenz and Wackernagel, 1994) ;
- ii) fratricide: programmed lysis towards siblings (Berg *et al.*, 2012; Claverys *et al.*, 2007) ;

Besides the selective mechanisms, bacteria have developed systems in order to limit the DNA transfers:

- iii) restriction/modification systems (RM): some bacterial species have restriction endonucleases that cleave DNA harboring specific short (often palindromic) sequences contained in foreign DNA internalized by an infectious agent like bacteriophages. These systems generally cut dsDNA, but not ssDNA (Seitz and Blokesch, 2013). Cutting the exogenous dsDNA could protect the cells by inactivating the infection process. Methyl groups are added on the bacterial genomic restriction sequences in order to prevent the cells from degrading their own DNA (Jeltsch and Pingoud, 1996; Johnston *et al.*, 2013b; Johnston *et al.*, 2013c). Pneumococci have in almost equal proportions either one of two RM systems: DpnI or DpnII (see **Figure 9**).

- **DpnI** is an atypical RM system, since it cuts methylated (me^+) dsDNA. The host genome is then unmethylated in order to be protected from its own RM system.
- **DpnII** is more classical, since its endonucleases cut unmethylated (me^0) dsDNA. As a consequence, the DpnII-containing strains have a methylase (DpnM) to protect their own genomic DNA. Moreover, these DpnII-containing strains have a special methylase called DpnA, which protects foreign ssDNA by methylation (Johnston *et al.*, 2013a). Indeed, if a piece of exogenous ssDNA is not methylated, then the $me^{+/0}$ heteroduplex (resulting from recombination of foreign and origin

DNA) will lead to a me^0 dsDNA after resolution by replication, which will be cut by DpnII, annihilating the effect of transformation. The role of DpnA is thus likely to allow transformation in DpnII-containing pneumococci, as corroborated by the facts that:

- i) DpnA is under control of a competence-induced promoter, and
- ii) DpnA has been shown to be essential for efficient competence development (Johnston *et al.*, 2013a). So, contrary to the classical view in which transformation and RM systems are antagonistic, the DpnII RM system (harboring the DpnA ssDNA protection) can be compatible with competence (see **Figure 9**).

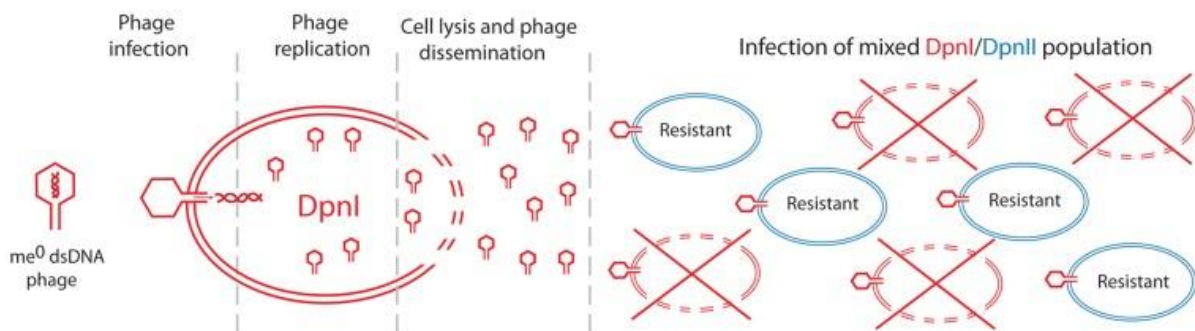


Figure 9. Comparison of DpnI and DpnII RM systems. When a pneumococcal population comprising both DpnI and DpnII cells gets infected by me^0 dsDNA phages (red hexagons), the DpnI isolates, which can only cleave methylated DNA, undergoes phage replication (releasing new me^0 phages) and lysis, while the DpnII isolates (blue ovals) survive thanks to its RM system (figure from (Johnston *et al.*, 2013c)).

- iv) **acquired immunity:** the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system implements a kind of prokaryotic immune system in ~40% of bacterial species, protecting them from parasitic DNA (Barrangou *et al.*, 2007; Barrangou and Horvath, 2012; Bikard *et al.*, 2012; Garneau *et al.*, 2010; Horvath *et al.*, 2009; Horvath and Barrangou, 2010; Makarova *et al.*, 2011; Makarova *et al.*, 2015). When infected by foreign DNA (phages, plasmids or naked DNA), the cell keeps a memory of the infection by integrating in its genome short sequences of the infecting DNA. If the cell gets reinfected by the same agent, the crRNA containing the “identity card” of the intruder will form a hybrid (DNA-RNA) complex with the foreign DNA, before its degradation by specific proteins (like Cas9). The presence of CRISPR in competence strains prevents the integration of exogenous DNA which has already entered the cell, impeding transformation by “CRISPR-listed” DNA. The co-existence inside a strain of competence

and CRISPR functionalities could confer a fitness advantage by preventing useless redundant transformations events. This hypothesis is supported by the fact that evolutionary losses of competence are often followed by CRISPR degeneration (Jorth and Whiteley, 2012).

E. Functions

Competence for natural transformation is an energetically expensive process for many reasons.

First, the entry in the competence state is very costly for the cell from the energetical point of view: the cell has to express a series of specific proteins, which represent an important metabolic burden, and DNA uptake by itself is consuming ATP.

Second, it has been shown that competence development inhibits the cellular growth rate (Gardan *et al.*, 2013; Haijema *et al.*, 2001; Mirouze *et al.*, 2015), hence reducing the fitness of the global population. In *Bsu* for instance, the ComGA protein, which is involved in pseudopilus assembly, interacts with MreB in order to prevent cell elongation (Mirouze *et al.*, 2015), and with RelA in order to slow down the cellular growth by impeding replisome assembly and decreasing the rRNA synthesis rate (Hahn *et al.*, 2015; Haijema *et al.*, 2001). This decrease in cellular growth can also be seen as a fitness cost for the competent cells.

Third, natural transformation represents a risk for the transforming cells: for instance, they could integrate a damaged piece of DNA, or the piece of DNA could be inserted in a place where it inactivates an essential gene.

And fourth, competence development is not always adequate: it is more useful in certain environmental conditions (depending on many external signals), when free DNA is very likely to be present in the vicinity, or under a certain level of stress (Prudhomme *et al.*, 2006).

If there was no fitness gain, bacteria would certainly have lost the competence genes. But the advantages of competence still raise questions among the experts. Here, we present three possible functions of competence, which can explain its utility (Johnsborg *et al.*, 2007). Of course, it can be a combination of these three advantages and vary from species to species, since initiation of competence requires different environmental conditions from one species to another.

I. Genetic plasticity

According to the “**sex hypothesis**” (Redfield, 2001; Szollosi *et al.*, 2006), the advantage of natural transformation is the genetic diversity which is conferred to the recipient cells, allowing them to acquire new traits (like antibiotic resistances, for instance) during transformation. Except the species *H. influenzae* and *N. gonorrhoeae*, which use specific signal sequences to select their recipient DNA,

bacteria can in theory take up DNA from any species. According to their environment, it allows them to access the set of all genomes that are present around them, making possible the exchange of a tremendous amount of information between species that share the same ecological niche. This is without any doubt a huge evolutionary advantage for the competent cells. Still, other accessory advantages cannot be definitively ruled out.

II. Nutrition

It had been shown that *Escherichia coli* can grow with DNA as sole carbon source (Finkel and Kolter, 2001). But the hypothesis of competence for nutrition is far from being accepted for the model species *Spn*. Indeed, two main observations invalidate this hypothesis. First, the fact that the degraded strand ends up in the external medium, hence being wasted outside the cell is not in favor of the hypothesis. If nutrition was the main purpose of transformation, this would not occur. Second, the fact that the internalized strand is protected in the “eclipse complex” by a series of Ssb proteins (among which is RecA, which plays an important role in recombination) is not coherent with a nutritional purpose (Finkel and Kolter, 2001; Redfield, 2001; Dubnau, 1999).

III. Repair of DNA

Another hypothesis could be that transformation is used by bacteria in order to repair their damaged DNA if needed (Hoelzer and Michod, 1991; Michod *et al.*, 1988; Prudhomme *et al.*, 2006). Indeed, DNA damages (single or double-strand breaks, punctual mutations, or hazardous recombinations) occur often in nature, resulting from various endogenous (reactive oxygen species from the cell's metabolism for instance) or exogenous (UV light, ionizing radiations, chemicals, replication errors, and so forth) processes. Competence could then offer a way to repair damages by using the information carried on homologous DNA coming from phylogenetically close dead cells (Prudhomme *et al.*, 2006).

F. Prevalence

While natural transformation has been extensively observed in many (Gram-positive and Gram-negative) bacterial species, it is not clear whether Archaea can be competent for natural transformation. In 1987, Bertani and Baresi reported natural transformation events in the species *Methanococcus voltae* (Bertani and Baresi, 1987). It is true that HGT events have been reported in some archaeal species (Patel *et al.*, 1994; Worrell *et al.*, 1988), but to date there is no proof that these gene transfers occurred by natural transformation.

Moreover, we know that in bacteria, natural transformation requires the presence of a pilus (in Gram-negative bacteria) or of a pseudo-pilus (in Gram-positive bacteria) (Chen *et al.*, 2006). Archaea do not have such (pseudo)-pilus, but have a flagellum that is more closely related to the bacterial type IV pilus than to the bacterial flagellum. However, the archaeal flagellum is a rotating device, in opposite to the

bacterial pilus, which is a twitching device. Twitching motility and transformability have been shown to be closely related (Wolfgang *et al.*, 1998). But it is somehow difficult to imagine that the archaeal flagellum could be involved in DNA uptake during transformation. So, to the best of our knowledge, we do not actually know whether Archaea are able of competence development.

In contrary, natural transformation has been observed in more than 60 bacterial species (Johnsborg *et al.*, 2007), and is very well documented for several model species. Of course, species for which transformation has not been reported yet are not definitively non-competent. Since competence is triggered by precise environmental conditions (which differ from species to species), it is likely that a lot of species are competent under conditions that have not been simulated in the laboratories. This census is obviously a poor underestimation of the actual prevalence.

A few examples of model competent species described in Gram-negative and Gram-positive bacteria are reported here below.

I. Gram-negative bacteria

a. Proteobacteria

Competent species have been reported in the alpha, beta, epsilon and gamma subdivisions of the proteobacteria. In the gamma-proteobacteria, the pathogenic species *H. influenzae* (Macfadyen, 2000) and *Vibrio cholerae* (Lo and Blokesch, 2012;Lo and Blokesch, 2013) are the best-described model species, while in the beta and epsilon groups, *N. gonorrhoeae* (Aas *et al.*, 2002) and *H. pylori* (Dorer *et al.*, 2013;Humbert *et al.*, 2011;Moore *et al.*, 2014) have been largely investigated.

b. Deinococcus-Thermus

The model species *Thermus thermophilus* has been shown to be competent for natural transformation (Koyama *et al.*, 1986), though it was mostly studied for its heat-stable proteins rather than its transformability.

c. Cyanobacteria

Several species from the genera *Synechococcus* and *Thermosynechococcus* are competent, but there is no study really focusing on this phenomenon (Johnsborg *et al.*, 2007).

II. Gram-positive bacteria

a. Actinobacteria

Very little is known about the prevalence of natural transformation in this phylum. So far, solely one species, *Mycobacterium smegmatis*, has been shown to be competent (Bhatt *et al.*, 2002; Norgard and Imaeda, 1978).

b. Firmicutes

A lot of species of the phylum Firmicutes are competent. The two well-known model species, Bsu and Spn, belong to this phylum. In the *Streptococcaceae* family, almost all species have homologs of the competence genes (Martin *et al.*, 2006). Besides Spn, for instance, *Streptococcus gordonii* (Lunsford and Roble, 1997), *Streptococcus mutans* (Smu) (Desai *et al.*, 2012; Desai *et al.*, 2012; Merritt *et al.*, 2005), *Streptococcus bovis* (Mercer *et al.*, 1999), *Streptococcus mitis* (Bensing *et al.*, 2001) and *Streptococcus thermophilus* (Sth) (Fontaine *et al.*, 2010a; Gardan *et al.*, 2009; Gardan *et al.*, 2013), *Streptococcus pyogenes* (Marks *et al.*, 2014), *Streptococcus infantarius* (Morrison *et al.*, 2013), *Streptococcus salivarius* (Fontaine *et al.*, 2010a), have been shown to be competent.

G. Our model organism: *S. thermophilus*

As reported above, Sth is a Gram-positive bacterium from the phylum Firmicutes. It displays an ovococcal shape (see **Figure 10**), is non-motile, non-sporulating and facultative anaerobe. It is found in fermented milk products (like cheese and yogurt) and in the human gastro-intestinal tract as a commensal species (Hols *et al.*, 2005). It is a moderate thermophilic bacterium, growing at around 40°C and forming cell chains (see **Figure 10**). It has a unique circular chromosome of approximately 1.8 Mb, and generally no plasmids, except for instance the LMD-9 strain which has two small endogenous plasmids (Makarova *et al.*, 2006).



Figure 10. Morphology of *S. thermophilus*. Micrograph of Sth obtained by scanning electron microscopy (SEM). Sth cells display an ovococcal shape and are forming chains (photo from Lloyd G. Simonson, ASM MicrobeLibrary from NCBI).

Among streptococci, it belongs to the salivarius group, as illustrated in **Figure 11**. Sth is the sole non-pathogenic streptococcal species, and has the GRAS status, meaning that it is Generally Recognized As Safe. In opposite, all its phylogenetically close neighbors are considered as pathogenic. Here is a few:

- i. Spn is the well-known causative agent of pneumonia, but it can also cause a long series of various infections, such as endocarditis, meningitis, bronchitis, rhinitis, conjunctivitis (Mehr and Wood, 2012).
- ii. *S. salivarius* is an early colonizer of the human body (mouth, skin, gut, urogenital tract), and an opportunistic pathogen. Once in the bloodstream, it can cause sepsis (Carley, 1992).
- iii. Smu is commonly found in the oral cavity, and contributes to the dental decay, causing the formation of caries (Kojima et al., 2012).
- iv. *S. pyogenes* is also an important human pathogen, causing various infections like pharyngitis, impetigo and necrotizing fasciitis (Lamagni et al., 2009).

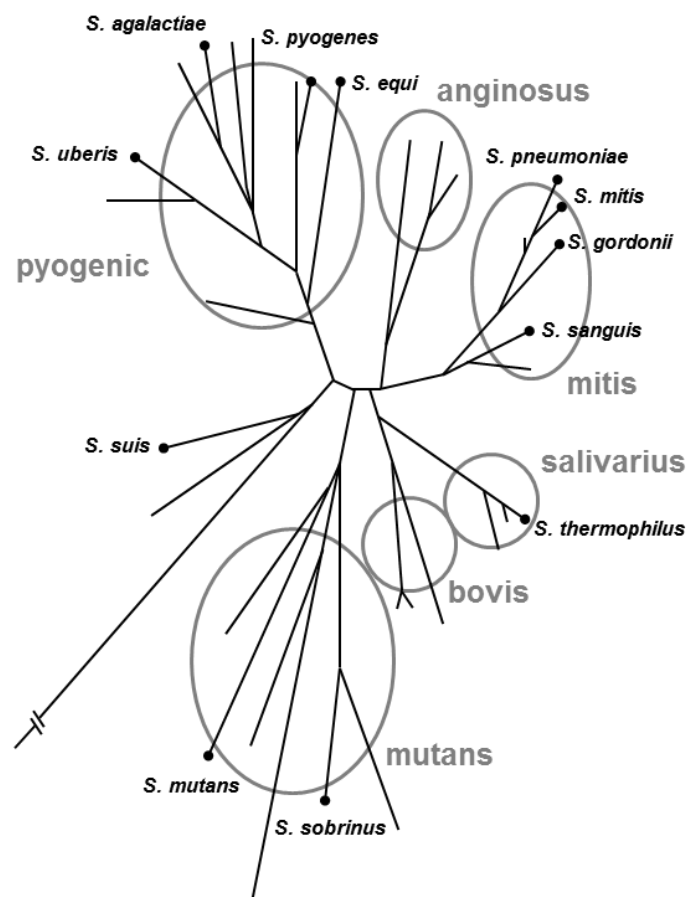


Figure 11. Phylogenetic organization of streptococcal groups. The sequences used to build the tree are those of the 16S rRNA. Our model species, Sth, belongs to the salivarius group (figure adapted from (Tettelin, 2004)).

Sth presents the striking particularity to be the only non-pathogenic species among all streptococci, making it especially interesting and safe. This outstanding particularity can be explained by the genomic adaptation it has undergone in order to adapt to its natural niche (Bolotin *et al.*, 2004; Hols *et al.*, 2005; Pastink *et al.*, 2009): the hypothesis is that Sth, initially coming from the mouth of humans, has adapted to milk as a new natural habitat when humans domesticated bovines and hence started to manipulate milk products, some 10,000 years ago. Once “fallen” in milk, there was no point anymore for Sth in maintaining its virulence genes, which have thus all been eliminated by evolution (by frameshift, truncation, deletion or mutation), rendering it totally inoffensive. The sole leftovers from its previous virulence are a series of pseudogenes (truncated, or mutated), coding for inactivated virulence factors. The recent sequencing of different strains of Sth (LMD-9, LMG18311, JIM 8232, CNRZ1066, ND03, and MN-ZLW-002) revealed that while it has an overall 39% G+C content (low G+C content), it contains a lot of genomic islands with a significantly higher G+C content, indicating that a series of genes were acquired horizontally from neighboring species through HGT events. Natural transformation is likely to have contributed to this large amount of genetic acquisitions (Hols *et al.*, 2005).

Choosing Sth as a model species for competence regulation presents a lot of interests and advantages:

- i. This species is a lactic acid bacterium (LAB) having a huge industrial interest, since it is widely used as a starter to produce various cheeses (e.g. Emmental, Comté, parmesan, camembert, Brie, mozzarella) and yogurt. It is co-cultured along with *Lactococcus lactis* and *Lactobacillus helveticus* in order to produce Italian-type and Swiss-type cheeses, respectively. The name “yogurt” can only be given to products fermented by the two species Sth and *Lactobacillus delbrueckii* subsp. *bulgaricus*, which grow synergistically together. To produce cheese and yogurt, Sth converts lactose, the sugar of milk and its main carbon source, in lactic acid. It also secretes aromas, giving a particular flavor to the fermented dairy products, and exopolysaccharides (EPS), substances that give to reduced-fat cheeses the same texture as full-fat cheeses. Sth is thus of an extremely important interest for food industrials, who constantly try to improve their starter strains in order to obtain new flavors and textures (and better profitability) (Delcour *et al.*, 2000).
- ii. Sth also has a health interest, being used as a probiotic and helping to restore a healthy intestinal flora (Mahmood *et al.*, 2013).
- iii. Since it is GRAS, Sth is not hazardous to manipulate in the laboratory.
- iv. Moreover, since Sth is phylogenetically close to pathogenic species involved in a lot of diseases, it is interesting from a medical point of view to study this non-pathogenic species in order to get more information about its close virulent neighbors.

- v. *Sth* is naturally competent for transformation, and the ComRS signaling pathway was discovered with this species some years ago, giving it the status of model species for the ComRS system (see section 2.3.1.1.) (Fontaine *et al.*, 2010a).
- vi. This species comprises different strains exhibiting strikingly different behaviors concerning competence: for instance, in certain conditions, the strain LMD-9 can reach a relatively high transformation rate of 10%, meaning that when a population of LMD-9 cells is put in contact of transformable DNA under favorable conditions, up to 10% of the cells effectively integrate the DNA in their genome. This transformation frequency is sometimes noted τ . In contrast, the strain LMG18311 is poorly competent ($\tau = 10^{-6}$, meaning one transformant per 10^6 cells). Having various transformation behaviors inside the same species is very useful in order to understand the determinants of competence development.
- vii. The inducible competence of *Sth* and its efficient homologous recombination process make it very easy to handle genetically, which is really useful for rapidly creating mutant strains (Fontaine *et al.*, 2010b).
- viii. This species grows quite fast (the doubling time is around 30 minutes), which is also useful and experimenter-friendly in the laboratory.
- ix. Last but not least, the whole genome of *Sth* has been sequenced (Hols *et al.*, 2005), allowing bioinformatic predictions and analyses to be performed.

2. Regulation of competence development

A. Observations

For most species, competence for natural transformation is a particular physiological state, occurring only under very specific conditions. Therefore, development of competence is tightly regulated. We start from the three following observations:

- i) First, in many species, competence development is only a **transient** phenomenon, occurring during a brief time-window and then disappearing. Exceptions to this rule are the species *N. gonorrhoeae* and *Neisseria meningitidis*, for which the competence state is constitutive (Johnsborg *et al.*, 2007): proteins needed for transformation are always expressed, without any obvious regulatory mechanism. But in the vast majority of bacterial species, competence development is time-restricted.

The induction timing and duration of the competence window vary a lot between species. In the model species *Bsu*, competence development starts when the cells enter the stationary phase, and lasts for some hours (up to 5 hours) (Berka *et al.*, 2002; Ogura *et al.*, 2002). In opposite, in *Spn*, competence is triggered during the exponential phase, and only lasts for approximately 15 minutes (Alloing *et al.*, 1998; Peterson *et al.*, 2004). In our model species *Sth*, the regulation mechanism acts more like a timing-device, since competence initiation does not depend on the initial cellular growth state, and always begins in the early log phase, 80 to 100 minutes after inoculation of the culture and lasts for ~1 hour (Gardan *et al.*, 2013).

- ii) Second, many species have all the genes needed to perform transformation but have never been observed to develop it. For instance, some strains of *L. lactis* present all the genes responsible for the integration of DNA by transformation (B. David *et al.*, manuscript in preparation), but to date, researchers have never identified the competence-permissive **environmental conditions** (e.g. pH, temperature, stress level, growth medium composition) for that species. Moreover, in our model species *Sth*, spontaneous competence activation seems restricted to some rare strains, although all strains encode the genetic requirements for it (Fontaine *et al.*, 2010a; Gardan *et al.*, 2009), suggesting that competence only turns on when very specific environmental conditions are met.

- iii) Third, it has been observed in several species, that competence development often occurs only in a **fraction of a monoclonal population** (*i.e.* a population in which all the cells have the same genetic background). The occurrence of two stable phenotypes (ON and OFF) among a monoclonal population is called “bistability”. For instance, it has been shown in *Smu* that only a fraction (~10%) of the cells simultaneously turn on competence, resulting in a population heterogeneity (Aspiras *et al.*, 2004; Li *et al.*, 2001; Qi *et al.*, 2005; Son *et al.*, 2012). The same kind of heterogeneity can be observed in monoclonal populations of *S. salivarius* (J. Mignolet, data not published). In contrary, the whole population synchronizes and develops competence in *Spn* cultures (Claverys and Havarstein, 2002). This population heterogeneity is discussed in more details in section 2.5.

The fact that competence development seems to be restricted at three different levels (at the environmental level (*i.e.*, according to environmental conditions), at the subpopulation level, and at the time level) shows that competence is regulated by species-specific mechanisms, that control its initiation and exit (for a recent review, see (Fontaine *et al.*, 2015)). This regulation mechanism can possibly act at three different levels: transcriptional, translational, and post-translational (Fontaine *et al.*, 2015).

B. Regulation of competence by communication peptides

Bacterial cells have the capability to communicate through the exchange of secreted pheromones, which allows the population to **behave in a coordinated way**, a bit like a “super-organism”.

In Gram-positive bacteria, communication pheromones are small oligopeptides that are produced inside the cells, matured and exported in the extracellular medium where they can diffuse¹ (see **Figure 12**). Upon contact with another cell, a peptide can either trigger a response extracellularly by interacting with a membrane receptor (like in the ComCDE system), or it can be reimported inside the cell where it can interact with a specific cytoplasmic receptor (like in the ComRS system). Pheromones generally induce their own expression, resulting in a positive feedback loop. This is why they are sometimes called “auto-inducers”, or AI.

¹ The peptides do not always diffuse freely in the medium. In *Spn* for instance, it has been shown that the peptide mediating competence remains mainly attached to the secreting cell (Prudhomme *et al.*, 2016). However, the diffusion/stationarity (on the cell) of the peptides is most likely not on/off: there exists probably a continuum of diffusion rates, varying among species and individuals, and depending on the physical and biochemical properties of the growth medium.

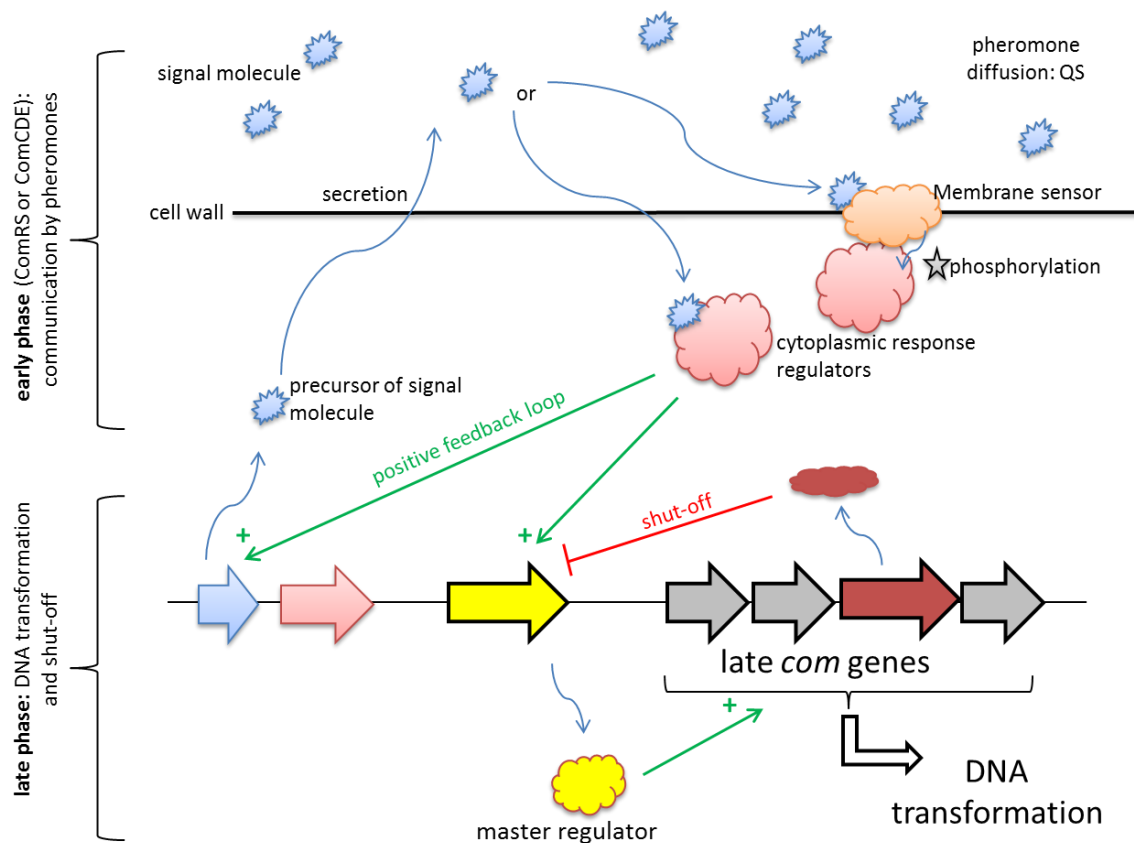


Figure 12. Competence regulation by pheromones in streptococci. A signal molecule is produced, matured and secreted. It then either directly interacts with a membrane-bound sensor resulting in the activation by phosphorylation of a response regulator, or is reimported and interacts directly with the response regulator in the cytoplasm. The activated regulator then in turn activates *comX* transcription, as well as the signal molecule transcription (positive loop of the early phase). Finally, ComX activates the transcription of the late competence genes, resulting in DNA transformation (late phase). The shut-off mechanism then turns the competence down by directly or indirectly inhibiting ComX production or activity. The green arrows represent transcriptional activations, the red arrows represent inhibitions (at any level).

Often, pheromones provide a mean for a bacterial population to assess its cellular density². Indeed, the signal molecules accumulate in the medium as the population size increases. Once a threshold concentration of peptide is reached (and hence a threshold cellular concentration, or “quorum”, is reached), a specific behavioral change is initiated. This cell-density-related communication was called “**quorum-sensing**”, or QS (Fuqua *et al.*, 1994). Thus, QS refers to systems in which

² The ability of a bacterial population to assess its own density can be challenged when complex growth media alter the diffusion properties of the peptides. Some authors have thus proposed alternatives to QS, such as “diffusion-sensing” (Trovato *et al.*, 2014; West *et al.*, 2012) and “efficiency sensing” (Hense *et al.*, 2007).

individuals “sense” their cellular density by the mean of a signal molecule in order to coordinate a multi-cellular level response. For instance, the response to a cell density increase can trigger biofilm formation, sporulation, or competence development (Castillo-Juarez *et al.*, 2015; Rutherford and Bassler, 2012).

QS was first described in *Vibrio fischeri*, a marine bacterium capable of bioluminescence (Lupp *et al.*, 2003; Dunlap, 1999; James *et al.*, 2000; Stevens and Greenberg, 1997). These cells can either live in a planktonic state in nearly all oceans, but also in symbiosis with deep-sea fish like the Hawaiian bobtail squid. A QS system allows the bacteria to activate the transcription of its luminescence-encoding genes (*lux*) when their cellular density reaches a critical threshold. While the hosting fish offers an habitat to the bacteria, the light emitted by the bacteria is beneficial for the fish, for example by suppressing the fish’s shadow during nights with intense moonlight, hence making it more discreet for its predators.

In streptococci and bacilli species, competence regulatory networks are intimately related to the notion of QS. However, they also integrate a series of external, environmental stimuli (e.g. pH, temperature, nitrogen concentration) and of internal, physiological parameters (e.g. cellular growth, metabolic state) in order to trigger natural transformation at the right moment. This allows a population to coordinate in order to maximize the benefits of DNA uptake (Fontaine *et al.*, 2015).

Once the conditions are favorable for competence, the regulatory network turns the competence on by activating the expression of the master regulator of competence: alternative sigma factor σ^X (also named ComX or SigX) in streptococci or ComK in bacilli. This in turn activates among others the expression of a series of genes implicated in the transformation *sensu stricto*. This regulatory cascade leading to the increase in concentration of the master regulator is called the « **early phase** » of competence, while the steps leading to DNA capture, uptake, and integration are said to be part of the « **late phase** » of competence (see **Figure 12**).

In order for the competence state to be transient, the late phase of competence often involves a “**shut-off**” mechanism, which aims at turning the competence down (see **Figure 12**). This mechanism can act upon the early phase, or directly upon ComX.

Although the late genes (coding for the transformation machinery) are quite well conserved among species (see section 1.c.), the early competence genes regulating the initiation of competence are species-specific, as described in the following sections (see **Figure 13** for a simplified view of the main early phase signalling systems).

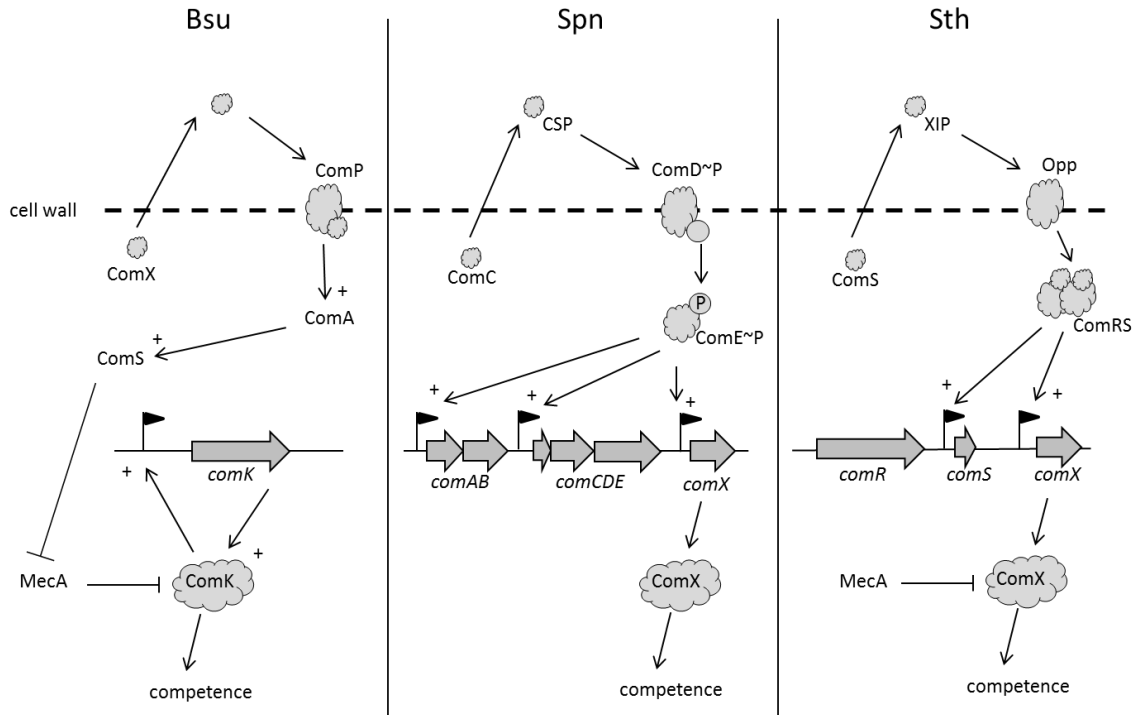


Figure 13. Comparison of the main early phase signalling systems in the model species. Left: in Bsu, the peptide ComX is exported and then interacts with the ComP histidine kinase, which in turn activates ComA. ComA then activates the expression of ComS, which competes with the MecA/ClpC/ClpP machinery for ComK degradation. ComK auto-activates, creating a positive feedback loop (Maamar *et al.*, 2007). Center: In Spn, the accumulation in the medium of CSP make the ComD histidine kinase to auto-phosphorylate. ComD then transfers its phosphate group to ComE, which becomes active and promotes the expression of the whole *comABCDE* operon as well as of *comX*, leading to competence induction (Fontaine *et al.*, 2015). Right: in Sth, the mature peptide XIP is reimported through the Opp channel, and interacts with ComR to form the ComRS complex, which activates the expression of *comS* and of *comX* (Fontaine *et al.*, 2015).

C. Competence regulation in Bsu

I. The ComK transcription factor

In the model species Bsu, the competence state is also named the “K-state” in reference to ComK, its master transcription factor, which was discovered and sequenced more than twenty years ago (Hahn *et al.*, 1996; van Sinderen D. *et al.*, 1994). ComK production is induced by QS and different starvation signals, and in particular by nitrogen limiting conditions (Jarmer *et al.*, 2002), as well as by a high cellular density. It then activates the expression of ~150 genes, among which are the late genes of competence, responsible for the transformation *sensu stricto*, as well as genes involved in DNA repair and in the uptake and utilization of nutritional sources (Berka *et al.*, 2002). Active ComK molecules are tetrameric and bind the so-called

ComK-boxes. Overexpressing *comK* results in a more effective transformation of the cells (Zhang *et al.*, 2014). ComK also activates its own transcription, resulting in an auto-stimulatory loop (see **Figure 14**). Competence in Bsu is regulated by multiple pathways, which are in turn interconnected with regulatory pathways of other differentiation programs (e.g. sporulation), in order for the cells to harmoniously switch from one program to another (Hamoen *et al.*, 2003). Here, we will discuss the different levels of the main regulatory pathways, having the most important impact on ComK expression and abundance.

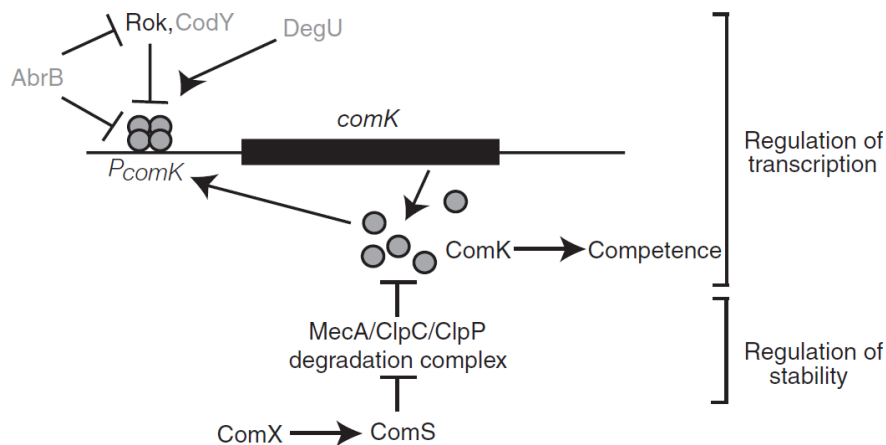


Figure 14. Competence regulation in Bsu. The master regulator of competence is ComK, which auto-regulates its own expression, resulting in bistability. The protease machinery MecA/ClpC/ClpP keeps the level of ComK low during the exponential growth phase. Moreover, the expression of *comK* is under control of repressors, including among others Rok, AbrB and CodY (figure from (Maamar *et al.*, 2007)).

II. Transcriptional control

Although *comK* expression is controlled by several repressors (e.g. Rok and CodY), ComK is produced at a low basal rate, and is hence always present in the cell at very low concentrations, which can vary stochastically. When the amount of ComK in one cell exceeds a certain threshold, the auto-amplification loop turns on, which activates *comK* expression as a positive feedback. This results in ComK accumulation inside the cell, which will enter in the K-state (Leisner *et al.*, 2007). The determinant for this entry into competence is the initial abundance of ComK, which is highly affected by noise (Maamar *et al.*, 2007). The helper protein DegU stimulates the binding of ComK on its own promoter when the ComK concentration is low (Hamoen *et al.*, 2000).

III. Post-transcriptional control

The stability of the *comK* mRNA is affected by the recently discovered protein Kre (for ComK-Repressor), which binds to the *comK* transcript and inhibits its translation (Gamba *et al.*, 2015), which might play a significant role in the control of competence development. The Kre protein does not have any homology with any known protein. It is only present in *comK*-containing genomes and its production is downregulated by ComK.

IV. Post-translational control

ComK is specifically interacting with the adaptor protein MecA, which conveys it to the ClpCP protease complex for its degradation (Turgay *et al.*, 1998).

Another player of the direct regulation of ComK is ComS (Turgay *et al.*, 1997), a small protein encoded in the *srfA* operon, that impedes ComK degradation by directly binding to MecA, hence acting by allosteric competition (D'Souza *et al.*, 1994; Hamoen *et al.*, 1995), and releasing ComK from the proteolytic machinery. The production of ComS is controlled by a series of quorum-sensing systems, depending on the cell density (Auchtung *et al.*, 2006; Hamoen *et al.*, 2003).

D. Competence regulation in streptococci

In streptococci, the entry in the competence state is directly dictated by the amount of σ^X (encoded by *comX/sigX* gene) (Fontaine *et al.*, 2015). This alternative sigma-factor triggers competence by redirecting the RNA polymerase towards the transcription of late competence genes, by directly recruiting σ^X binding to a specific promoter sequence, called the Cin-box, or ComX-box (consensus sequence in *Spn* is TACGAATA), located at position -10 of the late competence genes (Claverys and Havarstein, 2002). The σ^X core regulon varies from species to species, but it includes a set of 14 genes that are essential for natural transformation (Bolotin *et al.*, 2004; Hols *et al.*, 2005). The genome of all streptococci contains either one or two copies of the *comX/sigX* gene (Claverys and Martin, 2003).

The early phase of regulation directly impacts on the level of *comX* expression, so it is said to be a « proximal » and « transcriptional » regulation. As we will discuss below, σ^X is also subjected to indirect « post-translational » regulation, and to « distal » regulation. “Indirect” regulation refers to a regulation which involves some effectors that act on some intermediate elements which in turn impact on σ^X , so that the effect on ComX is indirect.

I. Proximal regulation of ComX abundance

a. Transcriptional control of *comX*

Streptococci species have developed two proximal regulatory systems in order to control *comX* expression: the ComCDE system controls competence in the *mitis* and *anginosus* groups, while the ComRS system is present in all other streptococcal groups (see **Figure 15**). As explained in detail below, these control systems use specific peptide pheromones to co-ordinate *comX* transcription among the population (Johnston *et al.*, 2014).

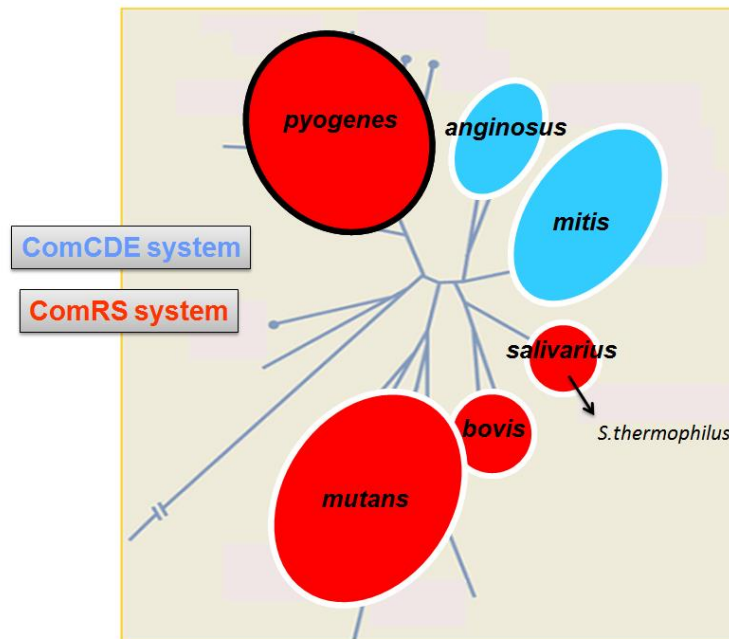


Figure 15. Distribution of ComCDE and ComRS systems among streptococcal groups. Our model species, Sth, belongs to the salivarius group, harboring the ComRS system.

i. The ComCDE system: signal transduction by phosphorelay

The well-known pathogenic agent Spn is part of the *mitis* group, and harbors the ComCDE system for competence regulation (Alloing *et al.*, 1998; Claverys and Havarstein, 2002; Havarstein *et al.*, 1995; Havarstein *et al.*, 1996; Pestova *et al.*, 1996).

In this system (see **Figure 16**), the precursor signalling peptide ComC is produced in the cytoplasm, matured and secreted in the external medium (Havarstein *et al.*, 1995). The mature form of ComC is called CSP (for Competence Stimulating Peptide). It can be exchanged between neighboring cells by cell contact, but does not freely diffuse in the medium (Prudhomme *et al.*, 2016).

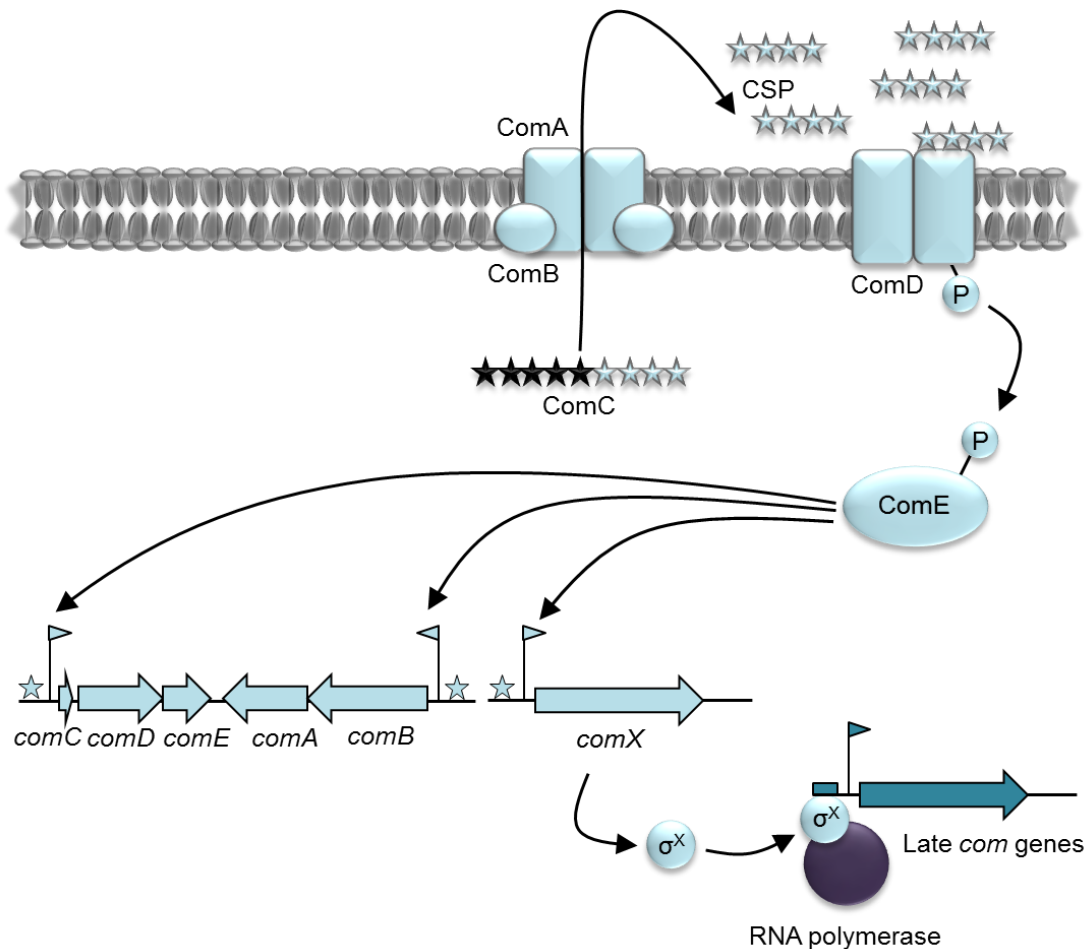


Figure 16. Topology of the ComCDE system. Signal transduction by phosphorelay: *comC* encodes a pre-peptide called pre-CSP, which is matured and exported. The resulting extracellular peptide, CSP, is sensed by the histidine kinase ComD, which in turn phosphorylates ComE and activates the transcription of genes having a ComE-box (represented by stars) in their promoter sequence (represented by flags) (figure adapted from (Boutry, 2013)).

Simultaneously to its secretion, the immature pre-CSP undergoes a cleavage by the ComA ABC transporter, yielding the mature competence pheromone CSP. The secreted peptides accumulate in the medium, and when a critical concentration is reached (the threshold concentration is typically around the nanomolarity), they interact and stimulate the ComDE Two-Component-System (TCS): the CSP is « sensed » by the membrane-anchored histidine kinase ComD, which in response auto-phosphorylates on an histidine residue (Havarstein *et al.*, 1996). ComD~P then transfers its phosphoryl group to the aspartate residue of its cognate response regulator ComE. Then, ComE~P dimerizes and become active. The active ComE~P

in turn activates the expression of the early *com* genes, including *comAB*, *comCDE*, *comX* and *comW* (ComW is an early gene product involved in σ^X stabilization and activation (Luo *et al.*, 2004; Sung and Morrison, 2005), see below), by binding to their promoter sequences. To do this, the AgrA/AlgR/LytTR domain of ComE binds to the so-called ComE-boxes, which consist in two 9 bp-direct repeats, separated by two helix turns, and located upstream of the early *com* genes. The auto-activation of the *comCDE* operon creates a positive feedback loop, which is essential to coordinate the entry into competence, and to maintain the cells in that state (Martin *et al.*, 2013).

The *comC* (encoding CSP)-*comD* pair of genes is quite polymorphic since it varies a lot between species and strains. For instance, in *Spn*, two different pairs CSP1-ComD1 and CSP2-ComD2 are found dominantly, and are called pherotypes. It has been shown that pherotypes are not immutable and can be exchanged between strains (Havarstein *et al.*, 1997). Due to differences in the CSP-ComD affinity, the communication between different pherogroups is relatively poor (Iannelli *et al.*, 2005), hence limiting the cross-talk between them. The effect of the existence of such pherotypes on the genetic diversity is still subjected to debates. However, the most recent studies suggest that the predation mechanism³ happening through sobrinicide between subpopulations harboring different pherotypes is actually promoting diversity by privileging exchanges between pherogroups (Carrolo *et al.*, 2009). Pherotypes also influence biofilm formation (Carrolo *et al.*, 2014).

In *Spn*, the genes coding for the early proteins are encoded in two operons (*comAB* and *comCDE*), located next to the replication origin *oriC* (Pestova *et al.*, 1996). This *oriC*-proximal localization has probably been selected by evolution due to the fact that it creates a genetic over-dosage of the *com* genes when the replication fork is stalled, hence helping the cells to cope with this issue by entering in the competence state (Slager *et al.*, 2014). As a consequence, when antibiotics (like mitomycin C and norfloxacin) cause the replication fork to stop, the copy number of the *comCDE* genes rises, yielding an increase of their transcription, hence triggering competence development (Slager *et al.*, 2014), and bacteriocin production (Kjos *et al.*, 2016). Remarkably, streptococci from these groups lack a SOS response system (Prudhomme *et al.*, 2006). This suggests that natural transformation could be used, among others, as a way to repair damaged DNA in these species.

In *Spn*, competence typically lasts for 15 minutes. Besides the intrinsic degradation of σ^X , at least three different mechanisms are responsible for the rapid shut-off of competence development in *Spn* (Weng *et al.*, 2013):

- i) Competition between active, phosphorylated ComE~P and non-active, dephosphorylated ComE: the latter competes with ComE~P to bind the the ComE-boxes.

³ In an ecosystem, predation is a biological interaction in which a predator kills a prey in order to consume its tissues.

- ii) A late gene product (activated by σ^X or by ComW) is likely to antagonize σ^X (Piotrowski *et al.*, 2009).
- iii) Besides its role in DNA transformation by recruiting RecA, the late ssDNA binding protein DprA is involved in the shut-off of competence by forming a complex with ComE~P, hence directly or indirectly inhibiting ComE~P activity (Mirouze *et al.*, 2013).

The extended shut-off plethora of regulations in *Spn* indicates that in this species, competence is a tool strictly reserved for very specific conditions, and that has to be inhibited as soon as transformation has occurred.

ii. The ComRS system: signal transduction by pheromone reimportation

The ComRS system has initially been discovered in *Sth* (Fontaine *et al.*, 2010a), and is present in many streptococcal species (Fontaine *et al.*, 2013; Havarstein, 2010; Mashburn-Warren *et al.*, 2010). It is found in the pyogenes, salivarius, bovis and mutans groups (Mashburn-Warren *et al.*, 2010; Mashburn-Warren *et al.*, 2012).

In the ComRS system, a precursor oligopeptide pheromone called pre-ComS is produced in the cytoplasm (see **Figure 17**), then secreted by a yet unknown exporter (no secretion signature except a Sec signal sequence in some cases), and concomitantly matured: the C-terminal part is released, yielding to a smaller, mature, ComS peptide, also named XIP (for ComX/SigX-Inducing-Peptide). XIP can then diffuse a bit in the medium (Fontaine *et al.*, 2010a), but it has been shown that it mainly remains attached to the cell surface in *Sth* (Gardan *et al.*, 2013). Contact between cells facilitates the exchange of XIP molecules between them, hence allowing the individuals to communicate. XIP peptides are then reinternalized through the Opp oligopeptide transporter system (from the superfamily of ABC transporters), also known as Ami in *Sth*. This transporter has been shown to be involved in the sensing and reimportation of a series of nutritional and pheromonal peptides in different Gram-positive bacteria (Gardan *et al.*, 2009). It can be saturated when a lot of free peptides are present in the medium, like in rich growth media where numerous casein or tryptone-encrypted peptides are present. These peptides can then hog the transporter, hence impeding the reinternalization of XIP, and preventing competence development (Gardan *et al.*, 2009; Gardan *et al.*, 2013).

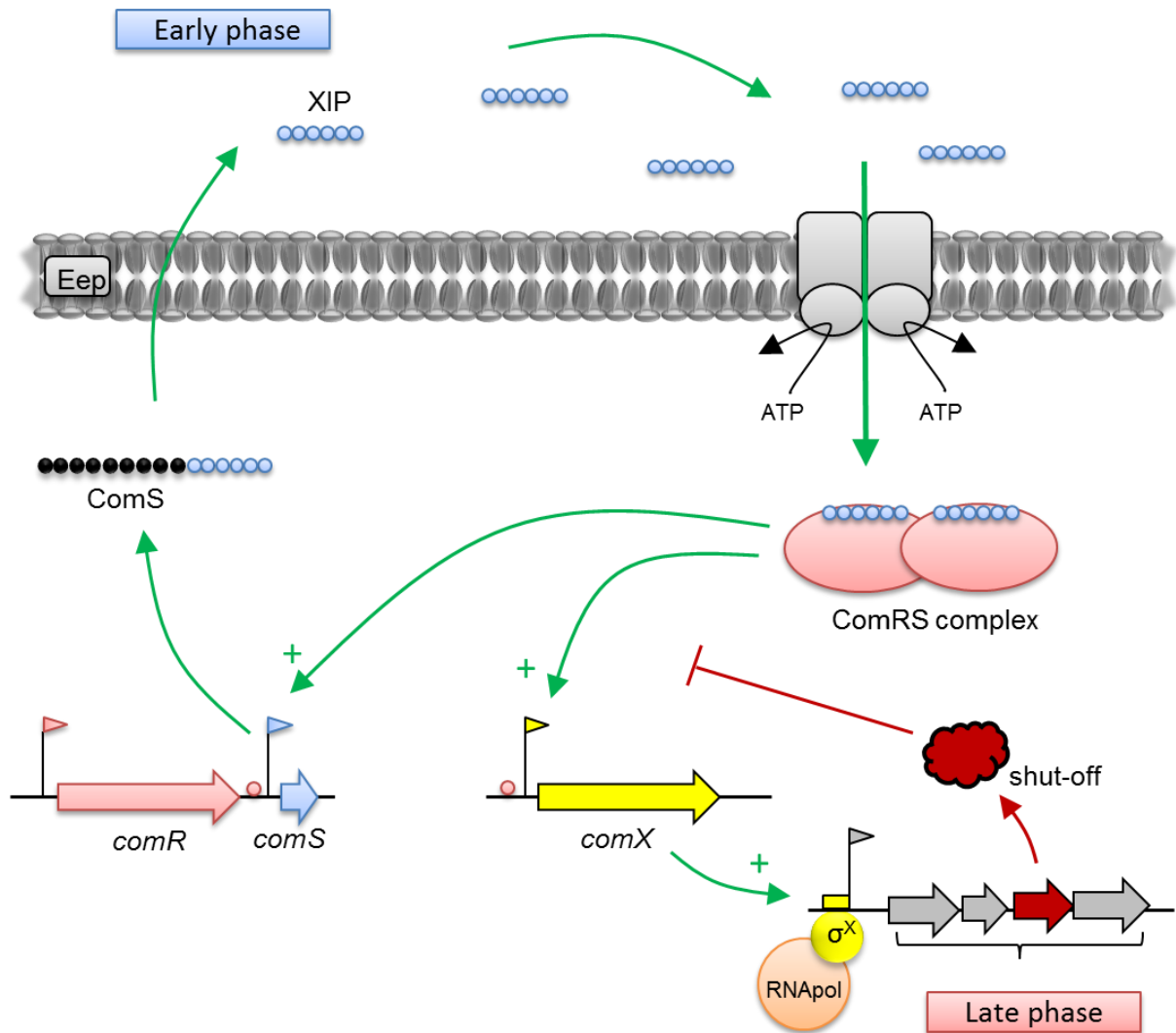


Figure 17. Topology of the ComRS system. Signal transduction by pheromone reimportation. The precursor of ComS is secreted, and exported outside the cell by a still unknown mechanism. It is concomitantly matured by the Eep protease. Mature ComS, or XIP, accumulates in the medium. It is then reimported through the Opp/Ami channel. Once reinternalized, it activates ComR, which dimerizes via a conformational change. This active ComRS complex activates the expression of *comS* and *comX*, the competence alternative sigma-factor, which in turn activates the expression of the late genes of competence by recruiting the RNA polymerase. The positive feedback loop on *comS* is essential for the synchronization of competence development. The *comR* gene is not included in the loop: it acts as the main integrator of signals and stimuli of the system (figure adapted from (Boutry, 2013)).

Once in the cytoplasm, mature ComS can bind to its cognate cytoplasmic response regulator ComR, forming an active complex named ComRS. The binding of mature ComS activates ComR, which in turn transcriptionally activates the expression of *comS* and *comX* genes, hence creating a positive feedback loop, which has been shown to be essential for competence synchronization (Fontaine *et al.*, 2013), and activating the transcription of *comX* (Fontaine *et al.*, 2010a).

The ComS signaling peptide

The mature ComS molecule or XIP is a small peptidic pheromone, which enables bacteria to communicate. The immature, pre-ComS peptide of Sth is 24-aa long (Fontaine *et al.*, 2010a). The Eep peptidase is likely to be involved in the maturation process (Gardan *et al.*, 2013). The minimal length of XIP that could activate competence in Sth is an heptamer (Fontaine *et al.*, 2010a), but its length can vary between strains and species (Fontaine *et al.*, 2013). According to the different species using the ComRS system, XIP can be classified into two main groups, depending on the specificity of sequence of their C-terminal part (Mashburn-Warren *et al.*, 2010). Streptococci from the salivarius group (including Sth) have a type I XIP (exhibiting the P(F/Y)F motif, and having no charged residues), and streptococci from pyogenes, bovis and mutans groups have a type II XIP (harboring a WW motif and charged residues). The suis group displays specific characteristics (and a WG(T/K)W motif), so the peptides from this group have been classified into a new particular type III (Zaccaria *et al.*, 2014; Fontaine *et al.*, 2015). XIP can easily be synthesized *in vitro*, and added to a liquid culture in order to trigger competence development and to maximize the transformation rate.

The ComR response regulator

ComR orthologs are closely related to the Rgg family (Fleuchot *et al.*, 2011; Fleuchot *et al.*, 2013; Zaccaria *et al.*, 2014) of transcriptional regulators. Its closest phylogenetically related homologs of known function are PlcR from *Bacillus thuringiensis* and PrgX from *Enterococcus faecalis* (Fontaine *et al.*, 2010a; Mashburn-Warren *et al.*, 2010), which belong to the RNPP (Rap/NprR/PlcR/PrgX) superfamily, containing pheromone-responsive regulators. In their inactive apoform, ComR monomers display a compact conformation, with a closed C-terminal peptide-binding pocket that interacts with the N-terminal Helix-Turn-Helix DNA binding domain. Upon ComS binding, ComR monomers undergo a conformational change, allowing them to dimerize, forming 2-2 complexes named ComRS, and to release the DNA-binding domain (Talagas *et al.*, 2016). ComRS complexes act as transcriptional activators, by binding to the so-called « ComR-boxes », which are 20 bp long and have the following consensus sequence of two 9-bp inverted repeats: 5'-TAGTGACATNTATGTCACTA-3' in case of Sth. Its core regulon is composed of bacteriocin encoding loci, *comS*, and *comX* (Fontaine *et al.*, 2013).

It is noteworthy that in opposite to the setting of the ComCDE system where ComE activates its own expression together with those of *comC*, in the ComRS system, the ComR regulator is not included in the positive feedback loop itself. Rather, the production of ComR has been proposed to be under the influence of a long series of external stimuli and signals, acting a bit like the « antenna » of the system (Fontaine *et al.*, 2013). This particular configuration could allow for an enhanced sensitivity of the system, since ComR production is only influenced by exogenous environmental signals, and not by the endogenous self-activation of the positive feedback loop.

b. Post-translational control of ComX abundance

As described above, ComX abundance is up-regulated at the transcriptional level by ComCDE or ComRS. Besides this first control, ComX abundance is also down-regulated at the post-translational level by the Clp proteolytic system in different streptococcal species (Biornstad and Havarstein, 2011; Chastanet *et al.*, 2001; Wahl *et al.*, 2014), similarly to ComK in Bsu. This degradation/stabilization system always involves the ClpP serine protease (Robertson *et al.*, 2002; Opdyke *et al.*, 2003), which can degrade the σ^X alternative sigma-factor by its association to ClpC (e.g. Sth) or ClpE (e.g. Spn), two ATPases subunits belonging to the AAA+ family.

In Spn, ComX is degraded by the ClpEP machinery (Biornstad and Havarstein, 2011). Interestingly, an early gene product, the small protein (80 aa) ComW, is involved in the stabilization of ComX (Sung and Morrison, 2005), and is (nearly) essential for DNA transformation (Luo *et al.*, 2004). Recent studies have suggested that ComW could act by enhancing the σ^X availability for the RNA core polymerase (Tovpeko and Morrison, 2014; Tovpeko *et al.*, 2016). In addition, the activation/stabilization of ComX by ComW seems to be very specific to Spn and its close relatives, since no ComW orthologs have been found in other streptococcal groups than the mitis group (Biornstad and Havarstein, 2011; Boutry *et al.*, 2012).

In Sth, it is ClpCP which is involved in ComX degradation as reported for the degradation of the master regulator ComK in Bsu (see section 2.2.) (Boutry *et al.*, 2012; Tian *et al.*, 2013). Recent studies show that ClpCP forms a ternary complex along with ComX and the adaptor protein MecA. The latter addresses ComX to the ClpCP degradation machinery by interacting with both ComX and ClpC (Wahl *et al.*, 2014). The inactivation of MecA leads to ComX stabilization and activation of the late phase of competence in non-permissive competence conditions (as well as higher transformation efficiency), hence showing that MecA is strongly involved (Wahl *et al.*, 2014; Boutry *et al.*, 2012). It has also been demonstrated that ATP hydrolysis is also needed for ComX degradation. It is thought that this MecA-related locking device of competence is used as a supplementary security to strictly control competence development, besides the positive activation performed at the transcriptional level by ComRS (Wahl *et al.*, 2014).

II. Distal regulation of ComX abundance

While standard laboratory growth conditions are quite static and homogeneous, real environmental conditions are dynamic, fluctuating with space and time, and often hostile because of the tough inter-species competition. This is why bacteria have developed sophisticated sensor systems in order to adapt their metabolism, their physiology, and to orient their transcriptional program according to the environmental cues. The regulation of natural transformation is one example of such transcriptional control. This signal transduction allows bacteria to trigger competence only when it is likely to bring them a fitness advantage. To do so, multiple sensors capture different

kinds of signals (see **Figure 18**) and modulate their action on the regulatory cascade of the early phase or on the σ^x alternative sigma-factor (Fontaine *et al.*, 2015).

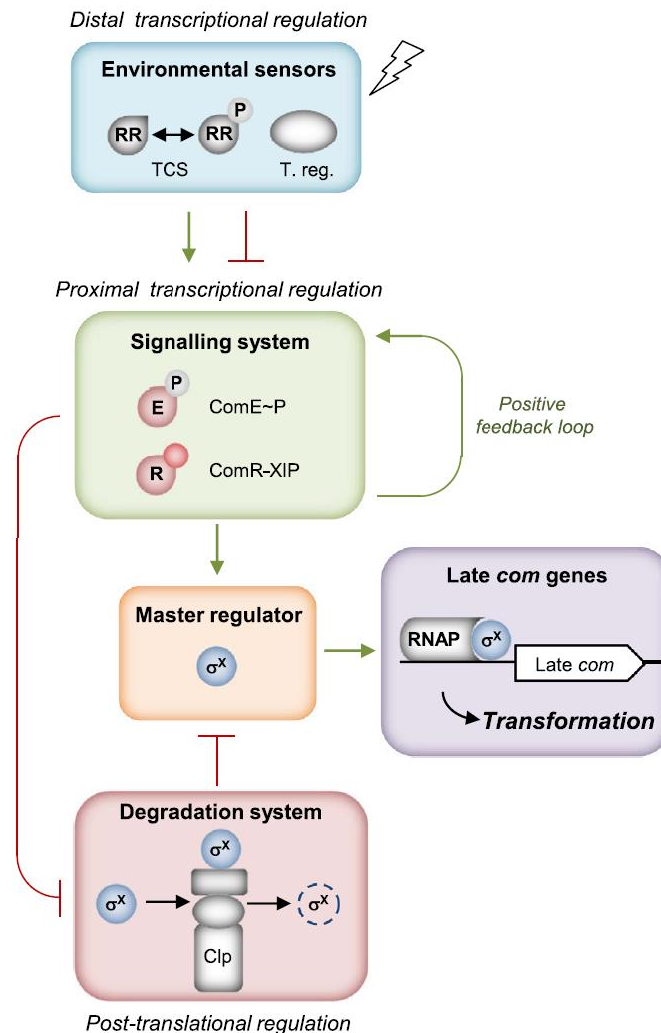


Figure 18. General organization of competence regulation in streptococci. The entry in the competence state is dictated by the amount of the σ^x master regulator (ComX) in the cell. This amount is regulated by the proximal systems ComCDE and ComRS (early phase), and by distal regulators, which are often implemented by TCS. ComX activation leads to the late phase of competence, resulting in DNA transformation. The ComX concentration is regulated by post-translational degradation systems. Green arrows denote activations, and red arrows denote inhibitions (figure from (Fontaine *et al.*, 2013)).

These signals are mainly related to stress (for instance, when environmental conditions threaten the cell surface). A series of TCS are able to sense abiotic (e.g. pH, temperature, oxygen concentration, ion content) as well as biotic signals (e.g. antibiotics, bacteriocins). Moreover, cellular growth and competence development have to be finely coordinated, since replication and transformation both imply consequent genetic rearrangements.

Here we will focus our attention on the current knowledge about distal regulation processes in the main models of streptococcal species: Smu, Spn and the model of this study, Sth.

a. Distal regulatory pathway specific to Smu

Streptococci from the mutans group have the particularity to possess both ComRS and a ComCDE-like regulatory system (see **Figure 19**). But while ComRS is the main, proximal system acting on *comX* (Mashburn-Warren *et al.*, 2010), the ComCDE-like system of Smu is not the paralog to the ComCDE proximal system present in Spn, as it was first thought (Aspiras *et al.*, 2004; Li *et al.*, 2001). It is rather more closely related to the pneumococcal cluster *blp* (this is why it has been recently renamed **BlpCHR**) (Martin *et al.*, 2006), and is linked to biofilm formation (Li *et al.*, 2002). Its main function is the regulation of the production of mutacins, that are small unmodified bacteriocins (Kreth *et al.*, 2005). It only indirectly impacts on ComX by inducing *comR* expression through a still unclear mechanism (Khan *et al.*, 2012; Wenderska *et al.*, 2012).

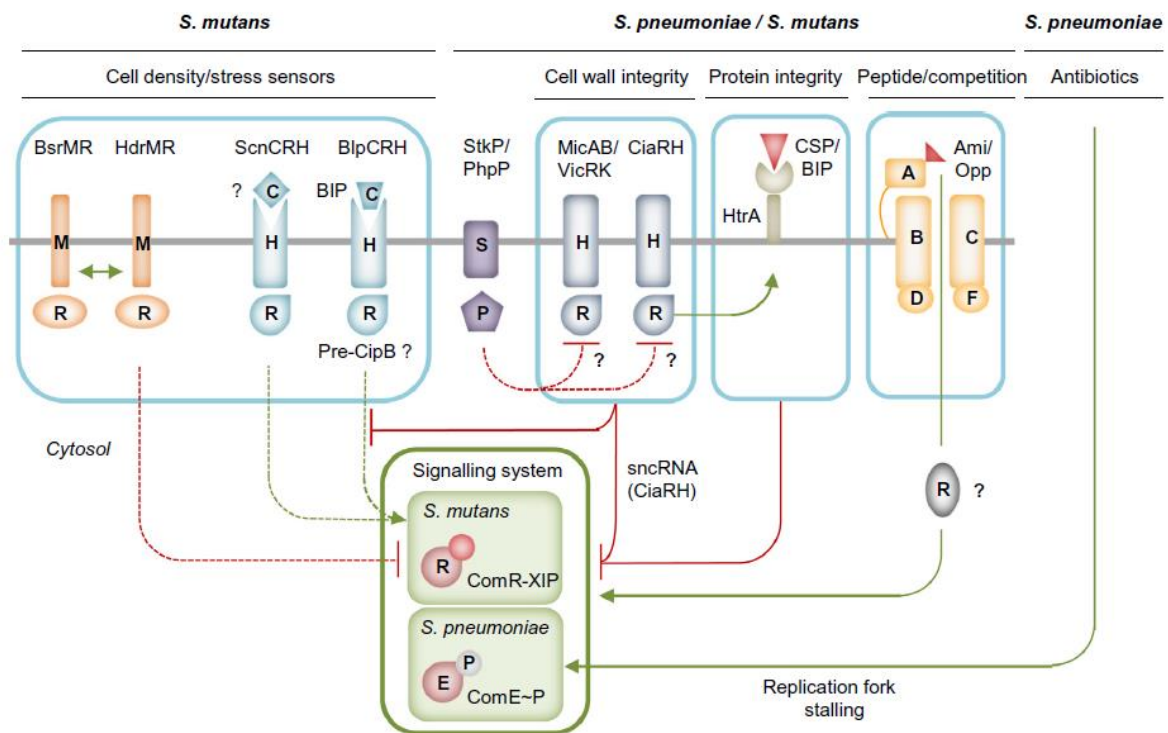


Figure 19. Distal regulation of pheromone based signalling systems in *S. mutans* and *S. pneumoniae*. In *S. mutans*, three specific signal transduction systems, HdrMR (interconnected with BsrMR), ScnCRH, and BlpCHR, responding to cell density or stressful conditions are believed to control competence through activation or repression of the ComRS signalling system. The BlpCHR system (previously called ComCDE) is activated by the bacteriocin inducing peptide (BIP, previously called CSP) which results in the intracellular accumulation of the pre-mutacin V (pre-CipB) which is postulated to activate the ComRS system. In *S. mutans* and *S. pneumoniae*, two additional signal transduction systems, MicAB/VicRK and CiaRH (RR/HK), sensing cell wall integrity, are negative effectors (red line) of the signalling system in both species and of the activity of the BlpCHR system in *S. mutans*. These two systems are hypothesized to be modulated by a eukaryotic-like serine/threonine

kinase (StkP or PknB) and its associated phosphatase (PhpP or PppL). The CiaRH system was also shown to induce the production of the cell wall protease HtrA which degrades CSP and BIP thereby reducing competence development. In *S. pneumoniae*, CiaRH was also shown to repress CSP production through a post-transcriptional mechanism involving a small non-coding RNA (sncRNA). In both species, the sensing of environmental peptides, which reflect the ecological competition status, by the Opp/Ami transporter also modulates the signalling systems, possibly via one or more unknown regulator(s) (R). In *S. pneumoniae*, antibiotics that target DNA replication, resulting in replication fork stalling and increase in the copy number of the comCDE operon located close to the replication origin, are responsible for competence induction. Green arrows: activation, red lines: inhibition, and dotted lines: hypothetical (figure from (Fontaine *et al.*, 2015)).

The BlpCHR signaling system is composed of the BlpHR TCS, and of the small 18-aa BIP peptidic pheromone (for Bacteriocin Inducing Peptide), encoded by *blpC*, and which directly regulates the production of mutacins and indirectly the expression of *comX* according to stress and cell density (Aspiras *et al.*, 2004; Li *et al.*, 2002; Merritt *et al.*, 2005). The ComRS is the proximal signaling system of competence regulation in *Smu*, since the deletion of the *blpCHR* operon does not impede XIP-induced competence activation (Mashburn-Warren *et al.*, 2010). However, both BIP and XIP pheromones have a complementary function in order to fine-tune competence regulation depending on the environmental cues: while a complex growth medium impedes competence activation by ComRS but allows the BlpCHR system to activate competence, a Chemically Defined Medium (CDM) will have the opposite effect (Son *et al.*, 2012). It is noteworthy that besides this dependence to the growth medium, the BIP-XIP interplay is also largely strain-dependent (Ahn *et al.*, 2006; Mashburn-Warren *et al.*, 2010; Prudhomme *et al.*, 2006). Evolution would not have retained the burden of this special regulatory setting if there was no fitness benefit linked to it. The double ComRS-ComCDE regulatory cascade could confer some advantage by allowing cells to finely adapt to diverse, time-and-space fluctuating conditions like those encountered in natural biofilm structures.

Moreover, **ScnC**, a small 61 aa-long peptide, is thought to act at the level of ComR production or activity. Indeed, the deletion of the *scnC* gene leads to a 10-fold decrease of the transformation frequency. The exact mechanism is still unclear, but it is likely to involve the **ScnRK** system, a TCS linked to the production of virulence factors (Kim *et al.*, 2013).

Another distal system of competence regulation is **HdrRM**, which is composed of a membrane-bound protein HdrM that represses the expression of its cognate regulator HdrR, belonging to the LytTR family of transcription activators (Merritt *et al.*, 2007). HdrR is induced at high cellular concentrations, and stimulates the expression of both the early (*comR* and *comX*) and late competence genes as well as the expression of genes encoding bacteriocins (Okinaga *et al.*, 2010). The **BsrRM** TCS also belongs to the LytTR family and has an indirect effect on competence, via stimulation of HdrR production (Xie *et al.*, 2010).

b. Distal regulatory pathway common to Smu and Spn

The distal TCS **CiaRH** (see **Figure 19**) was found to indirectly modulate *comX* expression in both Spn (Halfmann *et al.*, 2011) and Smu (Mazda *et al.*, 2012). CiaR~P, the phosphorylated form of the regulator, represses competence development by affecting CSP (BIP in Smu) concentration (Dagkessamanskaia *et al.*, 2004) through two different processes: (i) through a post-transcriptional repression of CSP production promoted by base-pairing between *comC* transcript and small non coding RNA (sncRNA) induced by CiaRH (Schnorpfeil *et al.*, 2013) and (ii) through CSP degradation via the extracellular protease **HtrA** (housekeeping chaperone), which is a direct target of CiaRH regulation (Cassone *et al.*, 2012). Recently, it has been proposed that the multiple translational errors induced by some antibiotics could saturate HtrA, annihilating its repression on CSP, hence triggering competence under antibiotic stresses (Stevens *et al.*, 2011). In Smu (Senadheera *et al.*, 2012; Senadheera *et al.*, 2005; Senadheera *et al.*, 2007) and Spn (Echenique and Trombe, 2001), transcription of *comCDE/blpCHR* is also repressed by the MicAB/**VicRK** TCS, an essential system involved in cell division, cell wall integrity, oxygen and redox sensing, which regulates the production of peptidoglycan hydrolases involved in cell division (Deng *et al.*, 2007; Duque *et al.*, 2011; Liu and Burne, 2011; Ng *et al.*, 2004; Patenge *et al.*, 2013; Senadheera *et al.*, 2005; Wayne *et al.*, 2012). In both *S. mutans* (Banu *et al.*, 2010; Hussain *et al.*, 2006; Reck *et al.*, 2011) and *S. pneumoniae* (Echenique *et al.*, 2004; Saskova *et al.*, 2007), the serine/threonine kinase StkP/PknB, and their associated phosphatase (PhpP or PppL) positively controls natural transformation by regulating the basal level of *comCDE/BlpCHR* expression. It is probable that StkP/PknB acts indirectly by controlling the phosphorylation states of CiaR and VicR repressors, independently from their cognate histidine kinase. Indeed, VicK and StkP have overlapping regulons (Banu *et al.*, 2010; Reck *et al.*, 2011).

Besides the distal regulators cited above, the global **CodY** nitrogen sensor (Caymaris *et al.*, 2010; Stenz *et al.*, 2011) is also involved in competence regulation in Spn (Claverys *et al.*, 2000) and Smu. Interestingly, this global regulator of nitrogen metabolism (Groot *et al.*, 2012) has been shown to be involved in competence and sporulation in Bsu, and in virulence in *Staphylococcus aureus* (Pohl *et al.*, 2009).

c. Distal regulatory pathway in our model Sth

In opposite to the *comCDE* operon, the *comRS* cluster is not located next to the origin of replication (~270 kb in the salivarius group). Hence, no gene over-dosage is created when replication forks are stalled. Besides, antibiotics have the opposite effect in Sth: they impede competence development (Boutry *et al.*, 2013). It is worth to note that in contrary to Spn, almost all streptococcal groups using the ComRS system have an equivalent to the SOS response system regulated by LexA (HdiR) (Varhimo *et al.*, 2007).

The ComRS system in Sth and in Spn is often referred to as a « timing device », since its activation always initiates after the same time post inoculation of the culture (Gardan *et al.*, 2013). This characteristic timing suggests that competence onset depends on metabolic conditions, linked to the growth phase. The exact signaling pathway involved is still unknown.

The oligopeptide content of the growth medium highly influences competence in Sth. Indeed, a medium enriched with tryptone extracts inhibits its development, either directly by saturating the Opp reimportation machinery, or indirectly via another distal regulator that would sense nitrogen components (Gardan *et al.*, 2009). In contrast, a CDM growth medium supplemented with specific casein-encrypted peptides triggers competence by mimicking the structure of XIP peptide (Fontaine *et al.*, 2013).

Competence development in Sth is also influenced by different TCS, which would deserve further investigation (L. Fontaine, unpublished results). For instance, the CodY transcriptional regulator, which is widely present in the low-GC Gram⁺ bacteria, is a major regulator of the different metabolic pathways in Sth: it was shown to control the amino acid and carbon metabolisms, and is also involved in nutrient transport and stress response (Lu *et al.*, 2015). Another well studied distal regulator is CovRS, a TCS implied in the control of many virulence-related genes (Olsen *et al.*, 2009).

E. Subpopulation heterogeneity for competence development

In spite of the population-scaled synchronization of competence through the production and sensing of auto-inducers, behavioral differences remain between individual (though monoclonal) cells: they do not behave exactly the same way. Indeed, single-cell level observations in many species show that only a fraction of the whole population actually enters into the competence state. The cells that become competent are said to be in the « **ON** » state, while the remaining individuals are said to be in the « **OFF** » state. The phenomenon during which a monoclonal population differentiate into two different subpopulations in homogenous growth conditions is called « bistability » or « bimodality » (Dubnau and Losick, 2006).

The **bistability phenomenon** exists in many other areas than in competence development (Dubnau and Losick, 2006; Angeli *et al.*, 2004). Many well-documented examples exist for bistability: for instance, Bsu cells « can decide » between the swarming and the swimming motility (Calvio *et al.*, 2005; Kearns and Losick, 2005), and whether to express extracellular proteases or not (Veening *et al.*, 2008a); in many species, only a fraction of the population becomes persister cells (Lewis, 2010; Lewis, 2007): their growth is transiently arrested in order to survive to a difficult situation ; and in many species, the sporulation program is undertaken by a subpopulation of the cells (Govin and Berger, 2009).

The occurrence of bistability relies on specific conditions regarding the topology of the regulatory network (see **Figure 20**). The topology of the regulatory circuit

controlling the entry in the ON state has either to (Ferrell, Jr., 2002; Smits *et al.*, 2006):

- i) include a positive feedback loop reinforcing the expression of the gene responsible for the ON phenotype, or
- ii) include a double nested negative loop (this kind of loop is called a “toggle-switch”) influencing the expression of the gene responsible for the ON phenotype, or
- iii) include some combination of the above (for example, an even number of nested negative loops), allowing some strong non-linearity between the inputs (the internal state of the cell as well as all environmental signals affecting competence) and the critical output (the expression of the target gene, *i.e.* *comK* or *comX* in the case of competence development).

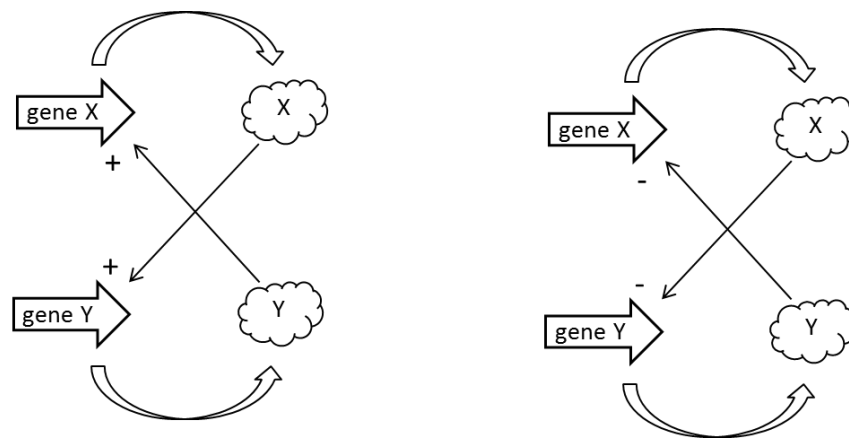


Figure 20. Topological requirements for bistability in genetic regulatory networks. The positive feedback loop (left: two genes activate each other) or the toggle switch (right: two genes repress each other) is needed for the bistability to occur. Straight arrow: activation/inhibition, curved arrow: protein production.

Besides the necessity of a suitable topological regulatory setting to promote a bistable response, small **inter-individual variations** are required. Indeed, these single-cell differences are the starting point for the larger differentiation between the ON and OFF states, through their amplification by the non-linearity implemented in the regulatory pathway. (Dubnau and Losick, 2006). In their work on the ComK master regulator of competence in Bsu, Dubau *et al.* decomposed these variations, into intrinsic (random events of transcription, and stochastic decay of mRNA) and extrinsic (linked to the amount of regulators in the cell) noise (Maamar *et al.*, 2007).

It is important to realize that **stochasticity** is inherent to the molecular level of genetic regulatory networks. Indeed, the transcriptional regulatory mechanisms typically involve low-concentrations of the key molecules (signal molecules and transcriptional activators and/or repressors), and are as a consequence subjected to a high level of stochastic noise (Elowitz *et al.*, 2002; McAdams and Arkin, 1997). Each

cell has its own metabolic and physiological state, possibly subjected to stochastic fluctuations (Schwabe *et al.*, 2013), and largely depending on its growth phase.

Moreover, in nature bacteria are often confronted to dynamic, fluctuating growth conditions in complex and changing structures like biofilms and host tissues, subjecting them to varying (according to time and space) physical parameters. These variations can affect any intensive⁴ property of the medium, but the most relevant are: temperature, pressure, mechanical stress (for instance, shearing stress), and concentration. The latter parameter often varies according to gradients, and can concern a long series of different substances (biotic and abiotic), among which are bacterial signal molecules. This heterogeneity in environmental settings could, at least partially, contribute to the heterogeneity in gene expression between neighboring monoclonal cells, and hence to the heterogeneity of competence development. However, the fact that bistability can be observed in laboratory conditions (in which the well-stirred growth media are in general quite spatially homogenous) indicates that heterogeneity in environmental conditions is not a necessary condition for heterogeneity in competence development.

Differences between individuals can also stem from the epigenetic level: even with the same genetic background, bacteria can differ by the expression levels of their genes, which can depend on their past history and even on the history of their parent cell (epigenetic inheritance).

It results from these small inter-individual variations that two monoclonal cells, living next to each other, can react slightly differently to the same signals. This tiny difference will be amplified by the **non-linear activations** of the regulatory pathway, and can be the starting point for differentiation, leading to bistability at the population scale-level.

The most extensively studied species for bistability in competence development is *Bsu*. It has been shown that the positive feedback loop affecting *comK* expression is necessary (Smits *et al.*, 2005) and sufficient for bistability to occur, since the Rok-ComK toggle switch (in which Rok and ComK inhibit each other) is not required (Maamar and Dubnau, 2005). It also requires a certain noise level on ComK concentration (Kampf and Stulke, 2015; Gamba *et al.*, 2015).

Bistability was also investigated in *Smu*. The regulation of competence development in this species involves the two signaling peptides CSP (or BIP) and XIP. However, in *Smu*, the CSP belongs to the system which is orthologous to the BIPCHR system of *Spn* (see section 2.4.2) (Martin *et al.*, 2006). In 2012, the team of Son *et al.* studied the interplay between these two systems regulating competence (Son *et al.*, 2012). They observed that the composition of the growth medium influenced which system

⁴ An intensive property is a property which can be defined in each point of a system, in opposite to an extensive property, which is additive in subsystems. Examples of extensive properties are volume, mass, energy, entropy.

was controlling competence development: while CDM allowed XIP to elicit a uniform development of competence across the population (in a *comS*-independent way, but requiring the reimportation of XIP through the permease Opp), a rich growth medium (containing a lot of free peptides) did not. In this complex medium, CSP was able to trigger competence (in a *comS*-dependent way, but without requiring the reimportation of XIP) with a bimodal distribution (Son *et al.*, 2012). Since competence development in Spn (harboring only the paralog ComCDE system) does not give rise to bimodality, Son *et al.* propose that the ComRS system of Smu is accountable for its bimodal distribution of competent cells in complex medium. They suggest that the composition of the rich medium could favor the ComRS system to autoactivate, hence promoting a bistable behavior.

As pointed out by Wei *et al.*, the bistability stemmed from stochastic noise is robust to fluctuations affecting the signal transduction system (Wei *et al.*, 2014), meaning that small perturbations in the system parameters do not abolish bistability. Moreover, the work of Goryachev *et al.* suggested that the dimerization of the transcriptional regulator is very important to get an efficient ON-OFF genetic switch (Goryachev *et al.*, 2006). Indeed, the dimerization reinforces the switch-like behavior of the system, as well as the stability of its steady states in the presence of stochastic noise.

Like with other bistable states, heterogeneity in competence development confers a **global advantage** to the whole population (Wylie *et al.*, 2010): this kind of bet-hedging strategy spreads the risks under challenging conditions (Veening *et al.*, 2008b). Indeed, the coexistence of various phenotypes maximizes the probability that at least one of the offsprings will survive to the threatening situation (Cohen, 1966). Obviously, even if all phenotypically different offsprings are not optimal, this heterogeneity will confer a global fitness advantage to the whole population, especially when confronted to hostile and changing conditions (Kussell and Leibler, 2005; Thattai and van, 2004). Moreover, it has been suggested with *Saccharomyces cerevisiae* that inter-phenotypic variation rates could be tuned to match the frequency of environmental changes (Acar *et al.*, 2008).

The percentage of cells entering in the ON state varies largely, depending on the concerned program (e.g. bioluminescence, sporulation, persistence, swimming motility, competence development). In the case of competence development, it also depends on the species and on environmental conditions. Bsu populations typically count 10 to 20% of competent cells (Leisner *et al.*, 2008; Schultz *et al.*, 2007; Leisner *et al.*, 2007; Leisner *et al.*, 2009); Smu around 10% (Lemme *et al.*, 2011), and Sth barely achieves 10%. The pathogenic species Spn is remarkably transformable, since almost 100% of its population is transformed during a competence event (Claverys and Havarstein, 2002). This can partially be explained by the fact that in Spn, fratricide is co-programmed with competence: non-competent cells are killed by the competent ones, so the only surviving cells are the transformed ones after a competence window (Martin *et al.*, 2010; Steinmoen *et al.*, 2002).

F. Open questions about the ComRS system

The ComRS system has been well studied and described, especially in the *Sth* (Fontaine *et al.*, 2013;Fontaine *et al.*, 2010a;Gardan *et al.*, 2009;Gardan *et al.*, 2013;Talagas *et al.*, 2016) and *Smu* (Desai *et al.*, 2012;Son *et al.*, 2012) species. However, a series of questions remain open about it. Here are some of the main remaining interrogations:

- i) First, the exportation step of pre-ComS is still unclear, since the exportation machinery has not been yet identified. Moreover, the maturation step by species-specific proteases has not been unraveled so far, except the known role of Eep in (Gardan *et al.*, 2013).
- ii) Second, the various signals that influence the transcription of the ComR-encoding gene could be investigated.
- iii) Third, the shut-off mechanism is still unknown.
- iv) Fourth, the regulation of bistability in competence development still deserves further research.

3. Mathematical modelling of regulatory networks

The major part of this thesis consisted in modelling the regulatory network controlling competence development in Sth. In this chapter, we will basically introduce what mathematical models are, how they are built, and why they can be useful in system biology.

A. What is a system ? What is a model ?

Let's see some basic definitions in order to share a common language. Any scientific activity requires the **quantification** of real magnitudes. Hence, mathematics are used everywhere. We build theories and models in order to gain understanding, prediction and control. But what are models ?

A **model** is a simplified, conceptual or mathematical representation of a system (Torres and Santos, 2015), that serves to understand and quantify it. Models are truly used in every scientific sector, ranging from chemistry, biochemistry, physics, biology, economy, aeronautics, ecology, genetics, astronomy, meteorology – just to name a few. The modelling process absolutely requires a **scientific method**, which is « *the procedure consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses* ». The subject of the model is in general complex, and can be seen as a system.

A **system** is composed of a set of elements, physical objects, usually numerous and diverse, that influence each other and that are physically and functionally separated from their environment (Torres and Santos, 2015). Here, the “environment” simply signifies everything that is not included inside the system. The frontier can be real (like a cell membrane) or imaginary (like the frontiers of a state), and can be permeable to matter and energy – the system is then said to be « open ». In the case of biological systems, the separation is a consequence of the fact that such systems are far from thermodynamical equilibrium, in opposite to the environment. The interchange of matter and energy is a requisite to sustain the chemical-physical processes that occur in living systems, sustaining homeostasis.

Examples of systems are: a set of interacting preys and predators (it could be cats and mice, for instance), evolving in the same habitat ; a whole living cell (with its metabolic reactions) ; an articulated arm used to replace a disabled one ; a synchronized shoal of fishes looking for food ; a set of poker players; and so on...

B. Why building a mathematical model ?

Nowadays, mathematical models are flourishing in every sector, and especially in biology. Indeed, the recent technological progresses have enabled very high-throughput analyses, performed at the DNA, mRNA, protein and metabolite levels. The so-called “Omics era”, in reference to all the recently developed fields like the metabolomics (Goodacre *et al.*, 2004), the proteomics and the interactomics (Pandey and Mann, 2000). Initially, mathematical models in systems biology were built to understand metabolic pathways (Heinrich *et al.*, 1991; Heinrich and Schuster, 1998; Schilling *et al.*, 1999). But mathematical models are increasingly used to harvest data from expression profiles (microarrays) (Bansal *et al.*, 2007). The research in these different fields has created such a terrific amount of available data about any measurable thing, that mathematical models have become inescapable in order to extract useful information from these tons of numbers.

From a pragmatic point of view, a model is used to *sum up* and to *quantify* the functioning of a system, to better *understand* it, to test *hypotheses* about it, to find the *minimal* set of hypotheses explaining the data, to *predict* its behavior in non-tested conditions, to help *controlling* it, to raise good *questions* about it, to find the *lacks of knowledge* in our knowledge of the system (it is especially interesting when the model does not fit the data), to discover *counter-intuitive* facts, to perform experiments *in silico* (less expensive and risk-free), and to *get ideas* for new « real-world » experimentations.

For instance, in the case of the Malaria disease, also known as paludism, which is a major issue in subsaharian Africa. People get infected only through mosquito bites (there is no inter-human risk), and mosquitos become infected when they take the blood of infected humans. This system, consisting in a set of inter-related humans and mosquitos, potentially affected by the disease, can be modelled in order to help preventing and controlling outbreaks: the model can be used to test *in silico* different scenarios or conditions, in order to decide which actions should be taken. For instance, we could be interested in knowing how the heterogeneity of the human population affects the outbreak (Chitnis *et al.*, 2008), or how the movements of people (Silal *et al.*, 2015) and how the seasonal variations of temperature and rainfall (Abiodun *et al.*, 2016) affect the transmission dynamics. The mathematical formalization can help considering which actions are the most effective to get rid of an epidemic (Silal *et al.*, 2014). For a review of the different mathematical formalisms used to model malaria, see (Fonseca and Voit, 2015).

C. Descriptive versus mechanistic models

We can classify models in two types: descriptive and mechanistic models (Torres and Santos, 2015; Gardner and Faith, 2005).

A **descriptive** model is a quantitative summary of the observed relationships among a set of measured variables (Torres and Santos, 2015). For instance, we could study how temperature changes affect the Malaria disease transmission rate: in this case, we just observe the relationship between one variable (the air temperature) and another variable (the rate of transmission). It allows to detect **correlations** between those variables. But it tells nothing about how the variables are connected: we cannot directly infer **causality** from descriptive models. Moreover, a given descriptive model can only describe one particular system in a given situation (particular conditions): such models are always very specific. Hence, they are not easily adaptable. They mainly consist in all kinds of regressions, and some marginal models like Neural Networks. In one sense, they are more a statistical analysis rather than a true model. Moreover, descriptive models are damned by the curse of dimensionality: a huge amount of data is needed to build statistical models (it grows exponentially with the number of variables).

A **mechanistic** model is built by explicitly considering and transcribing the processes that produce our observations: we really model *how* the system functions (Torres and Santos, 2015). In contrast to descriptive models, a mechanistic model explains the observations, and can be used to infer causality dependences. Besides, it can be modified or generalized to other variants of the system or to other conditions and scenarios. Mechanistic models often consist in a set of « state variables » inter-related by dynamic equations.

Indeed, mechanistic models are generally **dynamic**: since the mechanisms and processes ruling the system functioning are taken into account through the causal relations, the model allows to reproduce the system behavior, which means the succession of the different states of the system as time passes by (Torres and Santos, 2015). For instance, a SIR-type model (which simulates the evolution over time of the Susceptible, Infected and Recovered individuals, depending on the parameter setting), explicitly describing how one particular disease (e.g. Malaria) is spreading among humans (and mosquitos), can be used to *simulate* the evolution of an outbreak through time, offering the possibility to test (virtually) different intervention programs beforehand, without taking any risk, and to see what is the best option to implement in order to improve the situation on the ground (Kato *et al.*, 2011).

Dynamic (mechanistic) models are more difficult to create than descriptive (statistical) models. But they are much more informative and yield a better scientific understanding (Torres and Santos, 2015). Moreover, sometimes experiments are only feasible *in silico*: when it involves very large time or space scales, when the cost of the real on the ground experiment is too high, when the experiment could be hazardous (endangered species management by trial and error for instance), or when it involves humans (setting dosages for clinical trials).

D. Particularities of biological models

Dynamic models of biological systems can be very diverse, depending on the area of biology (in transcriptomics: a gene's expression level, in ecology: the abundance of a (endangered) species, in physiology: the mercury level in different organs, in epidemiology: the prevalence of an infectious disease...), on the mathematical setting (continuous *versus* discrete-time models, deterministic *versus* stochastic models...), on the method to study the model (mathematical analysis *versus* computer simulation), the purpose of the model (fundamental science, medicine, environmental management...).

Biological systems often involve non-linearities, complex interactions, and different interconnected layers: proteins, mRNA, DNA (see **Figure 21**). Models are really simplified views compared to the complex reality of biological systems. Unraveling the complexity of the dynamic and intertwined genetic regulatory networks (GRN) is one of the main challenges of systems biology.

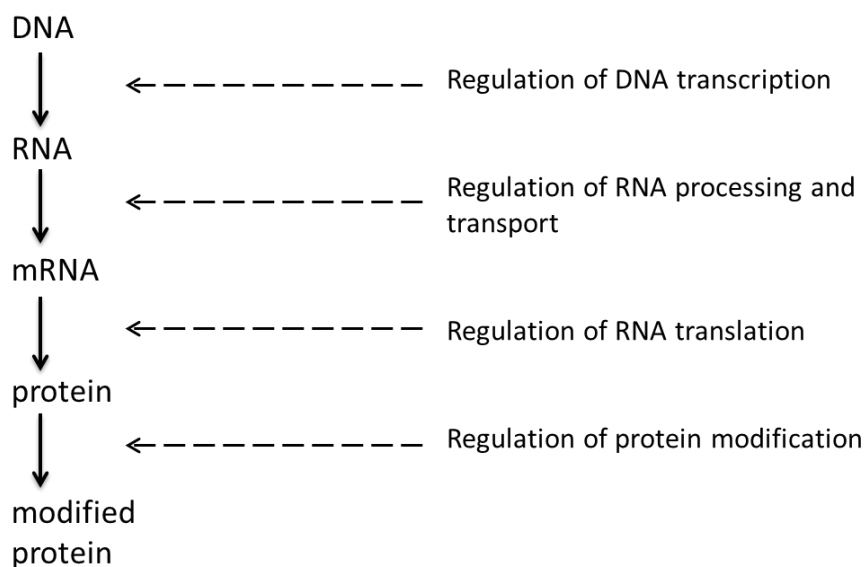


Figure 21. Interconnected layers in system biology. A final protein amount can be regulated at different stages: transcriptional, post-transcriptional, translational, and post-translational.

E. How do we build a model ?

Model building is a complex process, representing an area in its own right, and requiring to follow some basic machine learning rules in order to avoid common modelling mistakes. Here, we briefly explain the successive steps for model building (Torres and Santos, 2015), in a very simplified way. These steps are illustrated in **Figure 22**, and can vary from model to model.

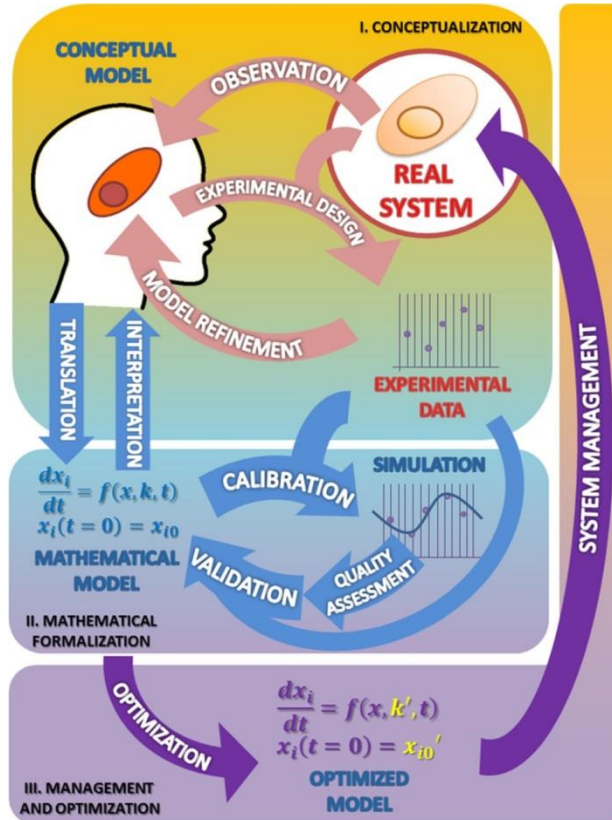


Figure 22. Successive steps in model building. A conceptual model is translated into equations, constituting the mathematical model, which is calibrated and then validated with independent data sets. The model can be further optimized by the use of new experimental data (Torres and Santos, 2015).

I. Conceptualization

The first and most obvious step of the process is the conceptualization of the model. It consists in establishing statements about the system, after observation (and potentially experimentation), by formulating hypotheses about its functioning. Taking a very simple example, these observations could be for instance: « the more glucose we put in the growth medium, the more the cells grow ». The following statement would then logically be proposed: « cells use glucose to grow ». We could then set an experiment to measure the quantitative relation between the amount of glucose

available to the cells, and the number of cells that have grown on it, in order to refine the model.

The iteration [hypothesis forming – experimenting] can be repeated as long as the experimenter has ideas (and time and resources) for new experiments, in order to get the best conceptual model as possible. In brief, a conceptual model is thus a collection of clear, accurate sentences (or statements) about how we think the system works.

II. Choice of the formalism

Once the conceptual model has been set up, one has to decide which formalism best fits to it. There are many formalisms from which to choose in order to model a (dynamic) system, but the most used ones are, from the most basic to the most complex (approximately):

- **Boolean Networks:** a set of boolean (logical) variables are defined by discrete-time equations (Albert and Thakar, 2014);
- **Dynamic Bayesian Networks:** a set of random variables are related to each other through some joint probability distribution (Tian, 2016);
- **Ordinary Differential Equations (ODE):** a set of continuous variables are related to each other through a set of equations (Vanlier *et al.*, 2013);
- **Delay Differential Equations (DDE):** the system of differential equations contains at least one time delay (Wang *et al.*, 2008);
- **Artificial Neural Networks:** different layers of “neurons” (typically three) are interconnected through “synapses”, that use specific activation functions. This formalism requires the most technical learning procedures, and is a kind of “black-box” model, with no direct physical interpretation of it (Zou *et al.*, 2008; Manning *et al.*, 2014).

The choice mainly depends on the type of variable we have and the type of relationships between them. For instance, a system where the state variables are logical will be well-suited for a boolean formalism, while continuous variables with continuous relationships will require a continuous-time formalism like differential equations.

III. Transcription into equations

When the conceptual model and the formalism have been fixed, it can appear to be straightforward to put everything down into equations: it simply consists in translating the sentences of the conceptual model into the chosen language, *i.e.*, into the chosen formalism.

But while the translation in itself is really direct, attention must be paid to the level of detail we chose to integrate in the model. Indeed, if this level is too high, then we have the risk of « **overfitting** », which means that the « resolution » of the model is too high: noise and other undesired, non-relevant features are being modeled. On the opposite, if the level of detail is too low (i.e., if the model does not take enough features into account), then the model will be really basic: it will not be able to simulate interesting behaviors, and its utility will then be reduced. So, it is a compromise to find between simplicity and informativity (Wessels *et al.*, 2001). The sparseness of a model reflects its simplicity. It refers to the fact that each node (each gene) has a low in-degree, i.e., not too much regulators (Arnone and Davidson, 1997). On the contrary, there is no restriction on the out-degree, since some regulators may affect a long series of genes.

IV. Learning step

Once the conceptual model has been translated into a set of equations, the crucial part consists in finding values for each parameter of the system. Indeed, at this stage all the parameters are still « symbolic », meaning that they have no real value yet. Though such a system is conceptually interesting, it cannot be used to deduce any quantitative information.

So, the learning step consists in finding values for all the parameters, in such a way that the model outputs fit with the experimental data. This set of data is called the « learning set ». The idea is to explore the space of parameters in order to find the combination giving the best matching between the model simulations and the experimental data. This can sometimes be done by minimizing a scalar value of the error (i.e., the difference between the model and the measured reality), for instance the least squares error (LSE) (van Someren *et al.*, 2006). But this is only feasible for simple cases, when the error can be calculated as a scalar value. Unfortunately, this is not always the case: depending on the type and amount of available data, it is often more complicated. Indeed, sometimes a lot of diverse outputs have to be compared to noisy measures, making it almost impossible to sum up the learning error into a neat, unique scalar value. The learning step can then be done by trial-and-error, possibly with a brute-force search algorithm when the network is very small and the dynamics simple (like in (Chen *et al.*, 1999) for instance), in order to approach a solution which is globally and qualitatively satisfying all the various criteria. Compromises and common-sense are often the most useful tools in this difficult task.

So, there is always a compromise to be made between having a lot of experimental data, hence making the model more reliable, but increasing the experimental costs, and minimizing the experimental burden. It is believed that the relation between the complexity of the network (the number of nodes) and the amount of necessary data to learn the model is logarithmic (Yeung *et al.*, 2002).

V. Validation step

Once calibrated, the model parameters have all been assigned values that make the model output match with the data. But at this stage, we cannot say if the model is **robust**, *i.e.*, if it can be used to predict behaviors that have not been tested during the learning phase. The notion of robustness is linked to the ability of the model to adapt to new situations.

To test the robustness, a second, *independent*, set of experimental data is needed (see **Figure 23**). This second set is called the « validation set ». The idea is to test whether the model is able reproduce behaviors that were not used to calibrate it. It is absolutely mandatory to have independent learning and validation sets, because validating a model with data of the learning set would have no sense: it would always work !

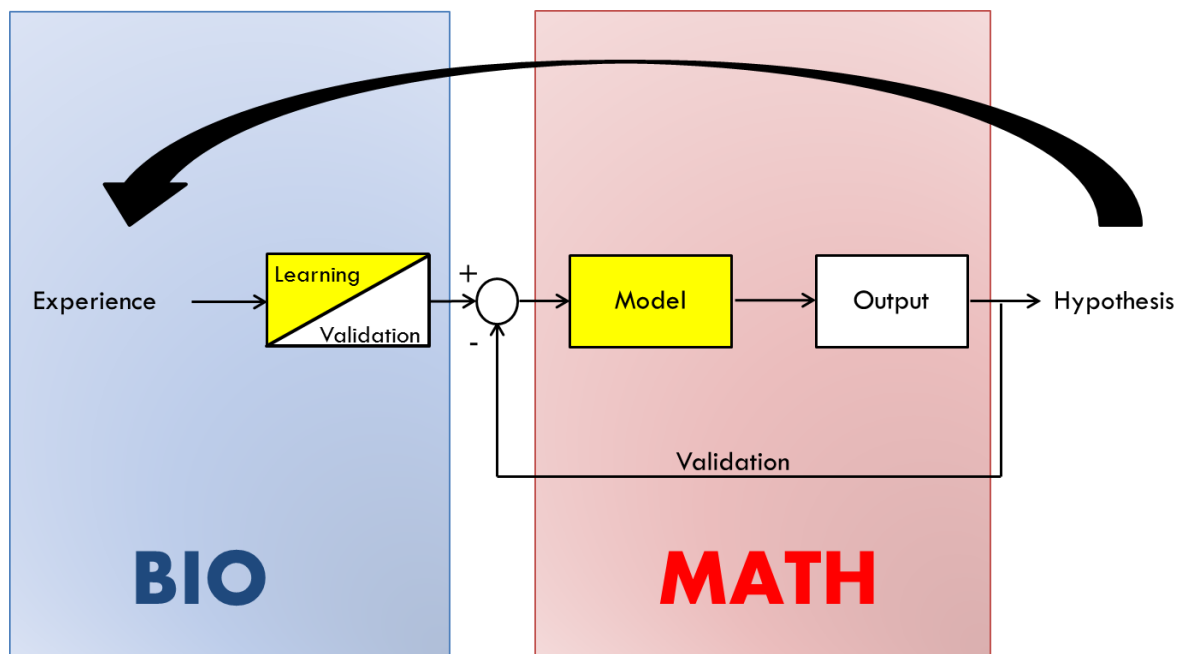


Figure 23. Iteration process [experience → hypothesis → experience]. The available experimental data have to be handled carefully (independent learning and validation sets). The model is validated when the model outputs match the experimental evidences.

A lot of different techniques exist to perform the learning and validation steps. For instance, the bootstrap algorithm is designed to optimize the repartition of data between the two sets, with random distribution, allowing to learn and test the model for each information.

VI. Optimization and use of the model

The validated model can be further optimized by integrating the information coming from new experimental tests. The iteration [test new things – improve the model] can be repeated, again, as long as the experimenter has the resources for it. Of course, each new component of the model will have to be independently validated.

The final model can be used in order to infer new knowledge about the system and its functioning, by testing new hypothesis and scenarios. If the model is particularly well calibrated, quantitative information can be inferred. For instance, if the parameters have a physical signification (which is not always the case), and if the calibration was really accurate, then it will allow to get values for unknown physical parameters.

F. State of the art: what kind of models do exist ?

Many mechanistic mathematical models have been built in order to describe biological systems. For instance, ODE have been used to model the cellular physiology in different organisms (Atlas *et al.*, 2008; Browning *et al.*, 2004; Castellanos *et al.*, 2004; Castellanos *et al.*, 2007; Chubukov *et al.*, 2012). A whole-cell model has even been proposed by Karr *et al.* for the human pathogen *Mycoplasma genitalium*: by assembling 28 different sub-models (using different formalisms), they managed to model all its molecular components and how they interact (Karr *et al.*, 2012). Their meta-model describes the whole life-cycle of the bacterium, taking into account every known gene product, and displays various experimentally observed cellular behaviors. Two quorum-sensing systems in *Pseudomonas aeruginosa*, *las* and *rhl* (Dockery and Keener, 2001), were also modeled, revealing the existence of two stable steady points in this regulatory system.

Fewer models exist in the narrower field of competence regulation. The regulatory network of competence of Bsu has been more extensively modeled than in streptococci. In 2005, Maamar and Dubnau showed that the positive feedback loop is necessary for bistability (Maamar and Dubnau, 2005). In 2009, Leisner *et al.* developed a non-linear dynamic model which described the stochastic aspect of switching into the competence state (Leisner *et al.*, 2009). The decision-making process between competence development and sporulation has also been studied (Veening *et al.*, 2006) and reviewed (Liebal *et al.*, 2010), since sporulation is another very important sub-population differentiation program in Bsu, also relying on a quorum-sensing system, which was also mathematically described (Bischofs *et al.*, 2009; de Jong *et al.*, 2010). A simple two-dimensional model of competence development has been proposed by Nguyen *et al.* (Nguyen *et al.*, 2016).

The discovery of the ComCDE phosphorelay system, which controls competence development in the mitis and anginosus streptococcal groups, was made some twenty years ago (Havarstein *et al.*, 1995). However, very few mathematical models have been built in order to unravel the complexity of this regulatory network. The ComRS system which is present in all the other streptococcal groups (Fontaine *et al.*, 2015) has only been recently discovered (Fontaine *et al.*, 2010a). While conceptual models (see section 3.e.i.) of the ComRS system have been thoroughly discussed, very few studies used an actual mathematical formalism. The existing competence-related mathematical models in streptococci are presented here below.

I. Model of the ComCDE system in *Spn*

The regulatory network of competence development in the model species *Spn* has been mathematically modeled (Karlsson *et al.*, 2007). In their work of 2006, Karlsson *et al.* used the formalism of ODE with the quasi-steady state assumption for mRNAs and Michaelis-Menten dynamics in order to integrate all their prior knowledge about the system in their model in a transparent way.

The analysis of the output of their model, compared to experimental data, suggested that the mechanism responsible for the shutdown of competence in *Spn* was likely to rely on a *comX*-induced repressor, which could act on the expression of both the *comCDE* operon and *comX*. This suggestion was later confirmed by the work of Weng *et al.* (Weng *et al.*, 2013), which showed that the shut-off actually relies on three different levels, among which a late gene product repressing σ^X (Piotrowski *et al.*, 2009).

In brief, they first built an initial model which did not include nor *comX* expression, nor the putative repressor. Their initial parameter set was chosen so that hysteresis occurs for the CSP concentration. The term hysteresis, well known in physics, is used to describe the lag between making a change and the response or effect of that change. In simple words, it means here that the state of the system (*i.e.*, the amount of activated ComE) does not only depend on CSP concentration in the cell, but also on the past history of the system. The bifurcation analysis they performed on their initial model revealed the existence of a saddle-node bifurcation⁵ upon the CSP parameter, the control parameter of their initial model (see **Figure 24**).

⁵ A saddle-node bifurcation occurs when two fixed points (two equilibria of the system) collide and annihilate each other.

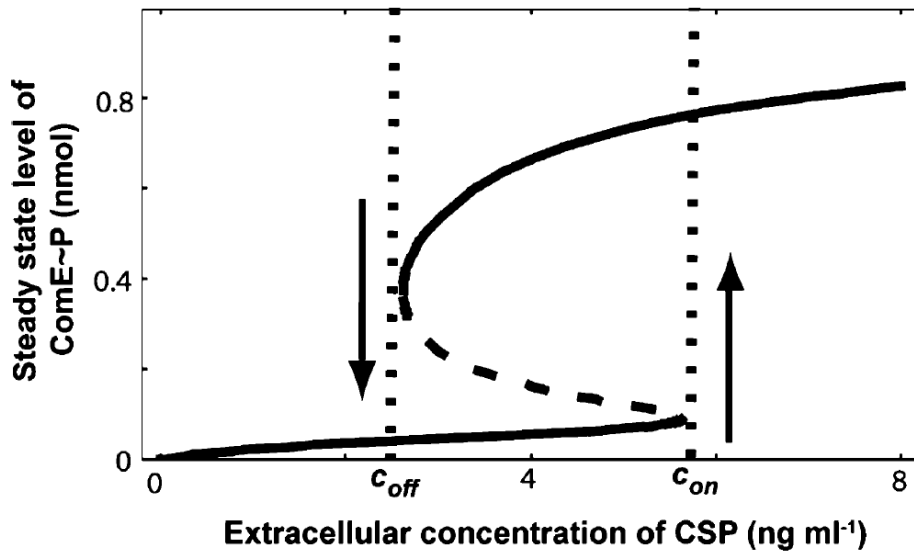


Figure 24. Steady states (amount of ComE~P, activated regulator) of the ComCDE system. The CSP concentration has to be higher than a threshold value C_{on} in order to activate the system (high values of ComE~P), but in order to shut-down, the CSP concentration has to be lower than C_{off} , with $C_{off} < C_{on}$, creating hysteresis. The plain lines represent the stable states, the dashed lines represent the unstable states (figure from (Karlsson *et al.*, 2007)).

A second model was then built by integrating an equation for *comX* and for the putative repressor. This extension of the model did not change the output from a qualitative point of view, and the hysteresis was similar. The CSP concentration was also modelled, and the control parameter was then chosen to be the cellular density. An elasticity analysis was performed in order to determine which parameters were the most sensitive and hence were the most likely to be responsible for the shut-off, since evolution is more likely to have selected the most effective mechanism in order to reduce the cost for the cell. Their elasticity analysis suggested that the shut-off mechanism was likely to be implemented by a late gene product repressing the *comCDE* operon and the *comX* expression.

Moreover, the extension of the initial model allowed it to predict successive waves of competence, as observed experimentally for *Spn* (Chen and Morrison, 1987; Morrison, 1997).

II. Model of the BlpCHR/ComRS system in Smu

The team of Son *et al.* has mathematically investigated the bistability phenomenon occurring in Smu. Based on the experimental observation that the growth medium composition determined the uni- or bimodality of competence development, they built up a simple ODE model in order to verify if the bimodality occurring in rich media could stem from the modulation of the feedback gain of the positive feedback loop acting on the ComRS system. Their *in silico* results were supporting the medium-dependent distribution of competence development inside the population of cells.

III. What is missing ? What do we want ?

Even though the BlpCHR, ComCDE and ComRS systems have been thoroughly discussed from a conceptual point of view, very few mathematical models were used to study these networks from the quantitative and qualitative points of view. In particular, no formal and detailed model integrating all the prior knowledge available about the ComRS system had been built.

Part II

Objectives of the thesis

As presented in this introduction, competence for natural DNA transformation is a widely occurring HGT mechanism, representing a major driving force for bacterial evolution. The onset of the transformation window is finely regulated. The ComRS system, controlling competence initiation in pyogenes, salivarius, mutans and bovis streptococcal groups, has been initially discovered in our laboratory (Fontaine *et al.*, 2010a). The functioning of the system in *S. thermophilus* has then been studied from a mechanistic point of view, through classical approaches of molecular biology (Fontaine *et al.*, 2013; Gardan *et al.*, 2013). The system was shown to behave like a "timing device", exhibiting remarkable kinetic patterns over the short time-window of competence development. Moreover, a series of distal regulators, mainly TCS, were known to influence the ComRS system in Sth (L. Fontaine, unpublished results).

However, many questions regarding the ComRS system remained open at the beginning of this thesis work.

First, all the **mechanistic hypotheses** of the "conceptual" model about the ComRS system, drawn on the basis of experimental data, had to be confirmed. For instance, the topology of the network had to be validated, and the level at which the different distal regulators are acting had to be determined.

Second, the **dynamics** of the system deserved further investigations, in order to explain the specific kinetic patterns observed during competence development.

Third, while the ComRS system was known to be the proximal determinant of the initiation of competence, the **shut-off mechanism** was totally unknown. The only cue regarding the exit of competence in *S. thermophilus* was the involvement of one or more late competence gene(s).

Finally, the **variability** in competence development between individuals of a monoclonal population, and between Sth strains, was poorly understood.

By combining molecular genetics (collecting experimental evidences) and mathematical modeling approaches (performing *in silico* analyses), this thesis aimed at:

- i) confirming our conceptual working hypotheses about the ComRS system functioning, concerning:
 - a. the dynamics of the system: reproduce the different experimentally observed time-courses profiles and verify their time-course succession ;
 - b. its topological hierarchy: verifying the structure (*i.e.*, the incidence matrix of the system, which is determined by all the relations between all the elements) of our conceptual model ;

- c. the mechanisms of interaction between its components: each interaction is translated into one equation term, which has to represent the physical type of interaction (sequestration, dimerization, and so on...) by its mathematical form (Hill or Michaëlis-Menten, power law...)
- ii) shedding light on the shut-off mechanism of competence.

To achieve these goals, the strategy will be to use molecular genetics tools to produce new experimental data in order to substantiate a mechanistic model which would integrate all the prior information at our disposal in a transparent and adaptable way, in order to allow us to test diverse hypotheses about different aspects of the biological system, and to infer new information about the complex ComRS regulatory cascade.

At the experimental level, the first step will consist in creating a series of mutant strains (both in LMG18311 and LMD-9 backgrounds) in a systematic way, in order to get the most information about the network (Bansal *et al.*, 2007). The genetic changes will be directed in order to either delete each of the genes of interest (resulting in knock-out (KO) strains), or to increase their expression (resulting in overexpression strains). In order to be measurable, these mutations will be performed in reporter strains, which possess a luciferase transcriptional fusion under the control of the promoter of interest.

The second step will consist in data generation, which will mainly concern two population-scaled aspects: cellular growth estimates of the different strains via the real-time measurement of the varying optical density, and estimates of expression profiles of the genes of interest via real-time measurement of the light emission produced by the luciferase fusions in the corresponding strain.

At the mathematical level, the first step will be to create a validated dynamical model of the ComRS system, by means of basic machine learning procedures. Once its robustness will be assessed, the second step will consist in challenging the model by testing different scenarios in order to i) confirm our *a priori* hypotheses, and ii) extract new, useful information about the different pre-cited shadow zones of knowledge about the system.

Part III

Results

The results section is divided into three chapters.

To improve the in-depth understanding and readability of our work, it is worth mentioning that the three chapters are not presented in their chronological order. Indeed, the more recent article (published in “Frontiers”), introducing our final and best-fitting model of the ComRS system, is here presented first, in order to introduce our modeling approach. The second chapter describes a general methodology for stability analysis of regulatory circuits. In fact, this article was published in the “MATHMOD” proceedings before our Frontiers manuscript. Finally, the last chapter of this section describes the application of the stability analysis methodology to the initial ComRS model that was built in the beginning of our thesis.

The main differences between our initial and final ComRS models lies in the mathematical formalism used for the equations. First, at the beginning of the modeling process, the equations of our model included time delays in order to take the transcription and translation steps duration into account. This formalism is called Delay Differential Equations (DDE). Second, all the activation terms of equations were modeled with Hill-functions (instead of Michaelis-Menten kinetics in the final model). These two upfront modeling choices were made for the sake of generality and adaptability. However, the learning procedure later revealed that they were not adding significant value to the model quality, while increasing its complexity. Consequently, time delays and Hill-functions were removed, hence simplifying the equations, and yielding our final model, which is presented here in the first chapter of the results section.

The second chapter presents two theoretical approaches that can be used to analyze the stability of the steady points of any genetic regulatory network, such as the ComRS regulation system. This methodology was presented at the MATHMOD Conference, and published as a conference proceeding. This article was thus published before the simplification of our model (removal of time delays and Hill coefficients). The methodology presented in chapter two was thus designed for Delay Differential Equations, which are a more general case than the Ordinary Differential Equations. As discussed in the article, the method is also valid for very general forms of activation functions (including Hill-functions), making it really adaptable to a variety of models.

The last chapter is a direct application of the methodology presented in chapter two to our initial model of the ComRS system (including time delays).

CHAPTER I

Mathematical modeling of the ComRS system reveals the limiting factors controlling competence in *Streptococcus thermophilus*

The following chapter presents the mathematical model that was built to investigate the dynamics and structure of the ComRS regulatory system.

The first step was the model construction, based on all a priori knowledge we had about the network. The learning step was performed by using an original forth-and-back approach.

As very little experimental data was available regarding the fluctuation of mRNA levels inside the cells, we decided to focus our model on the protein levels only.

We then present the validation procedure used to assess the model quality. Finally, we present the main results of the model, which helped us to infer new information about the system topology and limiting factors.

Modeling of the ComRS Signaling Pathway Reveals the Limiting Factors Controlling Competence in *Streptococcus thermophilus*

Laurie Haustenne¹, Georges Bastin², Pascal Hols^{1†} and Laetitia Fontaine^{1*†}

¹ Biochimie, Biophysique et Génétique des Microorganismes, Institut des Sciences de la Vie, Université catholique de Louvain, Louvain-la-Neuve, Belgium, ² Center for Systems Engineering and Applied Mechanics, ICTEAM, Université catholique de Louvain, Louvain-la-Neuve, Belgium

OPEN ACCESS

Edited by:

Eugénie Huillet,
Institut National de la Recherche
Agronomique, France

Reviewed by:

Lilia Macovei,
The Forsyth Institute, USA
Michael J. Federle,

In streptococci, entry in competence is dictated by ComX abundance. In *Streptococcus thermophilus*, production of ComX is transient and tightly regulated during growth: it is positively regulated by the cell-cell communication system ComRS during the activation phase and negatively regulated during the shut-off phase by unidentified late competence gene(s). Interestingly, most *S. thermophilus* strains are not or weakly transformable in permissive growth conditions (i.e., chemically defined medium, CDM), suggesting that some players of the ComRS regulatory pathway are limiting. Here, we combined mathematical modeling and experimental approaches to identify the components of

University of Illinois at Chicago, USA
 Pierre Nicolas,
 Institut National de la Recherche
 Agronomique, France

***Correspondence:**
 Laetitia Fontaine
 laetitia.fontaine@uclouvain.be

† These authors have contributed
 equally to this work.

Specialty section:
 This article was submitted to
 Antimicrobials, Resistance and
 Chemotherapy,
 a section of the journal
 Frontiers in Microbiology

Received: 21 August 2015
Accepted: 27 November 2015
Published: 22 December 2015

Citation:
 Haustienne L, Bastin G, Hols P and
 Fontaine L (2015) Modeling of the
 ComRS Signaling Pathway Reveals
 the Limiting Factors Controlling
 Competence in *Streptococcus*
thermophilus.
 Front. Microbiol. 6:1413.
 doi: 10.3389/fmicb.2015.01413

the ComRS system which are critical for both dynamics and amplitude of ComX production in *S. thermophilus*. We built a deterministic, population-scaled model of the time-course regulation of specific ComX production in CDM growth conditions. Strains LMD-9 and LMG18311 were respectively selected as representative of highly and weakly transformable strains. Results from *in silico* simulations and *in vivo* luciferase activities show that ComR concentration is the main limiting factor for the level of comX expression and controls the kinetics of spontaneous competence induction in strain LMD-9. In addition, the model predicts that the poor transformability of strain LMG18311 results from a 10-fold lower comR expression level compared to strain LMD-9. In agreement, comR overexpression in both strains was shown to induce higher competence levels with deregulated kinetics patterns during growth. In conclusion, we propose that the level of ComR production is one important factor that could explain competence heterogeneity among *S. thermophilus* strains.

Keywords: timing device, feedback regulation, competence shut-off, ordinary differential equations, kinetics model, oscillation

INTRODUCTION

Genetic plasticity is at the core of the adaptation and evolution strategies in bacteria. Horizontal gene transfer (HGT) events allow individuals to rapidly acquire/lose phenotypic traits and have a major impact on bacterial evolution. Conjugation, transduction, and competence are the main HGT mechanisms, and define genetic exchange communities (Skippington and Ragan, 2011).

Competence is defined as a transient physiological state that enables bacteria and archaea to take up exogenous naked DNA and to stably integrate it in their genome by homologous recombination (Johnsborg et al., 2007; Seitz and Blokesch, 2013). The transforming DNA can be passively or actively released from surrounding dead cells or siblings via a programmed cell lysis mechanism promoted by competent cells, respectively (Claverys et al., 2007; Berg et al., 2012; Wei and Havarstein, 2012; Borgeaud et al., 2015). Competence is increasingly viewed as a stress response that may increase adaptability and fitness in hostile conditions (Claverys et al., 2006; Perry et al., 2009; Charpentier et al., 2012; Dufour and Levesque, 2013).

DNA transformation is an energy-consuming process with possible deleterious effects on the population fitness due, for instance, to the acquisition of deleterious mutations and a slower growth rate of competent cells (Hajjema et al., 2001; Johnsen et al., 2009; Moradigaravand and Engelstadter, 2013). To optimize the benefits vs. cost of natural transformation, bacteria have implemented complex and interconnected regulatory systems to turn on competence only when particular environmental and physiological parameters reach a critical point (for recent reviews, see Seitz and Blokesch, 2013; Johnston et al., 2014; Fontaine et al., 2015). For instance, streptococci use cell-cell communication systems to synchronize competence induction between individuals, allowing them to behave such as a multicellular organism. This coordination relies on the production and sensing of peptidic competence pheromones. Their production and secretion is generally

initiated in response to specific environmental stresses (Claverys et al., 2006; Fontaine et al., 2015). At a critical extracellular concentration, competence pheromones activate the production of the master competence regulator in streptococci, ComX. ComX is a sigma factor (σ^X) that associates to the RNA polymerase to redirect transcription toward genes required for DNA transport and processing (also named late *com* genes). Competence pheromones also activate a positive feedback loop on their production that is thought to be necessary to coordinate and maintain competence induction in the population. Two main communication systems have been reported to directly control *comX* expression. The ComCDE three-component system is prevalent in species belonging to the mitis and anginosus taxonomic groups such as *Streptococcus pneumoniae* and *Streptococcus anginosus*, respectively. This system will not be extensively detailed here (For a recent review, see Fontaine et al., 2015). Briefly, the secreted competence pheromones (encoded by *comC*) are detected in the medium by the histidine kinase ComD, which transmits the information by phosphorylation to ComE. ComE~P then activates the transcription of *comX* and *comCDE*, thereby activating a positive feedback loop. The second system, named ComRS, is present in species belonging to the salivarius, mutans, pyogenic, suis, and bovis groups. Together with *comX*, the encoding genes of both ComCDE and ComRS communication systems belong to the so-called early *com* genes (Fontaine et al., 2015).

Growth in a chemically defined medium (CDM) devoid of complex oligopeptides seems to be optimal for ComRS activation (Mashburn-Warren et al., 2010; Desai et al., 2012; Son et al.,

2012; Wenderska et al., 2012; Gardan et al., 2013; Morrison et al., 2013; Zaccaria et al., 2014). The best studied models are *Streptococcus mutans* (mutans group) and *Streptococcus thermophilus* (salivarius group) (Figure 1). In CDM, the ComS precursor is produced and secreted by as yet undiscovered transporter(s). During or after its secretion, ComS is matured by species-specific protease(s) (Khan et al., 2012; Gardan et al., 2013) to release the C-terminal pheromone domain named XIP (for σ^X inducing peptide; ComS^{ext} in Figure 1) (minimum active size is between 7 and 11 aa, see Fontaine et al., 2015). At a critical extracellular concentration, XIP peptides are re-imported into the cytoplasm through the Opp/Ami transporter (Gardan et al., 2009, 2013; Mashburn-Warren et al., 2010; Fontaine et al., 2010a). XIP then interacts with the transcriptional activator ComR and stimulates the binding of the ComR-XIP complex as a dimer to the ComR-box located in the promoter of *comX* and *comS*, thereby promoting their transcription. The ComR-XIP complex induces a positive feedback loop on ComS production, but not on ComR. In agreement, the *comR* promoter is devoid of any ComR-box (Fontaine et al., 2013). The control of ComR production is believed to be the target of distal regulatory systems in response to environmental cues (Mashburn-Warren et al., 2010; Fontaine et al., 2013, 2015). In *S. mutans*, distal *comR* regulators are the BsrRM-HdrRM relay in response to cell density (Okinaga et al., 2010; Xie et al., 2010) and the ScnRK two-component system involved in oxidative stress resistance (Kim et al., 2013).

The shut-off mechanism of competence in ComRS-containing streptococci has been poorly investigated. In *S. thermophilus*,

deletion of *comX* increases the level and duration of early *com* genes expression (e.g., *comS*) (Boutry et al., 2013), similarly to the case of *S. pneumoniae* (Mirouze et al., 2013; Weng et al., 2013). This suggests that some ComX-regulated genes are involved in this process. In *S. mutans*, active proteolysis of σ^X by the MecA-ClpCP machinery plays an important role in the kinetics of competence shut-off (Tian et al., 2013).

Kinetics of *comX* induction in CDM differs remarkably between species and strains (Desai et al., 2012; Gardan et al., 2013; Morrison et al., 2013; Zaccaria et al., 2014) and is greatly influenced by growth parameters such as pH (Guo et al., 2014; Son et al., 2015). For instance, the ComRS system of *S. thermophilus* is viewed as a timing device, rather than a true quorum-sensing system. Indeed, spontaneous expression of early *com* genes starts to increase at the early log phase, whatever the size of inocula, and lasts for about 1 h. In addition, induction of competence by XIP in *S. thermophilus* is less efficient in the lag and stationary phases (Gardan et al., 2013). This suggests that regulatory components need to be synthesized, or alternatively, degraded for optimal ComRS signaling (Fontaine et al., 2015).

As the ComRS system was recently discovered, many questions about its regulatory cascade still remain open. For instance, it is not clear how environmental parameters are integrated to fine-tune the activity of the ComRS complex. In this context, few ComRS-containing strains turn out to be spontaneously transformable in CDM, but most are responsive to artificial induction (Gardan et al., 2009; Fontaine et al., 2010a,b; Desai et al., 2012; Son et al., 2012; Morrison et al., 2013; Zaccaria et al., 2014). This suggests that some important steps of the

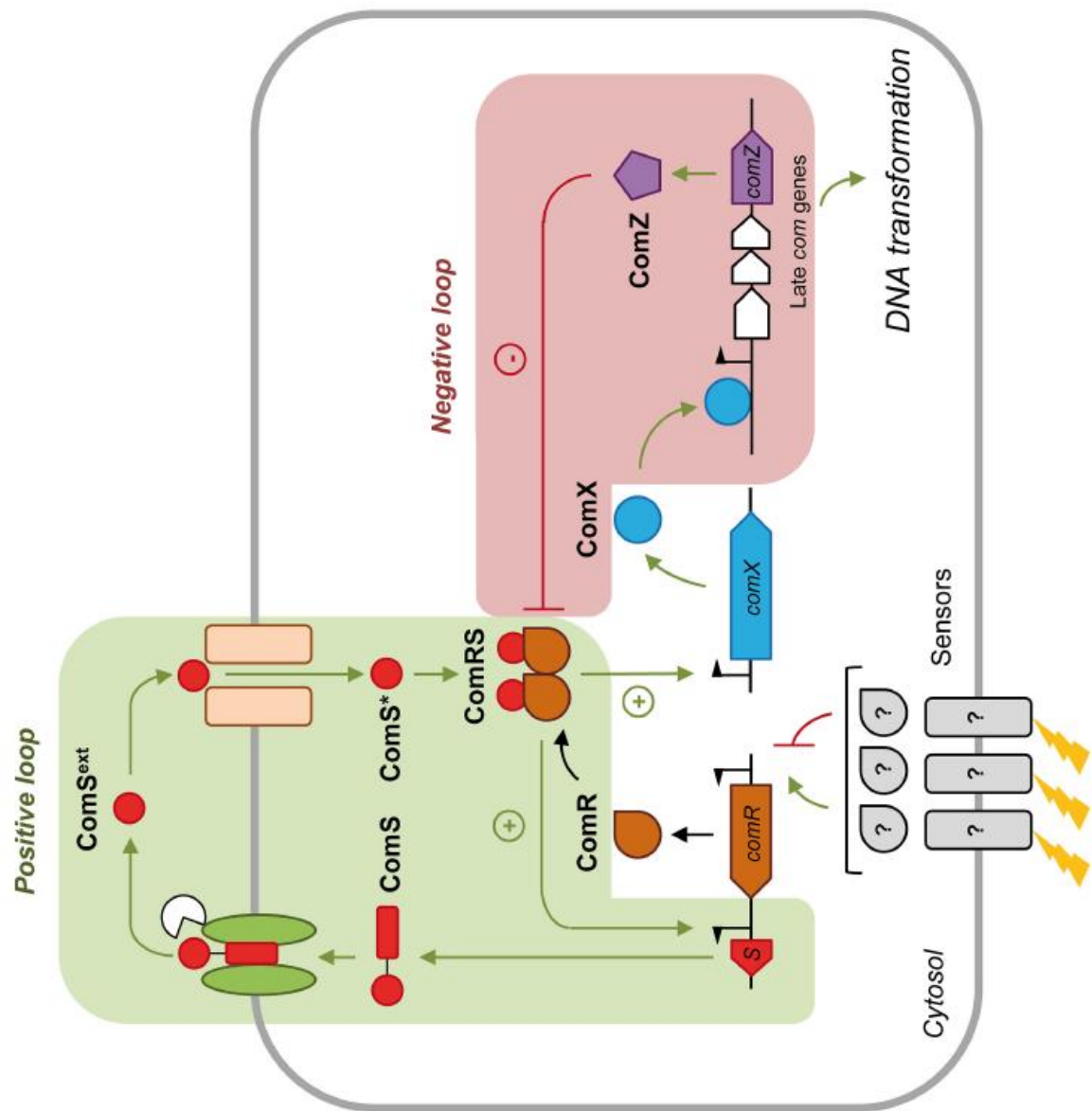


FIGURE 1 | Regulation of *comX* expression by the ComRS signal transduction system in streptococci. Competence development is initiated by the production of the precursor of the competence pheromone ComS. ComS is then exported by an unknown transporter and matured by species-specific protease(s). At a critical concentration, the mature competence peptide ComS^{ext} (also named XIP) is sensed by a specific oligopeptide-binding protein OppA/AmiA and reimported through the Opp/Ami transporter. In the cytoplasm, the intracellular mature peptide ComS* interacts with the transcriptional regulator ComR. The resulting dimeric ComRS complex activates the transcription of early competence genes (green arrows) including *comS* (plus sign; positive feedback loop) and *comX* by binding to their promoter. Expression of *comR* is not activated by ComRS and is thought to be regulated by other regulators/sensors (?) in response to specific environmental cues (represented by lightning). ComX or σ^X induces transcription of late competence genes necessary for DNA transformation. Competence shut-off (red lines) in *S. thermophilus* involves at least one or more late gene product(s), named here ComZ. ComZ is thought to interact with and inactivate the ComRS complex (minus sign; negative feedback loop). The names of modeled peptidic/proteic species in system S1 are in black bold.

regulatory cascade may either be limiting or missing in these growth conditions, which are far from those encountered in natural habitats.

The aim of this study is to deeply re-examine the ComRS regulatory cascade and identify the critical components for ComX production in species belonging to the salivarius group, using a combination of mathematical modeling and experimental molecular approaches. Based on the previous knowledge acquired on the ComRS regulatory cascade, *S. thermophilus* was chosen as a model bacterium. Strains LMD-9 and LMG18311 were respectively selected as representative of highly and weakly transformable strains (Fontaine et al., 2010b). A dynamic and mechanistic model of the ComRS system of strain LMD-9, partially calibrated by experimental data, was first built. This model, that integrates growth parameters, was then used to study qualitative and quantitative aspects (topology and dynamics)

of the ComRS system in order to identify the regulatory determinants of competence development in *S. thermophilus*.

MATERIALS AND METHODS

Experimental Procedures

Bacterial Strains, Plasmids, and Growth Conditions
The bacterial strains and plasmids used in the present study are listed in Supplementary Table S1. Plasmids derived from pGICB004 were specifically constructed in *Escherichia coli* EC1000. *E. coli* was grown in LB medium with shaking at 37°C (Sambrook et al., 1989). *S. thermophilus* was grown at 37°C in M17 broth (Difco Laboratories Inc., Detroit, MI) or in CDM, as described by Letort and Juillard (2001). All synthetic media contain 1% (wt vol⁻¹) lactose (M17L and CDML broth,

respectively). When required, spectinomycin ($200\ \mu\text{g ml}^{-1}$ for *E. coli* and *S. thermophilus*), erythromycin ($250\ \mu\text{g ml}^{-1}$ for *E. coli*, $2.5\ \mu\text{g ml}^{-1}$ for *S. thermophilus*) or chloramphenicol ($20\ \mu\text{g ml}^{-1}$ for *E. coli*, $5\ \mu\text{g ml}^{-1}$ for *S. thermophilus*) was added to the media. Solid agar plates were prepared by adding 2% (wt vol⁻¹) agar to the medium. Solid plates inoculated with *S. thermophilus* cells are incubated anaerobically (BBL GasPak systems, Becton Dickinson, Franklin lakes, NJ) at 37°C .

Detection of Absorbance and Luminescence

Small volumes ($300\ \mu\text{l}$) of culture samples (OD_{600} of 0.05) were incubated in the wells of a sterile covered white microplate with a transparent bottom (Greiner, Alphen a/d Rijn, The Netherlands). Growth (OD_{600}) and luciferase (Lux) activity (expressed in relative light unit; RLU) were monitored at 10-min intervals during 5 h in a Varioskan Flash multi-mode reader (ThermoFisher Scientific, Zellik, Belgium) as previously described (Fontaine et al., 2010a, 2013). In the supplementation experiments, different concentrations of synthetic forms of ComS₁₇₋₂₄ (LPYFAGCL) (purity >95%) supplied by Peptide 2.0 (Chantilly, VA) were added to the $300\ \mu\text{l}$ -culture samples after 1 h of growth at 37°C .

Natural Transformation Experiments

Experiments were performed as previously described (Fontaine et al., 2010a,b, 2013). The DNA, either pGIUD0855cat ($1\ \mu\text{g}$), pGILFspec::P_{comR} ($1\ \mu\text{g}$) or purified overlap PCR products ($25\ \text{ng}$), were added to $300\ \mu\text{l}$ culture samples. The transformation frequency was calculated as the number of antibiotic-resistant

CFUs per ml divided by the total number of viable CFUs per ml. Antibiotic-resistant CFUs were selected on chloramphenicol in case the transforming DNA is pGIUD0855cat or a P32-cat-encompassing PCR product, or on spectinomycin when the transforming DNA encodes a P_{spec-spec} cassette (plasmid pGILFspec::P_{comR}). Integration of the antibiotic resistance cassette at the right location in the chromosome of transformants was checked by PCR (primer pairs used are listed in Supplementary Table S2).

DNA Techniques and Electrotransformation

General molecular biology techniques were performed according to the instructions given by Sambrook et al. (1989). *S. thermophilus* LMD-9 and *S. thermophilus* LMG18311 chromosomal DNA was prepared as described by Ferain et al. (1996). Electrotransformation of *E. coli* was performed as described by Dower et al. (1988). Electrocompetent *S. thermophilus* cells were prepared as previously described (Blomqvist et al., 2006). After transformation with $1\ \mu\text{g}$ of plasmid DNA, cells were immediately resuspended in $1\ \text{ml}$ M17L and incubated anaerobically during 6 h at 37°C (pMG36e derivatives) (van de Guchte et al., 1989) or 29°C (pGICB001) (Law et al., 1995). PCRs were performed with Phusion high-fidelity DNA polymerase (Finnzymes, Espoo, Finland) in a GeneAmp PCR system 2400 (Applied Biosystems, Foster City, CA). The primers used in this study were purchased from Eurogentec (Seraing, Belgium) and are listed in Supplementary Table S2.

Deletion of DprA by Natural Transformation

The DprA⁻ derivatives of LMD-9 reporter strains were constructed by exchanging the *dprA* open reading frame (ORF) (sequence between the start and stop codons) with the chloramphenicol resistance cassette *lox66-P32-cat-lox71*, as previously described (Fontaine et al., 2010b). Primers used are listed in Supplementary Table S2.

Construction of Reporter Strains

Luminescent LMG18311 reporter strains CB009 and LF146 were constructed by replacing part of the *blp* locus (from *blpU* to *blpX*) by the transcriptional fusion *P_{comX}-luxAB* and *P_{comR}-luxAB-spec*, respectively. This was achieved by transforming pGICB004 derivatives, respectively pGICB001 and pGILF::P_{comR}, which carries *blp* recombination fragments, as previously described (Fontaine et al., 2010a; Fleuchot et al., 2011). Plasmid pGILFspec::P_{comR} was obtained in two steps. First, the spectinomycin resistance cassette was cloned between *luxAB* and the *blp* recombination sequence of pGICB004 in order to easily select the desired transformants on spectinomycin. To this end, P_{spec-spec} was amplified by PCR from plasmid pR412 with primer pairs LoxSpec-SmaI-F/LoxSpec-Pvu-R, digested by *SmaI* and *PvuII* and cloned in pGICB004 linearized by *SmaI*, yielding pGILFspec. Second, a PCR product corresponding to the P_{comR} sequence from LMG18311 and digested by *EcoRI* and *SpeI* was cloned in similarly digested plasmid pGILFspec (see Supplementary Tables S1, S2).

EMSA Experiments

The double-stranded DNA fragment Cy3-P_{comS} used in the EMSA experiments (approximately 200 bp) was amplified by PCR from *S. thermophilus* LMD-9 (Supplementary Table S2). The 5' end of the forward primer used was coupled to the Alexa 555 fluorophore. ComR-Strep from *S. thermophilus* LMD-9 was purified as previously described (Fontaine et al., 2013). Typically, a gel shift reaction (20 µl) was performed in a binding buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 10% glycerol, 1 mg ml⁻¹ BSA) and contained 150 ng labeled probe and 4 µM ComR-Strep proteins. When necessary, 0.2 or 2 µM of ComS₁₇₋₂₄ peptides are added. The reaction is incubated at 37°C for 10 min prior to loading of the samples on a native TBE 5% gel. The gel is next subjected to 80 V for approximately 1 h in TBE buffer. DNA complexes were detected by fluorescence on the Ettan DIGE Imager with bandpass excitation filters (nm) of 540/25 (Cy3) and bandpass emission filters of 595/25 (Cy3) (GE Healthcare, Waukesha, WI).

Mathematical Modeling

Our model describes the mean behavior of a competent cell in a homogenous *S. thermophilus* LMD-9 population grown in CDML conditions. The model focuses on the timing device function of the ComRS system, solely. In this context, ComS exportation/maturation and importation are modeled to create a time delay in the signaling cascade. The choice of not modeling the cell-cell communication aspect of ComS in *S. thermophilus* is here justified by experimental evidences that (i) secreted mature ComS peptides (XIP) do not freely diffuse in CDML medium

and do rather remain associated to the surface of secreting cells (Gardan et al., 2013) and that (ii) the competence cycle (on- and off-steps) takes place at low cell densities (at least in CDML growth conditions) (Gardan et al., 2009, 2013; Fontaine et al., 2010a).

System S1 is composed of seven deterministic ordinary differential equations (ODE) describing the time-course evolution of the seven proteic/peptidic species involved in the (transcriptional) control of ComX production: precursor peptides ComS, secreted peptides ComS^{ext}, intracellular mature peptides ComS*, ComR, active complexes ComRS, ComX, and ComZ. These variables are expressed in number of molecules per cell (Table 1). In addition, an equation representing the time-evolution of the cell density $X(t)$ is added (Equation 1, expressed in number of cells per ml), since the development of competence is tightly coupled to the growth process. ComZ is a hypothetical actor that represents the protein(s) of the late phase responsible for competence shut-off. The description of variables and parameters of all equations is presented in Tables 1, 2, respectively.

According to the structure of the regulatory network illustrated in Figure 1, the production of some variables (e.g., ComR, ComS, ComX) is controlled by a basal term $b(t)$ and activation and/or repression terms. The time-course evolution of all variables is also negatively affected by a degradation term, which depends on a specific degradation rate d (we assume first-order degradation kinetics), and a dilution term which accounts

for cell division and is thus directly related to the time-dependent growth rate $\mu(t)$.

ODE Equations of System S1

$$\frac{dX(t)}{dt} = \mu(t)X(t) \quad (1)$$

Equation (1) models the growth kinetics of LMD-9 cells in CDML conditions. It allows to model competence development during the different growth phases (lag, exponential, and stationary). The time-varying cell density $X(t)$ (expressed in cell ml⁻¹) was computed from a mean experimental growth curve. To convert experimental OD₆₀₀ units in number of cells per ml, we assumed

TABLE 1 | Description of variables in system S1.

Variable	Description	Unit*
$X(t)$	Cellular density	cell ml ⁻¹
$ComS(t)$	Concentration of precursor peptide ComS (24 aa)	mol. cell ⁻¹
$ComS^{ext}(t)$	Concentration of external ComS peptides (mature pheromone)	mol. cell ⁻¹
$ComS^*(t)$	Concentration of reimported mature ComS	mol. cell ⁻¹
$ComR(t)$	Concentration of ComR regulator	mol. cell ⁻¹
$ComRS(t)$	Concentration of active ComRS complex	mol. cell ⁻¹
$ComX(t)$	Concentration of alternative sigma factor ComX	mol. cell ⁻¹
$ComZ(t)$	Concentration of putative competence repressor ComZ	mol. cell ⁻¹

*mol., molecules.

that a OD_{600} unit of $1 = 5 \times 10^8$ cells ml^{-1} . This conversion was deduced from a plating and CFU (colony forming unit) counting experiment (data not shown). The varying growth rate $\mu(t)$ (expressed in min^{-1}) was then computed from $X(t)$ according to Equation (1) (see Supplementary Figure S1A)

$$\frac{dComS(t)}{dt} = b_s(t) + V_{max}^S \underbrace{\left(\frac{ComRS(t)}{K_m^S + ComRS(t)} \right)}_{\text{activation by ComRS}} - \underbrace{r_{out}ComS(t)}_{\text{exportation}} - \underbrace{\mu(t)ComS(t)}_{\text{growth}} \quad (2)$$

Equation (2) describes how the specific number (i.e., per cell) of cytoplasmic precursor ComS peptides (Figure 1) varies with time. ComS is produced in the cell at a non-constant basal rate $b_s(t)$, and its production is activated by the ComRS complex according to a Michaelis-Menten function, as modeled in Son et al. (2012). This activation term represents the positive feedback loop which activates and maintains ComS production. The term “exportation” on the right side represents the secretion of ComS in the extracellular medium. The latter is modeled as a non-saturating step: it depends on a constant exportation rate r_{out} and increases directly with $ComS(t)$ concentration. ComS maturation was also assumed to be concomitant with exportation. The last term models ComS dilution due to cell division. The degradation rate of the cytoplasmic precursor ComS was assumed to be null, reflecting the hypothesis that ComS peptides are directly secreted into the extracellular medium upon synthesis.

$$\frac{dComS^{ext}(t)}{dt} = \underbrace{r_{out}ComS(t)}_{\text{exportation}} - \underbrace{r_{in}ComS^{ext}(t)}_{\text{importation}} - \underbrace{d_{sex}ComS^{ext}(t)}_{\text{degradation}} - \underbrace{\mu(t)ComS^{ext}(t)}_{\text{growth}} \quad (3)$$

Equation (3) describes the time-course evolution of number of $ComS^{ext}$, the secreted mature form of ComS, referred to as XIP (Mashburn-Warren et al., 2010). As noted above, $ComS^{ext}$ peptides were assumed to remain closely associated to the cell surface, as evidenced in the literature (Gardan et al., 2013). This hypothesis is compatible with the timing-device mechanism of competence regulation in *S. thermophilus* proposed by Gardan et al. (2013). Expressing $ComS^{ext}$ values in molecules $cell^{-1}$ is thus relevant in this species. Nevertheless, it is important to note that the specific units (molecules $cell^{-1}$) can be directly converted in volumetric units (molecules ml^{-1}) at any time t based on the time-dependent value of $X(t)$ (expressed in number of cells ml^{-1}) (see Equation 1). In Equation 3, $ComS^{ext}$ concentration at the cell surface is positively affected by the rate of ComS exportation (same term as above) and negatively affected by the rates of (i) mature $ComS^{ext}$ re-importation in the cytoplasm, (ii) $ComS^{ext}$ degradation in the extracellular medium and (iii) cell division (related to Equation 1). Re-importation of $ComS^{ext}$ through the Opp system was also assumed to depend on a first-order importation constant r_{in} . This simplification can be justified by the fact that in CDM conditions, the Opp transporter

TABLE 2 | Description and numerical values of parameters in system S1.

Symbol	Description	Cell-level unit	Standard unit*
r_{out}	Exportation rate constant	1 min^{-1}	1 min^{-1}
r_{in}	Importation rate constant	0.01 min^{-1}	0.01 min^{-1}
n	Hill coefficient (Son et al., 2012)	2.5 (adim.)	2.5 (adim.)
c_{RS}	Complex formation rate constant	$0.1 \text{ mol.}^{-4} \text{ cell}^4 \text{ min}^{-1}$	$1.2 \times 10^{-3} \text{ nM}^{-4} \text{ min}^{-1}$
V_{max}^Z	Maximal activation rate of ComZ by ComX	$1 \text{ mol. cell}^{-1} \text{ min}^{-1}$	3 nM min^{-1}
n_Z	Hill coefficient	3 (adim.)	3 (adim.)
K_m^Z	Required concentration of ComX for half-maximum synthesis rate of ComZ	$15 \text{ mol. cell}^{-1}$	45 nM
V_{max}^{RSZ}	Maximal inactivation rate of ComRS by ComZ	$0.04 \text{ mol. cell}^{-1} \text{ min}^{-1}$	0.12 nM min^{-1}
K_m^{RSZ}	Required concentration of ComRSComZ for half-maximum inactivation rate of ComRS	$0.5 \text{ mol.}^2 \text{ cell}^{-2}$	4.5 nM^2
d_{Sext}	Degradation rate constant of ComS ^{ext}	0.07 min^{-1}	0.07 min^{-1}
d_{S*}	Degradation rate constant of ComS*	0.07 min^{-1}	0.07 min^{-1}
d_R	Degradation rate constant of ComR	0.01 min^{-1}	0.01 min^{-1}
d_{RS}	Degradation rate constant of ComRS	0.01 min^{-1}	0.01 min^{-1}
d_X	Degradation rate constant of ComX (Karlsson et al., 2007)	0.2 min^{-1}	0.2 min^{-1}
d_Z	Degradation rate constant of ComZ	0.01 min^{-1}	0.01 min^{-1}
V_{max}^S	Maximal activation rate of ComS by ComRS	$4 \text{ mol. cell}^{-1} \text{ min}^{-1}$	12 nM min^{-1}
K_m^S	Required concentration of ComRS for half-maximum synthesis rate of ComS	$0.5 \text{ mol. cell}^{-1}$	1.5 nM
A	Amplitude of the gaussian activation of ComR	$27 \text{ mol. cell}^{-1}$	81 nM
s	Standard deviation of the activation of ComR	35 min	35 min
m	Center of the gaussian activation of ComR	90 min	90 min
V_{max}^X	Maximal activation rate of ComX by ComRS	$5.3 \text{ mol. cell}^{-1} \text{ min}^{-1}$	15.9 nM min^{-1}
K_m^X	Required concentration of ComRS for half-maximum synthesis rate of ComX	1 mol. cell^{-1}	3 nM

*Units of the systems are converted in standard units as follows: 1 molecule per cell corresponds to a molar concentration of 3 nM, assuming that a *S. thermophilus* cell is spherical with a diameter of 1 μm , which corresponds to a volume of 0.5 fl. mol., molecules; adim., adimensional.

is not saturated by nutritive oligopeptides, in contrast to complex medium like M17 (Gardan et al., 2013; Son et al., 2012).

$$\frac{dComS^*(t)}{dt} = \underbrace{r_{in} ComS^{ext}(t)}_{\text{importation}} - \underbrace{c_{RS}(ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{d_S^* ComS^*(t)}_{\text{degradation}} - \underbrace{\mu(t)ComS^*(t)}_{\text{growth}} \quad (4)$$

Equation (4) describes the evolution of number of mature intracellular $ComS^*$ peptides. In this equation, the importation flux increases linearly with specific $ComS^{ext}$ concentration i.e., the number of $ComS^{ext}$ peptides that are physically associated to one cell. The mechanistic effect of cell density on the importation flux is not modeled in Equation (4) as the cell-cell signaling function of $ComS$ is not considered in the $ComRS$ model. We fit the degradation constants in order to keep the concentrations of $ComR$ and $ComS$ inside the cells close to a 1:1 ratio. The negative term $c_{RS}(ComR(t)ComS^*(t))^n$ corresponds to the formation of the $ComRS$ complex, according to the mass action law, as modeled in Son et al. (2012).

$$\frac{dComR(t)}{dt} = \underbrace{b_R(t) + N(m, \sigma, t)}_{\text{distal regulators}} - \underbrace{c_{RS}(ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{d_R ComR(t)}_{\text{degradation}} - \underbrace{\mu(t)ComR(t)}_{\text{growth}} \quad (5)$$

In Equation (5), $ComR$ is produced at a time-dependent basal rate $b_R(t)$ and its expression is activated by unknown distal regulators whose actions are summed up in deterministic Gaussian function $N(m, \sigma, t)$: the effects of all the activators/inhibitors of $ComR$ are embedded in this bell-shaped function:

$$N(m, \sigma, t) = A \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{t-m}{\sigma} \right)^2}$$

Formation of the $ComRS$ complex, as well as $ComR$ degradation and cell division, negatively influences $ComR(t)$ production.

$$\frac{dComRS(t)}{dt} = \underbrace{\frac{c_{RS}}{2} (ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{V_{max}^{RSZ} \left(\frac{ComRS(t)ComZ(t)}{K_m^{RSZ} + ComRS(t)ComZ(t)} \right)}_{\text{shut-off by ComZ}} - \underbrace{d_{RS} ComRS(t)}_{\text{degradation}} - \underbrace{\mu(t)ComRS(t)}_{\text{growth}} \quad (6)$$

Equation (6) describes the evolution of active $ComRS$ complexes. A $\frac{1}{2}$ stoichiometric coefficient was introduced to describe the formation rate of the active complex since 2 $ComS$ and 2 $ComR$ molecules are needed to produce 1 $ComRS$ molecule (Fontaine

et al., 2013). Production of ComRS is negatively influenced by its sequestration in an inactive complex, ComRSZ (for the role of ComZ, see Equation 8 below), which forms according to a Michaelis-Menten function. As mentioned above, this negative term models the negative feedback loop that operates during competence shut-off (see description of Equation 8 for detailed information).

$$\frac{d\text{ComX}(t)}{dt} = \underbrace{b_X(t) + V_{\max}^X \left(\frac{\text{ComRS}(t)}{K_m^X + \text{ComRS}(t)} \right)}_{\text{activation by ComRS}} - \underbrace{d_X \text{ComX}(t)}_{\text{degradation}} - \underbrace{\mu(t) \text{ComX}(t)}_{\text{growth}} \quad (7)$$

Equation (7) has the same structure as Equation (2): ComX is produced at a basal rate $b_X(t)$, and its expression is activated by the ComRS complex.

$$\frac{d\text{ComZ}(t)}{dt} = \underbrace{V_{\max}^Z \left(\frac{\text{ComX}(t)^{n_z}}{(K_m^Z)^{n_z} + \text{ComX}(t)^{n_z}} \right)}_{\text{activation by ComX}} - \underbrace{d_Z \text{ComZ}(t)}_{\text{degradation}} - \underbrace{\mu(t) \text{ComZ}(t)}_{\text{growth}} - \underbrace{V_{\max}^{\text{RSZ}} \left(\frac{\text{ComRS}(t) \text{ComZ}(t)}{K_m^{\text{RSZ}} + \text{ComRS}(t) \text{ComZ}(t)} \right)}_{\text{inactivation of ComRS}} \quad (8)$$

Equation (8) describes the evolution over time of the ComZ protein, which represents the putative actor(s) of the late phase responsible for competence shut-off. We made the hypothesis of

a null basal production rate for ComZ, which is assumed to be produced only in the presence of ComX. The model proposes that ComZ associates with ComRS to inactivate it (Figure 1). This hypothesis is based on the shut-off mechanism described in *S. pneumoniae* in which the late protein DprA was shown to interact with ComE~P to block ComE-driven transcription of *comX* (Mirouze et al., 2013).

These eight ODEs form the system S1. At any time t , the state $s(t)$ of the system is determined by the cell density $X(t)$, see Equation 1) and the number of each molecular species $(\xi_i(t))$, see Equations 2–8), and is defined by the following vector, evolving in the positive orthant (9):

$$s(t) = \{X(t), \xi_1(t), \xi_2(t), \dots, \xi_N(t)\} \in \mathfrak{R}_+^{N+1} \text{ with } N = 7 \quad (9)$$

Determination of Model Parameters

The determination of numerical values for the 22 parameters of system S1 was performed in a stepwise fashion, as described in details below. Briefly, system S1 was first simplified by replacing the production terms for *ComR*(t), *ComS*(t), and *ComX*(t) with lumped production terms derived from experimental measurements. This reduced system (termed S2) only contained 15 of the 22 parameters of system S1. In a second step, the 15 parameters of system S2 were determined, according to several mechanistic hypotheses and constraints. In a final step, the lumped, experiment-derived production terms for *ComR*(t), *ComS*(t), and *ComX*(t) were replaced with their mechanistic equivalents in system S1, and the remaining seven parameters were fitted.

Step 1—model simplification and inclusion of experimental data on ComS, ComR, and ComX production rates

To determine the time-dependent production rates (due to basal production rates and activation production rates) $prod(t)$ of variables $ComR(t)$, $ComS(t)$, and $ComX(t)$, we used data on the specific luciferase activities driven from CMDL cultures of LMD-9 derivative strains containing $P_{comR-luxAB}$, $P_{comS-luxAB}$, or $P_{comX-luxAB}$ reporter fusions (Fontaine et al., 2013; Supplementary Figure S1B), based on the assumption that the specific production rate of luciferase in a given reporter strain ($prod_{LuxAB}(t)$; molecules $cell^{-1} min^{-1}$) is identical to the specific production rate of the peptide or protein under control of the same promoter ($prod(t)$; molecules $cell^{-1} min^{-1}$):

$$prod(t) = prod_{LuxAB}(t) \quad (10)$$

In order to determine $prod_{LuxAB}(t)$ from the experimentally observed light emission profiles ($L(t)$; RLU OD_{600}^{-1}) of the reporter strains, the following assumptions were made:

(i) the observed light emission L was assumed to be directly proportional to the amount of luciferase $LuxAB$:

$$L(t) = kLuxAB(t) \quad (11)$$

(ii) the net production rate of luciferase $dLuxAB(t)/dt$ was assumed to be the difference between its absolute production rate ($prod_{LuxAB}(t)$), its degradation according to a constant degradation rate d_{LuxAB} and its dilution due to cell growth:

$$\frac{d_{LuxAB}(t)}{dt} = prod_{LuxAB}(t) - (d_{LuxAB}(t) + \mu(t))LuxAB(t) \quad (12)$$

Hence, by combining Equations (10–12), the production rates of ComS, ComR, and ComX can be reformulated as a function of the observed light emission as follows:

$$prod(t) = \frac{1}{k} \left(\frac{dL(t)}{dt} + d_{LuxAB}L(t) + \mu(t)L(t) \right) \quad (13)$$

For the dilution term, $\mu(t)$ was calculated from the experimental growth curve of *S. thermophilus* LMD-9 in CDML, as described in Equation (1) (Supplementary Figure S1A). The degradation rate of luciferase d_{LuxAB} was set to $0.015 min^{-1}$, as observed experimentally in *S. thermophilus* (data not shown). The value of the proportionality constant k ($2.197 \cdot 10^{-4}$ RLU molecule $^{-1}$) was chosen so as to mimic the experimentally observed light emission profiles of the three reporter strains (LMD-9 derivatives containing $P_{comS-luxAB}$, $P_{comX-luxAB}$, and $P_{comR-luxAB}$). Parameter k was thus assumed to be an intrinsic property of the luciferase $LuxAB$, independent on the promoter driving its expression. For determining k , an additional constraint was set on the system, by imposing a maximal production rate for ComS ($prod_S(t)$) of 20 molecules per minute, in accordance with the estimated maximal synthesis rate of small peptides (Bolouri and Davidson, 2003).

Step 2—determination of parameters for system S2

Since the seven parameters of the modeled production rates for $ComS(t)$, $ComS(t)$, and $ComX(t)$ are hidden behind the experimental production rates, the number of parameters to estimate in system S2 is only $m = 15$, instead of 22 in system S1. Based on mechanistic assumptions and literature information, specific values were assigned to 2 of the 15 parameters:

- (i) The Hill coefficient for formation of the ComRS complex was set to $n = 2.5$, according to Son et al. (2012), to capture the dimeric character of the complex (Fontaine et al., 2013).
- (ii) The degradation rate of ComX (d_X) was set to 0.2 min^{-1} , according to Karlsson et al. (2007). This corresponds to a half-life of around 3 min, which is in the same range as the experimental half-life estimated for ComX in *S. pneumoniae* ($\sim 4 \text{ min}$) (Luo and Morrison, 2003).

In addition, the degradation rates of the other variables were related by the following inequality:

$$0.01 \text{ min}^{-1} < d_R = d_{RS} = d_Z < d_{S^*} = d_{S_{ext}} < 0.69 \text{ min}^{-1}$$

which is the mathematical expression of imposing (i) that the half-lives of the peptides/proteins should lie between around 1 min and 1 h, and (ii) that, due to the respective molecular weights of the molecules, the degradation of small peptides ($ComS^{ext}$ and $ComS^*$) is faster than the degradation of bigger proteins. Because of their similar size (~ 10 amino acids; Gardan et al., 2013), the degradation rates of $ComS^{ext}$ and $ComS^*$ were assumed to be identical. The possible influence of $ComS^*$ on the stability of the ComRS complex was neglected, as

previously assumed by Voigt et al. (2005) for a similar protein-peptide complex. Hence, the degradation rates of ComR and ComRS were assumed to be identical. Finally, since ComZ is a hypothetical protein, its degradation rate cannot be anticipated on rational grounds: by default, it was constrained to be identical to that of ComR and ComRS.

Numerical values for the remaining ten parameters of system S2 were set with the requirement that the resulting model adequately describes the following experimentally observed behaviors:

- (a) Since we initially assumed that luciferase production in the reporter strains is proportional to the production of the corresponding peptides (see Step 1—Model Simplification and Inclusion of Experimental Data on ComS, ComR, and ComX production Rates), two qualitative requirements were set on the simulated profiles of intracellular concentrations of ComR, ComS, and ComX. First, the time sequence of the initial increase in concentration of the three peptides/proteins (Figure 2A) must qualitatively match the time sequence of the onset of light production (Supplementary Figure S1B), i.e., if luciferase activity is first detected in the $P_{comR-luxAB}$ reporter strain, then the simulated ComR accumulation must start before that of ComS and ComX. The production of the late competence protein ComZ must be delayed compared to ComX, as observed experimentally between the activation of $P_{comX-luxAB}$ and $P_{comGA-luxAB}$, another late *com* gene (data

not shown). Second, the time sequence of the maximum intracellular concentration of the three peptides/proteins (Figure 2A) (null derivative, where the production term equals the consumption, degradation, and dilution terms) must qualitatively match the time sequence of experimentally observed maximum luciferase activity in the corresponding reporter strains (Supplementary Figure S1B). When we compare Figure 2A (modeled kinetics) and Supplementary Figure S1B (experimental kinetics), we see that the two qualitative requirements are globally satisfied. The shape and amplitudes of the simulated curves must however not necessarily match those of luciferase activity, since except for the dilution term, consumption terms for ComR, ComS, and ComX differ between each other and from the degradation term of LuxAB.

- (b) It was required that in the modeled system, the concentration of ComX in simulated wild-type LMD-9 cells is not at saturation. In other words, we imposed that the maximum concentration of ComX in simulated cells can be further increased upon addition of extracellular ComS (i.e., non-zero initial values of the variable $ComS^{ext}(t)$), in agreement with experimental observations showing that natural transformation frequencies and the activity of $P_{comX-luxAB}$ in wild-type LMD-9 can be further increased by addition of the synthetic mature ComS peptide ComS₁₇₋₂₄ (Fontaine et al., 2010a and data not shown).

Step 3—determination of the remaining parameters for system S1

Using the 15 parameter values determined in the previous step, we then moved back to system S1, and replaced the experimentally-derived production rates $prod_R(t)$, $prod_S(t)$, and $prod_X(t)$ with their explicit mechanistic equivalents in Equations (5), (2), and (7) respectively. In all three cases, the mechanistic production terms are composed of a basal production rate term and an activation term.

The time-dependent basal expression rate of ComS ($b_S(t)$) was derived from an experimental dataset of luciferase activity in a CDML culture of an LMD-9 $\Delta comR$ strain containing a $P_{comS-luxAB}$ reporter fusion (Supplementary Figure S1C). Indeed, since the deletion of *comR* prevents formation of the ComRS complex, it therefore suppresses the activation of ComS production by ComRS, and the remaining luciferase activity thus represents the basal activity of promoter P_{comS} (Fontaine et al., 2013). Specific luciferase activity (RLU OD_{600}^{-1}) was converted into the basal production rate $b_S(t)$ (molecules $cell^{-1} min^{-1}$) as described in step 1 for determination of the production rates of ComR, ComS, and ComX, using the formula:

$$b_S(t) = \frac{1}{k} \left[\frac{dL_{\Delta comR}(t)}{dt} + (d_{Lux} + \mu(t))L_{\Delta comR}(t) \right] \quad (14)$$

where $L_{\Delta comR}$ is the luciferase activity of the $\Delta comR$ strain containing a $P_{comS-luxAB}$ reporter fusion. We used the same value of 2.197×10^{-4} for the proportionality constant k for ComR, ComS, and ComX. For simplicity, we also assumed that the basal expression rates of ComR and ComX are identical

to $b_S(t)$: $b_R(t) = b_X(t) = b_S(t)$. It is worth noting that these basal production functions are needed in the model to initiate the competence process from zero initial concentrations. But their amplitudes remain negligible, with respect to the activation terms, all along the culture. Therefore, the convenient approximation of taking identical functions is not a really restrictive condition here.

Numerical values were then set for the seven additional parameters of system S1, which define the activation rates of ComS (K_m^S and V_{max}^S) and ComX (K_m^X and V_{max}^X) by ComRS and the activation rate of ComR by distal regulators (m , σ , A). Besides the qualitative requirements already used for system S2 (see Step 2—determination of Parameters for System S2), we also imposed the requirement that the simulated profiles of ComR, ComS, and ComX production rates (basal rate + activation rate) (Figure 2B) are similar in shape to the corresponding luciferase production rates ($prod_L(t)$) deduced from experimental observations (Supplementary Figure S2). Constraints were also imposed on the parameter values for K_m^S , K_m^X , V_{max}^S , and V_{max}^X . Since our model does not explicitly deal with processes such as transcription and translation, K_m and V_{max} values cannot be directly compared to existing biological values of other systems. Nevertheless, we constrained those parameters to biologically relevant values, i.e., in the nM range as reported by Jabbari et al. (2011). We also imposed that K_m^S is lower than K_m^X because the binding affinity of ComR to its DNA motif (ComR-box) was previously predicted to be higher in P_{comS} compared to P_{comX} (Fontaine et al., 2013).

The resulting parameter values for the 22 parameters of system S1 are provided in Table 2.

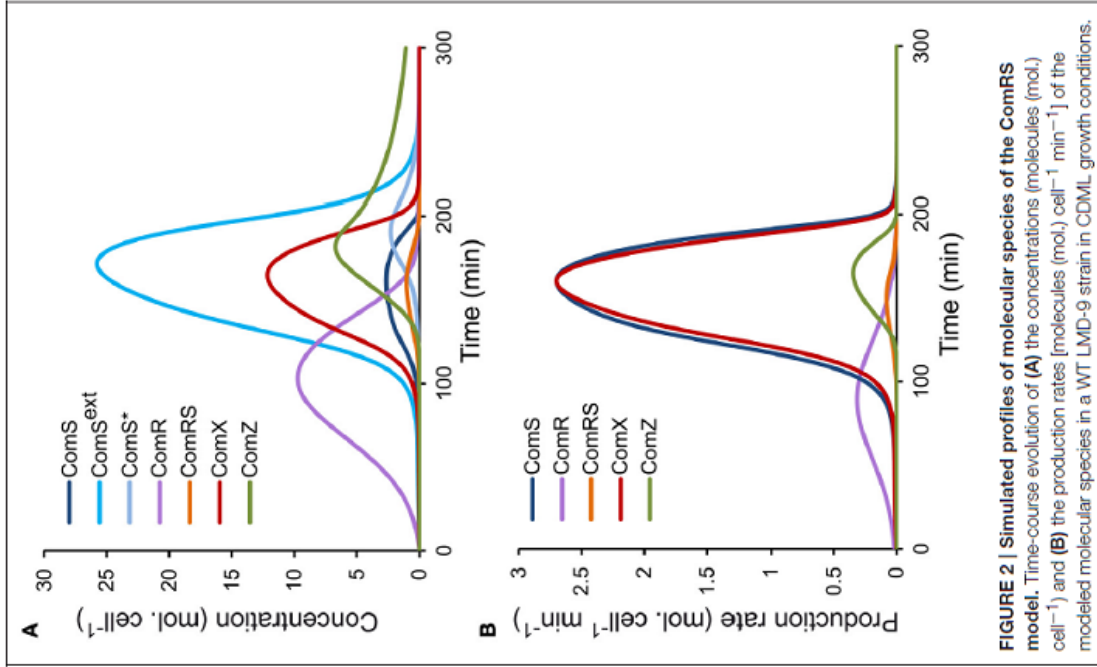


FIGURE 2 | Simulated profiles of molecular species of the ComRS model. Time-course evolution of (A) the concentrations (mol.) and (B) the production rates [mol. cell⁻¹ min⁻¹] of the modeled molecular species in a WT LMD-9 strain in CDML growth conditions.

Simulation Protocols

All computations and simulations were done using the MATLAB (R2012a) software. In particular, the ode23 routine was used for the integrations over time, with a null initial condition (except for dose-response simulations). Our full MATLAB source code is available as Supplementary Data Sheet 1. Experimental data necessary to run the code are available as Supplementary Data Sheet 2.

The mean behavior of a wild-type cell was simulated *in silico* by integrating system S1 over time (from $t = 0$ to $t = 400$ min), with the parameters listed in Table 2. A null initial condition was used: at $t = 0$, all cells were assumed to be devoid of any molecule of the modeled variables: $s(t_0) = \{X(t_0), 0, 0, 0, 0, 0, 0, 0\}$.

Simulation of knocked-out strains

To mimic a strain deleted for a gene of interest, the corresponding equation was set to zero in S1, and that modified S1 system was integrated over time with the same ode23 routine.

Simulation of dose-responses to ComS

To mimic the supplementation of ComS₁₇₋₂₄ peptides in the growth medium, the variable $ComS^{ext}(t)$ was given a non-zero initial value, between 0 and 2000 molecules per cell, at $t = 0$: $s(t_0) = \{X(t_0), 0, V_0, 0, 0, 0, 0, 0\}$.

To study the consequence of decreasing the affinity between ComR and ComS on dose-response, the complex formation constant c_{RS} was decreased by a factor ranging from 10^1 to 10^5 .

To simulate supplementation of ComS₍₁₇₋₂₄₎ at different times T of growth, the initial condition of $ComS^{ext}(t)$ was set to 0 until time T , then at time T ($T = [30, 60, 90, 120, 150, 180, 210]$ in min), an initial condition of ComS^{ext} molecules (i.e., 150 molecules per cell) was set. For each condition (various ComS^{ext} concentrations or induction times), the system S1 was integrated until $t_{end} = 400$ min with the ode23 routine and the maximal amount of ComX molecules per cell reached over time was computed.

Simulation of ComR or ComS overproduction

The different levels of ComR or ComS production were simulated by replacing their respective basal rate $b_R(t)$ and $b_S(t)$ by a constant value R : $R = \max(b(t))$ IF.

When not otherwise stated, the induction factor (IF) was set to 50.

Sensitivity analysis

Our sensitivity analysis was performed by varying the value of each parameter within an interval ranging of 0.1–10 times its assigned value (listed in Table 2) unless otherwise stated. The maximal value of ComX obtained over time in this range was then computed.

RESULTS

Modeling the Control of the Master Competence Regulator ComX By ComRS

In CDM conditions, the ComRS signaling system dominantly controls *comX* expression in *S. thermophilus* (Fontaine et al., 2013) and *S. mutans* (Mashburn-Warren et al., 2010; Son et al., 2012; Reck et al., 2015). We thus hypothesized that the modeling of the ComRS regulatory pathway would recapitulate regulation of ComX production in those growth conditions.

The LMD-9 strain of *S. thermophilus* was chosen as model strain since the competence regulating network of this strain has been extensively studied and a range of experimental data are available on the expression of key competence genes in wild type (WT) and competence-deficient backgrounds (Fontaine et al., 2010a, 2013; Gardan et al., 2013). Importantly, LMD-9 is also naturally competent in CDM, in contrast to most strains tested so far, where exogenous addition of micromolar concentrations of mature ComS peptides is necessary to reach high levels of competence (Fontaine et al., 2010b).

In CDM growth conditions, competence for natural transformation in *S. thermophilus* is a transient phenomenon. To take into account this expression dynamics, ordinary differential equations (ODEs) were used to model the time-course evolution of (known) principal molecular species involved in *comX* induction and shut-off (Figure 1). The ODE formalism is widely used in systems biology and presents several advantages: it allows to set up structural and dynamical models with an adjustable level of detail, allowing to capture various physical phenomena

in a transparent way, to use experimental time-course data, and to test different hypotheses easily (de Jong, 2002).

The modeled peptide/proteins (expressed in number of molecules per cell) are: (i) cytoplasmic prepeptides ComS; (ii) mature secreted pheromones ComS^{ext} (also named XIP); (iii) reimported mature pheromones ComS*; (iv) unbound cytoplasmic ComR transcriptional regulators; (v) active ComRS complexes; (vi) competence sigma factor ComX and (vii) ComZ, which embeds the actor(s) controlled by ComX that are responsible for competence shut-off. ComZ is assumed to inhibit the activity of ComRS by forming a ternary ComRSZ complex (Figure 1). An equation modeling the population cell density over time ($X(t)$), expressed in number of cells per ml) was also added to the system, to couple competence development to the growth process. At any time t , the state of the system consequently depends on the cell density.

The numerical values of the n ($n = 22$) parameters of the eight ODE equations were empirically determined on the basis of (i) literature; (ii) mechanistic hypotheses; and (iii) experimental data (see Materials and Methods for detailed information). The latter consist of the growth parameters of strain LMD-9 (Supplementary Figure S1A) and specific luciferase activities driven from promoters P_{comX} , P_{comS} , and P_{comR} in WT and ComR⁻ reporter strains grown in CDM conditions (Supplementary Figures S1B,C).

Meta-Hypotheses

Three meta-hypotheses were assumed to build the mathematical model of the ComRS regulatory cascade. First, our model describes the behavior of a homogenous population of cells in which all the individuals are assumed to act exactly the same way, i.e., we made a mean-field approximation of individual cells. The modeled behavior can thus be seen as the mean behavior over a large sample of cells, neglecting the intercellular variations. This simplification was made because the experimental data used to calibrate the model parameters derive from luciferase activities measured at the population scale level (see Materials and Methods). This explains why the values of some parameters are not whole numbers in our simulations (Table 2). Second, we used the quasi-steady-state assumption for mRNA dynamics, hence focusing on the protein level only. This simplification is generally accepted because RNAs fluctuate much more rapidly, and hence, can be assumed to be in quasi-steady-state compared to the slower varying protein level. Third, we assumed that the transcription and translation steps are instantaneous: as previously discussed (Young and Bremer, 1975; Proshkin et al., 2010), delays can be neglected in prokaryotes due to the lack of compartmentalization.

Model Validation

To get confident in our parameter setting and in the predictive power of our model, the output of different sets of simulation

were compared to existing or original experimental data. The chosen simulations aimed to test reactivity, dynamics and topology of the system.

Reactivity to ComS Induction

It was previously shown that the competence default of ComS⁻ derivatives or poorly transformable strains may be supplemented by exogenous synthetic ComS₁₇₋₂₄ peptides (Fontaine et al., 2010a,b, 2013). This behavior is typically observed in pheromone-based signaling systems (Sturme et al., 2002). One first set of validation consisted thus in testing if our model responded similarly to extracellular mature ComS (experimental synthetic ComS₁₇₋₂₄ or simulated ComS^{ext}) in a ComS⁻ background. To this end, the ComS dose-response curves of experimental and simulated ComS⁻ backgrounds were compared. Experimental curves were obtained by measuring the maximum luciferase activity emitted by the ComS⁻ P_{comX-luxAB} and P_{comS-luxAB} reporter strains in response to increasing synthetic ComS₁₇₋₂₄ concentrations. The simulated curves were obtained by computing the maximum number of ComX molecules per cell obtained by integrating our system with an increasing initial concentration for ComS^{ext}. To mimic a ComS⁻ background in our model, Equation (2), describing the production of ComS, was set to zero (see Materials and Methods). Experimental molar ComS₁₇₋₂₄ concentrations were converted into molecules per cell to facilitate the comparison. As

respectively illustrated in Figures 3A,B, experimental (blue) and simulated (red) curves are both qualitatively and quantitatively comparable. Indeed, the reactivity of the *in silico* and *in vivo* systems is in the same range: ComX concentrations were respectively induced ~ 280 - and 560 -fold by saturating peptide concentrations. Interestingly, the minimal pheromone concentration that is needed to experimentally turn on the system is $\sim 10^4$ -fold higher than the concentration needed to activate the system *in silico*. This suggests that all the synthetic ComS₁₇₋₂₄ peptides that are added to the cell culture are not effectively reimported inside the cells. This could be due to ComS degradation by cell surface proteases (e.g., HtrA), as described for other pheromones (Cassone et al., 2012), ComS aggregation in the growth medium or a limitation in the diffusion of ComS molecules through the cell wall.

Previous studies showed that the reactivity of streptococci to ComS induction is species-specific, suggesting that it is strongly dependent on the specificity of the ComR-ComS interaction. In *S. thermophilus*, some mutations in ComS₁₇₋₂₄ were shown to decrease its biological activity, probably by affecting the formation of active ComRS complexes (Fontaine et al., 2013). Indeed, the efficiency of ComS₁₇₋₂₄ variants to induce $P_{comS-luxAB}$ *in vivo* (see the dose-response curves on Figure 3C) is strongly related to its ability to induce the formation of ternary P_{comS} -ComRS complexes *in vitro* (Supplementary Figure S3). A second set of simulations aimed thus to test whether ComX production was similarly dependent on the rate of ComRS complex formation in our model. The value of the ComRS complex formation rate constant c_{RS} was decreased (see Materials

and Methods) and dose-response curves were simulated by computing the maximum number of ComX molecules produced in a ComS⁻ simulated cell (Figure 3D). As expected, our model was able to simulate the effect of ComS mutations that affect the reactivity of ComRS *in vivo*: the amplitudes and profiles of simulated dose-response curves resemble those observed with mutated ComS₁₇₋₂₄ variants in the reporter strain (compare Figure 3C and Figure 3D).

Dynamics of Induction

In CDM medium, the timing of induction by mature ComS peptides (i.e., ComS₍₁₇₋₂₄₎) was previously shown to be critical to reach optimal competence levels (Gardan et al., 2013; Figure 4A). To test the dynamics of ComX production over time in our model, we simulated the supplementation of mature ComS peptides (i.e., ComS^{ext}) at different moments during growth of a ComS⁻ strain (see Materials and Methods). Maximum ComX concentrations computed at each induction time is in good accordance with experimental data (compare Figure 4A and Figure 4B) (Gardan et al., 2013). Indeed, ComX production is inducible at all times of growth but both the rate and amplitude are weaker at time 0 and reach an optimum between 60 and 90 min after inoculation. The apparent discrepancy between modeled and experimental output when cells reach the stationary phase (i.e., at induction times of 180 and 210 min) is probably due to the loss of luciferase activities resulting from a depletion of FMNH₂ (data not shown and Gardan et al., 2013).

Topology of the ComRS Model

The topology of model S1 was tested by simulating the behavior observed in a ComX⁻ background, in which the activation of ComRS-controlled early gene *comS* is stronger and lasts longer, as a probable consequence of shut-off deregulation (Figure 4C) (Boutry et al., 2013). The ComX⁻ background was simulated by setting equation describing the evolution of *ComX(t)* to zero in the model (Equation 7). As expected, the kinetics and amplitude of the simulated evolution of ComS concentrations in a ComX⁻ background are similar to those of the luciferase activity reporting the activation of the P_{comS} promoter in the ComX⁻ reporter strain LF122 (compare Figure 4C and Figure 4D). Competence shut-off is indeed delayed in both cases. It is noteworthy that this delay appears shorter in experiments compared to simulations. However, this difference is artefactual since the well-documented loss of luciferase activity in the stationary phase (Duncan et al., 1994; Koga et al., 2005) is largely responsible for the decrease in *comS* expression (Figure 4C). The delayed drop in simulated ComS concentration (Figure 4D) solely results from the negative terms modeled in the corresponding equation, which account for ComS exportation, degradation and cell division (see Materials and Methods).

ComR Abundance Limits the Competence Level of *S. thermophilus* LMD-9

The level of competence of *S. thermophilus* LMD-9 is directly dictated by the amount of ComX molecules produced (Fontaine et al., 2010a; Boutry et al., 2013). The spontaneous activation of

comX expression in that strain is not at its maximum in CDM conditions since it may be further increased by adding saturating concentration of ComS (Fontaine et al., 2010a,b and data not shown). The ComRS model was used to determine which known actor limits *comX* expression in strain LMD-9 and, consequently, the competence level. To this end, a sensitivity analysis was exploited to identify the variable(s) and/or parameter(s) of the ComRS model having the most important impact on the amplitude of the model output. Parameters were grouped according to their role in equations (Supplementary Figure S4). Most sensitive parameters influence the specific amount of variables ComS (all peptidic forms: precursor, exported, and reimported mature peptide) and ComR. They account for their degradation rate (Supplementary Figure S4G), the rate of ComRS complex formation (Supplementary Figure S4B), and production rates (Supplementary Figures S4C,E,F,H). To further compare how ComS and ComR concentrations impact the competence level of LMD-9 cells, we tested *in silico* the effect of a ComS vs. ComR overproduction (respectively designated ComS⁺ and ComR⁺ in Figure 5A) on the maximum number of ComX molecules. Simulations were performed by multiplying the maximum basal production rate of ComR and ComS by a constant induction factor (IF) ranging from 1 to 100 (see Materials and Methods).

At IF ranging from 1 to 15, the maximum cellular ComX concentration increases similarly in both ComS⁺ and ComR⁺ simulated cells (Figure 5A). However, stronger induction resulted in a ~30% higher output in the case of ComR

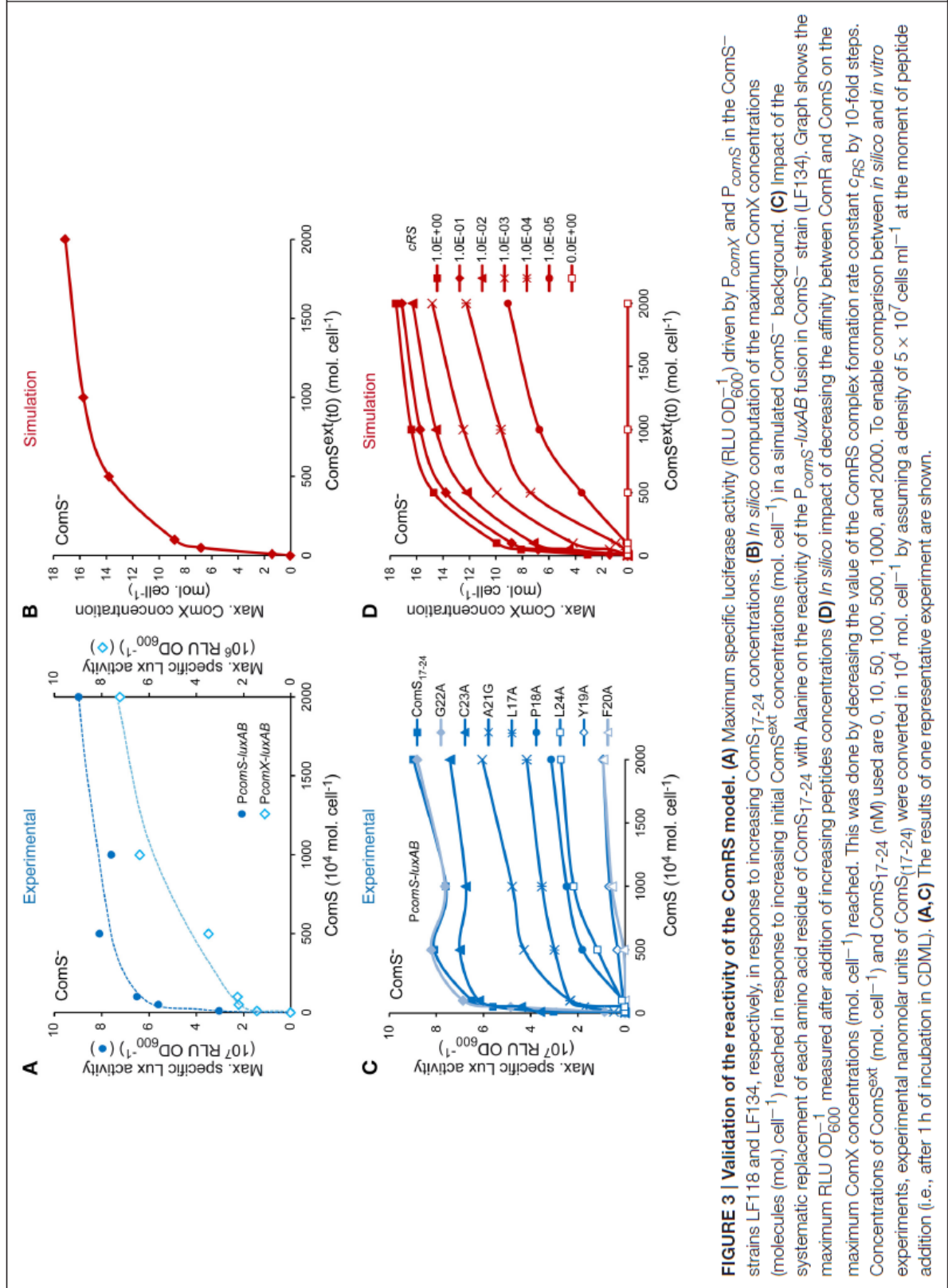


FIGURE 3 | Validation of the reactivity of the ComRS model. (A) Maximum specific luciferase activity (RLU OD_{600}^{-1}) driven by P_{comX} and P_{comS} in the ComS⁻ strains LF118 and LF134, respectively, in response to increasing ComS₁₇₋₂₄ concentrations. **(B)** *In silico* computation of the maximum ComX concentrations (molecules (mol.) cell⁻¹) reached in response to increasing initial ComS^{ext} concentrations (mol. cell⁻¹) in a simulated ComS⁻ background. **(C)** Impact of the systematic replacement of each amino acid residue of ComS₁₇₋₂₄ with Alanine on the reactivity of the $P_{comS-luxAB}$ fusion in ComS⁻ strain (LF134). Graph shows the maximum RLU OD_{600}^{-1} measured after addition of increasing peptides concentrations **(D)** *In silico* impact of decreasing the affinity between ComR and ComS on the maximum ComX concentrations (mol. cell⁻¹) reached. This was done by decreasing the value of the ComRS complex formation rate constant c_{RS} by 10-fold steps. Concentrations of ComS^{ext} (mol. cell⁻¹) and ComS₁₇₋₂₄ (nM) used are 0, 10, 50, 100, 500, 1000, and 2000. To enable comparison between *in silico* and *in vitro* experiments, experimental nanomolar units of ComS₁₇₋₂₄ were converted in 10^4 mol. cell⁻¹ by assuming a density of 5×10^7 cells ml⁻¹ at the moment of peptide addition (i.e., after 1 h of incubation in CDML). **(A,C)** The results of one representative experiment are shown.

overproduction. At saturation, a 2- vs. 1.5-fold increase of ComX concentration is observed in ComR⁺ vs. ComS⁺ cells. Importantly, similar outputs were obtained when the value of each parameter was varied separately within an interval ranging of 0.5–2 times its assigned value, indicating that this prediction is robust (Supplementary Figure S5). Those results suggest that increasing the cellular concentration of ComR vs. ComS will lead to stronger levels of *comX* expression in strain LMD-9. Interestingly, simulations also predict that the abundance of ComR controls the kinetics of competence in CDM. Indeed, the increased production rate of ComR is directly correlated to an earlier induction of ComX production (Figure 5B). To test those predictions experimentally, the competence level and kinetics of a WT and a ComR⁺ strain, which contains a *comR* expression vector (Boutry et al., 2013), were compared. The effect of increasing ComS abundance in both backgrounds was tested by monitoring P_{comX} in response to increasing ComS₁₇₋₂₄ concentrations. Results obtained (Figures 5C,D) show that compared to WT cells, P_{comX} activity in ComR⁺ cells is (i) no more inducible by ComS₁₇₋₂₄, indicating that ComS does not limit competence in those conditions; (ii) is stronger at all

ComS₁₇₋₂₄ concentrations tested (Figure 5C); and (iii) increases from the beginning of growth and is thus no more dependent on the growth phase (Figure 5D).

Altogether, our results strongly suggest that ComR abundance is the major limiting factor which controls the level of *comX* expression and the kinetics of competence induction in the naturally transformable strain LMD-9.

Low Competence of Strain LMG18311 is Due to Weak *comR* Expression

The ability of *S. thermophilus* to spontaneously turn on high competence levels under CDM growth conditions seems restricted to rare strains (Fontaine et al., 2010b). Since all of the key early competence genes to date are encoded in the genome of the poorly transformable strains, we assumed that differences in competence efficiency between *S. thermophilus* strains could be due to different production levels of actors of the ComRS signaling pathway. As ComR abundance was shown to limit competence in strain LMD-9, we studied the activity of P_{comR} and P_{comX} in the poorly transformable LMG18311

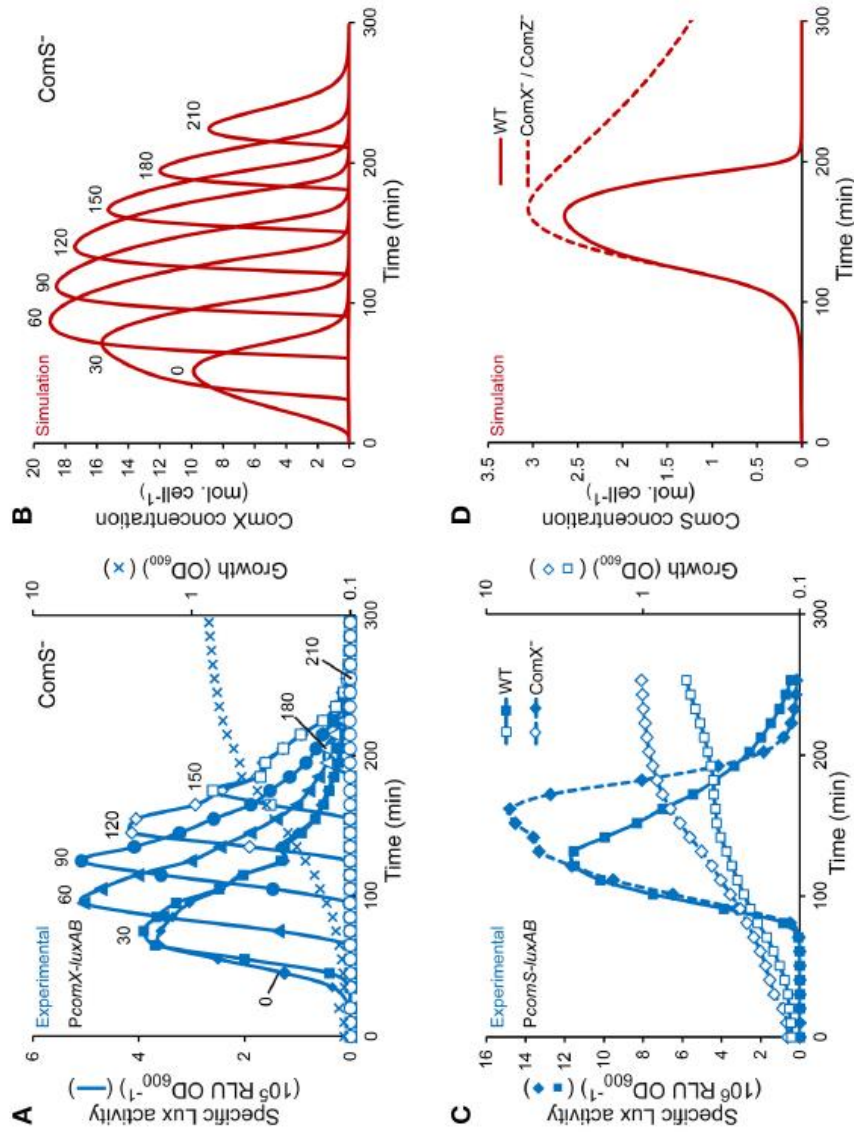


FIGURE 4 | Validation of the dynamics and topology of the ComRS model. (A) Specific activity of P_{comX} upon ComS₁₇₋₂₄ addition at different times of growth of a ComS⁻ luminescent reporter strain (TIL 1391). The curves show the specific luciferase activity (RLU OD₆₀₀⁻¹) of the reporter strain measured directly after addition of 1 μ M ComS₁₇₋₂₄ at different times of growth. The growth curve (OD₆₀₀) of the reporter strain grown in CDML in absence of ComS₁₇₋₂₄ is shown. Data were adapted from Gardan et al. (2013). (B) Modeled kinetics of ComX concentrations (molecules (mol.) cell⁻¹) in a ComS⁻ background obtained by simulating the addition of ComS^{ext} (150 mol. cell⁻¹) at different times of growth. The time T expressed in min (T = 0, 30, 60, 90, 120, 150, 180, and 210) at which ComS₁₇₋₂₄ and ComS^{ext} were experimentally and *in silico* added, respectively, is indicated above each curve. (C) Comparison of the kinetics of the specific activity of P_{comS} between WT (LF121) and ComX⁻ (LF122) reporter strains (RLU OD₆₀₀⁻¹) in CDML growth conditions. (D) Comparison of modeled kinetics of ComS concentrations (mol. cell⁻¹) between WT and ComX⁻ or ComZ⁻ strains. (A, C) The results of one representative experiment are shown.

strain. As expected, maximum P_{comX} activity is very low in LMG18311 but might be increased 345-fold upon ComS₁₇₋₂₄ supplementation (Figure 6A). Similar results were obtained for P_{comS} activity (data not shown), which confirms that similarly to other weak competent strain, ComS production is limiting in strain LMG18311. Interestingly, P_{comR} activity is also ~10-fold lower in strain LMG18311 compared to strain LMD-9 (Figure 6A).

The ComRS model was next used to test whether the observed difference in *comR* expression accounts for low expression of *comX*. To mimic strain LMG18311, the production rate of *ComR*(t) was divided by 10. The latter is modeled as the sum of rates due to ComR basal production ($b_R(t)$) and activation/repression by -as yet- unidentified regulator(s) ($N(t)$). As shown on Figure 6B, simulations done after dividing both terms reproduced the behavior of strain LMG18311: maximum cellular ComX concentration reaches high levels only in presence of initial ComS^{ext} concentrations. In addition, a similar four-fold difference in simulated and experimental outputs (respectively maximum ComX concentration and P_{comX} activity) was observed between strains LMG18311 and LMD-9 in ComS-induced conditions. Similar results were obtained when the rate of *ComR*(t) activation/repression solely was divided in the ComRS model, while it was not the case when only the

basal rate was divided by 10. This suggests that the contribution of ComR basal production to ComX production is minor in *S. thermophilus* (Figure 6B).

In conclusion, the weak transformability of strain LMG18311 results from a low expression of *comR*, solely. This may prevent the activation of the positive feedback loop acting on ComS production.

Increasing ComR Concentration Induces Competence in the Weakly Transformable Strain LMG18311

To further confirm the key role of ComR concentration on the transformability of strain LMG18311, we tested the consequence of *comR* overexpression on this phenotype. As expected, both the mean transformation frequency ($4.4 \times 10^{-4} \pm 4.0 \times 10^{-4}$ vs. $2.1 \times 10^{-6} \pm 5.0 \times 10^{-5}$) and the mean maximum specific luciferase activity (RLU OD₆₀₀⁻¹) of P_{comX} -*luxAB* fusions ($6.7 \times 10^6 \pm 3.5 \times 10^6$ vs. $4.4 \times 10^4 \pm 1.2 \times 10^3$) increased significantly in the reporter derivative strain carrying the *comR* overexpression vector compared to those carrying the empty vector. Interestingly, besides a high variability in the amplitude of P_{comX} activity (Figure 6C and data not shown), the various LMG18311 ComR⁺ clones also display variable (but reproducible) kinetics patterns,

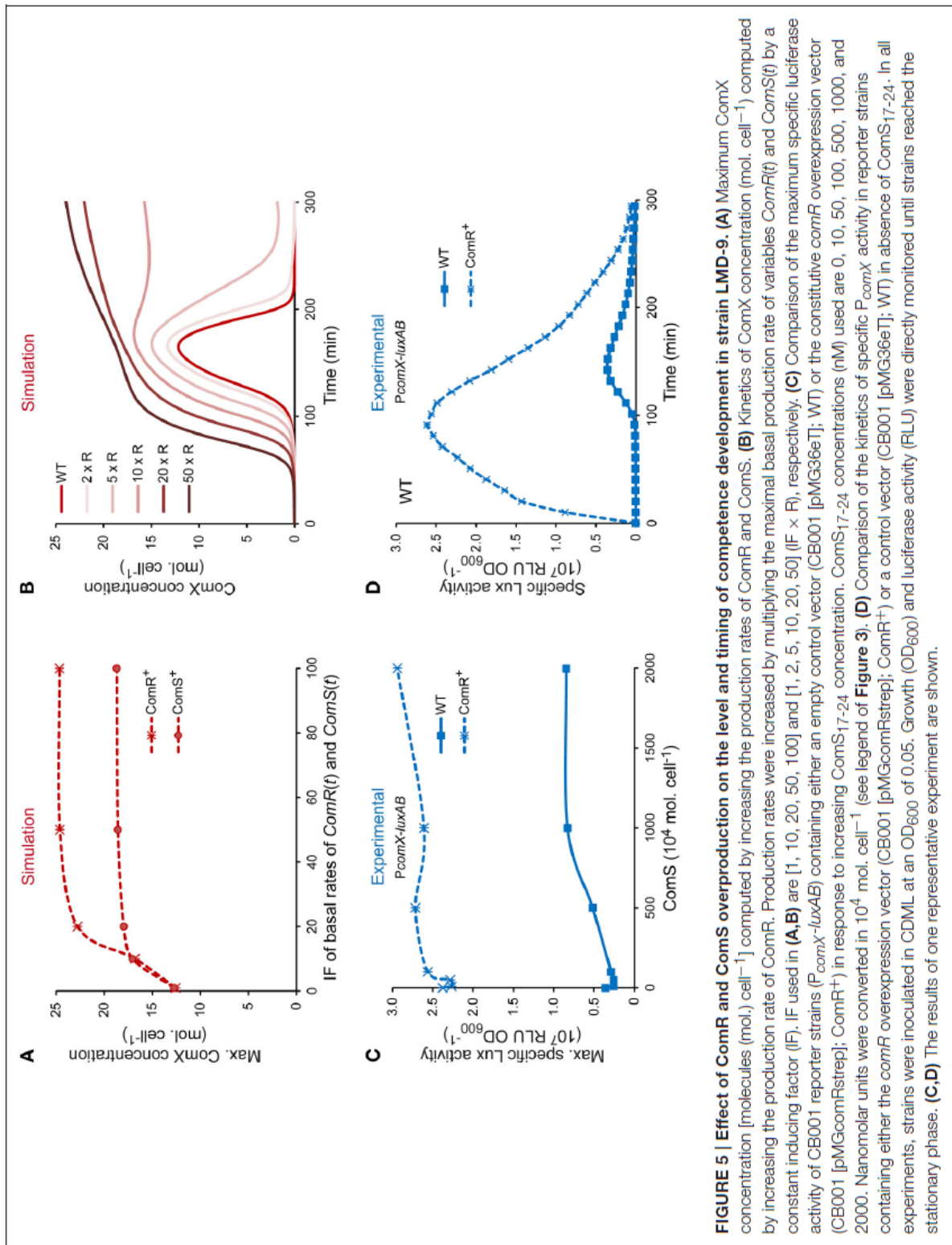


FIGURE 5 | Effect of ComR and ComS overproduction on the level and timing of competence development in strain LMD-9. (A) Maximum ComX concentration [molecules (mol.) cell⁻¹] computed by increasing the production rates of ComR and ComS. **(B)** Kinetics of ComX concentration (mol. cell⁻¹) computed by increasing the production rate of ComR. Production rates were increased by multiplying the maximal basal production rate of variables *ComR(t)* and *ComS(t)* by a constant inducing factor (IF). IF used in **(A,B)** are [1, 10, 20, 50, 100] and [1, 2, 5, 10, 20, 50] (IF x R), respectively. **(C)** Comparison of the maximum specific luciferase activity of CB001 reporter strains (*P_{comX}-luxAB*) containing either an empty control vector (CB001 [pMG36eT]; WT) or the constitutive *comR* overexpression vector (CB001 [pMGcomRstrep]; ComR⁺) in response to increasing ComS₁₇₋₂₄ concentration. ComS₁₇₋₂₄ concentrations (nM) used are 0, 10, 50, 100, 500, 1000, and 2000. Nanomolar units were converted in 10⁴ mol. cell⁻¹ (see legend of **Figure 3**). **(D)** Comparison of the kinetics of specific *P_{comX}* activity in reporter strains containing either the *comR* overexpression vector (CB001 [pMGcomRstrep]; ComR⁺) or a control vector (CB001 [pMG36eT]; WT) in absence of ComS₁₇₋₂₄. In all experiments, strains were inoculated in CDML at an OD₆₀₀ of 0.05. Growth (OD₆₀₀) and luciferase activity (RLU) were directly monitored until strains reached the stationary phase. **(C,D)** The results of one representative experiment are shown.

which is either monophasic (4/10 clones), biphasic (5/10 clones), or in rare instances triphasic (1/10 clone) (the specific luciferase activity of representative clones are presented in Figure 6C). To rationally interpret these results, we simulated increasing levels of ComR production by multiplying its maximum basal rate by a range of increasing values (1.5–20). For IF ranging from 3.25 to 7 (Figure 6D and data not shown), the time-course evolution of ComX concentration displays an oscillating behavior. The number of phases is directly proportional to the level of ComR production. At higher IF (Figure 6D), the phases merge while maximum amplitude continues to increase until saturation. Those predictions suggest that the different behaviors observed among LMG18311 ComR⁺ clones result from different levels of *comR* expression, which could for instance be due to different copy numbers of the corresponding expression vector.

DISCUSSION

Competence for natural transformation is a major driving force in bacterial evolution. Hence, it is of key importance to understand how this transient phenomenon takes place and

is regulated. In Gram-positive bacteria, the best documented competence regulation systems are those of the model species *S. pneumoniae* (ComCDE-ComX relay system) (Claverys et al., 2006) and *Bacillus subtilis* (ComK master regulator) (Hamoen et al., 2003). During the last decade, the regulatory cascades governing competence in these two species have been mathematically modeled for helping to solve some important aspects such as the entry in the competence state (Leisner et al., 2009; Liebal et al., 2010), competence shut-off (Karlsson et al., 2007; Schultz et al., 2007), competence bistability (Maamar and Dubnau, 2005; Dubnau and Losick, 2006; Son et al., 2012; Xi et al., 2013), decision-making process between developmental programs (competence vs. sporulation) (Schultz et al., 2009), or benefits and cost of competence development (Moradigaravand and Engelstadter, 2013). These modeling approaches were very useful to decipher these processes and to obtain in some cases counter-intuitive explanations.

The ComRS-ComX relay system, which regulates competence in a majority of streptococcal groups, was recently discovered (for a review, see Fontaine et al., 2015). Even if the available experimental data begin to accumulate, a series of questions

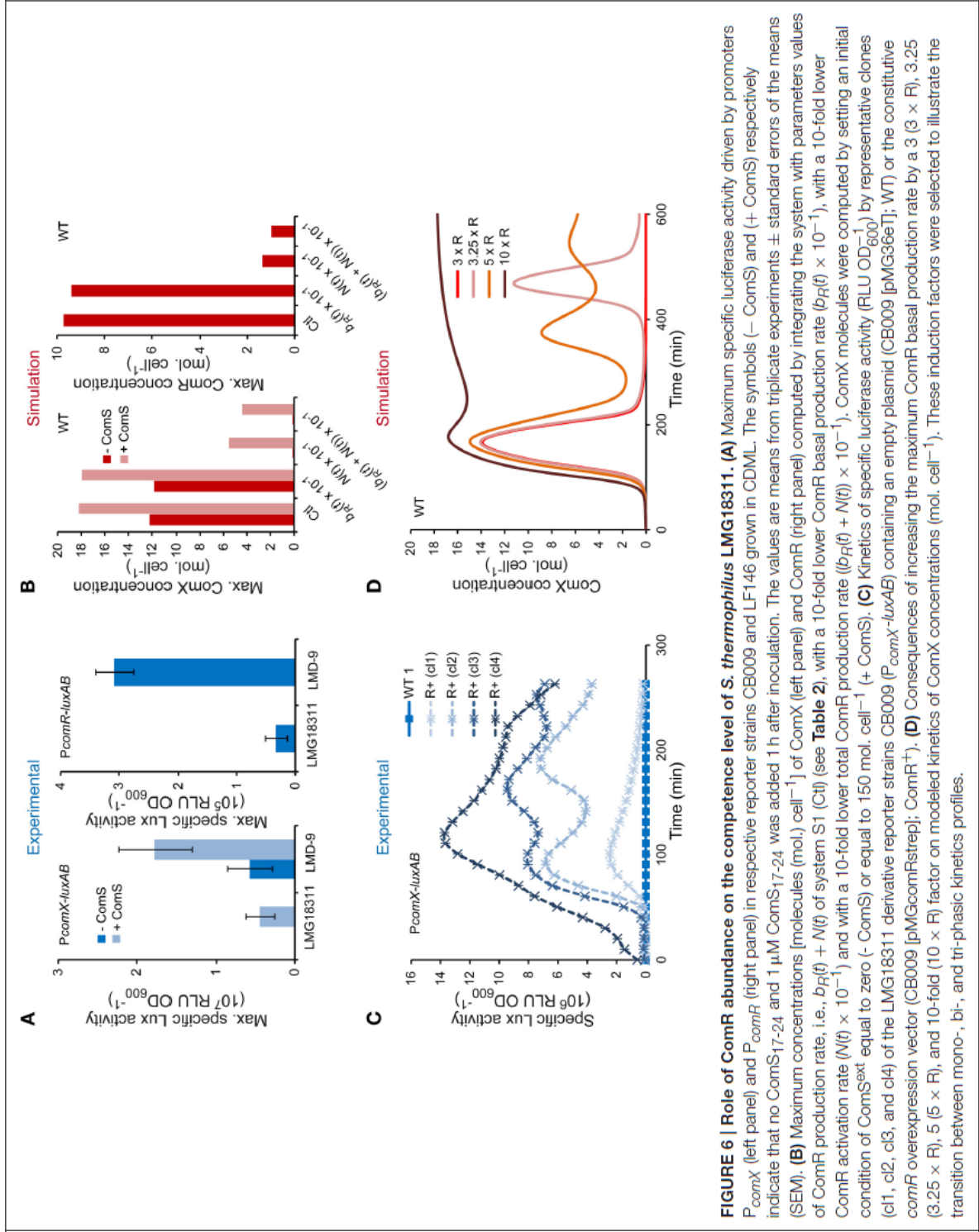


FIGURE 6 | Role of ComR abundance on the competence level of *S. thermophilus* LMG18311. (A) Maximum specific luciferase activity driven by promoters P_{comX} (left panel) and P_{comR} (right panel) in respective reporter strains CB009 and LF146 grown in CDML. The symbols (– ComS) and (+ ComS) respectively indicate that no ComS17-24 and 1 μ M ComS17-24 was added 1 h after inoculation. The values are means from triplicate experiments $\pm$ standard errors of the means (SEM). (B) Maximum concentrations [molecules (mol.) cell $^{-1}$] of ComX (left panel) and ComR (right panel) computed by integrating the system with parameters values of ComR production rate, i.e., $b_R(t) + N(t)$ of system S1 (Ctl) (see Table 2), with a 10-fold lower ComR basal production rate ($b_R(t) \times 10^{-1}$), with a 10-fold lower ComR activation rate ($N(t) \times 10^{-1}$) and with a 10-fold lower total ComR production rate ($b_R(t) + N(t) \times 10^{-1}$). ComX molecules were computed by setting an initial condition of ComS ^{ext} equal to zero (– ComS) or equal to 150 mol. cell $^{-1}$ (+ ComS). (C) Kinetics of specific luciferase activity (RLU OD_{600}^{-1}) by representative clones (cl1, cl2, cl3, and cl4) of the LMG18311 derivative reporter strains CB009 ($P_{comX-luxAB}$) containing an empty plasmid (CB009 [pMG386T]; WT) or the constitutive *comR* overexpression vector (CB009 [pMGcomRstrepi]; ComR $^{+}$). (D) Consequences of increasing the maximum ComR basal production rate by a 3 (3 $\times$ R), 3.25 (3.25 $\times$ R), 5 (5 $\times$ R), and 10-fold (10 $\times$ R) factor on modeled kinetics of ComX concentrations (mol. cell $^{-1}$). These induction factors were selected to illustrate the transition between mono-, bi-, and tri-phasic kinetics profiles.

remain open on its functioning such as the topology of the regulatory cascade or the critical components of this system responsible for the on-off state of competence development. Although a simple mathematical model integrating ComRS was built to study environmental inputs on competence bistability in *S. mutans* (Son et al., 2012), a kinetic model that integrates all known steps of *comX* regulation by ComRS was missing until now.

The ComRS Dynamic Model: Limitations and Future Developments

At this stage, our model presents some limitations partially due to mechanistic or meta-hypotheses that were done to construct the model. First, the model neglects the putative cell-to-cell communication role of ComS. However, although “self”-communication dominates over “others”-communication in peptide diffusion-limiting conditions (such as modeled here), the combination of specific parameters may influence the output of social behaviors mediated by pheromone-based regulatory systems, such as the rate of peptide secretion and importation,

the strength of the positive feedback loop (Youk and Lim, 2014) and/or the possibility of communication mediated by physical contacts between cells. In the future, both aspects (egoist and altruistic) of ComS signaling will be integrated in our model in order to study their relative contribution in competence regulation in various growth conditions.

Other limitations of our current model are that (i) the simulated cell growth is always assumed to be the same between WT and mutants while altered growth profiles were experimentally observe in mutant strains, (ii) the model has been calibrated and evaluated only for CDML growth conditions (iii) the model only considers the effect of proximal actors on *comX* expression while the effects of all distal regulators are summed up in the temporal gaussian term ($N(t)$) of ComR due to a lack of knowledge on distal regulation, and (iv) the model is not calibrated on the true abundance of ComR, ComS, and ComX proteins, which is unknown in ComRS-containing streptococci, but on expression data based on the activity of the luciferase reporter system. This last point implicates that the model remains qualitative and could be used to

predict behaviors and to compare quantitative effects, but not to directly infer quantitative data. So, future improvements of the model will aim to solve some of these limitations such as integrating detailed distal regulations (transcriptional and post-transcriptional), improving quantitative predictions by measuring the time-course production of key players of the system during growth, or taking into account both RNA and protein levels.

Finally, the question of the portability of the model developed for *S. thermophilus* to other ComRS-containing streptococci is an important aspect. Although differences exist between species regarding for instance the diffusion properties of ComS and the mechanism of competence shut-off (see below), the topology of the ComX induction cascade as well as growth conditions allowing its activation (i.e., CDM conditions) are strongly conserved among ComRS-containing streptococci (for a review, see Fontaine et al., 2015). This suggests that the adaptation of the model to other phylogenetically close streptococci (e.g., *S. mutans*, *S. salivarius*) could be achieved without major adjustments.

ComR Abundance is Critical for Competence Activation

One of the interesting outputs of the use of the model is that, besides the ComS pheromone, ComR is a limiting factor for maximum competence development of *S. thermophilus* in CDM conditions (Figure 5A). This was experimentally validated: the *comR* overexpression in strain LMD-9, which develops competence spontaneously, and strain LMG18311, which is

weakly competent, resulted in a mean 10- and 100-fold higher level of *comX* expression, respectively (Figures 5C, 6C). It is noteworthy that increasing ComS abundance in strains LMD-9 and LMG18311 (through *in silico* or *in vivo* exogenous peptide supplementation) (Figures 6A,B) induces a similar positive effect on competence development but the amplitude of the response is correlated to the number of ComR molecules produced in each strain. Indeed, the induction of *comX* by saturating ComS concentrations remains stronger in strain LMD-9 vs. LMG18311 (Figures 6A,B). In regard to the topology of the competence induction cascade, ComR is indeed pivotal as its amount will determine the number of ComRS complexes which will be formed. A critical amount of ComRS complexes is probably required to activate the positive feedback loop acting on *comS* expression and hence induce competence in the population. Our results strongly suggest that this minimal amount is not reached in strain LMG18311 because of a too low *comR* expression level. The model predicts that increasing the ComR concentration in this strain shifts the balance toward formation of enough ComRS complexes to spontaneously induce competence. Interestingly, competence also turns on much earlier during growth of ComR⁺ vs. WT cells because of the immediate production of high amount of ComRS (data not shown), indicating that the timing device is constrained by the kinetics and amplitude of ComR production, solely. Such a critical role of ComR abundance does not seem to be restricted to the *S. thermophilus* species since it was previously reported that ComR overexpression in *S. mutans* UA159 resulted in a 10-fold increase in transformability

(Mashburn-Warren et al., 2010). Increasing ComR production could thus provide a simple way to render weakly or non-competent strains highly competent without the requirement of adding synthetic ComS.

The role of ComR as a limiting factor of competence may be related to the structure of the ComRS signaling cascade in which ComR production is independent of the positive feedback loop (Figure 1). Indeed, this contrasts with the topology of most cell-cell communication systems in which productions of master regulator and pheromone are auto-regulated and coordinated at the transcriptional level, as for instance in the ComCDE system controlling competence in *S. pneumoniae* (Fontaine et al., 2015) and the streptococcal Rgg/Shp systems (Pérez-Pascual et al., 2015). This ensures that the same amount of regulator and associated pheromone are produced in these systems. The evolutionary consequence of uncoupling ComR and ComS production may be advantageous to fine-tune both kinetics and amplitude of competence activation in response to various environmental inputs. Currently, the modulation of *comR* transcription by distal regulatory systems (e.g., BsrRM-HdrRM relay and ScnRK) has only been shown in *S. mutans* (Okinaga et al., 2010; Xie et al., 2010; Kim et al., 2013). In addition, other signaling systems (i.e., *S. pyogenes* Rgg/Shp, and *S. pneumoniae* TprA/PhrA) of the RNPP family (the latter includes ComR) have also been shown as transcriptionally regulated by distal systems sensing either metal availability (Chang et al., 2015) or carbon source (Chang et al., 2015; Hoover et al., 2015). In *S. thermophilus*, no distal regulation acting on *comR* expression has been reported so far.

The comparative study of strains LMD-9 and LMG18311 indicate that competence heterogeneity between *S. thermophilus* strains (Fontaine et al., 2010b) could partially be due to differences in the *comR* expression level. The origin of these differences (i.e., LMD-9 vs. LMG18311) is not clear. Our model predicts that the activation or de-repression of *comR* expression by distal regulators, rather than the basal expression, is affected in strain LMG18311 compared to strain LMD-9 (Figure 6B). It may either be due to sequence polymorphism of the promoter region that affects binding of regulator(s), or to the presence of mutations in some *comR* distal regulators in poorly competent strains. Alternatively, in regard to the exceptional transformability of strain LMD-9 within the *S. thermophilus* species, it is possible that LMD-9 harbors specific mutations in *comR* regulator(s) allowing competence induction in otherwise non-permissive growth conditions. Consistent with this hypothesis of competence deregulation, strains of *Bacillus licheniformis* displaying efficient transformability carry mutations in competence regulator genes as a probable consequence of domestication (Jakobs and Meinhardt, 2015). Swapping *comR* promoter regions and determination of the sub-regions important for *comR* expression could help to begin answering this question.

Insights into Competence Shut-Off

In 2007, the use of an ODE mathematical model of the ComCDE system has suggested that the competence shut-off mechanism in *S. pneumoniae* was likely to involve a late phase repressor

protein (Karlsson et al., 2007). A couple of years later, the shut-off mechanism in *S. pneumoniae* was experimentally unraveled and takes place at multiple levels: (i) competition between ComE and ComE~P (Martin et al., 2013), (ii) inhibition of ComX activity (Weng et al., 2013), and (iii) direct inhibition of ComE~P by DprA (Mirouze et al., 2013). The latter is a late phase protein that was initially proven as involved in DNA integration by recruiting RecA (Bergé et al., 2003; Mortier-Barrière et al., 2007).

In ComRS-containing streptococci, the shut-off mechanism is poorly understood and seems to differ among species (Fontaine et al., 2015). In *S. mutans*, it was shown that this mechanism was at least due to a post-translational control of ComX stability by the concerted action of the MecA adaptor protein and the ClpCP proteolytic system (Tian et al., 2013; Dong et al., 2014). In contrast, it has been shown that the MecA-ClpCP machinery does not contribute to the shut-off in *S. thermophilus* (Boutry et al., 2012; Wahl et al., 2014). However, it has been observed that ComX-negative strains exhibit a longer and stronger activation of genes belonging to the ComRS core regulon (i.e., *comS* and *comX*), without affecting *comR* expression (Boutry et al., 2013). This behavior could be simulated with our predictive model (Figures 4C,D), confirming that at least one late gene product is partially responsible for competence shut-off and acts by negatively affecting ComRS activity.

In our model, the feedback inhibition loop via the putative effector of shut-off ComZ is directed on the ComRS complex in order to form a ternary complex ComRSZ (Figure 1). By analogy with *S. pneumoniae*, the orthologue of DprA (which interacts and inhibits ComE activity in the pneumococcus)

was an obvious putative candidate for this function. However, the early competence phase of *S. thermophilus* LMD-9 is not affected in a *dprA* deletion mutant compared to the WT (see Supplementary Figure S6), suggesting that the DprA orthologue is not involved in competence exit in this species. Although other late competence gene(s) could substitute DprA for sequestering active ComRS complexes in *S. thermophilus*, we cannot rule out alternative inactivation mechanisms such as (i) degradation of ComR and/or ComS, (ii) inactivation by post-translational modification of ComR and/or ComS (e.g., ComR phosphorylation), or (iii) inhibition of formation or activity of the ComRS complex (e.g., inhibitory peptide, competition for DNA binding). Changing the target of the negative loop in the model combined with modifications of parameters affecting the abundance of ComR and/or ComS could be useful to define at which level(s) the shut-off is taking place.

Oscillatory Behavior of Competence Activation

Besides the critical role of ComR in the strength of competence activation, ComR abundance also regulates an oscillatory behavior to competence activation. Indeed, an oscillating profile of *comX* expression was observed *in silico* and experimentally in the *comR*-overexpressing strain LMG18311 (Figures 6C,D). This provides an additional validation of the topological organization and the dynamical calibration of our model. Indeed, Novák and Tyson (2008) showed that the four minimal requirements for oscillations are: (i) a negative feedback loop, which is

made in our setting of $\text{ComRS} \rightarrow \text{ComX} \rightarrow \text{ComZ} \dashv \text{ComRS}$, (ii) a delay in this negative loop, implemented here by the positive loop made of $\text{ComS} \rightarrow \text{ComS}^{\text{ext}} \rightarrow \text{ComS}^* \rightarrow \text{ComRS} \rightarrow \text{ComS}$, (iii) sufficient non-linearity, which arises here from our non-linear activation/repression functions and from the mass action law of the ComRS assembly with $n = 2.5$, and (iv) appropriate time-scales of the opposite forces, confirming that our parameter set is coherent. Moreover, according to the notation of Lenz and Sogaard-Andersen (2011), our system can be classified as a relaxation oscillator due to its structure (a time-delayed negative feedback loop coupled to a positive feedback loop), hence operating in a broad frequency range. Indeed, both amplitude and frequency of oscillations can be modulated by varying the ComR production level (Figures 6C,D).

Temporal oscillations have been reported in a wide range of biological networks and organisms: from bacterial gene circuits (Schultz et al., 2009, 2013) to eukaryotic circadian rhythms (Lenz and Sogaard-Andersen, 2011) and quorum-sensing systems (Sgro et al., 2015). In addition, synthetic oscillators have been successfully built (Elowitz and Leibler, 2000). Interestingly, cycles in the competence state have already been reported for *S. pneumoniae* (Chen and Morrison, 1987) and *Helicobacter pylori* (Baltrus and Guillemin, 2006). In *S. pneumoniae*, the activation of these multiple competence cycles (up to three during growth in batch cultures) in a WT strain was observed at high initial pH values (>7.25) while competence is inhibited in acidic conditions (Chen and Morrison, 1987). In WT strains of *Helicobacter pylori*, multiple competence phases occurred spontaneously but appeared to be strain-dependent (Baltrus and Guillemin,

2006). Although a similar cycling behavior is observed here in an artificial genetic background of *S. thermophilus*, oscillation in competence gene activation was also seen in response to ComS addition in the closely related species *S. salivarius* (data not shown), indicating that competence oscillation could be an intrinsic property of the ComRS signaling system. It has been previously shown that oscillations are synchronized in the whole population through quorum-sensing inducing molecules (McMillen et al., 2002; Schwab et al., 2012; Singh and Parmananda, 2013). However, it remains to be determined here if oscillations are taking place at the whole population level or if successive subpopulations are activating a single competence episode. Future time-lapse experiments at the single cell level with a GFP-reporter strain of competence activation will be needed to answer this question.

Finally, the question of the biological advantage conferred by an oscillating competence state has not yet been addressed. We can speculate that the benefit could be the lengthening of the transformation process in conditions where competence is highly beneficial for the population. Oscillations would then allow recovery phases that could be needed to relieve the inhibition of cell division resulting from competence activation, hence implementing a compromise between competence and growth, two processes that could be mutually exclusive (Haijema et al., 2001; Qi et al., 2005). However, one must keep in mind that the current observations were obtained in artificial laboratory conditions in which bacteria could exhibit exotic behaviors that would not or rarely be seen in natural conditions.

CONCLUDING REMARKS

To conclude, we designed an original approach combining mathematical modeling with molecular biology tools to build a robust data-based model of the ComRS signaling system, which helped us to confirm the topology of the competence regulatory network and to identify its critical components in *salivarius* streptococci. We demonstrated the key hierarchical importance of ComR in transformation efficiency and timing of competence activation, compatible with its putative role of master integration-sensor of distal regulations by environmental signals. In addition, our model shed light on the competence shut-off mechanism and proposed that the topology of the regulatory network based on ComRS allows an oscillatory behavior of competence activation. Future work will consist in exploiting our model to identify more finely at which level the shut-off is taking place in order to identify putative key player(s) involved in this mechanism. A possible oscillating competence state is also an exciting aspect that will require

further investigation such as its occurrence at the single cell level and its biological relevance.

ACKNOWLEDGMENTS

We warmly thank P. Soumillon and P. Goffin for fruitful discussions regarding mathematical modeling and critically reading of the manuscript. We gratefully thank R. Gardan for providing the raw data of Figure 4A. This research has been funded by FNRS and the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office. LH holds a doctoral fellowship from FRiA. LF is a postdoctoral researcher at FNRS. PH is Senior Research Associate at FNRS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.01413>

REFERENCES

- Baltrus, D. A., and Guillemin, K. (2006). Multiple phases of competence occur during the *Helicobacter pylori* growth cycle. *FEMS Microbiol. Lett.* 255, 148–155. doi: 10.1111/j.1574-6968.2005.00066.x
- Berg, K. H., Bjørnstad, T. J., Johnsborg, O., and Håvarstein, L. S. (2012). Properties and biological role of streptococcal fratricins. *Appl. Environ. Microbiol.* 78, 3515–3522. doi: 10.1128/AEM.00098-12
- Bergé, M., Mortier-Barrière, I., Martin, B., and Claverys, J. P. (2003). Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol. Microbiol.* 50, 527–536. doi: 10.1046/j.1365-2958.2003.03702.x
- Blomqvist, T., Steinmoen, H., and Håvarstein, L. S. (2006). Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 72, 6751–6756. doi: 10.1128/AEM.01156-06

- Bolouri, H., and Davidson, E. H. (2003). Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9371–9376. doi: 10.1073/pnas.1533293100
- Borgeaud, S., Metzger, L. C., Scignari, T., and Blokesch, M. (2015). The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science* 347, 63–67. doi: 10.1126/science.1260064
- Boutry, C., Delplace, B., Clippe, A., Fontaine, L., and Hols, P. (2013). SOS response activation and competence development are antagonistic mechanisms in *Streptococcus thermophilus*. *J. Bacteriol.* 195, 696–707. doi: 10.1128/JB.01605-12
- Boutry, C., Wahl, A., Delplace, B., Clippe, A., Fontaine, L., and Hols, P. (2012). Adaptor protein MecA is a negative regulator of the expression of late competence genes in *Streptococcus thermophilus*. *J. Bacteriol.* 194, 1777–1788. doi: 10.1128/JB.06800-11
- Cassone, M., Gagne, A. L., Spruce, L. A., Seeholzer, S. H., and Seibert, M. E. (2012). The HtrA protease from *Streptococcus pneumoniae* digests both denatured proteins and the competence-stimulating peptide. *J. Biol. Chem.* 287, 38449–38459. doi: 10.1074/jbc.M112.391482
- Chang, J. C., Jimenez, J. C., and Federle, M. J. (2015). Induction of a quorum sensing pathway by environmental signals enhances group A streptococcal resistance to lysozyme. *Mol. Microbiol.* 97, 1097–1113. doi: 10.1111/umi.13088
- Charpentier, X., Polard, P., and Claverys, J. P. (2012). Induction of competence for genetic transformation by antibiotics: convergent evolution of stress responses in distant bacterial species lacking SOS? *Curr. Opin. Microbiol.* 15, 570–576. doi: 10.1016/j.mib.2012.08.001
- Chen, J. D., and Morrison, D. A. (1987). Modulation of competence for genetic transformation in *Streptococcus pneumoniae*. *J. Gen. Microbiol.* 133, 1959–1967. doi: 10.1099/00221287-133-7-1959
- Claverys, J. P., Martin, B., and Håvarstein, L. S. (2007). Competence-induced fratricide in streptococci. *Mol. Microbiol.* 64, 1423–1433. doi: 10.1111/j.1365-2958.2007.05757.x
- Claverys, J. P., Prudhomme, M., and Martin, B. (2006). Induction of competence regulons as a general response to stress in gram-positive bacteria. *Annu. Rev. Microbiol.* 60, 451–475. doi: 10.1146/annurev.micro.60.080805.142139
- de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: a literature review. *J. Comput. Biol.* 9, 67–103. doi: 10.1089/10665270252833208
- Desai, K., Mashburn-Warren, L., Federle, M. J., and Morrison, D. A. (2012). Development of competence for genetic transformation of *Streptococcus mutans* in a chemically defined medium. *J. Bacteriol.* 194, 3774–3780. doi: 10.1128/JB.00337-12
- Dong, G., Tian, X. L., Gomez, Z. A., and Li, Y. H. (2014). Regulated proteolysis of the alternative sigma factor SigX in *Streptococcus mutans*: implication in the escape from competence. *BMC. Microbiol.* 14:183. doi: 10.1186/1471-2180-14-183
- Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16, 6127–6145. doi: 10.1093/nar/16.13.6127
- Dubnau, D., and Losick, R. (2006). Bistability in bacteria. *Mol. Microbiol.* 61, 564–572. doi: 10.1111/j.1365-2958.2006.05249.x
- Dufour, D., and Lévesque, C. M. (2013). Bacterial behaviors associated with the quorum-sensing peptide pheromone ('alarmone') in streptococci. *Future Microbiol.* 8, 593–605. doi: 10.2217/fmb.13.23
- Duncan, S., Glover, L. A., Killham, K., and Prosser, J. I. (1994). Luminescence-based detection of activity of starved and viable but nonculturable bacteria. *Appl. Environ. Microbiol.* 60, 1308–1316.
- Elowitz, M. B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338. doi: 10.1038/35002125
- Ferain, T., Hobbs, J. N. Jr., Richardson, J., Bernard, N., Garmyn, D., Hols, P., et al. (1996). Knockout of the two *ldh* genes has a major impact on peptidoglycan precursor synthesis in *Lactobacillus plantarum*. *J. Bacteriol.* 178, 5431–5437.
- Fleuchot, B., Gitton, C., Guillot, A., Vidic, J., Nicolas, P., Besset, C., et al. (2011). Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol. Microbiol.* 80, 1102–1119. doi: 10.1111/j.1365-2958.2011.07633.x
- Fontaine, L., Boutry, C., de Frahan, M. H., Delplace, B., Fremaux, C., Horvath, P., et al. (2010a). A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* 192, 1444–1454. doi: 10.1128/JB.01251-09

- Fontaine, L., Dandoy, D., Boutry, C., Delplace, B., de Frahan, M. H., Fremaux, C., et al. (2010b). Development of a versatile procedure based on natural transformation for marker-free targeted genetic modification in *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 76, 7870–7877. doi: 10.1128/AEM.01671-10
- Fontaine, L., Goffin, P., Dubout, H., Delplace, B., Baulard, A., Lecat-Guillet, N., et al. (2013). Mechanism of competence activation by the ComRS signalling system in streptococci. *Mol. Microbiol.* 87, 1113–1132. doi: 10.1111/mmi.12157
- Fontaine, L., Wahl, A., Fléhard, M., Mignolet, J., and Hols, P. (2015). Regulation of competence for natural transformation in streptococci. *Infect. Genet. Evol.* 33, 343–360. doi: 10.1016/j.meegid.2014.09.010
- Gardan, R., Besset, C., Gitton, C., Guillot, A., Fontaine, L., Hols, P., et al. (2013). Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J. Bacteriol.* 195, 1845–1855. doi: 10.1128/JB.02196-12
- Gardan, R., Besset, C., Guillot, A., Gitton, C., and Monnet, V. (2009). The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J. Bacteriol.* 191, 4647–4655. doi: 10.1128/JB.00257-09
- Guo, Q., Ahn, S. J., Kaspar, J., Zhou, X., and Burne, R. A. (2014). Growth phase and pH influence peptide signaling for competence development in *Streptococcus mutans*. *J. Bacteriol.* 196, 227–236. doi: 10.1128/JB.00995-13
- Haijema, B. J., Hahn, J., Haynes, J., and Dubnau, D. (2001). A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol. Microbiol.* 40, 52–64. doi: 10.1046/j.1365-2958.2001.02363.x
- Hamoen, L. W., Venema, G., and Kuipers, O. P. (2003). Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology* 149, 9–17. doi: 10.1099/mic.0.26003-0
- Hoover, S. E., Perez, A. J., Tsui, H. C., Sinha, D., Smiley, D. L., DiMarchi, R. D., et al. (2015). A new quorum-sensing system (TprA/PhrA) for *Streptococcus pneumoniae* D39 that regulates a lantibiotic biosynthesis gene cluster. *Mol. Microbiol.* 97, 229–243. doi: 10.1111/mmi.13029
- Jabbari, S., Heap, J. T., and King, J. R. (2011). Mathematical modelling of the sporulation-initiation network in *Bacillus subtilis* revealing the dual role of the putative quorum-sensing signal molecule PhrA. *Bull. Math. Biol.* 73, 181–211. doi: 10.1007/s11538-010-9530-7
- Jakobs, M., and Meinhardt, F. (2015). What renders Bacilli genetically competent? A gaze beyond the model organism. *Appl. Microbiol. Biotechnol.* 99, 1557–1570. doi: 10.1007/s00253-014-6316-0
- Johnsborg, O., Eldholm, V., and Håvarstein, L. S. (2007). Natural genetic transformation: prevalence, mechanisms and function. *Res. Microbiol.* 158, 767–778. doi: 10.1016/j.resmic.2007.09.004
- Johnsen, P. J., Dubnau, D., and Levin, B. R. (2009). Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. *Genetics* 181, 1521–1533. doi: 10.1534/genetics.108.099523
- Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J. P. (2014). Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* 12, 181–196. doi: 10.1038/nrmicro3199
- Karlsson, D., Karlsson, S., Gustafsson, E., Normark, B. H., and Nilsson, P. (2007). Modeling the regulation of the competence-evoking quorum sensing network in *Streptococcus pneumoniae*. *Biosystems* 90, 211–223. doi: 10.1016/j.biosystems.2006.08.005
- Khan, R., Rukke, H. V., Ricomini Filho, A. P., Fimland, G., Arntzen, M. Ø., Thiede, B., et al. (2012). Extracellular identification of a processed type II ComR/ComS pheromone of *Streptococcus mutans*. *J. Bacteriol.* 194, 3781–3788. doi: 10.1128/JB.00624-12
- Kim, J. N., Stanhope, M. J., and Burne, R. A. (2013). Core-gene-encoded peptide regulating virulence-associated traits in *Streptococcus mutans*. *J. Bacteriol.* 195, 2912–2920. doi: 10.1128/JB.00189-13
- Koga, K., Harada, T., Shimizu, H., and Tanaka, K. (2005). Bacterial luciferase activity and the intracellular redox pool in *Escherichia coli*. *Mol. Genet. Genomics* 274, 180–188. doi: 10.1007/s00438-005-0008-5
- Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G., and Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J. Bacteriol.* 177, 7011–7018.

- Leisner, M., Kuhr, J. T., Rädler, J. O., Frey, E., and Maier, B. (2009). Kinetics of genetic switching into the state of bacterial competence. *Biophys. J.* 96, 1178–1188. doi: 10.1016/j.bpj.2008.10.034
- Lenz, P., and Søgaard-Andersen, L. (2011). Temporal and spatial oscillations in bacteria. *Nat. Rev. Microbiol.* 9, 565–577. doi: 10.1038/nrmicro2612
- Letort, C., and Juillard, V. (2001). Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. *J. Appl. Microbiol.* 91, 1023–1029. doi: 10.1046/j.1365-2672.2001.01469.x
- Liebal, U. W., Millat, T., de Jong, I. G., Kuipers, O. P., Völker, U., and Wolkenhauer, O. (2010). How mathematical modelling elucidates signalling in *Bacillus subtilis*. *Mol. Microbiol.* 77, 1083–1095. doi: 10.1111/j.1365-2958.2010.07283.x
- Luo, P., and Morrison, D. A. (2003). Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*. *J. Bacteriol.* 185, 349–358. doi: 10.1128/JB.185.1.349-358.2003
- Maamar, H., and Dubnau, D. (2005). Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol. Microbiol.* 56, 615–624. doi: 10.1111/j.1365-2958.2005.04592.x
- Martin, B., Soulet, A. L., Mirouze, N., Prudhomme, M., Mortier-Barrière, I., Granadel, C., et al. (2013). ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence). *Mol. Microbiol.* 87, 394–411. doi: 10.1111/mmi.12104
- Mashburn-Warren, L., Morrison, D. A., and Federle, M. J. (2010). A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an Rgg regulator. *Mol. Microbiol.* 78, 589–606. doi: 10.1111/j.1365-2958.2010.07361.x
- McMillen, D., Kopell, N., Hasty, J., and Collins, J. J. (2002). Synchronizing genetic relaxation oscillators by intercell signaling. *Proc. Natl. Acad. Sci. U.S.A.* 99, 679–684. doi: 10.1073/pnas.022642299
- Mirouze, N., Bergé, M. A., Soulet, A. L., Mortier-Barrière, I., Quentin, Y., Fichant, G., et al. (2013). Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1035–E1044. doi: 10.1073/pnas.1219868110
- Moradigaravand, D., and Engestrådt, J. (2013). The evolution of natural competence: disentangling costs and benefits of sex in bacteria. *Am. Nat.* 182, E112–E126. doi: 10.1086/671909
- Morrison, D. A., Guédon, E., and Renault, P. (2013). Competence for natural genetic transformation in the *Streptococcus bovis* group streptococci *S. infantarius* and *S. macedonicus*. *J. Bacteriol.* 195, 2612–2620. doi: 10.1128/JB.00230-13
- Mortier-Barrière, I., Velten, M., Dupaigne, P., Mirouze, N., Piétrement, O., McGovern, S., et al. (2007). A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* 130, 824–836. doi: 10.1016/j.cell.2007.07.038
- Novák, B., and Tyson, J. J. (2008). Design principles of biochemical oscillators. *Nat. Rev. Mol. Cell Biol.* 9, 981–991. doi: 10.1038/nrm2530
- Okinaga, T., Xie, Z., Niu, G., Qi, F., and Merritt, J. (2010). Examination of the hdrRM regulon yields insight into the competence system of *Streptococcus mutans*. *Mol. Oral Microbiol.* 25, 165–177. doi: 10.1111/j.2041-1014.2010.00574.x
- Pérez-Pascual, D., Gaudy, P., Fleuchot, B., Besset, C., Rosinski-Chupin, I., Guillot, A., et al. (2015). RovS and its associated signaling peptide form a cell-to-cell communication system required for *Streptococcus agalactiae* pathogenesis. *MBio* 6:e02306-14. doi: 10.1128/mBio.02306-14
- Perry, J. A., Jones, M. B., Peterson, S. N., Cvitkovitch, D. G., and Lévesque, C. M. (2009). Peptide alarmone signalling triggers an auto-active bacteriocin necessary for genetic competence. *Mol. Microbiol.* 72, 905–917. doi: 10.1111/j.1365-2958.2009.06693.x
- Proshkin, S., Rahmouni, A. R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328, 504–508. doi: 10.1126/science.1184939
- Qi, F., Kretz, J., Lévesque, C. M., Kay, O., Mair, R. W., Shi, W., et al. (2005). Peptide pheromone induced cell death of *Streptococcus mutans*. *FEMS Microbiol. Lett.* 251, 321–326. doi: 10.1016/j.femsle.2005.08.018
- Reck, M., Tomasch, J., and Wagner-Döbler, I. (2015). The alternative sigma factor SigX controls bacteriocin synthesis and competence, the two quorum

- sensing regulated traits in *Streptococcus mutans*. *PLoS. Genet.* 11:e1005353. doi: 10.1371/journal.pgen.1005353
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring, NY: Cold Spring Harbor Laboratory.
- Schultz, D., Ben, J. E., Onuchic, J. N., and Wolynes, P. G. (2007). Molecular level stochastic model for competence cycles in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17582–17587. doi: 10.1073/pnas.0707965104
- Schultz, D., Lu, M., Stavropoulos, T., Onuchic, J., and Ben-Jacob, E. (2013). Turning oscillations into opportunities: lessons from a bacterial decision gate. *Sci. Rep.* 3:1668. doi: 10.1038/srep01668
- Schultz, D., Wolynes, P. G., Ben, J. E., and Onuchic, J. N. (2009). Deciding fate in adverse times: sporulation and competence in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21027–21034. doi: 10.1073/pnas.0912185106
- Schwab, D. J., Baetica, A., and Mehta, P. (2012). Dynamical quorum-sensing in oscillators coupled through an external medium. *Physica D* 241, 1782–1788. doi: 10.1016/j.physd.2012.08.005
- Seitz, P., and Blokesch, M. (2013). Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol. Rev.* 37, 336–363. doi: 10.1111/j.1574-6976.2012.00353.x
- Sgro, A. E., Schwab, D. J., Noorbakhsh, J., Mestler, T., Mehta, P., and Gregor, T. (2015). From intracellular signaling to population oscillations: bridging size- and time-scales in collective behavior. *Mol. Syst. Biol.* 11, 779. doi: 10.15252/msb.20145352
- Singh, H., and Parmananda, P. (2013). Quorum sensing via static coupling demonstrated by Chua's circuits. *Phys. Rev. E. Stat. Nonlin. Soft. Matter Phys.* 88:040903. doi: 10.1103/PhysRevE.88.040903
- Skippington, E., and Ragan, M. A. (2011). Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol. Rev.* 35, 707–735. doi: 10.1111/j.1574-6976.2010.00261.x
- Son, M., Ahn, S. J., Guo, Q., Burne, R. A., and Hagen, S. J. (2012). Microfluidic study of competence regulation in *Streptococcus mutans*: environmental inputs modulate bimodal and unimodal expression of *comX*. *Mol. Microbiol.* 86, 258–272. doi: 10.1111/j.1365-2958.2012.08187.x
- Son, M., Ghoreishi, D., Ahn, S. J., Burne, R. A., and Hagen, S. J. (2015). Sharply tuned pH Response of genetic competence regulation in *Streptococcus mutans*: a microfluidic study of the environmental sensitivity of *comX*. *Appl. Environ. Microbiol.* 81, 5622–5631. doi: 10.1128/AEM.01421-15
- Sturme, M. H., Kleerebezem, M., Nakayama, J., Akkermans, A. D., Vaughan, E. E., and de Vos, W. M. (2002). Cell to cell communication by autoinducing peptides in gram-positive bacteria. *Antonie Van Leeuwenhoek* 81, 233–243. doi: 10.1023/A:1020522919555
- Tian, X. L., Dong, G., Liu, T., Gomez, Z. A., Wahl, A., Hols, P., et al. (2013). MecA protein acts as a negative regulator of genetic competence in *Streptococcus mutans*. *J. Bacteriol.* 195, 5196–5206. doi: 10.1128/JB.00821-13
- van de Guchte, M., van der Vossen, J. M., Kok, J., and Venema, G. (1989). Construction of a lactococcal expression vector: expression of hen egg white lysozyme in *Lactococcus lactis* subsp. *lactis*. *Appl. Environ. Microbiol.* 55, 224–228.
- Voigt, C. A., Wolf, D. M., and Arkin, A. P. (2005). The *Bacillus subtilis* *sin* operon: an evolvable network motif. *Genetics* 169, 1187–1202. doi: 10.1534/genetics.104.031955
- Wahl, A., Servais, F., Drubert, A. S., Foulon, C., Fontaine, L., and Hols, P. (2014). Control of natural transformation in salivarius *Streptococci* through specific degradation of sigmaX by the Meca-ClpCP protease complex. *J. Bacteriol.* 196, 2807–2816. doi: 10.1128/JB.01758-14
- Wei, H., and Håvarstein, L. S. (2012). Fratricide is essential for efficient gene transfer between pneumococci in biofilms. *Appl. Environ. Microbiol.* 78, 5897–5905. doi: 10.1128/AEM.01343-12
- Wenderska, I. B., Lukenda, N., Cordova, M., Magarvey, N., Cvitkovitch, D. G., and Senadheera, D. B. (2012). A novel function for the competence inducing peptide, XIP, as a cell death effector of *Streptococcus mutans*. *FEMS Microbiol. Lett.* 336, 104–112. doi: 10.1111/j.1574-6968.2012.02660.x
- Weng, L., Piotrowski, A., and Morrison, D. A. (2013). Exit from competence for genetic transformation in *Streptococcus pneumoniae* is regulated at multiple levels. *PLoS ONE* 8:e64197. doi: 10.1371/journal.pone.0064197

- Xi, H., Duan, L., and Turcotte, M. (2013). Point-cycle bistability and stochasticity in a regulatory circuit for *Bacillus subtilis* competence. *Math. Biosci.* 244, 135–147. doi: 10.1016/j.mbs.2013.05.002
- Xie, Z., Okinaka, T., Niu, G., Qi, F., and Merritt, J. (2010). Identification of a novel bacteriocin regulatory system in *Streptococcus mutans*. *Mol. Microbiol.* 78, 1431–1447. doi: 10.1111/j.1365-2958.2010.07417.x
- Youk, H., and Lim, W. A. (2014). Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science* 343:1242782. doi: 10.1126/science.1242782
- Young, R., and Bremer, H. (1975). Analysis of enzyme induction in bacteria. *Biochem. J.* 152, 243–254. doi: 10.1042/bj1520243
- Zaccaria, E., van Baarlen P., de Greef, A., Morrison, D. A., Smith, H., and Wells, J. M. (2014). Control of competence for DNA transformation in *Streptococcus suis* by genetically transferable phenotypes. *PLoS ONE* 9:e99394. doi: 10.1371/journal.pone.0099394

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Haustenne, Bastin, Hols and Fontaine. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Supplementary Material

Modeling of the ComRS signaling pathway reveals the limiting factors controlling competence in *Streptococcus thermophilus*

Laurie Haustenne¹, Georges Bastin², Pascal Hols^{1§} and Laetitia Fontaine^{1*§}

¹ Biochimie, Biophysique et Génétique des Microorganismes, Institut des Sciences de la Vie, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

² Center for Systems Engineering and Applied Mechanics, ICTEAM, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

***Correspondence:** Laetitia Fontaine; Biochimie, Biophysique et Génétique des Microorganismes; Institut des Sciences de la Vie, Université catholique de Louvain, 4-5 Place Croix du Sud (L7.07.06), 1348, Louvain-la-Neuve. laetitia.fontaine@uclouvain.be

[§] Both authors contributed equally to this work

1. Supplementary Data

1.1. Supplementary Data Sheets

Supplementary Data Sheet 1. | MATLAB file. This file contains the MATLAB code (as a text file) used for model building, validation, and outputs. Please copy-paste all the text in a new MATLAB file named 'ODE_model.m'. Supplementary Data Sheet 2 is required to run the model.

Supplementary Data Sheet 2. | Experimental data sheet. This xlsx file contains all the experimental data needed (growth and luciferase curves) to run the model. Please rename it as "Experimental_data.xlsx" and copy this file in the same folder as the MATLAB file 'ODE_model.m'.

2. Supplementary Figures and Tables

2.1. Supplementary Tables

Supplementary Table 1. | Bacterial strains and plasmids used in this study.

Strain	Genotype	Characteristic(s) ^a	Source or reference
<i>Streptococcus thermophilus</i>			
LMD-9	Wild type		ATCC ^b
CB001	LMD-9 (<i>blpD-blpX</i>):P _{comX} - <i>luxAB</i>		(Fontaine et al., 2010a)
LF125	CB001 <i>comX</i> ::P32- <i>cat</i>	Cm ^R	(Boutry et al., 2013)
LF118	CB001 <i>comS</i> ::P32- <i>cat</i>	Cm ^R	(Fontaine et al., 2013)
TIL1391	CB001 <i>comS</i> :: <i>spec</i>	Spc ^R	(Gardan et al., 2013)
LH001	CB001 <i>dprA</i> ::P32- <i>cat</i>	Cm ^R	This study
LF121	LMD-9 (<i>blpD-blpX</i>):P _{comS} - <i>luxAB</i>		This study
LF122	LF121 <i>comX</i> ::P32- <i>cat</i>	Cm ^R	(Boutry et al., 2013)
LF134	LF121 <i>comS</i> ::P32- <i>cat</i>	Cm ^R	(Fontaine et al., 2013)
LF135	LF121 <i>comR</i> ::P32- <i>cat</i>	Cm ^R	(Fontaine et al., 2013)
LH002	LF121 <i>dprA</i> ::P32- <i>cat</i>	Cm ^R	This study
LF123	LMD-9 (<i>blpD-blpX</i>):P _{comR} - <i>luxAB</i>		(Boutry et al., 2013)
LMG18311	Wild type		BBCM LMG ^c
CB009	LMG18311 (<i>blpU-blpX</i>):P _{comX} - <i>luxAB</i>		This study
LF146	LMG18311 (<i>blpU-blpX</i>):P _{comR} - <i>luxAB-spec</i>	Spc ^R	This study
<i>Escherichia coli</i>			
EC1000	MC1000 containing a copy of the <i>repA</i> gene of pWV01 in its chromosome	Kn ^R RepA ⁺	(Law et al., 1995)

Plasmid	Description	Characteristic(s) ^a	Source or reference
pNZ5319	pACYC184 derivative containing the <i>lox66-P32-cat-lox71</i> cassette	Cm ^R Em ^R	(Lambert et al., 2007)
pGIUD0855cat	pUC18 derivative used to assess natural transformation rates of <i>S. thermophilus</i> LMG18311 derivative strains	Ap ^R Cm ^R	(Fontaine et al., 2010b)
pR412	plasmid used to amplify the spectinomycin resistance cassette P _{spec} - <i>spec</i>	Spc ^R	(Martin et al., 2000)
pGICB004	pG+host9 derivative used to integrate the <i>luxAB</i> reporter genes by double cross-over at the <i>blp</i> locus of <i>S. thermophilus</i> LMG18311	Em ^R Ts	(Fleuchot et al., 2011)
pGICB001	pGICB004 derivative used to integrate the P _{comR} - <i>luxAB</i> transcriptional fusion by double cross-over at the <i>blp</i> locus of <i>S. thermophilus</i> LMG18311	Em ^R Ts	(Fontaine et al., 2010a)
pGILFspec	pGICB004 derivative containing the spectinomycin resistance cassette P _{spec} - <i>spec</i> downstream of <i>luxAB</i>	Em ^R Spc ^R Ts	This study
pGILFspec::P _{comR}	pGILFspec derivative used to introduce a P _{comR} - <i>luxAB</i> - <i>spec</i> transcriptional fusion at the <i>blp</i> locus of <i>S. thermophilus</i> LMG18311	Em ^R Spc ^R Ts	This study
pMG36eT	pMG36e derivative used for the constitutive expression of genes under the control of the P32 promoter	Em ^R	Fontaine and Hols, 2008
pMGcomRstrep	pMG36eT derivative used for the constitutive expression of <i>comR</i> _{LM9} ::streptagII fusion under the control of the P32 promoter	Em ^R	(Boutry et al., 2013)

^a Cm^R, Em^R, Spc^R and Ts indicate resistance to chloramphenicol, erythromycin and spectinomycin, and that the plasmid encodes a thermosensitive RepA protein, respectively.

^b ATCC, American Type Culture Collection, Rockville, MD.

^c BCCM Belgian Coordinated Collections of Microorganisms, LMG Laboratory of Microbiology and Genetics, Ghent, Belgium.

Supplementary Table 2. | Primers used for plasmid and strain construction.

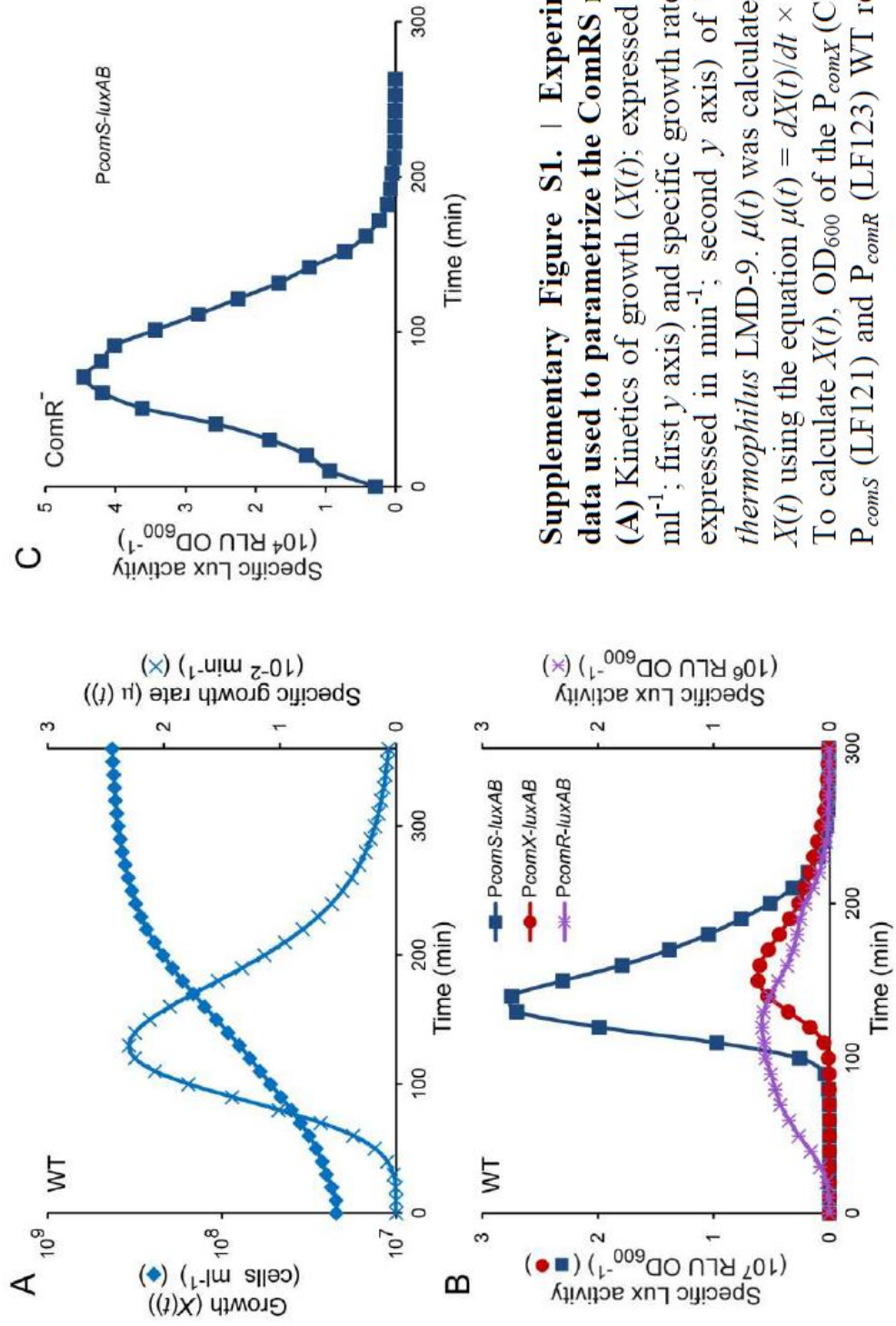
Primer	Sequence (5' → 3')	DNA target ^a	restriction sites ^b	name of constructed strain or plasmid	Source or reference
Primers used to amplify the <i>lox66-P32-cat-lox71</i> cassette ^c					
Uplox66	TAAGGAAGATAAATCCCATAGG	upstream <i>lox66</i> (plasmid pNZ5319)			(Fontaine et al., 2010b)
DNlox71	TTCACGTTACTAAAGGGAATGTA	downstream <i>lox71</i> (plasmid pNZ5319)			
Primers used to create overlap PCR products containing the <i>lox66-P32-cat-lox71</i> cassette. The PCR products are then transformed by competence to create mutant strains ^c					
UpdprA_A	CTCTGTCGGTTACCCAAATATTGCGTGCTC	upstream <i>dprA</i>			This study
UpdprA_B	CCTTATGGGATTATCTTCCTTATTCAAAGT TATTCATCTAAC			LH002	
DNdprA_A	TACATTCCCTTTAGTAACGTGAATCAGAATT TTCATAAAAATC	downstream <i>dprA</i>			
DNdprA_B	TTCGGTCCAAACACGACGAGCTTGCTGAG				
Primers used for the validation of mutant or reporter strains obtained by natural transformation					
ChdprA_A	TAAGAGTGCTATTGGTGTTCTCTTGC	upstream <i>dprA</i> (strains LH001 and LH002)			This study
ChdprA_B	TCATGGAAATTTACCTCAATTTCTTGC	downstream <i>dprA</i> (strains LH001 and LH002)		LH002	
INDISBAC1	TTAATGATAAACCAAGAAGAGTGG	<i>blpH</i> (CB009, LF123, LF146)		CB009, LF123, LF146	(Fontaine et al., 2007)
INTDISTIX2	TTATAACCAGTTCTGGCATGACCG	<i>pepX</i> (CB009, LF123, LF146)			This study

Primers used for plasmid construction					
LoxSpec-Sma-F	AAACCCGGGATAAGGAAGATAAATCCCAT AAGGATACCGTTCGTATAATGTATGCTATA CGAAGTTATTTCGATACCGTCGACCTCG	P _{spec-spec} cassette (plasmid pR412)	SmaI	pGILFspect	This study
LoxSpec-Pvu-R	AAATCAGCTGTTACGTTACTAAAGGGAAT GTAATACCGTTCGTATAGCATACATTATAC GAAGTTATAGCTCGAATTGACGCGGAATGG	P _{vuII}			
JUDPster0316A	GGACTAGTTGCATATTTTGTGGATAATCA AG	P _{comR} (S. <i>thermophilus</i> LMG18311)	SpeI	pGILFspect::PcomR	(Fontaine et al., 2013)
JUDPster0316B	CGGAATTCAAGTTCAAGAGAATCTCCTTTA		EcoRI		(Fontaine et al., 2013)
Primers used to amplify the Cy3-P _{comS} probe					
Cy3PcomS1A	CAGGAAAATTGGCAGATGGTTTATAG	P _{comS} (S. <i>thermophilus</i> LMG18311)			(Fontaine et al., 2013)
pMGSHPI6sur2	TTTCTGCAGTTACATTTTGGCATGATGGCTC C				

^a Refer to the hybridization target of primers. Unless otherwise stated (between brackets), template chromosomal DNA is from *S. thermophilus* LMD-9 WT.

^b Restriction sites introduced in primers for cloning purposes are underlined.

^c The sequence in the primer that allows overlapping with the *lox66-P32-cat-lox71* cassette is in bold type.

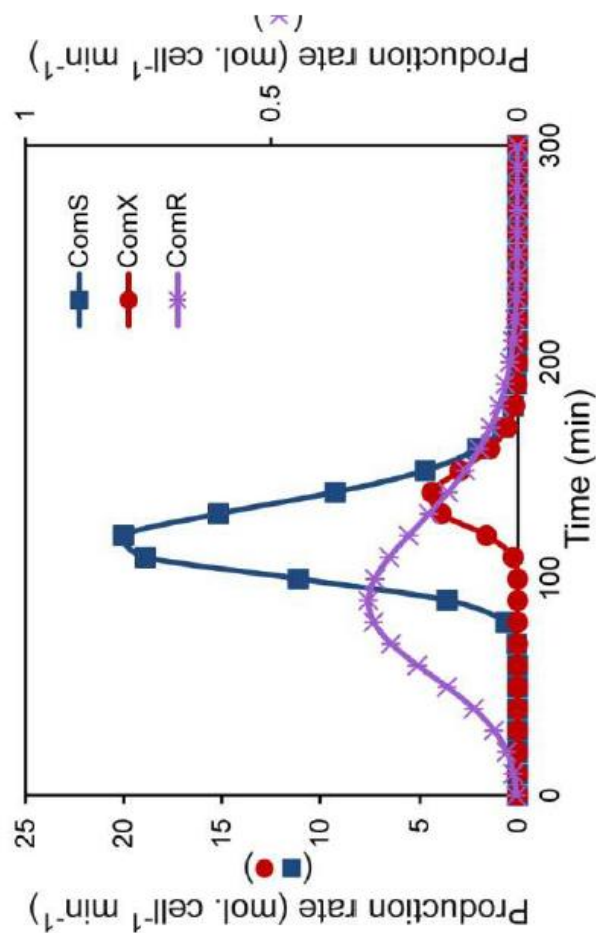


Supplementary Figure S1. | Experimental data used to parametrize the ComRS model.

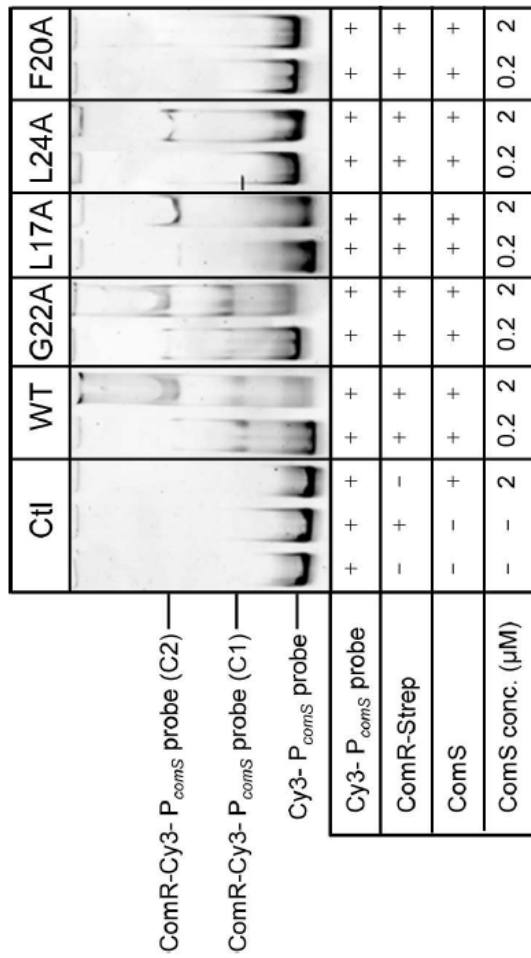
(A) Kinetics of growth ($X(t)$); expressed in cell ml^{-1} ; first y axis) and specific growth rate ($\mu(t)$); expressed in min^{-1} ; second y axis) of WT *S. thermophilus* LMD-9. $\mu(t)$ was calculated from $X(t)$ using the equation $\mu(t) = dX(t)/dt \times 1/X(t)$. To calculate $X(t)$, OD_{600} of the P_{comX} (CB001), P_{comS} (LF121) and P_{comR} (LF123) WT reporter

strains was measured every 10 min. The mean OD₆₀₀ values measured at each time (i.e. mean for the 3 reporter strains) were then converted in # cells ml⁻¹ by assuming that a OD₆₀₀ unit of 1 = 5×10^8 cells ml⁻¹ (see DataSheet 2). This conversion was deduced from a plating and CFU (colony forming unit) counting experiment. Culture samples were taken at regular intervals during growth, diluted, plated on M17L, and incubated at 37°C until CFU counting. These curves were used to model growth in system S1. **(B)** Kinetics of specific luciferase activity (RLU OD₆₀₀⁻¹) during growth of the LMD-9 reporter strains bearing the transcriptional fusion $P_{comS}\text{-}luxAB$ (LF121), $P_{comX}\text{-}luxAB$ (CB001) (first y axis) and $P_{comR}\text{-}luxAB$ (LF123) (second y axis). These curves were used to replace theoretical

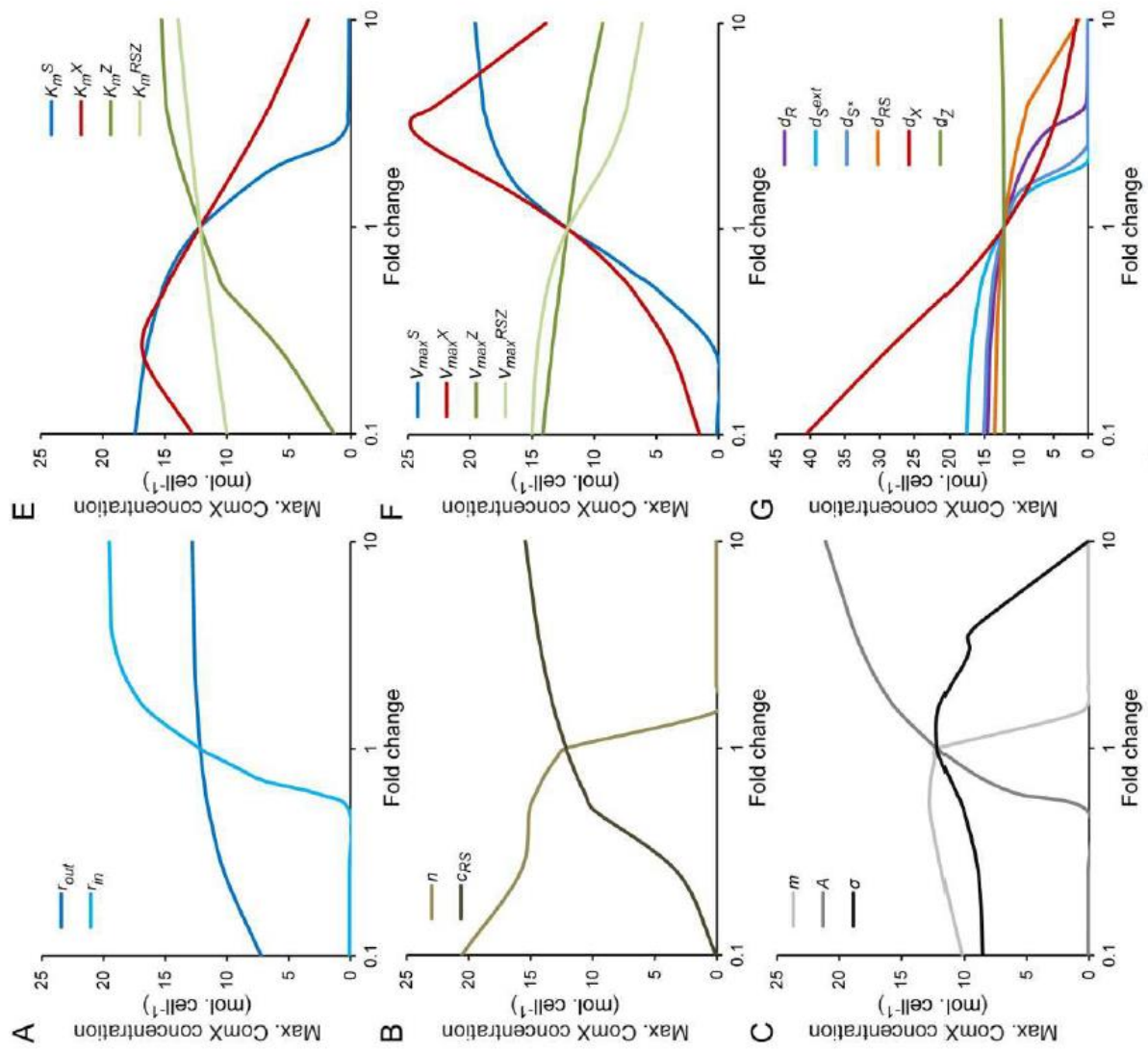
time-varying production rates of ComS, ComX and ComR, respectively, in system S2, and to parametrize the modeled activation terms of system S1 (see Materials and Methods). **(C)** Kinetics of specific luciferase activity (RLU OD₆₀₀⁻¹) of the LMD-9 ComR⁺ reporter strain encoding a transcriptional $P_{comS}\text{-}luxAB$ fusion (strain LF135). This curve was used to calculate the theoretical time-varying basal production rate of ComS ($b_S(t)$), ComX ($b_X(t)$) and ComR ($b_R(t)$) in system S1 using equation (14) (see Materials and methods). All strains were grown in CDML. Each curve represents the average from three independent repeats. Curves of panels A, B and C were adapted from Fontaine et al. (2013).

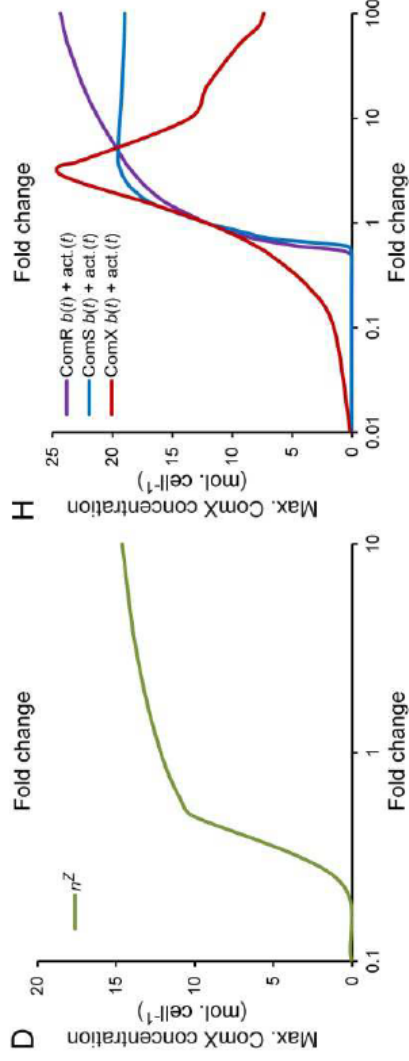


Supplementary Figure S2. | Experimental production rates. Kinetics of experimental production rates (expressed in molecules (mol.) cell⁻¹ min⁻¹) of ComS, ComX (first y axis), and ComR (second y axis). These production rates were deduced from the specific luciferase activities driven from promoters P_{comS} , P_{comX} , and P_{comR} , respectively (see Supplementary Figure S1B) using equation (13). These experimental production rates were used to replace modeled time-varying production rates of ComS, ComX and ComR, respectively, in system S2, and to parametrize the modeled activation terms of system S1 (see Materials and Methods for detailed information).

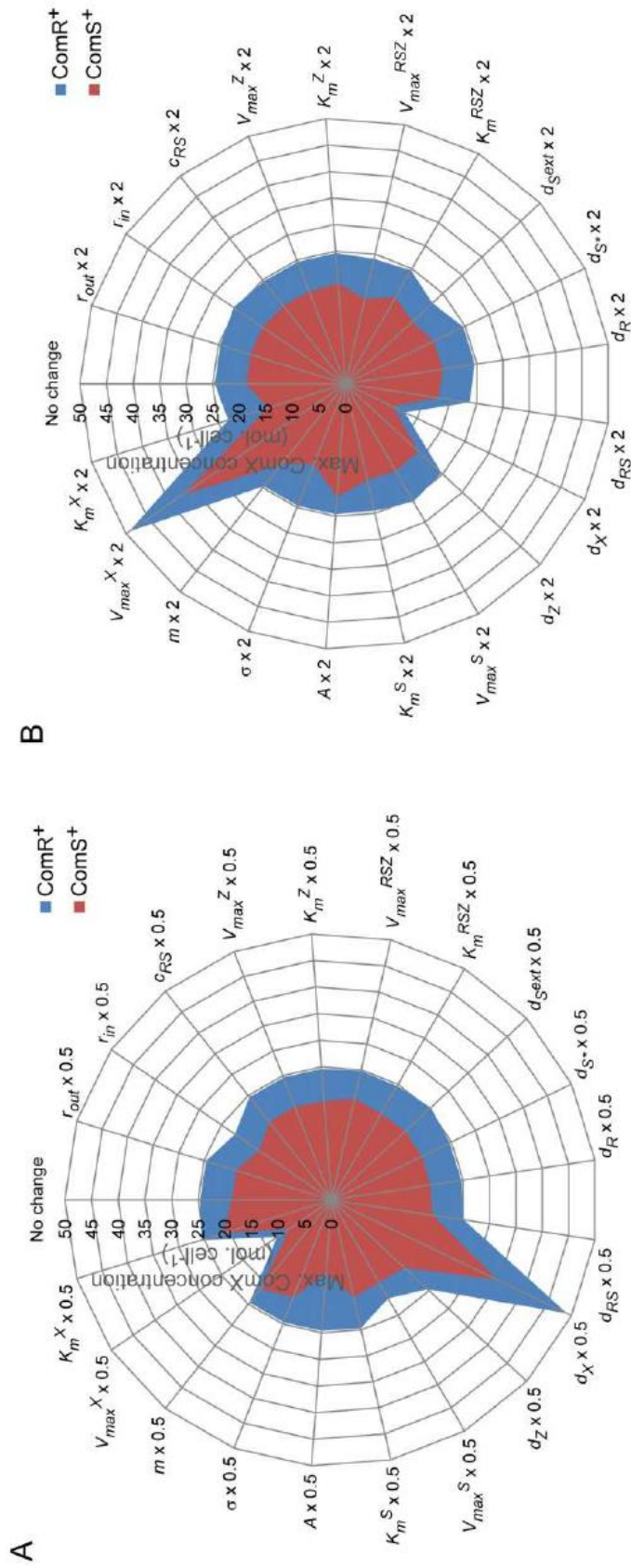


Supplementary Figure S3. | Efficiency of ComR binding to the promoter region of *comS* in presence of different ComS₁₇₋₂₄ variants. EMSA experiments showing the binding of purified ComR-Strep to promoter P_{comS}. 150 ng Cy3-P_{comS} probe (encompassing the entire P_{comS} region) was incubated in the presence (+) or absence (-) of purified ComR_{MD-9}-Strep (4 μM) and ComS₁₇₋₂₄ (0.2 and 2 μM) before separation on a TBE gel under non-denaturing conditions. Unmodified (WT) or modified ComS₁₇₋₂₄ octapeptide containing one alanine substitution at position 22 (G22A), 17 (L17A), 24 (L24A) and 20 (F20A) of the full-length ComS peptide. These variants (from left to right) display a decreasing ability to stimulate binding of ComR to P_{comS}. Ctl, negative controls; C1: protein-DNA complex 1; C2: protein-DNA complex 2.

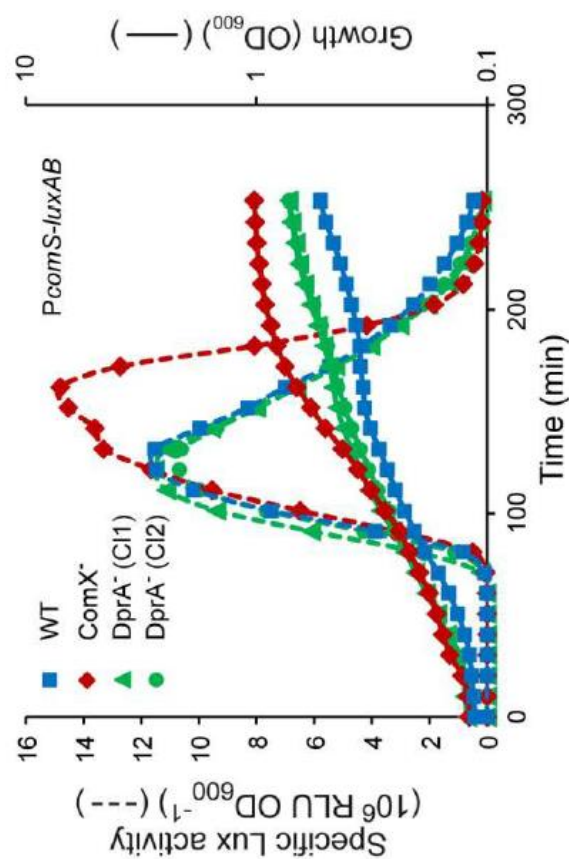




Supplementary Figure S4. | Sensitivity analysis of system S1. Maximum ComX concentration (molecules (mol.) cell⁻¹) computed when varying a given parameter value in system S1 within a fold change interval ranging of 0.1- to 10 times its assigned value (0.01 to 100 in panel H). Parameters were grouped according to their nature or specific role in ODE equations. They describe **(A)** ComS importation and exportation, **(B)** ComRS complexes formation, **(C)** the activation of ComR production, **(D)** the cooperativity of complex formation (hill coefficient), **(E)** the affinity of interactions (K_m constants), **(F)** the maximum rate of reactions (V_{max}), **(G)** degradation of modeled molecular species, and **(H)** production of ComR, ComS and ComX due to basal production $b(t)$ and activation ($\text{act.}(t)$). Parameters and their assigned value in system S1 are summarized in Table 2.



Supplementary Figure S5. | Impact of parameter values of system S1 on the maximum ComX concentration (molecules (mol.) cell⁻¹) computed for the simulated ComR⁺ and ComS⁺ strains. ComR⁺ (blue surface) and ComS⁺ (red surface) strains were simulated by increasing the maximal basal production rate of variables ComR(*t*) and ComS(*t*) by a constant inducing factor (IF) of 100. Then, the numerical value of 20 parameters of system S1 was individually varied within a fold change interval ranging from 0.5 to 2 times its assigned value. **(A)** Numerical value of each parameter multiplied by 0.5 (× 0.5). **(B)** Numerical value of each parameter multiplied by 2 (× 2). Parameters and their values in system S1 are listed in **Table 2**. This figure shows that higher ComX concentrations are reached in the ComR⁺ vs. ComS⁺ cells for each modified numerical value.



Supplementary Figure S6. | Role of DprA in competence shut-off. Kinetics of specific luciferase activity (first y axis; RLU OD_{600}^{-1}) and growth (second y axis; OD_{600}) of different derivatives of strain LF121 (carrying the *PcomS-luxAB* transcriptional fusion) grown in CDM medium: WT (LF121), DprA⁻ (LH002, clones C11 and C12) and ComX⁻ (LF122).

References

- Boutry, C., Delplace, B., Clippe, A., Fontaine, L., and Hols, P. (2013). SOS response activation and competence development are antagonistic mechanisms in *Streptococcus thermophilus*. *J. Bacteriol.* 195, 696-707.
- Fleuchot, B., Gitton, C., Guillot, A., Vidic, J., Nicolas, P., Besset, C., et al. (2011). Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol. Microbiol.* 80, 1102-1119.
- Fontaine, L., Boutry, C., de Frahan, M. H., Delplace, B., Fremaux, C., Horvath, P., et al. (2010a). A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* 192, 1444-1454.
- Fontaine, L., Boutry, C., Guedon, E., Guillot, A., Ibrahim, M., Grossiord, B., et al. (2007). Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. *J. Bacteriol.* 189, 7195-7205.
- Fontaine, L., Dandoy, D., Boutry, C., Delplace, B., de Frahan, M. H., Fremaux, C., et al. (2010b). Development of a versatile procedure based on natural transformation for marker-free targeted genetic modification in *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* 76, 7870-7877.
- Fontaine, L., Goffin, P., Dubout, H., Delplace, B., Baulard, A., Lecat-Guillet, N., et al. (2013). Mechanism of competence activation by the ComRS signalling system in streptococci. *Mol. Microbiol.* 87, 1113-1132.
- Gardan, R., Besset, C., Gitton, C., Guillot, A., Fontaine, L., Hols, P., et al. (2013). Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J. Bacteriol.* 195, 1845-1855.
- Lambert, J. M., Bongers, R. S., and Kleerebezem, M. (2007). Cre-lox-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. *Appl. Environ. Microbiol.* 73, 1126-1135.
- Law, J., Buist, G., Haandrikman, A., Kok, J., Venema, G., and Leenhouts, K. (1995). A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J. Bacteriol.* 177, 7011-7018.
- Martin, B., Prudhomme, M., Alloing, G., Granadel, C., and Claverys, J. P. (2000). Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. *Mol. Microbiol.* 38, 867-878.

In this second chapter, we present a general methodology aiming at analyzing the stability of equilibria in genetic networks.

This original methodology can be applied to the most general case in which both the mRNA and protein levels are modeled. Moreover, the methodology takes into account the potential time delays that occur with transcription and translation processes. It is also valid for very general forms of activation functions (including Hill-functions).

Two complementary approaches are presented to study the properties of the steady points of any set of interacting genes.

CHAPTER II

**On stability analysis of genetic regulatory networks represented by
delay-differential equations**

On stability analysis of genetic regulatory networks represented by delay-differential equations

Laurie Haustenne* Georges Bastin** Jean-Michel Coron***
Laetitia Fontaine* Pascal Hols*

* *Biochimie et Génétique Moléculaire Bactérienne,
Institut des Sciences de la Vie,
Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium.*

** *Department of Mathematical Engineering,
ICTEAM, Université catholique de Louvain,
4, Avenue G. Lemaitre, 1348 Louvain-La-Neuve, Belgium.*

*** *Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie,
Institut Universitaire de France,
4, Place Jussieu, 75252 Paris Cedex, France.*

Abstract: The purpose of this note is to give a short tutorial presentation of general methods for the stability analysis of the steady-states of genetic regulatory networks represented by delay-differential equations. Two complementary ways are investigated: a frequency domain approach which applies to linearized models and gives necessary and sufficient conditions for the local steady-state stability, a Lyapunov approach derived from the viewpoint of hyperbolic PDE systems which gives sufficient conditions for global stability. These methods are illustrated with the example of the “toggle switch”.

Keywords: Genetic regulatory network, Hyperbolic system, Lyapunov function, Stability.

1. INTRODUCTION

We are concerned with continuous time models of genetic networks when “time delays are included to allow for the time required for transcription, translation, and transport” (Smolen et al. [2000]). The network dynamics are therefore represented by delay-differential equations whose solutions evolve in the positive orthant. It is well known that these systems may have multiple steady-states or equilibria. A critical issue is to determine the stability of these equilibria. The purpose of this paper is to give a concise, but instrumental, presentation of general methods which can be used for the stability analysis of the steady-states of genetic regulatory networks represented by delay-differential equations. Two complementary ways are investigated: a frequency domain approach which applies to linearized models and gives necessary and sufficient conditions for the local steady-state stability, a Lyapunov approach derived from the viewpoint of hyperbolic PDE systems which gives sufficient conditions for global stability. These methods are illustrated with the example of the “toggle switch”.

★ GB and JMC are partially supported by the ERC advanced grant 266907 (CPDENL) of the 7th Research Framework Programme (FP7). GB is also partially supported by the Belgian Programme on Interuniversity Attraction Poles (IAP VII/19).

2. MATHEMATICAL MODEL

We consider a genetic regulatory network which involves n genes interconnected through activator or repressor proteins. As illustrated in Fig.1, the expression of the i -th gene in the network ($i = 1, \dots, n$) is represented by the following standard delay-differential system (see e.g. Bernot et al. [2013]):

$$\begin{aligned} \frac{dM_i(t)}{dt} &= b_i + h_i(P_k(t - \tau_k)) - \delta_i M_i(t), \\ \frac{dP_i(t)}{dt} &= \alpha_i M_i(t - \tau_{n+i}) - \beta_i P_i(t), \end{aligned} \quad (1)$$

where, at time t , $M_i(t)$ is the density of mRNA molecules and P_i the density of proteins expressed by the i -th gene. The constants b_i and α_i denote respectively the basal transcription rate and the specific translation rate. The constants β_i and δ_i are the natural degradation rate coefficients. The constant delays τ_i and τ_{n+i} are the times needed for transcription and translation respectively. The function $h_i(P_k)$ describes how the transcription of the i -th gene is activated or repressed by the density P_k of the protein expressed by the k -th gene of the network. It may therefore be either an activation Hill function of the form

$$h_i(P_k) = \frac{v_i P_k^{m_i}}{k_i^{m_i} + P_k^{m_i}}, \quad (2)$$

or an inhibition Hill function of the form

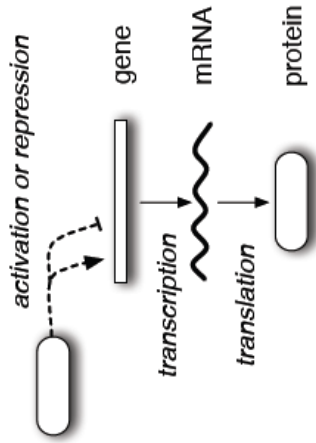


Fig. 1. Scheme of genetic transcription and translation with activation or repression.

$$h_i(P_k) = \frac{v_i k_i^{m_i}}{k_i^{m_i} + P_k^{m_i}}, \quad (3)$$

where v_i , k_i and m_i are constant parameters with v_i the maximal transcription rate, k_i the half-saturation coefficient and m_i the so-called Hill coefficient.

A steady-state of the system is a constant solution M_i^* , P_i^* , $i = 1, \dots, n$, of the dynamical system (1) i.e. a solution of the algebraic system

$$\begin{aligned} b_i + h_i(P_k^*) - \delta_i M_i^* &= 0, \\ \alpha_i M_i^* - \beta_i P_i^* &= 0. \end{aligned} \quad (4)$$

Let us define the deviations of P_i and M_i with respect to a steady-state M_i^* , P_i^* :

$$m_i(t) = M_i(t) - M_i^*, \quad p_i(t) = P_i(t) - P_i^*. \quad (5)$$

With these coordinates, the model (1) is alternatively written under the form

$$\frac{dm_i(t)}{dt} = g_i(p_k(t - \tau_k))p_k(t - \tau_k) - \delta_i m_i(t), \quad (6)$$

$$\frac{dp_i(t)}{dt} = \alpha_i m_i(t - \tau_{n+i}) - \beta_i p_i(t),$$

where the function g_i is defined such that

$$g_i(p)p = h_i(P^* + p) - h_i(P^*). \quad (7)$$

The linearization, around zero, of the system (6) is then given by

$$\begin{aligned} \frac{dm_i(t)}{dt} &= g_i(0)p_k(t - \tau_k) - \delta_i m_i(t), \\ \frac{dp_i(t)}{dt} &= \alpha_i m_i(t - \tau_{n+i}) - \beta_i p_i(t). \end{aligned} \quad (8)$$

In the next section we study the stability of the linear system (8) in the frequency domain.

3. FREQUENCY DOMAIN STABILITY

We introduce the following vector notations:

$$\mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}, \quad (9)$$

Using the Laplace transform, the linear system (8) is written in the frequency domain as

$$\begin{aligned} \mathbf{m} &= (sI_n + D)^{-1} F(s) \mathbf{p}, \\ \mathbf{p} &= (sI_n + B)^{-1} A E(s) \mathbf{m}, \end{aligned} \quad (10)$$

with the following matrix definitions:

$$\begin{aligned}
A &= \text{diag}\{\alpha_1, \dots, \alpha_n\}, \\
B &= \text{diag}\{\beta_1, \dots, \beta_n\}, \\
D &= \text{diag}\{\delta_1, \dots, \delta_n\}, \\
E(s) &= \text{diag}\{e^{-s\tau_{n+1}}, \dots, e^{-s\tau_{2n}}\}, \\
F(s) &= \text{matrix with entry } [F(s)]_{ik} = g_i(0)e^{-s\tau_k} \text{ if } k \sim i \\
&\quad \text{and 0 otherwise.}
\end{aligned}$$

The notation “diag” means that the matrix is diagonal. The notation $k \sim i$ means that the protein expressed by the k -th gene is an activator or a repressor of the i -th gene transcription.

The system (10) is a feedback system as shown in Fig.2. The poles of the system are the roots of the characteristic

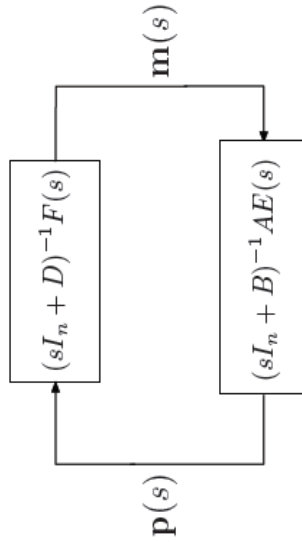


Fig. 2. The linearized model of a genetic network viewed as a feedback system.

equation

$$\det [I - (sI_n + B)^{-1}AE(s)(sI_n + D)^{-1}F(s)] = 0. \quad (11)$$

The following theorem gives the condition for the system stability.

Theorem 1. *The linear system (8) is exponentially stable if and only if the poles of the system have strictly negative real parts and are bounded away from zero.*

This theorem directly follows, as a special case, from [Hale and Verduyn-Lunel, 2002, Section 3] and [Michiels and Niculescu, 2007, Section 1.2]).

In the next section, as an example of how this theorem can be used, we consider the model of a “toggle switch”.

4. EXAMPLE: THE TOGGLE SWITCH

A toggle switch is a system of two genes that repress each other (see e.g. Smits et al. [2008], Veening et al. [2008] and the references therein). In the case of the toggle switch, the general model (1) is specialized as follows:

$$\begin{aligned}
\frac{dM_1(t)}{dt} &= b_1 + h_1(P_2(t - \tau_2)) - \delta_1 M_1(t), \\
\frac{dP_1(t)}{dt} &= \alpha_1 M_1(t - \tau_3) - \beta_1 P_1(t), \\
\frac{dM_2(t)}{dt} &= b_2 + h_2(P_1(t - \tau_1)) - \delta_2 M_2(t), \\
\frac{dP_2(t)}{dt} &= \alpha_2 M_2(t - \tau_4) - \beta_2 P_2(t),
\end{aligned} \quad (12)$$

with

$$\begin{aligned}
h_1(P_2) &= \frac{v_1 k_1^{m_1}}{k_1^{m_1} + P_2^{m_1}}, \\
h_2(P_1) &= \frac{v_2 k_2^{m_2}}{k_2^{m_2} + P_1^{m_2}}.
\end{aligned} \quad (13)$$

Depending on the numerical values of the constant coefficients and parameters, the system may have 1, 2 or 3 steady-states. The linearization around any of these steady-states is written

$$\begin{aligned}
\frac{dm_1(t)}{dt} &= g_1(0)p_2(t - \tau_2) - \delta_1 m_1(t), \\
\frac{dp_1(t)}{dt} &= \alpha_1 m_1(t - \tau_3) - \beta_1 p_1(t), \\
\frac{dm_2(t)}{dt} &= g_2(0)p_1(t - \tau_1) - \delta_2 m_2(t), \\
\frac{dp_2(t)}{dt} &= \alpha_2 m_2(t - \tau_4) - \beta_2 p_2(t).
\end{aligned} \tag{14}$$

Then, using formula (11), it is a matter of a few computations to show that the characteristic equation is

$$\underbrace{s^4 + a_3 s^3 + a_2 s^2 + a_1 s^1 + a_0 - \alpha_1 \alpha_2 g_1(0) g_2(0)}_{\psi(s)} e^{-s\tau} = 0$$

with

$$\begin{aligned}
\tau &= \tau_1 + \tau_2 + \tau_3 + \tau_4, \\
a_0 &= \beta_1 \beta_2 \delta_1 \delta_2, \\
a_1 &= (\beta_1 + \beta_2) \delta_1 \delta_2 + \beta_1 \beta_2 (\delta_1 + \delta_2), \\
a_2 &= \beta_1 \beta_2 + \delta_1 \delta_2 + (\beta_1 + \beta_2) (\delta_1 + \delta_2), \\
a_3 &= \beta_1 + \beta_2 + \delta_1 + \delta_2.
\end{aligned}$$

In order to analyze the stability of the poles in function of the time-delay τ , we follow the procedure of Walton and Marshall [1987].

The first step is to examine the stability when $\tau = 0$. In that case, it follows from a straightforward application of the Routh-Hurwitz criterion that the poles are stable if and only if

$$a_0 - \alpha_1 \alpha_2 g_1(0) g_2(0) > 0, \tag{15}$$

$$a_1 a_2 a_3 > a_1^2 + a_3^2 (a_0 - \alpha_1 \alpha_2 g_1(0) g_2(0)). \tag{16}$$

In the second step, we compute the following polynomial in ω^2 :

$$\begin{aligned}
W(\omega^2) &= \psi(j\omega)\psi(-j\omega) - \chi^2 \\
&= \omega^8 + b_3 \omega^6 + b_2 \omega^4 + b_1 \omega^2 + b_0
\end{aligned}$$

with

$$\begin{aligned}
b_0 &= a_0^2 - \chi^2, \\
b_1 &= a_1^2 - 2a_0 a_2, \\
b_2 &= a_2^2 - 2(a_0 + a_1 a_3), \\
b_3 &= 2a_2 + a_3.
\end{aligned}$$

Then, if conditions (15) and (16) are satisfied and if $W(\omega^2)$ has no positive real roots, the system (14) is stable for all $\tau \geq 0$. In contrast, if $W(\omega^2)$ has no negative and positive real roots, there is a maximum value of τ beyond which the system becomes unstable.

5. LYAPUNOV STABILITY

Since a genetic regulatory network is inherently a non-linear system having, possibly, multiple steady-states, the

frequency domain approach, based on linearization, has a fundamental limitation: it gives conditions that ensure only the *local* stability of the steady-states. By this, it is meant that if the initial conditions are slightly perturbed, then the system returns to the equilibrium. In order to determine larger domains of convergence (and even to guarantee *global* stability in the case where the genetic network has a single steady-state), it is more appropriate (and less easy !) to use the Lyapunov approach. For this purpose, we introduce the auxiliary independent variable $z \in [0, 1]$, and we consider that the protein and mRNA densities $p_i = P_i - P_i^*$ and $m_i = M_i - M_i^*$ now depend on the two variables t and z and satisfy the partial differential equations

$$\begin{aligned} \partial_t p_i(t, z) + \frac{1}{\tau_i} \partial_z p_i(t, z) &= 0, \\ \partial_t m_i(t, z) + \frac{1}{\tau_{n+i}} \partial_z m_i(t, z) &= 0, \end{aligned} \quad (17)$$

under the boundary conditions

$$\begin{aligned} \frac{dm_i(t, 0)}{dt} &= g_i(p_k(t, 1))p_k(t, 1) - \delta_i m_i(t, 0), \\ \frac{dp_i(t, 0)}{dt} &= \alpha_i m_i(t, 1) - \beta_i p_i(t, 0), \end{aligned} \quad (18)$$

The model (17)-(18) is totally identical to the delay-differential model (6) which has been used above since equations (17) imply readily that

$$p_i(t, 1) = p_i(t - \tau_i, 0), \quad m_i(t, 1) = m_i(t - \tau_{n+i}, 0). \quad (19)$$

We introduce the following candidate Lyapunov function

$$\begin{aligned} V = \int_0^1 & \left[\sum_{i=1}^n q_i \tau_i p_i^2(t, z) e^{-\mu \tau_i z} \right. \\ & + \sum_{i=1}^n q_{n+i} \tau_{n+i} m_i^2(t, z) e^{-\mu \tau_{n+i} z} \left. \right] dz \\ & + \frac{1}{2} \sum_{i=1}^n w_i p_i^2(t, 0) + \frac{1}{2} \sum_{i=1}^n w_{n+i} m_i^2(t, 0), \end{aligned} \quad (20)$$

with positive parameters μ , q_i , q_{n+i} , w_i , w_{n+i} to be determined. With matrix notations, this function is rewritten

$$\begin{aligned} V = \int_0^1 & [\mathbf{p}^T(t, z) Q_1 T_1 E_1(\mu, z) \mathbf{p}(t, z) \\ & + \mathbf{m}^T(t, z) Q_2 T_2 E_2(\mu, z) \mathbf{m}(t, z)] dz \\ & + \frac{1}{2} [\mathbf{p}^T(t, 0) W_1 \mathbf{p}(t, 0) + \mathbf{m}^T(t, 0) W_2 \mathbf{m}(t, 0)], \end{aligned} \quad (21)$$

with

$$\begin{aligned} Q_1 &= \text{diag}\{q_i, i = 1, \dots, n\}, \\ Q_2 &= \text{diag}\{q_{n+i}, i = 1, \dots, n\}, \\ T_1 &= \text{diag}\{\tau_i, i = 1, \dots, n\}, \\ T_2 &= \text{diag}\{\tau_{n+i}, i = 1, \dots, n\}, \\ E_1(\mu, z) &= \text{diag}\{e^{-\mu \tau_i z}, i = 1, \dots, n\}, \\ E_2(\mu, z) &= \text{diag}\{e^{-\mu \tau_{n+i} z}, i = 1, \dots, n\}, \\ W_1 &= \text{diag}\{w_i, i = 1, \dots, n\}, \\ W_2 &= \text{diag}\{w_{n+i}, i = 1, \dots, n\}. \end{aligned} \quad (22)$$

Using integration by parts, it can be shown that the time derivative of V , along the system solutions, is

$$\frac{dV}{dt} = -\mu V - \mathbf{X}^T \mathcal{M}(\mu, \mathbf{p}(t, 1)) \mathbf{X}, \quad (23)$$

with the vector $\mathbf{X}$ and the matrix $\mathcal{M}(\mu, \mathbf{p})$ defined as

$$\mathbf{X}^T = (\mathbf{p}^T(t, 0) \quad \mathbf{m}^T(t, 0) \quad \mathbf{p}^T(t, 1) \quad \mathbf{m}^T(t, 1))^T \quad (24)$$

$$\mathcal{M}(\mu, \mathbf{p}) = \begin{pmatrix} W_2 B - Q_1 & 0 & 0 & -W_2 A \\ 0 & W_2 D - Q_2 & -W_1 G(\mathbf{p}) & 0 \\ 0 & 0 & Q_1 E_1(\mu, 1) & 0 \\ 0 & 0 & 0 & Q_2 E_2(\mu, 1) \end{pmatrix} \quad (25)$$

We have the following theorem.

Theorem 2. *For $\eta > 0$, let $\mathcal{B}_\eta$ be an open ball of radius η in $\mathbb{R}^n$. There exists $\mu > 0$ sufficiently small such that V is a strict exponentially decreasing Lyapunov function along the solutions the system (17)-(18) with $\mathbf{p} \in \mathcal{B}_\eta$ if there exist $q_i > 0$, $q_{n+i} > 0$, $w_i > 0$, $w_{n+i} > 0$ such that the matrix $\mathcal{M}(0, \mathbf{p})$ is positive definite for all $\mathbf{p} \in \mathcal{B}_\eta$.*

In the next section, we use again the toggle switch as an example of how this theorem can be used.

6. EXAMPLE: THE TOGGLE SWITCH

For the toggle switch example, the general model (17)-(18) is specialized as follows:

$$\partial_t p_1(t, z) + \frac{1}{\tau_1} \partial_z p_1(t, z) = 0,$$

$$\partial_t p_2(t, z) + \frac{1}{\tau_2} \partial_z p_2(t, z) = 0,$$

$$\partial_t m_1(t, z) + \frac{1}{\tau_3} \partial_z m_1(t, z) = 0$$

$$\partial_t m_2(t, z) + \frac{1}{\tau_4} \partial_z m_2(t, z) = 0,$$

$$\frac{dm_1(t, 0)}{dt} = g_1(p_2(t, 1))p_2(t, 1) - \delta_1 m_1(t, 0),$$

$$\frac{dp_1(t, 0)}{dt} = \alpha_1 m_1(t, 1) - \beta_1 p_1(t, 0),$$

$$\frac{dm_2(t, 0)}{dt} = g_2(p_1(t, 1))p_1(t, 1) - \delta_2 m_2(t, 0),$$

$$\frac{dp_2(t, 0)}{dt} = \alpha_2 m_2(t, 1) - \beta_2 p_2(t, 0).$$

For this system, the matrix $\mathcal{M}(0, \mathbf{p})$ is

$$\mathcal{M}(0, \mathbf{p}) = \begin{pmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ 0 & \mathcal{M}_{22} \end{pmatrix}, \quad (26)$$

with

$$\mathcal{M}_{11} = \begin{pmatrix} w_3 \beta_1 - q_1 & 0 & 0 & 0 \\ 0 & w_4 \beta_2 - q_2 & 0 & 0 \\ 0 & 0 & w_1 \delta_1 - q_3 & 0 \\ 0 & 0 & 0 & w_2 \delta_2 - q_4 \end{pmatrix},$$

$$\mathcal{M}_{12} = \begin{pmatrix} 0 & 0 & -w_3 \alpha_1 & 0 \\ 0 & 0 & 0 & -w_4 \alpha_2 \\ 0 & -w_1 g_1(p_2) & 0 & 0 \\ -w_2 g_2(p_1) & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{M}_{22} = \begin{pmatrix} q_1 & 0 & 0 & 0 \\ 0 & q_2 & 0 & 0 \\ 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & q_4 \end{pmatrix}.$$

This matrix is positive definite if and only if the leading principal minors of the symmetric matrix $\mathcal{M}(0, \mathbf{p}) + \mathcal{M}^T(0, \mathbf{p})$ are all positive. This leads to the following inequalities:

$$\begin{aligned} 0 &< w_3\beta_1 - q_1 \\ 0 &< w_4\beta_2 - q_2 \\ 0 &< w_1\delta_1 - q_3 \\ 0 &< w_2\delta_2 - q_4 \\ 0 &< (w_1\delta_1 - q_3)q_2 - (1/4)w_1^2g_1^2(p_2) \\ 0 &< (w_2\delta_2 - q_4)q_1 - (1/4)w_2^2g_2^2(p_1) \\ 0 &< (w_3\beta_1 - q_1)q_3 - (1/4)w_3^2\alpha_1^2 \\ 0 &< (w_4\beta_2 - q_2)q_4 - (1/4)w_4^2\alpha_1^2. \end{aligned}$$

Hence, the system is stable for any τ_i if there exist positive values of q_i and w_i such that these inequalities are satisfied.

7. FINAL REMARKS

Remark 1. In this paper, we have presented two ways for analyzing the stability of steady-states in genetic regulatory networks represented by delay-differential equations: a frequency domain approach and a Lyapunov approach. The two methods are complementary. The advantage of the frequency domain method is to give necessary and sufficient conditions but the stability is local. In contrast, the Lyapunov method may give global stability results but the stability conditions are only sufficient and can be conservative in some instances.

Remark 2. In this paper, for simplicity, we have restricted the presentation to networks where each gene can be controlled by one protein at most. The analysis can easily be extended to situations where a gene can be controlled by several proteins simultaneously. In fact, both Theorems 1 and 2 hold exactly as they are, provided the definitions of the matrices $F(s)$ and $G(\mathbf{p})$ are adequately extended.

Remark 3. The Lyapunov method of this paper is derived from the viewpoint of hyperbolic systems and easy to implement (see Bastin and Coron [2014] for more details). An alternative approach, based on Lyapunov-Krasovskii functionals, has been used in other papers with a special emphasis on the robustness issue, see e.g. Wang et al. [2008], Ren and Cao [2008].

Remark 4. In the oral presentation of this communication, it is intended to give an application to the genetic network of the basic mechanism for the competence development in *Streptococcus thermophilus*, see Haustenne et al. [2015] and Fontaine et al. [2013].

REFERENCES

- G. Bastin and J.-M. Coron. *Stability and Boundary Stabilization of One-Dimensional Hyperbolic Systems*. Book in preparation, 2014.
- G. Bernot, J.-P. Comet, A. Richard, M. Chaves, J.-L. Gouzé, and F. Dayan. *Modeling and Analysis of Gene Regulatory Networks*, chapter 2 in Modeling in Computational Biology and Biomedicine: A Multidisciplinary Endeavor, F. Cazals and P. Kornprobst eds. Springer, 2013.
- L. Fontaine, P. Goffin, H. Dubout, B. Delplace, A. Baulard, N. Lecat-Guillet, E. Chambellon, R. Gardan, and P. Hols. Mechanism of competence activation by the comrs signalling system in streptococci. *Molecular Microbiology*, 87(6):1113–1132, 2013.
- J.K. Hale and S.M. Verduyn-Lunel. Strong stabilization of neutral functional differential equations. *IMA J. Math. Control Inform.*, 19(1-2):5–23, 2002.
- L. Haustenne, L. Fontaine, G. Bastin, and P. Hols. Mathematical modeling of the regulatory network governing competence for natural transformation in streptococcus thermophilus. Technical report, Université catholique de Louvain, 2015.
- W. Michiels and S.-I. Niculescu. *Stability and Stabilization of Time-Delay Systems*. Advances in Design and Control. SIAM, 2007.
- F. Ren and J. Cao. Asymptotic and robust stability of genetic regulatory networks with time-varying delays. *Neurocomputing*, 71:834–842, 2008.
- W.K. Smits, J.-W. Veening, and O.P. Kuipers. *Phenotypic Variation and Bistable Switching in Bacteria*, chapter 12 in Bacterial Physiology: A Molecular Approach, W. El-Sharoud ed. Springer-Verlag, 2008.
- P. Smolen, D.A. Baxter, and J.H. Byrne. Modeling transcriptional control in gene network - methods, recent results and future directions. *Bulletin of Mathematical Biology*, (62):247–292, 2000.
- J.W. Veening, W.K. Smits, and O.P. Kuipers. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Reviews in Microbiology*, (62):193–210, 2008.
- K. Walton and J.E. Marshall. Direct method for TDS stability analysis. *IEEE Proceedings - Control Theory and Applications*, 134(2):101–107, 1987.
- Z. Wang, H. Gao, J. Cao, and X. Liu. On delayed genetic regulatory networks with polytopic uncertainties: robust stability analysis. *IEEE Transactions on Nanobioscience*, 7(2):154–163, 2008.

In this last chapter, we apply the general methodology presented in chapter two to the equilibrium points of our initial model of the ComRS system consisting of DDE equations.

The analysis here presented shows that the system has two equilibrium points, one ON state (corresponding to the competence state) which is unstable, and one OFF state (corresponding to the normal, vegetative state) which is stable.

As explained before, our initial model is a DDE (Delay Differential Equations) model, with a slightly different set of equations and parameters.

Obviously, the differences in equations and parameter set between our model of chapter one and our model of chapter three could lead to different results for the stability analysis.

CHAPTER III

**Application of the stability analysis methodology to the ComRS
system in *Streptococcus thermophilus***

1. Framework

In order to set a formal framework for our stability analysis methodology of equilibria, let's first introduce some mathematical notations about the two theoretical concepts of "equilibrium" and "stability".

We define our **dynamic system** as:

$$x(t)' = f(x(t))$$

where t is the time, $x(t)$ is the vectorial notation for the state variables of the system, i.e. $x(t) = \{x_1(t), x_2(t), \dots, x_n(t)\}$, f is a map $R^n \rightarrow R^n$, and the prime superscript " ' " denotes the time derivative.

Since we have 8 equations in our model, we thus have $n = 8$. These state variables are all functions of the time (t), but for the sake of simplicity, we omit to write the time-dependence of each variable. We can thus re-write the system as:

$$x' = f(x) \tag{1}$$

For each time point t , the **state** of this dynamical system is totally defined by the vector x of state variables. It means that all the information about the system current state is included in this vector.

An **equilibrium** (or "steady-point") of a dynamic system is a state x^* of the system for which:

$$f(x^*) = 0$$

In other words, all the time-derivatives of the system are null at the steady points. It implies that if the system state is equal to x^* at time t_0 , then in absence of perturbations, it will remain in this equilibrium for all times:

$$x(t_0) = x^* \rightarrow x(t) = x^* \text{ for all } t \geq t_0$$

The steady-point x^* is a **stable equilibrium** of the system (1) if:

for all $\varepsilon > 0$, there is $\delta > 0$ such that $\|x(t_0) - x^*\| < \delta$ implies that $\|x(t) - x^*\| < \varepsilon$ for all $t \geq t_0$

2. The analyzed ODE system

In order to analyze the stability of the equilibria of our ComRS model, we applied the frequency domain approach to our initial DDE system:

$$\frac{dX(t)}{dt} = \mu(t)X(t) \quad (1)$$

$$\begin{aligned} \frac{dComS(t)}{dt} = & b_S(t) + \underbrace{V_{max}^S \left(\frac{ComRS^{n_s}(t - \tau_1)}{K_m^{S^{n_s}} + ComRS^{n_s}(t - \tau_1)} \right)}_{\text{activation by ComRS}} - \underbrace{r_{out}ComS(t)}_{\text{exportation}} - \underbrace{d_SComS(t)}_{\text{degradation}} \\ & - \underbrace{\mu(t)ComS(t)}_{\text{growth}} \end{aligned} \quad (2)$$

$$\frac{dComS^{ext}(t)}{dt} = \underbrace{r_{out}ComS(t)}_{\text{exportation}} - \underbrace{r_{in}ComS^{ext}(t)}_{\text{importation}} - \underbrace{d_{S^{ext}}ComS^{ext}(t)}_{\text{degradation}} - \underbrace{\mu(t)ComS^{ext}(t)}_{\text{growth}} \quad (3)$$

$$\frac{dComS^*(t)}{dt} = \underbrace{r_{in}ComS^{ext}(t)}_{\text{importation}} - \underbrace{c_{RS}(ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{d_{S^*}ComS^*(t)}_{\text{degradation}} - \underbrace{\mu(t)ComS^*(t)}_{\text{growth}} \quad (4)$$

$$\frac{dComR(t)}{dt} = \underbrace{b_R(t) + N(m, \sigma, t)}_{\text{distal regulators}} - \underbrace{c_{RS}(ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{d_RComR(t)}_{\text{degradation}} - \underbrace{\mu(t)ComR(t)}_{\text{growth}} \quad (5)$$

$$\begin{aligned} \frac{dComRS(t)}{dt} = & \underbrace{\frac{c_{RS}}{2}(ComR(t)ComS^*(t))^n}_{\text{ComRS formation}} - \underbrace{V_{max}^{RSZ} \left(\frac{(ComRS(t)ComZ(t))^{n_{RSZ}}}{K_m^{RSZ^{n_{RSZ}}} + (ComRS(t)ComZ(t))^{n_{RSZ}}} \right)}_{\text{shut-off by ComZ}} \\ & - \underbrace{d_{RS}ComRS(t)}_{\text{degradation}} - \underbrace{\mu(t)ComRS(t)}_{\text{growth}} \end{aligned} \quad (6)$$

$$\frac{dComX(t)}{dt} = b_X(t) + \underbrace{V_{max}^X \left(\frac{ComRS^{n_x}(t - \tau_2)}{K_m^{X^{n_x}} + ComRS^{n_x}(t - \tau_2)} \right)}_{\text{activation by ComRS}} - \underbrace{d_XComX(t)}_{\text{degradation}} - \underbrace{\mu(t)ComX(t)}_{\text{growth}} \quad (7)$$

$$\begin{aligned} \frac{dComZ(t)}{dt} = & b_Z(t) + \underbrace{V_{max}^Z \left(\frac{ComX^{n_z}(t - \tau_3)}{K_m^{Z^{n_z}} + ComX^{n_z}(t - \tau_3)} \right)}_{\text{activation by ComX}} - \underbrace{d_ZComZ(t)}_{\text{degradation}} - \underbrace{\mu(t)ComZ(t)}_{\text{growth}} \\ & - \underbrace{V_{max}^{RSZ} \left(\frac{(ComRS(t)ComZ(t))^{n_{RSZ}}}{K_m^{RSZ^{n_{RSZ}}} + (ComRS(t)ComZ(t))^{n_{RSZ}}} \right)}_{\text{inactivation of ComRS}} \end{aligned} \quad (8)$$

This first model is more general than our final model: the activation functions in Equations 1, 5, 6 and 7 take the form of **Hill functions** (the Michaelis-Menten functions of our final model are a particular case of Hill functions with $n = 1$).

Our initial model takes into account the time required for the transcription and translation steps. Indeed, if we look for instance at Eq1, the variation at time t of the amount in ComRS complexes does not immediately yield a variation in the amount of ComS, since the transcriptional activation takes some time for the binding of the ComRS complex on the promoter sequence of *comS*, for the transcription of *comS*, and finally for the translation of the resulting mRNA transcript. In mathematical form, it means that the activation functions of the newly produced proteins include **explicit time delays** denoted τ_1 , τ_2 and τ_3 , in order to represent transcription and translation delays.

The model calibration lead to the following parameter set, in which all exponents are equal to 1, except for n_Z .

r_{out}	Exportation rate of ComS	1	min^{-1}
r_{in}	Importation rate of ComS ^{ext}	1.10^{-2}	min^{-1}
n	Affinity between ComR and ComS	2.5	adim.
c_{RS}	Complex formation rate	10^{-1}	$\text{cell}^{n-1}.\text{mol}^{1-n}.\text{min}^{-1}$
b_Z	Basal synthesis rate of ComZ	0	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
v_{max}^Z	Maximal rate of induced synthesis of ComZ by ComX	1	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
n_Z	Hill coefficient	3	adim.
K_m^Z	Required concentration of ComX for half-maximum synthesis rate of ComZ	15	$\text{mol}.\text{cell}^{-1}$
v_{max}^{RSZ}	Maximal rate of inactivation of ComRS by ComZ	0.04	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
n_{RSZ}	Hill coefficient	1	adim.
K_m^{RSZ}	Required concentration of ComZ for half-maximum inactivation rate of ComRS	0.5	$\text{mol}.\text{cell}^{-1}$
d_s	Degradation rate of ComS	0	min^{-1}
$d_{S^{ext}}$	Degradation rate of ComS ^{ext}	0.07	min^{-1}
d_{S^*}	Degradation rate of ComS*	0.07	min^{-1}
d_R	Degradation rate of ComR	0.01	min^{-1}
d_{RS}	Degradation rate of ComRS	0.01	min^{-1}
d_X	Degradation rate of ComX	0.2	min^{-1}
d_Z	Degradation rate of Z	0.01	min^{-1}
τ_A	Time delay for ComZ activation	16	min
A	Amplitude of the gaussian activation of ComR	27	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
c	Center (time) of the gaussian activation of ComR	90	min
σ	Standard deviation of the gaussian activation of ComR	35	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
v_{max}^S	Maximal rate of activation of ComS by ComRS	4	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
n_S	Hill coefficient	1	adim.
K_m^S	Required concentration of ComRS for half-maximum activation rate of ComS	0.5	$\text{mol}.\text{cell}^{-1}$
v_{max}^X	Maximal rate of activation of ComX by ComRS	5.5	$\text{mol}.\text{min}^{-1}.\text{cell}^{-1}$
n_X	Hill coefficient	1	adim.
K_m^X	Required concentration of ComRS for half-maximum activation rate of ComX	1	$\text{mol}.\text{cell}^{-1}$

3. Finding the equilibria of the system

Let us re-write the system (1) as:

$$x' = f(x) = F(x) - Dx \quad (2)$$

where D is a notation for all diagonal terms, linked to dilution and growth, while $F(x)$ represents the kinetic terms. We solve this system in steady-state by setting all inputs constant:

$$F(x^*) - Dx^* = 0 \quad (3)$$

We will discard the first equation of the model for our analysis, as it only depends on an experimental measure of the cell density $X(t)$. We thus only keep the intracellular concentrations.

The state of the system can now be defined as:

$$x(t) = \{ X_1, X_2, X_3, X_4, X_5, \mathbf{X_6}, X_7 \}$$

$$= \{ComS(t), ComS^{ext}(t), ComS^*(t), ComR(t), ComRS(t), \mathbf{ComX(t)}, ComZ(t)\}$$

Numerical solving of (3) gives many solutions. We only keep those having a physical sense, i.e., lying in the non-negative orthant⁹. Here are the two steady-points:

$$x^{*1} = \{ 10,6 \quad 106 \quad 10,8 \quad 0,33 \quad 0,48 \quad \mathbf{7,9} \quad 3,2 \}$$

$$x^{*2} = \{ 0,05 \quad 0,53 \quad 0,05 \quad 3,3 \quad 0,001 \quad \mathbf{0,07} \quad 0 \}$$

The first solution x^{*1} corresponds to a steady-state in which the intracellular concentration in ComX is equal to 7,9 molecules per cell, while the second solution corresponds to a ComX concentration of 0,07 molecules per cell. Since the competence level of cells is directly dictated by the amount in the ComX alternative sigma-factor, we will say that the first equilibrium is the ON state, and the second equilibrium is the OFF state:

$$x^{*1} = x^{ON} : ComX(t)^* \uparrow$$

$$x^{*2} = x^{OFF} : ComX(t)^* \approx 0$$

4. Stability analysis of the steady-points

Since the frequency domain approach only applies to linearized systems, the first step is to linearize the system (2) around its equilibria. We obtain:

$$x(t)' = Mx(t-\tau) - Dx(t) \quad (4)$$

where M is the linearization of $F(x)$, i.e. the Jacobian matrix of the function F evaluated at the equilibrium x^* . By taking the Laplace transform and using the frequency notation, we can write:

$$[sI + D - M \text{diag}\{\exp(-s\tau_i)\}] x = 0$$

where s is the Laplace complex variable.

Theorem (Hale and Verduyne-Lunel, 1993)

The linear system (4) is exponentially stable if and only if the poles of the system have strictly negative real parts and are bounded away from zero.

⁹ The non-negative orthant in R^n is defined as the set of points $\{x_1, x_2, \dots, x_n\}$ such that $x_i \geq 0$ for $i = \{1, 2, \dots, n\}$.

To use this theorem, we have to find the poles of the system. We can find them by calculating the roots of the characteristic equation:

$$g(s, \tau) = \det[sI + D - M \text{diag}\{\exp(-s\tau_i)\}] = 0$$

It can be shown that the function $g(s, \tau)$ can be written in the following form:

$$g(s, \tau) = \Psi(s) - \xi(s)\exp(-s\tau) \quad (5)$$

where the times delays τ_i are aggregated in the term τ :

$$\tau = \sum_i \tau_i$$

and where $\Psi(s)$ and $\xi(s)$ are polynomials in s . This expression of $g(s, \tau)$ allows to use the procedure of Walton and Marshall (Walton and Marshall, 1987).

1) $\tau = 0$

We compute the roots of $g(s, 0)$ by numerical methods:

$$x_g^{\text{ON}} = \{ -1,02 \quad -0,5 \quad -0,26+0,12i \quad -0,26-0,12i \quad -0,06+0,11i \quad -0,06-0,11i \quad \underline{\underline{0,05}} \}$$

$$x_g^{\text{OFF}} = \{ -1,01 \quad -0,3 \quad -0,22 \quad -0,19 \quad -0,017+0,08i \quad -0,0017-0,08i \quad -0,03 \}$$

Since we have one non-negative root for $g(s, 0)$ evaluated at x^{ON} , the ON equilibrium is unstable for all values of τ . The OFF equilibrium is stable when $\tau=0$.

2) $\tau \neq 0$

Then, to study the stability of the OFF state when $\tau \neq 0$, we compute the second polynomial:

$$W(\omega^2) = \Psi(j\omega) \Psi(-j\omega) - \xi(j\omega) \xi(-j\omega)$$

and calculate its roots:

$$x_\psi^{\text{OFF}} = \{ -7,4 \quad -2,28+7,04i \quad -2,28-7,04i \quad \underline{\underline{5,99+4,35i}} \quad \underline{\underline{5,99-4,35i}} \quad \underline{\underline{0}} \quad \underline{\underline{0}} \}$$

Since $W(\omega^2)$ has roots with positive real parts, we know that there is a maximum value of τ beyond which the OFF steady-state becomes unstable.

5. Biological relevance of the analysis

Let us analyze these results from a biological point of view.

First, concerning the two equilibria of the system, we see that the number of ComS molecules (produced, exported and reinternalized) is much higher in the ON state than in the OFF state, with an external concentration of ~100 molecules per cell. The ComR concentration is however lower in the ON cell, since ComR molecules are recruited to form ComRS complexes, which are more abundant in that steady-point. The ComZ concentration is null in the OFF state, which is coherent with the fact that the late phase of competence is not activated in that equilibrium.

Second, regarding the stability of the steady-states, the fact that the ON state is unstable while the OFF state is stable (under certain conditions on the time-delays) suggests that the regulatory system controlling competence is not bistable in *Sth*, in opposite to *Spn* (Karlsson *et al.*, 2007).

Finally, the fact that the stability of the system equilibria depends on the values of the time-delays shows that the regulatory function of the system relies on a finely tuned network layout which requires adequate time-constants.

Altogether, these results shed light on the system behavior around the equilibrium points: the majority of the cells remain in the stable OFF state, while only a minority of the population can transiently enter the unstable ON state under competence-permissive conditions, depending on stochastic inter-cellular fluctuations.

Part IV

Discussion & perspectives

Natural transformation is one of the **strongest driving forces** of bacterial evolution (Johnsborg *et al.*, 2007; Skippington and Ragan, 2011). It confers to competent bacteria the ability to acquire advantageous traits in one shot, while punctual mutations only allow a very slow genetic evolution (creating a new function from scratch takes a while). The benefits of one transformation event thus represent an extremely long period of conventional evolutionary time. Moreover, pathogenic (*i.e.*, disease causing) species (like *S. pneumoniae*) thrive on this lateral gene transfer mechanism, constantly evolving by changing their extracellular recognition patterns in order to hide and escape from the host immune system (Carrolo *et al.*, 2009; Zaccaria *et al.*, 2014). Competence for natural transformation is a major evolutionary advantage, present in a long series of bacterial species, hence strongly impacting the prokaryotic fate (Skippington and Ragan, 2011). Besides this powerful genetic plasticity, competence could be beneficial by providing an alternative source of energy (Johnsborg *et al.*, 2007), as it was demonstrated that *E. coli* could grow with DNA as sole carbon source (Finkel and Kolter, 2001). Moreover, integration of exogenous DNA can also help the recipient bacterium to repair its genome when needed (Johnsborg *et al.*, 2007). As a counterpart, the cost of competence development is not negligible for the cells: it slows down the cellular growth, represents an important energetic burden, temporarily stops DNA replication, and yields hazardous genetic modifications.

This is why natural transformation is **tightly regulated** in almost all competent species, and hence restricted at different levels: i) it only appears during a short time window, ii) in precise growth conditions, and iii) in only a fraction of the population (bistability) in most species. In streptococci, competence for natural transformation is regulated by the ComCDE and ComRS systems, depending on the streptococcal group (Fontaine *et al.*, 2015). The latter system was recently discovered in our model species *S. thermophilus* (Fontaine *et al.*, 2010a; Fontaine *et al.*, 2013). As a dairy starter and probiotic, this bacterial species has a huge **industrial** and **commercial interest**. It is used in cocultures with *L. delbrueckii* subsp. *bulgaricus* for yogurt and various cheeses production, and represents the second most important dairy starter after *L. lactis*, totalizing an annual market of \$40 billion (Bolotin *et al.*, 2004).

The **aim of this thesis** was to **analytically examine** competence regulation in *Sth*, since the recent discovery of the ComRS system had not been thoroughly analyzed yet. The main investigation points were briefly the following: i) understanding the dynamic behavior of the transient activation of the ComRS system, ii) confirming its topological organization, iii) determining its critical components, iv) finding how the distal regulations are integrated by the system, and v) shedding light on its shut-off mechanism. To do this, an original approach combining classic **molecular genetic** tools and **mathematical modeling** was used. Indeed, a detailed, mechanistic, data-based mathematical model of the ComRS system, including all known molecular steps of the regulatory cascade, was missing in order to quantitatively analyze this particular genetic network.

The **challenge** thus consisted in integrating all our *a priori* knowledge and mechanistic hypotheses about the system, in a very detailed and transparent mathematical model, which had to reproduce all the observed behaviors concerning competence, and then to be able to **predict** new features by testing conditions and scenarios that were not yet tested at the experimental level. Moreover, the model had to be easily modified and adapted to new observations and hypotheses, and to be “expendable” for the introduction of new components, like for instance the shut-off plug-in. Our model was based on experimental data that were already available, reproduced or generated during the course of this thesis. These data were collected from luciferase reporter strains in various mutant backgrounds, following a **systematic perturbation** approach, which is an efficient way to **infer knowledge** about a system organization.

1. Modeling choices

We focused our model on the **timing-device** functionality of the ComRS system in the model organism *Sth*. Especially, the LMD-9 strain was chosen as initial model strain, since it is naturally one of the most competent strains of *Sth* and since its competence level can still be further activated upon ComS supplementation (Fontaine *et al.*, 2010a).

We finally chose the formalism of **Ordinary Differential Equations** (ODE) in order to obtain a transparent, easily tunable model. Other types of models exist, like the Boolean Networks, the Dynamic Bayesian Networks or the Artificial Neural Networks, for instance. But all of these have the drawback of non-transparency: they are black-box models, lacking a clear correspondence between the terms of their equations and the physical components and mechanisms of the real system. Moreover, it would have been more complicated to reproduce interesting behaviors with such models that remain quite basic and are mainly discrete-time. So, the user-friendly, time-continuous setting of ODE seemed much more suitable for our purpose. Its transparency and ease of use allowed us to translate into equations the fact that, for instance, the ComR regulator does not make part of the positive auto-regulatory loop itself, in contrary to the ComE regulator of the ComCDE system.

The level of detail of a biological model is always a matter of compromise between its informativeness (high informativeness requires high complexity, leading to a risk of overfitting) and scarcity (less noise but less informativeness). Here, we have decided to build a model which had to be able to reproduce the main experimental features, and to integrate the main mechanistic hypotheses we have about its functioning. Adopting a lower degree of detail would have resulted in a less interesting output, probably not informative enough to be semi-quantitative as our model is, and certainly not detailed enough to output subtle behaviors like damped oscillatory profiles. On the other hand, adopting a higher level of detail would have implied an hazardous learning step, since our parameter set was already quite large compared

to the data we had at disposal to calibrate it. So, a more complex model would have required additional data to feed it.

Besides our “mechanistic” working hypotheses, which aim to integrate all our *a priori* knowledge about the system into the equations, we made some commonly accepted “meta-hypotheses”. These **meta-hypotheses** are **high-level choices** that have to be done before beginning to write the model’s equations as well as the formalism choice. We have made three major meta-hypotheses.

- i) First, we assumed that the cell population is homogenous, meaning that in our model, all individuals behave exactly the same way. We know it is not the case in reality, since inter-individuals variations exist, leading sometimes to a bistability phenomenon. The two main reasons for this choice are: 1) our experimental data mainly consisted in luciferase profiles, which are averaged light emissions over the population, and 2) modeling the inter-individuals variations would have required more advanced mathematical tools, with stochasticity simulators, which was not our prime purpose.
- ii) Second, we decided to focus our model output on the proteic level only, hence neglecting the mRNA fluctuations. This assumption is sometimes referred to as the “quasi-steady state” for mRNA (Sengupta and Kar, 2015). It implies that the time-course variations of mRNA levels are assumed to be null, like if mRNA were constant over time. This assumption is actually made because the mRNA levels vary according to a very small time-scale compared to the proteic time-scale. Hence, these very rapid variations are considered as noise, and neglected. Taking the mRNA layer into account would require the model to include two different layers (the proteic and the mRNA layers) with totally different time-scales, which can be very difficult from a technical point of view.
- iii) Third, our final model uses the assumption that the transcription and translation steps are instantaneous: the variation at time T of an activator concentration directly increases the concentration of the activated component, instantaneously at the same time T . This is obviously not the case in nature, since transcription and translation do take some time. We tried to include delays into our model in order to take these durations into account. But the resulting Delay Differential Equations (DDE) model was not better than our initial ODE model. So, for scarcity concerns, we decided to stick with ODE, since the delays were increasing the model’s complexity (DDE models are actually of infinite order) without improving its quality nor predictiveness.

We used an original “**forth-and-back**” modeling approach for parameter estimation. This two-step approach is innovative and has allowed us to reduce the degrees of freedom of our initial system. This particular and complex learning step can be

considered as a result by itself. The model was then **thoroughly validated** at different levels (dynamics, reactivity to ComS, and topology of the network), giving us a fair confidence in its quality and robustness.

2. The critical role of ComR on competence level

Our modeling approach revealed that the concentration of the σ^X alternative sigma-factor, and hence the competence level, was largely determined by the concentration of ComR in Sth LMD-9. We already knew that the ComS peptide was limiting, since the competence level could be further increased upon ComS addition, but we did not know that ComR was even more critical for competence development. As already discussed in the Results section, the ComRS system has the particularity that the regulator, ComR, is not included itself inside the positive feedback loop. This topological feature allows ComR to behave like the “master regulator” of the early phase of competence, integrating a series of external and internal signals to fine-tune competence regulation, in order to take the decision to develop this important but costly transient program according to various environmental cues.

The apparent competence level discrepancies between strains can thus be explained by differences in *comR* expression. These can stem from variation in the promoter sequences of *comR* or distal regulations, for instance. So, It suggested us to overexpress *comR* in the low competent strain LMG18311, confirming the initial hypothesis of ComR limitation.

This main result has an important experimental signification. From a biological point of view, it pinpoints a crucial determinant for the competence level, located upstream of ComX. It opens the door to a series of interesting experiments: one could try to identify the differences in the sequence of different promoters of *comR* in order to explain the differences in expression levels. It would also be interesting to identify distal regulators acting on *comR* expression, and to localize their binding sites on the promoter of *comR* by a mutational approach.

Concerning the development of new applications, this result also represents a major step since it suggests a new way to increase the competence level in poorly competent strains of *S. thermophilus*, but also probably in other phylogenetically close streptococci. Converging results had been obtained with *S. mutans* (Mashburn-Warren *et al.*, 2010). So, this finding has many implications both from biological and industrial points of view.

This result is a remarkable illustration of the interest of building mathematical models. In our case, the model *a priori* predicted the critical role of ComR, which was then a *posteriori* confirmed by experimental data. This clearly illustrates the utility to have a robust, predictive model at hand in order to analyze (*i.e.*, to quantify and understand

observed data) and to investigate (to discover new features *in silico* and to orient experimental work accordingly) a complex system.

3. Oscillations in competence development

The *in silico* simulations had predicted an oscillating development of competence in the simulated *comR* overexpression strain, which was experimentally observed afterwards. Of course, our model had not been calibrated on such oscillating data, since they were only observed after the model's prediction.

These oscillations are an intrinsic characteristic of the system, which presents the technical requirements for them (*i.e.*, interconnected positive and negative feedback loops with appropriate time scales). However, the model only predicted them for a very narrow range of parameter values. Indeed, only a very small sub-space of the parameter space was prompting them. This sensitivity of the oscillatory behavior in regard to the system parameters suggests that the oscillating development of competence only takes place under particular conditions, when the bacterial population has an interest for implementing this special program.

Since the *in silico* scenario giving rise to oscillations was for a simulated strain overexpressing *comR*, it suggests that oscillating competence could be a response triggered when the conditions are particularly favorable for natural transformation, implementing a compromise between a longer competence window (in order for the population to maximize the benefits from the transformation) and a minimal cellular growth under conditions that are normally non permissive for division. The biological purpose of oscillations in competence development could thus be to conciliate two processes that are generally seen as antagonistic: on one hand, competence development normally redirects all the transcription program of the cells towards its specific competence machineries during the short duration of natural transformation; and on the other hand, cellular growth implies replication of the genome and cell division, that are hardly compatible with the genetic rearrangements that are entailed with natural transformation (Gardan *et al.*, 2013; Haijema *et al.*, 2001; Mirouze *et al.*, 2015). So, oscillating between the ON state in which cells can actually undergo transformation but not growth, and the OFF state in which transformation is transiently stopped while growth is resumed, could provide competent strains with a new physiological state that allow them to take advantages from both the exceptionally favorable conditions for natural transformation, while not totally bypassing growth.

While environmental conditions permissive for oscillations in competence development seem to be very restricted both in computational and laboratory settings, we cannot exclude the possibility that natural conditions like those encountered in (sometimes very complex) biofilms could be sensed as extremely favorable for competence development. Both the DNA (and nutrients) proficiency and

the relatively secured habitat could contribute to trigger a particular behavior which would allow the privileged population to maximize the advantages it can take from the “heavenly” environment.

Experimentally, it would be interesting, in order to confirm this hypothesis of a compromise between competence and growth, to monitor the cellular growth under conditions permissive for an oscillating development of competence. This would allow us to verify whether division is actually blocked and resumed at each oscillatory phase.

Besides, future work could investigate the heterogeneity of the oscillatory behavior inside the population: at this point, we can not definitely exclude the possibility that several subpopulations develop competence at different time points, resulting in an overall impression of oscillations. To test this hypothesis, a microfluidic approach could be used with an appropriate reporter system: a sufficient time-wise resolution is required in order to detect the waves of competence. Hence, the reporter molecule should have a high turn-over (*i.e.*, a short half-life). Some GFP have been intentionally developed for this purpose (Houser *et al.*, 2012). Alternatively, one could take advantage of the photobleaching phenomenon.

However, the high reproducibility of the experimental occurrences of oscillations, as well as the striking correspondence between the *in silico* levels of *comR* overexpression and the time-frequencies of the waves strongly favors the hypothesis that the whole population undergoes successive waves of transformation. In order to strengthen this hypothesis, a simple experimental way could be to study the effect of varying the *comR* expression level (by using an inducible promoter with a range of inducer concentrations, for instance) on the shape and number of oscillations.

Besides the very informative point of this *in silico* result, the prediction of these oscillations can be seen as an additional, totally independent, validation level of our model. Indeed, since oscillations had never been observed experimentally, their simulation was not expected at all, but finally proved to be valuable.

4. Competence shut-off

Concerning the shut-off of competence, our current model told us that it was likely to take place at the level of the ComRS complex, as it was the case with the model of Karlsson *et al.* (Karlsson *et al.*, 2007), but was not able to discriminate the exact level at which it is taking place. Indeed, we do not know whether the putative repressor(s) inhibits or inactivate the ComRS complex, or if its action is towards ComS or ComR, or if it represses the expression of *comS* or *comR*, or if it impedes the complex formation, and so on... So, it would be interesting to improve the shut-off module efficiency by implementing all the possible scenarios for the shut-off (and their combinations) with a sufficient level of detail (hence requiring to have additional

quantitative experimental data concerning the shut-down), in order to infer new information about the exit from competence.

The attractive candidate DprA (Berge *et al.*, 2013; Mirouze *et al.*, 2013; Mortier-Barriere *et al.*, 2007) was tested in mutant strains deleted for *dprA*, but no major effect on the length and strength of the early phase was detected, ruling out this tempting hypothesis. The shut-off mechanism of competence in Sth thus remains to be explored.

In addition to the *in silico* approach with the improvement of the shut-off module, an experimental approach could be very helpful, for instance consisting in a systematic deletion and/or overexpression of all the late operons with the monitoring of the dynamics of exit of competence via luciferase reporters.

5. Bistability in competence development

The observation of LMD-9 fluorescent reporter cells for P_{comS} activity by epifluorescence microscopy suggested that it could be possible that only a fraction of the population actually activates competence (A. Wahl, unpublished results), which, if proved true, would indicate a bistable behavior of competence activation in Sth.

Since our model is a homogeneous average of a whole cell population, it could not directly predict this bistability. However, knowing this experimental fact, we could *a posteriori* use the model to better understand this phenomenon. For this purpose, we have proposed a general methodology in our paper (Haustenne *et al.*, 2014), which could help to analyze the (local and global) **stability** of the **steady points** of the system, based on two complementary approaches: a frequency domain approach, based on the linearization of the system (Bastin and Coron, 2016) and a Lyapunov approach, which uses an auxiliary function, exponentially decreasing along its solutions (Bastin and Coron, 2016). We have applied this stability analysis which allowed us to mathematically investigate the properties of the steady states (see chapter three in the Results section) of our *in silico* model of competence regulation in Sth. It revealed that the system OFF state, corresponding to the vegetative growth (low level of ComX), is stable, while the ON state, corresponding to the competence window (high levels of ComX), is unstable. This analysis suggests that competence in *S. thermophilus* is not a bistable phenomenon. However, it is worth recalling that this stability analysis was NOT performed on our final and best-fitting ComRS model and that the numerical values of the modelling parameters were not the same between this model and the initial model used. It would thus be very interesting to re-analyze the stability of the equilibria of the final ComRS model in order to draw a final conclusion to this part of our work.

Moreover, it would be interesting to see if the system can exhibit a hysteretic behavior (*i.e.*, the system behavior does not only depend on the current state of the

variable, but also on their past history, resulting in a time-gap between one cause and its effects), like it was the case for the ComCDE model of Karlsson *et al.* (Karlsson *et al.*, 2007).

From an experimental point of view, future work could help testing the bistability hypothesis by observing reporter strains for late competence promoters (like P_{comGA}), and to test other strains than LMD-9 (including LMG18311).

In order to trigger the potential bistable behavior, one might try to disturb key-components and parameters of the system such as, for instance, the expression level of *comR* (overexpression), the reimportation rate of ComS, the suppression of the negative feedback loop (by deleting *comX* for instance), the deletion of distal regulators, or the strength of the positive feedback loop (by modifying the ComR-ComS affinity for instance).

Moreover, it could be very interesting to perform time-lapse experiments with adequate reporter strains to test whether the total population is indeed divided into a ON and a OFF subpopulations, or whether different subpopulations undergo successive waves of transformation.

6. Model limitations and future improvements

Our model presents some limitations, linked to the various simplifications and meta-hypotheses that had to be done. All these could be alleviated in future works.

First, the social aspect of the cell-cell communication could be further investigated by taking into account both the “self” and “others-communication” (Youk and Lim, 2014). It would be interesting to test the effect of medium variations (concentration of free peptides, diffusion of the mature ComS peptide) on the inter-cellular communication, and to see how these environmental parameters can influence the social behaviors of the population.

Besides, the reimportation step of our model does not take into account the fact that the speed of reimportation is proportional to the cellular density. Indeed, our model takes the peptide reimportation into account (see **Eq.4**) through a term which is linear in the variable $ComS^{ext}$, but which does not integrate the cellular density. It would then be interesting to see which impact it would have to include the variable $X(t)$ in that reimportation term, in order to take the QS functionality into account.

Another interesting point to test in future developments would be to adapt the model to other growth media, like rich media for instance (e.g. M17 broth) where competence development is less permissive. Also linked to the bacterial growth, the model could be slightly modified in order to take into account the negative effect of competence development on the cellular growth rate. For this purpose, the time-varying growth rate ($\mu(t)$) could be multiplied by a supplementary term accounting for

the level of developed competence, which should increase when ComX concentration increases.

A more important change to the model could be to model more precisely the action of the different distal regulators acting on the ComRS module. For now, all the positive and negative influences are summed up in one gaussian term. But with enough knowledge and new experimental data, the different signals could be decorrelated and modeled separately in a more informative model. Maybe another formalism, possibly discrete, could be advantageous for future work focusing on integrating all the distal regulations acting on the ComRS system (see **Figure 25**).

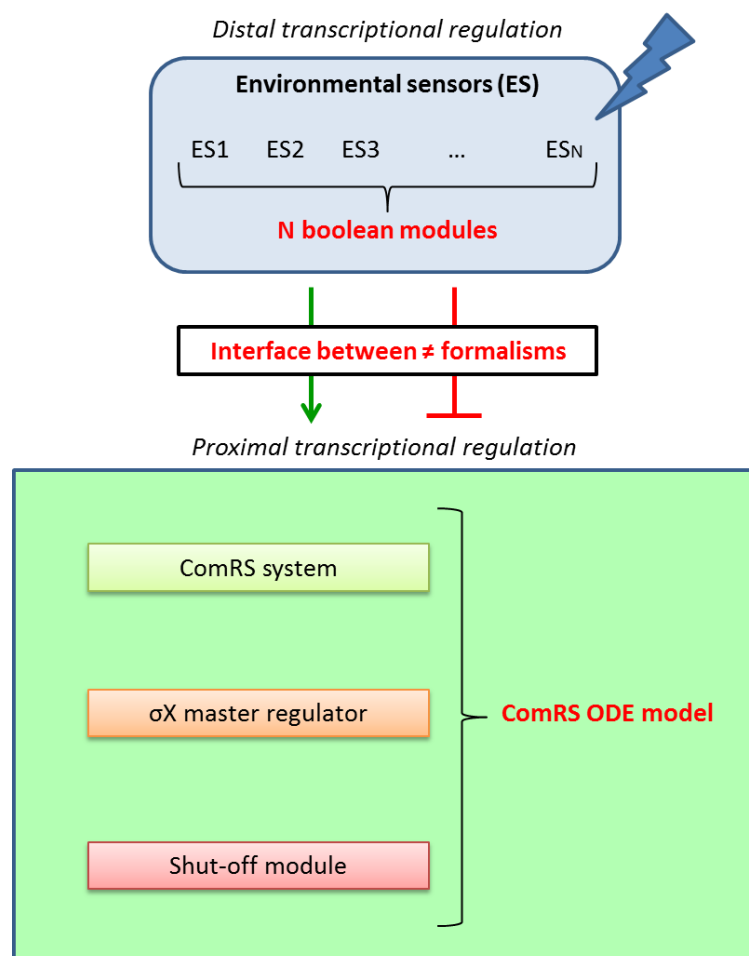


Figure 25. Boolean – ODE hybrid model for the ComRS system. The distal transcriptional regulators (environmental sensors) are here embedded in boolean modules interacting positively or negatively with the ComRS ODE model. The connection between the two types of formalisms is called the “interface”, representing a challenging computational issue.

Possibly, a “hybrid” model (for an example, see (Chang and Tomlin, 2012)) could be imagined, consisting in a core ComRS-module described with the ODE formalism, and additional distal modules, written in other simpler formalisms, such as Boolean Networks or Regression models. With such a composite model, the technical difficulty would be the interfaces between the different modules.

Moreover, it would be very interesting to re-learn the model with experimental data of protein concentrations inside the cells. This would really be more accurate than the current approximations based on luciferase profiles, and would allow the model to be predictive from a quantitative point of view, hence allowing to infer unknown intracellular concentrations.

Another interesting development could be to adapt the model to other ComRS containing streptococci. For instance, an extended model, able to simulate the intertwined functionings of both the ComRS and BlpCHR systems of *S. mutans*, could be very informative about these two interconnected signaling systems.

Finally, the heterogeneity of competence development between the different subpopulations deserves further developments. The model could be extended in order to integrate the stochastic effects and the noise that are necessary for the bistability phenomenon (Zimmer and Sahle, 2016; Gupta *et al.*, 2014). In order to keep a differential formalism, one could try to use the Gillespie or Monte-Carlo algorithms¹¹ on the most sensitive parameters of the model in order to simulate the inter-individuals variations (Albert, 2016; Huang *et al.*, 2013). In the light of our results, it could probably be done when varying the ComR-related parameters (*i.e.*, all parameters appearing in the equation describing ComR evolution).

When the critical parameters for bistability will be identified, it would be interesting to study the stability of the steady points of the final model with the methodology we proposed in our MATHMOD article, and to perform a bifurcation analysis in order to see how the variation of each parameter value affects the distribution of cells into the ON and OFF states.

The combination of mathematical analyses (stability of steady points and bifurcation diagrams) with time-lapse experiments could be helpful to better understand the mechanisms of subpopulations differentiation and its advantages. A meta-model of

¹¹ Monte-Carlo methods are a kind of computational algorithms that perform repeated random sampling to solve problems. It can be used to solve deterministic problems, when the number of degrees of freedom gives rise to too much uncertainty. It is often used to simulate physical interactions in cellular structures. Gillespie algorithms involve a Monte-Carlo step: after the initiation phase (in which the state of the system is determined according to initial conditions), the Gillespie algorithm generates random numbers in order to "toss out" which action will happen next, and when. The probability for a reaction to be chosen is generally proportional to the number of substrates of this reaction present in the system. The algorithm then updates the state of the system, and iterates.

the different subpopulations could provide some information about the global fitness advantage conferred to the whole population by such a bet-hedging strategy.

7. Conclusion

To conclude, during this thesis, we have designed and executed an original approach combining molecular genetics with mathematical modeling. The result is a detailed, well validated model, which has allowed us to better understand competence regulation in *Sth*. Indeed, except the simplest model of Son *et al.*, such a model was missing in order to objectively and analytically investigate the complex ComRS regulatory network (Son *et al.*, 2012).

Biological experimentations and mathematical modeling steps were feeding each other continuously during this thesis, one field giving clues about how to go ahead in the other field.

This work brings a significant contribution to the current knowledge of the ComRS system, by shedding light on the critical role of ComR, by explaining the inter-strain variability in the competence level, by suggesting a simple and efficient way to increase the competence rate in poorly transformable strains, and by identifying the technical requirements about the system topology and parameters for the observed behaviors.

Moreover, this work gives an original and general methodology for the learning, validation, use, and analysis of genetic regulatory networks. It also opens new doors for future work and developments, especially concerning the particular characteristics of the network, like its bimodality and oscillations. The shut-off enigma was not broken, and still remains to be solved.

Reference List

- Aas,F.E., Wolfgang,M., Frye,S., Dunham,S., Lovold,C., and Koomey,M. (2002) Competence for natural transformation in *Neisseria gonorrhoeae*: components of DNA binding and uptake linked to type IV pilus expression. *Mol Microbiol* **46**: 749-760.
- Abiodun,G.J., Maharaj,R., Witbooi,P., and Okosun,K.O. (2016) Modelling the influence of temperature and rainfall on the population dynamics of *Anopheles arabiensis*. *Malar J* **15**: 364.
- Acar,M., Mettetal,J.T., and van,O.A. (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nat Genet* **40**: 471-475.
- Ahn,S.J., Wen,Z.T., and Burne,R.A. (2006) Multilevel control of competence development and stress tolerance in *Streptococcus mutans* UA159. *Infect Immun* **74**: 1631-1642.
- Albert,J. (2016) A Hybrid of the Chemical Master Equation and the Gillespie Algorithm for Efficient Stochastic Simulations of Sub-Networks. *PLoS One* **11**: e0149909.
- Albert,R., and Thakar,J. (2014) Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. *Wiley Interdiscip Rev Syst Biol Med* **6**: 353-369.
- Alloing,G., Martin,B., Granadel,C., and Claverys,J.P. (1998) Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. *Mol Microbiol* **29**: 75-83.
- Angeli,D., Ferrell,J.E., Jr., and Sontag,E.D. (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc Natl Acad Sci U S A* **101**: 1822-1827.
- Arnone,M.I., and Davidson,E.H. (1997) The hardwiring of development: organization and function of genomic regulatory systems. *Development* **124**: 1851-1864.
- Aspiras,M.B., Ellen,R.P., and Cvitkovitch,D.G. (2004) ComX activity of *Streptococcus mutans* growing in biofilms. *FEMS Microbiol Lett* **238**: 167-174.
- Assalkhou,R., Balasingham,S., Collins,R.F., Frye,S.A., Davidsen,T., Benam,A.V. *et al.* (2007) The outer membrane secretin PilQ from *Neisseria meningitidis* binds DNA. *Microbiology* **153**: 1593-1603.
- Atlas,J.C., Nikolaev,E.V., Browning,S.T., and Shuler,M.L. (2008) Incorporating genome-wide DNA sequence information into a dynamic whole-cell model of *Escherichia coli*: application to DNA replication. *IEET Syst Biol* **2**: 369-382.
- Attaiech,L., Olivier,A., Mortier-Barriere,I., Soulet,A.L., Granadel,C., Martin,B. *et al.* (2011) Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: maintenance of a reservoir for genetic plasticity. *PLoS Genet* **7**: e1002156.
- Auchtung,J.M., Lee,C.A., and Grossman,A.D. (2006) Modulation of the ComA-dependent quorum response in *Bacillus subtilis* by multiple Rap proteins and Phr peptides. *J Bacteriol* **188**: 5273-5285.
- Avery,O.T., Macleod,C.M., and McCarty,M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types : induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J Exp Med* **79**: 137-158.
- Balaban,M., Battig,P., Muschiol,S., Tirier,S.M., Wartha,F., Normark,S., and Henriques-Normark,B. (2014) Secretion of a pneumococcal type II secretion system pilus correlates with DNA uptake during transformation. *Proc Natl Acad Sci U S A* **111**: E758-E765.
- Bansal,M., Belcastro,V., Ambesi-Impiombato,A., and di,B.D. (2007) How to infer gene networks from expression profiles. *Mol Syst Biol* **3**: 78.

- Banu,L.D., Conrads,G., Rehrauer,H., Hussain,H., Allan,E., and van der Ploeg,J.R. (2010) The *Streptococcus mutans* serine/threonine kinase, PknB, regulates competence development, bacteriocin production, and cell wall metabolism. *Infect Immun* **78**: 2209-2220.
- Barrangou,R., Fremaux,C., Deveau,H., Richards,M., Boyaval,P., Moineau,S. *et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709-1712.
- Barrangou,R., and Horvath,P. (2012) CRISPR: new horizons in phage resistance and strain identification. *Annu Rev Food Sci Technol* **3**: 143-162.
- Bastin,G., and Coron,J.M. (2016) Stability and Boundary Stabilization of 1-D Hyperbolic Systems.
- Bensing,B.A., Rubens,C.E., and Sullam,P.M. (2001) Genetic loci of *Streptococcus mitis* that mediate binding to human platelets. *Infect Immun* **69**: 1373-1380.
- Berg,K.H., Biørnstad,T.J., Johnsborg,O., and Havarstein,L.S. (2012) Properties and biological role of streptococcal fratricins. *Appl Environ Microbiol* **78**: 3515-3522.
- Berge,M., Mortier-Barriere,I., Martin,B., and Claverys,J.P. (2003) Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol Microbiol* **50**: 527-536.
- Berge,M.J., Kamgoue,A., Martin,B., Polard,P., Campo,N., and Claverys,J.P. (2013) Midcell recruitment of the DNA uptake and virulence nuclease, EndA, for pneumococcal transformation. *PLoS Pathog* **9**: e1003596.
- Berka,R.M., Hahn,J., Albano,M., Draskovic,I., Persuh,M., Cui,X. *et al.* (2002) Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**: 1331-1345.
- Bertani,G., and Baresi,L. (1987) Genetic transformation in the methanogen *Methanococcus voltae* PS. *J Bacteriol* **169**: 2730-2738.
- Bhatt,A., Kieser,H.M., Melton,R.E., and Kieser,T. (2002) Plasmid transfer from *Streptomyces* to *Mycobacterium smegmatis* by spontaneous transformation. *Mol Microbiol* **43**: 135-146.
- Bikard,D., Hatoum-Aslan,A., Mucida,D., and Marraffini,L.A. (2012) CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12**: 177-186.
- Biørnstad,T.J., and Havarstein,L.S. (2011) ClpC acts as a negative regulator of competence in *Streptococcus thermophilus*. *Microbiology* **157**: 1676-1684.
- Bischofs,I.B., Hug,J.A., Liu,A.W., Wolf,D.M., and Arkin,A.P. (2009) Complexity in bacterial cell-cell communication: quorum signal integration and subpopulation signaling in the *Bacillus subtilis* phosphorelay. *Proc Natl Acad Sci U S A* **106**: 6459-6464.
- Bolotin,A., Quinquis,B., Renault,P., Sorokin,A., Ehrlich,S.D., Kulakauskas,S. *et al.* (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* **22**: 1554-1558.
- Boutry,C. Regulation of competence for natural transformation in *Streptococcus thermophilus*. 1-5-2013.
- Boutry,C., Delplace,B., Clippe,A., Fontaine,L., and Hols,P. (2013) SOS response activation and competence development are antagonistic mechanisms in *Streptococcus thermophilus*. *J Bacteriol* **195**: 696-707.
- Boutry,C., Wahl,A., Delplace,B., Clippe,A., Fontaine,L., and Hols,P. (2012) Adaptor protein MecA is a negative regulator of the expression of late competence genes in *Streptococcus thermophilus*. *J Bacteriol* **194**: 1777-1788.
- Browning,S.T., Castellanos,M., and Shuler,M.L. (2004) Robust control of initiation of prokaryotic chromosome replication: essential considerations for a minimal cell. *Biotechnol Bioeng* **88**: 575-584.
- Burton,B., and Dubnau,D. (2010) Membrane-associated DNA transport machines. *Cold Spring Harb Perspect Biol* **2**: a000406.
- Calvio,C., Celandroni,F., Ghelardi,E., Amati,G., Salvetti,S., Cecilian, F. *et al.* (2005) Swarming differentiation and swimming motility in *Bacillus subtilis* are controlled by *swrA*, a newly identified dicistronic operon. *J Bacteriol* **187**: 5356-5366.

- Carley,N.H. (1992) Streptococcus salivarius bacteremia and meningitis following upper gastrointestinal endoscopy and cauterization for gastric bleeding. *Clin Infect Dis* **14**: 947-948.
- Carrolo,M., Pinto,F.R., Melo-Cristino,J., and Ramirez,M. (2009) Pherotypes are driving genetic differentiation within Streptococcus pneumoniae. *BMC Microbiol* **9**: 191.
- Carrolo,M., Pinto,F.R., Melo-Cristino,J., and Ramirez,M. (2014) Pherotype influences biofilm growth and recombination in Streptococcus pneumoniae. *PLoS One* **9**: e92138.
- Cassone,M., Gagne,A.L., Spruce,L.A., Seeholzer,S.H., and Sebert,M.E. (2012) The HtrA protease from Streptococcus pneumoniae digests both denatured proteins and the competence-stimulating peptide. *J Biol Chem* **287**: 38449-38459.
- Castellanos,M., Kushiro,K., Lai,S.K., and Shuler,M.L. (2007) A genomically/chemically complete module for synthesis of lipid membrane in a minimal cell. *Biotechnol Bioeng* **97**: 397-409.
- Castellanos,M., Wilson,D.B., and Shuler,M.L. (2004) A modular minimal cell model: purine and pyrimidine transport and metabolism. *Proc Natl Acad Sci U S A* **101**: 6681-6686.
- Caymaris,S., Bootsma,H.J., Martin,B., Hermans,P.W., Prudhomme,M., and Claverys,J.P. (2010) The global nutritional regulator CodY is an essential protein in the human pathogen Streptococcus pneumoniae. *Mol Microbiol* **78**: 344-360.
- Chang,Y.H., and Tomlin,C. (2012) Modeling Response of Biological Signal Pathways Using a Hybrid Boolean Framework. *Biochemical systems*.
- Chastanet,A., Prudhomme,M., Claverys,J.P., and Msadek,T. (2001) Regulation of Streptococcus pneumoniae clp genes and their role in competence development and stress survival. *J Bacteriol* **183**: 7295-7307.
- Chen,I., Christie,P.J., and Dubnau,D. (2005) The ins and outs of DNA transfer in bacteria. *Science* **310**: 1456-1460.
- Chen,I., and Dubnau,D. (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol* **2**: 241-249.
- Chen,I., Provvedi,R., and Dubnau,D. (2006) A macromolecular complex formed by a pilin-like protein in competent Bacillus subtilis. *J Biol Chem* **281**: 21720-21727.
- Chen,J.D., and Morrison,D.A. (1987) Modulation of competence for genetic transformation in Streptococcus pneumoniae. *J Gen Microbiol* **133**: 1959-1967.
- Chen,T., He,H.L., and Church,G.M. (1999) Modeling gene expression with differential equations. *Pac Symp Biocomput*: 29-40.
- Chitnis,N., Smith,T., and Steketee,R. (2008) A mathematical model for the dynamics of malaria in mosquitoes feeding on a heterogeneous host population. *J Biol Dyn* **2**: 259-285.
- Chubukov,V., Zuleta,I.A., and Li,H. (2012) Regulatory architecture determines optimal regulation of gene expression in metabolic pathways. *Proc Natl Acad Sci U S A* **109**: 5127-5132.
- Claverys,J.P., Grossiord,B., and Alloing,G. (2000) Is the Ami-AliA/B oligopeptide permease of Streptococcus pneumoniae involved in sensing environmental conditions? *Res Microbiol* **151**: 457-463.
- Claverys,J.P., and Havarstein,L.S. (2002) Extracellular-peptide control of competence for genetic transformation in Streptococcus pneumoniae. *Front Biosci* **7**: d1798-d1814.
- Claverys,J.P., and Havarstein,L.S. (2007) Cannibalism and fratricide: mechanisms and raisons d'etre. *Nat Rev Microbiol* **5**: 219-229.
- Claverys,J.P., and Martin,B. (2003) Bacterial "competence" genes: signatures of active transformation, or only remnants? *Trends Microbiol* **11**: 161-165.
- Claverys,J.P., Martin,B., and Havarstein,L.S. (2007) Competence-induced fratricide in streptococci. *Mol Microbiol* **64**: 1423-1433.

- Claverys, J.P., Martin, B., and Polard, P. (2009) The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiol Rev* **33**: 643-656.
- Cohen, D. (1966) Optimizing reproduction in a randomly varying environment. *J Theor Biol* **12**: 119-129.
- Collins, R.F., Davidsen, L., Derrick, J.P., Ford, R.C., and Tonjum, T. (2001) Analysis of the PilQ secretin from *Neisseria meningitidis* by transmission electron microscopy reveals a dodecameric quaternary structure. *J Bacteriol* **183**: 3825-3832.
- Collins, R.F., Ford, R.C., Kitmitto, A., Olsen, R.O., Tonjum, T., and Derrick, J.P. (2003) Three-dimensional structure of the *Neisseria meningitidis* secretin PilQ determined from negative-stain transmission electron microscopy. *J Bacteriol* **185**: 2611-2617.
- Collins, R.F., Frye, S.A., Balasingham, S., Ford, R.C., Tonjum, T., and Derrick, J.P. (2005) Interaction with type IV pili induces structural changes in the bacterial outer membrane secretin PilQ. *J Biol Chem* **280**: 18923-18930.
- Collins, R.F., Frye, S.A., Kitmitto, A., Ford, R.C., Tonjum, T., and Derrick, J.P. (2004) Structure of the *Neisseria meningitidis* outer membrane PilQ secretin complex at 12 Å resolution. *J Biol Chem* **279**: 39750-39756.
- Craig, L., and Li, J. (2008) Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* **18**: 267-277.
- D'Souza, C., Nakano, M.M., and Zuber, P. (1994) Identification of *comS*, a gene of the *srfA* operon that regulates the establishment of genetic competence in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **91**: 9397-9401.
- Dagkessamanskaia, A., Moscoso, M., Henard, V., Guiral, S., Overweg, K., Reuter, M. *et al.* (2004) Interconnection of competence, stress and *CiaR* regulons in *Streptococcus pneumoniae*: competence triggers stationary phase autolysis of *ciaR* mutant cells. *Mol Microbiol* **51**: 1071-1086.
- de Jong, I.G., Veening, J.W., and Kuipers, O.P. (2010) Heterochronic phosphorelay gene expression as a source of heterogeneity in *Bacillus subtilis* spore formation. *J Bacteriol* **192**: 2053-2067.
- Delcour, J., Ferain, T., and Hols, P. (2000) Advances in the genetics of thermophilic lactic acid bacteria. *Curr Opin Biotechnol* **11**: 497-504.
- Deng, D.M., Liu, M.J., ten Cate, J.M., and Crielaard, W. (2007) The *VicRK* system of *Streptococcus mutans* responds to oxidative stress. *J Dent Res* **86**: 606-610.
- Desai, K., Mashburn-Warren, L., Federle, M.J., and Morrison, D.A. (2012) Development of competence for genetic transformation of *Streptococcus mutans* in a chemically defined medium. *J Bacteriol* **194**: 3774-3780.
- Dockery, J.D., and Keener, J.P. (2001) A mathematical model for quorum sensing in *Pseudomonas aeruginosa*. *Bull Math Biol* **63**: 95-116.
- Dorer, M.S., Cohen, I.E., Sessler, T.H., Fero, J., and Salama, N.R. (2013) Natural competence promotes *Helicobacter pylori* chronic infection. *Infect Immun* **81**: 209-215.
- Draskovic, I., and Dubnau, D. (2005) Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* **55**: 881-896.
- Dubnau, D. (1999) DNA uptake in bacteria. *Annu Rev Microbiol* **53**: 217-244.
- Dubnau, D., and Losick, R. (2006) Bistability in bacteria. *Mol Microbiol* **61**: 564-572.
- Duque, C., Stipp, R.N., Wang, B., Smith, D.J., Hofling, J.F., Kuramitsu, H.K. *et al.* (2011) Downregulation of GbpB, a component of the *VicRK* regulon, affects biofilm formation and cell surface characteristics of *Streptococcus mutans*. *Infect Immun* **79**: 786-796.
- Echenique, J., Kadioglu, A., Romao, S., Andrew, P.W., and Trombe, M.C. (2004) Protein serine/threonine kinase StkP positively controls virulence and competence in *Streptococcus pneumoniae*. *Infect Immun* **72**: 2434-2437.

- Echenique,J.R., and Trombe,M.C. (2001) Competence repression under oxygen limitation through the two-component MicAB signal-transducing system in *Streptococcus pneumoniae* and involvement of the PAS domain of MicB. *J Bacteriol* **183**: 4599-4608.
- Elowitz,M.B., Levine,A.J., Siggia,E.D., and Swain,P.S. (2002) Stochastic gene expression in a single cell. *Science* **297**: 1183-1186.
- Ferrell,J.E., Jr. (2002) Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol* **14**: 140-148.
- Findlay,W.A., and Redfield,R.J. (2009) Coevolution of DNA uptake sequences and bacterial proteomes. *Genome Biol Evol* **1**: 45-55.
- Finkel,S.E., and Kolter,R. (2001) DNA as a nutrient: novel role for bacterial competence gene homologs. *J Bacteriol* **183**: 6288-6293.
- Fischer,W., Windhager,L., Rohrer,S., Zeiller,M., Karnholz,A., Hoffmann,R. *et al.* (2010) Strain-specific genes of *Helicobacter pylori*: genome evolution driven by a novel type IV secretion system and genomic island transfer. *Nucleic Acids Res* **38**: 6089-6101.
- Fleuchot,B., Gitton,C., Guillot,A., Vidic,J., Nicolas,P., Besset,C. *et al.* (2011) Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol Microbiol* **80**: 1102-1119.
- Fleuchot,B., Guillot,A., Mezange,C., Besset,C., Chambellon,E., Monnet,V., and Gardan,R. (2013) Rgg-associated SHP signaling peptides mediate cross-talk in Streptococci. *PLoS One* **8**: e66042.
- Fonseca,L.L., and Voit,E.O. (2015) Comparison of mathematical frameworks for modeling erythropoiesis in the context of malaria infection. *Math Biosci* **270**: 224-236.
- Fontaine,L., Boutry,C., de Frahan,M.H., Delplace,B., Fremaux,C., Horvath,P. *et al.* (2010a) A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J Bacteriol* **192**: 1444-1454.
- Fontaine,L., Dandoy,D., Boutry,C., Delplace,B., de Frahan,M.H., Fremaux,C. *et al.* (2010b) Development of a versatile procedure based on natural transformation for marker-free targeted genetic modification in *Streptococcus thermophilus*. *Appl Environ Microbiol* **76**: 7870-7877.
- Fontaine,L., Goffin,P., Dubout,H., Delplace,B., Baulard,A., Lecat-Guillet,N. *et al.* (2013) Mechanism of competence activation by the ComRS signalling system in streptococci. *Mol Microbiol* **87**: 1113-1132.
- Fontaine,L., Wahl,A., Flechard,M., Mignolet,J., and Hols,P. (2015) Regulation of competence for natural transformation in streptococci. *Infect Genet Evol* **33**: 343-360.
- Frye,S.A., Assalkhou,R., Collins,R.F., Ford,R.C., Petersson,C., Derrick,J.P., and Tonjum,T. (2006) Topology of the outer-membrane secretin PilQ from *Neisseria meningitidis*. *Microbiology* **152**: 3751-3764.
- Fuqua,W.C., Winans,S.C., and Greenberg,E.P. (1994) Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol* **176**: 269-275.
- Furuya,E.Y., and Lowy,F.D. (2006) Antimicrobial-resistant bacteria in the community setting. *Nat Rev Microbiol* **4**: 36-45.
- Gamba,P., Jonker,M.J., and Hamoen,L.W. (2015) A Novel Feedback Loop That Controls Bimodal Expression of Genetic Competence. *PLoS Genet* **11**: e1005047.
- Gardan,R., Besset,C., Gitton,C., Guillot,A., Fontaine,L., Hols,P., and Monnet,V. (2013) Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J Bacteriol* **195**: 1845-1855.
- Gardan,R., Besset,C., Guillot,A., Gitton,C., and Monnet,V. (2009) The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J Bacteriol* **191**: 4647-4655.
- Gardner,T.S., and Faith,J.J. (2005) Reverse-engineering transcription control networks. *Phys Life Rev* **2**: 65-88.

- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P. *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67-71.
- Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B. (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* **22**: 245-252.
- Goryachev, A.B., Toh, D.J., and Lee, T. (2006) Systems analysis of a quorum sensing network: design constraints imposed by the functional requirements, network topology and kinetic constants. *Biosystems* **83**: 178-187.
- Govin, J., and Berger, S.L. (2009) Genome reprogramming during sporulation. *Int J Dev Biol* **53**: 425-432.
- Griffith, F. (1928) The Significance of Pneumococcal Types. *Journal of Hygiene* **27**: 113-159.
- Grohmann, E., Muth, G., and Espinosa, M. (2003) Conjugative plasmid transfer in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**: 277-301, table.
- Groot, K.T., Koenders, E., Hagemeijer, Y., Overmars, L., Siezen, R.J., de Vos, W.M., and Francke, C. (2012) Comparative genome analysis of central nitrogen metabolism and its control by GlnR in the class Bacilli. *BMC Genomics* **13**: 191.
- Gupta, C., Lopez, J.M., Azencott, R., Bennett, M.R., Josic, K., and Ott, W. (2014) Modeling delay in genetic networks: from delay birth-death processes to delay stochastic differential equations. *J Chem Phys* **140**: 204108.
- Hahn, J., Luttinger, A., and Dubnau, D. (1996) Regulatory inputs for the synthesis of ComK, the competence transcription factor of *Bacillus subtilis*. *Mol Microbiol* **21**: 763-775.
- Hahn, J., Maier, B., Haijema, B.J., Sheetz, M., and Dubnau, D. (2005) Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell* **122**: 59-71.
- Hahn, J., Tanner, A.W., Carabetta, V.J., Cristea, I.M., and Dubnau, D. (2015) ComGA-RelA interaction and persistence in the *Bacillus subtilis* K-state. *Mol Microbiol* **97**: 454-471.
- Haijema, B.J., Hahn, J., Haynes, J., and Dubnau, D. (2001) A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol* **40**: 52-64.
- Hale, and Verduyne-Lunel (1993) Introduction to functional differential equations. *Applied Mathematical Science* **99**.
- Halfmann, A., Schnorpfel, A., Muller, M., Marx, P., Gunzler, U., Hakenbeck, R., and Bruckner, R. (2011) Activity of the two-component regulatory system CiaRH in *Streptococcus pneumoniae* R6. *J Mol Microbiol Biotechnol* **20**: 96-104.
- Hamoen, L.W., Eshuis, H., Jongbloed, J., Venema, G., and van Sinderen D. (1995) A small gene, designated comS, located within the coding region of the fourth amino acid-activation domain of *srfA*, is required for competence development in *Bacillus subtilis*. *Mol Microbiol* **15**: 55-63.
- Hamoen, L.W., Van Werkhoven, A.F., Venema, G., and Dubnau, D. (2000) The pleiotropic response regulator DegU functions as a priming protein in competence development in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **97**: 9246-9251.
- Hamoen, L.W., Venema, G., and Kuipers, O.P. (2003) Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology* **149**: 9-17.
- Haustenne, L., Bastin, G., Coron, J.M., Fontaine, L., and Hols, P. On stability analysis of genetic regulatory networks represented by delay-differential equations. 2014.
- Havarstein, L.S. (2010) Increasing competence in the genus *Streptococcus*. *Mol Microbiol* **78**: 541-544.
- Havarstein, L.S., Coomaraswamy, G., and Morrison, D.A. (1995) An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **92**: 11140-11144.
- Havarstein, L.S., Gaustad, P., Nes, I.F., and Morrison, D.A. (1996) Identification of the streptococcal competence-pheromone receptor. *Mol Microbiol* **21**: 863-869.

- Havarstein, L.S., Hakenbeck, R., and Gaustad, P. (1997) Natural competence in the genus *Streptococcus*: evidence that streptococci can change phenotype by interspecies recombinational exchanges. *J Bacteriol* **179**: 6589-6594.
- Heinrich, R., and Schuster, S. (1998) The modelling of metabolic systems. Structure, control and optimality. *Biosystems* **47**: 61-77.
- Heinrich, R., Schuster, S., and Holzhutter, H.G. (1991) Mathematical analysis of enzymic reaction systems using optimization principles. *Eur J Biochem* **201**: 1-21.
- Hense, B.A., Kuttler, C., Muller, J., Rothballer, M., Hartmann, A., and Kreft, J.U. (2007) Does efficiency sensing unify diffusion and quorum sensing? *Nat Rev Microbiol* **5**: 230-239.
- Hoelzer, M.A., and Michod, R.E. (1991) DNA repair and the evolution of transformation in *Bacillus subtilis*. III. Sex with damaged DNA. *Genetics* **128**: 215-223.
- Hols, P., Hancy, F., Fontaine, L., Grossiord, B., Prozzi, D., Leblond-Bourget, N. *et al.* (2005) New insights in the molecular biology and physiology of *Streptococcus thermophilus* revealed by comparative genomics. *FEMS Microbiol Rev* **29**: 435-463.
- Horvath, P., and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**: 167-170.
- Horvath, P., Coute-Monvoisin, A.C., Romero, D.A., Boyaval, P., Fremaux, C., and Barrangou, R. (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int J Food Microbiol* **131**: 62-70.
- Houser, J.R., Ford, E., Chatterjea, S.M., Maleri, S., Elston, T.C., and Errede, B. (2012) An improved short-lived fluorescent protein transcriptional reporter for *Saccharomyces cerevisiae*. *Yeast* **29**: 519-530.
- Huang, G., Wan, J., and Chen, H. (2013) Density-based Monte Carlo filter and its applications in nonlinear stochastic differential equation models. *Comput Biol Med* **43**: 135-143.
- Humbert, O., Dorer, M.S., and Salama, N.R. (2011) Characterization of *Helicobacter pylori* factors that control transformation frequency and integration length during inter-strain DNA recombination. *Mol Microbiol* **79**: 387-401.
- Hussain, H., Branny, P., and Allan, E. (2006) A eukaryotic-type serine/threonine protein kinase is required for biofilm formation, genetic competence, and acid resistance in *Streptococcus mutans*. *J Bacteriol* **188**: 1628-1632.
- Iannelli, F., Oggioni, M.R., and Pozzi, G. (2005) Sensor domain of histidine kinase ComD confers competence phenotype specificity in *Streptococcus pneumoniae*. *FEMS Microbiol Lett* **252**: 321-326.
- Ilangovan, A., Connery, S., and Waksman, G. (2015) Structural biology of the Gram-negative bacterial conjugation systems. *Trends Microbiol* **23**: 301-310.
- Jarmer, H., Berka, R., Knudsen, S., and Saxild, H.H. (2002) Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. *FEMS Microbiol Lett* **206**: 197-200.
- Jeltsch, A., and Pingoud, A. (1996) Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. *J Mol Evol* **42**: 91-96.
- Johnsborg, O., Eldholm, V., and Havarstein, L.S. (2007) Natural genetic transformation: prevalence, mechanisms and function. *Res Microbiol* **158**: 767-778.
- Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.P. (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* **12**: 181-196.
- Johnston, C., Martin, B., Granadel, C., Polard, P., and Claverys, J.P. (2013a) Programmed protection of foreign DNA from restriction allows pathogenicity island exchange during pneumococcal transformation. *PLoS Pathog* **9**: e1003178.
- Johnston, C., Martin, B., Polard, P., and Claverys, J.P. (2013b) Postreplication targeting of transformants by bacterial immune systems? *Trends Microbiol* **21**: 516-521.
- Johnston, C., Polard, P., and Claverys, J.P. (2013c) The DpnI/DpnII pneumococcal system, defense against foreign attack without compromising genetic exchange. *Mob Genet Elements* **3**: e25582.

- Jorth,P., and Whiteley,M. (2012) An evolutionary link between natural transformation and CRISPR adaptive immunity. *MBio* **3**.
- Kampf,J., and Stulke,J. (2015) Minor Cause-Major Effect: A Novel Mode of Control of Bistable Gene Expression. *PLoS Genet* **11**: e1005229.
- Karlsson,D., Karlsson,S., Gustafsson,E., Normark,B.H., and Nilsson,P. (2007) Modeling the regulation of the competence-evoking quorum sensing network in *Streptococcus pneumoniae*. *Biosystems* **90**: 211-223.
- Karr,J.R., Sanghvi,J.C., Macklin,D.N., Gutschow,M.V., Jacobs,J.M., Bolival,B., Jr. *et al.* (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* **150**: 389-401.
- Kato,F., Tainaka,K., Sone,S., Morita,S., Iida,H., and Yoshimura,J. (2011) Combined effects of prevention and quarantine on a breakout in SIR model. *Sci Rep* **1**: 10.
- Kearns,D.B., and Losick,R. (2005) Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev* **19**: 3083-3094.
- Khan,R., Rukke,H.V., Ricomini Filho,A.P., Fimland,G., Arntzen,M.O., Thiede,B., and Petersen,F.C. (2012) Extracellular identification of a processed type II ComR/ComS pheromone of *Streptococcus mutans*. *J Bacteriol* **194**: 3781-3788.
- Kim,J.N., Stanhope,M.J., and Burne,R.A. (2013) Core-gene-encoded peptide regulating virulence-associated traits in *Streptococcus mutans*. *J Bacteriol* **195**: 2912-2920.
- Kjos,M., Miller,E., Slager,J., Lake,F.B., Gericke,O., Roberts,I.S. *et al.* (2016) Expression of *Streptococcus pneumoniae* Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System. *PLoS Pathog* **12**: e1005422.
- Kojima,A., Nakano,K., Wada,K., Takahashi,H., Katayama,K., Yoneda,M. *et al.* (2012) Infection of specific strains of *Streptococcus mutans*, oral bacteria, confers a risk of ulcerative colitis. *Sci Rep* **2**: 332.
- Koyama,Y., Hoshino,T., Tomizuka,N., and Furukawa,K. (1986) Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other *Thermus* spp. *J Bacteriol* **166**: 338-340.
- Kreth,J., Merritt,J., Shi,W., and Qi,F. (2005) Co-ordinated bacteriocin production and competence development: a possible mechanism for taking up DNA from neighbouring species. *Mol Microbiol* **57**: 392-404.
- Kung,S.H., Retchless,A.C., Kwan,J.Y., and Almeida,R.P. (2013) Effects of DNA size on transformation and recombination efficiencies in *Xylella fastidiosa*. *Appl Environ Microbiol* **79**: 1712-1717.
- Kussell,E., and Leibler,S. (2005) Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**: 2075-2078.
- Lacks,S., Greenberg,B., and Neuberger,M. (1974) Role of a deoxyribonuclease in the genetic transformation of *Diplococcus pneumoniae*. *Proc Natl Acad Sci U S A* **71**: 2305-2309.
- Lamagni,T.L., Neal,S., Keshishian,C., Powell,D., Potz,N., Pebody,R. *et al.* (2009) Predictors of death after severe *Streptococcus pyogenes* infection. *Emerg Infect Dis* **15**: 1304-1307.
- Laurenceau,R., Krasteva,P.V., Diallo,A., Ouarti,S., Duchateau,M., Malosse,C. *et al.* (2015) Conserved *Streptococcus pneumoniae* spiroosomes suggest a single type of transformation pilus in competence. *PLoS Pathog* **11**: e1004835.
- Laurenceau,R., Pehau-Arnaudet,G., Baconnais,S., Gault,J., Malosse,C., Dujeancourt,A. *et al.* (2013) A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog* **9**: e1003473.
- Leisner,M., Kuhr,J.T., Radler,J.O., Frey,E., and Maier,B. (2009) Kinetics of genetic switching into the state of bacterial competence. *Biophys J* **96**: 1178-1188.
- Leisner,M., Stingl,K., Frey,E., and Maier,B. (2008) Stochastic switching to competence. *Curr Opin Microbiol* **11**: 553-559.
- Leisner,M., Stingl,K., Radler,J.O., and Maier,B. (2007) Basal expression rate of *comK* sets a 'switching-window' into the K-state of *Bacillus subtilis*. *Mol Microbiol* **63**: 1806-1816.

- Lemme,A., Grobe,L., Reck,M., Tomasch,J., and Wagner-Dobler,I. (2011) Subpopulation-specific transcriptome analysis of competence-stimulating-peptide-induced *Streptococcus mutans*. *J Bacteriol* **193**: 1863-1877.
- Lewis,K. (2007) Persister cells, dormancy and infectious disease. *Nat Rev Microbiol* **5**: 48-56.
- Lewis,K. (2010) Persister cells. *Annu Rev Microbiol* **64**: 357-372.
- Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P., and Cvitkovitch,D.G. (2001) Natural genetic transformation of *Streptococcus mutans* growing in biofilms. *J Bacteriol* **183**: 897-908.
- Li,Y.H., Tang,N., Aspiras,M.B., Lau,P.C., Lee,J.H., Ellen,R.P., and Cvitkovitch,D.G. (2002) A quorum-sensing signaling system essential for genetic competence in *Streptococcus mutans* is involved in biofilm formation. *J Bacteriol* **184**: 2699-2708.
- Liebal,U.W., Millat,T., de Jong,I.G., Kuipers,O.P., Volker,U., and Wolkenhauer,O. (2010) How mathematical modelling elucidates signalling in *Bacillus subtilis*. *Mol Microbiol* **77**: 1083-1095.
- Liu,Y., and Burne,R.A. (2011) The major autolysin of *Streptococcus gordonii* is subject to complex regulation and modulates stress tolerance, biofilm formation, and extracellular-DNA release. *J Bacteriol* **193**: 2826-2837.
- Lo,S.M., and Blokesch,M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet* **8**: e1002778.
- Lo,S.M., and Blokesch,M. (2013) A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res* **41**: 3644-3658.
- Lorenz,M.G., and Wackernagel,W. (1994) Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* **58**: 563-602.
- Lu,W.W., Wang,Y., Wang,T., and Kong,J. (2015) The global regulator CodY in *Streptococcus thermophilus* controls the metabolic network for escalating growth in the milk environment. *Appl Environ Microbiol* **81**: 2349-2358.
- Lunsford,R.D., and Roble,A.G. (1997) comYA, a gene similar to comGA of *Bacillus subtilis*, is essential for competence-factor-dependent DNA transformation in *Streptococcus gordonii*. *J Bacteriol* **179**: 3122-3126.
- Luo,P., Li,H., and Morrison,D.A. (2004) Identification of ComW as a new component in the regulation of genetic transformation in *Streptococcus pneumoniae*. *Mol Microbiol* **54**: 172-183.
- Maamar,H., and Dubnau,D. (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol Microbiol* **56**: 615-624.
- Maamar,H., Raj,A., and Dubnau,D. (2007) Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* **317**: 526-529.
- Macfadyen,L.P. (2000) Regulation of competence development in *Haemophilus influenzae*. *J Theor Biol* **207**: 349-359.
- Mahmood,T., Masud,T., Imran,M., Ahmed,I., and Khalid,N. (2013) Selection and characterization of probiotic culture of *Streptococcus thermophilus* from dahi. *Int J Food Sci Nutr* **64**: 494-501.
- Makarova,K., Slesarev,A., Wolf,Y., Sorokin,A., Mirkin,B., Koonin,E. *et al.* (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**: 15611-15616.
- Makarova,K.S., Haft,D.H., Barrangou,R., Brouns,S.J., Charpentier,E., Horvath,P. *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**: 467-477.
- Makarova,K.S., Wolf,Y.I., Alkhnbashi,O.S., Costa,F., Shah,S.A., Saunders,S.J. *et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**: 722-736.
- Manning,T., Sleator,R.D., and Walsh,P. (2014) Biologically inspired intelligent decision making: a commentary on the use of artificial neural networks in bioinformatics. *Bioengineered* **5**: 80-95.

- Marks, L.R., Mashburn-Warren, L., Federle, M.J., and Hakansson, A.P. (2014) Streptococcus pyogenes biofilm growth in vitro and in vivo and its role in colonization, virulence, and genetic exchange. *J Infect Dis* **210**: 25-34.
- Martin, B., Granadel, C., Campo, N., Henard, V., Prudhomme, M., and Claverys, J.P. (2010) Expression and maintenance of ComD-ComE, the two-component signal-transduction system that controls competence of Streptococcus pneumoniae. *Mol Microbiol* **75**: 1513-1528.
- Martin, B., Quentin, Y., Fichant, G., and Claverys, J.P. (2006) Independent evolution of competence regulatory cascades in streptococci? *Trends Microbiol* **14**: 339-345.
- Martin, B., Soulet, A.L., Mirouze, N., Prudhomme, M., Mortier-Barriere, I., Granadel, C. *et al.* (2013) ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence). *Mol Microbiol* **87**: 394-411.
- Mashburn-Warren, L., Morrison, D.A., and Federle, M.J. (2010) A novel double-tryptophan peptide pheromone controls competence in Streptococcus spp. via an Rgg regulator. *Mol Microbiol* **78**: 589-606.
- Mashburn-Warren, L., Morrison, D.A., and Federle, M.J. (2012) The cryptic competence pathway in Streptococcus pyogenes is controlled by a peptide pheromone. *J Bacteriol* **194**: 4589-4600.
- Mazda, Y., Kawada-Matsuo, M., Kanbara, K., Oogai, Y., Shibata, Y., Yamashita, Y. *et al.* (2012) Association of CiaRH with resistance of Streptococcus mutans to antimicrobial peptides in biofilms. *Mol Oral Microbiol* **27**: 124-135.
- McAdams, H.H., and Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A* **94**: 814-819.
- Mehr, S., and Wood, N. (2012) Streptococcus pneumoniae--a review of carriage, infection, serotype replacement and vaccination. *Paediatr Respir Rev* **13**: 258-264.
- Mercer, D.K., Melville, C.M., Scott, K.P., and Flint, H.J. (1999) Natural genetic transformation in the rumen bacterium Streptococcus bovis JB1. *FEMS Microbiol Lett* **179**: 485-490.
- Merritt, J., Qi, F., and Shi, W. (2005) A unique nine-gene comY operon in Streptococcus mutans. *Microbiology* **151**: 157-166.
- Merritt, J., Zheng, L., Shi, W., and Qi, F. (2007) Genetic characterization of the hdrRM operon: a novel high-cell-density-responsive regulator in Streptococcus mutans. *Microbiology* **153**: 2765-2773.
- Miao, R., and Guild, W.R. (1970) Competent Diplococcus pneumoniae accept both single- and double-stranded deoxyribonucleic acid. *J Bacteriol* **101**: 361-364.
- Michod, R.E., Wojciechowski, M.F., and Hoelzer, M.A. (1988) DNA repair and the evolution of transformation in the bacterium Bacillus subtilis. *Genetics* **118**: 31-39.
- Mirouze, N., Berge, M.A., Soulet, A.L., Mortier-Barriere, I., Quentin, Y., Fichant, G. *et al.* (2013) Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. *Proc Natl Acad Sci U S A* **110**: E1035-E1044.
- Mirouze, N., Ferret, C., Yao, Z., Chastanet, A., and Carballido-Lopez, R. (2015) MreB-Dependent Inhibition of Cell Elongation during the Escape from Competence in Bacillus subtilis. *PLoS Genet* **11**: e1005299.
- Moore, M.E., Lam, A., Bhatnagar, S., and Solnick, J.V. (2014) Environmental determinants of transformation efficiency in Helicobacter pylori. *J Bacteriol* **196**: 337-344.
- Morrison, D.A. (1997) Streptococcal competence for genetic transformation: regulation by peptide pheromones. *Microb Drug Resist* **3**: 27-37.
- Morrison, D.A., Guedon, E., and Renault, P. (2013) Competence for natural genetic transformation in the Streptococcus bovis group streptococci S. infantarius and S. macedonicus. *J Bacteriol* **195**: 2612-2620.
- Morrison, D.A., Mortier-Barriere, I., Attaiech, L., and Claverys, J.P. (2007) Identification of the major protein component of the pneumococcal eclipse complex. *J Bacteriol* **189**: 6497-6500.

- Mortier-Barriere,I., Velten,M., Dupaigne,P., Mirouze,N., Pietrement,O., McGovern,S. *et al.* (2007) A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* **130**: 824-836.
- Ng,W.L., Kazmierczak,K.M., and Winkler,M.E. (2004) Defective cell wall synthesis in *Streptococcus pneumoniae* R6 depleted for the essential PcsB putative murein hydrolase or the VicR (YycF) response regulator. *Mol Microbiol* **53**: 1161-1175.
- Nguyen,A., Prugel-Bennett,A., and Dasmahapatra,S. (2016) A Low Dimensional Approximation For Competence In *Bacillus Subtilis*. *IEEE/ACM Trans Comput Biol Bioinform* **13**: 272-280.
- Nielsen,K.M., van Weerelt,M.D., Berg,T.N., Bones,A.M., Hagler,A.N., and van Elsas,J.D. (1997) Natural transformation and availability of transforming DNA to *Acinetobacter calcoaceticus* in soil microcosms. *Appl Environ Microbiol* **63**: 1945-1952.
- Norgard,M.V., and Imaeda,T. (1978) Physiological factors involved in the transformation of *Mycobacterium smegmatis*. *J Bacteriol* **133**: 1254-1262.
- Ogura,M., Yamaguchi,H., Kobayashi,K., Ogasawara,N., Fujita,Y., and Tanaka,T. (2002) Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J Bacteriol* **184**: 2344-2351.
- Okinaga,T., Xie,Z., Niu,G., Qi,F., and Merritt,J. (2010) Examination of the *hdrRM* regulon yields insight into the competence system of *Streptococcus mutans*. *Mol Oral Microbiol* **25**: 165-177.
- Olsen,R.J., Shelburne,S.A., and Musser,J.M. (2009) Molecular mechanisms underlying group A streptococcal pathogenesis. *Cell Microbiol* **11**: 1-12.
- Opdyke,J.A., Scott,J.R., and Moran,C.P., Jr. (2003) Expression of the secondary sigma factor sigmaX in *Streptococcus pyogenes* is restricted at two levels. *J Bacteriol* **185**: 4291-4297.
- Pandey,A., and Mann,M. (2000) Proteomics to study genes and genomes. *Nature* **405**: 837-846.
- Pastink,M.I., Teusink,B., Hols,P., Visser,S., de Vos,W.M., and Hugenholtz,J. (2009) Genome-scale model of *Streptococcus thermophilus* LMG18311 for metabolic comparison of lactic acid bacteria. *Appl Environ Microbiol* **75**: 3627-3633.
- Patel,G.B., Nash,J.H., Agnew,B.J., and Sprott,G.D. (1994) Natural and Electroporation-Mediated Transformation of *Methanococcus voltae* Protoplasts. *Appl Environ Microbiol* **60**: 903-907.
- Patenge,N., Fiedler,T., and Kreikemeyer,B. (2013) Common regulators of virulence in streptococci. *Curr Top Microbiol Immunol* **368**: 111-153.
- Perry,J.A., Cvitkovitch,D.G., and Levesque,C.M. (2009) Cell death in *Streptococcus mutans* biofilms: a link between CSP and extracellular DNA. *FEMS Microbiol Lett* **299**: 261-266.
- Pestova,E.V., Havarstein,L.S., and Morrison,D.A. (1996) Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. *Mol Microbiol* **21**: 853-862.
- Peterson,S.N., Sung,C.K., Cline,R., Desai,B.V., Snesrud,E.C., Luo,P. *et al.* (2004) Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol Microbiol* **51**: 1051-1070.
- Piotrowski,A., Luo,P., and Morrison,D.A. (2009) Competence for genetic transformation in *Streptococcus pneumoniae*: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW. *J Bacteriol* **191**: 3359-3366.
- Pohl,K., Francois,P., Stenz,L., Schlink,F., Geiger,T., Herbert,S. *et al.* (2009) CodY in *Staphylococcus aureus*: a regulatory link between metabolism and virulence gene expression. *J Bacteriol* **191**: 2953-2963.
- Provvedi,R., Chen,I., and Dubnau,D. (2001) NucA is required for DNA cleavage during transformation of *Bacillus subtilis*. *Mol Microbiol* **40**: 634-644.
- Prudhomme,M., Attaiech,L., Sanchez,G., Martin,B., and Claverys,J.P. (2006) Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science* **313**: 89-92.

- Prudhomme,M., Berge,M., Martin,B., and Polard,P. (2016) Pneumococcal Competence Coordination Relies on a Cell-Contact Sensing Mechanism. *PLoS Genet* **12**: e1006113.
- Qi,F., Kreth,J., Levesque,C.M., Kay,O., Mair,R.W., Shi,W. *et al.* (2005) Peptide pheromone induced cell death of *Streptococcus mutans*. *FEMS Microbiol Lett* **251**: 321-326.
- Reck,M., Rutz,K., Kunze,B., Tomasch,J., Surapaneni,S.K., Schulz,S., and Wagner-Dobler,I. (2011) The biofilm inhibitor carolacton disturbs membrane integrity and cell division of *Streptococcus mutans* through the serine/threonine protein kinase PknB. *J Bacteriol* **193**: 5692-5706.
- Redfield,R.J. (2001) Do bacteria have sex? *Nat Rev Genet* **2**: 634-639.
- Robertson,G.T., Ng,W.L., Foley,J., Gilmour,R., and Winkler,M.E. (2002) Global transcriptional analysis of *clpP* mutations of type 2 *Streptococcus pneumoniae* and their effects on physiology and virulence. *J Bacteriol* **184**: 3508-3520.
- Saskova,L., Novakova,L., Basler,M., and Branny,P. (2007) Eukaryotic-type serine/threonine protein kinase StkP is a global regulator of gene expression in *Streptococcus pneumoniae*. *J Bacteriol* **189**: 4168-4179.
- Schilling,C.H., Schuster,S., Palsson,B.O., and Heinrich,R. (1999) Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnol Prog* **15**: 296-303.
- Schnorpfeil,A., Kranz,M., Kovacs,M., Kirsch,C., Gartmann,J., Brunner,I. *et al.* (2013) Target evaluation of the non-coding csRNAs reveals a link of the two-component regulatory system CiaRH to competence control in *Streptococcus pneumoniae* R6. *Mol Microbiol* **89**: 334-349.
- Schultz,D., Ben,J.E., Onuchic,J.N., and Wolynes,P.G. (2007) Molecular level stochastic model for competence cycles in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **104**: 17582-17587.
- Schwabe,A., Maarleveld,T.R., and Bruggeman,F.J. (2013) Exploration of the spontaneous fluctuating activity of single enzyme molecules. *FEBS Lett* **587**: 2744-2752.
- Seitz,P., and Blokesch,M. (2013) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* **37**: 336-363.
- Seitz,P., Pezeshgi,M.H., Borgeaud,S., Bulushev,R.D., Steinbock,L.J., Radenovic,A. *et al.* (2014) ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable *Vibrio cholerae* cells. *PLoS Genet* **10**: e1004066.
- Senadheera,D.B., Cordova,M., Ayala,E.A., Chavez de Paz,L.E., Singh,K., Downey,J.S. *et al.* (2012) Regulation of bacteriocin production and cell death by the *VicRK* signaling system in *Streptococcus mutans*. *J Bacteriol* **194**: 1307-1316.
- Senadheera,M.D., Guggenheim,B., Spatafora,G.A., Huang,Y.C., Choi,J., Hung,D.C. *et al.* (2005) A *VicRK* signal transduction system in *Streptococcus mutans* affects *gtfBCD*, *gbpB*, and *ftf* expression, biofilm formation, and genetic competence development. *J Bacteriol* **187**: 4064-4076.
- Senadheera,M.D., Lee,A.W., Hung,D.C., Spatafora,G.A., Goodman,S.D., and Cvitkovitch,D.G. (2007) The *Streptococcus mutans* *vicX* gene product modulates *gtfB/C* expression, biofilm formation, genetic competence, and oxidative stress tolerance. *J Bacteriol* **189**: 1451-1458.
- Sengupta,D., and Kar,S. (2015) Are Quasi-Steady-State Approximated Models Suitable for Quantifying Intrinsic Noise Accurately? *PLoS One* **10**: e0136668.
- Silal,S.P., Little,F., Barnes,K.I., and White,L.J. (2014) Towards malaria elimination in Mpumalanga, South Africa: a population-level mathematical modelling approach. *Malar J* **13**: 297.
- Silal,S.P., Little,F., Barnes,K.I., and White,L.J. (2015) Hitting a Moving Target: A Model for Malaria Elimination in the Presence of Population Movement. *PLoS One* **10**: e0144990.
- Skipington,E., and Ragan,M.A. (2011) Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol Rev* **35**: 707-735.

- Slager,J., Kjos,M., Attaiech,L., and Veening,J.W. (2014) Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell* **157**: 395-406.
- Smits,W.K., Eschevins,C.C., Susanna,K.A., Bron,S., Kuipers,O.P., and Hamoen,L.W. (2005) Stripping Bacillus: ComK auto-stimulation is responsible for the bistable response in competence development. *Mol Microbiol* **56**: 604-614.
- Smits,W.K., Kuipers,O.P., and Veening,J.W. (2006) Phenotypic variation in bacteria: the role of feedback regulation. *Nat Rev Microbiol* **4**: 259-271.
- Son,M., Ahn,S.J., Guo,Q., Burne,R.A., and Hagen,S.J. (2012) Microfluidic study of competence regulation in Streptococcus mutans: environmental inputs modulate bimodal and unimodal expression of comX. *Mol Microbiol* **86**: 258-272.
- Steinmoen,H., Knutsen,E., and Havarstein,L.S. (2002) Induction of natural competence in Streptococcus pneumoniae triggers lysis and DNA release from a subfraction of the cell population. *Proc Natl Acad Sci U S A* **99**: 7681-7686.
- Stenz,L., Francois,P., Whiteson,K., Wolz,C., Linder,P., and Schrenzel,J. (2011) The CodY pleiotropic repressor controls virulence in gram-positive pathogens. *FEMS Immunol Med Microbiol* **62**: 123-139.
- Stevens,K.E., Chang,D., Zwack,E.E., and Seibert,M.E. (2011) Competence in Streptococcus pneumoniae is regulated by the rate of ribosomal decoding errors. *MBio* **2**.
- Sung,C.K., and Morrison,D.A. (2005) Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in Streptococcus pneumoniae. *J Bacteriol* **187**: 3052-3061.
- Szollosi,G.J., Derenyi,I., and Vellai,T. (2006) The maintenance of sex in bacteria is ensured by its potential to reload genes. *Genetics* **174**: 2173-2180.
- Talagas,A., Fontaine,L., Ledesma-Garca,L., Mignolet,J., Li de la Sierra-Gallay, Lazar,N. *et al.* (2016) Structural Insights into Streptococcal Competence Regulation by the Cell-to-Cell Communication System ComRS. *PLoS Pathog* **12**: e1005980.
- Tettelin,H. (2004) Streptococcal genomes provide food for thought. *Nat Biotechnol* **22**: 1523-1524.
- Thattai,M., and van,O.A. (2004) Stochastic gene expression in fluctuating environments. *Genetics* **167**: 523-530.
- Thomas,C.M., and Nielsen,K.M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* **3**: 711-721.
- Tian,T. (2016) Bayesian Computation Methods for Inferring Regulatory Network Models Using Biomedical Data. *Adv Exp Med Biol* **939**: 289-307.
- Tian,X.L., Dong,G., Liu,T., Gomez,Z.A., Wahl,A., Hols,P., and Li,Y.H. (2013) MecA protein acts as a negative regulator of genetic competence in Streptococcus mutans. *J Bacteriol* **195**: 5196-5206.
- Torres,N.V., and Santos,G. (2015) The (Mathematical) Modeling Process in Biosciences. *Front Genet* **6**: 354.
- Tovpeko,Y., Bai,J., and Morrison,D.A. (2016) Competence for Genetic Transformation in Streptococcus pneumoniae: Mutations in sigmaA Bypass the ComW Requirement for Late Gene Expression. *J Bacteriol* **198**: 2370-2378.
- Tovpeko,Y., and Morrison,D.A. (2014) Competence for genetic transformation in Streptococcus pneumoniae: mutations in sigmaA bypass the comW requirement. *J Bacteriol* **196**: 3724-3734.
- Trovato,A., Seno,F., Zanardo,M., Alberghini,S., Tondello,A., and Squartini,A. (2014) Quorum vs. diffusion sensing: a quantitative analysis of the relevance of absorbing or reflecting boundaries. *FEMS Microbiol Lett* **352**: 198-203.
- Turgay,K., Hahn,J., Burghoorn,J., and Dubnau,D. (1998) Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor. *EMBO J* **17**: 6730-6738.
- Turgay,K., Hamoen,L.W., Venema,G., and Dubnau,D. (1997) Biochemical characterization of a molecular switch involving the heat shock protein ClpC, which controls the activity of ComK, the competence transcription factor of Bacillus subtilis. *Genes Dev* **11**: 119-128.

- van Sinderen D., ten, B.A., Hayema, B.J., Hamoen, L., and Venema, G. (1994) Molecular cloning and sequence of comK, a gene required for genetic competence in *Bacillus subtilis*. *Mol Microbiol* **11**: 695-703.
- van Someren, E.P., Vaes, B.L., Steegenga, W.T., Sijbers, A.M., Dechering, K.J., and Reinders, M.J. (2006) Least absolute regression network analysis of the murine osteoblast differentiation network. *Bioinformatics* **22**: 477-484.
- Vanlier, J., Tiemann, C.A., Hilbers, P.A., and van Riel, N.A. (2013) Parameter uncertainty in biochemical models described by ordinary differential equations. *Math Biosci* **246**: 305-314.
- Varhimo, E., Savijoki, K., Jalava, J., Kuipers, O.P., and Varmanen, P. (2007) Identification of a novel streptococcal gene cassette mediating SOS mutagenesis in *Streptococcus uberis*. *J Bacteriol* **189**: 5210-5222.
- Veening, J.W., Igoshin, O.A., Eijlander, R.T., Nijland, R., Hamoen, L.W., and Kuipers, O.P. (2008a) Transient heterogeneity in extracellular protease production by *Bacillus subtilis*. *Mol Syst Biol* **4**: 184.
- Veening, J.W., Smits, W.K., Hamoen, L.W., and Kuipers, O.P. (2006) Single cell analysis of gene expression patterns of competence development and initiation of sporulation in *Bacillus subtilis* grown on chemically defined media. *J Appl Microbiol* **101**: 531-541.
- Veening, J.W., Smits, W.K., and Kuipers, O.P. (2008b) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**: 193-210.
- Wahl, A., Servais, F., Drucbert, A.S., Foulon, C., Fontaine, L., and Hols, P. (2014) Control of natural transformation in salivarius *Streptococci* through specific degradation of sigmaX by the MecA-ClpCP protease complex. *J Bacteriol* **196**: 2807-2816.
- Walton, K., and Marshall, J.E. (1987) Direct method for TDS stability analysis. *IEE proceedings - Control Theory and Applications*.
- Wang, Z., Gao, H., Cao, J., and Liu, X. (2008) On delayed genetic regulatory networks with polytopic uncertainties: robust stability analysis. *IEEE Trans Nanobioscience* **7**: 154-163.
- Wayne, K.J., Li, S., Kazmierczak, K.M., Tsui, H.C., and Winkler, M.E. (2012) Involvement of WalK (VicK) phosphatase activity in setting WalR (VicR) response regulator phosphorylation level and limiting cross-talk in *Streptococcus pneumoniae* D39 cells. *Mol Microbiol* **86**: 645-660.
- Wei, K., Moinat, M., Maarleveld, T.R., and Bruggeman, F.J. (2014) Stochastic simulation of prokaryotic two-component signalling indicates stochasticity-induced active-state locking and growth-rate dependent bistability. *Mol Biosyst* **10**: 2338-2346.
- Wenderska, I.B., Lukenda, N., Cordova, M., Magarvey, N., Cvitkovitch, D.G., and Senadheera, D.B. (2012) A novel function for the competence inducing peptide, XIP, as a cell death effector of *Streptococcus mutans*. *FEMS Microbiol Lett* **336**: 104-112.
- Weng, L., Piotrowski, A., and Morrison, D.A. (2013) Exit from competence for genetic transformation in *Streptococcus pneumoniae* is regulated at multiple levels. *PLoS One* **8**: e64197.
- Wessels, L.F., van Someren, E.P., and Reinders, M.J. (2001) A comparison of genetic network models. *Pac Symp Biocomput*: 508-519.
- West, S.A., Winzer, K., Gardner, A., and Diggle, S.P. (2012) Quorum sensing and the confusion about diffusion. *Trends Microbiol* **20**: 586-594.
- Wolfgang, M., Lauer, P., Park, H.S., Brossay, L., Hebert, J., and Koomey, M. (1998) PilT mutations lead to simultaneous defects in competence for natural transformation and twitching motility in pilated *Neisseria gonorrhoeae*. *Mol Microbiol* **29**: 321-330.
- Worrell, V.E., Nagle, D.P., Jr., McCarthy, D., and Eisenbraun, A. (1988) Genetic transformation system in the archaeobacterium *Methanobacterium thermoautotrophicum* Marburg. *J Bacteriol* **170**: 653-656.
- Wylie, C.S., Trout, A.D., Kessler, D.A., and Levine, H. (2010) Optimal strategy for competence differentiation in bacteria. *PLoS Genet* **6**: e1001108.

- Xie,Z., Okinaga,T., Niu,G., Qi,F., and Merritt,J. (2010) Identification of a novel bacteriocin regulatory system in *Streptococcus mutans*. *Mol Microbiol* **78**: 1431-1447.
- Yeung,M.K., Tegner,J., and Collins,J.J. (2002) Reverse engineering gene networks using singular value decomposition and robust regression. *Proc Natl Acad Sci U S A* **99**: 6163-6168.
- Youk,H., and Lim,W.A. (2014) Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science* **343**: 1242782.
- Zaccaria,E., van,B.P., de,G.A., Morrison,D.A., Smith,H., and Wells,J.M. (2014) Control of competence for DNA transformation in *streptococcus suis* by genetically transferable pherotypes. *PLoS One* **9**: e99394.
- Zhang,X.Z., You,C., and Zhang,Y.H. (2014) Transformation of *Bacillus subtilis*. *Methods Mol Biol* **1151**: 95-101.
- Zimmer,C., and Sahle,S. (2016) Comparison of approaches for parameter estimation on stochastic models: Generic least squares versus specialized approaches. *Comput Biol Chem* **61**: 75-85.
- Zou,J., Han,Y., and So,S.S. (2008) Overview of artificial neural networks. *Methods Mol Biol* **458**: 15-23.